Untying Gordian knots: The evolution and biogeography of the large European apomictic polyploid *Ranunculus auricomus* plant complex

Dissertation
for the award of the degree
"Doctor rerum naturalium" (Dr.rer.nat)
of the Georg-August University Göttingen

within the doctoral program Biology
of the Georg-August University School of Science (GAUSS)

submitted by
Master of Science (M.Sc.) in Evolution, Ecology, and Systematics
Kevin Karbstein

from Gotha

Göttingen, August 2021

Thesis Committee

Referee: Prof. Dr. Elvira Hörandl
Department of Systematics, Biodiversity, and Evolution of Plants (with Herbarium), Albrecht-von-Haller Institute for Plant Sciences

Second Referee: PD Dr. Sven Bradler
Department of Animal Evolution and Biodiversity, Johann-Friedrich-Blumenbach Institute for Zoology and Anthropology

Members of the Examination Board

Reviewer: Prof. Dr. Elvira Hörandl
Department of Systematics, Biodiversity, and Evolution of Plants (with Herbarium), Albrecht-von-Haller Institute for Plant Sciences

Second Reviewer: PD Dr. Sven Bradler
Department of Animal Evolution and Biodiversity, Johann-Friedrich-Blumenbach Institute for Zoology and Anthropology

Further Members of the Examination Board

Prof. Dr. Christoph Bleidorn, Department of Animal Evolution and Biodiversity, Johann-Friedrich-Blumenbach Institute for Zoology and Anthropology

Prof. Dr. Holger Kreft, Department of Biodiversity, Macroecology, and Biogeography, Centre for Biodiversity and Sustainable Land Use

Prof. Dr. Stefan Scheu, Department of Animal Ecology, Johann-Friedrich-Blumenbach Institute for Zoology and Anthropology

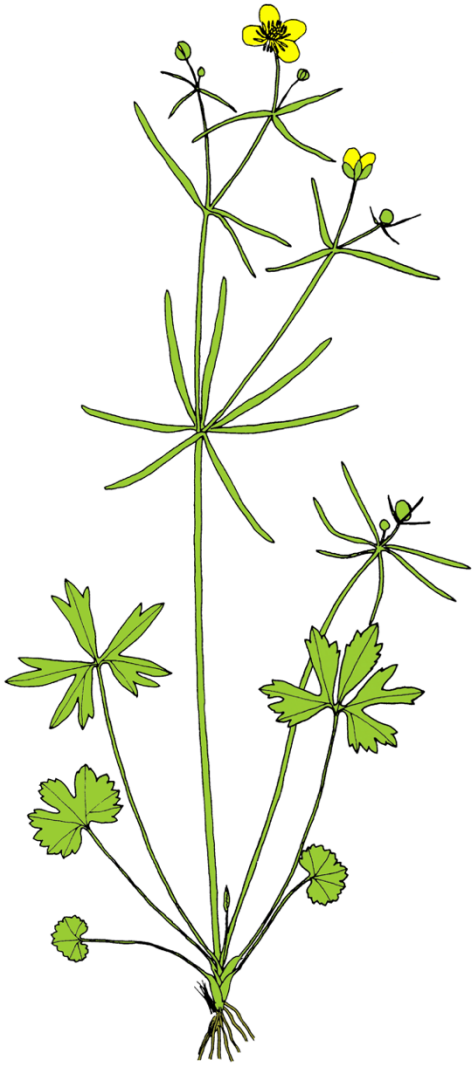
PD Dr. Christian Roos, German Primate Center, Leibniz Institute for Primate Research

Date of Oral Examination

8th October 2021

"Nothing is more extraordinary in the history of the Vegetable Kingdom, as it seems to me, than the apparently very sudden or abrupt development of the higher [eudicotylous flowering] plants.
I have sometimes speculated whether
there did not exist somewhere
during long ages an extremely
isolated continent, perhaps
near the South Pole."

(Letter written by Charles Darwin on the 6th August 1881 to Joseph Hooker; Friedman, 2009; Buggs, 2021)



In recent years, polyploidy and hybridization have been considered to explain the described phenomenon. Nevertheless, the speciation process shaped by both processes is still poorly understood, for example, in the mainly polyploid asexual, evolutionary young, eudicotolyous flowering plant complex *Ranunculus auricomus* [goldilock buttercups].

Figure: Drawing of a *Ranunculus auricomus* individual.
Figure source: Hörandl & Gutermann (1995), edited by Hodač et al. (2018) and M. Daubert

Curriculum Vitae

Kevin Karbstein

*26th December 1992 in Gotha, Thuringia

12/2017 – 10/2021	**PhD student** at the Albrecht-von-Haller Institute for Plant Science, Department of Systematics, Biodiversity, and Evolution of Plants (with Herbarium), Georg-August University Göttingen, Germany **PhD thesis** (DFG, SPP 1991 TaxonOmics): "The biodiversity of apomictic polyploid plants: The *Ranunculus auricomus* complex"
10/2014 – 09/2017	**Master of Science** in Evolution, Ecology, and Systematics, Institute of Systematic Botany with Herbarium Haussknecht and Botanical Garden, Friedrich-Schiller University Jena, Germany (grade 1.3) **Master thesis**: "Moving beyond assumptions: Relationships between population size, genetic and functional diversity, and habitat heterogeneity in *Trifolium montanum*"
08/2016 – 09/2017	Research assistant (student) at the Max-Planck Institute of Human History, Jena, Germany
10/2015 – 03/2016	Internship and following working contracts, Forest Research and Competence Center (FFK), Gotha, Germany Project: "Forest adaptation in the middle, high und ridge altitudes of the Thuringian Forest"
10/2011 – 09/2014	**Bachelor of Science** in Biology, Institute of Systematic Botany with Herbarium Haussknecht and Botanical Garden, Friedrich-Schiller University Jena, Germany (grade 1.7) **Bachelor thesis**: "Biologisch-ökologische Untersuchungen an *Thymus praecox* Opiz in den *Badlands* und in den umliegenden Halbtrocken- und Trockenrasen Thüringens mit Etablierung von Mikrosatellitenmarkern"
06/2011	**Abitur** (A level), Gustav-Freytag-Gymnasium Gotha, Thuringia, Germany (grade 1.3)

Conference Contributions

04/2021	SPP 1991 (DFG) TaxonOmics Kick-Off Meeting, online talk (*Ranunculus auricomus*)
10/2019	XVI OPTIMA Meeting, Athens (Greece), talk (*Ranunculus auricomus*)
06/2019	3rd annual SPP 1991 (DFG) TaxonOmics Meeting, Munich (Germany), talk (*Ranunculus auricomus*)
08/2018	II Joint Congress on Evolutionary Biology, Montpellier (France), poster presentation (*Ranunculus auricomus*)
02/2018	3rd Annual Meeting in Conservation Genetics, Vienna (Austria), talk (*Trifolium montanum*)
02/2017	Schutz bedrohter Pflanzenarten in Mitteleuropa, Berlin (Germany), poster presentation (*Thymus praecox*)

Teaching

2021	Practical bachelor course and excursions "Structure and Diversity of plants"
2018 – 2021	Involved in the supervision of bachelor (Elisabeth Rahmsdorf, Ellen Lorberg, and Lea Hansen) and master theses (Mareike Daubert)
2020 – 2021	Practical bachelor course "Plant Systematics and Basic Course Botany"
2018 – 2020	Lab and bioinformatic bachelor course "Methods of Systematic Botany"
2019	Alpine excursion master course
2014 – 2017	Practical bachelor course "Mathematics and Biostatistics in R"

Further Activities

2021	Member of the Thuringian Botanical Society (TBG), German Ecological Society (GFÖ), and German Botanical Society (DBG)
2020 – 2021	Reviewer for scientific journals (Ecology and Evolution, Taxon, Journal of Biogeography, American Journal of Botany)
2016 – 2021	Red List mapping of endangered plant species in Thuringia

Publication List

Peer-Reviewed Articles

Karbstein, K., Tomasello, S., Hodač, L., Natascha, W., Marinček, P., Barke, B. H., Pätzold, C., & Hörandl, E. (2021). Unraveling phylogenetic relationships, reticulate evolution, and genome composition of polyploid plant complexes by Rad-Seq and Hyb-Seq. ***BioRxiv***, 1–86. https://doi.org/10.1101/2021.08.30.458250

Karbstein, K., Tomasello, S., Hodač, L., Lorberg, E., Daubert, M., & Hörandl, E. (2021). Moving beyond assumptions: Polyploidy and environmental effects explain a geographical parthenogenesis scenario in European plants. ***Molecular Ecology***, *30*, 2659–2675. https://doi.org10.1111/mec.15919

Karbstein, K., Gockel, S., Frischbier, N., Kahlert, K., Konnert, M., & Profft, I. (accepted). "High-altitude spruces" in Central Europe - a summarizing contribution to phenotypic and (epi)genetic differentiation within *Picea abies* (L.) H.KARST. ***Allgemeine Forst- und Jagdzeitschrift***.

Strecker, T., Jesch, A., Bachmann, D., Jüds, M., **Karbstein, K.**, Ravenek, J., Roscher, C., Weigelt, A., Eisenhauer, N., & Scheu, S. (2021). Incorporation of mineral nitrogen into the soil food web as affected by plant community composition. ***Ecology and Evolution***, *11*, 4295–4309. https://doi.org/10.1002/ece3.7325

Karbstein, K., Rahmsdorf, E., Tomasello, S., Hodač, L., & Hörandl, E. (2020). Breeding system of diploid sexuals within the *Ranunculus auricomus* complex and its role in a

geographical parthenogenesis scenario. ***Ecology and Evolution***, *10*, 14435–14450. https://doi.org/10.1002/ece3.7073

Barke, B. H., **Karbstein, K.**, Daubert, M., & Hörandl, E. (2020). The relation of meiotic behaviour to hybridity, polyploidy and apomixis in the *Ranunculus auricomus* complex (Ranunculaceae). ***BMC Plant Biology***, *20*, 523. https://doi.org/10.1186/s12870-020-02654-3

Karbstein, K., Tomasello, S., Hodač, L., Dunkel, F. G., Daubert, M., & Hörandl, E. (2020). Phylogenomics supported by geometric morphometrics reveals delimitation of sexual species within the polyploid apomictic *Ranunculus auricomus* complex (Ranunculaceae). ***Taxon***, *69*, 1191–1220. https://doi.org/10.1002/tax.12365

Tomasello, S., **Karbstein, K.**, Hodač, L., Pätzold, C., & Hörandl, E. (2020). Phylogenomics unravels Quarternary vicariance and allopatric speciation patterns in temperate-montane plant species: A case study on the *Ranunculus auricomus* species complex. ***Molecular Ecology***, *29*, 2031–2049. https://doi.org/10.1111/mec.15458

Karbstein, K., Prinz, K., Hellwig, F., & Römermann, C. (2020). Plant intraspecific functional trait variation is related to within-habitat heterogeneity and genetic diversity in *Trifolium montanum* L. ***Ecology and Evolution***, *10*, 5015–5033. https://doi.org/10.1002/ece3.6255

Karbstein, K., Tomasello, S., Prinz, K. (2019). Desert-like badlands and surrounding (semi-)dry grasslands of Central Germany promote small-scale phenotypic and genetic differentiation in *Thymus praecox*. ***Ecology and Evolution***, *9*, 14066–14084. https://doi.org/10.1002/ece3.5844

Articles in Preparation

Karbstein, K., Barke, B. H., Pätzold, C., Hörandl, E. A draft of the *Ranunculus auricomus* genome.

Karbstein, K., Bradican, J. P., Hodač, L., Tomasello, S., Hörandl, E. Geometric morphometrics combined with nuclear-gene data elucidates complex differentiation and evolution of diagnostic traits within the *Ranunculus auricomus* complex.

Karbstein, K., Tomasello, S., Bradican, J. P., Hodač, L., Hörandl, E. A new tetraploid, sexual-apomictic species of the *Ranunculus auricomus* complex.

Schubrt, D., **Karbstein, K.**, & Hodač, L. Relationships between leaf shape and environment gradients within the European *Ranunculus auricomus* complex.

Contents

Summary

Polyploidy, the presence of two or more full genomic complements, repeatedly occurs across the tree of life. In plants, not only the economic but particularly the evolutionary importance is overwhelming. Polyploidization events, probably connected to key innovations (e.g., vessel elements or the carpel), occurred frequently in the evolutionary history of flowering plants, which are the most species-rich group in the plant kingdom (ca. 370,000 species) and contain 30–70% neopolyploids. Polyploidy and hybridization (i.e., allopolyploidy) are particularly considered to create biotypes with novel genomic compositions and to be key factors for subsequent speciation and macroevolution. In plants, both processes are frequently connected to apomixis, i.e., the reproduction via asexually-formed seeds. However, the enigmatic phenomenon of plant speciation accompanied by polyploidy and apomixis is still poorly understood despite tremendous progress in the field of genomics.

The question of "What is a species?" is of highest priority for evolutionary biologists: Species are the fundamental units for biodiversity, and further evolutionary and ecological research relies on well-defined entities. Evolutionarily young plant species complexes offer a unique opportunity to study plant speciation and accompanying processes. They usually comprise a few sexual progenitor species, and numerous polyploid, partly apomictic, hybrid derivatives. In apomictic lineages, the lack of recombination and cross-fertilization can result in numerous clonal lineages with fixed morphological and ecological traits (agamospecies). Nevertheless, even recognizing and delimiting the sexual progenitors of species complexes is methodically challenging due to low genetic divergence, possible hybrid origins, ongoing gene flow, and/or incomplete lineage sorting (ILS). Integrative approaches using both genomic and morphometric data for disentangling the young progenitors are still lacking so far.

The biogeography and evolution of those plant complexes is even more challenging. Apomicts frequently occupy larger areas or more northern regions compared to their sexual relatives, a phenomenon called geographical parthenogenesis (GP). GP patterns usually have a Pleistocene context because climatic range shifts in temperate to boreal zones offered frequent opportunities for interspecific hybridization, probably giving rise to apomixis in the Northern Hemisphere. Factors shaping GP patterns are still controversially discussed. GP has been widely attributed to advantages of apomicts caused by polyploidy and uniparental reproduction, i.e., fixed levels of high heterozygosity leading to increased stress tolerance, and self-fertility leading to better colonizing capabilities. On the one hand, complex interactions of genome-wide heterozygosity, ploidy, reproduction mode (sexual versus asexual), and climatic environmental factors shaping

GP have not been studied enough. On the other hand, potential disadvantages of sexual progenitors due to their breeding system on fitness and genetic diversity have received even less attention. Finally, alongside biogeography, the reticulate relationships and genome composition and evolution of young, large polyploid plant species complexes have not yet been deciphered comprehensively. Besides challenges attributed to numerous numbers of polyploidization and hybridization events, bioinformatic analyses are also often hampered by missing information on progenitors, ploidy levels, and reproduction modes.

The European apomictic polyploid *Ranunculus auricomus* (goldilock buttercup) plant complex is well-suited to study all the aforementioned issues. The majority of goldilock buttercups probably arose from hybridization of a few sexual progenitors, leading to more than 800 described, morphologically highly diverse agamospecies. Sexuals are estimated to have speciated less than 1.0 million years ago, and agamospecies are probably much younger. In this thesis, using *R. auricomus* as a model system, I examined the recalcitrant and hitherto poorly understood phylogenetic, genomic, and biogeographical relationships of young polyploid apomictic plant complexes.

I developed a comprehensive theoretical and bioinformatic workflow, starting with analyzing the evolution of the sexual progenitor species, continuing with unraveling reproduction modes and biogeography of apomictic polyploids, and ending up with revealing the reticulate origins and genome composition and evolution of the polyploid complex. Spanning up to 251 populations and 87 *R. auricomus* taxa Europe-wide, this work gathered data of 97,312 genomic loci (RADseq), 663 nuclear genes (target enrichment), and 71 plastid regions, and 1,474 leaf ploidy, 4,669 reproductive seed, 284 reproductive crossing (seed sets), as well as 1,593 geometric morphometric measurements.

First of all, phylogenomics based on RADseq, nuclear gene, and geometric morphometric data supported the lumping of the twelve described sexual morphospecies into five newly circumscribed progenitor species. These species represent clearly distinguishable genetic main lineages or clusters, which are both well geographically isolated and morphologically differentiated: *R. cassubicifolius* s.l., *R. envalirensis* s.l., *R. flabellifolius*, *R. marsicus*, and *R. notabilis* s.l. Mainly within-clade reticulate relationships, missing geographical isolation, and a lack of distinctive morphological characters led to this taxonomic treatment. Interestingly, allopatric speciation events took place ca. 0.83–0.58 million years ago during a period of severe climatic oscillations, and were probably triggered by vicariance processes of a widespread European forest-understory ancestor. Sexual species re-circumscriptions were additionally

supported by population crossing experiments. Besides inbreeding depression, outbreeding benefits, and sudden self-compatibility, crossings also revealed a lack of reproductive barriers among some of the formerly described morphospecies.

Moreover, flow cytometric ploidy and reproductive, RADseq, and environmental data were combined into a genetically informed path analysis based on Generalized Linear Mixed Models (GLMMs). The analysis unveiled a complex European GP scenario, whereby diploids compared to polyploids showed significantly higher sexuality (percent of sexual seeds), more petals (petaloid nectary leaves), and up to three times less genome-wide heterozygosity. Surprisingly, sexuality was positively associated with solar radiation and isothermality, and heterozygosity was positively related to temperature seasonality. Results fit the southern distribution of diploid sexuals and suggest a higher resistance of polyploid apomicts to more extreme climatic conditions.

Finally, a self-developed, multidisciplinary workflow incorporating all previously gathered data demonstrated, for the first time, the predominantly allopolyploid origin, genome composition, and post-origin genome evolution of the *R. auricomus* complex. Taxa were organized in only three to five supported, north-south distributed clades or cluster, each usually containing diploid sexual progenitor species. Allopolyploidizations involved two to three different diploid sexual subgenomes per event. Only one autotetraploid event was detected. Allotetraploids were characterized by subgenome dominance and enormous post-origin evolution, i.e., Mendelian segregation of hybrid generations, back-crossing to parents, and/or gene flow due to facultative sexuality of apomicts. Four diploid sexual progenitors and a previously unknown, nowadays extinct progenitor, probably gave rise to the more than 800 taxa of the European *R. auricomus* complex. Analyses also showed that the majority of analyzed polyploid agamospecies are non-monophyletic and similar morphotypes probably originated multiple times. The lack of monophyly suggests a comprehensive taxonomic revision of the entire complex.

In the General Discussion, I combine my thesis results with existing plant studies on diploid sexual and polyploid apomictic phylogenetics, biogeography, and composition and genome evolution of young species complexes. I explain the taxonomic conclusions and how species complexes link micro- and macroevolutionary processes. Finally, I give conclusions of my thesis and an outlook of the project and the field of polyploid phylogenetics.

Zusammenfassung

Polyploidie, das Vorhandensein von zwei oder mehr vollständigen Chromosomensätzen, tritt wiederholt über den gesamten Baum des Lebens auf. Bei Pflanzen ist die wirtschaftliche, aber vor allem auch die evolutionäre Bedeutung überwältigend. Polyploidisierungen, wahrscheinlich verbunden mit Schlüsselinnovationen (z.B. die Entwicklung der Gefäßelemente oder des Fruchtblattes), traten in der Evolution der Blütenpflanzen häufig auf. Blütenpflanzen sind die artenreichste Gruppe im Pflanzenreich mit ca. 370,000 Arten und umfassen 30–70% Neopolyploide. Es wird angenommen, dass Polyploidie und Hybridisierung (Allopolyploidie) besonders zur Entstehung von Biotypen mit neuartiger genomischer Zusammensetzung beitragen und damit Schlüsselfaktoren für nachfolgende Artbildungen und Makroevolution sind. Bei Pflanzen sind beide Prozesse häufig mit Apomixis, der Reproduktion über asexuell gebildete Samen, verbunden. Das rätselhafte Phänomen der von Polyploidie und Apomixis begleiteten Artbildung ist jedoch trotz enormer Fortschritte auf dem Gebiet der Genomik noch immer kaum verstanden.

Die Frage „Was ist eine Art?“ hat für Evolutionsbiologen höchste Priorität: Arten sind die Grundlage der Biodiversitätsforschung, und die evolutionäre und ökologische Forschung stützt sich auf gut definierte Einheiten. Evolutionär junge Artkomplexe bieten eine einzigartige Möglichkeit die Artbildung bei Pflanzen und deren begleitende Prozesse zu erforschen und zu verstehen. Sie umfassen meist wenige sexuelle Stammarten und zahlreiche polyploide, teilweise apomiktische, hybridogene Derivate. Das Fehlen von Rekombination und Kreuzbestäubung in apomiktischen Linien kann zu einer Vielzahl klonaler Hybridlinien mit fixierten morphologischen und ökologischen Merkmalen führen (Agamospezies). Selbst das Erkennen und Abgrenzen der sexuellen Stammarten ist aufgrund geringer genetischer Divergenz, eventuellen hybridogenen Ursprüngen, stetigem Genfluss und/oder unvollständiger genetischer Auftrennung der Abstammungslinien (ILS) methodisch herausfordernd. Integrative Ansätze, die sowohl genomische als auch morphometrische Daten verwenden, um die jungen Stammarten aufzutrennen, fehlen bisher.

Die Biogeographie und Evolution der Artkomplexe ist weitaus komplexer. Apomikten besetzen im Vergleich zu ihren sexuellen Verwandten häufig größere Areale oder sind in nördlicheren Regionen verbreitet, ein Phänomen, das als Geographische Parthenogenese (GP) bezeichnet wird. GP-Muster haben meist einen pleistozänen Kontext. Klimatische Schwankungen in den gemäßigten und borealen Zonen boten häufig Möglichkeiten zur interspezifischen Hybridisierung, was wahrscheinlich auch zur Entstehung von Apomixis auf der Nordhalbkugel

geführt hat. Faktoren, die diese Muster erzeugen, werden immer noch kontrovers diskutiert. GP-Muster wurden bisher oft den Vorteilen apomiktischer Populationen aufgrund von (Allo)polyploidie und uniparentaler Fortpflanzung zugeschrieben: Fixierte, hohe Heterozygotie führt zu einer erhöhten Stresstoleranz, und Selbstfertilität bedingt eine bessere Kolonisierungsfähigkeit. Einerseits sind die komplexen Wechselwirkungen von genomweiter Heterozygotie, Ploidie, Reproduktionssmodi (sexuell versus asexuell) und klimatischer Umweltfaktoren auf GP-Muster nicht ausreichend untersucht worden, andererseits wurden potentielle Nachteile sexueller Stammarten aufgrund ihres Fortpflanzungssystems auf Fitness und genetische Vielfalt bisher kaum betrachtet. Schließlich sind neben der Biogeographie die retikulate Evolution und die genomische Zusammensetzung und Evolution junger, großer polyploider Pflanzenartenkomplexe noch nicht detailliert entschlüsselt worden. Neben Herausforderungen, die auf eine hohe Anzahl an Polyploidisierungs- und Hybridisierungsereignissen zurückzuführen sind, werden bioinformatische Analysen oft durch fehlende Informationen zu sexuellen Stammarten, Ploidiegraden und Reproduktionsmodi erschwert.

Der europäische, polyploid-apomiktische *Ranunculus auricomus* (Gold-Hahnenfuß) Pflanzenkomplex ist gut geeignet, um alle aufgeworfenen Fragestellungen zu untersuchen. Der Komplex entstand wahrscheinlich durch unzählige Hybridisierungen weniger sexueller Stammarten. Bisher wurden mehr als 800 morphologisch sehr diverse Agamospezies (Derivate) beschriebenen. Die sexuellen Stammarten werden weniger als 1.0 Millionen Jahren alt geschätzt, und die Agamospezies sind wahrscheinlich noch viel jünger. In meiner Dissertation habe ich unter Verwendung des *R. auricomus* Komplexes als Modellsystem die bisher wenig verstandenen phylogenetischen, genomischen und biogeographischen Beziehungen junger, polyploider Pflanzengruppen untersucht.

Ich habe einen umfassenden theoretischen und bioinformatischen Workflow entwicklelt, beginnend mit der Untersuchung der Evolution der sexuellen Stammarten, über die Entschlüsselung der Reproduktionsmodi und Biogeographie polyploid-apomiktischer Derivate bis hin zur Aufdeckung der retikulaten Ursprünge und Genomzusammensetzung und -evolution des Polyploidkomplexes. Diese Arbeit umfasst 251 Populationen und 87 *R. auricomus* Taxa europaweit. Die Analysen basieren auf 97,312 genomischen Loci (RADseq), 663 Kerngenen (target enrichment) und 71 Plastidenregionen, und 1,474 Blattploidie-, 4,669 Reproduktions-Samen-, 284 Kreuzungs- (Samenansatz), und 1,593 Morphometrie-Messungen.

Phylogenomische Daten basierend auf RADseq, Kerngenen und geometrischer Morphometrie unterstützten die Zusammenlegung der zwölf sexuellen Morphospezies in fünf neu klassifizierte Stammarten. Diese Arten stellen klar unterscheidbare genetische Hauptlinien oder Cluster dar, die sowohl geographisch gut isoliert als auch morphologisch klar differenziert sind: *R. cassubicifolius* s.l., *R. envalirensis* s.l., *R. flabellifolius*, *R. marsicus* und *R. notabilis* s.l. Enorme retikulate Beziehungen innerhalb der Kladen, die nicht-vorhandene geographische Isolation und das Fehlen markanter morphologischer Merkmale haben zu diesem taxonomischen Konzept geführt. Allopatrische Artbildungsereignisse fanden interessanterweise vor ca. 0.83–0.58 Millionen Jahren während enormer klimatischer Schwankungen statt und wurden wahrscheinlich durch Vikarianzprozesse aus einer weit verbreiteten europäischen Stammart ausgelöst. Darüber hinaus wurde die neue Umschreibung der sexuellen Stammarten durch Populationskreuzungsexperimente unterstützt. Kreuzungen zeigten neben Inzuchtdepression, Auszuchtvorteilen und plötzlicher Selbstkompatibilität auch völlig fehlende Reproduktionsbarrieren zwischen einigen Morphospezies.

Darüber hinaus wurden durchflusszytometrische Ploidy- und Reproduktions-, genomweite RADseq- und klimatische Umweltdaten in einer genetisch-informierten Pfadanalyse basierend auf Generalisierten Linearen Gemischten Modellen (GLMMs) kombiniert. Die Analyse hat ein komplexes europäisches GP-Szenario aufgedeckt, in der Diploide im Vergleich zu Polyploiden eine signifikant höhere Sexualität (Prozent sexueller Samen), mehr Blütenblätter (petaloide Nektarblätter) und bis zu dreimal weniger genomweite Heterozygotie zeigten. Die Sexualität war überaschenderweise positiv mit Sonneneinstrahlung und Isothermalität verbunden, und die Heterozygotie zeigte einen positiven Zusammenhang mit der Temperatursaisonalität. Die Ergebnisse stimmen mit der südlichen Verbreitung diploid-sexueller Populationen überein und deuten auf eine höhere Resistenz polyploid-apomiktischer Populationen gegenüber extremeren klimatischen Bedingungen hin.

Ein neu entwickelter, multidisziplinärer Workflow, der alle bisherigen Daten einbezieht, deckte zum ersten Mal den weitestgehend allopolyploiden Ursprung und die Genomzusammensetzung und -evolution des *R. auricomus* Komplexes auf. Die Taxa waren in nur drei bis fünf unterstützten, nord-süd verbreiteten Kladen oder Clustern organisiert, die jeweils meistens diploid-sexuelle Stammarten enthielten. Allopolyploidisierungsereignisse bezogen jeweils zwei bis drei verschiedene, diploid-sexuelle Subgenome ein. Es wurde nur ein autotetraploides Ereignis nachgewiesen. Allotetraploide Genome sind gekennzeichnet durch Subgenomdominanz und einer enormen Evolution nach ihrer Entstehung (z.B. Mendelsche Segregation der Hybridgenerationen, Rückkreuzungen zu Elternarten und Genfluss aufgrund

fakultativer Sexualität der Apomikten). Die über 800 Taxa des europäischen *R. auricomus*-Komplexes sind vermutlich aus vier diploiden Stammarten und eine bisher unbekannte, aktuell wahrscheinlich ausgestorbene Stammart, entstanden. Analysen zeigten auch, dass die Mehrzahl der beschriebenen polyploiden Agamospezies nicht monophyletisch ist und ähnliche Morphotypen wahrscheinlich mehrfach entstanden sind. Eine umfassende taxonomische Überarbeitung des gesamten Komplexes ist daher angebracht.

In der Allgemeinen Diskussion kombiniere ich die Ergebnisse meiner Dissertation mit bereits existierenden Pflanzenstudien zur diploid-sexuellen und polyploid-apomiktischen Phylogenetik, Biogeographie und Genomzusammensetzung und -evolution junger Artkomplexe. Ich gebe zudem taxonomische Schlussfolgerungen und erkläre wie Artkomplexe mikro- und makroevolutionäre Prozesse miteinander verbinden. Abschließend gebe ich ein Fazit über die Ergebnisse meiner Dissertation und einen Ausblick für das laufende Forschungsprojekt und der Forschungsdisziplin der polyploiden Phylogenetik.

Thesis Contribution Declarations

- Declaration of the author´s contribution to manuscripts -

Chapter 1

Phylogenomics supported by geometric morphometrics reveals delimitation of sexual species within the polyploid apomictic *Ranunculus auricomus* complex (Ranunculaceae)

Conception and design of the study; data sampling; data analysis; data interpretation; writing and figures

Chapter 2

Breeding system of diploid sexuals within the *Ranunculus auricomus* complex and its role in a geographical parthenogenesis scenario

Conception and design of the study (supporting); data sampling (supporting); data analysis; data interpretation; writing (equal) and figures

Chapter 3

Phylogenomics unravels Quaternary vicariance and allopatric speciation patterns in temperate-montane plant species: A case study on the *Ranunculus auricomus* species complex

Data sampling (supporting); data analysis (supporting); data interpretation; writing (supporting)

Chapter 4

Moving beyond assumptions: Polyploidy and environmental effects explain a geographical parthenogenesis scenario in European plants

Conception and design of the study; data sampling; data analysis; data interpretation; writing and figures

Chapter 5

Unraveling phylogenetic relationships, reticulate evolution, and genome composition of polyploid plant complexes by RAD-Seq and Hyb-Seq

Conception and design of the study; data sampling; data analysis; data interpretation; writing and figures

I General Introduction

Extensively used, humid mixed forest along a small stream close to Weimar (Thuringia, Germany), occupied by individuals of the *R. auricomus* complex. This habitat represents the locus classicus of "*R.* × *opimus*", a dissected basal leaf, polyploid asexual taxon described by O. Schwarz in 1949 (Schwarz, 1949; location "KK021", 26th April 2018). Photograph: Kevin Karbstein

The Evolutionary and General Significance of Polyploidy

"The rapid development as far as we can judge of all the higher [dicotylous flowering] plants within recent geological times is an abominable mystery" (Letter written by Charles Darwin on the 22nd July 1879 to Joseph Hooker; Friedman, 2009; Buggs, 2021)

Polyploidy is the presence of two or more full genomic complements in a cell (whole genome duplication; Comai, 2005; Wendel, 2015). It is of tremendous evolutionary but also ecological and economical significance (Van de Peer, et al. 2017; Fox et al., 2020). Polyploidy occurs across the tree of life, spanning bacteria, archaea, plants, fungi, and animals (mainly fish and amphibians; Mable et al., 2011; Van de Peer et al., 2017; Blischak et al., 2018; Jaron et al., 2021). Modern genomic analyses indicate the occurrence of at least one ancient polyploidization event in the common ancestor of all land plants, and, particularly in flowering (angiosperm) plant lineages, several have occurred (Adams & Wendel, 2005; Van de Peer et al., 2017; Leebens-Mack et al., 2019). Flowering plants are the largest plant group with more than 370,000 species (ca. 400,000–500,000 land plant species; Corlett, 2016; Lughadha et al., 2016). Key innovations of flowering plants, e.g., the carpel, double fertilization, or vessel elements, have been hypothesized to be connected to polyploidy (Soltis et al., 2015; Leebens-Mack et al., 2019). Moreover, ca. 30–70% of flowering plants are of recent polyploid origin, and neopolyploid formation potentially caused upshifts in species diversification rates (Wood et al., 2009; Soltis et al., 2015; Landis et al., 2018).

The occurrence of polyploid cells and tissues is ubiquitous in nature (including humans), and is considered a strategy for higher abiotic and biotic stress tolerance, particular in harsh or disturbed environments (Schoenfelder & Fox, 2015; Fox et al., 2020; Van de Peer et al., 2020). Interestingly, the long-term establishment of polyploidy seems to be linked to environmental conditions during the evolution of plants and to coincide with periods of global change or mass extinction (Van de Peer et al., 2020). Polyploids have multiple gene copies, allowing for gene sub- and neofunctionalizations, epigenetic changes, and different gene expressions (Comai, 2005; Blischak et al., 2018; Qiu et al., 2020). The colonization of various different ecosystems is facilitated by the larger physiological and phenotypic flexibility in relation to environmental conditions (Hörandl, 2006; Te Beest et al., 2012; Blaine Marchant et al., 2016; Rice et al., 2019; Meudt et al., 2021). For example, many polyploids are invasive weeds (e.g., spotted knapweed *Centaurea stoebe* in North America, Asian knotweed *Reynoutria japonica* or common cordgrass *Sporobolus anglicus* in Europe) or frequently occupy northern or past glaciated areas

(Hörandl, 2006; Te Beest et al., 2012; Qiu et al., 2020). Polyploidy-driven cell changes not only impact earth's biodiversity and biocomplexity, but also affect agriculture, biotechnology, and medicine (Fox et al., 2020). For example, artificially-induced polyploidy has been exploited by plant breeders to produce seedless watermelons, larger strawberries, or drought-resistant triticale, and in medicine, polyploidy is considered a major factor promoting or suppressing tumors (Fox et al., 2020; Gordon et al., 2020).

Polyploid Types, Genome Evolution, and Phenotypic Consequences

Different polyploid formation types enormously influence subsequent genome evolution. Allopolyploids originate by hybridization of different evolutionary lineages or species, whereas autopolyploids arise within individuals or from crossings of different populations within an evolutionary linage or species, followed by polyploidization (Stebbins, 1950; Comai, 2005; Blischak et al., 2018). Autopolyploids frequently form multivalents during meiosis (i.e., four chromosomes of each type) in contrast to allotetraploids, which usually form bivalents allowing a stable meiosis due to homologous and not homeologous (between diverged subgenomes) chromosome pairing (Comai, 2005). However, lineage or species boundaries are sometimes blurred, especially in evolutionarily young groups (e.g., along a speciation continuum, De Queiroz, 2007). Consequently, a third category was defined: Segmental allopolyploids, which originate out of differentiated lineages within a species and show an inheritance pattern intermediate between allo- and autopolyploids (Stebbins, 1950; Soltis & Soltis, 2000; Blischak et al., 2018). Considering genome composition, autopolyploids contain genetically similar whereas allopolyploids unify genetically diverged progenitor (parental) subgenomes.

In nature, allopolyploidy is the most frequently observed type, although many autopolyploids remained unnamed and thus unrecognized in these comparisons (Barker et al., 2016; Spoelhof et al., 2017). Allopolyploids, compared to autopolyploids, exhibit higher levels of genomic, transcriptomic, and epigenetic changes (Comai, 2005; Soltis et al., 2015; Spoelhof et al., 2017). Consequently, the combination of hybridization and polyploidization is considered particularly likely to create novel (epi)genomic and phenotypic features (Liu & Wendel, 2003; Abbott et al., 2013; Van de Peer et al., 2020; Rothfels, 2021). In addition to genomic shocks accompanying both formation types, polyploid genomes begin to revert to a functionally diploid state after polyploidization over evolutionary time scales (diploidization; Wolfe, 2001; Soltis

et al., 2015; Van de Peer et al., 2017). Several processes are responsible for this pattern, and in allopolyploids, these processes frequently lead to expression and sequence dominance of one progenitor subgenome over the other(s) (Soltis et al., 2015; Zhao et al., 2017; Blischak et al., 2018; Alger & Edger, 2020). Before polyploidization, Mendelian segregation of hybrid generations and backcrosses to progenitor species can distort the original equal subgenome contributions (e.g., Barke et al., 2018; Hodač et al., 2018). After polyploidization, a complex combination of epigenetic modifications, homeolog expression bias, homeologous gene loss towards one progenitor subgenome (biased fractionation), and gene flow among polyploid lineages lead to subgenome dominance (Blischak et al., 2018; Melichárková et al., 2020; Alger & Edger, 2020; Wagner et al., 2020).

Polyploidy and Apomixis

In plants, polyploidization and hybridization are often associated with apomixis, the reproduction via asexually-formed seeds (Asker & Jerling, 1992; Hörandl, 2006; Brukhin et al., 2019; Hojsgaard & Hörandl, 2019). In flowering plants, ca. 52% of orders, 19% of families, and 2% of genera contain apomicts, although new apomictic species are steadily being recognized (Hojsgaard et al., 2014a; Zhang et al., 2021). Apomixis is scattered along the phylogenetic tree and occurs particular often in young plant families (e.g., Asteraceae, Poaceae, and Rosaceae; Richards, 1997; Hojsgaard et al., 2014a). Hybridization is probably the main trigger of apomixis, and positive side-effects of polyploidy can foster the establishment of young apomictic lineages (Ernst, 1918; Asker & Jerling, 1992; Hand & Koltunow, 2014; Barke et al., 2018; Hojsgaard & Hörandl, 2019). Some asexual polyploid animals are also of hybrid origin (Jaron et al., 2021). However, in general, not all apomicts are (neo)polyploids, and not all (neo)polyploids show apomixis (e.g., Masci et al., 1994; Kantama et al., 2007; Brukhin et al., 2019; Jaron et al., 2021).

Research of the past few decades indicates that apomixis is (epi)genetically controlled and thus heritable, although the exact molecular regulation is still not well understood (Nogler, 1984; Hojsgaard & Hörandl, 2019; Hojsgaard, 2021). Interestingly, the introduction of apomixis-like features into crop plants is of high interest economically: The uniparental reproduction due to retainment of maternal genotypes and the thus possibility to exploit heterosis and conserve favorable allele combinations is important for agricultural purposes (Koltunow & Grossniklaus,

2003; Hojsgaard & Hörandl, 2019). Moreover, apomixis is a modification of the sexual pathway and thus usually facultative (Asker & Jerling, 1992; Ozias-Akins & Van Dijk, 2007; Hojsgaard & Hörandl, 2019). In plants, typical sexual development results in a reduced megaspore after meiosis (Fig. 1). This megaspore develops into a reduced embryo sac, which after double fertilization of the egg cell and polar nuclei forms a diploid embryo and triploid endosperm tissue (Fig. 1; e.g., Hojsgard et al., 2014b; Klatt et al., 2016). In contrast, apomictic plants skip meiosis and develop the embryo without fertilization (Fig. 1; Koltunow & Grossniklaus, 2003; Grimanelli, 2012; Barke et al., 2018; Hojsgaard & Hörandl, 2019; Ulum et al., 2020): Apomictic development begins with a somatic unreduced cell, which results in an unreduced embryo sac (functional female gametophyte, apomeiosis). This pathway is called apospory (Fig. 1), whereas in the alternative pathway, diplospory, the development starts from restitutional meiosis. Then, the unreduced egg cell develops via parthenogenesis into an embryo (parthenogenesis; the combination of apomeiosis and parthenogenesis is called gametophytic apomixis). Subsequently, the endosperm is formed via autonomous apomixis, where polar nuclei are unfertilized, or via pseudogamy where polar nuclei are fertilized by one or both sperm nuclei. These processes lead to endosperm, which is tetra- (autonomous apomixis), and penta- or hexaploid (pseudogamous apomixis), respectively (e.g., Barke et al., 2018; Ulum et al., 2020).

Therefore, functional apomixis basically involves the molecular regulation of apomeiosis, parthenogenesis, and endosperm development as a deviation from the sexual pathway (Hojsgaard, 2021). Using flow cytometric seed screening (FCSS; Matzk et al., 2000), the respective ratios of endosperm to embryo ploidies for each seed are measurable. The ratio is defined as peak index (PI, last column of Fig. 1), and directly indicates the sexual or the various apomictic reproduction modes as already demonstrated in a magnitude of plant species and species complexes (e.g., Cosendai & Hörandl, 2010; Šarhanova et al., 2012; Barke et al., 2018; Karunarathne et al., 2018; Ulum et al., 2020).

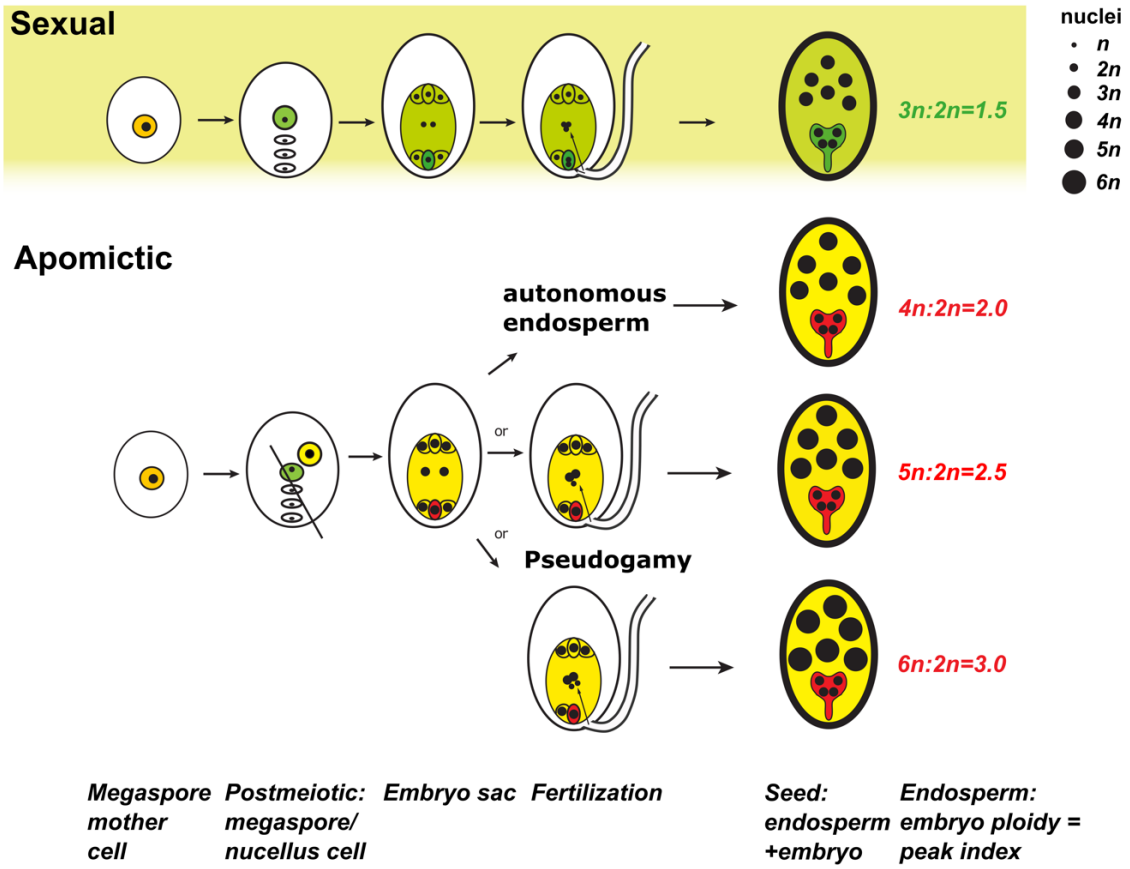


Figure 1. Sexual versus apomictic reproduction modes (developmental pathways) using the example of the *R. auricomus* complex (edited figure from Hojsgaard & Hörandl, 2019 and Karbstein et al., 2021; both published under Attribution 4.0 International, CC-BY 4.0). We symbolized ploidy levels of nuclei (n) by the size of black dots. In the *R. auricomus* complex, the unreduced gametophyte develops via apospory (gametophytic apomixis). Please see the General Introduction, section Ploidy and Apomixis for detailed explanations.

The Evolution of Polyploid Apomictic Species Complexes

Much breathtaking technical and theoretical progress has been made in the current era of genomics (Wendel, 2015; Sukumaran & Knowles, 2017; Rothfels et al., 2017; Rothfels, 2021). However, the enigmatic phenomenon of plant speciation accompanied by polyploidy and apomixis including innumerable reticulate relationships and enormous genomic changes is still poorly understood. The question of "What is a species?" is also of highest priority for evolutionary biologists. Species are the fundamental units in biodiversity and evolutionary research, but their definition and delimitation remain a persistent challenge (Hörandl, 1998, 2018; Agapow et al., 2004; Sukumaran & Knowles, 2017; Sukumaran et al., 2021). Nowadays, species are considered to be ancestor-descendant lineages, which applies to both sexual and

asexual lineages (De Queiroz, 2007; Sukumaran & Knowles, 2017; Hörandl, 2018). The challenge is to define a lineage's evolutionary role and circumscription, for example, the persistence in time and space, ecological niche, geographical distribution, or a shared phenotype (Freudenstein et al., 2017; Hörandl, 2018). Further evolutionary and ecological research relies on well-defined entities, but different species concepts can tremendously impact species richness and potentially bias further biodiversity data analyses and conservation actions (Agapow et al., 2004; Freudenstein et al., 2017; Sukumaran & Knowles, 2017; Coates et al., 2018).

To study plant speciation and accompanying processes, evolutionarily young plant species complexes offer a unique opportunity (Fig. 2). They usually comprise a few sexual diploid progenitor species, and numerous polyploid apomictic hybrid derivatives (Fig. 2; Šarhanová et al., 2012; Sochor et al., 2015; Hörandl, 2018; Carter et al., 2019). Residual sexuality enabled by the facultative character of apomixis allows for backcrossing of apomicts to progenitors and gene flow among apomicts, resulting in large reticulate structures (i.e., networks) of hundreds to thousands of hybridogenetic lineages (Weber, 1996; Sochor et al., 2015; Hörandl, 2018). Young species complexes of abundant plant genera are for example dandelions (*Taraxacum*), hawkweeds (*Hieracium* s.l.), brambles (*Rubus*), *Citrus*, and *Paspalum*. Nowadays, the diversity of polyploid apomictic complexes is either lumped into a single artificial "aggregate" largely underestimating biodiversity or overemphasized in thousands of indistinguishable "morphospecies" lacking a genetic background (Richards, 2003; Hörandl, 2018; Żabicka et al., 2020). Therefore, Hörandl (2018) proposed a concept to classify species within apomictic plant complexes. Four main types of sexual and asexual lineages are classified according to reproduction mode (sexual, facultative or obligate apomictic) and lineage stability (Fig. 2).

At the beginning of the species complex evolution, an outcrossing, sexual progenitor lineage (or species; green framing, sexual populations represented by white circles) diverged into several lineages due to (e.g., geographical or ecological) isolation over evolutionary time scales. Even recognizing and delimiting these sexual progenitors is methodically challenging because of low genetic divergence, present incomplete lineage sorting (ILS), ongoing gene flow, and/or partly hybrid origins (Hörandl, 2018; Pease et al., 2018; Wagner et al., 2019). So far, integrative approaches using both genomic and morphometric data for disentangling such young progenitors and proposing well-defined entities for subsequent analyses are lacking.

Range fluctuations can lead to secondary contact of previously isolated sexual plant and animal lineages. For instance, climatic oscillations enabled range fluctuations in Central Europe due to

cold glacials and warm interglacials during the Pleistocene (Stebbins, 1984; Taberlet et al., 1998; Brewer et al., 2002; Abbott et al., 2013; Ebdon et al., 2021). The secondary contact of previously isolated plant lineages caused the formation of hybrid clusters (red framing, mainly sexual populations represented by grey circles) and the rise of facultative or obligate apomictic populations (represented by dark grey and black circles; Carman, 1997; Dobeš et al., 2004; Pellino et al., 2013). In obligate apomictic lineages, the lack of recombination and cross-fertilization can result in fixed morphological and ecological traits (agamospecies), whereas facultative apomictic lineages are expected to form large, morphologically diverse, instable clusters because novel genotypes steadily arise via facultative apomixis within lineages, or via crossing of sexual mother plants with apomictic pollen donors (Hörandl, 2018). Occasionally, a few apomictic individuals can arise within sexual lineages (blue framing; e.g., in polyploid sexuals, Masci et al., 1994). However, the practical implementation of the proposed species concept of Hörandl (2018) including multi-disciplinary (phylo)genomic approaches and taxonomic re-circumscriptions have not yet been conducted due to sequencing, bioinformatic, and data (ploidy and reproduction mode) limitations for analyzing young, reticulate evolutionary processes within species complexes.

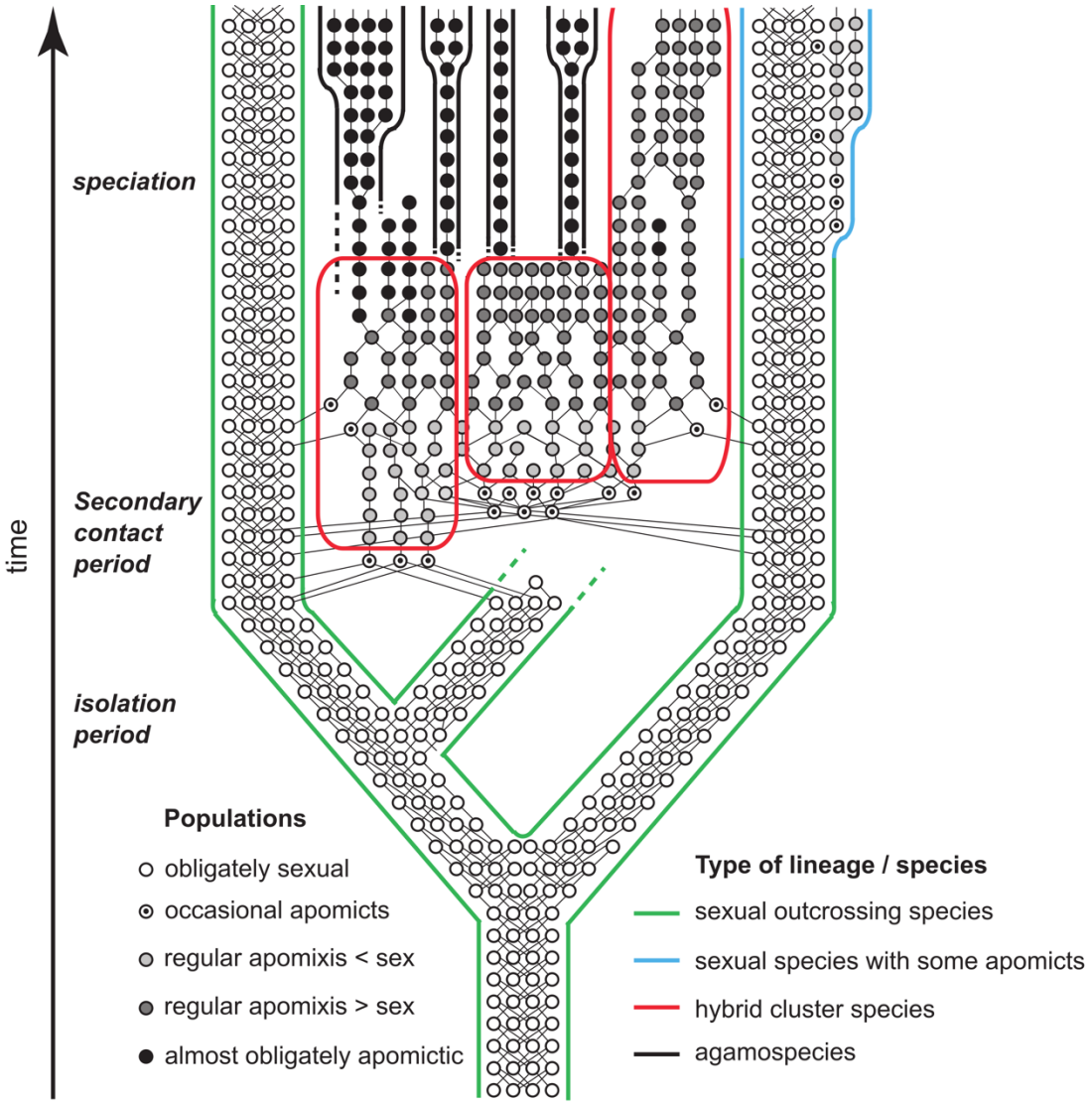


Figure 2. Scheme illustrating the rise of a (polyploid) apomictic species complex including accompanying evolutionary processes (slightly modified figure from Hörandl 2018; published under Attribution-ShareAlike 4.0 International, CC BY-SA 4.0). Populations are represented by circles. Please see the General Introduction, section The Evolution of Polyploid Apomictic Species Complexes for detailed explanations.

Relationships between Polyploidy, Apomixis, and Environment

Apomicts frequently occupy larger areas, more northern regions (in the Northern Hemisphere), higher altitudes, or more extreme environments compared to their sexual relatives, a phenomenon called geographical parthenogenesis (GP; Vandel, 1928; Bierzychudek, 1985; Hörandl, 2006; Lo et al., 2013; Hojsgaard & Hörandl, 2019; see Kearny, 2005 for animals). GP patterns usually have a Pleistocene context: Climatic range shifts in temperate to boreal zones

offered frequent opportunities for interspecific hybridization, which probably triggered the rise of apomixis (Carman, 1997; Dobeš et al., 2004; Pellino et al., 2013; Barke et al., 2018). Apomicts often colonize habitats severely affected by past glacial periods (Late Pleistocene; Bierzychudek, 1985; Mráz et al., 2009; Hartmann et al., 2017). Factors shaping GP patterns are complex (Hörandl, 2006) and still controversially discussed.

On the one hand, GP has been widely attributed to the advantages of apomictic polyploids. As already mentioned above, the presence of multiple gene copies in polyploids provides more physiological and phenotypic flexibility to respond to different environmental conditions (Comai, 2005; Hörandl, 2006; Van de Peer et al., 2017). Apomicts frequently show high proportions of heterozygosity not only on the populational but also on the individual level, attributable to their allopolyploid origin, fixed heterozygosity at duplicated loci at homeologous chromosomes, and allelic sequence divergence in highly polyploid obligate apomictic lineages (i.e., Meselson effect; Grant, 1981; Gornall, 1999; Welch & Meselson, 2000; Brochmann et al., 2004; Hörandl, 2006; Hörandl & Paun, 2007; Pellino et al., 2013; Mohammadin et al., 2018). Nevertheless, mutation accumulation in asexual genomes has been thought to drive these individuals or species to extinction in the long term (Muller's Ratchet and the "dead-end" hypothesis; Muller, 1964; Mayrose et al., 2015). Interestingly, no genome-wide mutation accumulation in asexual oribatid mites, and in facultative apomictic polyploid plants was observed (Pellino et al., 2013; Brandt et al., 2017; Hodač et al., 2019). Moreover, plants can benefit from increased heterozygosity, for example, by novel genetic combinations, heterosis effects, masking of deleterious mutations, or changes in secondary metabolites relevant to biotic interactions (Soltis & Soltis, 2000; Comai, 2005; Qiu et al., 2020). Increased heterozygosity might enable the spreading of polyploids towards more variable or extreme climatic conditions (Hörandl, 2006; Rice et al., 2019).

Uniparental reproduction is also an important factor considered for range expansion of apomicts. Due to self-compatibility and no need for pollination vectors, apomicts are good colonizers (Stebbins, 1950; "Baker´s law": Baker, 1967; Hörandl et al., 2008; Tilquin & Kokko, 2016). A single or few individuals can found a new population after long-distance dispersal, without the detrimental effects of inbreeding depression. Uniparental reproduction is therefore particularly advantageous in extreme environments with pollinator limitations or fluctuations (e.g., in the Alps or badlands; Hörandl, 2006; Schinkel et al., 2016; Karbstein et al., 2019). In addition, environmental niche shifts caused by (allo)polyploidy effects also explain the differing distributions of sexuals and apomicts. Apomicts are able to shift towards more anthropogenic habitats (Asker & Jerling, 1992; Hörandl & Paun, 2007), colder high-mountain

areas (Kirchheimer et al., 2016; Schinkel et al., 2016), or drier and more temperately seasonal sites (Coughlan et al., 2017).

However, in general, complex interactions of heterozygosity based on genome-wide data, ploidy, reproduction modes (sexual versus apomictic), and climatic environmental factors shaping GP patterns of large species complexes have not yet been investigated. Despite the manifold advantages of polyploid apomicts in GP scenarios, the breeding system of sexual species potentially also explains their small, restricted ranges. Sexual progenitors of apomictic derivatives are usually self-incompatible (Hörandl, 2010; Cosendai et al., 2013; Alonso-Marcos et al., 2018; Vašková & Kolarčik, 2019). In small and/or isolated populations, the restricted availability of conspecific mating partners and pollen vectors, risk of inbreeding depression, and loss of genetic diversity are problematic (Schleuning et al., 2009; Freeland et al., 2011). Nevertheless, potential negative effects of the breeding system of sexual species on fitness and genetic diversity in GP scenarios remain to be thoroughly researched.

Current NGS Developments and Approaches to Unravel Evolutionary Processes in Polyploid Complexes

"Polyploid phylogenetics (...) revolves around inferring the full phylogenetic history of relatively recently formed polyploid lineages (...) and their relatives, and the application of those inferences to investigating the myriad processes related to polyploid macroevolution"
(Carl J. Rothfels; Rothfels, 2021)

It's an exciting era for plant biologists: The current transition from genetics to genomics has been accompanied by remarkable breakthroughs in plant evolutionary (phylo)genomics, due to novel approaches in massively parallel DNA sequencing, computation capabilities, and bioinformatics (Wendel, 2015; Rothfels et al., 2017; Rothfels, 2021). While only a few years ago, the application of 100, ten, or fewer markers was common (e.g., Hörandl & Paun, 2007; Freeland et al., 2011; Barke et al., 2018; Karbstein et al., 2019, 2020a), nowadays the transition toward analysis of far more than thousands of markers or the entire plastomes, transcriptomes, or genomes has been made ("OMICS" data; e.g., Pellino et al., 2013; Gordon et al., 2020; Wagner et al., 2020). At the moment, two main approaches of genome subsampling are commonly used: Restriction site-associated DNA sequencing (RADseq; and similar methods,

e.g., genotyping-by-sequencing, GBS) and target enrichment (hybrid capture) of nuclear genes. Both approaches have enabled the clarification of diploid (and some polyploid) below-species, species- and genus-level relationships, from less than a few million to 50 million years (Baird et al., 2008; Folk et al., 2015; Andrews et al., 2016; Schmickl et al., 2016; Eaton et al., 2017; Eriksson et al., 2020; Wagner et al., 2020). Both RADseq and target enrichment have their advantages and disadvantages. RADseq covers non-coding and coding regions across the genome, provides much more loci and SNPs, and is less costly and work-intensive allowing larger sample sets (Andrews et al., 2016; McKain et al., 2018). Nevertheless, with increasing species divergence, RADseq locus dropout due to mutations in enzyme cutting sites can become problematic in phylogenetic reconstructions (Andrews et al., 2016; Eaton et al., 2017). Target enrichment focuses on sequencing of hundreds of low-copy nuclear coding genes, which are more conservative and better suitable to resolve within-genus or among-genera relationships (Schmickl et al., 2016; McKain et al., 2018; Carter et al., 2019). Loci assembled from target enrichment methods are often longer, allowing for the segregation of alleles at a single locus (i.e., phasing) and thus enabling analytical methods based on the multi-species coalescent model (MSC; Rannala & Yang, 2003; Yang & Rannala, 2014; McKain et al., 2018).

Retrieving allelic information is particularly crucial for inferring evolutionary processes in highly reticulate, polyploid, and evolutionary young species complexes (Oxelman et al., 2017; Eriksson et al., 2018; Rothfels, 2021). Polyploid phylogenetic relationships and genome evolution are not shaped by the evolutionary origin and the genomic contributions of progenitors, they are also influenced by post-origin processes (e.g., biased fractionation or gene flow), resulting in a mosaic-like genomic pattern with parental, additive, and novel features (Sochor et al., 2015; Soltis et al., 2015; Hörandl, 2018; Carter et al., 2019). To unravel the evolutionary processes in polyploids, a multidisciplinary approach including tree, genetic structure, and network methods based on phased genes and SNP data has to be implemented (e.g., Lo et al., 2010; Brandrud et al., 2020).

Trees can build a first phylogenetic framework of the species complex. Autopolyploid evolution usually results in a tree-like, non-conflicted pattern whereas network-like allopolyploid evolution leads to decreased support values in tree reconstructions (Lo et al., 2010; Pease et al., 2018; Brandrud et al., 2020). Nevertheless, standard ML and coalescent tree reconstructions combine different evolutionary histories (subgenomes) in consensus sequences or trees, which under certain circumstances lead to unreliable results (McDade, 1992; Oxelman et al., 2017; Rothfels, 2021). Consequently, non-tree-like approaches should be applied for inferring evolutionary processes in young polyploid complexes (McDade, 1992; Huson & Bryant, 2006;

Rothfels, 2021). Several genetic structure and delimitation methods appropriate for polyploid analyses have appeared recently, for instance, a genetic similarity approach incorporating all SNPs across all phases per locus and across all loci (Malinsky et al., 2018) or a multispecies coalescent model using (un)phased genes to delimit genetic structure and species (Jones, 2017a). However, phylogenetic network methods are most suitable, modeling the reticulation based on maximum pseudolikelihood from SNP-based multilocus sequence data or gene trees under a coalescent model (e.g., PhyloNetworks or PhyloNet; Than et al., 2008; Solís-Lemus et al., 2017; Olave & Meyer, 2020). Despite recent efforts in subgenome assignments and accurate allopolyploid network reconstruction (e.g., Jones, 2017b; Lautenschlager et al., 2020; Freyman et al., 2020; Šlenker et al., 2021), more sophisticated approaches remain impractical for large datasets and young polyploid complexes. Low genetic divergence among or missing knowledge on progenitors (e.g., delimitation, contributing numbers, or extinction events), ploidy levels, and polyploid formation types of apomictic polyploid derivatives are potential pitfalls for these kinds of analyses. Therefore, young reticulate phylogenetic relationships, genome composition and evolution of large polyploid plant species complexes have not yet been elucidated.

The *Ranunculus auricomus* Complex

Ranunculus auricomus is a large polyploid apomictic plant species complex. It is well-suited to study the aforementioned issues and fill the gaps in our understanding of young polyploid evolution. A great deal of research has been conducted on the goldilock buttercups within the last decades, rendering it a model system in apomixis and in evolution and biogeography of young plant groups (e.g., Nogler, 1984; Hörandl, 1998; Hörandl et al., 2009; Hodač et al., 2014, 2018; Barke et al., 2018, 2020; Dunkel et al., 2018). The group is monophyletic within the genus *Ranunculus*, and close relatives are located in North America and Central Asia (Emadzade et al., 2011, 2015).

The *R. auricomus* complex is distributed across Eurasia, ranging from Greenland, Europe, to Siberia (and Alaska), and spanning arctic, boreal, temperate, and Mediterranean climatic zones (Jalas & Suominen, 1989; GBIF Secretariat, 2017). Species inhabit a broad range of habitats in Europe (Fig. 3a–c). They occupy stream- and riversides, alluvial and humid deciduous forests, forest edges, extensively used, swampy to humid (to semi-dry) meadows, hedges and shrubberies, and waysides. Population sizes range tremendously according to the habitat type

and quality, from a few to thousands of individuals (see legend of Fig. 3), whereby humid, semi-shaded, and extensively used habitats are preferred.

Currently, the complex comprises more than 840 taxa (Euro+Med database, Raab-Straube et al., 2014; Leslie, 2018), which were mainly morphologically recognized (e.g., Julin 1967, 1980; Marklund, 1961, 1965; Hörandl & Gutermann, 1998a,b, 1999; Dunkel et al., 2018). Linnaeus (1753) recognized the existence of two remarkably different morphotypes, *R. auricomus* from Western Europe with dissected basal leaves and *R. cassubicus* from North Poland or further east (Siberia) with undivided basal leaves (Kvist, 1987; see also Fig. 4 for basal leaf morphology). Afterward, hundreds of variants have been described, interconnecting these core morphotypes by endless hybridogenic intermediates. Consequently, a concept of "main species" and apomictic lineages as subspecies (Marklund, 1961, 1965) failed and was no longer formally accepted (Ericsson, 1992; Hörandl, 1998). The majority of species (ca. 600) were described in the 20th century from Northern Europe (Marklund, 1961, 1965; Julin, 1980; Ericsson, 1992), although many, partly regional, species have been described in Central, Eastern, and Southern Europe more recently (e.g., Soó, 1964; Borchers-Kolb, 1985; Hörandl & Gutermann, 1998a,b; Dunkel, 2010, 2015, 2021; Dunkel et al., 2018).

The evolutionary history of the *R. auricomus* complex is strongly shaped by hybridization and polyploidization: The great number of polyploid agamospecies (>800) has been hypothesized to originate from hybridization of sexual progenitor species (Paun et al., 2006b; Hörandl et al., 2009; Hodač et al., 2014, 2018). Only twelve sexual, mainly diploid morphospecies have been described so far, occurring allopatrically in small ranges across Southern, Central, and Eastern European mountain systems. The diploid species *R. cassubicifolius* W.Koch (Koch, 1939), *R. carpaticola* Soó (Soó, 1965), and *R. flabellifolius* Rchb. (Reichenbach, 1832) occur in Central and Eastern Europe and are characterized by undivided basal leaves (Fig. 5a,b). *Ranunculus flabellifolius* differs from the other species through the occasional appearance of three-lobed basal leaves and a fan-shaped stem leaf (connected segments; the other species have non-connected broad lanceolate segments, Fig. 5c). In detail, *R. cassubicifolius* is distributed along stream- and riversides and meadows at the northern and eastern edges of the Alps (also some autotetraploid populations; Hörandl & Gutermann, 1998a; Hörandl & Greilhuber, 2002; Dunkel, 2010); *R. carpaticola* inhabits humid deciduous forests in the Carpathians (Paun et al., 2006a,b; Hörandl et al., 2009); and *R. flabellifolius* is restricted to a few locations in the Southern Carpathians (and already hybridizes with adjacent populations of the *R. auricomus* complex; Reichenbach, 1832; Jalas & Suominen, 1989; Dunkel et al., 2018).

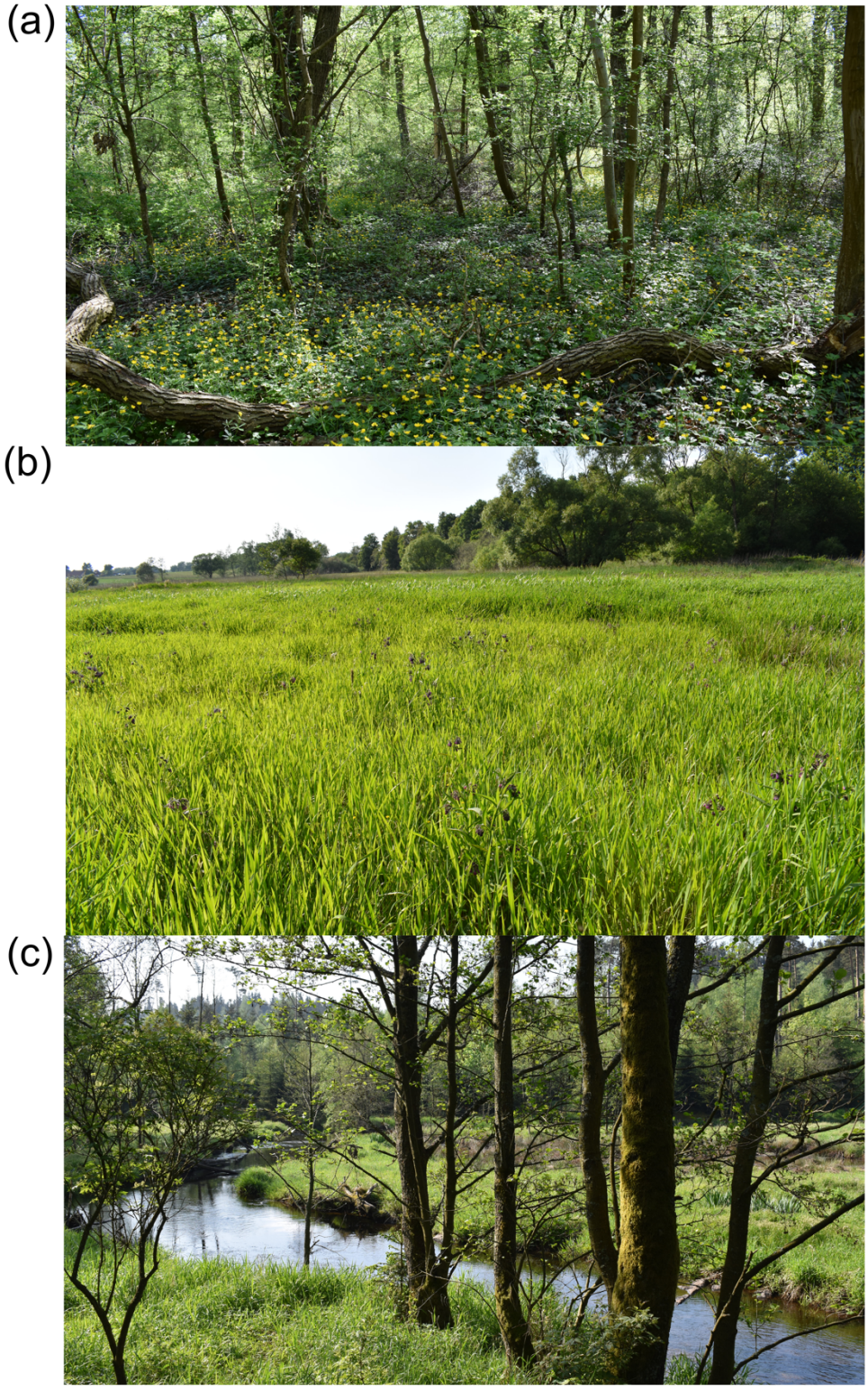


Figure 3. Common habitats of taxa within the *R. auricomus* complex across Central Europe. (a) Large population (>1000 individuals) growing in a semi-shaded, semi-humid mixed forest in Alsace, France (population ID "KK099", tetraploid apomict "*R. x alsaticus*"), (b) medium-sized population (<1000 individuals) occupying a swampy, extensively used meadow in Eastern Bavaria, Germany (population ID "KK216", tetraploid apomict "*R. x indecorus*"), and (c) small population (<50 individuals) along way- and streamsides in Northern Bavaria, Germany (population ID "KK055", tetraploid apomict "*R. x phragmiteti*"). Population sizes were recorded during field work. Photographs: Kevin Karbstein

The remaining sexual species possess a heterophyllous leaf cycle with dissected basal leaves at anthesis and narrow stem leaf segments (Fig. 5a–c). The diploids *R. envalirensis* Grau (Grau, 1984) and *R. cebennensis* Dunkel (Dunkel et al., 2018), and the only tetraploid sexual *R. marsicus* Guss. & Ten. (Tenore, 1835–1836) inhabit small, restricted ranges in the Southern European mountain systems. The dwarf species *R. envalirensis* occurs in subalpine meadows in the Eastern Pyrenees (Spain, Andorra, and France) and is characterized by basal leaves with three- to five-lobed or -dissected, roundish segments and linear stem leaf segments (Grau, 1984; Dunkel et al., 2018); *R. cebennensis* is only known from two meadows in the Massif Central (France) and is morphologically similar to *R. envalirensis* (differing, however, in plant height; Dunkel et al., 2018); the dwarf species *R. marsicus* inhabits subalpine meadows in the Central Apennines and differs from the other species through its linear stem leaf segments with sinuses (Fig. 4c; also tri- to heptaploid cytotypes known; Masci et al., 1994; Dunkel, 2011; Paule et al., 2018; Dunkel et al., 2018). The other diploid dissected basal leaf species are distributed, partly sympatrically, in humid deciduous forests and meadows in the Illyrian lowlands: *Ranunculus notabilis* Hörandl & Guterm. in Southeastern Austria (Hörandl & Gutermann, 1998b), and *R. austroslovenicus*, *R. calapius*, *R. mediocompositus*, *R. peracris*, and *R. subcarniolicus* Dunkel in Slovenia and Western Croatia (Dunkel et al., 2018). These species are larger and have rather narrowly lobed or dissected basal leaf segments compared to the previously mentioned high-mountain species.

Using traditional genetic markers (isoenzyme, nuclear ITS, plastid, and AFLP data), a deep split between *R. notabilis* and *R. cassubicifolius* (incl. *R. carpaticola*) was estimated to have occurred ca. 900,000 years ago, whereas the morphologically similar undivided basal leaf species *R. cassubicifolius* and *R. carpaticola* split ca 300,000 years ago (Hörandl, 2004; Hörandl et al., 2009; Hodač et al., 2014). However, all described morphospecies have not been investigated by (phylo)genomic methods so far. Moreover, the morphological differentiation among the recently described Illyrian species but also among all described morphospecies is unclear and no identification key is available. Geometric morphometric analyses including the most taxonomically informative traits, and species circumscriptions based on genetic lineage or clusters concepts are thus needed.

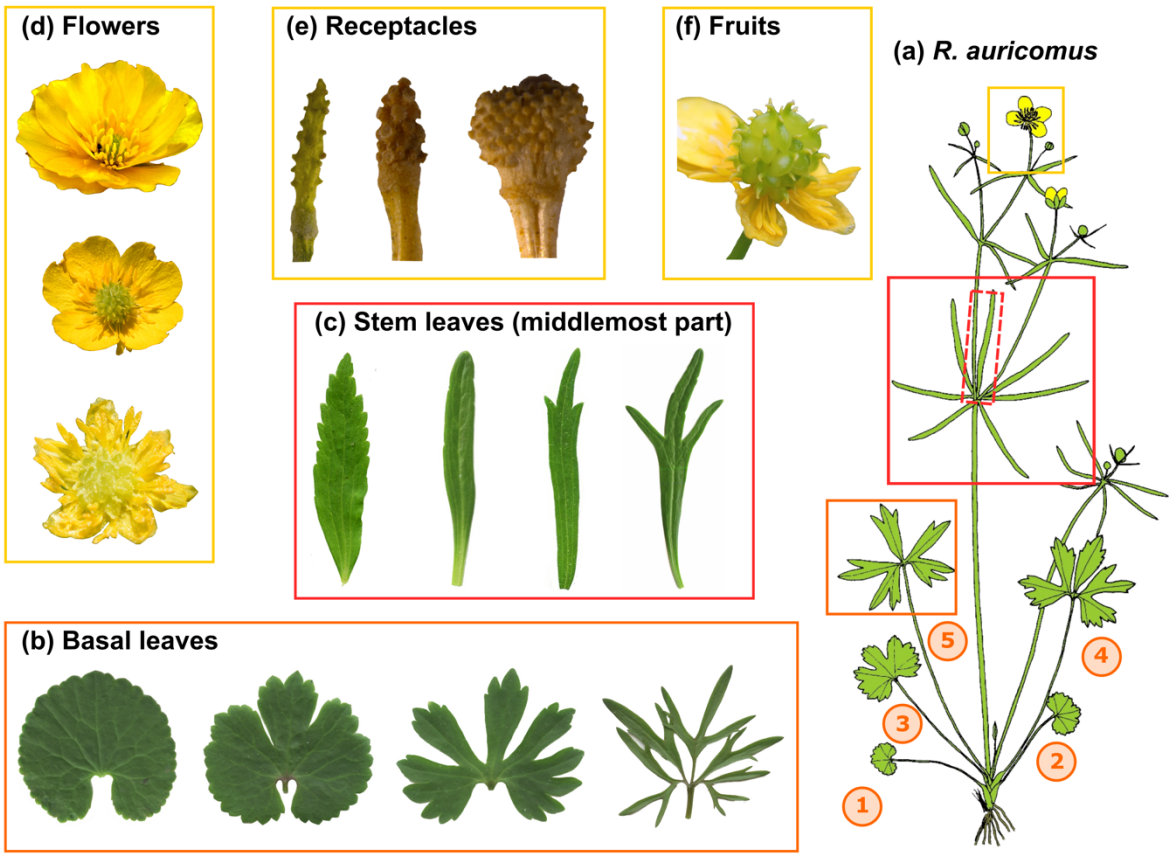


Figure 4. The most taxonomically informative morphological traits of *R. auricomus* individuals (see also Fig. S4 of Karbstein et al. 2020b for leaf venations and homologous regions per trait). (a) Scheme of a *R. auricomus* individual with highlighted (framed) morphological traits (figure source: Hörandl & Gutermann, 1995, edited by Mareike Daubert): basal leaf cycle (orange framing; 1–5, 5 = most dissected leaf at anthesis) and stem leaves (red framing) with middlemost segment (red-dotted framing), and reproductive structures (yellow framing; flower, receptacle at the fruit stage, and fruits). (b) Basal leaf variation from undivided (left) to broad three-lobed or dissected (center) to strongly dissected forms (right). (c) Variation of the middlemost segment of a stem leaf from broad lanceolate (with teeth; left) to narrowly linear forms (with dissection; right). (d) Petal or petaloid nectary leaf variation from five to 15 (rather sexuals; the uppermost ones) or reduced or absent forms (rather polyploid apomicts; the lowermost one). (e) Receptacle variation from narrowly large (left) to smaller roundish ones (right). (f) Fruit (achene) in ripening process: In polyploid apomicts, one fruit can contain sexually and apomictically produced seeds. Photographs (b–f): Kevin Karbstein.

In contrast to sexuals, the more than 800 described polyploid apomicts are frequently found across the entire range of the complex, and towards more northern or past glaciated regions (Jalas & Suominen, 1989). Apomicts are mainly tetra- or hexaploid (2n = 16 and 4n/6n = 30–34/44–48 chromosomes), whereas other ploidy levels are less frequent (tri-, penta-, hepta-, and docaploids; Paule et al., 2018). Gametophytic apomixis was confirmed by embryological and cytological studies (Häfliger, 1943; Rousi, 1956; Nogler, 1984; Barke et al., 2018), and

supported by pollen quality measurements (Hörandl et al., 1997; Hörandl & Gutermann, 1998a,b, 1999; Dunkel, 2010, 2011; Dunkel et al., 2018). The extraordinarily high morphological diversity of polyploid apomicts, ranging from undivided to strongly dissected basal leaves, from broad lanceolate to linear stem leaf segments, from narrow to roundish receptacles, and complete to reduced flowers (Fig. 5a–e), probably resulted from hybridization of sexual progenitor species. Hodač et al. (2018) experimentally showed by crossing experiments that F_2 hybrids of the morphologically most divergent sexual species segregate in a Mendelian manner, covering a broad spectrum of described polyploid apomictic leaf phenotypes. In addition, the apomicts *R. hungaricus* and *R. variabilis* are probably of allopolyploid origin (Hörandl et al., 2009; Hodač et al., 2014). Some apomicts are even less than 100,000 years old, as estimated from transcriptomic data (Pellino et al., 2013). However, despite recent efforts, the phylogenetic relationships, genome composition and evolution, and biogeography of the apomictic *R. auricomus* taxa is unkown.

Aims and Scope

In this thesis, I investigated the polyploid apomictic *R. auricomus* complex comprising twelve described sexual morphospecies and more than 800 polyploid agamospecies (Raab-Straube et al., 2014; Dunkel et al., 2018). Only a very limited knowledge about phylo(genetic) relationships among sexuals and the origin of polyploid derivatives existed at the beginning of the project (e.g., Hörandl et al., 2009; Hodač et al., 2014). For this, I employed **multidisciplinary approaches** ranging from field observations, crossing experiments, geometric morphometrics, flow cytometry, DNA sequencing to statistical modeling, to unravel the evolution and biogeography from the progenitors to the polyploid deviates of the complex. My **main research questions** are: (i) How many well-defined progenitor species are recognizable? (ii) Which genomic and environmental factors shape the European GP pattern of the complex? (iii) Are RADseq and nuclear gene supported by plastome data able to resolve the origin of the less than 1.0 million-year-old complex?

I mainly addressed the questions by using genomic RADseq (97,312 loci) and nuclear gene (663) data applying current concatenation- and coalescent-based methods. Plastome (71 regions), ploidy (1,474 leaves), reproductive (4,669 seeds), crossing (481 bagged flowers), and geometric morphometric (1,593 leaves and receptacles) data supported nuclear data. Finally, I

included 251 populations and 87 *R. auricomus* taxa sampled Europe-wide. The data were incorporated into **bioinformatic workflows** accounting for low genetic divergence, gene flow, ILS, hybridization, and polyploidy, and results were interpreted and discussed in the context of evolutionarily young (Pleistocene) speciation processes.

First, I developed a workflow to analyze the twelve described sexual morphospecies by RADseq loci, nuclear genes, and a concatenated trait geometric morphometric dataset, yielding a clear sexual species delimitation and a final taxonomic treatment (*Chapter 1*). Further, I analyzed data from population crossing experiments to elucidate the role of the sexual breeding system in GP scenarios and to justify the previous taxonomic treatment (*Chapter 2*). I supported analyses on absolute divergence time and ancestral range reconstruction based on nuclear genes to understand speciation patterns of the sexual progenitors (*Chapter 3*). Then, I developed a comprehensive path analysis incorporating sexuality of populations, ploidy, genome-wide heterozygosity, reproductive, and climatic environmental factors to unravel the factors shaping the European GP scenario of the complex (*Chapter 4*). Lastly, for the first time, I disentangled basic phylogenetic relationships, genome composition and evolution of the polyploid complex using all previously gathered multidisciplinary data combined in a workflow (*Chapter 5*).

More Specific Summaries of the Chapters

Chapter 1 – The Sexual Progenitor Species

Species are the fundamental units of biodiversity and evolution (Sukumaran & Knowles, 2017). However, their definition continues to present a persistent challenge for evolutionary biologists (Freudenstein et al., 2017; Hörandl, 2018). Delimiting the sexual progenitor species of young, fast-evolving species complexes is difficult due to low genetic divergence, ongoing gene flow, and ILS (e.g., Burgess et al., 2015). Here, I used concatenation-based methods for RADseq (RAXML, Quartet Sampling), coalescent-based methods for nuclear genes (ASTRAL), and genetic structure, distance, and delimitation methods (STRUCTURE, SPLITSTREE, and STACEY) supported by geometric morphometrics of the taxonomically most informative traits. I found five genetic main clades or clusters across datasets and methods, characterized by an allopatric European distribution and clearly distinguishable morphotypes. These main clades or clusters were accepted as species and used as progenitor species in polyploid analyses.

Chapter 2 – The Breeding System of Diploid Sexuals

Asexuals often have a larger distribution or are distributed towards more northern regions compared to their sexual relatives - a phenomenon called geographical parthenogenesis (GP; Vandel, 1928; Bierczychudek, 1985; Hörandl, 2006). Advantages of asexuals have been discussed in detail to explain GP patterns (Hörandl, 2006; Lo et al., 2013; Paule et al., 2018), whereas little attention has been paid to potential disadvantages of the sexual breeding system on fitness and genetic diversity. I evaluated fitness measurements based on crossing experiments combined with genetic distances and heterozygosity (genomic RADseq data). Experiments demonstrated full-crossability among the formerly described *R. notabilis* s.l. morphospecies, justifying the previous taxonomic treatment. Interestingly, I detected semi-self compatibility in sexuals, and negative inbreeding and positive outbreeding effects indicated by higher seed sets of more distant or heterozygous lineages. Small and/or isolated sexual populations can suffer from decreased female fitness in GP scenarios, probably limiting (among other factors) their range expansion.

Chapter 3 – The Biogeography of Diploid Sexuals

In the Northern Hemisphere, Pleistocene climatic fluctuations caused periodic species range expansions and contractions (e.g., Taberlet et al., 1998; Magri et al., 2006; Magri, 2008). However, the understanding of speciation patterns of temperate-montane herbs within this period is still limited. I supported analysis of absolute divergence time and ancestral range reconstructions based on nuclear genes (target enrichment). Speciation events occurred 830–580 thousand years ago within a period of severe climatic deteriorations (Mid-Pleistocene Transition). Vicriariance processes probably triggered allopatric speciation from a Europe-wide forest-understory ancestor. The study highlights the speciation processes in a period of tremendous climatic changes, where temperate-montane sexuals were potentially confined in refugia during glacial periods and hampered in spreading by apomictic derivatives (or other congeners) during interglacial periods.

Chapter 4 – A Geographical Parthenogenesis Scenario

In GP scenarios, the success of apomicts depends on a combination of internal (e.g., ploidy, facultative apomixis, or genetic diversity) and external factors (habitat or climate; Bierzychudek, 1984; Hörandl, 2006; Lo et al., 2013). Factors are still under debate, and the

complex effects of ploidy, genome-wide genetic diversity (RADseq), reproduction mode (sexual versus apomictic), and environment on plant sexuality (percent of sexual seeds) are still poorly understood. I combined all mentioned factors in a genetically informed path analysis based on Generalized Linear Mixed Models (GLMMs). This revealed higher sexuality, more petals, and up to three times less heterozygosity in diploids compared to polyploids. In contrast to other GP scenarios, sexuality was positively related to solar radiation and temperature stability (fitting the more southern distribution of sexuals), and heterozygosity showed an association with temperature seasonality (potentially higher climatic fluctuation tolerance). The study provides empirical evidence that GP scenarios are shaped by genomic features and large-scale climatic environmental gradients.

Chapter 5 – The Origin of the Polyploid Complex

Polyploidy and hybridization (i.e., allopolyploidy) are particularly likely to create biotypes with novel genomic compositions, and are important for subsequent speciation (Abbott et al., 2013; Blischak et al., 2018; Rothfels, 2021). However, up to now, large-scale phylogenetic relationships, genome composition and evolution of large polyploid complexes have yet not been deciphered. I used all previously gathered data, and created a bioinformatic workflow based on phylogenomic tree (RAXML-NG and ASTRAL), structure (RADPAINTER, SNMF, and STACEY), and network (PHYLONETWORKS and PHYLONET) methods applied to genomic dataset (RADseq, nuclear genes, and plastome regions). This revealed predominantly allopolyploid origins. Hybridizations involved each two to three different diploid progenitor subgenomes. Allopolyploids are characterized by enormous post-origin evolution, probably due to Mendelian segregation, backcrossings, and gene flow among facultative apomicts. Surprisingly, analyses unveiled a now extinct progenitor species. Most included agamospecies were non-monophyletic.

In the *General Discussion*, I combine the results of my thesis with a comprehensive and comparative overview of existing plant studies on phylogenetics, biogeography, and composition and genome evolution of young species complexes, regarding the mainly diploid sexual progenitors and the predominantly apomictic allopolyploid derivatives characterized by subgenome dominance. I explain how these processes shape the diversity of polyploid plant complexes including taxonomic challenges and further micro- and macroevolutionary

processes. Finally, I present conclusions and an outlook of the research project and the field of polyploid phylogenetics.

II Research Chapters

Chapter 1 – The Sexual Progenitor Species

Phylogenomics supported by geometric morphometrics reveals delimitation of sexual species within the polyploid apomictic *Ranunculus auricomus* complex (Ranunculaceae)

Karbstein, K., Tomasello, S., Hodač, L., Dunkel, F. G., Daubert, M., & Hörandl, E.

Taxon, *69*, 1191–1220 (2020b).
doi: 10.1002/tax.12365 (CC BY 4.0, changes in format and citation style)
BioRxiv, 1–58 (2020).
doi: 10.1101/2020.01.07.896902 (reuse allowed by author)

Species are the basic units of biodiversity and evolution. Nowadays, they are widely considered as ancestor-descendant lineages. Their definition remains a persistent challenge for taxonomists due to lineage evolutionary role and circumscription, i.e., persistence in time and space, ecological niche, or a shared phenotype. Recognizing and delimiting species is particularly methodically challenging in fast-evolving, evolutionary young species complexes often characterized by low genetic divergence, hybrid origin, introgression, and incomplete lineage sorting. *Ranunculus auricomus* is a large Eurasian apomictic polyploid complex that probably has arisen from the hybridization of a few sexual progenitor species. However, even delimitation of and relationships among diploid sexual progenitors are unclear, ranging from 2 to 12 species. Here, we present an innovative workflow combining phylogenomic methods based on 86,782 parameter-optimized RADseq loci and target enrichment of 663 nuclear genes accompanied by geometric morphometrics to delimit sexual species in this evolutionary young complex (<1 Mya). For the first time, we revealed a fully resolved and well-supported maximum likelihood tree phylogeny congruent to neighbor-net network and STRUCTURE results based on RADseq data. In a few clades, we found evidence of discordant patterns indicated by quartet sampling, and reticulation events in the neighbor-net network

probably caused by introgression and incomplete lineage sorting. Together with coalescent-based species delimitation approaches based on target enrichment data, we found five main genetic lineages, with an allopatric distribution in Central and Southern Europe. A concatenated geometric morphometric dataset including data from basal and stem leaves, as well as receptacles, revealed the same five main clusters. We accept those five morphologically differentiated, geographically isolated, genetic main lineages as species: *R. cassubicifolius* s.l. (incl. *R. carpaticola*), *R. envalirensis* s.l. (incl. *R. cebennensis*), *R. flabellifolius*, *R. marsicus*, and *R. notabilis* s.l. (incl. *R. austroslovenicus*, *R. calapius*, *R. mediocompositus*, *R. peracris*, and *R. subcarniolicus*). Our comprehensive workflow combining phylogenomic methods supported by geometric morphometrics proved to be successful in delimiting closely related sexual taxa and applying an evolutionary species concept. This workflow is also applicable to other evolutionarily young complexes.

Keywords

Europe; geometric morphometrics; RADseq, *Ranunculus auricomus*; sexual species; target enrichment

Introduction

Species are the basic units of biodiversity and evolution, but their definition remains a persistent challenge for taxonomists (Sukumaran & Knowles, 2017). Most modern authors consider species as ancestor-descendant lineages (De Queiroz, 2007; Sukumaran & Knowles, 2017), a concept that applies to both sexual and asexual lineages. Nevertheless, it is rather challenging to define their evolutionary role and circumscription, i.e., persistence in time and space, ecological niche, or a shared phenotype (Freudenstein et al., 2017; Hörandl, 2018). Hörandl (2018) proposed a species delimitation workflow based on four main principles particularly addressing apomictic polyploid complexes: First, separate and classify the obligate sexual progenitor species; second, merge sexual progenitors and highly facultative apomictic lineages into one species; third classify the main clusters of the remaining facultative apomicts as species; and fourth, treat obligate apomictic lineages as agamospecies.

However, recognizing and delimiting even the sexual progenitor species is methodically challenging in fast-evolving, evolutionary young species complexes (e.g., Burgess et al., 2015). Commonly, they are characterized by low genetic divergence, occasionally by hybrid origin, persistent gene flow, introgression, and incomplete lineage sorting (ILS). During speciation processes, ILS is rather the rule than an exception causing incongruences in phylogenetic reconstructions (see, e.g., Oliver, 2013; Pease et al., 2018). The progenitor species can also be sexual polyploids. Natural hybridization and polyploidy are important evolutionary processes and are regarded as key factors for diversification of angiosperms (Soltis & Soltis, 2000; Jiao et al., 2011; Alix et al., 2017). Reticulate evolution (e.g., hybridization, allopolyploidy, and ILS), however, results in the phenomenon that gene trees do not match species trees (e.g., McBreen & Lockhart, 2006; Pirie et al., 2009).

The application of molecular markers has helped to reconstruct not only relationships of sexual progenitor species but also hybrid origin, reticulate relationships, and population genetic structure of apomictic lineages within species complexes (Dobeš et al., 2004; Hörandl, 2004; Paun et al., 2006a,b; Kirschner et al., 2015). Nevertheless, most complexes are evolutionarily young, and many originated or diversified in the context of Pleistocene glaciations. Therefore, genetic divergence among taxa is usually quite low. Traditional genetic markers (for example, internal transcribed spacer [ITS] markers) often failed to reconstruct fully resolved phylogenies (e.g., Hörandl et al., 2009; Krak et al., 2013; Kirschner et al., 2015). Population genetic markers like microsatellites or amplified fragment length polymorphisms have been widely used, but are often not homologous among more divergent taxa (Freeland et al., 2011). Hence, species-level relationships in apomictic complexes often remained unresolved.

The application of phylogenomic methods provides a magnitude more loci compared to traditional genetic markers and can resolve relationships of taxa with less than a few million (or even less than 100,000) years of divergence (Pellino et al., 2013; Tripp et al., 2017; Tomasello et al., 2020). Currently, two main approaches are used in species-level systematics: Target enrichment and restriction site-associated DNA sequencing (RADseq; and the similar genotyping-by-sequencing [GBS] method). Target enrichment (i.e., nuclear genes selected out of a reference genome) is most useful for diverged species and genus-wide relationships (Folk et al., 2015; Schmickl et al., 2016; Tomasello, 2018). RADseq has repeatedly proven its efficiency in resolving intraspecific relationships, phylogenies of closely related species, but also relationships within genera (relationships between less than 100,000 and more than 50 million years; Baird et al., 2008; Hipp et al., 2014; Cavender-Bares et al., 2015; Tripp et al., 2017; Wagner et al., 2018; Pätzold et al., 2019, He et al., 2020). Although RADseq yields much

more loci and single nucleotide polymorphisms (SNPs) than target enrichment, it is considered to become less informative with increasing species divergence due to increasing mutations in enzyme cutting sites leading to loss of loci (Rubin et al., 2012; Eaton et al., 2017). Specific pipelines were developed for processing (filtering and assembling) of RADseq data, e.g., the IPYRAD pipeline that accounts for indels and therefore is suited for phylogenetic analysis (Eaton, 2014; Eaton & Overcast, 2020). Several parameters can be optimized to find homologous loci and to maximize the true phylogenetic signal leading to fully resolved topologies (Paris et al., 2017; McCartney-Melstad et al., 2019; Pätzold et al., 2019). Since incongruences due to reticulate evolution are hard to determine in concatenated alignments, new methods have been developed to examine the phylogenetic tree topologies for conflicted and/or poorly informed branches (e.g., quartet sampling [QS] method; Pease et al., 2018).

With a target enrichment approach, single-copy loci characterized by a certain amount of variation are selected beforehand based on their position in the genome. Loci obtained by target enrichment are usually longer compared to those from RADseq, allowing for gene tree estimation and allele phasing. The importance of the latter has been highlighted in recent studies (Eriksson et al., 2018; Andermann et al., 2019). They demonstrate the advantages derived by retrieving allele information for a correct phylogenetic inference, particularly in recently diverged species complexes where reticulate evolutionary processes play a tremendously important role. The lack of the need for concatenation makes the employment of coalescent-based methods for species tree and species delimitation inference possible. These methods are particularly useful because they are able to accommodate for all the stochastic processes responsible for incongruence among gene genealogies or between genealogies and species phylogenies (e.g., ILS, gene loss, and duplication) while reconstructing the correct species tree and species delimitation scenario. Model-based methods are preferable (Rannala, 2015), which can simultaneously estimate gene trees and the species tree, taking into account gene tree uncertainty and different models of sequence evolution (e.g., BEAST [Heled & Drummond, 2010] and BEST [Liu, 2008]). On the other hand, model-based methods apply computationally intensive algorithms, and their application on target enrichment datasets (i.e., on hundreds of loci) is still challenging. Several species tree inference methods have been described in the last decade using gene trees as input, and are therefore easily applicable in phylogenomic studies (e.g., PhyloNet [Than et al., 2008], STEM [Kubatko et al., 2009], MRL [Nguyen et al., 2012], ASTRAL [Mirarab et al., 2014], and ASTRID [Vachaspati & Warnow, 2015]).

The multi-species coalescent theory (Yang & Rannala, 2014) is also beneficial because it offers a framework to probabilistically estimate species boundaries and therefore delimit species in

taxonomically challenging complexes. Different methods capable to infer species delimitation under the coalescent model have been proposed in the last years (Pons et al., 2006; Leliaert et al., 2009; Zhang et al., 2013; Yang & Rannala, 2014; Jones et al., 2015). Although for more than a decade they have been applied to animal, fungal, and algal species complexes, only in the last years these methods have been increasingly employed on plant systems (Hu et al., 2015; Ruiz-Sanchez, 2015; Toprak et al., 2016; Naciri et al., 2017; Wagner et al., 2017; Hassanpour et al., 2018; Tomasello, 2018).

Phenotypic differentiation is still regarded as an important criterion for species delimitation (e.g., Stuessy, 2009; Freudenstein et al., 2017). In flowering plants, the shape of organs (e.g., of leaves) is an important criterion for species-level classification, and many quantitative approaches are available to discriminate species (e.g., Jensen et al., 2002). In the past, traditional morphological classification as a purely descriptive approach has led to the subjective descriptions of hundreds to thousands of morphotypes as species due to minor morphological differences, especially in apomictic polyploid complexes (e.g., Stace, 1998). Relying on single, "diagnostic" characters, in general bears the danger of erroneous classifications (Stuessy, 2009). Another challenge is the assessment of the phenotypic variation of characters, which is in plants usually very high due to phenotypic plasticity in response to environmental factors (Stuessy, 2009; Gratani, 2014). Geometric morphometrics makes progress in the exact, objective, and fine-scale assessment of differences in leaf shape via landmark methods and appropriate multivariate statistics (Hodač et al., 2014, 2018). Geometric morphometrics can support final taxonomic decisions based on next-generation sequencing data (Kilian et al., 2015), particularly in complexes where morphological differences are hard to assess with traditional morphological classification. So far, there have been only a few studies applying both genomic and morphometric approaches to disentangle phylogenetic relationships in plants (e.g., Jiang et al., 2019).

Ranunculus auricomus is one of the largest European apomictic polyploid complexes and a model system to study apomixis and the evolution of phylogenetically young groups (Hörandl, 1998; Hörandl et al., 2009; Hodač et al., 2014; Barke et al., 2018). The distribution ranges from Greenland to Europe, Western Siberia, and Alaska, and from the arctic, temperate to the Mediterranean zone (Jalas & Suominen, 1989; GBIF Secretariat, 2017). In Europe, species inhabit a broad range of habitats from stream- and riversides to forests and marshy, humid, and semi-dry anthropogenic meadows. The complex comprises more than 800 species (currently ca. 840 taxa in Euro+Med database; Raab-Straube et al., 2014), mainly described based on a morphological species concept. Hybridization and polyploidization strongly influenced the

evolutionary history of *R. auricomus* species: The huge number of morphologically diverse polyploid (mainly tetraploid) agamospecies probably arose from hybridization of few sexual progenitor species (Paun et al., 2006b; Hörandl et al., 2009; Hodač et al., 2014, 2018; Hojsgaard et al., 2014b). Already Linnaeus (1753) recognized the existence of two contrasting morphotypes. He described them as *R. auricomus* L., characterized by deeply dissected basal leaves, from Western Europe, and *R. cassubicus* L., characterized by non-dissected basal leaves, from Siberia (Kvist, 1987). Afterward, several authors realized that these core morphotypes include numerous variants and are interconnected by several hybridogenic intermediates. Because of too many intermediates, a concept of "main species" and agamic lineages as subspecies (Marklund, 1961, 1965) failed. Later on, these main species or morphogroups were no longer formally accepted (Ericsson, 1992; Hörandl, 1998). Phylogenetically, these morphogroups are not monophyletic (Hörandl et al., 2005, 2009). Most of the species (about 600) were described during the 20th century from Northern Europe (Marklund, 1961, 1965; Julin, 1980; Ericsson, 1992). In recent years, some new regional and local species have been recognized in Central Europe (e.g., Soó, 1964, 1965; Borchers-Kolb, 1985; Hörandl & Gutermann, 1998c, 1999; Dunkel, 2014, 2015; Dunkel et al., 2018).

Despite the huge diversity of apomictic taxa (> 800), only 12 sexual species have so far been described. *Ranunculus cassubicifolius* W.Koch (Koch, 1939) is a widespread but disjunctly distributed diploid species (with locally distributed sexual autotetraploid populations) occurring along stream-/riversides, in swampy or humid forests, and in meadows at the northern and southeastern edge of the Alps (Hörandl & Gutermann, 1998b; Hörandl & Greilhuber, 2002). *Ranunculus carpaticola* Soó (Soó, 1965) inhabits humid beech, hornbeam, and oak forests and has a disjunct distribution in the Carpathians (Paun et al., 2006a,b; Hörandl et al., 2009). The two species form tetra- and hexaploid, apomictic hybrids lineages distributed in central to northeastern Slovakia (Hörandl et al., 2009). *Ranunculus flabellifolius* Rchb. (Reichenbach, 1832) is a conspicuous species probably restricted to the Southern Carpathians (Jalas & Suominen, 1989; Dunkel et al., 2018; see Mráz & Ronikier, 2016 for geographical classification). Nearby, a few occurrences were also mentioned outside of the Carpathians in Serbia (Stevanović et al., 1991). Similar to *R. carpaticola* and *R. cassubicifolius*, *R. flabellifolius* is characterized by non-dissected basal leaves but differs from them in fan-shaped (connected segments) stem leaves and the infrequent occurrence of up to 3-lobed basal leaves.

The other sexual species share dissected basal leaves and a heterophyllous basal leaf cycle. *Ranunculus cebennensis* Dunkel (Dunkel et al., 2018), *R. envalirensis* Grau (Grau, 1984), and *R. marsicus* Guss. & Ten. (Tenore, 1835–1836) are high mountain species with small, restricted

geographical ranges. Diploid *R. envalirensis* inhabits subalpine meadows in the Eastern Pyrenees, *R. cebennensis* is only known from two different meadows in the Massif Central (France), and the only (mainly) tetraploid sexual species *R. marsicus* (also with penta- to heptaploid asexual cytotypes) occurs in the Central Apennines in subalpine meadows. The other species with dissected basal leaves are restricted to the Illyrian region, a warm, humid deciduous forest area ranging from southeastern Austria to central and southwestern Serbia (see, e.g., Košir et al., 2008). The diploid *R. notabilis* Hörandl & Guterm. (Hörandl & Gutermann, 1998c) was first described from southeastern Austria and is characterized by strongly dissected leaves. *Ranunculus austroslovenicus*, *R. calapius*, *R. mediocompositus*, *R. peracris*, and *R. subcarniolicus* grow in Slovenia and western Croatia, partly in close sympatry, and were recently described (Dunkel et al., 2018). They settle in floodplain forests, humid deciduous forests, and moist meadows. Obligate sexual reproduction was ascertained previously for *R. marsicus* (tetraploids) by Masci et al. (1994) based on embryological studies; for *R. carpaticola*, *R. cassubicifolius*, and *R. notabilis* using embryology and flow cytometric seed screening (FCSS) by Hojsgaard et al. (2014, and previous literature therein) and Barke et al. (2018); and for the other taxa included here by Dunkel et al. (2018) using FCSS and/or pollen quality measurements. For other European species of the complex, apomixis was confirmed using embryological (Häfliger, 1943; Nogler, 1984) or cytological data (Rousi, 1956), large-scale FCSSs on more than 200 populations (Karbstein et al., subm. a), and pollen quality measurements (Hörandl et al., 1997; Hörandl & Gutermann, 1998a,b, 1999; Dunkel, 2010, 2011).

Applying isozyme markers, the split between sexual species of both contrasting morphotypes (*R. carpaticola*/*cassubicifolius* and *R. notabilis*) was estimated ca. 900,000 years ago, whereas the split between morphologically similar *R. carpaticola* and *R. cassubicifolius* approximately occurred ca. 300,000 years ago (Hörandl, 2004). Recently published biogeographical analyses based on single-copy nuclear genes by Tomasello et al. (2020) revealed that vicariance has triggered allopatric speciation of a widespread European, probably forest understory, ancestor and that main speciation events took place ca. 830,000–580,000 years ago. The whole *R. auricomus* group is monophyletic within the genus and is related to North American and Central Asian taxa (Emadzade et al., 2011, 2015). Molecular phylogenetic investigations based on genomic data covering all diploid and tetraploid sexual species of the complex have not been conducted. Hence, information on the monophyly and the degree of genetic divergence of taxa is lacking. Thus, the delimitation of sexual species within this complex is taxonomically unclear.

Species within the *R. auricomus* complex exhibit various basal leaf shapes throughout the year, but only the most dissected spring basal leaf is the taxonomically most important one (Borchers-Kolb, 1985; Hörandl & Gutermann, 1995, 1998a). Furthermore, only the central part of the lowermost stem leaf is deemed to be of taxonomic relevance (Hörandl & Gutermann, 1995). Geometric morphometrics has been shown to elucidate morphological leaf diversity of experimental F_2 hybrids between the morphologically most divergent species (Hodač et al., 2018). However, a concatenated analysis using all taxonomically informative traits of all sexual species has not yet been conducted.

Consequently, we address the following questions in our study: Is a combination of RADseq and target enrichment able to resolve distinct evolutionary lineages in the sexual taxa of the *R. auricomus* complex? Do these lineages match described species, or do we have to revise the species delimitation? Do phylogenetic relationships match similarity patterns found by geometric morphometrics? Can we propose a well-founded species concept based on next-generation sequencing data supported by geometric morphometrics? Finally, we aim to establish an innovative workflow combining phylogenomics (RADseq and target enrichment) and geometric morphometrics to delimit species in evolutionary young complexes using the *R. auricomus* group as a model system.

Materials and Methods

Study Locations and Population Sampling

We included all 12 described sexual species of the *Ranunculus auricomus* complex in the present study (Table 1). From 1998 to 2018 (mainly 2017 and 2018), we sampled up to 20 individuals per population and one to three populations per species in Europe including locations in Andorra, Austria, Croatia, France, Germany, Hungary, Italy, Romania, Slovakia, and Slovenia (Table 1, Fig. 1). We used the diploid species *R. pygmaeus* and *R. sceleratus* as outgroup species according to the *Ranunculus* phylogeny of Emadzade et al. (2011). We sampled living plants and kept them in the Old Botanical Garden at the University of Göttingen under controlled environmental conditions for further flow cytometric, genetic, and geometric morphometric analyses. Herbarium specimens were also collected for all populations and deposited in GOET. Additionally, we recorded altitude, GPS coordinates, and habitat characteristics of populations. We dried fresh leaves (collected freshly in spring 2017 to 2019)

in silica gel for further flow cytometric and genetic laboratory work. Moreover, we sampled fresh basal and stem leaves, and receptacles of garden individuals for geometric morphometric analyses (see section Geometric morphometrics for details). Individuals were cultivated in 1.5 l pots with Fruhstorfer Topferde LD 80 (garden beds with similar solar radiation and water supply).

Flow Cytometry and Single-Seed Flow Cytometric Seed Screening

We checked ploidy levels of all samples by conducting flow cytometry (FC) of silica-gel dried leaf material. Leaf material (~0.5–1 cm^2) was ground to small pieces by one steel ball (Ø 4 mm) in a 2 ml Eppendorf tube with Tissue Lyzer II (Qiagen, Hilden, Germany; 30 Hz, time 7–14 s), basically following Klatt et al. (2016) and Barke et al. (2018) (see Table S1). Analyses were carried out for 1 to 15 individuals per population yielding 191 FC measurements in total (FC and FCSS Table available on Figshare, https://doi.org/10.6084/-m9.figshare.12383351.v1). Data were quality-filtered according to the following settings: A maximum coefficient of variation of 6.0, a minimum number of parts within a particular peak of 500, and a minimum number of parts in a total measurement of 4000. We removed all samples that did not match the quality settings.

For confirmation of the sexual reproduction mode, we conducted single-seed flow cytometric seed screening (ssFCSS) using seeds harvested from summer 2017 to 2019. We analyzed seeds as described in FC measurements but with slight modifications: We added 100 µl extraction buffer Otto I to the seed and ground the seed for 5–7 s. Then, we put 100 µl extraction buffer Otto I to the ground seed material and immediately inverted samples for 30 s. Afterward, we filtered samples through CellTrics filters with a mesh of 30 µm into flow cytometric sample tubes and stained them with 800 µl Otto II. We conducted ssFCSS analysis for 5 to 10 seeds per individual and 1 to 3 individuals per population yielding 5 to 23 seeds per population (see FCSS Table on Figshare). In total, we performed 205 FCSS measurements. We set a maximum coefficient of variation of 7.0, a minimum number of parts within a particular peak to 100 (embryo) and 300 (endosperm), and a minimum number of parts in a total measurement of 3000 as quality thresholds.

All analyses were performed on a CyFlow Ploidy Analyzer (Sysmex, Norderstedt, Germany). We used the diploid *R. auricomus* individual "J6xF7/01" (Barke et al., 2018) as an external standard (performed before each measurement series at a day) and for gain adjustment, i.e., the voltage of the photomultiplier (Doležel et al., 2007). The software CUBE16 v.1.6 (Sysmex,

Norderstedt, Germany) was used to calculate the DNA content (as relative fluorescence intensity) by determining the mean peak values of standard and leaf sample, and coefficient of variation (CV, per peak) and number of parts (in total). To ascertain the reproductive pathway, we calculated the ratio between endosperm ploidy and embryo ploidy as peak index (PI). A PI around 1.5 indicates a sexual pathway because a diploid embryo and a triploid endosperm formed after double fertilization. Peak indices of 2 or higher indicate apomictic pathways, as the unreduced egg cell develops parthenogenetically, while the endosperm developed without fertilization or with contribution of one or two reduced sperm nuclei (PI = 2, 2.5, and 3.0, respectively; Matzk et al., 2000; Doležel et al., 2007; Klatt et al., 2016). If the embryo was still in development and hence the peak too small for a reliable FCSS measurement (but the endosperm peak clearly determinable), we used the leaf value of the mother plant to calculate the ratios (62 of 205, ca. 30% of measurements; see Fig. S1A, and FCSS Table on Figshare). We carried out statistical analyses in R v.3.6.1 (R Core Team, 2019).

We calculated mean leaf/embryo peak values for each population and translated them into ploidy levels of leaves (mother plant), embryo, and endosperm (see Table S2).

DNA Extraction, RADseq, and Target Enrichment

We extracted the total genomic DNA of 45 *R. auricomus* samples from ~1.5 cm2 silica-gel dried leaf material using the Qiagen DNeasy Plant Mini Kit (Qiagen, Hilden, Germany). Additionally, we also extracted silica-gel dried leaf material of the outgroup species *R. sceleratus* and *R. pygmaeus*. We followed the manufacturer's instructions but prolonged sample incubation in lysis buffer to 1 hour increasing DNA yield. We checked DNA concentration using the Qubit fluorometer and the Qubit dsDNA HS Assay Kit (ThermoFisher Scientific, Waltham, U.S.A.). We diluted DNA solutions to 30 ng/µl in a target volume of 55 µl, checked DNA quality by gel electrophoresis, and removed samples with low quality (fragmented bands). Samples packed in dry ice were sent to Floragenex (Portland, Oregon, U.S.A.) for generating RAD libraries based on the protocol of Baird et al. (2008). The enzyme PstI was used for digestion, and the multiplexed samples were sequenced on an Illumina HiSeq 4000 platform to produce 100 bp single-end reads at the University of Oregon Genomics Core Facility. We checked the quality of raw reads using FASTQC v.0.11.8 (Andrews, 2010).

Table 1. Locations of 34 studied populations of sexual diploid and tetraploid species within the *R. auricomus* complex across Europe and outgroup specimens *R. sceleratus* and *R. pygmaeus*. Pop. ID, Population ID; Map ID, see Fig. 1; RAD/TE, DNA samples used in RADseq and/or target enrichment analyses; Locality by country and ISO code 3166-2; Collection date means date of population sampling; Altitude, in meter above sea level, m.a.s.l.; F.G. Dunkel (Karlstadt, Germany) kindly provided garden individuals and habitat characteristics of populations with Pop. ID "Du". * locality is at locus classicus or nearby; [#] holotype; [+] private herbarium F.G. Dunkel.

Pop. ID	Map ID	RAD/TE	Taxon	Locality	Collection date	Alt.	Latitude Longitude	Habitat	Collector	Herbarium voucher
Du-30441*	(1)	RAD/TE	*R. austroslovenicus* Dunkel	Slovenia (SI)	23 Apr 2013	260	45.462527 14.817019	Shrubbery, forest edge	*F.G. Dunkel*	LJU[#], B, M
Du-30442	(2)	RAD	*R. austroslovenicus*	Slovenia (SI)	23 Apr 2013	470	45.684671 14.810336	Forest	*F.G. Dunkel*	M
LH012	(3)	RAD/TE	*R. austroslovenicus*	Slovenia (SI)	03 May 2017	469	45.677445 14.826140	Forest	*L. Hodač, K. Spitzer*	GOET
Du-34889*	(4)	RAD/TE	*R. calapius* Dunkel	Croatia (HR)	10 Apr 2017	141	45.518944 15.570917	Forest	*F.G. Dunkel*	ZA[#], B, FR, M, WB, ZT
Du-35351	(5)	RAD/TE	*R. calapius*	Croatia (HR)	23 Apr 2018	117	45.51806 15.612500	Forest	*F.G. Dunkel*	[+]
EH9126	(6)	RAD/TE	*R. carpaticola* Soó	Romania (RO)	07 Jul 1998	1360	45.432833 25.530708	Forest	*E. Hörandl, S. Hörandl, F. Hadacek*	GOET
Du-21047	(7)	RAD/TE	*R. carpaticola*	Hungary (HU)	19 Apr 2008	230	48.114122 20.653033	Forest edge	*F.G. Dunkel*	[+]
LH040*	(8)	RAD/TE	*R. carpaticola*	Slovakia (SK)	03 May 2018	387	48.67731 20.106684	Forest	*L. Hodač, K. Spitzer*	GOET
Du-28673	(9)	RAD	*R. cassubicifolius* W.Koch	Slovenia (SI)	17 May 2012	510	46.305528 14.612833	Forest	*F.G. Dunkel*	[+]
Du-15980	(10)	RAD	*R. cassubicifolius*	Italy (IT)	25 Apr 2006	330	46.248611 13.313056	Forest	*F.G. Dunkel*	[+]
LH006	(11)	RAD/TE	*R. cassubicifolius*	Germany (DE)	30 Apr 2017	539	47.811203 12.506822	Marshy forest	*L. Hodač, K. Spitzer*	GOET
LH007	(12)	RAD	*R. cassubicifolius*	Austria (AT)	30 Apr 2017	511	47.993123 13.103185	Marshy forest	*L. Hodač, K. Spitzer*	GOET
LH008	(13)	TE	*R. cassubicifolius*	Austria	01 May 2017	473	47.938828	Streamside	*L. Hodač, K. Spitzer*	GOET

				(AT)			14.94475			
LH009	(14)	TE	*R. cassubicifolius*	Austria (AT)	01 May 2017	548	47.924792 14.972606	Forest	*L. Hodač, K. Spitzer*	GOET
LH016	(15)	RAD/TE	*R. cassubicifolius*	Slovenia (SI)	04 May 2017	514	46.010798 14.102718	Marshy forest	*L. Hodač, K. Spitzer*	GOET
Du-33354*	(16)	RAD/TE	*R. cebennensis* Dunkel	France (FR)	06 May 2016	910	45.560557 2.656944	Steep meadow	*F.G. Dunkel*	LY[#], B, M
Du-29983	(17)	RAD/TE	*R. envalirensis* Grau	Andorra (AD)	23 May 2013	1750	42.591333 1.671583	Shrubbery, alpine meadow	*F.G. Dunkel*	M
Du-29988*	(18)	RAD/TE	*R. envalirensis*	France (FR)	23 May 2013	1600	42.463966 2.074866	Alpine meadow	*F.G. Dunkel*	M
EH10316*	(19)	RAD	*R. envalirensis*	Andorra (AD)	17 Jul 2016	2370	42.542083 1.715333	Alpine meadow	*E. Hörandl*	GOET
Du-25795*	(20)	RAD/TE	*R. flabellifolius* Rchb.	Romania (RO)	21 Apr 2010	885	45.049444 21.771306	Forest	*F.G. Dunkel*	M
LH023*	(21)	RAD	*R. flabellifolius*	Romania (RO)	10 Apr 2018	634	45.033952 21.833548	Forest	*L. Hodač, K. Spitzer*	GOET
LH025*	(22)	RAD/TE	*R. flabellifolius*	Romania (RO)	10 Apr 2018	811	45.032008 21.785110	Forest	*L. Hodač, K. Spitzer*	GOET
Du-23722*	(23)	RAD	*R. marsicus* Guss. & Ten.	Italy (IT)	25 Apr 2003	1390	42.033055 13.805500	Humid meadow	*F.G. Dunkel*	[+]
LH017*	(24)	RAD/TE	*R. marsicus*	Italy (IT)	26 Jun 2017	1535	41.845403 13.929692	Humid meadow	*L. Hodač, K. Spitzer*	GOET
LH018*	(25)	RAD/TE	*R. marsicus*	Italy (IT)	27 Jun 2017	1244	41.868211 14.021872	Humid meadow	*L. Hodač, K. Spitzer*	GOET
LH014*	(26)	RAD/TE	*R. mediocompositus* Dunkel	Slovenia (SI)	03 May 2017	434	45.849540 14.259460	Humid meadow	*L. Hodač, K. Spitzer*	GOET
LH015*	(27)	RAD/TE	*R. mediocompositus*	Slovenia (SI)	03 May 2017	436	45.848988 14.257088	Marshy meadow	*L. Hodač, K. Spitzer*	GOET
EH10137*	(28)	RAD/TE	*R. notabilis* Hörandl & Guterm.	Austria (AT)	08 May 2011	220	47.047778 16.432500	Forest edge, meadow	*E. Hörandl, S. Hörandl, F. Hadacek*	GOET
Hoe5615/01	(29)	TE	*R. notabilis* Hörandl & Guterm.	Austria (AT)	05 May 1994	210	47.052739 16.38406	Forest	*E. Hörandl, W. Gutermann*	WU

LH028*	(30)	RAD	*R. notabilis*	Austria (AT)	13 Apr 2018	227	47.053221 16.435162	Forest edge, meadow	*L. Hodač, K. Spitzer*	GOET
LH010*	(31)	RAD/TE	*R. peracris* Dunkel	Slovenia (SI)	02 May 2017	151	45.890750 15.370933	Forest edge	*L. Hodač, K. Spitzer*	GOET
LH011	(32)	RAD/TE	*R. peracris*	Slovenia (SI)	02 May 2017	148	45.880803 15.336522	Forest edge	*L. Hodač, K. Spitzer*	GOET
Du-33266*	(33)	RAD/TE	*R. subcarniolicus* Dunkel	Slovenia (SI)	26 Apr 2016	325	45.944056 14.653083	Forest edge	*F.G. Dunkel*	M
LH013*	(34)	RAD/TE	*R. subcarniolicus*	Slovenia (SI)	03 May 2017	324	45.945453 14.651618	Humid meadow	*L. Hodač, K. Spitzer*	GOET
EH10426		RAD/TE	*R. sceleratus* L.	Germany (DE)	15 Apr 2018	165	51.579070 10.145890	Moat	*E. Hörandl, S. Hörandl, F. Hadacek*	GOET
LG09		TE	*R. pygmaeus* Wahlenb.	U.S.A. (US)	07 Nov 2009	2	64.456100 −165.09610	–	*A.L.S. Gustafsson*	O

For target enrichment, we selected almost the same samples as in RADseq analyses (see Table 1; Tomasello et al., 2020). Probe design for target enrichment was done from RNA sequences of *R. carpaticola*, *R. cassubicifolius*, and *R. notabilis* published by Pellino et al. (2013), and is described in detail in Tomasello et al. (2020). In total, the baits system consisted of 17,988 probes (14,632 of which unique), used to in-liquido capture 736 target genomic regions. Library preparation and hybrid capture protocol are described in detail in Appendix S1. Two different paired-end sequencing runs (2× 250 bp) were performed on an Illumina MiSeq System (Illumina, San Diego, California, U.S.A.). The obtained raw reads were processed using the pipeline HYBPHYLOMAKER (all scripts available at: https://github.com/tomas-fer/HybPhyloMaker/; Fér & Schmickl, 2018). Before analyses, reads were phased using SAMTOOLS v.0.1.19 (Li et al., 2009) in order to retrieve allelic information. After processing data on HYBPHYLOMAKER and filtering for missing data, we were able to retrieve information from 663 of the initial 736 nuclear loci (2495 exons in total) excluding 73 loci (Appendix S1). Then, these alignments were used for the subsequent gene tree/species tree and coalescent-based species delimitation analyses. Although chloroplast sequences were generated during laboratory work, we excluded them because sequences were not informative to resolve species relationships.

Raw reads are stored in the National Center for Biotechnology Information Sequence Read Archive (SRA): BioProject IDs PRJNA627796 (RADseq) and PRJNA628081 (target enrichment).

De Novo Assembly of RAD Loci and Parameter Optimization in IPYRAD

The following analyses were performed with IPYRAD v.0.7.28 (Eaton, 2014; Eaton & Overcast, 2020) on the GWDG HPC Cluster (Göttingen, Germany). RADseq raw reads were demultiplexed allowing no mismatch in barcode sequence. Afterward, raw reads were cleaned by removing adapter sequences and restriction overhang (TGCAG), and trimmed if the Phred quality score was below 20 (maximum low-quality base calls in a read).

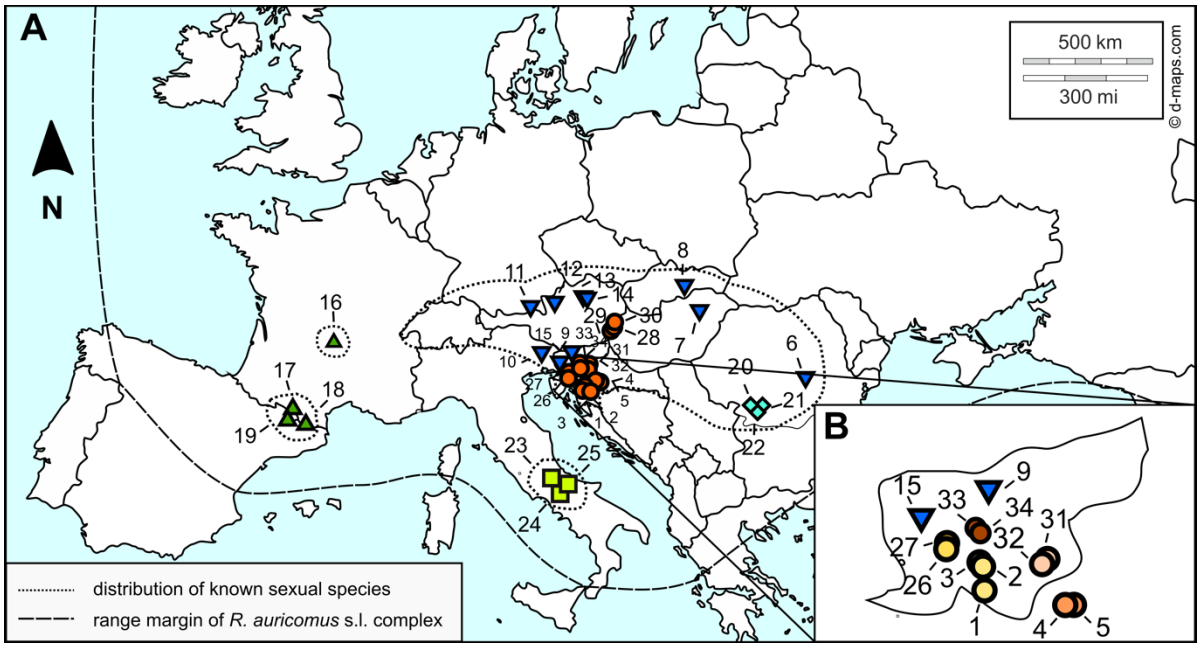


Figure 1. A, Locations of 34 studied *Ranunculus auricomus* populations across Europe. The range margin of the whole *R. auricomus* complex is indicated by a broad dashed line whereas the distribution of the known sexual species is marked with a narrow dotted line. Dark blue triangles: *R. cassubicifolius* s.l.; turquoise rhombs: *R. flabellifolius*; green triangles: *R. envalirensis* s.l.; light green squares: *R. marsicus*; and orange circles: *R. notabilis* s.l. **B**, Sampling locations of the sexual *R. auricomus* populations in the Illyrian region (Slovenia, Croatia). Numbers beside locations correspond to Map IDs in Table 1. Colors of circles are given according to membership to formerly accepted species within *R. notabilis* s.l. (light yellow: *R. austroslovenicus*; light orange: *R. calapius*; yellow: *R. mediocompositus*; light pink: *R. peracris*; and dark orange: *R. subcarniolicus*). We downloaded the original map from https://d-maps.com/.

Reads shorter than 35 bp were excluded after adapter trimming. For parameter optimization, we developed a workflow accounting for different ploidy levels of sexual *R. auricomus* species. First, we separated diploid and tetraploid (and higher ploidy-level) individuals into two datasets; second, we (separately) optimized the in-sample clustering threshold (ISCT, within individuals) of diploid and tetraploid (and higher ploidy-level) individuals; third, we optimized the between-sample clustering threshold (BSCT, between individuals) of merged assemblies; and fourth, we selected the optimal minimum number of samples per locus by choosing the maximum likelihood (ML) tree with the best bootstrap support and a stabilized topology. We included all 45 individuals in the optimization process and basically followed strategies of parameter optimization described in Paris et al. (2017) and Pätzold et al. (2019). For each dataset (2n, 4n + one 6n), we kept all parameters constant (default) except minimum depth for statistical/majority rule base calling and clustering threshold (minimal sequence similarity in

percent) for de novo assembly. We specified the minimum depth to 6 and 12, and the range of clustering threshold from 60% to 99% (60, 70, 80, 85, 90, 91, 92, …, 99).

To optimize ISCT, we carried out calculations for different minimum depths and clustering thresholds and analyzed the output regarding the number of clusters within samples, average cluster depth, and clusters rejected due to high heterozygosity (Fig. S2A–F). With decreasing ISCT, number of clusters (steepest between 0.99 and 0.96 [2n] / 0.99 and 0.93 [4n+6n]) and cluster depth (steepest between 0.99 and 0.96 [2n] / 0.99 and 0.93 [4n+6n]) decreased, and clusters rejected due to high heterozygosity increased (steepest between 0.94 and 0.60 [2n] / 0.91 and 0.60 [4n+6n]) (Fig. S2A–F). We selected an ISCT of 0.95 for diploid species and an ISCT of 0.92 for tetraploid (and hexaploid) individuals, minimizing rejected clusters and maximizing number of clusters and cluster depth (see also Pätzold et al., 2019). Generally, fewer clusters of mindepth12 compared to mindepth 6 setting were rejected due to high heterozygosity. Therefore, we selected the mindepth12 setting. Then, we merged the optimized diploid and tetraploid (and hexaploid) assemblies.

To optimize BSCT, we again conducted calculations for different minimum depths and clustering thresholds (setting maximum alleles to 2 because more than 90% of the individuals are diploid) and analyzed the output regarding number of polymorphic loci, SNPs, loci filtered by maxSNPs (maximum of allowed SNPs), removed duplicates, shared loci, and new polymorphic loci (Fig. S2G–L). Then, we selected the optimal BSCT. With decreasing BSCT, number of polymorphic loci, SNPs, and shared loci steeply increased (peak ~0.95–0.90) followed by a smooth decrease (Fig. S2G,I,J). Number of loci filtered by maxSNPs and removed duplicates raised in a sigmoidal manner with decreasing BSCT (turning point ~0.93–0.95; Fig. S2H, K). From BSCT 0.99 to 0.95, the merged assembly gained loci; from 0.94 to 0.90, the number of polymorphic loci pended around 45,000 followed by a smooth decrease (Fig. S2G,L). We chose a BSCT of 0.92 for the merged assembly minimizing number of removed duplicates and loci filtered by maxSNPs and maximizing number of polymorphic loci, SNPs, and shared loci. In general, fewer loci were filtered by maxSNPs and removed duplicates of mindepth12 compared to mindepth 6 setting.

To evaluate the impact of missing loci on the topology of phylogenetic inferences, we performed the final filtering step in IPYRAD specifying the minimum number of samples per locus to 2 (4%), 5 (10%), 14 (30%), 24 (50%), 33 (70%), and 42 (90%). We assessed the number of loci, number of missing data, and bootstrap support of the ML tree for each step (see paragraph below; Fig. S3A–F). With decreasing minimum number of samples per locus, the

amount of missing data (3.7%–85.3 %), number of loci (134–205,738), and bootstrap support (until "min10" setting, i.e., 10% available samples per locus) increased remarkably. Number of loci drastically ascends, but average bootstrap support slightly decreased from "min10" to "min4" setting. Moreover, the topology is similar from "min30" setting downwards.

We selected the alignment "min10" due to the highest mean bootstrap support among all ML trees and a stabilized topology (compare to "min50" and "min30", Fig. S3B–D; see also workflow Fig. 2). We are aware of the problems concerning inflated bootstrap values in concatenated phylogenomic analyses (see, e.g., Weisrock et al., 2012; Shen et al., 2017). We addressed the potential inflation of bootstrap support in concatenated datasets with the QS method (in particular the QC value as a ratio of concordant to discordant patterns, see below and Pease et al., 2018). However, the gene partition analysis option of the QS method was not appropriate for our dataset. Thus, the issue of a topology that is supported by only a fraction of the dataset is still unresolvable for concatenated RADseq alignments that are characterized by a high percentage of missing data.

Phylogenomic Analyses: Maximum Likelihood Tree of RAD Loci and Quartet Sampling

Maximum likelihood analyses were performed on all resulting alignments (4%, 10%, …, and 90%) with EXAML v.3.0.21 (Kozlov et al., 2015; see Fig. S3A–F). Here, we only used *R. sceleratus* as outgroup because extractions from *R. pygmaeus* from herbarium material did not reveal sufficient high-quality DNA for RADseq. We used a *.phy file (one sequence per individual and locus [majority-rule base calling], loci concatenated into a supermatrix) from IPYRAD as input, and set the new rapid hill-climbing algorithm, the gamma model of rate heterogeneity, and a random seed number. Additionally, we enabled the memory saving option for gappy alignments regarding datasets that contained more than 50% missing data (4%–30% minimum samples per locus; see Fig. S3A–C). We also ran RAXML v.8.2.4 (Stamatakis, 2014) to generate 100 bootstrap replicate alignments with their respective parsimony starting trees. Subsequently, we performed ML analyses with EXAML v.3.0.21 on every replicate alignment based on the above-specified options. Then, we plotted the results on the EXAML tree (full dataset). We visualized all resulting trees in FIGTREE v.1.4.3 (Rambaut, 2014). For the following analyses, we selected the alignment "min10" (see section above).

To examine the phylogenetic tree for conflicted and/or poorly informative branches, we conducted the QS method (Pease et al., 2018). We set 100 replicates per branch and the log-likelihood threshold cutoff to 2. The QS method outputs three different node scores (Pease et

al., 2018): The quartet concordance (QC) describes the ratio of concordant to both discordant quartets (1: all concordant, >0: more concordant, <0: more discordant); the quartet differential (QD) defines skewness of both discordant patterns (1: equal, 0.3: skewed, 0: all topology 1 or 2); and quartet informativeness (QI) characterizes the proportion of informative replicates (1: all informative, 0: none informative). We assigned QC values to four categories (QC = 1, $1 > QC \geq 0.3$, $0 \leq QC < 0.3$, and $0 > QC \geq -1.0$) to make QC values easier accessible in Fig. 3A (e.g., all blue points with $QC \geq 0.3$, i.e., a majority of concordant patterns). QD values around 1 can indicate ILS due to the presence of both discordant topologies (random pattern), whereas values towards 0 potentially hint at directional introgression events due to the presence of only one particular alternative topology (Pease et al., 2018). We drew results together with bootstrap values on the "min10" ML tree (Fig. 3A). The calculation procedure also gives the quartet fidelity score (QF, see Table S3) indicating whether a taxon is misplaced within the topology (1 = not misplaced, 0 = misplaced; see rogue taxon approaches, Pease et al., 2018).

Phylogenomic Analyses: Genetic Distance

In order to investigate reticulate (non-tree like) evolution of sexual *R. auricomus* species, we calculated a network analysis based on the neighbor-net algorithm using the program SPLITSTREE v.4.14.6 (Huson & Bryant, 2006). To estimate the genetic distance, we specified a general time reversible (GTR) model with estimated site frequencies and maximum likelihood (equal rates of site variation, default rate matrix). Then, we performed bootstrapping with 1000 replicates and showed values >80% for the main clusters.

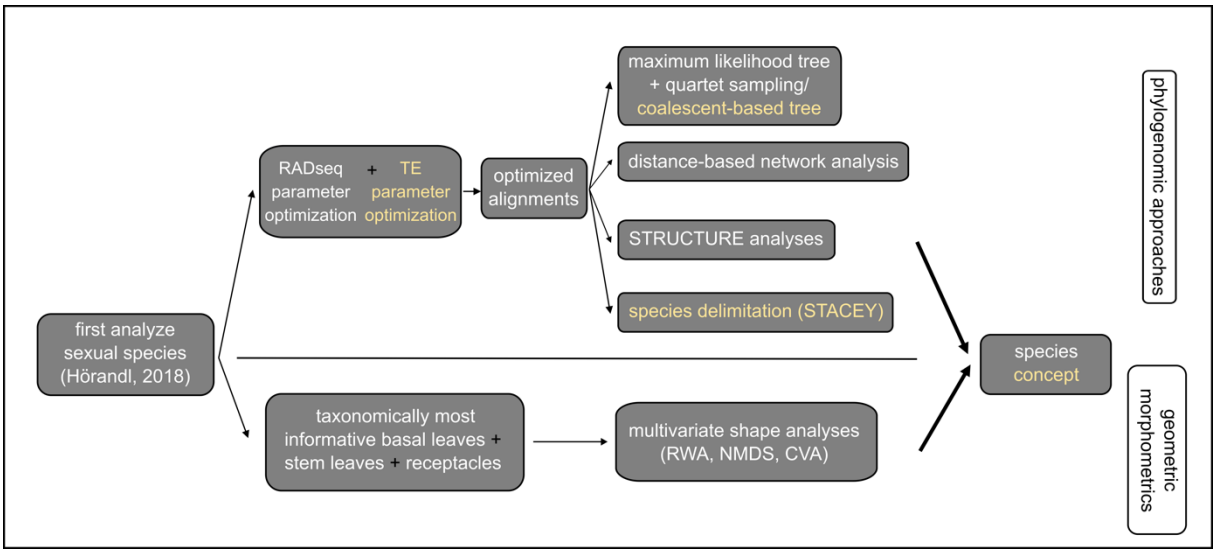


Figure 2. The basic workflow of this study: As proposed by Hörandl (2018) for polyploid apomictic species complexes, we first analyzed the sexual species of the *Ranunculus auricomus* complex. We carried out a dual-approach based on phylogenomic approaches supported by geometric morphometrics. In the first part, phylogenetic inference was mainly based on an optimized RADseq dataset (see Materials and Methods for details): We conducted a maximum likelihood analysis and investigated node support with the quartet sampling method (Pease et al., 2018). Distance-based network (Neighbor-net) and STRUCTURE analyses were applied to examine the clustering/grouping of individuals. We used an optimized target enrichment (TE) dataset to calculate a coalescent-based species tree and a species delimitation analysis (STACEY; Bouckaert & al., 2014). In the second part, we included the taxonomically most informative traits (basal and stem leaves, receptacles) to carry out multivariate shape analyses based on relative warps analysis (RWA). Phylogenomic approaches supported by geometric morphometrics were used to propose a well-defined species concept.

Phylogenomic Analyses: Genetic Structure

To infer the genetic composition of sexual *R. auricomus* species, we took the SNP-based format including one randomly selected variable site per locus from IPYRAD ("min10.ustr"; dataset without outgroup). We performed a genetic structure analysis in STRUCTURE v.2.3.4 (Pritchard et al., 2000), and set an admixture model, a burn-in of 5000, and a MCMC of 50,000 replicates. K was set from 1 to 12, and each K was replicated 5- to 10-times. To determine the optimal K value, we ran the Evanno method implemented in STRUCTURE HARVESTER (Earl & Holdt, 2012; see Table S4). We used CLUMPP v.1.1.2 (Jakobsson & Rosenberg, 2007) to merge all runs of the optimal K value, and plotted genetic structure with DISTRUCT v.1.1 (Rosenberg, 2004). Since we observed the deepest split between *R. carpaticola*/*cassubicifolius* and the other species, we conducted two additional IPYRAD analyses (same settings as "min10") considering the abovementioned groups (subsets). K was set from 1 to 6 (*R.*

carpaticola, *R. cassubicifolius*), 1 to 11 (*R. flabellifolius* to *R. peracris*), and each K was replicated 5- to 10-times. The choice of K, merging of runs, and illustration of results followed the procedure described above (see also Table S4). The second-most likely K values of the total dataset, and both subsets were usually characterized by a very low delta K value, and were therefore not drawn. We illustrated the results of both subsets together with the ML tree (Fig. 3B; other population genetic measures like heterozygosity will be presented elsewhere, Karbstein et al., subm. a).

Phylogenomic Analyses: Target Enrichment and Coalescent-Based Species Delimitation

In addition to the concatenated tree obtained with RADseq data, we estimated the species tree using the target enrichment dataset and the coalescent-based method ASTRAL (Mirarab et al., 2014). For the scope, gene trees (exon trees) were first reconstructed with ML as implemented in RAXML v.8.2.4 (Stamatakis, 2014). We run analyses with 100 standard bootstrap replicates, partitioned by exons, and applying the GTRGAMMA model to all partitions. RAXML trees were rooted, and branches with bootstrap values lower than 50 were collapsed using the script "Newick Utilities" (Junier & Zdobnov, 2010). Finally, the species tree was inferred using the coalescent-based algorithm implemented in ASTRAL-III v.5.6.3 (Zhang et al., 2018) with 100 multi-locus bootstrap replicates (Seo, 2008).

We used the BEAST2 package STACEY v.1.2.1 (Jones, 2017) to infer species boundaries among the sexual representatives of the *R. auricomus* complex (see Fig. 2). STACEY uses a Bayesian approach to infer species delimitation and species phylogeny based on the multispecies coalescent model. It is one of the few model-based programs able to perform this type of analysis. STACEY was demonstrated to outperform the other programs in estimating the correct ultrametric species tree (Andermann et al., 2019), or while inferring the best species delimitation scenario (Tomasello, 2018).

We ran BEAUTI v.2.5.2 (Bouckaert et al., 2014) to create an input file for STACEY. We used the 50 most informative loci (with the highest number of parsimony-informative sites) from those sequenced applying a target enrichment approach by Tomasello et al. (2020). The alignments (between 500 and 1000 bp long) consisted of a maximum of 56 OTUs (two alleles from 28 samples, outgroups were excluded from the analyses).

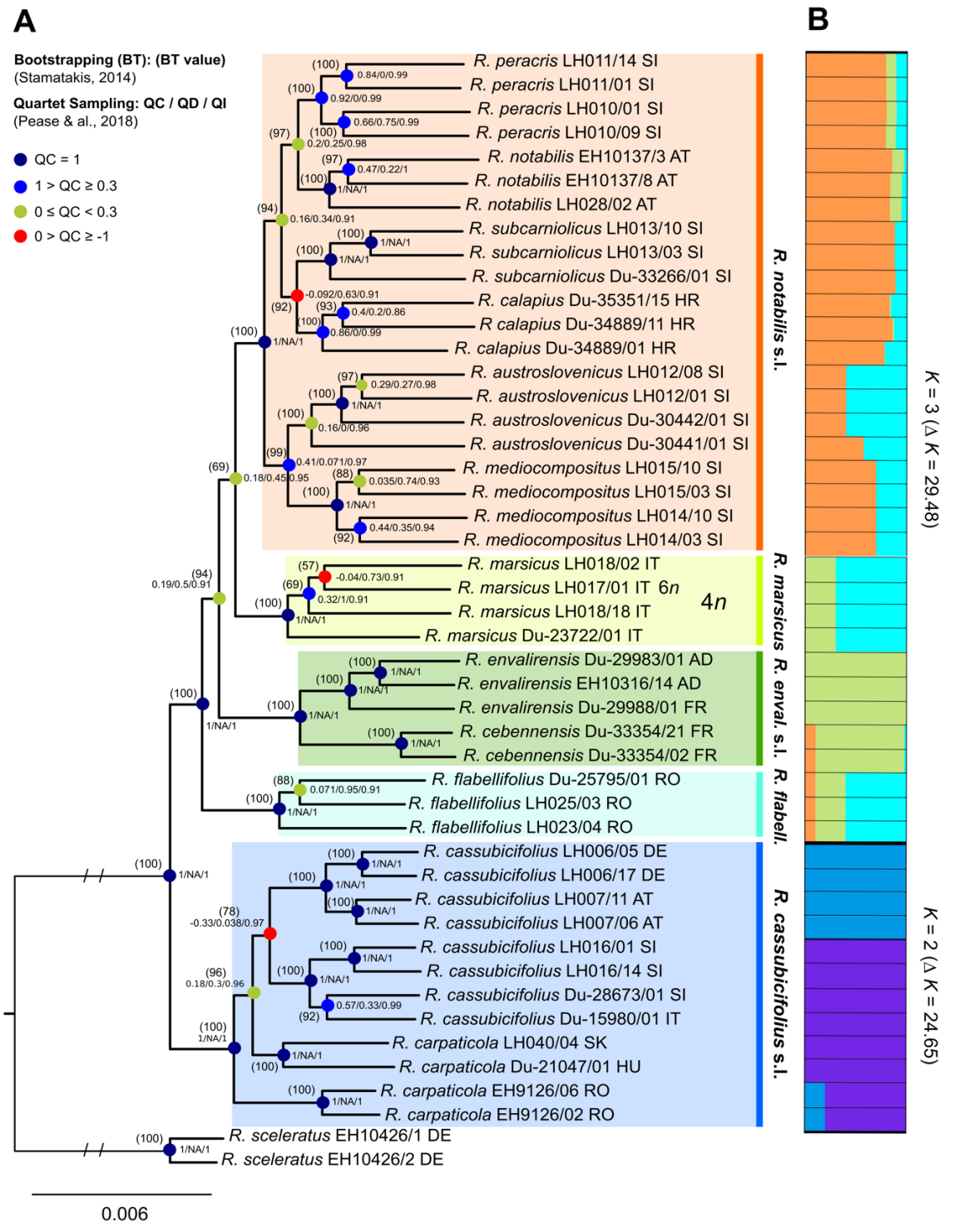


Figure 3. A, Maximum likelihood (ML) tree of 45 individuals and 12 sexual taxa within the *R. auricomus* complex (*R. sceleratus* as outgroup). Bootstrap (BT) values are given in brackets above nodes. Quartet sampling scores QC (quartet concordance), QD (quartet differential), and QI (quartet informativeness) are arranged to the right of the node (QC/QD/QI). If QC equals 1, then QD is NA (skewness of discordant patterns cannot be determined if there are no discordant patterns). See Table S3 for QF (quartet fidelity) score. Nodes are colored according to QC value (see figure legend). Boxes around clades specify the accepted species with their respective names. Country labels behind taxa are in ISO code 3166-2 format. **B**, Right to the ML tree, the bar graphs of STRUCTURE analyses for each of the two subsets with the most likely K values are illustrated (see Table S4 for results of Evanno tests).

During the inference, substitution models, clock models, and gene trees were treated as unlinked for all loci. Sequence substitution models were selected for each locus separately using the Bayesian Information Criterion (BIC) in jModeltest v.2.1.10 (Darriba et al., 2012). In the STACEY *.xml input file, parameters of the substitution models were fixed to those found in jModeltest. The strict clock was enforced for all loci fixing to 1.0 the average clock rate of one random locus while estimating all other clock rates in relation to this locus. We set the "Collapse Height" to 1×10^{-4}. This parameter has no biological meaning, and values between 0.000001 and 0.0001 usually produce similar results in similar times (Jones, 2017). The "Collapse Weight" parameter was estimated using a Beta prior with parameters $\alpha = 1.0$ and $\beta = 1.0$, which represent a flat distribution between 0 and 1 (i.e., values close to 1 indicate higher probability densities for a low number of species whereas values close to 0 indicate higher probability densities for a high number of clusters/species). We followed this approach in order to equally investigate all possible scenarios of species delimitation. Finally, we gave to the bdcGrowthRate prior a log-normal distribution (M = 4.6 and S = 1.5), a gamma shape ($\alpha = 0.1$ and $\beta = 3.0$) to the popPriorScale prior, and for the relativeDeathRate, we set a beta prior ($\alpha = 1.0$ and $\beta = 1.0$).

The input files were run for 1×10^9 iterations sampling every 100,000th generation. Two independent runs were performed to check convergence among independent analyses. We checked convergence and effective sample size (ESS > 200) in TRACER v.1.6 (Rambaut et al., 2018). We combined output files containing the trees sampled in the two independent runs using LOGCOMBINER v.2.5.1 (Bouckaert et al., 2014) after discharging 10% of the analyses as burn-in. The obtained file was processed with the "species delimitation analyser" (speciesDA.jar, Jones et al., 2015; available at: http://www.indriid.com/) using a "Collapse Height" of 1×10^{-4} and setting the similarity cut-off to 0.9. Finally, we produced a similarity matrix using a modified version of the R script provided in Jones et al. (2015).

Geometric Morphometrics

The geometric morphometric dataset comprised predominantly fresh material sampled from living plants cultivated under the same garden conditions and completed with leaves from herbaria (Table S5). We selected the most taxonomically informative traits based on previous species delimitation for geometric morphometric analyses, i.e., the most-dissected basal leaves in the leaf cycle (in comparison to type specimens), the central part of the lowermost stem leaf, and the receptacle at fruit stage (Borchers-Kolb, 1985; Hörandl & Gutermann, 1995, 1998a;

Hodač et al., 2018; see Fig. 2, Fig. S4A–C). In general, the most taxonomically informative leaves were comparable with type specimens, and thus differences due to phenotypic plasticity were minor. We collected 1 to 7 fresh basal and stem leaves per flowering plant from April to May 2018 and 2019. We scanned leaves immediately after harvesting in 400 dpi resolution using Epson Perfection V500 Photo (SEIKO Epson, Suwa, Japan) and CanoScan LiDE 220 (Canon, Ōta, Japan) scanners to avoid potential bias of morphometric analysis due to herbarization (Volkova et al., 2010; Klingenberg, 2015). We harvested receptacles at fruiting stage from June to August 2018 and 2019, and digitized them under a Leica M125 microscope (Leica Microsystems, Wetzlar, Germany) with 10–15-fold magnification. In total, we included 592 basal leaves, 580 stem leaves, and 421 receptacles in the following data analyses based on 215, 210, and 170 individuals, respectively (images stored in Figshare https://doi.org/10.6084/m9.figshare.12383351.v1). For the concatenated data analyses, we only used individuals that are characterized by all three morphological traits.

We followed Hodač et al. (2018) for image preprocessing and making of TPS files. Adjusted specimens were imported into TPSDIG v.1.40 (Rohlf, 2015) to perform digitization of 2D landmarks and semilandmarks (Bookstein, 1997a; MacLeod, 2008). We applied a different landmark approach to each dataset (see Fig. S4A–C). For basal leaves, we adopted the landmark configuration from Hodač et al. (2018) and placed 26 landmarks along the leaf outline according to the leaf venation assuring homology of landmarks among samples. Landmarks, i.e., point locations, have the advantage to objectively assess morphological differences between anatomically homologous positions of different samples (Klingenberg et al., 2012; Gunz & Mitteroecker, 2013; Hodač et al., 2018). For stem leaves, we developed a new application by putting eight landmarks along the outline of the middle leaf segment according to the venation of the central vein. Due to a gap of homologous locations, we filled the gap between landmarks 2 and 13, and 18 and 27 with 10 semilandmarks roughly situated in corresponding positions, respectively. Semilandmarks (or sliding landmarks) are able to capture length-to-width ratios (Gunz & Mitteroecker, 2013), which is important to stem leaves of sexual species within the *R. auricomus* complex. The spacing of semilandmarks is optimized by sliding along the leaf shape, and after sliding, landmarks and semilandmarks can be treated together for statistical analyses (Perez et al., 2006; Gunz & Mitteroecker, 2013). For receptacles, we also developed a new application by putting eight landmarks along the receptacle outline including locations at the androclinium, intervallum, gynoclinium, and central uppermost part. We filled the gap between landmarks 4 and 10, and 10 and 16 with five semilandmarks roughly situated

in corresponding positions, respectively. We also used semilandmarks (or sliding landmarks) to capture length to width ratios.

We applied the thin-plate-spline approach (Bookstein, 1997b) to extract shape variables from our landmark-based data. Because *R. auricomus* shows high phenotypic plasticity with extremely divergent forms, we tested the suitability of the thin-plate-spline approach using the program TPSSMALL v.1.34 (Rohlf, 2015). The first data standardization procedure comprised the separation of symmetric and asymmetric components of shape variation. The landmark configurations were symmetrized (after Klingenberg et al., 2002) with respect to the bilateral symmetry axis defined by landmarks 1 and 14 in basal leaves, landmarks 1 and 15 in stem leaves, and landmarks 1 and 10 in receptacles. The symmetrization followed the approach described in Klingenberg et al. (2002) (for details, see the caption of Fig. S4).

Symmetrized landmark configurations (each representing a single plant) were analyzed in the program TPSRELW v.1.70 (Rohlf, 2015) to search for group clustering within basal/stem leaf and receptacle morphospaces. The resulting scores of specimens (= mean shapes per plant) on the relative warps then represented input shape variables for further analyses. The intra- and interspecific morphological variability were analyzed separately for basal/stem leaves and receptacles. Specimens were plotted in the morphospace captured by the 1st and 2nd relative warp. Scores of specimens on up to 4 relative warps extracted from each morphological structure were concatenated and used as input datasets in exploratory and discriminant analyses of species differentiation.

The concatenated morphometric dataset included 11 shape variables (basal leaf, 3; stem leaf, 4; receptacle, 4), which were available without missing data in 108 plants (23 *R. cassubicifolius* s.l., 23 *R. envalirensis*, 7 *R. flabellifolius*, 3 *R. marsicus*, 52 *R. notabilis* s.l.). To assess morphological differentiation among species, we used the concatenated dataset and performed a non-metric multidimensional scaling (NMDS) analysis with Euclidean distances in the program PAST v.2.17c (Hammer et al., 2001). We used the same software to test for morphological differentiation among species with the non-parametric multivariate analysis of variance (NP-MANOVA) and 10,000 permutations. Important shape variables responsible for species differentiation were visualized together with morphological clusters as a biplot ordination diagram drawn from canonical variates analysis (CVA) aiming in maximizing the differences among a priori predefined groups (i.e., species).

Reconstructions of shape changes described by relative warps (shape principal components) were modeled with the thin-plate-spline method in the program TPSRELW v.1.70 (Rohlf,

2015). We used the same method for computation of species mean shapes averaged over plant individuals in the program MORPHOJ v.1.06d (Klingenberg, 2011). We finally examined how shapes vary with size (centroid size, related to length and width; allometry). To assess which shape variables exhibit allometry, we used centroid sizes extracted from original unaligned landmark configurations as an independent variable. Shape variables (only the symmetric component) were regressed on the independent variable (centroid size), and the significance of the regression model was computed with 1000 permutations in the program MORPHOJ v.1.06d (Klingenberg, 2011). In addition to the regression model, we tested the association between size and particular shape variables, and computed Spearman's rank correlation coefficients (ρ) for centroid sizes and scores on the four most important relative warps. Both, centroid sizes and scores on relative warps were exported from MORPHOJ v.1.06d (Klingenberg, 2011), and correlations were computed in the program PAST v.2.17c (Hammer et al., 2001).

Indumentum of Receptacles

Trichome density of the receptacle has been regarded as an important differential character by most taxonomists (e.g., Marklund, 1961, 1965; Julin, 1980; Borchers-Kolb, 1985; Hörandl & Gutermann, 1998a; Dunkel et al., 2018). We assessed trichome density as the number of trichomes in a 0.25 mm^2 transect. We placed a 0.5 mm transect on the receptacle using the image manipulation software GIMP 2 v.2.10.06 (The GIMP Team, 2018) and counted the number of trichomes within the transect for a total of 219 receptacles. We used R v.3.6.1 to illustrate trichome density per species/accepted species (mean/species) with boxplots and to examine group differences with Kruskal-Wallis rank-sum test followed by pairwise Wilcoxon rank-sum tests with Holm correction (R scripts available on request).

Results

Flow Cytometry and Single-Seed Flow Cytometric Seed Screening

All species were diploid except for *Ranunculus marsicus* exhibiting mainly tetraploid individuals (see Table S2). *Ranunculus marsicus* population LH017 comprised tetraploid and hexaploid individuals. FCSS results confirmed the sexual reproductive pathway of all diploid taxa and tetraploid *R. marsicus* (LH018). The *R. marsicus* population LH017 exhibited only

one sexual seed from a tetraploid individual and asexual seeds from tetraploid and hexaploid individuals (see Table S2). The sexual *R. cassubicifolius* population LH016 exhibited also two asexual seeds produced by only one individual (remaining seeds sexually formed, see Fig. S1B and Figshare).

Phylogenomic Analyses: Maximum Likelihood Tree of RAD Loci and Quartet Sampling

We retained 7,237 million raw reads on average per sample (range: 1,413 to 14,405 million raw reads). After filtering steps, analyses yielded 7,235 million reads (range: 1,412 to 14,401 million reads). After optimization of the ISCT and BSCT, we finally obtained 24,338 loci per sample on average (range: 8,950–38,435 loci), and 86,782 loci and 400,271 SNPs for the total dataset.

The ML analysis revealed a fully resolved and well-supported phylogeny of the sexual species within the *R. auricomus* complex (bootstrap values = BT, in brackets, Fig. 3A). Almost all described morphospecies are monophyletic (except *R. carpaticola*). The first (and deepest) split separates the clade containing *R. cassubicifolius* s.str. and *R. carpaticola* (*R. cassubicifolius* s.l., non-dissected-leaf morphotypes) from the clade containing *R. flabellifolius*, *R. envalirensis* s.str. and *R. cebennensis* (*R. envalirensis* s.l.), *R. marsicus*, and the Illyrian species (non-dissected- and dissected-leaf morphotypes). The first main clade is geographically east-west partitioned in a Carpathian group (*R. carpaticola*), a Southern Alpine-Illyrian group, and a Northern Alpine group of *R. cassubicifolius* s.str., the latter with low support (BT = 78%). Within the second main clade, surprisingly, the species with mostly non-dissected leaves, *R. flabellifolius*, mainly found in the Southern Carpathians, appears on the first, well-supported branch (BT = 100%). The following branch (BT = 100%) contains the high-altitude species *R. envalirensis* s.str. from the Eastern Pyrenees and *R. cebennensis* from the Massif Central in France. The split between *R. marsicus* and the Illyrian clade is less supported (BT = 69%). The only known tetraploid sexual *R. marsicus* from high-altitudes in the Apennines is thus positioned between the other high-altitude species and the Illyrian clades. The Illyrian cluster is subdivided into two well-supported subclades: First, *R. austroslovenicus* and *R. mediocompositus* (BT = 99%), and second, *R. calapius*, *R. subcarniolicus*, *R. notabilis* s.str., and *R. peracris* (BT = 94%). Within these subclades, bootstrap support is partially lowered (see Fig. 3A).

We examined the "placement" of taxa and level of node incongruence with the quartet sampling method (QC/QD/QI below bootstrap value; Fig. 3A): The taxon-specific QF score ranged

between 68% and 93% concordant replicates (mean 82%), i.e., taxa are widely "placed" correctly within the phylogeny (see Table S3). All nodes were highly informative indicated by QI values between 91% and 100% informative replicates per node (mean 98%). The QC value frequently revealed nodes with high amounts of concordant patterns (>0.3; 63.7% of all nodes) but also with particular amounts of discordant patterns (<0.3; 27.3% of all nodes, Fig. 3A). Within the first main clade (*R. carpaticola*/*cassubicifolius*), the node splitting the Northern Alpine *R. cassubicifolius* s.str. and the Southern Alpine-Illyrian *R. cassubicifolius* s.str. showed more discordant than concordant patterns (QC = −0.33) that are skewed towards one alternative topology (QD = 0.038) indicating directional introgression. Particularly the nodes separating the recently described Illyrian species are moderately to highly conflicting. The nodes splitting *R. austroslovenicus* and *R. mediocompositus*, and *R. notabilis* s.str. and *R. peracris* revealed moderate to low amounts of concordant patterns (QC: 0.41 and 0.2, respectively) but showed a skew towards one alternative topology potentially also indicating (directional) introgression (QD: 0.071 and 0.25, respectively). Moreover, the node between *R. calapius* and *R. subcarniolicus* revealed more discordant than concordant patterns (QC = −0.092) potentially caused by ILS rather than introgression (QD = 0.63). In general, low bootstrap values corresponded to low QC values. In contrast, some nodes with high bootstrap values (90%–100%) were moderately conflicting and possessing low QC values.

Phylogenomic Analyses: Genetic Distance

The clusters of the Neighbor-net analysis confirmed the topology found in the ML analysis, but also revealed discordant, network-like structures among particular taxa (Fig. 4; see also similar results based on single-copy nuclear genes of target enrichment data in Tomasello et al. 2020: Fig. S8). We found the deepest split among species between the *R. cassubicifolius*/*carpaticola* cluster and all other species. Within the *R. cassubicifolius*/*carpaticola* group, three clusters can be recognized: (1) *R. carpaticola* from the Southern Carpathians, (2) *R. carpaticola* (Hungary, Slovakia, Western Carpathians) plus *R. cassubicifolius* s.str. (Southern Alpine-Illyrian group), and (3) *R. cassubicifolius* s.str. (Northern Alpine group). The last two clusters particularly exhibit reticulation events at their basis. Three clusters with a similar degree of divergence and high BT support are located between the most distant groups: *R. flabellifolius* (BT = 100%), *R. cebennensis* plus *R. envalirensis* s.str. (BT = 100%), and *R. marsicus* (BT = 99.8%). Interestingly, the Illyrian cluster arises from one basis (BT = 88%), and its subclusters show a

low genetic distance with strong reticulations on the basis. Subclusters are *R. austroslovenicus*, *R. mediocompositus*, and *R. subcarniolicus*, and *R. calapius*, *R. peracris*, and *R. notabilis* s.str.

Phylogenomic Analyses: Genetic Structure

The likeliest K value of the total dataset (diploids and tetraploids) is K = 2 (ΔK = 2582.67) as ascertained by the Evanno method. The bar graph of the most likely K value exhibits two main genetic partitions (Fig. 3B, Fig. S5A): A cluster containing *R. cassubicifolius* s.l. (corresponding to the first main clade in the ML tree), and a partition containing *R. flabellifolius*, *R. envalirensis* s.l., *R. marsicus*, and the Illyrian species (corresponding to the second main clade in the ML tree).

The subset containing *R. cassubicifolius* s.l. is composed of K = 2 genetic partitions (ΔK = 24.65): *R. carpaticola* (Carpathians) and *R. cassubicifolius* s.str. (Southern Alpine-Illyrian group) form one genetic partition (violet), the Northern Alpine *R. cassubicifolius* s.str. the other one (blue; Fig. 3B, Fig. S5B). Two *R. carpaticola* individuals from the Southern Carpathians possess also a small fraction from the Northern Alpine partition. The second main clade revealed K = 3 genetic partitions (ΔK = 29.48) and almost all individuals exhibit admixture (Fig. 3B, Fig. S5B). *Ranunculus flabellifolius* is characterized by three genetic partitions, *R. marsicus* by two, and both species have a main genetic partition in common (cyan; Fig. 3B, Fig. S5B). *Ranunculus envalirensis* s.str. and *R. cebennensis* share a main genetic partition (green; Fig. 3B, Fig. S5B) although the latter showed a small proportion of admixture coming from the Illyrian group and *R. flabellifolius* or *R. marsicus*. The genotypes of the Illyrian species share a big proportion of the third genetic partition (orange in Fig. 3B, Fig. S5B). The other two genetic partitions of this group appear in smaller, varying proportions.

Phylogenomic Analyses: Target Enrichment and Coalescent-Based Species Delimitation

Results from the ASTRAL species tree estimation (Fig. 5) are comparable to the tree topology obtained with RADseq data. The main incongruences between results from the two datasets are (i) the position of *R. marsicus*, and (ii) relationships among described species in the Illyrian clade. *Ranunculus marsicus* is found to be sister to the Illyrian group in the RADseq analyses, whereas the same position is occupied by the *R. envalirensis* s.l. clade in the ASTRAL tree. However, these relationships are characterized by very short, weekly supported branches in

both analyses. Relationships among taxa of the Illyrian clade are also contrasting in the two reconstruction methods.

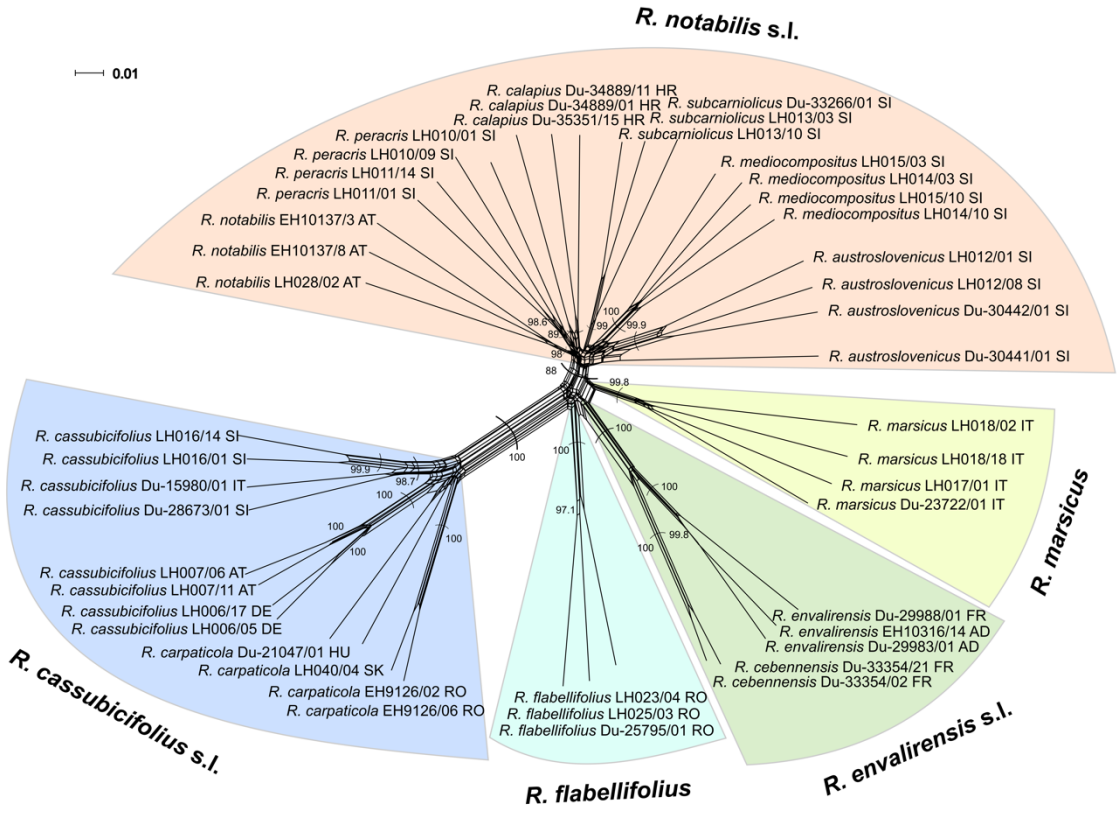


Figure 4. Neighbor-net analysis of 45 sexual *R. auricomus* individuals based on genetic distances (general time reversible [GTR] model with estimated site frequencies and ML). Bootstrap values above 80% are shown for major genetic clusters. We highlighted major clusters with colors (see Fig. 3A) and designated them with accepted names. Country labels behind taxa are in ISO code 3166-2 format.

Although samples from the same described species are found to be monophyletic (all except *R. austrosloveniсus* in the ASTRAL tree), branches describing relationships among these taxa received low support in both analyses. In the ASTRAL tree, *R. carpaticola* is also non-monophyletic.

The STACEY species delimitation analysis preferred scenarios with a reduced number of species. The "species delimitation analyser" found classifications with six genetic clusters (species) as the ones with the highest frequency (0.11). The similarity matrix (Fig. 5) showed

six well-defined clusters: Samples of (1) *R. cassubicifolius* s.l., (2) *R. flabellifolius* (sample Du-25795/02), (3) the remnant *R. flabellifolius*, (4) *R. marsicus*, (5) *R. envalirensis* s.l., and (6) all Illyrian taxa. In almost all cases, individuals from a cluster have zero posterior probability of belonging to another cluster. The only exception is the tetraploid *R. marsicus*, which samples show alternatively a discreetly high posterior probability of belonging to the cluster of the Illyrian taxa.

The cluster containing *R. cassubicifolius* s.l. is weakly differentiated and showed no grouping according to morphospecies or geographical regions. Interestingly, samples from *R. flabellifolius* formed two distinct, well-separated clusters although occurring mainly in a small, restricted area in the Southern Carpathians. The cluster containing the high-altitude species of the Pyrenees and Massif Central exhibits a weak differentiation between *R. envalirensis* s.str. and *R. cebennensis*. Even though samples from the Illyrian taxa are not clearly separable, a certain structure was recognizable within their cluster. For example, *R. mediocompositus*, *R. calapius*, and *R. peracris* formed respectively well-defined subclusters, whereas *R. mediocompositus* showed a remarkable higher similarity to *R. subcarniolicus* than to the other Illyrian taxa. In contrast, samples from the *R. notabilis* s.str. clustered separately.

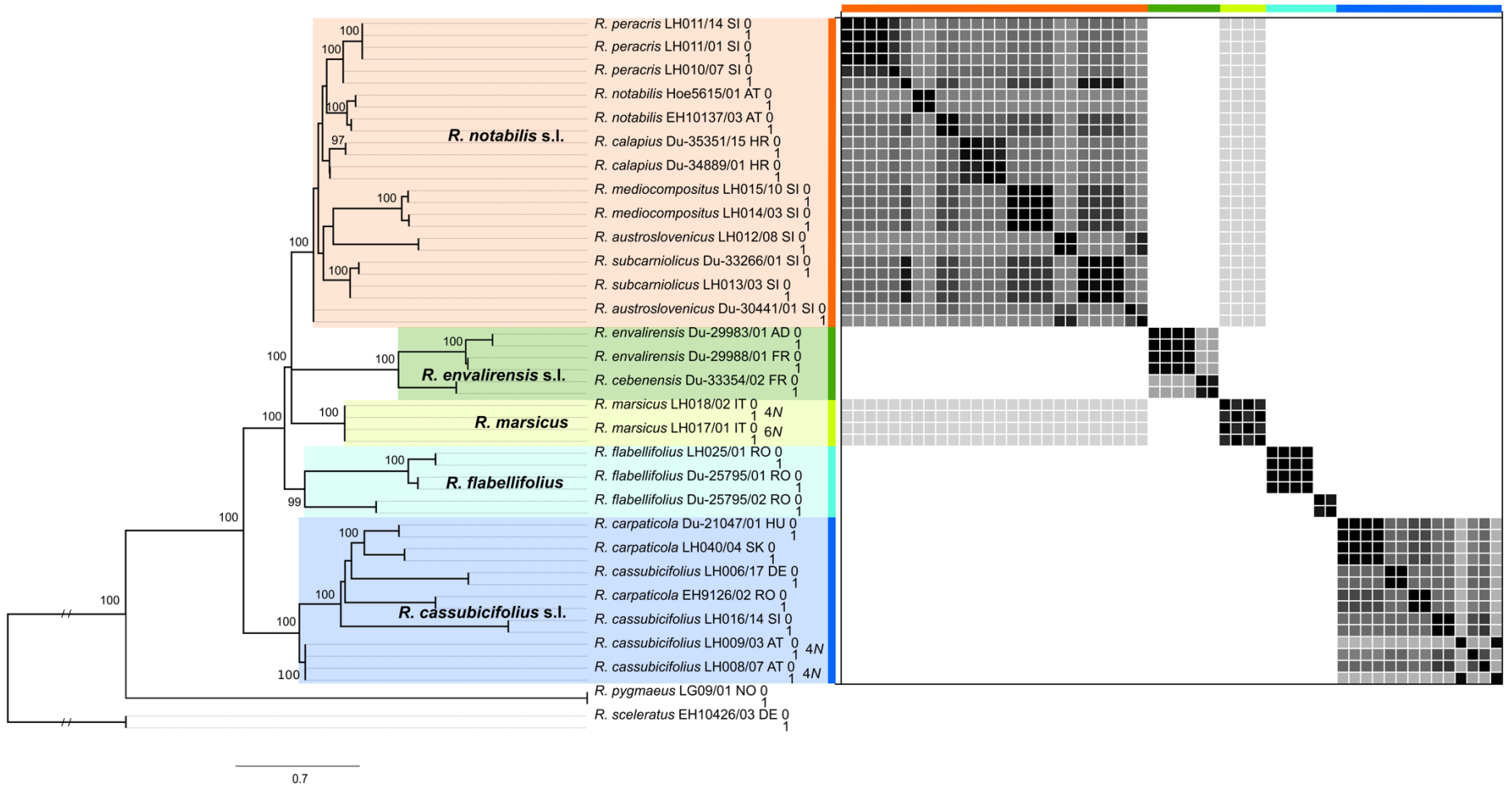


Figure 5. Coalescent-based species tree and species delimitation analyses of sexual species within the *R. auricomus* complex. The ASTRAL species tree (left) was obtained using information from 2495 exon trees. Bootstrap values above 95% are shown at nodes. The data used to infer the species tree (ASTRAL) and for species delimitation purposes (STACEY) include allelic information (phasing, is given by 0 and 1). Colored boxes around clades reflect the here proposed taxonomic treatment (as in Fig. 3A, the RADseq ML tree). Country labels behind taxa are in ISO code 3166-2 format. On the right part of the figure, we illustrated the similarity matrix obtained from the STACEY analyses. Posterior probabilities to belong to the same cluster (species) are shown for pairs of individuals. Black is for 1.0 posterior probability and white for 0.0.

Geometric Morphometrics: Morphological Gradients

RWA of basal leaves (Fig. 6A–C) pointed out a major differentiation among species following a shape gradient from dissected leaves (*R. envalirensis* s.l., *R. marsicus*, *R. notabilis* s.l.) to non-dissected leaves (*R. cassubicifolius* s.l., *R. flabellifolius*). Regarding great variation and overlap in morphospecies, we focus in the following on differentiation of the five main genetic lineages (see above). The best-separating warp (bl1) with 56% of the described variance was significantly correlated with leaf size (Spearman's $\rho = -0.6$; $p < .001$), i.e., dissected leaves were smaller and non-dissected ones were bigger. The second relative warp (bl2) with 18% of the described variance showed a leaf shape change from wide blade bases and narrow lateral segments to narrow blade bases and wide lateral segments. This morphological gradient was not associated with leaf size separating species with dissected leaves, namely *R. envalirensis* s.l. and *R. notabilis* (Fig. 6A). The third (bl3; 9%) and fourth (bl4; 5%) relative warps described shape changes concerning the middle and first lateral segment (Fig. S4A), and altogether the warps bl1–bl4 captured 88% of the variation in basal leaves. The major morphological gradient in stem leaves (sl1) with 72% of captured variation separated *R. cassubicifolius* s.l. from the most other species (Fig. 6B). The shape changes from narrow non-dissected to wide dissected stem leaves was positively correlated with size (Spearman's $\rho = 0.2$; $p < .001$). The second stem leaf warp (sl2; 15%) was not correlated with size. It described shape change from wide non-dissected to narrow dissected stem leaves and separated slightly *R. flabellifolius* from the remaining species. The warps sl1–sl4 explained 97% of the variation in stem leaves. Receptacles exhibited the lowest separating power, and neither the first (rt1; 46%) nor the second relative warp (rt2; 34%) could clearly separate the species (Fig. 6C). The first relative warp revealed a shape change from smaller wide receptacles to narrow bigger receptacles (Spearman's $\rho = 0.4$; $p < .001$). In addition, the second warp exhibited allometry (rt2; Spearman's $\rho = 0.6$; $p < .001$) describing the shape change from smaller receptacles with wide and high androclinium and short gynoclinium to bigger receptacles with short and narrow androclinium and high gynoclinium (Fig. 6C). The warps rt1–rt4 explained 92% of the variation in receptacles.

Geometric Morphometrics: Morphological Clustering

The concatenated morphometric dataset containing shape variables extracted from basal leaves, stem leaves, and receptacles enabled a better resolution of species differences. NMDS analysis of 11 shape variables supported the deep morphological split between *R. cassubicifolius* s.l. on

the one hand and *R. envalirensis* s.l., *R. marsicus*, and *R. notabilis* s.l. on the other hand (axis 1; Fig. 6D). This morphological split regards not only the most prominent character, i.e., the basal leaves, but also the stem leaves and, to a lesser extent, the receptacles. Concerning *R. flabellifolius*, the concatenated analysis confirmed its overall higher similarity to *R. cassubicifolius* s.l. than to other species. Moreover, *R. flabellifolius* was morphologically closer to southern than to northern populations of *R. cassubicifolius* s.l. (Fig. 6D). The shape of stem leaves and receptacles separated *R. flabellifolius* from *R. cassubicifolius* s.l. The second ordination axis (component 2) separated between species with dissected basal leaves (right part of the NMDS ordination plot; Fig. 6D), i.e., *R. envalirensis* s.l. and *R. notabilis* s.l. The NMDS analysis suggested that the separation between *R. envalirensis* s.l. and *R. notabilis* s.l.was mostly due to basal leaves and receptacles. Canonical variates analysis with five a priori determined groups (= species) supported the clustering of *R. envalirensis* s.l. populations apart from *R. notabilis* s.l. populations with main separating characters being the receptacles (rt1) and basal leaves (bl2) (Fig. 7). *Ranunculus notabilis* s.l. exhibited narrower and bigger receptacles than *R. envalirensis* s.l., and basal leaves finer dissected and with a wider blade base than *R. envalirensis* s.l. *Ranunculus marsicus* possessed a combination of traits: The basal leaves were *R. envalirensis* s.l.-like, the stem leaves resembled particular populations of *R. notabilis* s.l. with less-dissected leaves (*R. austroslovenicus*, *R. mediocompositus*) and the receptacles were generally *R. notabilis* s.l.-like. Non-parametric multivariate analysis of variance (NP-MANOVA) revealed significant morphological differentiation among *R. cassubicifolius* s.l., *R. envalirensis* s.l., and *R. notabilis* s.l. with overall ($p < .001$) and pairwise comparisons ($p_{cas/env} < .001$, $p_{cas/not} < .001$, and $p_{env/not} < .001$). *Ranunculus flabellifolius* and *R. marsicus* were not included here due to low sample sizes.

Indumentum of Receptacles

Trichome density of the indumentum of receptacles varied considerably within and among taxa, but it significantly differed among described and accepted taxa ($Chi^2 = 151.69$, $df = 9$, $p < .001$ and $Chi^2 = 129.97$, $df = 4$, $p < .001$; see Fig. S6A,B and Table S6A,B). *Ranunculus cassubicifolius* s.l. revealed a significant higher trichome density (12.12/0.25 mm^2) compared to the other species. Whereas *R. marsicus* is completely glabrous and *R. flabellifolius* showed almost no trichomes (0 and 0.22/0.25 mm^2), *R. envalirensis* s.str. and *R. notabilis* s.l. (1.01 and 2.25/0.25 mm^2, respectively) exhibited low receptacle trichome densities.

Ranunculus notabilis s.str. was the only taxon of the Illyrian clade that showed a remarkable trichome density (9.50/0.25 mm^2), not significantly different to *R. cassubicifolius* s.l. (see Fig. S6A and Table S6A).

Discussion

Here, we present the first comprehensive phylogenomic and geometric morphometric analyses of all hitherto known sexual species of the Eurasian *Ranunculus auricomus* complex. The delimitation of sexual species and reconstruction of a phylogenetic framework of progenitors is an important step for the classification of apomictic polyploid complexes (Grant, 1981; Burgess et al., 2015; Hörandl, 2018). We ascertained previously reported ploidy levels and a sexual mode of reproduction for all taxa by using flow cytometric methods. In general, we observed congruent patterns among ML, genetic distance in the neighbor-net network, and STRUCTURE analyses based on RADseq data (Figs. 3, 4). As an independent test, the coalescent-based species delimitation implemented in STACEY also supported scenarios with a reduced number of species. We investigated five main lineages/clusters that in some cases do not match described species. Six well-defined clusters are easily recognizable in the similarity matrix (Fig. 5). Of these lineages, five could be also characterized by morphometric data and geographical distribution, and represent species in an evolutionary sense (*R. cassubicifolius* s.l., *R. envalirensis* s.l., *R. flabellifolius*, *R. marsicus*, *R. notabilis* s.l.). For species delimitation, a combined workflow of phylogenomic analyses (Fig. 2) based on independent methods (ML, distance-based and STRUCTURE analyses, coalescent-based species delimitation) and geometric morphometrics provide a comprehensive picture of genetic and phenotypic differentiation among *R. auricomus* species.

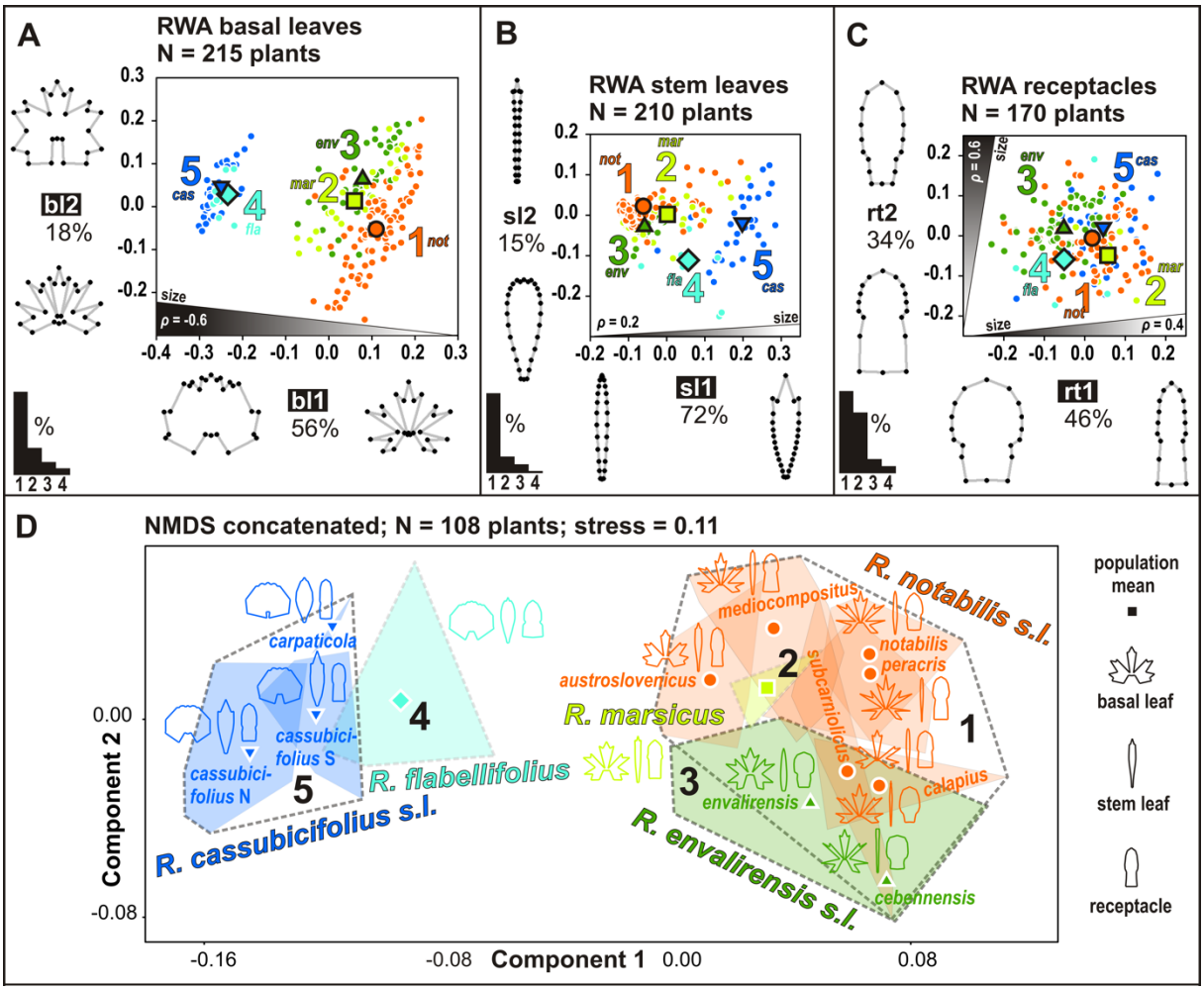


Figure 6. Morphological variation among species. **A–C**, Geometric morphometric analysis of basal leaves (**A**), stem leaves (**B**), and receptacles (**C**) based on relative warps analysis (RWA). Ordination plots of the first and second relative warps show different species clustering according to the three morphological characters. Percentages give variation explained by relative warps, and bar charts in the lower-left corners show the decrease in the explained variation by four major relative warps. The presence of significant positive or negative allometry is given as Spearman's rho (ρ) value in shaded areas close to the associated relative warps. The area width reflects the correlation strength. Landmark configurations close to relative warps illustrate shape changes. Relative warps bl1, bl2, sl1, and rt1 best differentiate among species. **D**, Morphological differentiation among species based on non-metric multidimensional scaling (NMDS) of the concatenated dataset. Convex hulls delimitate intraspecific morphological variation, and symbols show the position of population centroids. Silhouettes close to population centroids illustrate population mean shapes of all three morphological characters. The five accepted species are coded as follows (see also Fig. 3A): *R. notabilis* s.l. = 1, orange circle; *R. marsicus* = 2, light green square; *R. envalirensis* s.l. = 3, green triangle; *R. flabellifolius* = 4, turquoise diamond; *R. cassubicifolius* s.l. = 5, dark blue triangle.

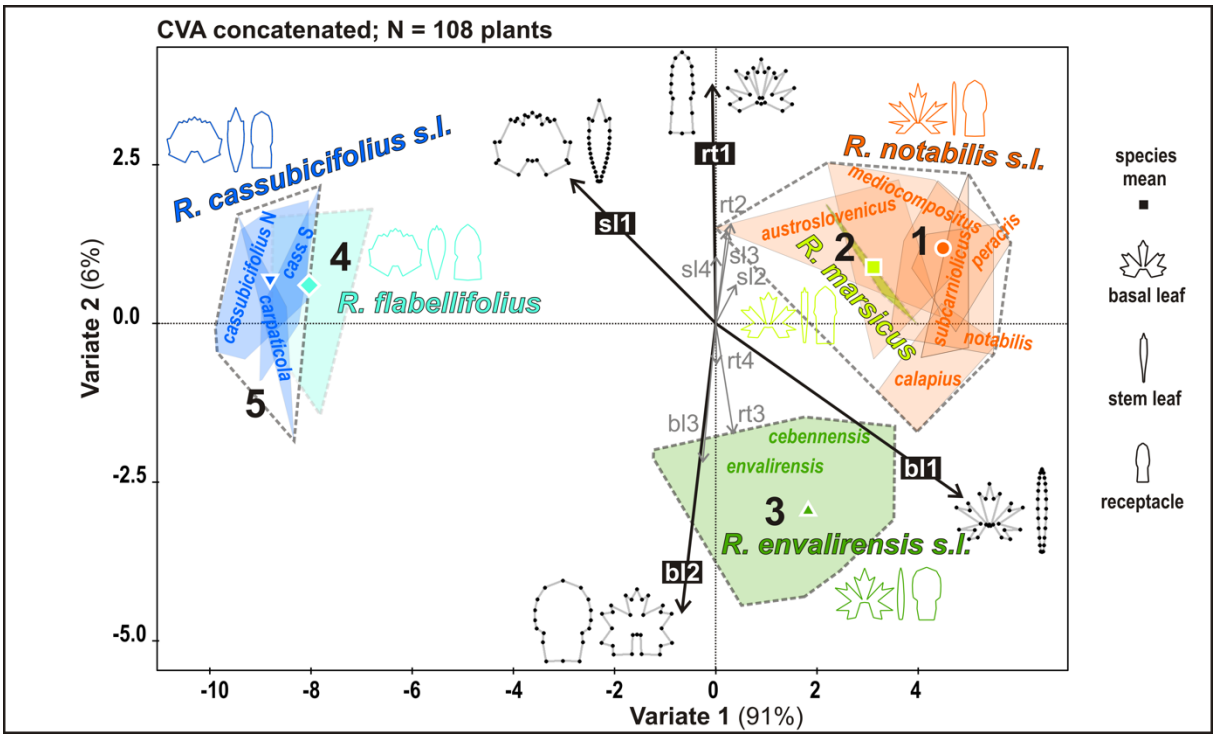


Figure 7. Morphological variation among species. Canonical variates analysis (CVA) highlights shape variables that best differentiate among species. Shape variables correspond to relative warp scores and the most important vectors (bl1, bl2, rt1, sl1) correspond to relative warps shown above (Fig. 6A–C). The five accepted species are coded as follows (see also Fig. 3A): *R. notabilis* s.l. = 1, orange circle; *R. marsicus* = 2, light green square; *R. envalirensis* s.l. = 3, green triangle; *R. flabellifolius* = 4, turquoise diamond; *R. cassubicifolius* s.l. = 5, dark blue triangle.

Phylogenomic Data

Phylogenomic analyses based on RADseq and target enrichment proved to be successful in disentangling phylogenetic relationships among sexual species in the less than 1.0 million-year-old *R. auricomus* complex. The herein presented comprehensive workflow is applicable to other evolutionary young species complexes that are shaped by low genetic divergence, hybrid origins, introgression, and ILS.

Both methods were able to resolve distinct evolutionary lineages in the sexual taxa of the *R. auricomus* complex. Results are in line with literature demonstrating that RADseq is able to unravel relationships among populations, closely related plant species, and genera more than 50 million years old (Baird et al., 2008; Hipp et al., 2014; Cavender-Bares et al., 2015; Tripp et al., 2017; Wagner et al., 2018; Pätzold et al., 2019). RADseq parameter optimization mainly aimed at the assembly of homologous loci and maximization of phylogenetic information in

concatenated ML trees resulting in a well-resolved phylogeny of the evolutionary young sexual species within the *R. auricomus* complex. However, de novo assembly of RADseq reads remains methodically challenging considering various genomic divergence levels within and among individuals/species (see, e.g., Guillardín-Calvo et al., 2019; Pätzold et al., 2019; Rancilhac et al., 2019).

Target enrichment of nuclear genes demonstrated its usefulness in resolving phylogenetic relationships above the species level, producing results comparable to those of RADseq and a well-resolved backbone of the tree topology. Phylogenetic reconstruction below species level loses power as demonstrated by the low bootstrap values obtained by the branches within the Illyrian clade. Target enrichment has in several studies shown its usefulness for resolving phylogenetic relationships above genus and family level (Mandel et al., 2014; Sass et al., 2016; Wanke et al., 2017; Vatanparast et al., 2018), and only in a few cases, it has been employed on recently radiating species groups (Stephens et al., 2015) or below the species level (Holliday et al., 2016). In *Ranunculus*, target enrichment of nuclear genes is also clearly superior over plastid markers for resolution of species-level relationships (e.g., Emadzade et al., 2015).

In general, we found congruent patterns among ML, genetic distance in the neighbor-net network, and STRUCTURE analyses based on RADseq data (Figs. 3, 4). The fully resolved and well-supported phylogeny (BT ~90%–100%) with fully informative nodes is the first one concerning all yet known sexual species within the *R. auricomus* complex. We observed the deepest split in the ML tree, the largest genetic distance in the neighbor-net network, and the most distinct genetic structuring between *R. cassubicifolius* s.l. with non-dissected basal leaves and all other species characterized by non-dissected and dissected basal leaves. Previous studies concerning isoenzyme and ITS marker supported the observed split between the most morphological diverged species *R. cassubicifolius* s.l. and *R. notabilis* s.l. (Hörandl, 2004; Hodač et al., 2014). Furthermore, the presence of the non-dissected basal-leaf species *R. flabellifolius* from the Southern Carpathians in the clade of the other species with dissected basal-leaf types clearly rejects the old morphological Linnaean classification of the complex in *R. auricomus* L. s.l. with only dissected basal leaves and *R. cassubicus* L. s.l. with only nondissected basal leaves (Linnaeus, 1753; but see Kvist, 1987). Almost all described species are monophyletic in the ML tree, but ML tree topology, neighbor-net network, and genetic structure indicate for less genetic lineages/groups than described species (Figs. 3, 4).

The clade of *R. cassubicifolius* s.l. in the ML tree is obviously genetically distant to the other species in the neighbor-net network possessing an own partition with a geographical east-west

gradient in the STRUCTURE analysis (see Fig. 3B). BT values and QS scores are high concerning the branch between Southern Carpathian populations and the other species but rather conflicted between the Northern Carpathian and the perialpine populations, suggesting introgression and/or ILS. The high conflict between northern and southern *R. cassubicifolius* s.str. populations also indicates introgression (see also Pease et al., 2018 for interpretation of QS scores). The network analysis also supported the presence of reticulation events between northern and southern *R. cassubicifolius* s.str. populations. STRUCTURE analyses revealed a northern and a southern cluster of *R. cassubicifolius* s.l. showing small fractions of admixture. Close genetic relationships between *R. cassubicifolius* s.str. and *R. carpaticola* were already investigated in Hörandl (2004), Hörandl et al. (2009), and Hodač et al. (2014), fitting to the minor morphological differences in basal and stem leaf traits observed in present geometric morphometric analyses. Intriguingly, tetraploid *R. cassubicifolius* formed together with diploid samples a specific clade in the coalescent-based species tree (Fig. 5; see for details Tomasello et al., 2020). Genetic distance and structure, and the presence of introgression and/or ILS despite geographical barriers within this disjunctly distributed species indicate the presence of one huge, though geographically partitioned, genetic lineage.

Ranunculus flabellifolius, *R. envalirensis* s.l., and *R. marsicus*, the closest species of the second main clade in the ML tree that are locally distributed in European mountains (Figs. 1, 3A), are genetically far distant from *R. cassubicifolius* s.l. and genetically close to the other Illyrian species (Fig. 4). Whereas the split between *R. flabellifolius* and the other species is fully supported (BT = 100%), the splits between *R. envalirensis* s.l. and the other species, and *R. marsicus* and the Illyrian species received decreased support in the RADseq ML tree (BT = 94% and 69%, respectively; see also the presence of discordant patterns in Fig. 3A). The split between *R. marsicus* and *R. envalirensis* s.l. plus the Illyrian species also received decreased support in the coalescent-based species tree (Fig. 5). Neighbor-net network and partitioning of genetic clusters in STRUCTURE analysis indicate three genetic clusters composed of different subclusters (except *R. envalirensis* s.str.), suggesting past reticulation events, i.e., hybrid origins and ILS, or ancient polymorphisms. QD node scores of 0.5 and 0.45 indicate both a more frequent and a less frequent alternative topology suggesting ILS, but also the possibility of hybridization with subsequent sequence evolution cannot be excluded. Despite full BT support with an absence of discordant patterns characterizing *R. flabellifolius* as a distinct lineage, the admixture in STRUCTURE and the occurrence of broad 3-lobed basal spring leaves in *R. flabellifolius* may hint at an ancient reticulate origin. Furthermore, Reichenbach (1832) and Dunkel et al. (2018) observed that *R. flabellifolius* potentially formed hybrids with other species

of the *R. auricomus* complex at the locus classicus. Hence, recent introgression may have influenced the genetic composition of *R. flabellifolius* species although individuals sampled in the present study possessed a characteristic leaf cycle and fan-shaped stem leaves. Introgression could also explain the strange pattern observed in the species delimitation analysis, in which a sample of *R. flabellifolius* (Du-25795/02) forms a distinct cluster, well separated from the other two accessions of the species. Even though from the locus classicus and morphologically identical to the others, this sample might be an introgressed individual.

The dwarf Southeastern European species *R. envalirensis* s.str. clusters together with the recently described species *R. cebennensis* occurring in the Massif Central. Both species are morphologically similar (except remarkable plant height differences) as already mentioned by Dunkel et al. (2018) and confirmed by our geometric morphometric analyses (Figs. 6D, 7). Together they form a clade in the ML tree and a cluster/partition in the other phylogenomic analyses (Figs. 3–5). Due to isolation-by-distance (~300 km, Fig. 1) that has been frequently observed in many plant species (Heywood, 1991; Ali et al., 2012), the *R. envalirensis* s.l. lineage has been probably diverged into geographically isolated, genetically weakly differentiated sublineages as indicated in the neighbor-net network (see Fig. 4).

Ranunculus marsicus, the only known sexual tetraploid species without a diploid cytotype, formed a particular clade in the ML tree and a distinct cluster in the neighbor-net network (Figs. 3, 4). Topological conflict and low support of branches involving this species suggest an allopolyploid origin. STRUCTURE analyses indicate that *R. envalirensis* s.l., *R. flabellifolius*, and a yet unknown species (Fig. 3B, Fig. S7C; cyan partitions) or even an Illyrian species may have contributed to the origin of *R. marsicus*. Interestingly, we did not observe this genetic partition of the yet unknown species in *R. cassubicifolius* s.l. (see K = 4 of the total dataset in Fig. S7C). Results from the coalescent-based species delimitation analysis indicate a possible contribution of the Illyrian clade to a putative allopolyploid origin of *R. marsicus*, as demonstrated by the moderately high posterior probability of belonging to the same genetic cluster shown by samples from both clades (Fig. 5). *Ranunculus marsicus* and *R. envalirensis* s.l. are the only sexual lineages of the *R. auricomus* complex adapted to subalpine environments. Contact zones between members of the Illyrian group and the other progenitors of a putative allopolyploid *R. marsicus* might have occurred in the Pleistocene due to distributional range shift towards south (of an Illyrian progenitor) and/or expansion (of the putative subalpine progenitor) during cold periods (see Feliner, 2011; Hewitt, 2011; Tomasello et al., 2020). A potential current or past present diploid species in the Central Apennines (Dunkel et al., 2018) may have also been a progenitor of *R. marsicus*.

Masci et al. (1994) observed diploid behavior, i.e., predominantly two alleles per isoenzyme locus in sexual *R. marsicus*, potentially suggesting an allopolyploid origin for *R. marsicus*. However, during the filtering processes, only two alleles were allowed (removing loci with more than two alleles), potentially biasing allelic diversity of *R. marsicus*. Further research is needed, and the application of species network reconstruction approaches (e.g., Yu & Nakhleh, 2015; Oberprieler et al., 2017; Solís-Lemus et al., 2017; Wen & Nakhleh, 2018) might help to unequivocally shed light on the evolution of this tetraploid species. Regarding the Illyrian species group that comprises *R. notabilis* s.str. and five recently described species (Dunkel et al., 2018), bootstrap support is decreased, whereas discordant patterns are remarkably increased in the RADseq ML tree (Figs. 3, 4). Genetic distance is low among these species sharing one characteristic genetic partition in varying proportions (orange part; Figs. 3B, 4). Discordant tree topologies and the lack of species-specific partitions suggest the presence of introgression and/or ILS. Species of the Illyrian group, except *R. notabilis* s.str., occur in close neighborhood (see Fig. 1; sometimes four species nearby [1 km^2], see Dunkel et al., 2018). At least some species have overlapping distributions (Dunkel et al., 2018), but ecological differentiation has not yet been investigated. However, the absence of crossing barriers (Rahmsdorf, 2019; Karbstein et al., subm. b) apparently allows ongoing gene flow. The genetic differentiation process of the Illyrian clade, which probably started 200,000 years ago and lasts until today, might have caused the observed ILS pattern (see Tomasello et al., 2020). Overall, the Illyrian species group of the *R. auricomus* complex represents one genetic lineage that is weakly geographically structured.

Geometric Morphometrics

Our results quantified the total phenotypic variation concerning basal and stem leaves, and receptacles exhibited by the five sexual species of the *R. auricomus* complex. In accordance with the phylogenomic data, morphometrics pointed out clear morphological clusters corresponding to *R. cassubicifolius* s.l., *R. envalirensis* s.l., and *R. notabilis* s.l. Within the latter, results revealed a morphological differentiation between the most distant morphotypes *R. austroslovenicus* (broad 3-dissected basal leaves) and *R. peracris* (5-dissected basal leaves) but all other described species as intermediates, indicating population structure of a highly variable species rather than morphologically and genetically well-separated taxa.

The three morphological clusters could be distinguished after combining information from basal leaves, stem leaves, and receptacles. When analyzed separately, the morphological

characters exhibited only limited species resolution due to overall high phenotypic variation (Fig. 6A–C). The other two species, *R. flabellifolius* and *R. marsicus*, occupy small and restricted distribution areas with only a few known populations, so that their morphological variation is low, particularly for *R. marsicus*. *Ranunculus flabellifolius* is morphologically close to *R. cassubicifolius* s.l. due to non-dissected basal leaves, whereas *R. marsicus* is more similar to *R. notabilis* s.l. due to roundish 3-lobed basal leaves and linear, sometimes dissected stem leaves in concatenated analyses (Figs. 6D, 7).

In general, leaf shapes provided better overall taxonomical resolution than receptacles. In concordance with morphological studies (e.g., Borchers-Kolb, 1983, 1985; Hörandl, 2002; Hörandl et al., 2009), geometric morphometrics confirmed that shape of basal spring leaves discriminates best among the sexual species in the *R. auricomus* complex. Non-dissected basal leaves were expectedly bigger than finely dissected ones (particularly within *R. notabilis* s.l.). Apart from the most prominent shape gradient distinguishing *R. cassubicifolius* s.l. from *R. envalirensis* s.l./*R. notabilis* s.l., the basal leaves showed further levels of finer and size-independent shape variation. These minor shape vectors possibly reflect intraspecific variation among populations belonging to the species *R. notabilis* s.l. The total variation of basal leaves extracted from all five sexual species together did not exceed the variation revealed in experimental crossings between *R. cassubicifolius* s.l. and *R. notabilis* s.l. (Hodač et al., 2018). Moreover, the first three major morphological gradients observed in the five sexual species (83% of the total variability) were identical with those revealed in the crossing study by Hodač et al. (2018). However, the genetic structure would not support a hypothesis of a hybrid origin of the species with "intermediate" leaf morphotypes (*R. envalirensis* s.l., *R. marsicus*, *R. flabellifolius*), from the *R. cassubicifolius* and *R. notabilis* s.l. lineages. Instead, it seems that the whole spectrum of leaf shapes is already present in progenitor species.

The major morphological trend in stem leaves distinguished between *R. envalirensis* s.l./*R. notabilis* s.l. and *R. cassubicifolius* s.l. The gradient from small narrow non-dissected stem leaf segments towards broader and dissected ones also showed a significant covariation with distinct basal leaf shape changes. The outstanding and species-specific fan-shaped stem leaf of *R. flabellifolius* distinguished this species from the other four ones. The associated morphological trend was independent of leaf size but again associated with basal leaf variation. Despite less basal and stem leaf differentiation to *R. notabilis* s.l. in geometric morphometric analyses, the subalpine *R. envalirensis* s.str. and *R. marsicus* have a remarkably lower plant height (5–15 cm) compared to the Illyrian lowland species (10–30 cm). We observed that *R. envalirensis* s.str. and *R. marsicus* keep their dwarfism that is typical for alpine plants (Körner, 2003) also under

garden culture (2017–2020). In *R. envalirensis* s.str., this tiny habit was kept even in garden offspring raised from seeds (pers. obs. of authors).

Interestingly, although basal and stem leaves covariate in their shape changes, they show different morphological gradients. Although the stem leaf middle segments are homologous to middle segments of basal leaves, however, we did not find any relationship between the variance of these structures. Furthermore, trichome density of receptacles remarkably varied within and among species but discriminated *R. cassubicifolius* s.l. with remarkably higher trichome density from the other accepted species. However, receptacles of the Illyrian species were not glabrous as described by Dunkel et al. (2018). Variation of the indumentum of the receptacle is in general much higher than it was described in previous taxonomic studies (Grau, 1984; Hörandl & Gutermann, 1998c; Dunkel et al., 2018).

From the three morphological shape characters under study, receptacles showed the lowest discriminating power among species, and its major shape changes were significantly affected by size variation. This allometry resulted in a general gradient from bigger receptacles (with mostly elongated gynoclinium plus narrow androclinium) towards smaller receptacles (with mostly spherical gynoclinium plus wider androclinium). As mentioned above, variation in receptacles had a limited taxonomic resolution when analyzed separately from other traits. However, distinct vectors of receptacle variation showed significant covariation with shape changes in other morphological characters. For example, a size-independent morphological trend from receptacles with elongate gynoclinium (plus wide androclinium) towards receptacles with spherical gynoclinium (plus narrow androclinium) covaried with shape changes of basal and stem leaves. In the Nordic species of the complex, characters of the receptacle were regarded as fairly constant, but their variation is also influenced by different environmental conditions (Ericsson, 2001). Phenotypic plasticity of all morphological characters, however, needs to be studied.

Species Delimitation

In general, the results of the present study raised the need for a re-classification of the described sexual species within the *R. auricomus* complex. On the one hand, an evolutionary ancestor-descendant lineage concept, including the evolutionary role and circumscription of lineages, should be preferred for species delimitation from the conceptual point of view (De Queiroz, 2007; Freudenstein et al., 2017; Sukumaran & Knowles, 2017; Hörandl, 2018). However, as pointed out by Sukumaran & Knowles (2017), lineages could also reflect an infraspecific

population structure, and lineages should not be classified as species without involving additional criteria. On the other hand, the genetic cluster concept, where a species is defined as a "morphologically or genetically distinguishable group of individuals that has few or no intermediates when in contact with other such clusters" (Mallet, 1995), is also well-applicable when considering the present distance-based genetic and morphometric analyses. A clustering criterion helps specifically for the recognition of distinct morphological entities in groups with a high phenotypic variation and a lack of exclusive diagnostic characters, as we observe it in the *R. auricomus* complex. Combinations of leaf and fruit characters rather than single diagnostic features characterize also species within the whole genus *Ranunculus* (Hörandl & Emadzade, 2012). The biological species concept (Mayr, 1942), requiring crossing barriers, is partly supported by the strongly reduced fertility of hybrids between *R. cassubicifolius* s.l. and *R. notabilis* s.l., i.e., the genetically most distant lineages (Hörandl, 2008). Disturbances of meiosis, megasporogenesis, and pollen formation in these hybrids are possible triggers for apomixis (Hojsgaard et al., 2014; Barke et al., 2018; Barke et al., acc.). Interestingly, experimental crossings of the Illyrian taxa revealed a comparable high seed set between inter- and intraspecific crossings, supporting our conclusion to merge them all into one species *R. notabilis* s.l. (Rahmsdorf, 2019; Karbstein et al., subm. b). Some tendencies to self-fertility in the Slovenian populations might result in local lineages and morphotypes (Rahmsdorf, 2019; Karbsteinet al., subm. b). Detailed heterozygosity investigations of taxa will be presented elsewhere (Karbstein et al., subm. a, b). Moreover, a rare occurrence of asexual seeds in single individuals of otherwise diploid sexual taxa does not justify a separate taxonomic recognition (Hörandl, 2018). The appearance of rare, asexual seed formation within diploid sexual populations has been observed in other species as well (Schinkel et al., 2016) and may be due to environmental influence (Klatt et al., 2018).

Geographical and altitudinal isolation are the most important factors separating the five main species from each other: *R. cassubicifolius* s.l. in the N and SE edges of the Alps and the Carpathians; *R. flabellifolius* restricted to the Southern Carpathians and Serbia; *R. envalirensis* s.l. to the Massif Central and SE Pyrenees; *R. marsicus* to the Central Apennines; and *R. notabilis* s.l. to the Illyrian region (Fig. 1). We suppose that allopatric speciation has shaped the evolution of species (Tomasello et al., 2020). Geographical isolation may have further led to genetic differentiation between the Western and Southern Carpathian and the northern perialpine populations of *R. cassubicifolius* s.l. and between the Pyrenees and Massif Central within *R. envalirensis* s.l. Low amounts of self-compatibility may have resulted in local differentiation of the sympatric *R. notabilis* s.l. populations in Slovenia. It might be useful to

recognize geographical/ecological entities as subspecies, but more information is needed on ecology and population genetics of the respective taxa. For instance, disjunct *R. cassubicifolius* s.l. might be better split into several regional subspecies (instead into two taxa, *R. cassubicifolius* s.str. and *R. carpaticola*). Considering polyploid *R. marsicus*, both ploidy-level and geographical isolation separate it from the other species.

We, therefore, propose a new circumscription for the sexual species of the *R. auricomus* complex by rejecting the old classification with purely morphological descriptions and by applying an evolutionary lineage concept based on phylogenetic trees combined with a cluster concept based on distinct genetic and morphological groups. Following this approach, we would have three options of lumping vs. splitting of species: We could define (1) two main genetic or morphological lineages/clusters, that are not congruent to each other due to the variable position of *R. flabellifolius*; (2) five genetic geographically isolated lineages/clusters with congruent results in phylogenomic and geometric morphometric data, neglecting particular weak genetic and morphological (or local, ecotypic) differentiation; or (3) twelve mostly monophyletic, lineages lacking genetic and morphological divergence.

Though monophyly is widely supported by the RADseq and target enrichment phylogenies (except *R. austroslovenicus* and *R. carpaticola*), the other phylogenomic approaches revealed a more comprehensive picture of relationships: We found the best congruence of all phylogenomic and morphometric data with five distinct lineages/clusters. Monophyletic lineages do not necessarily represent species (see, e.g., Sukumaran & Knowles, 2017), particularly in the *R. notabilis* s.l. group with estimated speciation events less than 100,000 to 200,000 years ago (Tomasello et al., 2020), extremely low genetic distances (see Fig. 4), and full crossability (Rahmsdorf, 2019; Karbstein et al., in subm. b).

We also showed that results of geometric morphometrics were more reliable than standard morphological approaches by capturing various dimensions of taxonomically important traits in a single analysis and by uncovering the same clusters as phylogenomic analyses. However, conclusions on species relatedness are slightly different between phylogenomics and geometric morphometrics (see *R. flabellifolius*). According to the applied species concepts, phylogenomics has to be favored (see, e.g., non-dissected leaves of *R. cassubicifolius* s.l. and *R. flabellifolius*). The fact that taxa within *R. notabilis* s.l. are mostly monophyletic may be also due to sampling of only two, neighboring populations (near locus classicus) sharing similar morphotypes. Sampling close to the locus classicus should ensure incorporation of correct taxa described by Dunkel et al. (2018) because taxon recognition in the field is almost impossible

(no identification key of Illyrian taxa is available, basal leaf cycle can vary according to site conditions). We would get probably more morphological intermediates if the sampling comprises the complete Illyrian region. However, genomic differences are minor between Illyrian lineages, supporting a concept lumping them into *R. notabilis* s.l. though apparent monophyly.

Therefore, we accept five, morphologically differentiated, geographically (and altitudinally) well-separated, genetic main lineages as species, to be named after the rules in the ICN (Turland et al., 2018): *R. cassubicifolius* s.l. (incl. *R. carpaticola*), *R. envalirensis* s.l. (incl. *R. cebennensis*), *R. flabellifolius*, *R. marsicus*, and *R. notabilis* s.l. (incl. *R. austroslovenicus*, *R. calapius*, *R. mediocompositus*, *R. peracris*, *R. subcarniolicus*; see also Fig. S8A–E). These five taxa will provide also a robust phylogenetic framework for the analysis of origin and evolution of their apomictic derivatives (Hörandl, 2018; Karbstein et al., in prep.).

Taxonomic Treatment

Ranunculus cassubicifolius W.Koch in Ber. Schweiz. Bot. Ges. 49: 553. 1939 – Holotype: Switzerland, Kanton Solothurn: Wald östlich Erlenmoss südöstlich Recherswil längs eines Baches am Waldrand, ca. 465 m, 03 May 1938, *W. Koch* 38/56 (ZT barcode ZT-00035121!).

= *Ranunculus carpaticola* Soó in Acta Bot. Acad. Sci. Hung. 10: 224, in clavi, 231. 1964 – Holotype: "in silvis montanis ap. Várhegy pr. opp. Huszt, 14. 8. 1960, R. de Soó" (BP!) – Epitype (designated by Hörandl et al. in Taxon 58(4): 1212. 2009): "*Ranunculus carpaticola*. Flora Slovaciae centralis, Skalka nad Revúca, *J. Májovsky s.n.*, 1972, cyto: 2n = 16" (SAV!).

Comments. – *Ranunculus cassubicus* auct. europ. p.p. (e.g., Jalas & Suominen, 1989; Tutin & Akeroyd, 1993) non L. (see Kvist, 1987) is a misapplied name to this taxon.

Distribution. – Austria, Germany, Hungary, Italy, Romania, Slovakia, Slovenia, Croatia, Switzerland (Borchers-Kolb, 1985; Hörandl et al., 2009; Dunkel, 2010; this paper, Table 1). This species turned out to be more widely distributed and more variable than anticipated by previous authors. An infraspecific classification recognizing its disjunct distribution appears desirable but requires further investigations, including the type area and also the autotetraploid sexual cytotypes in Lower Austria (Hörandl & Greilhuber, 2002). The polyploid hybridogenetic apomictic derivatives were excluded from the diploid sexual species (Hörandl et al., 2009).

Ranunculus envalirensis Grau in Mitt. Bot. Staatssamml. München 20: 14. 1984 ≡ *Ranunculus auricomus* subsp. *envalirensis* (Grau) Molero, J.Pujadas & Romo in Monogr. Inst. Piren. Ecol. 4: 274. 1988 – Holotype: Andorra. Alpine Matten zwischen Soldeu und dem Puerto de Envalira, ca. 1900 m, 23 May 1970, *Merxmüller & Gleisner* 25864 (M barcode M-0008123!).

= *Ranunculus cebennensis* Dunkel et al. in Willdenowia 48(2): 237. 2018 – Holotype: France, Auvergne, Puyde-Dôme, Verbindungsstraße zwischen D29 und D203, S La Chaleille [connecting road between D 29 and D 203, S La Chaleille], 45_33′38″N, 02_39′25″E, 910 m, Hangweide [steep meadow], 6 May 2016, F.G. Dunkel 33354 (LY!; isotype: M barcode M-0291696!).

Distribution. – Andorra, France, Spain (Grau, 1984; Dunkel et al., 2018; this paper, Table 1). Already Landolt in Jalas & Suominen (1989) supposed that the diploid populations in the Eastern Pyrenees and the Massif Central match *R. envalirensis*.

Ranunculus flabellifolius Heuff. ex Rchb., Fl. Germ. Excurs.: 723. 1832 ≡ *Ranunculus cassubicus* var. *flabellifolius* (Heuffel ex Rchb.) Borza, Consp. Fl. Romaniae 1: 103. 1947 – Lectotype (designated by Dunkel et al. in Willdenowia 48(2): 252. 2018): Romania, Banat, "in dumetis sylvisque montius [sic!] calc. Banatus", Apr/May ca. 1830, *J.A. Heuffel* (JE barcode JE00007661!).

Distribution. – Western Romania, Serbia (Dunkel et al., 2018; this paper, Table 1).

Comment. – *Ranunculus cassubicus* auct. europ. p.p. (e.g., Jalas & Suominen, 1989; Tutin & Akeroyd, 1993) non L. (see Kvist, 1987).

Ranunculus marsicus Guss. & Ten., Fl. Napol. 5, Appendix Quarta: IX. 1835–1836 – Lectotype (designated by Dunkel in Webbia 66(2): 171. 2011): [Italy, Abruzzo, province of L'Aquila], Piano de 5 miglia, M. Tenore s.n. (NAP!).

Distribution. – Central Italy (Masci et al., 1994; Dunkel, 2011; this paper, Table 1). Dunkel (2011) discriminated three varieties: var. *marsicus*, var. *incisior* Dunkel, and var. *approximans* Dunkel, whereby the typical variety represents the tetraploid, sexual populations that we analyzed here. The two other varieties have more deeply dissected basal leaves and were interpreted as introgressants or hybrids with other co-occurring species of the complex (Dunkel, 2011). Introgression may also explain the occurrence of apomixis in higher ploidy levels.

Ranunculus notabilis Hörandl & Gutermann in Phyton (Horn) 37: 268. 1998 – Holotype: Österreich, Burgenland, Süd-Burgenland: 8964/4: Stremtal, Moschendorfer Wald, am Rand 1,5 km ENE Strem, 220 m, Feuchtwiese, 1 May 1994, *E. Hörandl & W. Gutermann* Hö5613 = Gu27897 (WU No. 033476!; isotype: GOET barcode GOET014403!, M, W Nos. 2006-12661!, 2006-12662!, 2006-12663! & 2006-12664!).

= *Ranunculus austroslovenicus* Dunkel in Willdenowia 48(2): 230. 2018 – Holotype: Slovenia, Preddinarsko Območje, 0554.2, Kočevje, an der Kolpa 1 km vor Srobotnik, [on the rivulet Kolpa 1 km before Srobotnik], 45°29′43″N, 14°48′14″E, 250 m, Gebüsch, Waldrand [shrubbery, forest edge], 23 Apr 2013, *F.G. Dunkel* 30441 (LJU!; isotype: M barcode M-0291693!).

= *Ranunculus calapius* Dunkel in Willdenowia 48(2): 234. 2018 – Holotype: Croatia, Turopolje, 0459.3, Karlovac, Orlovac, 1.1 km SE des Ortes [of the village], W Sisačka ulica, 45°31′08.2″N, 15°34′15.0″E, 115 m, feuchter Hainbuchen-Auenwald [swampy hornbeam alluvial forest], 10 Apr 2017, *F.G. Dunkel* 34889 (ZA!; isotype: M barcode M-0291695!).

= *Ranunculus mediocompositus* Dunkel inWilldenowia 48(2): 243. 2018 – Holotype: Slovenia, Predalpsko Območje, 0151.4, Logatec Planinsko Polje, E Liplje, E der Straße Planina-Laze, N of P. 446, 50–100 m S Flüsschen Unica [E of road Planina–Laze, N of P. 446, 50–100 m S of little river Unica], 45°50′58″N, 14°15′32″E, 445 m, magere Wirtschaftswiese, Flachmoor [nutrient-poor meadow, mire], 24 Apr 2012, *F.G. Dunkel* 28639 (LJU!; isotype: M barcode M-0291697!).

= *Ranunculus peracris* Dunkel in Willdenowia 48(2): 246. 2018 – Holotype: Slovenia, Unterkrain (Dolenjska), 0158.1, 1,6 km S Gmajna an der Straße nach Zameško, O der Straße [on the road to Zameško, E side], 45°53′24″N, 15°22′09″E, 155 m, Wiese, Waldrand und angrenzender Hainbuchen-Wald [meadow, edge of the forest and adjacent hornbeam forest], 23 Apr 2013, *F.G. Dunkel* 30446 (LJU!; isotype: M barcode M-0291698!).

= *Ranunculus subcarniolicus* Dunkel in Willdenowia 48(2): 249. 2018 – Holotype: Slovenia, Preddinarsko območje, 0053.2, Grosuplje, W des Ortes, ca. 150 m N der Ponova Vas, 250 m W des Flüsschens Bičje [W of the village, 250 m W of the brook Bičje], 45°56′46″N, 14°38′54″E, 325 m, magere Wirtschaftswiese [nutrient-poor meadow], 8 Apr 2017, *F.G. Dunkel* 34772 (LJU!; isotype: M barcode M-0291699!).

Distribution. – Austria, Croatia, Slovenia (Hörandl & Gutermann, 1998c; Dunkel et al., 2018; this paper, Table 1).

Identification Key

(Only for di- and tetraploid sexual species; a general identification key for the whole complex is in progress. Pollen and reproductive data after Hörandl et al., 1997; Dunkel et al., 2014, 2018, Karbstein et al., subm. a.)

1. Pollen quality (stainability of grains) below 90% (mainly below 70%; tri- to octoploid, less than 50% asexually formed seeds within populations..... **apomictic lineages**

1. Pollen quality (stainability of grains) above 90%, di- to tetraploid, more than 85% sexually formed seeds within populations, flowers always complete, with 5(–8) petals .....**sexual species**....2

2. Basal spring leaves broad, non-dissected (rarely a 3-lobed one). The middle segment of the lowermost stem leaf (broad) lanceolate with a variable number of teeth at the leaf tip, or complete lowermost stem leaf fan-like 3

2. Basal spring leaves 3-lobed to dissected. The middle segment of the lowermost stem leaf linear to narrow lanceolate, without teeth at the leaf tip, but sometimes with deep sinuses 4

3. The middle segment of the lowermost stem leaf only at the basis connected to the other segments; therefore, the lowermost stem leaf segmented, not fan-like. Receptacles densely pilose ..... ***R. cassubicifolius* s.l.**

3. The middle segment of the stem leaf connate with the other segments; therefore, the lowermost stem leaf fanlike. Receptacles almost glabrous ..... ***R. flabellifolius***

4. Basal spring leaves lobed to dissected into 3–5 segments. Plants 15–30 cm high. Forest zone of the Illyrian region ..... ***R. notabilis* s.l.**

4. Basal spring leaves lobed to dissected into 3 (rarely 5), rather roundish, segments. Tiny plants (5–)10(–15) cm high. Subalpine zone of the Apennines, Massif Central, and Pyrenees 5

5. The middle segment of the lowermost stem leaf linear with (sometimes deep) sinuses. Central Apennines ..... ***R. marsicus***

5. The middle segment of the lowermost stem leaf linear without sinuses. Massif Central, Eastern Pyrenees ..... ***R. envalirensis* s.l.**

Author Contributions

K.K. and E.H. designed research; E.H., F.G.D., and L.H. sampled plant material. K.K., L.H., F.G.D., M.D., and S.T. collected data; K.K., L.H., and S.T. analyzed data. K.K. wrote the manuscript with assistance from L.H., M.D., S.T., and E.H. K.K., https://orcid.org/0000-0003-1424-6557; S.T., https://orcid.org/0000-0001-5201-1156; L.H., https://orcid.org/0000-0002-6885-1317; E.H., https://orcid.org/0000-0002-7600-1128

Acknowledgments

We thank the German Research Foundation for project funding (DFG, Ho4395/10-1) to EH, within the priority program "Taxon-Omics: New Approaches for Discovering and Naming Biodiversity" (SPP 1991). We also acknowledge Jennifer Krüger, Birthe H. Barke, and Julius Schmidt for technical help, Claudia Pätzold for the bioinformatic support, and Silvia Friedrich for curating the garden plants. We thank the Botanical Museum, University of Oslo (O) for silica-gel dried leaf material of *R. pygmaeus*. Open access funding enabled and organized by Projekt DEAL. Open access funding enabled and organized by Projekt DEAL.

Chapter 2 – The Breeding System of Diploid Sexuals

Breeding system of diploid sexuals within the *Ranunculus auricomus* complex and its role in a geographical parthenogenesis scenario

Karbstein, K., Rahmsdorf, E., Tomasello, S., Hodač, L., & Hörandl, E.

Ecology and Evolution, *10*, 14435–14450 (2020c).
doi: 10.1002/ece3.7073 (CC BY 4.0, changes in format and citation style)

The larger distribution area of asexuals compared with their sexual relatives in geographical parthenogenesis (GP) scenarios has been widely attributed to the advantages of uniparental reproduction and polyploidy. However, potential disadvantages of sexuals due to their breeding system have received little attention so far. Here, we study the breeding system of five narrowly distributed sexual lineages of *Ranunculus notabilis* s.l. (*R. auricomus* complex) and its effects on outcrossing, inbreeding, female fitness, and heterozygosity. We performed selfing and intra- and interlineage crossings by bagging 481 flowers (59 garden individuals) followed by germination experiments. We compared seed set and germination rates, and related them to genetic distance and genome-wide heterozygosity (thousands of RADseq loci). Selfings (2.5%) unveiled a significantly lower seed set compared with intra- (69.0%) and interlineage crossings (69.5%). Seed set of intra- (65%) compared to interpopulation crossings (78%) was significantly lower. In contrast, all treatments showed comparable germination rates (32%–43%). Generalized linear regressions between seed set and genetic distance revealed positive relationships in general and between lineages, and a negative one within lineages. Seed set was the main decisive factor for female fitness. Germination rates were not related to genetic distance at any level, but were positively associated with heterozygosity in interlineage crossings. Experiments confirmed full crossability and predominant outcrossing among sexual *R. notabilis* s.l. lineages. However, up to 5% (outliers 15%–31%) of seeds were formed by

selfing, probably due to semi-self-compatibility in a multi-locus gametophytic SI system. Less seed set in intrapopulation crossings, and higher seed set and germination rates from crossings of genetically more distant and heterozygous lineages (interlineage) indicate negative inbreeding and positive outbreeding effects. In GP scenarios, sexual species with small and/or isolated populations can suffer from decreased female fitness due to their breeding system. This factor, among others, probably limits range expansion of sexuals.

Keywords

experimental crossings, genetic distances, genome-wide heterozygosity, outcrossing, RADseq, *Ranunculus auricomus*, selfing

Introduction

Geographical parthenogenesis (GP) describes the phenomenon by which asexual taxa have larger distribution areas than their sexual relatives (Hörandl, 2006; Kearney, 2005; Vandel, 1928). Several factors may contribute to the pattern, whereby most hypotheses focus on the potential advantages of asexual taxa. In general, uniparental reproduction allows for a faster establishment of populations, and hence more efficient colonization of devastated areas (Hörandl, 2006). Reproductive assurance in the case of uniparental reproduction is a major advantage for asexuals in GP scenarios (e.g., Lo et al., 2013). Additionally or alternatively, positive side effects of polyploidy of asexual lineages may allow for niche shifts towards colder or more extreme climatic conditions (Bierzychudek, 1985; Lo et al., 2013; Paule et al., 2018). The success of asexuals can be also related to differential niche dynamics of clonal lineages vs. sexual species (general-purpose vs. niche-specific genotypes; Vrijenhoek & Parker, 2009).

So far little attention has been paid to the potential disadvantages of sexual species in GP scenarios. Sexuality in plants and animals is by far the most abundant and widespread mode of reproduction, and benefits from short-term advantages of recombination (Burt, 2000; Maynard-Smith, 1978). In most cases of GP scenarios, however, the sexual relatives of asexual complexes occur in relatively small distribution areas that are quite often interpreted as relict areas in glacial refugia during the Pleistocene (Bierzychudek, 1985; Cosendai & Hörandl, 2010;

Hörandl et al., 2008; Lo et al., 2009). It is still enigmatic why sexual populations fail to expand their range in postglacial recolonization scenarios despite the ample availability of suitable habitats (e.g., Kirchheimer et al., 2018; Tomasello et al., 2020). Although differential niche dynamics, the occurrence of apomictic conspecifics, and ecological preferences can explain a geographical separation of sexual and asexual lineages in some cases (Karunarathne et al., 2018; Kirchheimer et al., 2018; Nardi et al., 2020), others do not confirm a scenario of niche separation (Mau et al., 2015). Sexual species are otherwise (without asexual congeners) also the most abundant colonizers in extremely high altitudes and latitudes (Asker & Jerling, 1992; Brožová et al., 2019; Hörandl et al., 2011). It remains an open question why sexual species are unsuccessful in GP scenarios.

In flowering plants, the diversity of breeding systems in sexual species is an important factor for these considerations. Most angiosperms are hermaphroditic and hence, self-fertilization is a frequent option for reproduction. Selfing provides reproductive assurance even for isolated single plants or populations (Renner, 2014; Richards, 1997). Indeed, self-fertility has been recognized as a general short-term advantage for colonization and range expansion, when mating partners and pollinators are rare (Baker's law, Baker, 1967; Pannell, 2015). However, sexual selfing has the disadvantage of a rapid loss of heterozygosity over generations, which can cause inbreeding depression (Charlesworth & Charlesworth, 1987; Goldberg et al., 2010; Richards, 1997; Freeland et al., 2011). In contrast, asexual selfing in apomicts fixes the existing heterozygosity over generations due to the lack of meiosis and recombination (Brochmann et al., 2004; Hörandl, 2006). Most members of sexual-apomictic complexes are long-lived perennials (Asker & Jerling, 1992), for which the effects of inbreeding are expected to be more severe than in annuals. Purging effects of the mutational load from one generation to the next (Barrett & Charlesworth, 1991; Crnokrak & Barrett, 2002) are faster and more efficient in obligate selfing annuals than in perennials (Richards, 1997). Beside selfing, biparental inbreeding in small and isolated outcrossing populations could have also negative effects on fitness as has been suggested for example by experiments in *Hieracium* (Pinc et al., 2020). Haag and Ebert (2004) hypothesized that asexual colonizers in marginal populations would suffer less from genetic bottlenecks and subsequent drift than sexual populations. Fitness of sexuals would be reduced in such small colonizer populations due to inbreeding depression. Comprehensive population genetic studies on GP scenarios, however, either did not observe the respective genetic pattern (e.g., Cosendai et al., 2013) or found the colonization history and decline of genetic diversity in marginal sexual populations as less relevant for the GP pattern (e.g., Nardi et al., 2020). On the contrary, apomictic lineages can benefit from uniparental

reproduction either via pollen-independent (autonomous) apomixis (Mráz et al., 2018) or via pseudogamy and self-compatibility (Hörandl, 2010) without negative effects of inbreeding. However, little is known about the effects of breeding systems, inbreeding, and fitness of sexual species in geographical parthenogenesis scenarios.

Most sexual progenitors of apomictic plants are reported to be self-incompatible (Alonso-Marcos et al., 2018; Cosendai et al., 2013; Hörandl, 2010; Vašková & Kolarčik, 2019). Hence, they are trapped in a twofold dilemma: As outcrossers, they do need conspecific mating partners and pollen vectors, which is problematic in small and isolated populations. Followingly, the founding of populations from single or few individuals is hampered after long-distance dispersal (Baker's law; Pannell, 2015). However, also dispersal and gene flow with other, more adjacent conspecific populations will remain limited, if lineages are distributed disjunctly in refugial areas. Inbreeding, loss of genetic diversity via drift, loss of heterozygosity, and eventually inbreeding depression are to be expected in such small and isolated populations. On the one hand, facultative self-fertilization would severe the process of loss of heterozygosity in perennial populations. On the other hand, occasional selfing could aid range expansion of sexual lineages and foster gene flow with other adjacent populations or lineages of the species. Such intercrossing of slightly diverged, but still cross-compatible lineages might counteract the negative effects of inbreeding and preserve the genetic diversity of a species. However, effects of inbreeding, rare selfing, or intercrossing on genetic diversity of sexual lineages are poorly explored in the context of geographical parthenogenesis.

We focus here on *R. notabilis* s.l., a sexual species of the Eurasian *Ranunculus auricomus* complex, as a model system for studying the breeding system and the effects of genetic differentiation and genetic diversity (heterozygosity) among individuals/lineages on reproductive fitness. The *R. auricomus* complex shows a typical pattern of geographical parthenogenesis, with five allopatric sexual species in Southern to Central Europe (Karbstein et al., 2020; Tomasello et al., 2020), and a huge diversity of apomictic lineages occupying temperate to arctic Europe, Western Siberia, and Greenland (Hörandl, 2009; Karbstein et al., unpublished data). The sexual species are restricted to very small and/or disjunct distribution areas, and are sometimes confined to single populations (Dunkel et al., 2018; Karbstein et al., 2020). Biogeographical analyses revealed an allopatric speciation process from the mid-Pleistocene onwards, but could not answer the question why these plants of forest understory or meadow habitats failed to expand their range after postglacial reforestation of temperate Europe (Tomasello et al., 2020). Pollination experiments revealed self-incompatibility of sexual taxa, whereby mostly samples of the temperate-montane and Central European species

R. cassubicifolius s.l. were included in this study (Hörandl, 2008). Tetraploid and hexaploid (Hörandl, 2008) apomicts were found to be self-fertile. The seed set of sexual taxa was significantly higher than that of apomictic ones, while germination rates were not different between reproduction modes and ploidy levels (Hörandl, 2008). However, in this earlier study, very few samples of the other main sexual progenitor species in Central Europe, *R. notabilis*, could be included because only a few and very small populations were known from Southeastern Austria. Population genetic studies using isoenzymes suggested inbreeding in some of these small populations (Hörandl et al., 2000). However, population genetic studies applying more efficient DNA markers are missing so far. Recently, several local sexual lineages in Slovenia were detected (Dunkel et al., 2018), but these were also classified within *R. notabilis* s.l. according to comprehensive phylogenomic and geometric morphometric studies (Karbstein et al., 2020; Tomasello et al., 2020). The species *R. notabilis* s.l. is structured into six locally distributed lineages within the Illyrian region (incl. Southeastern Austria; Fig. 1), but lineage ranges are partially overlapping (Dunkel et al., 2018; Tomasello et al., 2020). *Ranunculus notabilis* s.l. represents an established model system appropriate for studying the breeding system, effects of crossability within and between local lineages, and reproductive fitness measures.

By using experimental crossings and analyses of key genetic measures, we aim at clarifying the role of breeding systems of sexual plant species in GP scenarios. We address the following questions: (a) Are lineages within *R. notabilis* s.l. self-incompatible, or does facultative self-fertility occur? (b) Are these lineages fully cross-compatible, as expected within a species? (c) Do interlineage crossings differ in their seed set from intralineage ones (in comparison with other sexual species of the *R. auricomus* complex to exclude general crossing barriers/inbreeding)? (d) Do germination rates differ in inter- and intralineage crossings? (e) How do genetic distance and genome-wide heterozygosity differ among lineages? (f) How do these measures relate to reproductive fitness?

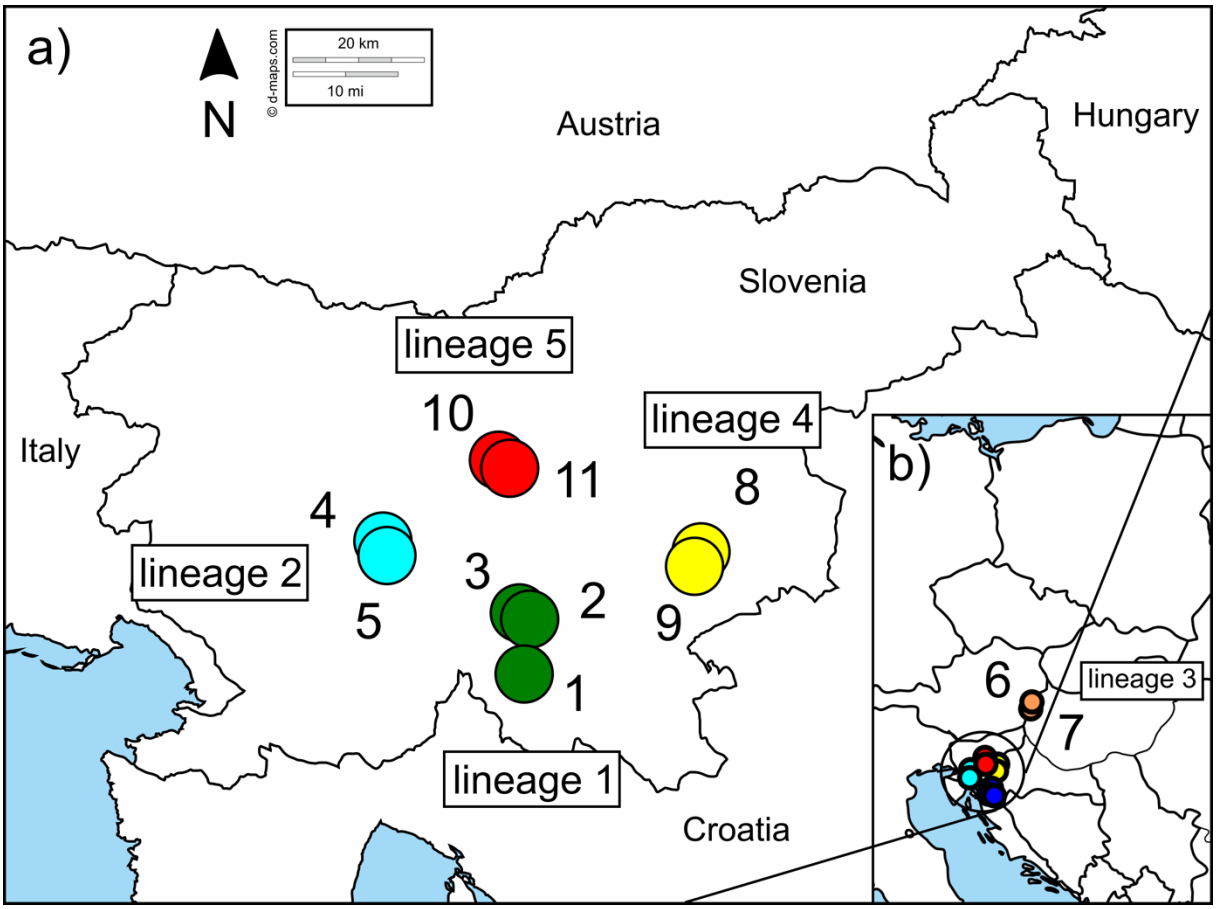


Figure 1. (a) Study locations of sexual *R. notabilis* s.l. populations in Slovenia (see also Karbstein et al., 2020), and (b) all study locations in a European view. Locations are indicated by circles with different numbers corresponding to Map IDs in Table 1. We downloaded the original map from https://d-maps.com/.

Table 1. Locations of sexual diploid lineages of *R. notabilis* s.l. in the Illyrian region (see also Karbstein et al., 2020). Population ID with map ID (see Fig. 1), species, lineage number, number of garden individuals involved in crossings (N_{Cross}), number of RADseq samples (N_{RAD}), locality by country, collection date of population sampling, altitude (in meter above sea level, m.a.s.l.), latitude (N, decimal), longitude (E, decimal), habitat, collector, and herbarium voucher specimens. Former and no longer accepted sexual species names after Dunkel et al. (2018) are described in Materials and methods or can be assessed via population IDs in Dunkel et al. (2018) and Karbstein et al. (2020). Populations Du-30441, Du-30442, EH10137, and Du-33266 were not involved in crossings. Herbarium voucher codes are given in Karbstein et al. (2020).

Pop ID	Map ID	Species	Lineage	N_{CROSS}	N_{RAD}	Locality	Collection date	Altitude	Latitude (N)	Longitude (E)	Habitat	Collector	Herbarium Voucher
Du-30441	(1)	*R. notabilis* s.l.	1	-	1	Slovenia	23.04.2013	260	45.462527	14.817019	shrubbery, forest edge	F.G.Dunkel	LJU, B, M
Du-30442	(2)	*R. notabilis* s.l.	1	-	1	Slovenia	23.04.2013	470	45.684671	14.810336	forest	F.G.Dunkel	M
LH012	(3)	*R. notabilis* s.l.	1	8	2	Slovenia	03.05.2017	469	45.677445	14.82614	forest	L.Hodač, K.Spitzer	GOET
LH014	(4)	*R. notabilis* s.l.	2	8	2	Slovenia	03.05.2017	434	45.84954	14.25946	humid meadow	L.Hodač, K.Spitzer	GOET
LH015	(5)	*R. notabilis* s.l.	2	7	2	Slovenia	03.05.2017	436	45.848988	14.257088	marshy meadow	L.Hodač, K.Spitzer	GOET
EH10137	(6)	*R. notabilis* s.l.	3	-	2	Austria	08.05.2011	220	47.047778	16.4325	forest edge, meadow	E.Hörandl, S.Hörandl, F.Hadacek	GOET
LH028	(7)	*R. notabilis* s.l.	3	4	1	Austria	13.04.2018	227	47.053221	16.435162	forest edge, meadow	L.Hodač, K.Spitzer	GOET
LH010	(8)	*R. notabilis* s.l.	4	12	2	Slovenia	02.05.2017	151	45.89075	15.370933	forest edge	L.Hodač, K.Spitzer	GOET
LH011	(9)	*R. notabilis* s.l.	4	11	2	Slovenia	02.05.2017	148	45.880803	15.336522	forest edge	L.Hodač, K.Spitzer	GOET
Du-33266	(10)	*R. notabilis* s.l.	5	-	1	Slovenia	26.04.2016	325	45.944056	14.653083	forest edge	F.G.Dunkel	M
LH013	(11)	*R. notabilis* s.l.	5	8	2	Slovenia	03.05.2017	324	45.945453	14.651618	humid meadow	L.Hodač, K.Spitzer	GOET

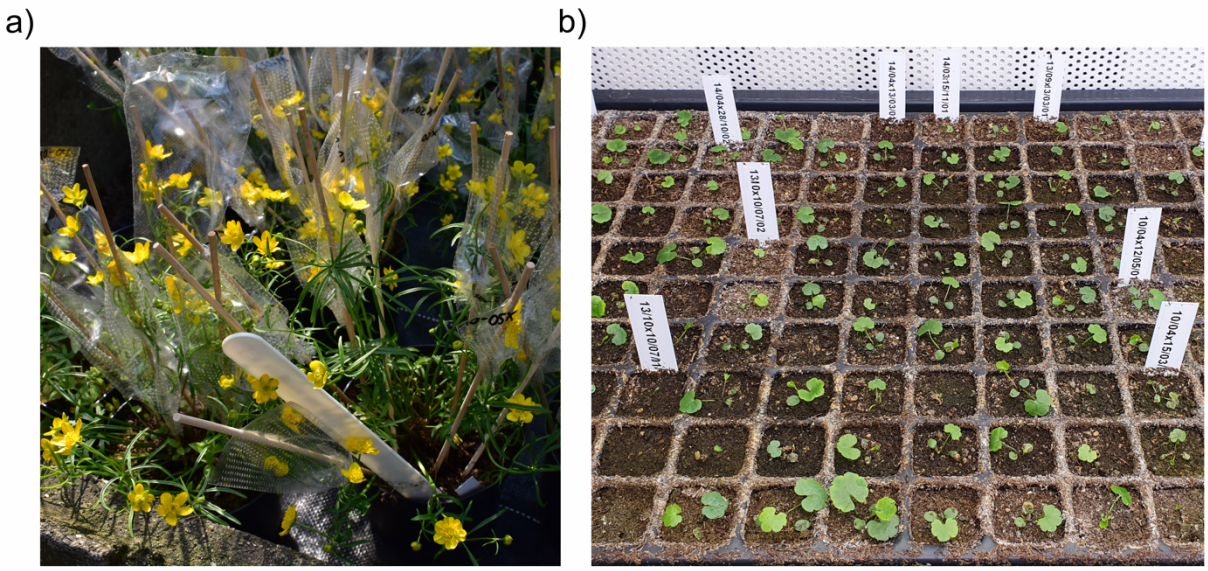


Figure 2. (a) Crossings experiments were conducted in the Old Botanical Garden at the University of Göttingen. *Ranunculus notabilis* s.l. individuals were bagged with porous plastic bags (bags sealed with tape and labeled) and cultivated under controlled environmental conditions. Viable seeds were sowed out and cultivated in (b) climate chambers for germination rate experiments at the University of Göttingen. Labels separate different crossing IDs.

Materials and Methods

Study Locations and Population Sampling

We included five lineages of *R. notabilis* s.l. in the present study, covering the whole range of this Illyrian species. Four of them were formerly described by Dunkel et al. (2018) as distinct species, but were lumped to *R. notabilis* s.l. by Karbstein et al. (2020) due to close phylogenetic relationships and similar morphology in leaf and receptacle traits (lineage 1 "*R. austroslovenicus,*" lineage 2 "*R. mediocompositus*," lineage 3 *R. notabilis* s.str. (Austria), lineage 4 "*R. peracris*," and lineage 5 "*R. subcarniolicus*"). Up to 16 individuals per wild population and one to three populations per lineage were sampled in Slovenia and Austria from 2011 to 2018 (Table 1, Fig. 1). We took living plants and cultivated them in the Old Botanical Garden at the University of Göttingen (controlled environmental conditions for solar radiation and water supply) for crossing experiments. As described in Karbstein et al. (2020), we collected herbarium specimens (deposited in GOET), and recorded altitude, GPS coordinates, and habitat characteristics of each population (Table 1, Fig. 1). All lineages prefer humid and semi-shaded habitats in the submontane forest zone (see Hörandl & Gutermann, 1998; Dunkel et al., 2018). Therefore, we assume that seed formation and germination under garden

conditions, where we also applied semi-shading, were not affected by different original habitat preferences. We used silica gel dried fresh leaves for further genetic laboratory work. Diploid ploidy level and sexual reproduction mode were previously confirmed by Karbstein et al. (2020).

Crossing and Germination Experiments

We conducted manual crossings in 2019 (after at least one year of plant cultivation under the same garden conditions) following two different treatments: Among individuals of a lineage including intra- and interpopulation crossings ("intralineage") and among individuals of different lineages ("interlineage"; see also Hörandl, 2008). Individual crossing pairs are shown in the Excel file "crossings_2019.xlsx" deposited on Dryad data repository. We isolated single flowers (not emasculated) in the bud stage using porous plastic bags (Baumann Saatzuchtbedarf GmbH, Waldenburg, Germany) to avoid undesired cross-pollination. Plastic bags were voluminous enough not to be touched by androecium and gynoecium. This avoids accidental pollination by insects "walking" on the bag surfaces (see Fig. 2a). Isolated flowers were not emasculated in intra- and interlineage crossings because this procedure injures the flowers severely and makes them susceptible to fungal infections, often resulting in flower abortion (pers. experience of the senior author). Moreover, sexual *R. auricomus* species were observed to be absolutely self-incompatible so far (Hörandl, 2008). When individuals were met in the flowering stage and pollen grains were visible, we conducted the above-mentioned cross-pollination treatments. Cross-pollination of flowers was repeated at least once to ensure successful pollination. Due to this crossing procedure, we ensured that intra- or interlineage pollen was the dominant part of crossings. We regard self-fertilization via mentor effects as unlikely (see Discussion). Additionally, we isolated single flowers before anthesis with porous plastic bags to test for self-incompatibility ("selfing"; see also Hörandl, 2008). We included eight, 15, four, 23, and eight individuals for lineage 1, 2, 3, 4, and 5, respectively (see Table 1). In total, we bagged 481 flowers and 59 garden individuals. Afterward, bags were sealed with tape to prevent seed loss during harvesting (see Fig. 2a).

We harvested ripened achenes from May to June 2019. Per bag (one flower per bag), achenes were categorized into viable vs. aborted achenes and counted separately. Viable achenes were recognized by color (brown to dark brown), endosperm development (hard and full achene bodies), size (larger), and removability from the receptaculum (easily falling off; see also Hörandl, 2008). Since *Ranunculus* has single-seeded achenes, the proportion of viable achenes

equals proportions of viable seeds from all seeds (= seed set). In 2020, we added the seed set of three additional individuals of lineage 3 to increase the sample size.

To preserve germinability for germination experiments, we stored seeds in plastic boxes at 7°C (fridge). Cool and dark storage should simulate autumn/winter conditions to overcome seed dormancy (see Lohwasser, 2001 and Vandelook, 2009 for *R. auricomus*). We sowed seeds on the 19th of September 2019 and stored pots under similar garden conditions (garden beds were protected against direct rain and sunlight). To prevent frost damage on young seedlings, we transferred pots into climate growth chambers and raised plants under conditions of 10 hr photoperiod and 12°C air temperature (see Fig. 2b). Pricking out of plants started on the 23rd of January 2020 and was finished on the 27th of March 2020. Within this period, we counted germinated seeds per pot.

Laboratory Work - DNA Extraction and RADseq

We extracted the DNA of 18 *R. notabilis* s.l. individuals (representing three to four samples per lineage) as described in Karbstein et al. (2020). We prolonged sample incubation in lysis buffer to one hour for more DNA yield. The DNA concentration was checked using Qubit fluorometer and Qubit dsDNA BR Assay Kit (ThermoFisher Scientific, Waltham, USA). We targeted a final DNA solution of 30 ng/µl in a volume of 55 µl and visually checked DNA quality by gel electrophoresis. Samples were analyzed by Floragenex Inc. (Portland, USA). The generation of RAD libraries was based on the protocol of Baird et al. (2008), and the enzyme PstI was used for digestion. Sequencing was performed on an Illumina® HiSeq 4,000 platform to produce 100 bp single-end reads at the University of Oregon Genomics Core Facility. The quality of raw reads was checked with FASTQC v.0.11.8 (Andrews, 2010).

Data Analyses

All statistical analyses were performed with R v.4.0.0 (R Core Team, 2020). We calculated seed set per bag, and germination rates as the ratio of germinated seeds to viable seeds. We averaged seed set and germination rate per crossing ID (e.g., LH010-4 × LH028-10). For statistical analyses, our final sample size was N = 284 for seed set and N = 260 for germination rates.

To examine differences between seed set or germination rates and different treatments (selfing, intralineage (intra- and interpopulation), and interlineage) of different *R. notabilis* s.l. lineages,

we performed Kruskal-Wallis tests (for non-normal distribution of proportional data) to check for significant group differences. If results were significant, we calculated pairwise Wilcoxon rank-sum tests with Holm correction to assess pairwise group differences.

To evaluate seed set and germination rates in relation to genetic distance, we performed de novo assembly of RADseq loci and parameter optimization in IPYRAD v.0.9.14 (Eaton, 2014; Eaton & Overcast, 2020) on the GWDG HPC Cluster (Göttingen, Germany). We created a subset of 18 sexual *R. notabilis* s.l. samples out of the total 45 sexual *R. auricomus* ones used by Karbstein et al. (2020; see Table 1). Raw read demultiplexing, removal of adapter sequences and restriction overhang, further quality filtering, and selection of in-sample clustering threshold (ISCT = 95%) and between sample clustering threshold (BSCT = 92%) exactly followed Karbstein et al. (2020; see also Paris et al., 2017; Pätzold et al., 2019). We selected the "min30" instead of the "min10" dataset, that is, a minimum of 30% available samples per locus (5 out of 18 samples), avoiding too high amounts of missing data in the final assembly. Analysis yielded 41,580 filtered loci, ranging from 13,873 to 27,457 loci/individual (mean 19,691 loci/individual by 52% missing sites in the final SNP matrix). In order to estimate the genetic distance among *R. notabilis* s.l. lineages, we computed a network analysis based on the IPYRAD *.u.snps output file (unlinked SNPs, one SNP per locus; converted into a *.nex file) and the Neighbor-Net algorithm using the program SPLITSTREE v.4.14.6 (Huson & Bryant, 2006). We specified a general time-reversible (GTR) model with estimated site frequencies and maximum likelihood (equal rates of site variation, default rate matrix; see also Karbstein et al., 2020). Genetic distances were exported from SPLITSTREE, and (mean) values were used in further data analyses.

To evaluate seed set and germination rates in relation to heterozygosity, we used genome-wide heterozygosity values based on the RADseq dataset of the same 18 samples already calculated by Karbstein et al. (unpublished data; final assembly "min50"). Raw values, mean values per crossing ID, and mean values per lineage (e.g., mean of lineage 1 and 2) were used in further data analyses.

We performed the Kruskal-Wallis test to check for significantly different genetic distances among lineage combinations (e.g., lineage 1 - lineage 1 (intralineage), lineage 1 - lineage 2 (interlineage)). The same tests were also conducted to assess significant differences in heterozygosity among other sexual diploid *R. auricomus* species (heterozygosity values were also taken from Karbstein et al., unpublished data; final assembly "min50"), and among sexual diploid *R. notabilis* s.l. lineages.

To evaluate relationships between seed set or germination rates and mean genetic distances (mean between parental lineage pairs; see Fig. S8 for simple distances) or mean heterozygosity (mean between parental mother lineage and pollen donor lineage) of parental lineages, we performed simple quasibinomial generalized linear regression models (GLMs). We used mean values of parental lineages because no full correspondence between crossed and sequenced individuals was available (see also Excel file "crossings_2019.xlsx" in Dryad data repository). Moreover, genetic cohesion in sexual (meta-)populations (see e.g., Hörandl et al., 2001; Freeland et al., 2011), and the fact that lineages are slightly genetically differentiated (i.e., have some own genetic features, see also ML tree in Karbstein et al., 2020), also justify the use of lineage means in our regressions. Finally, results were plotted, and regression curves were drawn according to GLM results.

To study isolation-by-distance (IBD) among individuals of *R. notabilis* s.l. lineages, we first calculated pairwise geographic distances among locations applying the Vincenty method implemented in the R package GEOSPHERE v.1.5–10 (Hijmans et al., 2019). Second, we performed a Mantel test (Spearman; 9,999 permutations) between genetic (previously used) and geographical distances using the R package APE v.5.3 (Paradis et al., 2019). We plotted data points and the regression line.

Results

We observed significant differences between seed set and different treatments ($Chi^2 = 108.33$, $df = 2$, $p < .001$; Fig. 3a). Selfings revealed significant lower (mean 2.47%, $p < .001$) seed set compared with intra- and interlineage crossings (means 68.97% and 69.49%; Table 2, Fig. 3a). Inter- and intralineage crossings showed no significant difference ($p = .55$). Seed set was significantly lower in intra- compared with interpopulation crossings (means 65.01% and 77.85%; $Chi^2 = 9.48$, $df = 1$, $p < .01$; Table 2, Fig. 3b). In contrast, selfings, intra- and interlineage crossings (means 42.84%, 32.22%, and 37.42%; $Chi^2 = 1.88$, $df = 2$, $p = .39$; Table 2, Fig. S1) as well as intra- and interpopulation crossings (means 32.66% and 31.14%; $Chi^2 = 0.05$, $df = 1$, $p = .82$; Table 2, Fig. S2) unveiled comparable germination rates.

Selfings of *R. notabilis* s.l. lineages did not differ significantly from each other in seed set (means 0.00%–5.32%; $Chi^2 = 5.42$, $df = 4$, $p = .25$; Fig. 4), but in germination rates (means 23.14%–100.00%, $Chi^2 = 8.36$, $df = 3$, $p < .05$; Table 2, Fig. S3). Nevertheless, we did not

detect significant differences in germination rates in pairwise comparisons ($p > .05$), and, in general, investigated differences might be due to the low sample size of lineage 1. Intralineage *R. notabilis* s.l. crossings revealed no significant differences in seed set among lineages (means 50.40%–77.88%; $Chi^2 = 7.75$, $df = 4$, $p = .10$; Table 2, Fig. S4). Germination rates of intralineage crossings showed significant differences among lineages (means 16.55%–72.41%; $Chi^2 = 12.07$, $df = 4$, $p < .05$; Table 2, Fig. S5) but also no significant differences in pairwise comparisons (probably due to large within lineage variation). In contrast to intralineage crossings, we observed significant differences in seed set among interlineage ones (means 57.17%–72.89%; $Chi^2 = 19.94$, $df = 4$, $p < .001$; Table 2, Fig. S6). However, lineage 5 was the only lineage that exhibited significant differences to other lineages (lineages 1, 2, and 4; $p < .01$). Interlineage crossings significantly differed also in relation to germination rates (means 22.03%–57.80%; $Chi^2 = 40.72$, $df = 4$, $p < .001$; Table S2, Fig. S7). Lineages 1 and 2 (22.03 and 29.56%) showed significantly lower germination rates than lineages 4 and 5 (42.77 and 57.80%; $p < .05$). Lineage 3 (46.70%) differed significantly from lineage 1 ($p < .05$), but not from lineages 2, 4, and 5 ($p > .05$).

Lineage comparisons revealed significantly different genetic distances (means 0.2173–0.3312; $Chi^2 = 111.30$, $df = 14$, $p < .001$; Table 3, Fig. S8). Individuals within lineages are mostly characterized by the lowest genetic distances (means 0.2173–0.2683) compared to individuals involved in interlineage crossings (means 0.3100–0.3312). Heterozygosity was significantly higher in *R. notabilis* s.l. compared to *R. cassubicifolius* s.l. (means 0.69 vs. 0.52, $p < .05$) but not compared to all other diploid sexual species of the *R. auricomus* complex (means 0.52–0.69; $Chi^2 = 9.57$, $df = 3$, $p < .05$; Fig. S9a). Nevertheless, *R. notabilis* s.l. lineages did not differ with respect to heterozygosity (means 0.54–0.80; $Chi^2 = 6.13$, $df = 4$, $p > .05$, Fig. S9b).

Moreover, we detected significant relationships between seed set and genetic distance (Table 4, Fig. 5a–c). In general, seed set and genetic distance were significantly positively related ($t = 15.20$, $df = 277/276$, $p < .001$; Fig. 5a). Whereas the intralineage relationship showed a significantly negative slope ($t = -2.58$, $df = 58/57$, $p < .05$; Fig. 5b), the interlineage relationship was also significantly positive ($t = 2.37$, $df = 175/174$, $p < .05$; Fig. 5c). Genetic distance did not statistically affect germination rates in general and intra- and interlineage ($p = .51$, $p = .74$, and $p = .38$, respectively; Table 4). Intra- and interpopulation crossings showed similar relationships (similar slope, though always $p > .05$) between seed set or germination rates and mean genetic distance compared to intralineage crossings (Table 4). In general, and for interlineage crossings, we observed no significant relationships between seed set and mean heterozygosity ($p = .07$ and $p = .33$, respectively; Table 5). In contrast, we found a significant

negative relationship between seed set and mean heterozygosity for intralineage crossings ($t = -2.27$, $df = 58/57$, $p < .05$; Table 5, Fig. 6). Germination rates were significantly positively related to mean heterozygosity in interlineage crossings ($t = 2.29$, $df = 181/180$, $p < .05$; Fig. 7) but not in general and in intralineage crossings ($p = .21$ and $p = .54$, respectively). Intra- and interpopulation crossings showed similar relationships (similar slope, though always $p > .05$) between seed set or germination rates and mean heterozygosity compared to intralineage crossings (Table 5). Moreover, we observed a significant isolation-by-distance among *R. notabilis* s.l. lineages ($R_{xy}^2 = 0.22$, $p_{xy} < 0.001$; Fig. S10).

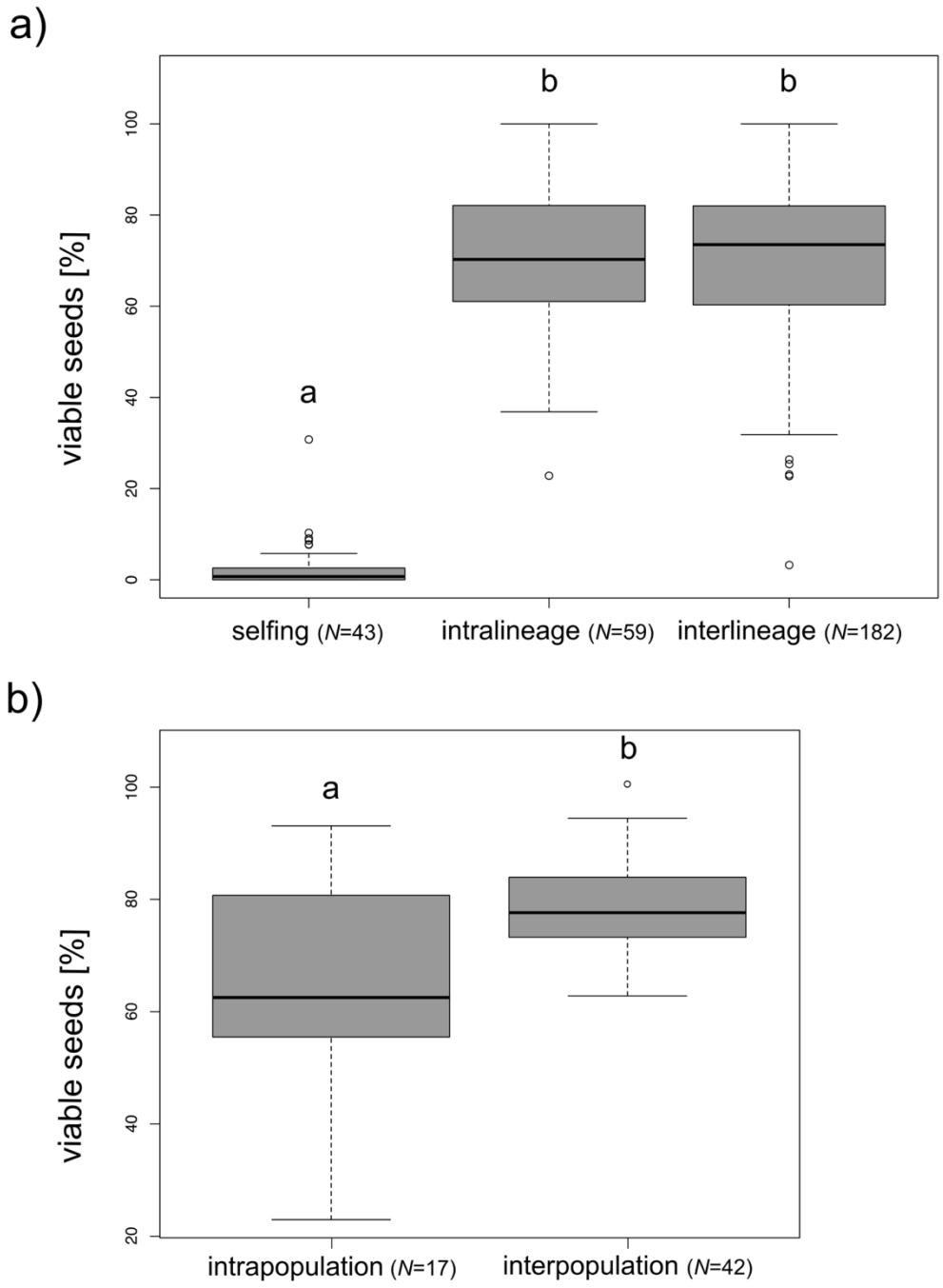


Figure 3. Boxplots showing (a) seed set (percent of viable seeds) of different treatments (selfing, intra-, and interlineage) and (b) seed set of inter- and intrapopulation crossings (intralineage). Letters above boxplots indicate significant/nonsignificant differences between groups. N = sample size (number of crossings/crossing IDs used to determine seed set).

Table 2. Mean percentages of seed set (percent of viable seeds) and germination rates (germinated seeds) for all lineages and for each lineage according to treatments (selfing, intralineage, and interlineage crossing), and for different intralineage crossing treatments (intra- and interpopulation). We marked significant ($p < .05$) Kruskal-Wallis test results in bold. The percentages of viable seeds or germinated seeds per crossing ID (see also Excel file "crossings_2019.xlsx" in Dryad data repository for individuals involved in crossings) were averaged. Superscript letters indicate significant/nonsignificant group differences in post hoc comparisons (e.g., within a dataset, groups with "a" belong together because the post hoc test did not reveal significant differences whereas groups with different letters "a" and "b" are significantly different). See Figs. 3, 4, S1–S7 for within-group variation. The decreased sample size of germination rates is due to crossings with no seed set. N_{CS} = sample size (number of crossings/crossing IDs used to determine seed set), N_S = number of seeds, N_{CG} = sample size (number of different crossings/crossing IDs used to determine germination rates), and NG = number of germinated seeds/seedlings.

dataset	treatment	N_{CS}	N_S	viable seeds [%]	N_{CG}	N_G	germinated seeds [%]
all lineages	selfing	43	**82**	**2.47^{b}**	22	40	42.84^{a}
	intralineage	59	**2528**	**68.97^{a}**	58	871	32.22^{a}
	interlineage	182	**7228**	**69.49^{a}**	182	2747	37.42^{a}
intralineage	intrapopulation	**17**	**738**	**65.01^{b}**	17	231	32.66^{a}
	interpopulation	**42**	**1790**	**78.75^{a}**	41	640	31.14^{a}
lineage 1	selfing	6	3	0.48^{a}	**2**	**3**	**100.00^{a}**
lineage 2	selfing	10	27	5.32^{a}	**6**	**15**	**38.38^{a}**
lineage 3	selfing	2	0	0.00^{a}	**-**	**0**	**-**
lineage 4	selfing	17	22	1.12^{a}	**9**	**5**	**23.15^{a}**
lineage 5	selfing	8	30	3.91^{a}	**5**	**17**	**60.79^{a}**
lineage 1	intralineage	10	250	67.44^{a}	10	44	16.55^{a}
lineage 2	intralineage	16	539	77.88^{a}	16	129	26.07^{a}
lineage 3	intralineage	2	33	50.40^{a}	1	21	72.41^{a}
lineage 4	intralineage	22	1204	66.18^{a}	22	376	33.37^{a}
lineage 5	intralineage	9	502	65.77^{a}	9	301	53.28^{a}
lineage 1	interlineage	35	**1257**	**70.90^{a}**	35	**261**	**22.03^{a}**
lineage 2	interlineage	57	**2040**	**72.89^{a}**	57	**667**	**29.56^{ab}**
lineage 3	interlineage	10	**290**	**67.96^{ab}**	10	**150**	**46.70^{bc}**
lineage 4	interlineage	49	**2611**	**72.98^{a}**	49	**1109**	**42.78^{c}**
lineage 5	interlineage	31	**1030**	**57.17^{b}**	31	**560**	**57.80^{c}**

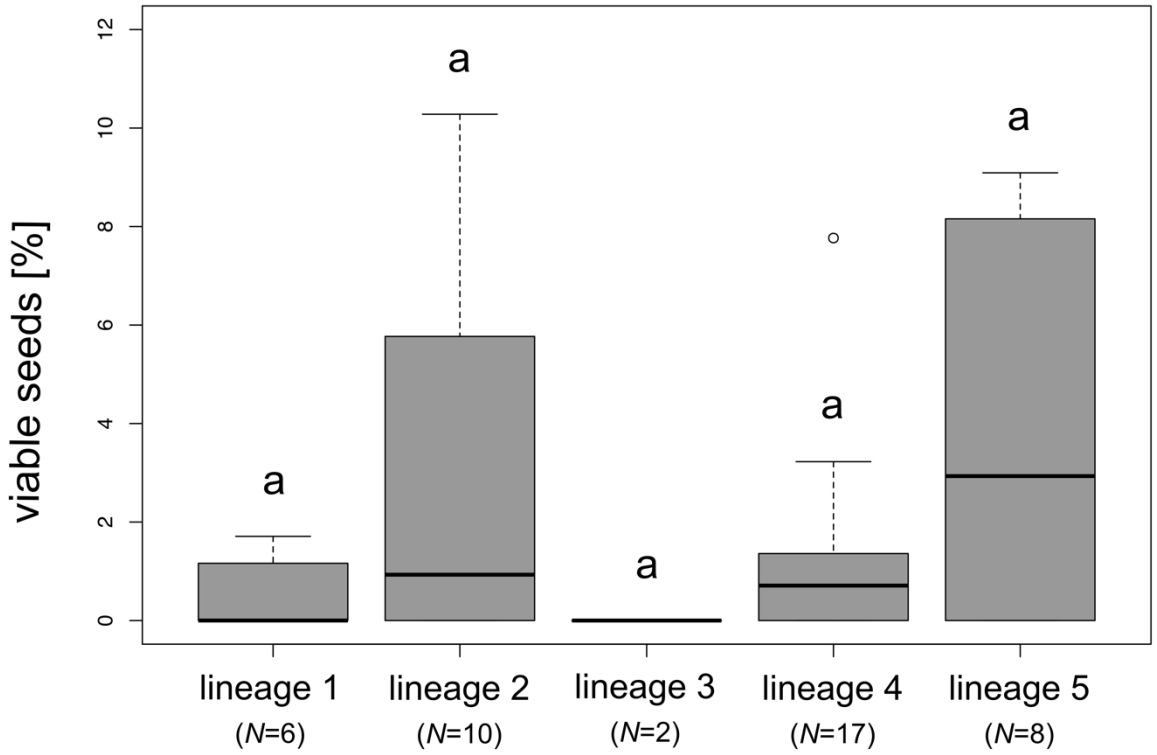


Figure 4. Boxplots showing seed set (percent of viable seeds) of different lineages (selfing treatment). Letters above boxplots indicate nonsignificant differences between groups. N = sample size (number of crossings/crossing IDs used to determine seed set).

Table 3. Mean genetic distances among diploid *R. notabilis* s.l. lineages based on a general time-reversible (GTR) model with estimated site frequencies and maximum likelihood (ML).

lineage	***R. notabilis* s.l. 1**	***R. notabilis* s.l. 2**	***R. notabilis* s.l. 3**	***R. notabilis* s.l. 4**	***R. notabilis* s.l. 5**
***R. notabilis* s.l. 1**	0.2535[a]	0.3110[c]	0.3298[d]	0.3312[d]	0.3293[d]
***R. notabilis* s.l. 2**	-	0.2173[a]	0.3254[cd]	0.3199[cd]	0.3127[c]
***R. notabilis* s.l. 3**	-	-	0.2683[abcd]	0.3053[bc]	0.3209[cd]
***R. notabilis* s.l. 4**	-	-	-	0.2604[a]	0.3100[abc]
***R. notabilis* s.l. 5**	-	-	-	-	0.2370[abcd]

Table 4. GLM results between mean seed set or germination rates for treatments all, intra- (intra- and interpopulation), and interlineage (response variables) and mean genetic distance (explanatory variable). Degrees of freedom (df) are given for null deviance and residual deviance (nd/rd).

		mean genetic distance				
	treatment	**estimate**	**standard error**	**degrees of freedom**	**t value**	**p value**
seed set [%]	all	10.95	0.72	277/276	15.20	**< 0.001**
	intralineage	-13.78	5.33	58/57	-2.58	**< 0.05**
	(intrapopulation)	-10.14	7.69	41/40	-1.32	0.19
	(interpopulation)	-4.72	3.57	19/18	-1.32	0.20
	interlineage	13.97	5.91	175/174	2.37	**< 0.05**
germination rates [%]	all	-0.53	0.81	261/260	-0.66	0.51
	intralineage	2.85	8.68	57/56	0.33	0.74
	(intrapopulation)	3.73	13.01	40/39	0.29	0.78
	(interpopulation)	3.71	7.06	19/18	0.53	0.61
	interlineage	-5.62	6.43	181/180	-0.87	0.38

Discussion

In scenarios of geographical parthenogenesis, the astonishing feature is not so much the wide distribution of asexuals, but the restricted distribution of sexual progenitor taxa. The need for mating partners and pollinators for outcrossing, the danger of inbreeding, and a loss of heterozygosity in small populations are potential obstacles for range expansions of sexual species (Baker, 1967; Haag & Ebert, 2004; Hörandl, 2006; Pinc et al., 2020). Here, we analyzed the breeding system of the diploid sexual species *R. notabilis* s.l., a narrowly distributed species within the Illyrian region, and one of the putative progenitors of the Eurasian *R. auricomus* complex. We tested hypotheses whether crossings of adjacent lineages within the species' range (Karbstein et al., 2020; Tomasello et al., 2020) are self-incompatible and whether intra- and interlineage crossings affect reproductive fitness differentially. We compared reproductive features with levels of genetic distance and genome-wide heterozygosity in natural *R. notabilis* s.l. lineages.

The test for self-incompatibility of *R. notabilis* s.l. via bagging experiments confirmed predominant seed formation via outcrossing in the sexual taxa of *R. auricomus* complex as already observed by Hörandl (2008; Table 2). However, the self-compatibility system in *R.*

notabilis is not perfect, as a low amount of seeds were formed also in bagged plants (Table 2; Figs. 3, 4). This low degree of self-fertility in the *R. auricomus* complex is probably due to a multi-locus gametophytic SI system (GSI) as reported from other congeneric species (Lundqvist, 1990, 1994, 1998), which can result in semi-self-compatibility due to dominance effects of loci (Richards, 1997). Rare occurrences of selfed seed were also observed in bagging experiments of the alpine diploid sexual *R. kuepferi* (Cosendai et al., 2013) and in alpine buttercups in Australia (Pickering, 1997). Spontaneous mutations of S-alleles in the pollen are likely not the reason for self-compatibility, as we observed selfed seeds in most lineages (except lineage 3 *R. notabilis* s.str.). Finally, pseudo-self-compatibility could influence the actual efficiency of SI systems: Extreme temperatures and early or late pollinations could reduce the expression of SI proteins, and hence allow for occasional self-pollen tube growth (de Nettancourt, 2001). Although we are not aware of exposures to extreme temperature or off-time pollinations in our experiments, we cannot rule them out completely, as plants were kept outside in an experimental garden. Additionally, late gametophytic self-compatibility was already reported for Ranunculaceae (Gibbs, 2014). Altogether, frequencies of selfed seeds are below 5% (with outliers up to ca. 15%–31%), which makes it questionable that selfing played a major role in range expansion scenarios. In past glacial times, selfing rates could have been potentially higher due to cold-induced pseudo-self-compatibility, but climate chamber experiments with cold stress treatments would have to verify this assumption. Eventually, occasional selfing helped to disperse lineages within the species range, but long-distance dispersal appears to play no major role for *R. notabilis* s.l. In the *R. auricomus* complex, achenes are mainly dispersed by ants, but potentially also by wind (due to a cavity between achene and seed; Müller-Schneider, 1986). Biogeographical analysis of sexual species of the complex, however, did not support a hypothesis of long-distance dispersal but suggested overall a vicariance scenario (Tomasello et al., 2020). In contrast, polyploid apomicts of the *R. auricomus* complex are fully self-fertile, which appears to be a major advantage for range expansions in GP scenarios (Kirchheimer et al., 2018) and also for long-distance dispersal (Cosendai & Hörandl, 2010). Polyploid apomicts of the *R. auricomus* complex indeed occur in isolated, remote areas (e.g., on Iceland, Svalbard, but also in Southern Europe, see maps in Hörandl, 2009), which have more likely been reached via long-distance dispersal.

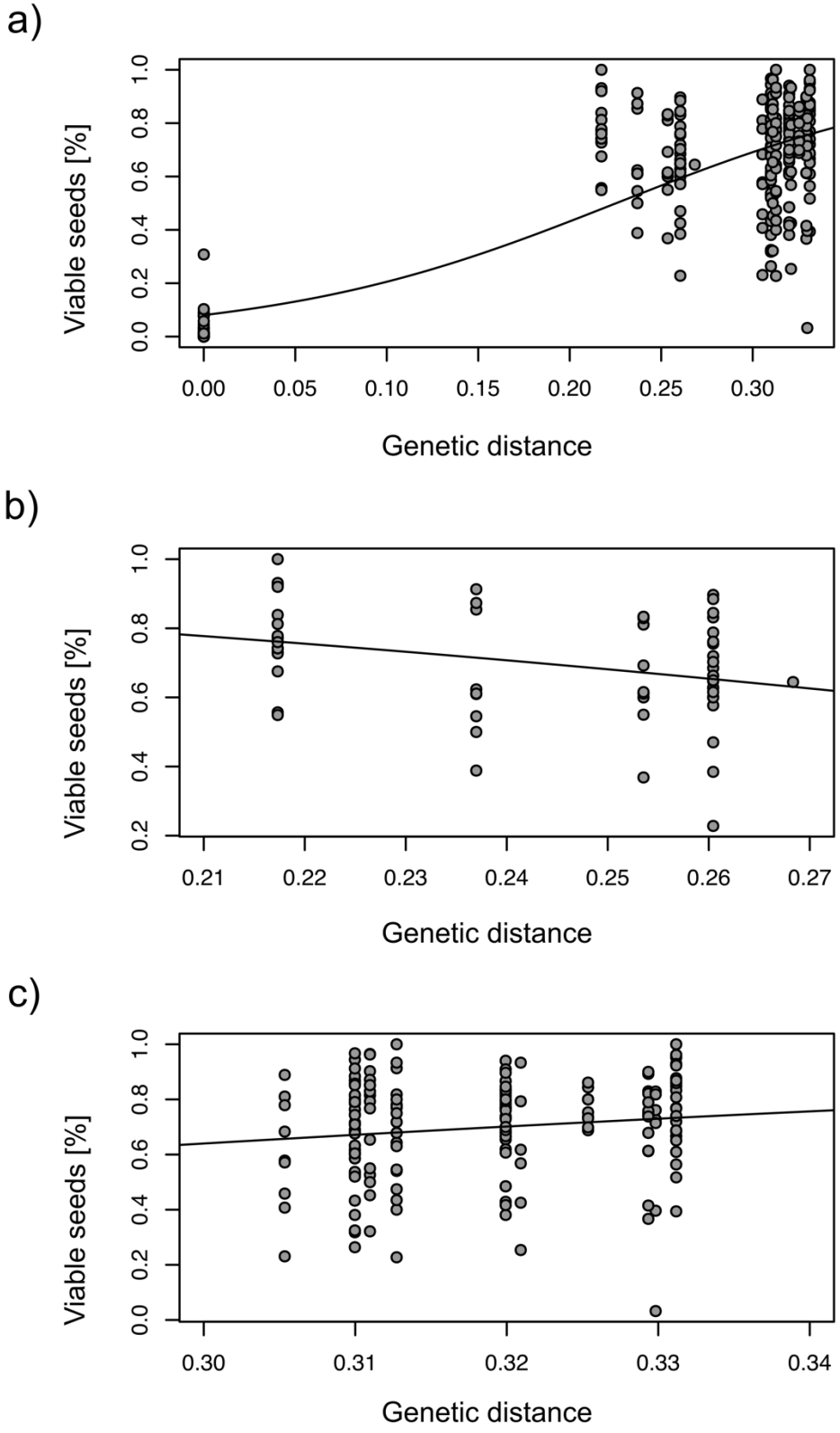


Figure 5. Scatter plots with regression lines based on GLM results of mean seed set (percent of viable seeds) as a function of mean (pairwise) genetic distance. (a) All treatments (selfing, intra- and interlineage crossings), (b) intra- and (c) interlineage crossings (see Table 3 for pairwise lineage comparisons). The genetic distance of selfings is zero. Regression lines are drawn for significant results ($p < .05$).

The inter- and intralineage crossings in *R. notabilis* s.l. revealed no significant differences in seed set (Table 2, Fig. 3a). Hence, full crossability is confirmed within *R. notabilis* s.l., which fits the observation of a shared gene pool, very low genetic distances, and a similar morphology of lineages, justifying the taxonomic treatment as a single species (Karbstein et al., 2020). Apparently, no crossing barriers exist within *R. notabilis* s.l., and adjacent lineages have a potential for gene flow. The variation of seed set (Figs. 3, 4, S4, S6) is similar to results for other sexual species of the *R. auricomus* complex (ca. 60% in Hörandl, 2008; 78% in Hojsgaard et al., 2014). In contrast, interspecific crossings between the genetically more diverged diploid species *R. notabilis* s.str. and *R. cassubicifolius* s.l. revealed a drastically reduced seed set (10%–20% viable seeds), disturbances of meiosis and sporogenesis, and shifts to apomixis (Barke et al., 2018; Hojsgaard et al., 2014; Hörandl, 2008). Cross-pollinations in non-emasculated flowers eventually can induce spontaneous selfing, when foreign pollen originated from a different species, from another cytotype, or was also partly aborted (Mentor effects; de Nettancourt, 2001; Hörandl & Temsch, 2009). However, none of these conditions applied to our crossings, as all lineages are diploid, conspecific, and do have excellent pollen quality (Dunkel et al., 2018; Hörandl et al., 1997). Mentor effects occurred in interploidal crossings of *R. auricomus* cytotypes at very low frequencies (1.3% of all seeds; Hörandl & Temsch, 2009). Therefore, we regard selfed seeds resulting from Mentor effects as unlikely and negligible in our crossing experiments on *R. notabilis* s.l.

Seed set is apparently the main decisive factor for female fitness as germination rates do not differ between selfed, intra-, or interlineage crossings. This observation and the range of variation of germination rates among treatments and lineages (Figs. S1–3, S5, S7) are also in accordance with previous observations in the *R. auricomus* complex (Barke et al., 2018; Hörandl, 2008; Lohwasser, 2001). Hence, we regard the differences in germination rates among *R. notabilis* lineages not necessarily as lineage-specific. Pollen quality is for all lineages of *R. notabilis* s.l. high with more than ca. 90% good pollen (Hörandl et al., 1997; Dunkel et al., 2018), and hence effects of male fitness on seed set are probably negligible. We measured genetic differentiation and genome-wide heterozygosity in natural populations to understand their effects on breeding systems. Genomic markers like RADseq loci provide a genome-wide estimate of genetic differentiation and heterozygosity, and hence can be informative also with a low number of sampled individuals (Lovell et al., 2014; Park & Donoghue, 2019). In general, we found a positive relationship between seed set and genetic distance (Table 4, Fig. 5a): The more genetically distant the individuals involved in crossings, the higher the seed set. However, different mechanisms are probably responsible for the reduction of seed set when genetically

identical or similar pollen is involved: Either the GSI system inhibiting self-pollination, or inbreeding depression resulting from crossings of genetically similar individuals (see e.g., intrapopulation crossings). This suggests a positive effect of outbreeding, probably due to increased genetic diversity and decreased inbreeding effects in the offspring (Gai & Lu, 2013; Wirth et al., 2011). Interestingly, we observed positive relationships between interlineage seed set and mean genetic distance, and interlineage germination rates and mean heterozygosity (Tables 4, 5, Figs. 5c, 7). Sufficient genetic distance and heterozygosity among crossed interlineage individuals might prevent seed abortion due to lethal alleles and might ensure embryo and endosperm development, and therefore allow better seed set and germination rates. Instead of inbreeding depression, the multi-locus SI system can also partly explain the decreased seed set. Gametophytic SI systems rely on a high diversity of S-alleles within a population (Richards, 1997), which might be eroded in small, genetically depauperated populations. We have no direct measures of S-allele diversity in our lineages. However, our genotypes of different lineages were genetically more distant than genotypes within lineages (Table 3, Fig. S8), suggesting sufficient allelic diversity in interlineage crossings. Thus, we consider inbreeding depression as a more reasonable explanation. Inbreeding effects may partly explain the remarkably lower seed set of intra- compared to interpopulation crossings (Fig. 3b). Outbred progeny may benefit from heterosis effects, potentially explaining the higher germination rates generated by crossings of more genetically diverse interlineage individuals. In diploid sexuals of the *Hieracium alpinum* complex also characterized by a GP pattern, the inbred progeny showed lower biomass than the outbred progeny, probably contributing to lower colonizing ability of sexuals compared with apomictic conspecifics (Pinc et al., 2020).

Table 5. GLM results between mean seed set or germination rates for treatments all, intra- (intra- and interpopulation), and interlineage crossings (response variables) and mean heterozygosity (explanatory variable). Degrees of freedom (df) are given for null deviance and residual deviance (nd/rd).

	treatment	**mean heterozygosity**				
		estimate	**standard error**	**degrees of freedom**	**t value**	**p value**
seed set [%]	all	-1.79	0.99	277/276	-1.82	0.07
	intralineage	-2.27	0.96	58/57	-2.27	**< 0.05**
	(intrapopulation)	-1.72	1.39	41/40	-1.24	0.22
	(interpopulation)	-1.34	1.13	19/18	-1.18	0.25
	interlineage	-1.19	1.23	175/174	-0.97	0.33
germination rates [%]	all	1.27	1.02	261/260	1.24	0.21
	intralineage	0.96	1.55	57/56	0.62	0.54
	(intrapopulation)	1.67	2.34	40/39	0.71	0.48
	(interpopulation)	0.58	2.18	19/18	0.27	0.79
	interlineage	3.62	1.58	181/180	2.92	**< 0.05**

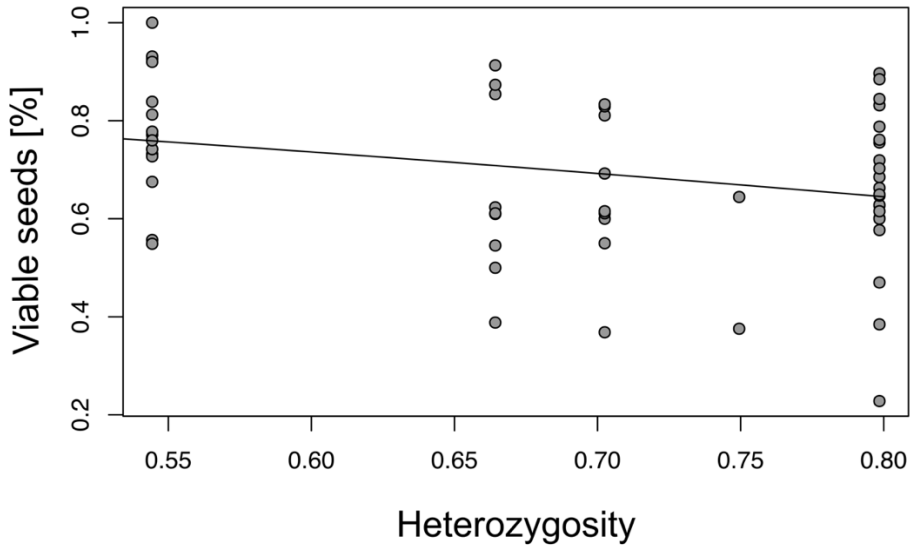


Figure 6. Scatter plot with regression line based on GLM results of mean seed set (percent of viable seeds) as a function of mean heterozygosity of lineages (intralineage treatment).

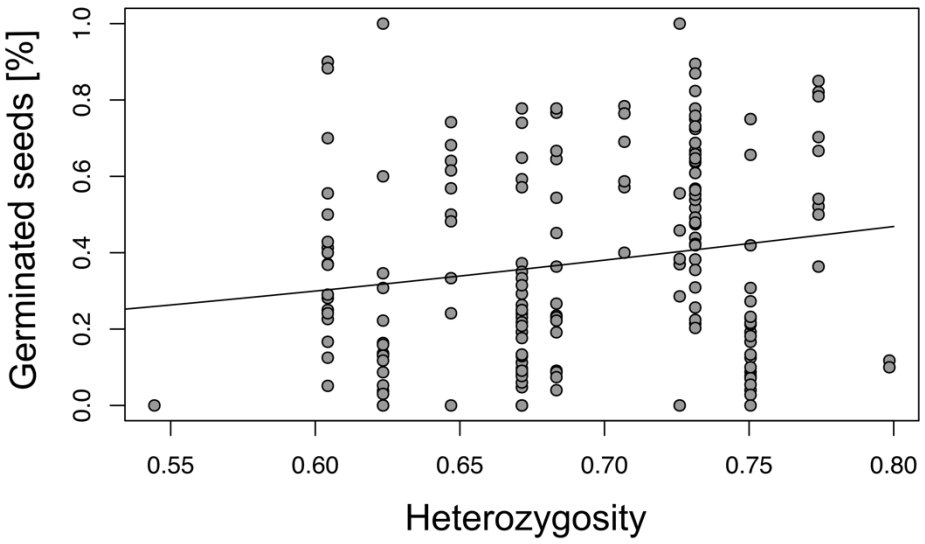


Figure 7. Scatter plot with regression line based on GLM results of mean germination rates (percent of germinated seeds) as a function of mean heterozygosity of lineages (interlineage treatment).

Within our intralineage crossings, in contrast, seed set was decreased, when genetically more distant individuals or individuals with high heterozygosity were crossed (Tables 3, 5, Figs. 5b, 6). We observed similar relationships between intra- and interpopulation crossings compared with intralineage crossings in total, indicating that probably similar processes act on these relationships. An explanation might be outbreeding depression by crossings of different individuals within small populations that was already demonstrated and is probably related to adaptation to environmental variation (microhabitats) or the coadaptation of genes at different loci (e.g., *Anchusa crispa*; Quilichini et al., 2001). The absence of inbreeding depression would have been more reasonable in relation to the other results. However, we observed a reduced seed set in intra- compared with interpopulation crossings, ruling out the explanation of absent inbreeding depression. Different processes may act on intra- and interlineage crossings. However, we could not detect a general positive effect of increased heterozygosity on seed set and germination rates probably due to sufficient genetic variation within and only minor genetic differences among parental *R. notabilis* s.l. individuals.

In a microevolutionary context, we suppose that interlineage crossings have repeatedly reshuffled the gene pool of the species, and avoided strong inbreeding in small populations. Gene flow could have happened during the evolutionary history of the species in the last cold

periods (see Tomasello et al., 2020). In phylogenetic reconstructions of the *R. auricomus* complex, incongruence and signals for introgression were found among lineages of *R. notabilis* s.l., and these lineages showed also a shared gene pool in STRUCTURE analyses (Karbstein et al., 2020). Gene flow might be expected in extant sympatric populations in Slovenia (Dunkel et al., 2018). However, the partial self-compatibility of *R. notabilis* potentially explains beside isolation-by-distance (Fig. S10) the slight genetic differentiation among lineages despite their partial sympatric occurrence. Prezygotic barriers, for example, differences in flowering time (nonoverlapping flowering periods), were not apparent among diploid sexual *R. notabilis* s.l. lineages from the collection dates of Dunkel et al. (2018) and Hörandl and Gutermann (1998), and lineages flowered synchronously at the beginning of our crossing experiments. Individuals of the *R. auricomus* complex form cup-shaped generalist flowers (Steinbach & Gottsberger, 1995), which makes pollinator specificity also unlikely. A population genetic study with a more dense sampling over the whole area, however, is needed to settle the actual amount of gene flow within and between lineages. In other cases of GP, geographically more isolated sexual species showed a much more pronounced geographical population structure and genetic differentiation, for example, in *R. kuepferi* in the Alps and *Crataegus* in North America (Cosendai et al., 2013; Lo et al., 2009). Genome-wide heterozygosity levels as measured here with RADseq data within *R. notabilis* s.l. are comparable to other species of the complex and even higher than in the widespread species *R. cassubicifolius* s.l. (Fig. S9a), another progenitor of the *R. auricomus* complex. The latter species has a strongly disjunct distribution around the Northern and Southern Alps and in the Carpathians, with a much stronger geographical isolation between lineages (see Tomasello et al., 2020), which may reduce heterozygosity and genetic diversity due to inbreeding.

We observed a decreased female fitness in intrapopulation crossings and in crossings between genetically similar or less heterozygous individuals of different lineages (Tables 2, 4, 5, Figs. 3b, 5c, 7). Self-incompatibility and the tendency of inbreeding depression in small and/or isolated populations are probably partly responsible for the limited geographic ranges of sexual lineages of *R. notabilis* s.l. However, here, the outcrossing system appears to be fully functional, and besides the reduced seed set in intrapopulation crossings, no signs of strong inbreeding are apparent from our data. Hence, it is plausible that the species can maintain and might even expand its distributional range and would not be at risk of extinction within the Illyrian region. Population genetic studies on the most isolated lineage of *R. notabilis* s.str. in Austria showed a population structure typical for outcrossers, with a tendency to inbreeding in one population (Hörandl et al., 2000). Natural forest habitats in this area are partly destroyed by spruce

forestation, which results in acidification of soils and devastation of the understory flora (pers. obs. of authors). However, *R. notabilis* s.l. also occurs on humid nutrient-poor anthropogenic meadows that are, despite recent nature conservation efforts, still under pressure due to habitat eutrophication, draining, and defragmentation. These recent factors probably limit not only range expansion, but contribute to fragmentation of the distribution area. In deeper time levels, that is, in the postglacial reforestation scenarios of Central Europe, apomictic polyploid congeners might have colonized faster newly available habitats, as shown by Kirchheimer et al. (2018) for *R. kuepferi*. However, also the sympatric emergence of apomicts can play a major role in GP scenarios (Nardi et al., 2020). Biotic interactions of diploids of the *R. auricomus* complex with polyploid apomictic conspecifics or crossings with sympatric tetraploid apomicts as pollen donors may result in introgression and loss of fertility in triploid offspring (see Hörandl et al., 2000), potentially also reducing fitness of *R. notabilis* s.l. populations. Hence, in mixed populations, diploid outcrossing sexuals would suffer from a minority cytotype disadvantage (Levin, 1975) while polyploid self-fertile apomicts would not be fertilized and hence not affected.

Conclusion

We confirm predominant self-incompatibility of sexual *R. notabilis* s.l., as a general disadvantage for rapid colonization and range expansions, as predicted by Baker's law. In GP scenarios, small and disjunct distribution areas of sexual species in relic areas confer the potential disadvantage of inbreeding of genetically similar individuals. We confirmed in our model system a reduced female fitness in intrapopulation (a part of intralineage) crossings. We also found higher female fitness (seed set) with increasing genetic distance in all and interlineage crossings, and higher germination rates with increasing heterozygosity in interlineage crossings. These results support an interpretation that inbreeding depression particularly affects intrapopulation (intralineage) crossings, and that positive outbreeding effects particularly influence interlineage crossings. Breeding systems and genetic properties of sexual species are important, although not the exclusive factor for the restricted distribution ranges of sexual species and have to be considered for understanding GP patterns.

Acknowledgments

We thank the German Research Foundation for project funding (DFG, Ho4395/10-1) to Elvira Hörandl, within the priority program "Taxon-Omics: New Approaches for Discovering and Naming Biodiversity" (SPP 1991). We acknowledge S. Friedrichs for gardening and taking care of germination experiments in climate chambers, and F. G. Dunkel for providing garden plants and silica gel dried leaf material. We thank the anonymous referees for valuable comments on the manuscript. Open access funding enabled and organized by Projekt DEAL.

Conflict of Interest

None declared.

Author Contributions

Kevin Karbstein: Conceptualization (supporting); Formal analysis (lead); Investigation (equal); Methodology (lead); Visualization (lead); Writing-original draft (lead); Writing-review & editing (lead). Elisabeth Rahmsdorf: Conceptualization (supporting); Formal analysis (supporting); Investigation (lead); Methodology (supporting); Visualization (supporting); Writing-original draft (equal); Writing-review & editing (supporting). Salvatore Tomasello: Conceptualization (supporting); Investigation (supporting); Supervision (supporting); Writing-original draft (supporting); Writing-review & editing (supporting). Ladislav Hodač: Investigation (supporting); Resources (supporting); Writingoriginal draft (supporting); Writing-review & editing (supporting). Elvira Hörandl: Conceptualization (lead); Formal analysis (supporting); Funding acquisition (lead); Methodology (equal); Project administration (lead); Resources (supporting); Supervision (lead); Visualization (supporting); Writing-original draft (lead); Writing-review & editing (lead).

Data Availability Statement

Code availability. The R script used in data analyses are deposited on Dryad (https://doi.org/10.5061/dryad.66t1g1k10). Basic data supporting the findings of this study are available within the manuscript and the Appendix. Counts and statistics of crossing and germination experiments are deposited on Dryad (https://doi.org/10.5061/dryad.66t1g1k10). Demultiplexed RADseq reads are stored on National Center for Biotechnology Information Sequence Read Archive (SRA): BioProject ID PRJNA627796.

Orcid

Kevin Karbstein https://orcid.org/0000-0003-1424-6557
Salvatore Tomasello https://orcid.org/0000-0001-5201-1156
Ladislav Hodač https://orcid.org/0000-0002-6885-1317
Elvira Hörandl https://orcid.org/0000-0002-7600-1128

Chapter 3 – The Biogeography of Diploid Sexuals

Phylogenomics unravels Quaternary vicariance and allopatric speciation patterns in temperate-montane plant species: A case study on the *Ranunculus auricomus* species complex

Tomasello, S., Karbstein, K., Hodač, L., Pätzold, C., & Hörandl, E.

Molecular Ecology, *29*, 2031–2049 (2020).
doi: 10.1111/mec.15458 (CC BY 4.0, changes in format and citation style)
BioRxiv, 1–41 (2020).
doi: 10.1101/2020.01.06.895904 (reuse allowed by author)

The time frame and geographical patterns of diversification processes in European temperate-montane herbs are still not well understood. We used the sexual species of the *Ranunculus auricomus* complex as a model system to understand how vicariance vs. dispersal processes in the context of Pleistocene climatic fluctuations have triggered speciation in temperate-montane plant species. We used target enrichment sequence data from about 600 nuclear genes and coalescent-based species tree inference methods to resolve phylogenetic relationships among the sexual taxa of the complex. We estimated absolute divergence times and, using ancestral range reconstruction, we tested if speciation was enhanced by vicariance or by dispersal processes. Phylogenetic relationships among taxa were fully resolved with some incongruence in the position of the tetraploid *R. marsicus*. Speciation events took place in a very short time at the end of the Mid-Pleistocene Transition (830–580 thousand years ago [ka]). A second wave of intraspecific geographical differentiation occurred at the end of the Riss glaciation or during the Eemian interglacial between 200 and 100 ka. Ancestral range reconstruction suggests a widespread European ancestor of the *R. auricomus* complex. Vicariance has triggered allopatric speciation in temperate-montane plant species during the climatic

deterioration that occurred during the last phase of the Mid-Pleistocene Transition. Vegetation restructuring from forest into tundra could have confined these forest species into isolated glacial macro- and microrefugia. During subsequent warming periods, range expansions of these species could have been hampered by apomictic derivatives and by other congeneric competitors in the same habitat.

Keywords

ancestral range, geodispersal, Mid-Pleistocene Transition, *Ranunculus auricomus* complex, target enrichment, vicariance

Introduction

Pleistocene climatic fluctuations have caused periodic range shifts (e.g., north-south) and/or range expansions and contractions in Northern Hemisphere plant species. The alternation of cold and warm periods has characterized the entire Quaternary (the last 1.8 million years). In Europe, cold periods have produced southward shifts in the distribution ranges of temperate species (in the so-called "Mediterranean refugia"; Brewer et al., 2002; Taberlet et al., 1998) and/or the survival of these in few, isolated Central European refugial areas (e.g., Magri et al., 2006; Naydenov et al., 2007). On the other hand, warmer interglacials have favoured range expansion towards the north and the formation of contact zones among diverged populations of the same species (Hewitt, 1999; Magri, 2008). As a result, the survival of formerly coherent population groups in isolated refugia promoted allopatric speciation. Range expansions and the consequent formation of secondary contact zones resulted in hybridization, and eventually in the formation of new taxa through homoploid hybrid speciation or allopolyploidization (Abbott et al., 2013; Stebbins, 1984).

In the last three decades, many studies have elucidated upon the distribution and diversification in European high-alpine plant species (see Schönswetter et al., 2005; Vargas, 2003). Although many studies have highlighted the importance of vicariance and postglacial recolonization for explaining phylogeographical patterns in alpine plant species (Comes & Kadereit, 1998; Kropf et al., 2003; Taberlet et al., 1998; Widmer & Lexer, 2001), sometimes the situation is rather

complicated and not easily explained by either vicariance or dispersal (Reisch, 2008; Sanz et al., 2014; Schönswetter & Tribsch, 2005; Tomasello & Oberprieler, 2017; Tribsch et al., 2002). Cold periods have not exclusively influenced the actual species distribution and intraspecific genetic patterns of alpine species, they have also actively promoted divergence and speciation, as documented in studies on animals (Knowles, 2001) and plants (Kadereit et al., 2004). In the latter study, the authors observed that speciation processes in *Primula* sect. *Auricula* occurred principally during glacial periods in geographically isolated refugia.

Much less information is available on montane and/or subalpine plant species of the Northern Hemisphere. Research of forest plants has focused on the dominating tree species, such as beech, oaks and pine (e.g., Brewer et al., 2002; Magri, 2008; Naydenov et al., 2007). However, herbaceous plants inhabiting the understorey of cool and deciduous mountain forests of the temperate biome have rarely been investigated (Alsos, 2005; Després et al., 2002). In cold periods of the Pleistocene, large parts of Central Europe were dominated by steppe vegetation, and deciduous forests were mostly restricted to southern and main refugia in the Balkan and Iberian peninsulas (Stuchlik & Wójcik, 2001). However, more distant local microrefugia outside the main southern glacial refugial areas should also be considered (Rull, 2009). For the boreo-montane species *Polygonatum verticillatum*, putative glacial refugia might have existed in the foothills of heavily glaciated mountain ranges, such as the Alps and the Carpathians (Kramp et al., 2008). In warmer periods, the ecoclimatic barriers for forests disappeared, leaving only the highest chains of the Alps and Carpathians as a geographical barrier. Hence, reforestation of the lower elevational zones around these mountain areas and in the lowlands has provided ample opportunities for herbaceous understorey plants to expand their range via geodispersal. In contrast to "jump" dispersal over a persisting barrier, geodispersal describes a more or less continuous range expansion after disappearance of a barrier, and would result in a single, widespread species (Ronquist & Sanmartín, 2011; Sanmartín, 2012). Geodispersal was found to be important for boreotropical forest plants (Couvreur et al., 2011), but little is known for temperate forest plants. Zhao et al. (2013), for example, found vicariance and allopatric speciation in a temperate deciduous forest plant in China.

The cosmopolitan genus *Ranunculus* L. with ~ 600 species originated and started diversifying in the early Miocene (Emadzade & Hörandl, 2011). Emadzade et al. (2011) have elucidated how both vicariance and long-distance dispersal have to be invoked to explain the present distribution patterns of lowland and montane species in the genus. Intercontinental and transoceanic dispersal of forest species occurred mostly in the *Ranunculus polyanthemos* clade (Emadzade et al., 2011). Other clades comprising forest species, such as *R.* sect. *Auricomus*,

have not yet been fully resolved. The Eurasian *Ranunculus auricomus* complex is nested as a monophyletic group within this large clade (*R.* sect. *Auricomus*) that comprises otherwise Asian, Arctic and North American species (Emadzade et al., 2015). Whereas apomictic polyploids of the *R. auricomus* complex are widespread and abundant in various temperate and boreal biomes, by contrast sexuals are montane and subalpine species of Mediterranean and temperate European mountains, inhabiting deciduous forest ecosystems or natural and anthropogenic meadows. Although these forests are widespread over temperate Europe, these taxa have restricted or disjunct distribution ranges, and usually occur at the foothills of the main European mountain ranges (Fig. 1). Both vicariance or jump dispersal could have shaped this pattern. It is also unclear why speciation occurred despite the lowland areas being connected during interglacial periods and providing habitats for geodispersal.

The whole *R. auricomus* complex shows a pattern of geographical parthenogenesis: Sexual progenitors have much smaller and more southern distribution areas than the numerous polyploid apomictic derivatives (Bierzychudek, 1985; Hörandl, 2006). Apomictic lineages can expand their range rapidly due to uniparental reproduction, both via long-distance dispersal and geodispersal (Baker, 1967; Cosendai et al., 2013; Hörandl, 2006; Kirchheimer et al., 2018). However, why sexual progenitors of such complexes do not manage to expand their distribution areas and fail to fill their potential range remain poorly understood (see e.g., simulations in Kirchheimer et al., 2018). Hence, it is important to disentangle the biogeographical histories and speciation processes of sexual species from their apomictic derivatives to understand the phenomenon of geographical parthenogenesis.

All taxa included herein are diploid or tetraploid sexuals (Dunkel et al., 2018; Hojsgaard et al., 2014; Masci et al., 1994). Hörandl (2004) estimated the divergence time between *R. notabilis* and *R. cassubicifolius* s.l. (including *R. carpaticola*) using isozyme data and Nei's genetic distances (Nei, 1975) as 914,000 years ago. Apomictic hybrid taxa are hypothesized to be even younger (<100,000 years; Pellino et al., 2013). Up to now, traditional markers have failed to resolve the phylogenetic relationships among taxa of the complex (Hörandl et al., 2009). Karbstein et al. (2020) applied phylogenomic methods (RADseq and target enrichment), together with geometric morphometrics to investigate species boundaries among the sexual taxa of the complex. The authors concluded that five species are recognizable within the complex (*R. cassubicifolius* s.l., *R. envalirensis* s.l., *R. flabellifolius*, *R. marsicus* and *R. notabilis* s.l.; Table 1).

In the present study, we address three major questions concerning the spatiotemporal diversification of temperate-montane species in Europe, using as study object the sexual species

of the *R. auricomus* complex. (a) Did colder periods during glaciation cycles enhance speciation processes in temperate-montane species? (b) Is vicariance, dispersal or geodispersal the most probable scenario to explain the present distribution of sexual taxa of *R. auricomus*? (c) Did the species survive glacial periods in southern macrorefugia or also in microrefugia, and (if the latter scenario is the most plausible) why are they still restricted to these areas? Therefore, we: (a) Applied target enrichment of nuclear genes to resolve phylogenetic relationships among taxa of the complex; (b) use coalescent-based methods to estimate the species divergence times in the context of past climatic oscillations; (c) apply ancestral range reconstruction methods to test vicariance vs. dispersal speciation scenarios and identify potential macro- and microrefugia; and (d) discuss the relevance of our findings for understanding the geographical parthenogenesis pattern.

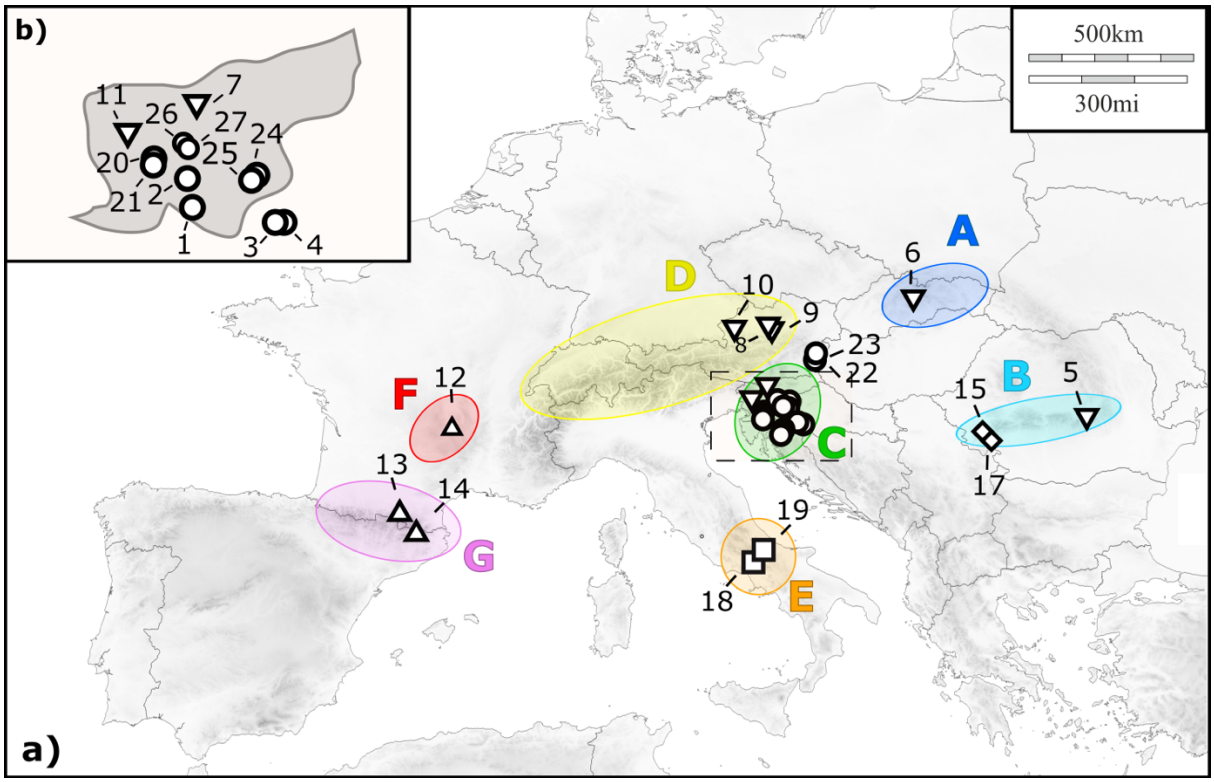


Figure 1. Distribution map of the accessions included in the present study (a) and detailed map of specimens collected in Slovenia (b). Numbers correspond to the map IDs in Table 1. Different symbols are for different species, following the treatment given in Karbstein et al. (2020): Circles are for *Ranunculus notabilis* s.l., diamonds for *R. flabellifolius*, squares for *R. marsicus*, triangles for *R. envalirensis* s.l. and overturned triangles for *R. cassubicifolius* s.l. Coloured ellipses refer to the areas defined for the ancestral range reconstruction analysis: A, Northern Carpathians; B, Southern Carpathians; C, Illyrian region; D, Northern Prealps; E, Central Apennines; F, Massif Central; and G, Pyrenees [Colour figure can be viewed at wileyonlinelibrary.com].

Materials and Methods

Plant Material

We included samples from all the described sexual species of the *Ranunculus auricomus* complex. Plant material used for DNA extraction consisted of silica-gel dried leaves or herbarium specimens (Table 1). We follow the taxonomy of Karbstein et al. (2020) (see synonymy in Table 1). Two additional *Ranunculus* species not belonging to the *R. auricomus* complex were included in the analyses as an outgroup: *R. pygmaeus* from the clade of *Ranunculus* sect. *Auricomus*, and *R. sceleratus* from the next sister clade (Emadzade et al., 2011).

Gene Selection and Probe Design

To find single-copy genes and select target regions for phylogenetic analyses, we used transcriptomes from two diploid species (*R. notabilis* and *R. cassubicifolius*, as synonym "*R. carpaticola*" in Pellino et al., 2013), a tetraploid sexual accession of *R. cassubicifolius* (Pellino et al., 2013), and *R. brotherusii* (Chen et al., 2015), an Asian species of *R.* sect. *Auricomus*. MARKERMINER v.1.0 (Chamala e t al., 2015) was used to identify putative single-copy orthologous loci for probe design with *Arabidopsis thaliana* as reference. We selected exons found in at least two of the four transcriptomes, with lengths ranging from 120 to 960 bp and a minimum variability of two single nucleotide polymorphisms (SNPs)/120 bp, using a custom python script (available at: https://github.com/ClaudiaPaetzold/MarkerMinerFilter.git). We identified 2,628 exonic regions belonging to 736 target genes (Table S1). The selected target regions (genes) ranged from 121 to 5,291 bp and included at least one exonic fragment each. Arbor Biosciences produced a MYbaits Target Enrichment kit with 20,000, 120-bp-long, in-solution, biotinylated baits based on target sequence information. The final bait panel consisted of 17,988 probes, 14,632 of which are unique, and tiling at 2 × density.

Table 1. Information for the 28 accessions belonging to sexual taxa of the *Ranunculus auricomus* complex, and the two outgroup samples. Accepted names following Karbstein et al. (2020) are in bold type, and younger synonyms are in inverted commas. Countries are abbreviated according to ISO code 3166-2. Altitude is given in metre above sea level (m a.s.l.). Range codes are those used in biogeobears for the ancestral range reconstruction and correspond to those in Fig. 1. Map IDs refer to the sample IDs given in Fig. 1. [a]After Masci et al. (1994), Hojsgaard et al. (2014), Dunkel et al. (2018) and Karbstein et al. (2020). [b]Samples from type material or collected from "locus classicus" and nearby.

ID	Map ID	Taxon	Locality	Collection date	Altitude	Latitude	Longitude	Range code	Herb. Voucher	Ploidy[a]
Du003-1[b]	1	*"R. austroslovenicus"* **_R. notabilis_ s.l.**	Slovenia (SI)	23.04.2013	260	45.495278	14.803889	C	Du-30441	2n
LH012-8	2	*"R. austroslovenicus"* **_R. notabilis_ s.l.**	Slovenia (SI)	03.05.2017	469	45.677445	14.82614	C		2n
2[b]	3	*"R. calapius"* **_R. notabilis_ s.l.**	Croatia (HR)	10.04.2017	141	45.518944	15.570917	C	Du-34889	2n
Du-35351-15	4	*"R. calapius"* **_R. notabilis_ s.l.**	Croatia (HR)	23.04.2018	117	45.51806	15.6125	C	Du-35351-15	2n
9126-2	5	*"R. carpaticola"* **_R. cassubicifolius_ s.l.**	Romania (RO)	07.07.1998	1360	45.432833	25.530708	AB		2n
Du013-1	6	*"R. carpaticola"* **_R. cassubicifolius_ s.l.**	Hungary (HU)	19.04.2008	230	48.114122	20.653033	AB	Du-21047	2n
LH040-4	7	*"R. carpaticola"* **_R. cassubicifolius_**	Slovakia (SK)	03.05.2018	387	48.67731	20.106684	AB		2n
LH008-7	8	**_R. cassubicifolius_**	Austria (AT)	01.05.2017	473	47.938828	14.94475	C		4n
LH009-3	9	**_R. cassubicifolius_**	Austria (AT)	01.05.2017	548	47.924792	14.972606	C		4n
LH006-17	10	**_R. cassubicifolius_**	Germany (DE)	30.04.2017	539	47.811203	12.506822	CD		2n
LH016-14	11	**_R. cassubicifolius_**	Slovenia (SI)	04.05.2017	514	46.010798	14.102718	CD		2n
Du-33354-2[b]	12	*"R. cebennensis"* **_R. envalirensis_ s.l.**	France (FR)	06.05.2016	910	45.5605566	2.656944	F	Du-33354-2	2n
Du018-1[b]	13	**_R. envalirensis_**	Andorra (AD)	23.05.2013	1750	42.591333	1.671583	G	Du-29983	2n
Du019[b]	14	**_R. envalirensis_**	France (FR)	23.05.2013	1595	42.473278	2.080333	G	Du-29988	2n
Du021-1[b]	15	**_R. flabellifolius_**	Romania (RO)	21.04.2010	885	45.049444	21.771306	B	Du-25795	2n
Du054-1[b]	16	**_R. flabellifolius_**	Romania (RO)	21.04.2010	885	45.049444	21.771306	B	Du-25796	2n

LH025-1[b]	17	***R. flabellifolius***	Romania (RO)	10.04.2018	811	45.032008	21.78511	B		2n
LH017-1[b]	18	***R. marsicus***	Italy (IT)	26.06.2017	1535	41.845403	13.929692	E		4n
LH018-2[b]	19	***R. marsicus***	Italy (IT)	27.06.2017	1244	41.868211	14.021872	E		4n
LH014-3[b]	20	*"R. mediocompositus"* ***R. notabilis* s.l.**	Slovenia (SI)	03.05.2017	434	45.84954	14.25946	C		2n
LH015-10	21	*"R. mediocompositus"* ***R. notabilis* s.l.**	Slovenia (SI)	03.05.2017	436	45.848988	14.257088	C		2n
10137-3[b]	22	***R. notabilis***	Austria (AT)	08.05.2011	220	47.047778	16.4325	C		2n
Hoe5615	23	***R. notabilis***	Austria (AT)	05.05.1994	210	47.052739	16.38406	C	Hoe5615	2n
LH010-7[b]	24	*"R. peracris"* ***R. notabilis* s.l.**	Slovenia (SI)	02.05.2017	151	45.89075	15.370933	C		2n
LH011-1	25	*"R. peracris"* ***R. notabilis* s.l.**	Slovenia (SI)	02.05.2017	148	45.880803	15.336522	C		2n
LH011-14	25	*"R. peracris"* ***R. notabilis* s.l.**	Slovenia (SI)	02.05.2017	148	45.880803	15.336522	C		2n
Du049-1[b]	26	*"R. subcarniolicus"* ***R. notabilis* s.l.**	Slovenia (SI)	26.04.2016	325	45.944056	14.653083	C	Du-33266	2n
LH013-3[b]	27	*"R. subcarniolicus"* ***R. notabilis* s.l.**	Slovenia (SI)	03.05.2017	324	45.945453	14.651618	C		2n
Outgroups										
LG09		*R. pygmaeus*	Alaska (US)	07.11.2009	2	64.4561	-165.0961		LG09-A-47-07	2n
10426-3		*R. sceleratus*	Germany (DE)	15.04.2018	165	51.5789	10.1458			2n

Library Preparation

Sequencing libraries were prepared using either the "NEBNext Ultra II DNA Library Prep Kit for Illumina" (E7645) or the "NEBNext Ultra II FS DNA Library Prep Kit for Illumina" (E7805) (New England BioLabs). Sequencing was conducted on an Illumina MiSeq System (Illumina Inc.) at the Transcriptome and Genome Analysis Laboratory (University of Goettingen). Pools were mixed equimolarly and sequenced in two different paired-end runs (six pools, 24 samples each) with a 2 × 250-bp (500 cycles) version 2 kit (see Appendix S1 for details on library preparation).

Read Processing and Alignment

We checked the quality of raw reads with FASTQC (available at: http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc). Further processing of the raw reads was done using the pipeline HYBPHYLOMAKER (Fér & Schmickl, 2018). Sequence adapters were removed, and reads were quality-trimmed using TRIMMOMATIC v.0.32 (Bolger et al., 2014), with the default settings used in HYBPHYLOMAKER. Duplicated reads were removed with FASTUNIQ v.1.1 (Xu et al., 2012). Quality-trimmed individual raw reads were mapped to a reference sequence and then merged into contigs that are aligned for each gene separately. As pseudoreference for read mapping, we used a sequence consisting of the concatenated target exonic sequences, separated by stretches of 800 Ns. Mapping to the pseudoreference genome was done with BWA (Li & Durbin, 2009), and consensus sequences of the mapped reads were produced with CONSENSUSFIXER (available at: https://github.com/cbg-ethz/ConsensusFixer), because this is the only approach available in HYBPHYLOMAKER able to call ambiguity DNA codes in case of multiple bases per site in the mapped reads. For CONSENSUSFIXER we used the following settings: Minimum relative abundance of the alternative base ("plurality" in the setting file of HYBPHYLOMAKER) of 0.2 and a minimum read coverage for ambiguity calling ("mincov") of 5. Consensus sequences were matched to sequences of the target exons to produce PSLX files using blat (Kent, 2002). They were therefore combined to produce exon-wise matrices with "assembled_exons_to_fastas.py" (Weitemier et al., 2014). Matrices were aligned with MAFFT v.7.029 (Katoh & Standley, 2013), using the default program settings, and then gene-wise concatenated with AMAS (Borowiec, 2016). We filtered potential paralogues by running the script "HybPhyloMaker4a2_selectNonHet.sh" and setting the maximum number of heterozygous sites per locus ("maxhet" in the HYBPHYLOMAKER settings file) to 10. We continued the

pipeline with both data sets, filtered and unfiltered. HYBPHYLOMAKER performs two consecutive steps of missing-data filtering. First, sequences with more than a certain percentage of Ns in an alignment ("missingpercent" in the settings file) are deleted. We set this option to 40. Secondly, alignments with less than a certain percent of sequences ("speciespresence" in the settings file) are filtered out. We set this value to 75, so that alignments with more than 25% of missing sequences (i.e., with fewer than 25 sequences) were excluded. Finally, we calculated alignments statistics for the selected genes with AMAS, TRIMAL v.1.2 (Capella-Gutiérrez et al., 2009) and MSTATX (https://github.com/gcoll et/Mstat X/), as implemented in HYBPHYLOMAKER.

Phasing

Because the importance of retrieving allele information for a correct phylogeny estimation (especially in a recently diverged group) has been emphasized in recent studies (Eriksson et al., 2018; Andermann et al., 2019), we tested the impact of allele phasing on our phylogenetic analyses. To avoid the loss of allelic information during the process of allele mapping and consensus sequence production, we extracted the .bam files produced after mapping the read to the pseudoreference sequence from the HYBPHYLOMAKER pipeline. These were phased with SAMTOOLS v.0.1.19 (Li et al., 2009), using the combination of the commands "samtools sort" and "samtools phase." The phased .bam and .bai files were then placed in a new HYBPHYLOMAKER working directory for further processing within the pipelines workflow. A new "/10rawreads/" folder was also placed in the HYBPHYLOMAKER working directory, containing only the file with the modified sample names. The pipeline was therefore resumed for computing of the consensus sequences (this time allele-wise consensus sequences) by calling the script "HybPhyloMaker2_readmapping.sh" but specifying "mapping = no" in the setting file. The alignments obtained are hereafter called allele alignments. We ran the pipeline further as explained above, with the only differences that (a) matrices of exons belonging to the same gene were not concatenated; and (b) exon trees (instead of gene trees) were inferred and used for further analyses.

Position Filtering

We evaluated the effect of filtering alignments for positions that could add phylogenetic noise to the phylogenetic reconstructions, the so-called "phantom" spike positions (ambiguous

positions, usually situated close to indel-rich regions of the alignment, with abnormally high substitution rates). We followed the procedure illustrated by Fragoso-Martínez et al. (2017), calculating first the Phylogenetic Informativeness (Townsend, 2007) and net Phylogenetic Informativeness profiles using HYPHY (Pond et al., 2005) in the web portal PhyDesign (López-Giráldez & Townsend, 2011). As tree inputs, we used trees (different trees for different data sets; consensus or allele alignments, paralogue filtering or not, etc.) obtained from HYBPHYLOMAKER by concatenating gene alignments and FASTTREE (Price et al., 2010). The phylograms were then transformed to ultrametric trees using Penalized Likelihood methods (Sanderson, 2002) as implemented in the R package APE (Paradis & Schliep, 2018), with correlated rates and smoothing parameter (λ) set to 1. The relative timescale was set to 1 at the root.

Second, we used the estimated substitution rate per locus from HYPHY and the R script from Fragoso-Martínez et al. (2017) to identify the specific positions in the alignments with an unusually high substitution rate. As substitution rate threshold, we use 20 (the highest value tried in Fragoso-Martínez et al., 2017) because we were just interested in deleting poorly aligned positions and artefacts, without running the risk of excluding naturally highly variable regions in the alignments. Detected sites were checked and removed manually from the alignments. Lists of the removed sites for each scenario are given in Table S2. The cleaned alignments were then reintroduced in the HYBPHYLOMAKER pipeline, just before the missing-data filtering step.

Gene Tree and Species Tree/Network Analyses

For all different data sets, species trees were estimated (a) using maximum likelihood (ML) as implemented in RAXML v.8.2.4 (Stamatakis, 2014) on the concatenated data sets (concatML) and (b) applying the coalescent-based method ASTRAL-III (Chao et al., 2018). For the concatML analyses, concatenated alignments were produced in HYBPHYLOMAKER using AMAS and saved in phylip format. Analyses were run on the Scientific Computer Cluster of the GWDG (https://www.gwdg.de/) using RAXML, with alignments partitioned by genes, the GTRGAMMA model and 100 bootstrap replicates. Concatenated analyses were only performed for the consensus alignments, because there was no way of knowing how the phased alleles of a sample combined across loci. We also quantified branch support and conflict on the concatML trees with the Quartet Sampling method (Pease et al., 2018).

For the coalescent-based analyses, first gene trees (or exon trees in the case of the allele alignments) were reconstructed with RAXML. Analyses were run with 100 standard bootstrap replicates, with the GTRGAMMA model and partitioning by exons. Gene tree characteristics (average bootstrap support, average branch length, etc.) and correlations among alignments and gene tree characteristics were calculated and plotted in R v.3.5.2 (R Core Team, 2018), using modified scripts from Borowiec (2016) as implemented in HYBPHYLOMAKER. Gene trees were rooted, collapsing branches with bootstrap values <50, and combined into single NEWICK files using Newick Utilities (Junier & Zdobnov, 2010). Species trees were inferred applying the coalescent-based algorithm implemented in ASTRAL-III v.5.6.3 (Chao et al., 2018) with 100 multilocus bootstrap replicates. To assess the amount of gene tree conflict on branches, we measured the quartet support on the ASTRAL trees (Sayyari & Mirarab, 2016) for the main, the first alternative and the second alternative topologies ("-t 8" option in ASTRAL).

In addition, and to test for the extent of reticulate evolution among the sexual representatives of the *R. auricomus* complex, we inferred neighbor-net networks on the concatenated data sets using SPLITSTREE v.4.14.6 (Huson & Bryant, 2006). We used the General Time Reversible (GTR) model to estimate the genetic distance with estimated site frequencies and maximum likelihood (equal rates of site variation). Support values were calculated based on 100 bootstrap replicates.

Divergence Time Estimation

To obtain absolute divergence times, we used two secondary calibration points for the analyses. The crown age of the *R. auricomus* complex was set according to Hörandl (2004). In this study, the divergence time between *R. notabilis* and *R. carpaticola* was estimated as 914,000 years, whereas the split between *R. notabilis* and *R. cassubicifolius* was 535,000 years. Consequently, we applied a normally distributed time to the most recent common ancestor (TMRCA) prior to the crown age of the complex, with prior distribution ranging between the two age estimates (0.7245 ± 0.189 million years ago [Ma]). The root of the tree (only analysis including the outgroup; see below) was calibrated according to the dated phylogeny published in Emadzade and Hörandl (2011). Therefore, we used a normally distributed TMRCA prior for the root centred at 12.53 Ma (±2.3 Ma).

Divergence times were estimated with BEAST v.2.5.2 (Bouckaert et al., 2019) using the 50 gene alignments with the highest number of parsimony-informative (PI) sites from those of the

consensus data set, without paralogue filtering and cleaned from "phantom spikes." Prior to analyses, the selected loci were tested for selection using an ad-hoc test based on dN/dS values. For the scope, we set the correct reading frame for each alignment using the script "HybPhyloMaker4b_correctframe_translate.sh" in HYBPHYLOMAKER and setting the maximum number of stop codons to 1. Therefore, we calculated gene-wise ω (dN/dS ratios) with codeML (implemented in PAML v.4.8; Yang, 2007), using the corrected alignments and the concatML as species tree. We did not find any genes showing signal of positive selection (which would have been excluded from further analyses; Tables S3 and S4).

We estimated the best fitting substitution model for the selected loci (Table S5) using the Bayesian Information Criterion (BIC) in JMODELTEST v.2.1.10 (Darriba et al., 2012). Input files for the BEAST analysis were prepared in BEAUTI v.2.5.2 (Bouckaert et al., 2019). We used the *beast template and unlinked substitution models, clock models and gene trees for all loci. A strict clock was enforced and the average clock rate for one random locus was fixed to 1.66×10^{-6} while estimating all other clock rates in relation to this locus. We opted for adding an informative prior on the clock rate given the broad interval used for the calibration priors and because estimations based on isozyme data and the general clock used in Hörandl (2004) might not have been sufficiently accurate. The clock rate was derived from standard substitution rates for plant nuclear genes as in Pellino et al. (2013). Assuming a standard substitution rate of $5e^{-9}$ (Wolfe et al., 1989) and a generation time of 3 years for members of the *R. auricomus* complex (Pellino et al., 2013), the beast clock rate translates to $1.66e^{-6}$ mutations per thousand years. Formerly described species were treated as independent lineages. Tetraploid accessions of *R. cassubicifolius* were treated separately from the diploid representatives of the species.

Because the main goal of the analyses was to estimate the time frame of speciation events, and because these analyses were based only on a reduced number of loci, we used a soft-constrained approach on the species tree topology in order to reach faster convergence and to make the topology of the *BEAST tree similar to the topology of the species trees obtained with the other methods. In particular, the monophyly of the clade with the Illyrian taxa was enforced. Substitution models for each locus were adjusted to those found with JMODELTEST, and the Birth-Death model was used as prior on the species tree. Two independent analyses were run for 2×10^9 generations, sampling every 100,000th generation. Effective sample size (ESS) values and convergence between independent analyses were checked in TRACER v.1.6 (Rambaut & Drummond, 2007). Results of the two analyses were merged using LOGCOMBINER v.2.5.2 (Bouckaert et al., 2019) applying a burn-in of 10%. Finally, the remaining 18,000 trees were used to construct a maximum clade credibility tree with a posterior

probability limit set to 0.5 and "Mean Heights" for node heights using TREEANNOTATOR v.2.5.2 (Bouckaert et al., 2019). Another set of analyses was performed as described above, but including only the ingroup taxa, as many authors have highlighted that using outgroups in coalescent-based analyses violates some of the model assumptions (Agudo et al., 2017; Drummond & Bouckaert, 2015).

An additional set of analyses was performed for the above-mentioned data sets (including and excluding the outgroup) using only clock rates or calibrations (and then applying an exponential prior with mean equal to 0.01 for clock rates) to obtain absolute time estimates. This was done to evaluate the effect of using both calibration and clock rate priors in the same analysis, and to test for congruence of the resulting time estimates.

Ancestral Area Reconstruction

To evaluate hypotheses about past distribution patterns and dispersal and/or vicariance events in the evolutionary history of the sexual representatives of the *R. auricomus* complex, we performed an ancestral area reconstruction with the R package BIOGEOBEARS v.1.1.2 (Matzke, 2018). This approach allows several models of biogeographical evolution to be compared via likelihood-based model selection (Matzke, 2013). Because sexual *R. auricomus* species grow within or near the main European mountain ranges, we assigned them to the seven following areas: Northern Carpathian (A), Southern Carpathian (B), Illyrian region (at the edge of the southeastern border of the Alps; C), Northern Prealps (D), Central Apennines (E), Massif Central (F) and Pyrenees (G; Fig. 1). As tree input, we used the *BEAST tree without the outgroup taxa, because use of widely distributed and nonuniformly sampled outgroups would have been counterproductive for the analyses. We specified a distance matrix (including scaled geographical distances among predefined areas), and activated the x parameter, which can be used to estimate dispersal probability as a function of distance (Van Dam & Matzke, 2016). Ancestral areas were estimated under all six available models DEC, DIVA-LIKE and BAYAREA-LIKE, with and without the jump-dispersal parameter j. We compared the different nested models (e.g., DEC + x against DEC) with a likelihood ratio test (LRT) and selected the overall best-fitting model using the Akaike Information Criterion (AIC).

Results

Gene Capture and Sequencing

The average number of obtained reads per sample was 1,373,158 (range 867,000–2,178,952). Read quality was relatively high, and on average only 3.27% of reads per sample were discarded after quality check (Table S6). The average percent of mapped reads was 70.14%, with mapping success ranging from 53% to 82%, and only one sample (*Ranunculus notabilis* s.l.; 2) exhibiting a considerably lower number of mapped reads (26.89%; Table S7). Mapping success was not taxon-dependent. The sample with the highest mapping percent belongs to the outgroup (*R. pygmaeus*, 82.48%). From the initial 736 target loci, 700 were successfully captured. On average, 612 loci per sample were captured, with LH008-7 (*R. cassubicifolius*) capturing the most (621 loci), and *R. sceleratus* the least (595 loci).

Alignments and the Impact of the Different Filtering Schemes

The number of alignments used for the phylogenetic analyses varied according to the filtering scheme applied (Table 2). For the "consensus data set," about 15% of loci were excluded after missing data filtering and more than 25% after paralogue filtering (more than 50% of the total alignment length). Concerning the "allele alignments," working with exons made it possible to retrieve more loci (more than 80% and 50% of the total alignment length before and after paralogue filtering, respectively). More details are given in Table 2.

Allele phasing generated alignments with a considerably higher proportion of PI sites compared to the consensus alignments (5.82% and 1.4% on average, respectively). Paralogue filtering resulted in a loss of a fraction of PI sites in both the consensus and the allele data sets. When selecting the 50 loci for the BEAST analyses, 14.02% (78,458 bp) of the total matrix length was retrieved, comprising >23% of the total PI sites (1,971). Lists of the alignments used in the different filtering schemes, their characteristics and substitution models of the loci selected for the BEAST analyses are given in Table S5.

Phylogenetic Analyses and Species Trees

Species trees obtained from the concatML and coalescent-based analyses of the consensus data set showed topological incongruences across most of the filtering scenarios. When using allele alignments, coalescent-based species trees obtained from different filtering treatments always

showed the same topology, differing only slightly in support values and branch lengths. When taking the average local posterior probability (LPP) of the ASTRAL bootstrap trees as a criterion to compare different treatments, the paralogue filtering tool offered in HYBPHYLOMAKER significantly worsened the results (in a post-hoc Tukey's test only differences between treatments with and without paralogue filtering were significant; Table S8), whereas position filtering improved support values slightly in some cases (Fig. S1). No significant differences were noticed between "consensus" and "allele" data sets. Accordingly, and as a matter of simplicity, only results of the "consensus" and "allele" analyses, without paralogue filtering and with position filtering, are shown in Figs. 2 and 3. The rest of the trees are available as Figs. S2–S7.

We detected two main incongruences among species trees. (a) The position of the taxa within *R. notabilis* s.l. This clade is always highly supported, and in general samples cluster according to the taxonomy proposed by Dunkel et al. (2018), with the exception of "*R. austroslovenicus*" and "*R. calapius*" that in some cases are found to be polyphyletic. However, support values are usually low within the clade and phylogenetic relationships among taxa change consistently across analyses. (b) The position of *R. marsicus*, which often forms a clade with *R. envalirensis* s.l. but is resolved as sister to *R. flabellifolius* in some analyses. Support values of these phylogenetic relationships are generally low and branches very short.

Apart from these considerations, relationships among the main clades in the *R. auricomus* group are constant across analyses. The first split separates *R. cassubicifolius* s.l. from the rest of the taxa. The tetraploid *R. cassubicifolius* was well separated from the diploid accessions of the same species. In general, samples do not seem to cluster according to geographical origin within this clade. In the other main clade, samples cluster according to geographical distribution. *R. flabellifolius* (Southern Carpathians) occupies a basal position, whereas the clade with samples from the Massif Central and the Pyrenees (*R. envalirensis* s.l.) is sister to the Illyrian clade. The position of the Apenninian *R. marsicus*, as mentioned above, is equivocal. The same patterns can be found in the Neighbor-net results estimates. As for the concatML and the ASTRAL species tree, relationships within *R. notabilis* s.l. and within *R. cassubicifolius* s.l. are blurred and phylogenetic relationships receive low support (Fig. 4; Fig. S9).

The crown age of the *R. auricomus* complex is estimated to be ~733.5 thousand years ago (ka) (Table 3), somewhat older (~750 ka) in the analyses including outgroups (Fig. S9). Most of the speciation events occurred between 830 and 580 ka. The only exception is the Apenninian *R. marsicus*, which diverged from *R. notabilis* s.l. more recently (266–107 ka). However, the latter

result does not seem to be supported by the concatML and by the species tree analyses, in which *R. marsicus* is rather found to be related to *R. envalirensis* (but see support values in Figs. 2 and 3), and its divergence seems to have occurred much earlier. Results of analyses including only calibrations or clock rates were congruent to those of the main analyses (Table S9).

Differentiation within *R. notabilis* s.l., *R. envalirensis* s.l. and *R. cassubicifolius* s.l. occurred in the last 200 thousand years, with the crown ages of these three species estimated as 176, 147 and 130 ka, respectively (166, 128 and 119 ka in the analysis including outgroups; Fig. S9).

Ancestral Area

The LRT shows that results of models including the x parameter were significantly better ($p < .05$) than those without this parameter (Table S10). The j parameter did not improve the model fit for DEC and DIVA-LIKE, whereas it considerably ameliorated the BAYAREALIKE model fit. Overall, the best-fitting model was the DIVA-like + x (Table 4).

Results of the DIVA-LIKE + x model support a vicariance scenario, with the existence of a widespread ancestor for the *R. auricomus* complex (Fig. 5a). The most probable ancestral range of *R. cassubicifolius* s.l. was the Northern Carpathians and the Northern Prealps, while the ancestor of the remaining species most probably had a southern range, being distributed in the Southern Carpathians, Apennines, and Pyrenees. However, some uncertainty is present in the estimation of the ancestral areas for the deepest nodes of the phylogeny, and statistical differences between the most probable and alternative ancestral ranges are relatively low for some nodes (Fig. 5b).

Table 2. Alignment characteristics and performance of different filtering scenarios for the consensus and the allele data set. For the Local Posterior Probability (LPP) of the ASTRAL bootstrap trees, mean values are given with standard deviation in parentheses.

		Initial	**Missing-data filtering**			
					Paralog filtering	
				Position filtering		**Position filtering**
Consensus	**Loci (exons)**	700 (2,628)	605	605	521	521
	Total length	727,475	562,824	559,781	302,963	301,868
	Average proportion of PI sites	0.0145	0.0151	0.0152	0.0127	0.0127
	Average LPP ASTRAL trees		0.776 (0.223)	0.776 (0.225)	0.728 (0.238)	0.729 (0.239)
Allele	**Loci (exons)**	700 (2,628)	663 (2,495)	662 (2,493)	588 (1,815)	587 (1,813)
	Total length	728,903	597,424	595,772	372,948	371,426
	Average proportion of PI sites	0.0585	0.0611	0.0611	0.0553	0.0553
	Average LPP ASTRAL trees		0.770 (0.229)	0.767 (0.227)	0.736 (0.239)	0.735 (0.239)

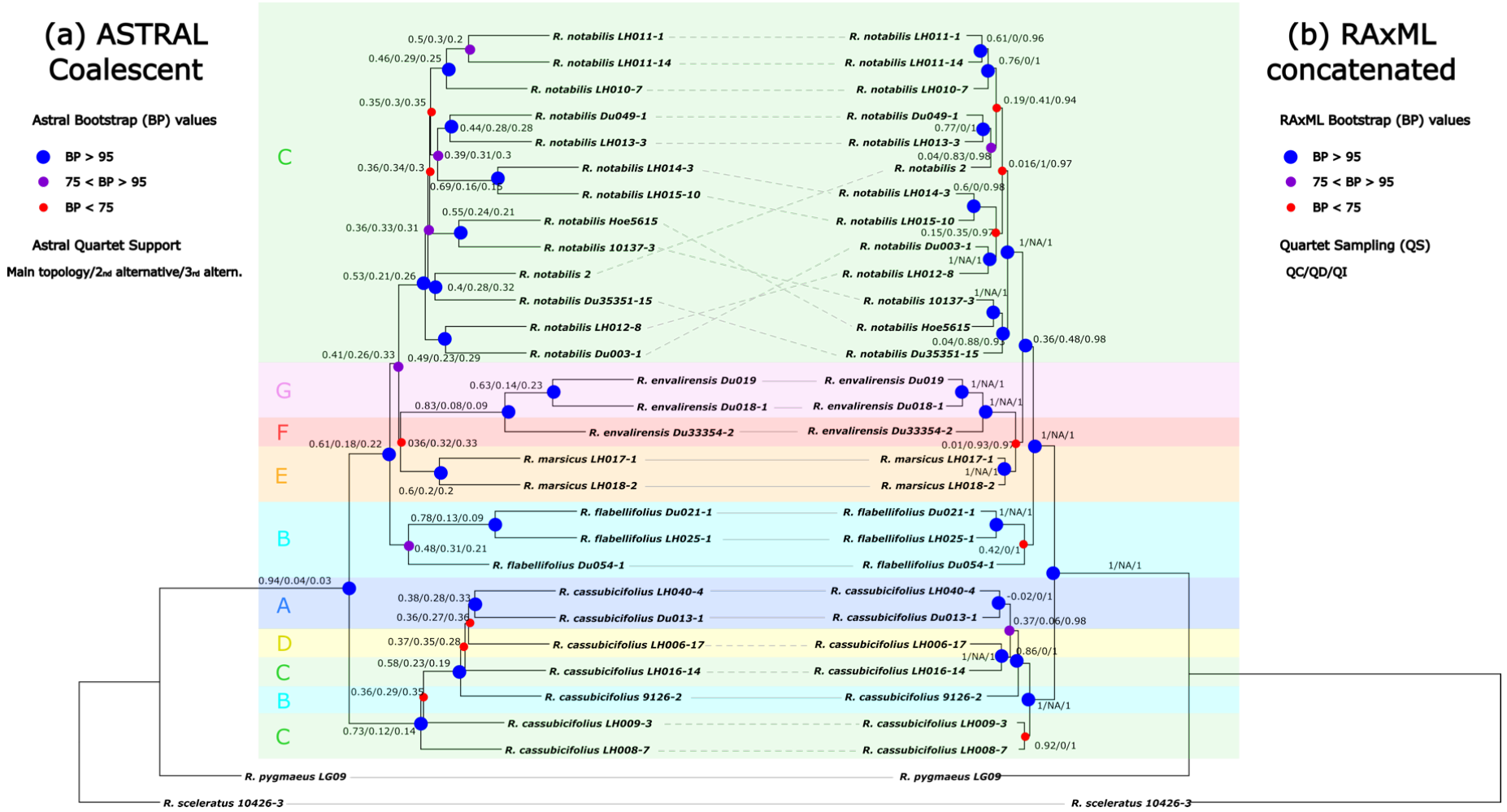


Figure 2. Tanglegram showing the coalescent-based ASTRAL species tree (a) and concatenated maximum likelihood tree (b), based on the consensus data set after applying position filtering and without paralogue filtering. Continuous lines indicate accessions with a congruent position, and dashed lines indicate incongruence. Circles on the nodes refer to bootstrap support values. Blue big circles are for support values above 95, purple medium-sized circles are for values between 75 and 95, and small red circles are for values below 75. Numbers at nodes indicate ASTRAL quartet support values (a) and Quartet Sampling results on the maximum likelihood tree (b). Colours and letters refer to the areas defined for the ancestral range reconstruction, as indicated in Fig. 1 [Colour figure can be viewed at wileyonlinelibrary.com].

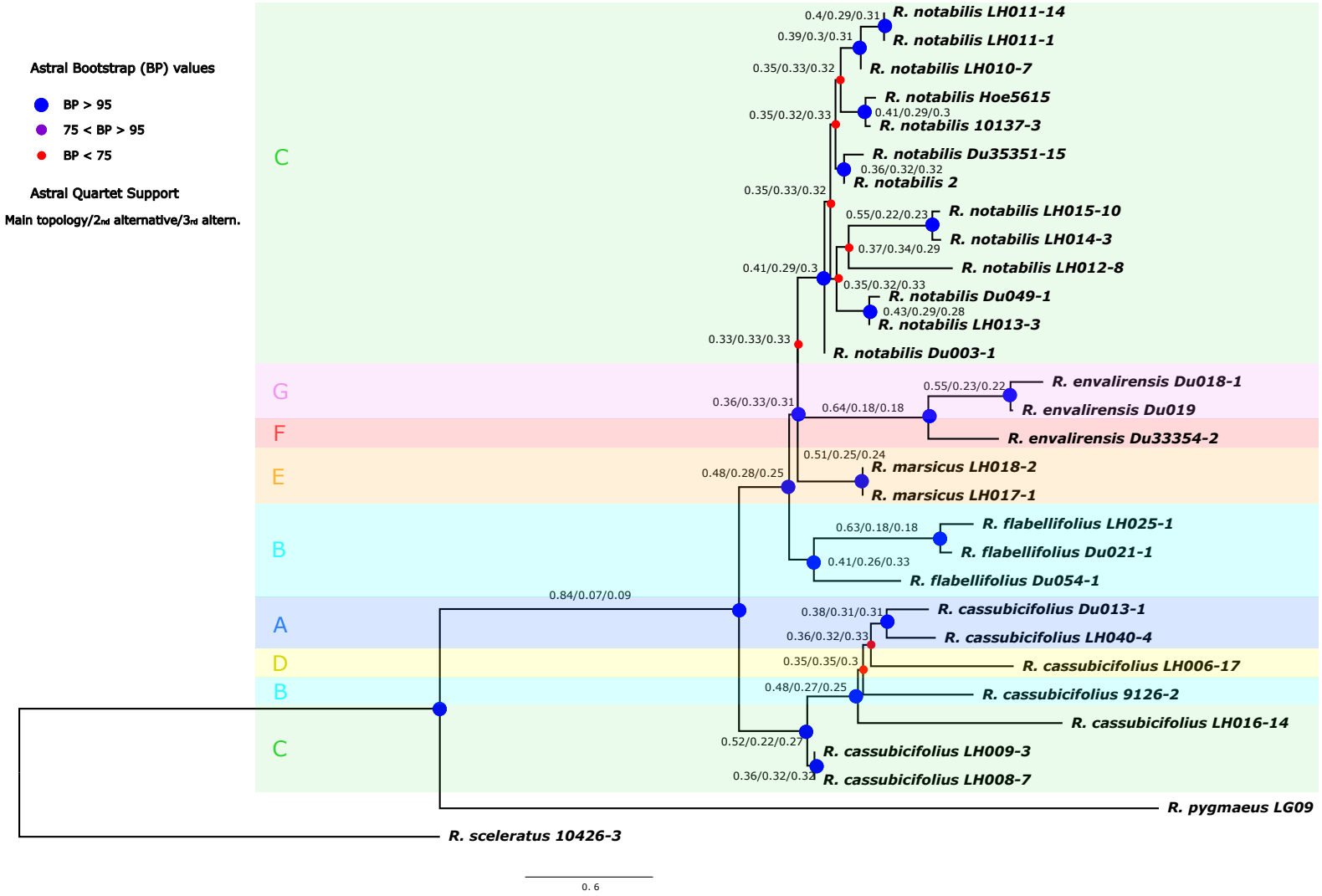


Figure 3. ASTRAL species tree based on the allele data set after applying position filtering and without paralogue filtering. Circles on the nodes of the trees refer to bootstrap support values. Blue big circles are for support values above 95, purple medium-sized circles are for values between 75 and 95, and small red circles are for values below 75. Numbers at nodes indicate ASTRAL quartet support. Colours and letters refer to the areas defined for the ancestral range reconstruction, as indicated in Fig. 1 [Colour figure can be viewed at wileyonlinelibrary.com].

Discussion

Phylogenetic Relationships among Sexual Species of the R. auricomus Complex

Disentangling phylogenetic relationships among recently and rapidly diverged lineages remains a challenge (Knowles & Chan, 2008). In the present study, we resolved the phylogeny of the sexual representatives of the less than 1-Myr-old *Ranunculus auricomus* complex. To this end, we analysed Target Enrichment data using concatenated ML and coalescent-based species-tree inference methods. Target Enrichment is considered to be a very suitable approach for resolving deep phylogenies, as the baits are usually designed for highly conserved regions (McCormack et al., 2013; McKain et al., 2018). Nevertheless, the potential of this method for resolving shallow phylogenies and application in population genetic studies has been highlighted recently (Harvey et al., 2016). Until now, this sequencing approach has been applied a few times to resolve phylogenetic relationships in recently radiating plant and animal groups (e.g., the 7.2-Myr-old genus *Heuchera* [Folk et al., 2015]; the ~3-Myr-old carnivorous plant *Sarracenia* [Stephens et al., 2015]; the rapid island radiation of the Philippine Shrews [<4.8 Ma; Giarla & Esselstyn, 2015]; and the lizard species group *Liolaemus fitzingerii* [2.6 Ma; Grummer et al., 2018]). The crown age of the *R. auricomus* complex was estimated to be 733.5 ka (829–639 ka) herein. In this sense, our study represents the most outstanding example of employing Target Enrichment data in recently and rapidly radiating species complexes.

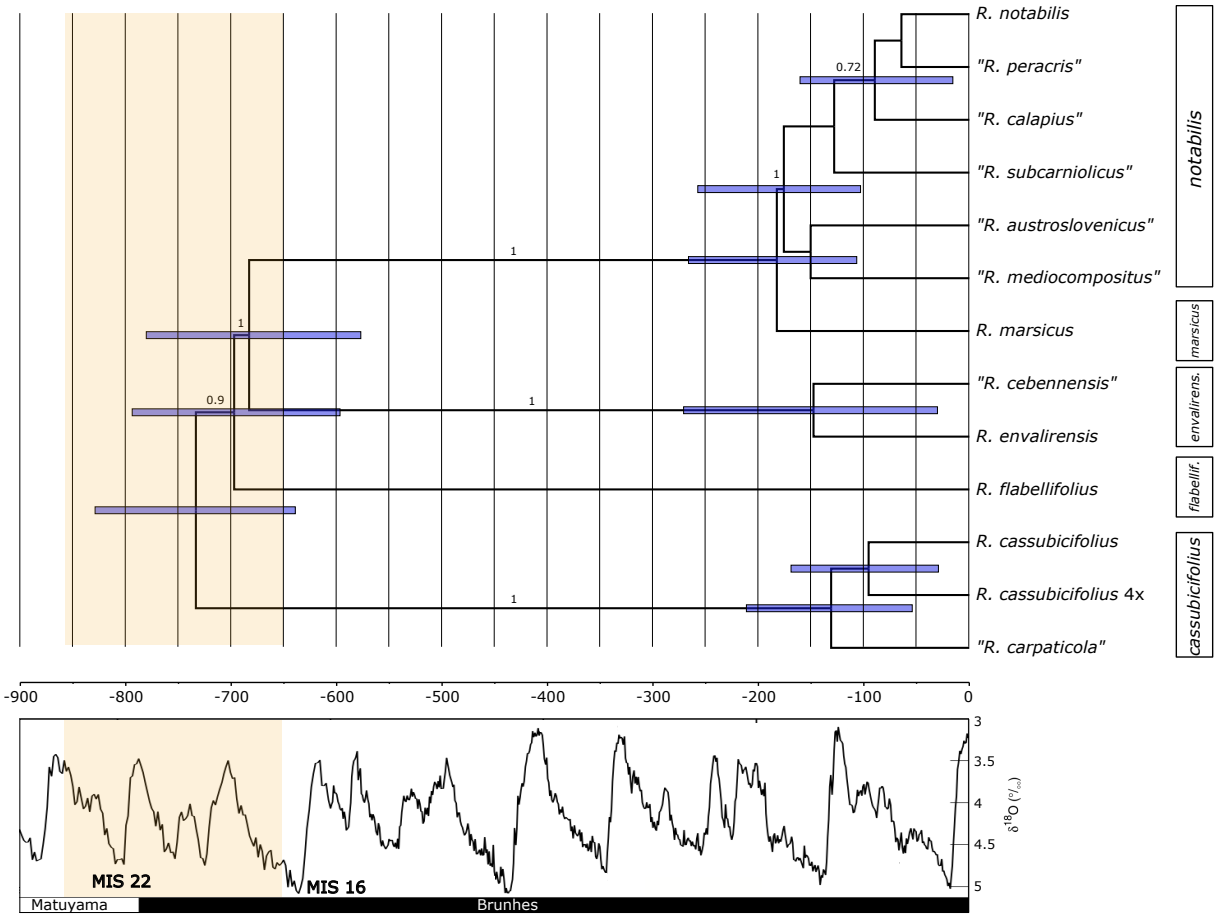


Figure 4. Maximum clade credibility tree estimated in *BEAST using the 50 most informative loci from the consensus data set. Blue bars indicate 95% highest posterior density (HPD) intervals of the age estimate (see Table 3). Numbers above branches indicate Bayesian posterior probabilities; only values above 0.7 are shown. On the right, bars indicate the current nomenclature according to Karbstein et al. (2020). The curve below the tree represents the $\delta^{18}O$ modified from Lisiecki and Raymo (2005). Valleys correlate with cold periods (glaciations) and peaks with warm periods (interglacials). Marine Isotope Stages (MIS) 22 and 16, representing the two most severe glaciations during the Mid-Pleistocene Transition (MPT), are also indicated. The orange region corresponds to the final phase of the MPT [Colour figure can be viewed at wileyonlinelibrary.com].

Table 3. Prior and posterior distributions of age estimates (ka) for the most important nodes of the *BEAST chronogram for *Ranunculus* (Fig. 4). HPD, highest posterior density.

Node description	Prior distribution		Posterior distribution	
	Median age	95% HPD interval	Mean age	95% HPD interval
root	724.5	914-535	733.5	829.1-639.2
R. flabellifolius stem age			697.2	793.9-596.8
R. envalirensis clade stem age			683	780.6-577.1
R. marsicus-R. notabilis s.l. split			182.5	266.3-106.7
R. notabilis s.l. crown age			175.9	257.5-103.1

Table 4. Summary of data likelihoods under each model, and results of statistical model choice. Estimates for the parameters dispersal parameter (d), extinction (e), x (dispersal probability as a function of distance) and j (jump speciation events at cladogenesis). The favoured model is in bold type.

Model	Max. No. areas	LnL	No. Parameters	d	e	x	j	AIC
DEC	7	-26.228	2	0.210	0.497	-	-	56.457
DEC+x	7	-22.757	3	2.275	1E-12	-2.020	-	51.513
DEC+j	7	-24.274	3	0.063	1.0E-12	-	0.054	54.548
DEC+x+j	7	-22.597	4	1.213	1.0E-12	-1.736	0.210	53.195
DIVA	7	-21.475	2	0.123	1.0E-12	-	-	46.950
DIVA+x	**7**	**-18.307**	**3**	**0.890**	**1.0E-12**	**-1.522**	**-**	**42.614**
DIVA+j	7	-21.473	3	0.121	1.0E-12	-	0.002	48.946
DIVA+x+j	7	17.718	4	2.450	1.0E-12	-2.243	1.0E-05	43.437
BAYAREA	7	-31.872	2	0.278	2.791	-	-	67.744
BAYAREA+x	7	-29.506	3	2.976	2.911	-1.771	-	65.012
BAYAREA+j	7	-26.982	3	0.059	0.536	-	0.063	59.963
BAYAREA+x+j	7	-25.354	4	0.336	0.218	-1.253	0.302	58.709

The main problems when working with rapidly radiating groups are: (a) Rampant incomplete lineage sorting (ILS) due to the fact that ancestral polymorphisms are carried through a series of nearly simultaneous speciation events; and (b) the lack of information due to the slow mutation rate of the (most often) coding regions used in studies employing a Target Enrichment approach (Giarla & Esselstyn, 2015). Using information from several (hundreds of) independent loci from multiple individuals per species, and applying coalescent-based species tree methods should improve phylogenetic inference and overcome the above-mentioned problems (Kumar et al., 2012). However, in Grummer et al. (2018) even 580 loci were not enough to provide significant support for interspecific relationships in the *L. fitzingerii* species

group. In *R. auricomus*, 500–600 loci were able to resolve phylogenetic relationships among the main lineages of the complex. Incongruence was found mostly within the Illyrian group and within the clade of *R. cassubicifolius* s.l. There, relationships among formerly described species were inconsistent across analyses and clades did not receive significant support, confirming the broader species delimitation provided by Karbstein et al. (2020). Diversification processes within these two species (i.e., *R. notabilis* s.l. and *R. cassubicifolius* s.l.) took place recently and in a very short time frame (i.e., the last 176 and 136 kyr, respectively). We believe that the lack of phylogenetic divergence might be the cause of the poorly resolved intraspecific relationships in these two species. As already observed in comparable complexes (Grummer et al., 2018; Stephens et al., 2015, for organellar genomes), there might have been too little time for the genomes to accumulate enough informative polymorphisms.

When looking at the backbone of the phylogenetic trees, the main source of incongruence is the tetraploid *R. marsicus*, which exhibits inconsistent positions in different concatML analyses, and in the ASTRAL and *BEAST species trees. However, all these sister relationships are not supported and the branches describing them are usually very short. Possible causes of these contrasting patterns can be either ILS or hybridization/introgression. An allotetraploid origin of *R. marsicus* with the parental contribution of ancestral relatives of *R. notabilis* and *R. envalirensis* might be a realistic and geographically plausible scenario, and structure analyses based on RADseq data in Karbstein et al. (2020) seem to support it to some extent. However, in the present analyses, the high quartet differential values of the quartet sampling (for the concatML; Fig. 2), and the comparable frequencies of the two alternative topologies in the quartet support of the ASTRAL trees, sustain the hypothesis of ILS as the cause of incongruence rather than hybridization (Fig. 3; Figs. S2–S7). Our Neighbor-net analyses do not support an allopolyploid origin of *R. marsicus* with the contribution of ancestral relatives of *R. notabilis* and R. *envalirensis*. In these analyses, *R. marsicus* seems to occupy a more basal position between *R. flabellifolius* and *R. envalirensis*, and no significant signal of reticulation can be observed. The two scenarios (ILS or reticulation producing incongruence in the data set) are not necessarily mutually exclusive and a definitive and unequivocal reconstruction of the processes giving rise to this polyploid taxon, which requires a different experimental design and more specific analyses (e.g., coalescent-based species network inference), is far beyond the scope of the present study.

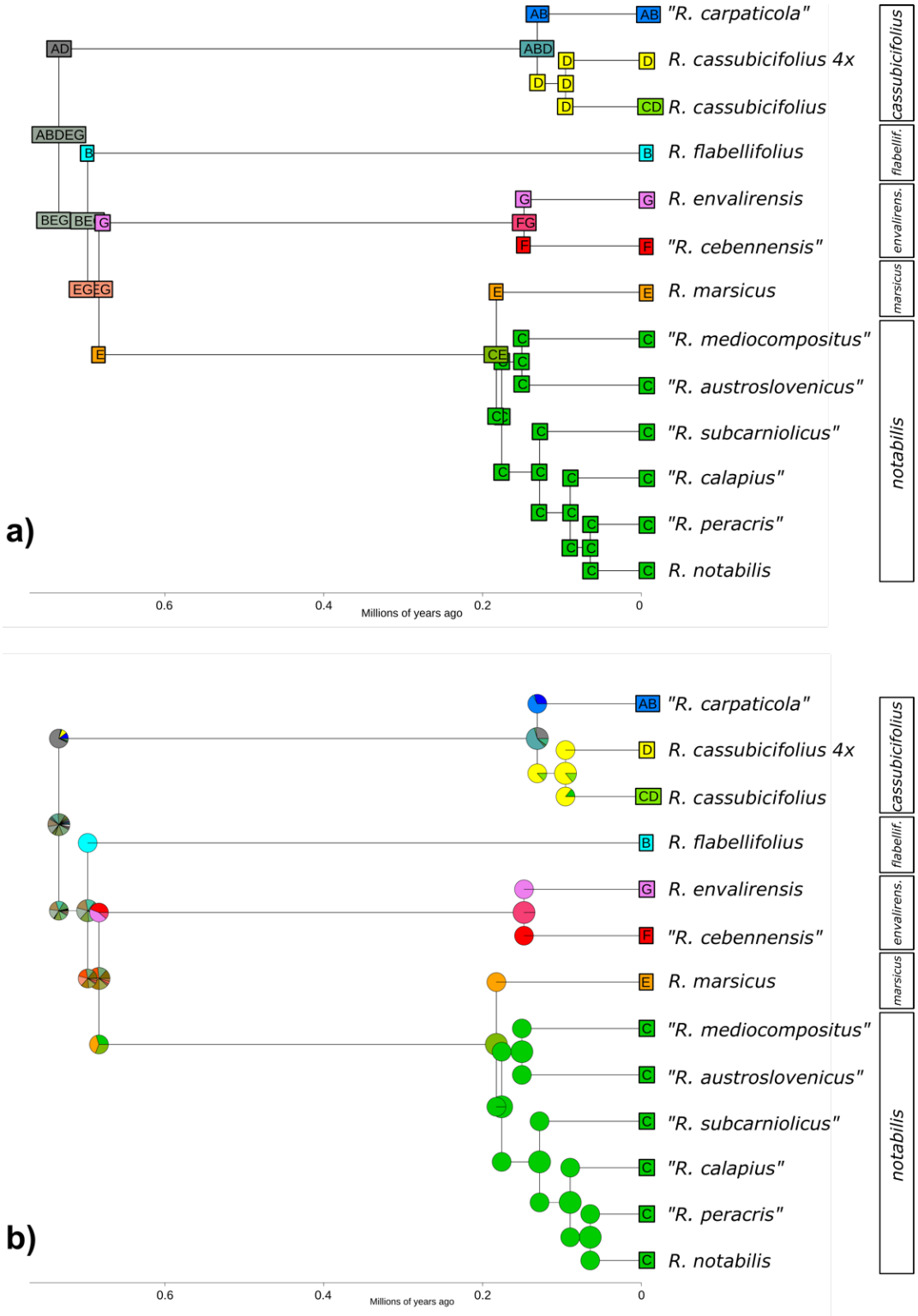


Figure 5. (a) Ancestral state reconstruction for the sexual representatives of the *Ranunculus auricomus* complex, as estimated in BIOGEOBEARS using the favoured DIVALIKE + x model. (b) Plot of the single-most-probable geographical range at each node (just before speciation) and post-split (just after speciation). Pie charts represent the probabilities of each possible geographical range per node. Capital letters indicate the areas defined for the analysis: A, Northern Carpathians; B, Southern Carpathians; C, Illyrian region; D, Northern Prealps; E, Central Apennines; F, Massif Central; and G, Pyrenees. Bars to the right of the diagrams indicate the species delimitation according to Karbstein et al. (2020) [Colour figure can be viewed at wileyonlinelibrary.com].

Concerning differences among tree inference methods (concatenated vs. coalescence-based) and performance of data sets from different filtering settings, we note that concatenation usually resulted in higher bootstrap values (compared to equivalent coalescent-based analyses) even though topologies were more discordant among trees. Similar patterns have already been observed in empirical and simulation studies (Herrando-Moraira & The Cardueae Radiations Group, 2018; Kubatko & Degnan, 2007; Weisrock et al., 2012). Remarkably, trees obtained by applying coalescent-based analyses to the phased data set were topologically identical, and minor differences concerned only branch lengths and support values (Figs. S5–S7). This is further evidence that using allelic information together with coalescent-based species tree approaches can considerably increase the trustworthiness of phylogenetic analyses (Andermann et al., 2019). Concerning the performance of different phasing and filtering settings, the paralogue filtering implemented in HYBPHYLOMAKER significantly worsened branch supports. Applying this option to diploids possibly resulted in the elimination of many non-paralogue (but variable) alignments. Moreover, paralogy might not be an issue in our data set, because the baits kit was specifically designed for members of the *R. auricomus* complex. Therefore, it is rather unlikely that the selected single-copy regions went through duplication in some of these very closely related taxa. No significant amelioration in support values of the bootstrap ASTRAL trees was produced when applying phasing (but see above considerations on tree topologies) or position filtering (Fig. S1).

Time Frame of Diversification Processes

The crown age of the *R. auricomus* complex was estimated to be ~730 ka (Table 3). Hörandl (2004) estimated the divergence time between three sexual species using isozyme allelic frequencies and genetic distances after Nei (1975). Ages of these divergence events were estimated to 914,000 and 535,000 years, respectively, even though they describe the same split in the phylogeny of the complex (the crown age of the whole group). Therefore, using a broad prior for the crown age of the complex and a relatively informative prior for the clock rate (according to the mutation rates estimated for the complex by Pellino et al., 2013), we were able to estimate the age of this node more precisely. Moreover, we obtained the time frame of all speciation events among sexual representatives of *R. auricomus*. Interestingly, almost all these events have taken place in a very short time between 830 and 580 ka (Fig. 4). The only exception is the tetraploid species *R. marsicus*, which seems to have diverged much more recently (but see considerations about uncertainty above). The divergence estimates for all these

speciation events coincide with the Mid-Pleistocene Transition (MPT; Clark et al., 2006; Lisiecki & Raymo, 2005), one of the most important climatic changes occurring in the Pleistocene. During this interval (1.2 Ma to 650 ka), the relatively low-amplitude 41-kyr climate cycles of the earlier Pleistocene were progressively replaced by high-amplitude 100-kyr cycles (Lisiecki & Raymo, 2005). Glaciations became longer and more severe, and the average global ice volume increased consistently, with larger increases at high and middle latitudes of Eurasia and North America (Elderfield et al., 2012; Raymo et al., 2006). This process became progressively stronger, and evidence from all northern continents indicates that Marine Isotope Stage (MIS) 22 (~870–880 ka) was the first of the major glaciation events characterizing the later Pleistocene (Head & Gibbard, 2005). At the end of the MPT (MIS 16; ~650 ka), the most substantial glaciation yet experienced in the Northern Hemisphere took place (Head & Gibbard, 2005). This tremendous climatic transition was probably the consequence of the increase in atmospheric moisture provided by warm subtropical water in the Western Mediterranean region that, combined with a decrease in boreal summer insolation, contributed to the feed of the ice-sheets in Central and Northern Europe (Bahr et al., 2018; Sánchez-Goñi et al., 2016). The onset of the global cooling phase and ice expansion caused a major restructuring of the global vegetation and faunal systems (Alroy et al., 2000; Gaboardi et al., 2005; Zhao et al., 2017; Zhou et al., 2018).

Diversification processes within species took place more recently, most probably within the last two glaciation cycles. Diverging populations of *R. envalirensis* s.l., *R. notabilis* s.l. and *R. cassubicifolius* s.l. differentiated allopatrically (in the first case) or parapatrically (in the latter two cases) within the last 160 ka (Fig. 4). As for the alpine taxa, and probably also for montane plant species, the last two glaciation cycles were a triggering force shaping intraspecific genetic patterns, but were not sufficiently long to complete speciation processes. Overall, our age estimates are concordant with those found in previous studies (Emadzade & Hörandl, 2011; Hörandl, 2004; Pellino et al., 2013). In Hörandl (2004) and Emadzade and Hörandl (2011), diversification processes within *R. cassubicifolius* were found to be slightly older, 317 and 640 ka, respectively. In our analyses, sexual autotetraploids in *R. cassubicifolius* (Hörandl & Greilhuber, 2002) were found to be ~95,000 years old. This result is in accordance with previous studies estimating the maximum age of hexaploid apomictic derivatives as 80,000 years (Pellino et al., 2013), and confirms the general assumption that the origin of polyploids is associated with the last glaciation cycle.

Biogeographical History

In our biogeographical analyses, AIC preferred models supporting the existence of a widespread ancestor for the *R. auricomus* complex, stressing the importance of vicariance in triggering speciation events among sexual representatives of the complex (Table 4). The ancestor of the whole complex was reconstructed to be distributed in the whole of Europe, although with some uncertainty (Fig. 5). A first vicariance event separated populations with a northern distribution (giving rise to *R. cassubicifolius* in the Northern Prealps and the Western Carpathians) from those with a southern distributional range. Further vicariance events then gave rise through allopatric speciation to the remaining sexual species of the complex (i.e., *R. flabellifolius*, *R. envalirensis*, *R. marsicus* and *R. notabilis*).

Almost all of these vicariance events occurred in a restricted period, coinciding with the final stage of the MPT. At the beginning of this period of drastic climatic change, humid and mild temperate forests with deciduous oaks and tertiary elements (such as *Carya* and *Zelkowa*) occupied a vast area of continental Europe (Magri & Palombo, 2013; Pastre et al., 2007; Stuchlik & Wójcik, 2001). Such environments must have been similar to the ecosystems inhabited nowadays by many of the sexual representatives of the complex. Climatic deterioration (the progressive cooling and drying of the climate) resulted in the fragmentation of such environments and the establishment of grass-dominated open vegetation adapted to cold and arid conditions (Magri & Palombo, 2013). Towards the end of the MPT, tundra vegetation was well established in Central and Northern Europe (Stuchlik & Wójcik, 2001). A similar situation was registered in temperate Eastern Asia, where broadleaved mixed forests dominated by *Quercus* and *Pinus* shrank significantly during the MPT, transforming into grass-dominated vegetation definitively by around 0.7 Ma (Zhou et al., 2018).

Areas at the foothills of mountain systems might have acted as microrefugia for species of deciduous temperate forest ecosystems. Indeed, for the boreo-montane *P. verticillatum*, putative glacial refugia were identified in the area surrounding the glaciated Alps and in the foothills of the Carpathian and Balkan mountain systems (Kramp et al., 2008). The ancestor of the *R. auricomus* complex might have therefore inhabited the deciduous temperate forests widespread in Europe during the first phases of the MPT. Climatic deterioration and the retreat of temperate forests to isolated refugial areas efficiently separated *R. auricomus* populations, which finally gave rise to the different sexual species of the complex by allopatric speciation. A similar pattern of allopatric speciation in the Mid-Pleistocene was inferred for the temperate-montane Asian species *Dysosma versipellis* (Qiu et al., 2009). Surprisingly, the sexual species

of the *R. auricomus* complex did not show long-distance dispersal, which has been observed in many other plant genera (Knapp et al., 2005), and also in other forest species from other clades of the genus *Ranunculus* (Emadzade et al., 2011). The achenes as diaspores exhibit no obvious specialized structures potentially suitable for long-distance dispersal via anemo- or zoochory. The absence of long-distance dispersal in the sexual *R. auricomus* species may also reflect limited abilities to adapt to new environments or novel niches. Outcrossing and self-sterility in the diploids may be other reasons for low colonization abilities (Hörandl, 2008). Nevertheless, moderate range expansions in warmer periods between 600 and 200 ka should be considered, especially in the *R. cassubicifolius*, *R. envalirensis* and *R. notabilis* clades, as their respective crown group areas are larger than their stem group areas (Fig. 5a). Such geodispersal during interglacials might have potentially resulted in secondary contact hybridization and an allopolyploid origin of the apomictic taxa. However, range expansions remained at a regional scale and did not allow sexual taxa to cover again the whole ancestral area of the *R. auricomus* complex. In fact, dispersal and vicariance scenarios are not mutually exclusive (Sanmartín, 2009).

A second wave of intraspecific vicariance occurred ~200–100 ka in the Alpine–Carpathian system (*R. cassubicifolius* s.l.), between the Pyrenees and the Massif Central (*R. envalirensis* s.l.) and, to a lesser extent, within the Illyrian clade (*R. notabilis* s.l.). Most taxa occur today disjunctly in very restricted areas with few populations, although appropriate habitats would be broadly available. The disjunct patterns suggest that survival of the last glacial periods occurred also in microrefugia where local conditions protected plant species from unfavourable regional conditions (Rull, 2009). A distribution pattern suggesting survival in proximal microrefugia is specifically apparent in *R. cassubicifolius* in the northern foothills and foreland of the Alps. Here, the species occurs today only in patchy areas that remained ice-free between the large glacier tongues of the Last Gacial Maximum (LGM; see map in Paun et al., 2006a).

Strikingly, the species are still distributed just in their very restricted areas, despite post-glacial reforestation of temperate Europe. The failure of the sexual species in post-glacial geodispersal may be explained by biotic barriers, namely congeneric competitors. The origin of polyploid apomictic derivatives is presumably younger than 100 thousand years (Pellino et al., 2013), and they rapidly colonized various habitats following the LGM (Paun et al., 2006a). The rapid range expansion of younger, polyploid apomictic derivatives of the *R. auricomus* complex could have blocked range expansion of sexual progenitors (Hörandl, 2006). Polyploid apomicts could be more competitive, but they could also benefit from having no minority cytotype disadvantage (Levin, 1975): Sexual outcrossers need a conspecific pollinator for reproduction, while the

apomicts can reproduce as single individuals (Kirchheimer et al., 2018). Fertilization of sexual individuals by sympatric apomictic pollen donors can even result in the formation of aneuploid, introgressed offspring. Such processes were observed in diploid *R. notabilis* populations mixed with tetraploid apomicts (Hörandl et al., 2000) and may also occur between sexual and apomictic cytotypes of *R. marsicus* (Dunkel, 2011). However, other distantly related and widespread species of the genus occupying the same habitats within the forest zone (e.g., *R. polyanthemos* group, or *R. lanuginosus*) could also have been competitors for sexual taxa of *R. auricomus*. *R. marsicus* and *R. envalirensis* established small populations in subalpine grassland habitats in the Apennines and Pyrenees, respectively. Here, members of the widespread, distantly related *R. montanus* group are very abundant (E. Hörandl, L. Hodač, pers. obs.) and are probably more successful. Hence, geographical parthenogenesis is resulting not only from the colonization success of apomicts, but also from failure of sexual species to expand ranges.

We conclude that biogeographical histories in temperate forest plants are complex, as area configurations changed over time due to climatic fluctuations. Vicariance, allopatric speciation, and fragmentation of habitats during colder climatic periods shaped the evolutionary history of the sexual species of the *R. auricomus* complex. Our observations fit a hierarchical vicariance model (sensu Ronquist & Sanmartín, 2011). However, even after disappearance of climatic barriers after the LGM, biotic barriers of congeneric competitors might have blocked range expansions of sexual species of the *R. auricomus* complex.

Acknowledgements

We thank the German Research Foundation for project funding (DFG, Ho4395/10-1) to E.H., within the priority programme “Taxon-Omics: New Approaches for Discovering and Naming Biodiversity” (SPP 1991), and the Botanical Museum, University of Oslo (O) for the material of *R. pygmaeus*. We also thank Mr. F. G. Dunkel for providing herbarium vouchers and living material for the recently discovered diploid populations from the Illyrian region and France.

Author Contributions

S.T. and E.H. conceived the ideas; S.T., K.K., L.H. and E.H. collected materials and data; S.T., K.K. and C.P. analysed the data; and S.T. wrote the paper with contributions from all co-authors.

Data Availability Statement

The raw reads used in this study are deposited in the Sequence Read Archive (SRA) under the BioProject PRJNA628081.

Orcid

Salvatore Tomasello https://orcid.org/0000-0001-5201-1156
Kevin Karbstein https://orcid.org/0000-0003-1424-6557
Ladislav Hodač https://orcid.org/0000-0002-6885-1317
Claudia Paetzold https://orcid.org/0000-0002-4128-6645
Elvira Hörandl https://orcid.org/0000-0002-7600-1128

Chapter 4 – A Geographical Parthenogenesis Scenario

Moving beyond assumptions: Polyploidy and environmental effects explain a geographical parthenogenesis scenario in European plants

Karbstein., K., Tomasello., S., Hodač., L., Lorberg, E., Daubert, M., & Hörandl, E.

Molecular Ecology, *30*, 2659–2675 (2021a).
doi: 10.1111/mec.15919 (CC BY 4.0, changes in format and citation style)

Geographical parthenogenesis (GP) describes the phenomenon that apomicts tend to have larger distribution areas and/or occur at higher altitudes or latitudes compared to sexual relatives. However, the complex effects of genome-wide heterozygosity, ploidy, reproduction mode (sexual vs. apomictic), and environment shaping GP of plants are still not well understood. We ascertained ploidy and reproduction mode by flow cytometry of 221 populations, and added genomic RADseq data (maximum 33,165 loci) of 80 taxa of the *Ranunculus auricomus* polyploid plant complex in temperate Europe. We observed 7% mainly diploid sexual, 28% facultative apomictic (mean sexuality 7.1%), and 65% obligate apomictic populations. Sexuals occupied a more southern, smaller distribution area, whereas apomicts expanded their range to higher latitudes. Within the complex, we detected three main genetic clusters and highly reticulate relationships. A genetically-informed path analysis using GLMMs revealed several significant relationships. Sexuality of populations (percent of sexual seeds) was higher in diploids compared to polyploids, associated with more petals, and similar between forests and open habitats. In contrast to other apomictic plant complexes, sexuality was mainly positively correlated to solar radiation and isothermality, which fits the southern distribution. We found up to three times higher heterozygosity in polyploids compared to diploids, and generally more heterozygous individuals in forests compared with open habitats. Interestingly, we revealed a previously unknown positive association between heterozygosity and

temperature seasonality, suggesting a higher resistance of polyploids to more extreme climatic conditions. We provide empirical evidence for intrinsic and extrinsic factors shaping the GP pattern in a polyploid plant complex.

Keywords

flow cytometric seed screening, genome-wide heterozygosity, geographical parthenogenesis, RADseq, *Ranunculus auricomus*, sexuality

Introduction

Geographical parthenogenesis (GP) is the phenomenon that closely related sexual and asexual organisms show different distributional ranges (Vandel, 1928). Asexual organisms usually occupy larger areas, tend to be distributed at higher altitudes, and are biased towards northern regions or more extreme environments compared to their sexual relatives (Bierzychudek, 1985; Hojsgaard & Hörandl, 2019; Hörandl, 2006; Kearney, 2005; Lo et al., 2013). Often, the more northern distribution of asexual lineages can be explained by the colonization of habitats severely affected by past glacial periods (Late Pleistocene; Bierzychudek, 1985; Kearney, 2005; Mráz et al., 2009; Hartmann et al., 2017). Moreover, sexuals are frequently distributed within the range of apomicts (Hörandl, 2006).

In plants, sexual relatives of apomictic lineages (which reproduce via asexually formed seeds) predominantly occur in temperate to boreal zones (Hörandl, 2006). High climatic fluctuations during the Pleistocene led to range shifts and offered frequent opportunities for interspecific hybridization, probably giving rise to apomixis in these areas (Carman, 1997; Dobeš et al., 2004; Pellino et al., 2013; Tomasello et al., 2020). GP has long been recognized in plants (Bierzychudek, 1985; Kearney, 2005; Vandel, 1928), but reasons for this pattern have been continuously debated. The success of apomicts depends on the complex combination of several factors (Hojsgaard & Hörandl, 2019; Hörandl, 2006).

Apomixis is frequently connected to polyploidy and/or hybridization in plants (Asker & Jerling, 1992; Hojsgaard & Hörandl, 2019). Apomixis is a modification of the sexual pathway,

(epi)genetically controlled and thus heritable. Intriguingly, hybridization rather than polyploidization seems to activate apomixis (Barke et al., 2018; Hand & Koltunow, 2014; Kantama et al., 2007; Kumar, 2017; Nogler, 1984). In plants, apomixis is usually facultative, meaning that an individual can produce both sexual and apomictic seeds (Asker & Jerling, 1992). A single dominant Mendelian factor in the heterozygous state with a wild type allele controls egg cell formation without meiosis, and thus the ability for sexual reproduction cannot be completely eliminated (Nogler, 1984; Ozias-Akins & van Dijk, 2007). Evolution of new genotypes via facultative sexuality enhances genotypic diversity and maintains genetic variation for selection processes on apomictic populations (Cosendai et al., 2013; Gornall, 1999; Hörandl & Paun, 2007; Lushai et al., 2003). In facultative apomicts, frequencies of sexual ovules or seeds vary with ploidy level, pollination characteristics, and environmental conditions (Asker & Jerling, 1992; Espinoza, 2002; Naumova et al., 1999; Schinkel et al., 2016; Ulum et al., 2020). However, the degree of facultative sexuality under different environmental conditions across distributional ranges of apomictic species complexes has been less investigated.

On the one hand, polyploidy and/or hybrid origin can be advantageous to flowering plants and particularly to apomicts (Alix et al., 2017; Bierzychudek, 1985; Leebens-Mack et al., 2019; Soltis & Soltis, 2000). Multiple genome copies allow for gene subfunctionalizations, epigenetic change, and thus a differential expression of parental genes, providing larger physiological and phenotypic flexibility to respond to environmental stress (Comai, 2005; Hörandl, 2006; Chen, 2007; Van de Peer et al., 2017). On the other hand, known defects of polyploidization and hybridization exist; for example, consequences of extra gene dosage such as gene redundancy or deleterious effects, maintained incompatibility between parental subgenomes, mitotic loss of chromosomes, or abnormal meiosis reducing fertility (Comai, 2005; Qiu et al., 2020). In the long term, the accumulation of deleterious mutations in apomictic polyploids could drive them to extinction (Muller's Ratchet; Muller, 1964). However, several polyploidization events occurred in the evolutionary history of flowering plants, which are the most species-rich group of land plants (ca. 370,000 species; Lughadha et al., 2016) and exhibit several key innovations (One Thousand Plant Transcriptomes Initiative, 2019).

Particularly the combination of polyploidy and hybridization (allopolyploidy) is considered to create biotypes with a novel genomic composition (Abbott et al., 2013). Using traditional genetic markers, allopolyploids often showed higher proportions of heterozygous individuals within populations than autopolyploids (Coughlan et al., 2017; Hörandl, 2004; Hörandl & Paun, 2007; Štorchová et al., 2002). On an individual level, genomic heterozygosity has important

benefits such as buffering effects of recessive deleterious mutations, novel genetic combinations, heterosis effects, differential expression of homeologous genes, increased fitness compared to sexual progenitors, and changes in secondary metabolites potentially important in biotic interactions (Comai, 2005; Chen, 2007; Qiu et al., 2020). Apomictic allopolyploids exhibit fixed heterozygosity at duplicated loci at homeologous chromosomes (Gornall, 1999; Hörandl & Paun, 2007). Because of the lack of meiosis and segregation, heterozygosity is retained over generations, probably maintaining favorable gene combinations. Allelic sequence divergence in the absence of recombination further increases heterozygosity and can occur particularly in high polyploid genomes (Pellino et al., 2013). Since asexual genomes are inherited as big linkage groups, selection against deleterious mutations is less efficient than in sexual lineages (Felsenstein, 1974). Interestingly, no genome-wide mutation accumulation was found in facultative apomictic, hexaploid *Ranunculus auricomus* plants, probably prevented by low levels of facultative recombination (Pellino et al., 2013; Hodač et al., 2019). However, detailed investigations of intraindividual genome-wide heterozygosity over large distributional areas are so far missing.

Uniparental reproduction can additionally promote extensive range expansion: Apomicts have improved colonizing abilities due to self-compatibility, and thus no need for pollination vectors (Stebbins, 1950; "Baker´s law": Baker, 1967; Hörandl, 2008, Cosendai et al., 2013). This ability is most effective when single or few apomictic individuals found a new population after long-distance dispersal, or in extreme environments at high altitudes/latitudes, where pollinator density is expected to be fluctuating. Nevertheless, little is known about differential pollinator attraction of sexuals and apomicts.

In plants, niche shifts of apomicts towards pastures, anthropogenic meadows, and moderately disturbed habitats have been frequently observed for Central and Northern Europe (Asker & Jerling, 1992; Hörandl & Paun, 2007). Pronounced climatic niche shifts to colder climates explained GP patterns of the alpine species *Ranunculus kuepferi* (Kirchheimer et al., 2016; Schinkel et al., 2016). In alpine *Potentilla puberula*, apomicts occupy human-made habitats with higher water availability more frequently than sexual tetraploids (Alonso-Marcos et al., 2019). In lower altitudes, climatic niche shifts were observed in North American *Crataegus* (Coughlan et al., 2017) and South American *Paspalum* (Karunarathne et al., 2018).

Niche shifts could be enabled by “general-purpose genotypes”, i.e., by asexual clones that are well-suited to a wide array of environments allowing them to occupy large distributional ranges under varying environmental conditions (GPG hypothesis; Baker, 1965; Lynch, 1984; van Dijk,

2003). In contrast, the "frozen niche-variation" hypothesis (FNV, Vrijenhoek 1979; Vrijenhoek & Parker, 2009) claims that several specialized clones with limited geographic ranges inhabit specific niches across the species distributional range. Both patterns have been found in plants (Hörandl, 2006; Hörandl & Paun, 2007; Cosendai et al., 2013; Coughlan et al., 2017). Geographic structure and isolation-by-distance (IBD) were found in sexual populations of *R. kuepferi* and *Boechera*, but not in the apomictic ones (Cosendai et al., 2013; Lovell et al., 2014).

However, separate and combined effects of intrinsic and extrinsic factors shaping GP patterns are still not well understood for large apomictic polyploid species complexes across wide geographical scales. The pattern is specifically enigmatic for temperate lowlands, where strong climatic gradients are missing.

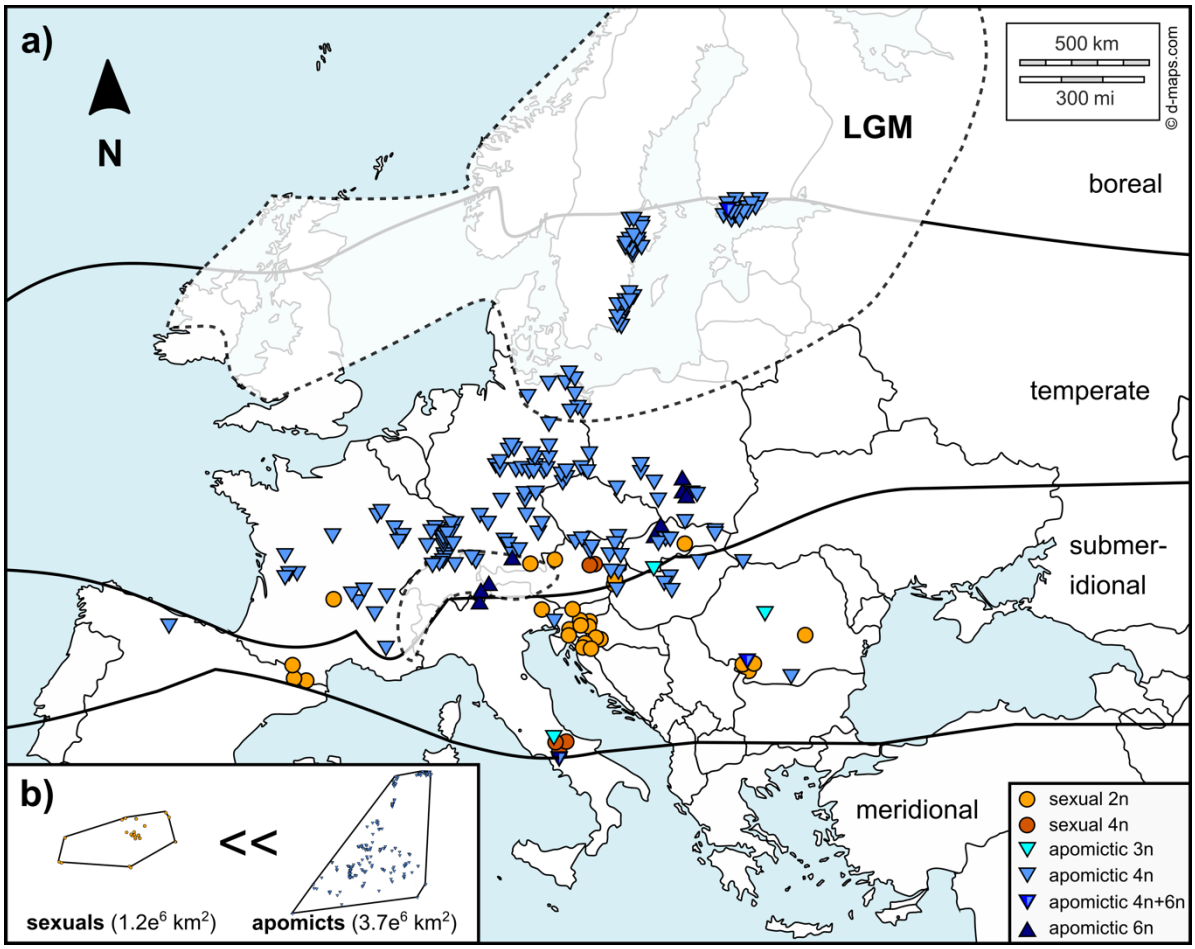


Figure 1. (a) Sampling locations of 251 *R. auricomus* populations across Europe. Symbols indicate main ploidy levels and reproduction modes (see also Tables S1 and S2). Floristic zones (Meusel et al., 1978) and extent of the Last Glacial Maximum (LGM, 20,000 years ago, Becker et al., 2015) were drawn. (b) Range area (km^2) of sexuals and apomicts. The original map was downloaded from https://d-maps.com/.

The *R. auricomus* complex comprises more than 840 taxa (Hörandl & Raab-Straube, 2015) that are mainly distributed from Greenland, temperate to submeridional Europe to Western Siberia (GBIF Secretariat, 2017; Jalas & Suominen, 1989; Fig. 1). Due to the huge number of described morphospecies, it is one of the largest European apomictic polyploid complexes and a model system to study apomixis (Barke et al., 2018; Hodač et al., 2014; Hojsgaard et al., 2014; Hörandl et al., 2009; Nogler, 1984). The complex comprises five, geographically distinct, mainly diploid sexual progenitor species (Karbstein et al., 2020b, Tomasello, et al., 2020). The high diversity of polyploid agamospecies originated most probably from hybridization and polyploidization between sexual progenitor species (allopolyploidization; Paun et al., 2006; Hörandl et al., 2009; Hodač et al., 2014, 2018; Hojsgaard et al., 2014). The crown age of *R. auricomus* is estimated ca. 733.5 kya with speciation events of sexuals between 830–580 kya and age of apomicts less than 100,000 years (Hörandl, 2004; Pellino et al., 2013; Tomasello et al., 2020). Phylogenetic relationships and genome evolution among described apomictic polyploid taxa of the *R. auricomus* complex are still unknown. Sexual species are predominantly self-incompatible, whereas pseudogamous apomicts are self-fertile (Hörandl, 2008, 2010; Karbstein et al., 2020b, Karbstein et al., 2020b). Apomicts often exhibit no or a variable number of petals (petaloid nectary leaves) whereas sexuals usually possess five of them. It is still unknown whether the number of petals is related to the degree of sexuality, potentially improving pollinator attraction.

Ranunculus auricomus populations occupy a variety of habitats ranging from lake-, stream-, and riversides to marshy, humid, semi-dry forests, waysides, and anthropogenically formed meadows. Niche modelling across the complex range showed that polyploids are distributed in slightly drier and colder habitats than diploids (Paule et al., 2018). However, since large-scale reproductive and genomic data are missing so far, there is only limited knowledge for the complex about the occurrence of a GP pattern and potential factors shaping it.

Due to varying ploidy levels, reproduction modes, heterozygosity, habitat types, and distribution between sexual and apomictic taxa across Europe, the *R. auricomus* complex is an appropriate model system to study the intrinsic and extrinsic factors leading to GP. In the present study, we addressed the following questions: (i) Is there a GP pattern in temperate Europe? (ii) How is the complex genetically structured? (iii) Is the degree of sexuality related to ploidy, habitat, or climate? (iv) Is sexuality a good predictor for the number of petals? (v) How is the genome-wide heterozygosity related to sexuality, ploidy, reproduction mode, habitat, or climate?

Materials and Methods

Study Locations and Population Sampling

The study comprised five sexual species and ca. 80 apomictic polyploid taxa of the *R. auricomus* complex (Table S1). From 1998 to 2018, we collected 1–72 individuals per population, 1–13 populations per taxon, and 230 populations in total across temperate and submeridional Europe (Fig. 1, Fig. S1a–i, Table S1). In situ, we recorded altitude, GPS coordinates, habitat types, and range of individual petal number for each population. Most living plants were sampled during flowering or early fruiting stage. Since apomictic vs. sexual development is determined before flowering, during early bud stage (Hojsgaard et al., 2014; Klatt et al., 2016; Nogler, 1984), sexual vs. apomictic reproduction reflected wild conditions. Plants were transferred to the Old Botanical Garden at the University of Göttingen, and cultivated in 1.5 L pots under controlled environmental conditions. Cross-pollination, flower isolation, and seed harvesting were described in Text S1. We added leaf material collected by F. G. Dunkel ("Du") to the genetic analysis increasing total population sample size to N = 251. For some populations (mainly "Du") without available living individuals, we took ploidy levels and mean values of sexuality from literature, or, in case of sexual populations, we calculated population mean values per species (see Table S2, and the reference list herein).

Flow Cytometry (FC) and Flow Cytometric Seed Screening (FCSS)

Flow cytometry and FCSS measurements and subsequent peak estimation and filtering followed Karbstein et al. (2020a) (see Fig. S2 and also Barke et al. 2018 for sexual vs. apomictic ovule and seed development, and the respective embryo to endosperm ploidy levels). To ascertain the ploidy level of each individual, we screened silica gel dried leaf material collected in spring 2017 to 2019 by FC analyses on 1–36 individuals per population and 234 populations in total (mean six individuals/population), yielding 1,474 FC measurements (Tables S1 and S2). We calculated the ratio between leaf peak index and the diploid standard. We found variation in ratios due to genome size variation between diploid species such as in Paule et al. (2018). The observed peak range variation also affects polyploid genome sizes because polyploids probably originated from hybridization events (Hodač et al., 2014; Hörandl et al., 2009) involving parents with different genome sizes. We also noticed a remarkable genome downsizing effect, similar to Hörandl and Greilhuber (2002), when comparing hexa- to tetraploids (Fig. S3). Therefore, individual ploidy level calculations were adjusted to mentioned effects.

To investigate the reproduction mode, we performed FCSS (after Matzk et al., 2000). We analysed up to 17 seeds per individual, 1–14 individuals per population (mean of five individuals per population, 1,075 individuals in total), yielding 11–64 seeds per population (mean 21 seeds/population, Table S2; see comparable minimum thresholds in Klatt et al. 2016 and Schinkel et al. 2016). High sample sizes in some populations were due to merging of apomictic subpopulations, previous measurements of sexual species, or assurance of facultative sexuality measurements. In total, we screened 221 populations and 4,669 seeds.

To ascertain sexual and/or apomictic reproduction modes, the respective ratios of endosperm ploidy and embryo ploidy were calculated as peak index (PI; Fig. S2). The sexual reproduction mode observed in this study involved di-, tetra-, and hexaploid embryos and the respective tri-, hexa-, and nonaploid endosperms formed after double fertilization, yielding PI values of around 1.5. We excluded embryo polyploidizations, haploid parthenogenesis (no observed BIII hybrids, Text S2), and defined an upper threshold of 1.7 for sexual seeds (see also Schinkel et al., 2016). Apomictic reproduction modes are indicated by a PI ≥ 1.9 (Figs. 2, S2, Text S2 for classifications and filtering, Fig. S4 for measurements). Reproduction modes were defined as follows: Sexual populations ≥85% (Karbstein et al., 2020b, Supporting Information), facultative apomicts <85% and >0%, and obligate apomicts 0% sexual seeds. Sample sizes and percentages of sexual seeds per population were listed in Table S2. Degree of sexuality (percent of sexual seeds, continuous variable) is in the following termed as “sexuality”.

Laboratory Work, RADseq, De Novo Assembly, and Parameter Optimization

We performed DNA extractions of 303 *R. auricomus* samples from ~1.5 cm^2 silica-dried leaf material using the Qiagen DNeasy Plant Mini Kit (Qiagen), following the manufacturer’s instructions. Adjustment of DNA concentration, DNA quality check, RADseq workflow (Baird et al., 2008), sequencing (Floragenex Inc.), raw read quality check (FASTQC; Andrews, 2010), raw read demultiplexing, removal of adapters and restriction overhang, and further quality filtering in IPYRAD (Eaton, 2014; Eaton & Overcast, 2020) are described in Karbstein et al. (2020a).

De novo assembly of RADseq loci and parameter optimization were performed with IPYRAD v.0 .9.14, and different output formats were created with IPYRAD v.0.9.52 on the local GWDG HPC Cluster (Göttingen, Germany). For parameter optimization, we applied the previously developed workflow optimizing within- and among-sample clustering threshold separately and accounting for different ploidy levels of sexual *R. auricomus* taxa (Karbstein et al., 2020a) (see

Text S3 for assembly details, assurance of orthologous locus assembly, and potential degree of chloroplast sequences in the final assembly, Figs. S5 and S6).

Data Analyses

All statistical analyses were performed in R v.4.0.1 (R Core Team, 2020). We calculated sexuality of populations as the ratio of sexual seeds to all seeds (Table S2). We selected the maximum number of petals per population (as petals easily fall off) for further analyses. To simplify later modelling procedure, we classified habitats into "forests" and "open habitats". We also ensured comparability of petals (2018 vs. 2019) and sexuality of facultative apomicts across data sets (in situ vs. garden; Text S1). In addition, geographical range area (polygon area) of each reproduction mode was calculated using the R package GEOSPHERE v.1.5–10 (Hijmans, Williams, et al., 2019).

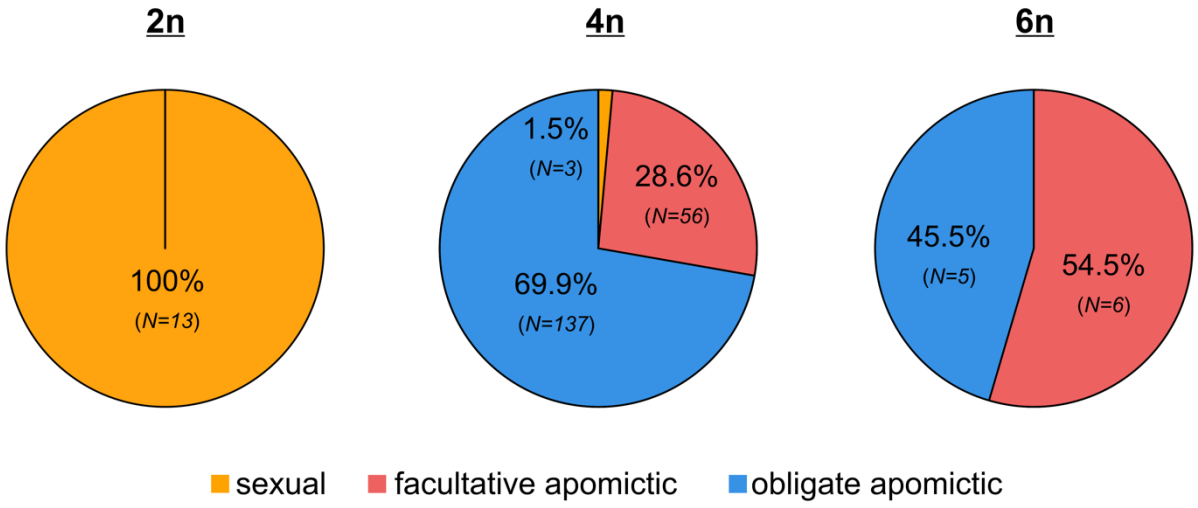


Figure 2. Pie charts illustrating percentages of sexual, and facultative and obligate apomictic reproduction modes per ploidy level of 220 *R. auricomus* populations (only own measurements). N = number of populations per group.

To determine genome-wide heterozygosity of individuals, we used the final min50 RADseq assembly (balancing missing data and number of loci, see also Karbstein et al. 2020a, Fig. S3). IPYRAD analysis yielded 14,489 filtered RADseq loci (only 21 without genetic variation) with 43% missing data and in mean 8,353 loci/individual. We transformed information about

reference and alternative alleles of the *.vcf file (diploid SNP calls, four allowed alleles per locus). Homozygous sites were assigned a value of 0, and heterozygous sites a value of 1. Then, we calculated the ratio of heterozygous sites to all sites per individual (Table S3). In apomictic complexes, the high intraindividual diversity assessed by genomic markers is informative even in the presence of a single individual per population (e.g., Lovell et al., 2014; see Text S4 including further statistical tests). We excluded populations 10, 18, 42, EH10334, LH001, and LH002 (slightly different laboratory workflow). The final maximal sample size was N = 280 individuals (N = 220 populations).

To investigate genetic structure, we carried out sNMF analyses provided by the R package LEA v.3.0.0 (Frichot et al., 2014, Frichot & François, 2015, 2020) and based on the min30 alignment (see Karbstein et al., 2020b for calculations of genetic distance and heterozygosity, and comparability). We used the *.ugeno output file (1 SNP/locus; 33,165 loci) as input and set the number of genetic clusters (K) from 1 to 80, ploidy to four, and repetitions to seven. The cross-entropy criterion was employed to choose the likeliest K. To detect phylogenetic relationships and reticulations, we conducted NeighborNet analyses based on the min30 alignment using SPLITSTREE v.4.14.6 (Huson & Bryant, 2006). We used the *.u.snps file (1 SNP/locus, converted to *.nex file) to estimate genetic distances with a general time-reversible (GTR) model, estimated site frequencies, and maximum likelihood (equal rates of site variation, default rate matrix).

sNMF analyses revealed three genetic partitions with moderately to highly admixed individuals falling into three genetic clusters according to their predominant partition (I–III, Fig. 3). The network analysis (Fig. S7) indicated reticulate relationships and polyphyly of described agamospecies. Standard phylogenetic generalized linear mixed models (GLMMs) or phylogenetic path analysis require well-resolved phylogenetic trees (see Lefcheck, 2015; Pegoraro et al., 2020) and are not appropriate here due to the lack of a tree-like, hierarchical phylogenetic structure (see also Hörandl et al., 2009; Hodač et al., 2014, 2018) and artificial circumscriptions of agamospecies.

Instead, we performed a genetically-informed structural equation model (SEM, confirmatory path analysis) based on GLMMs with main genetic sNMF clusters as random effect to account for potential genetic dependence of model variables (data set without missing values, N = 205 populations). We specified the following models/paths: (i) Sexuality explained by ploidy levels (without 3n due to low sample size), habitat types, and climatic environmental factors, (ii) maximum number of petals explained by sexuality, (iii) genome-wide heterozygosity explained

by sexuality, ploidy levels, reproduction modes, habitat types, and climatic environmental factors. Although sexuality and reproduction mode are nonindependent due to calculation procedure, we included both in the SEM (iii). The categorical variable reproduction mode probably better detects scarce differences between facultative and obligate apomicts in relation to heterozygosity. GLMMs were consequently specified: Binary (sexuality, proportional data), poisson (petals, count data), and gaussian (heterozygosity, proportional data but log-transformed to normal distribution due to nonbinary range between 0.003 and 0.03). We only allowed "sexuality - reproduction mode" as independence claim (correlation thus included in SEM calculations). The other 11 unspecified relationships were classified as correlated errors (not causal and unidirectional; e.g., petals-heterozygosity, $p < .05$).

To investigate IBD, we computed pairwise geographic distances among sampling locations using the Vincenty method implemented in the R package GEOSPHERE v.1.5–10 (Hijmans, Williams, et al., 2019). We used the pairwise genetic distance matrix of SPLITSTREE (see above), and obtained pairwise ecological distances among study locations by Euclidean distances based on non-autocorrelated standardized climatic environmental factors (see below) and by applying the R package ECODIST v.2.0.7 (Goslee & Urban, 2020). Mantel tests were performed between genetic or ecological and geographical distances (isolation-by-distance [IBD]/environment [IBE]) for different data sets using the R package APE v.5.3 (Paradis et al., 2020). We plotted data points, and drew regression lines based on linear models.

To investigate relationships between sexuality or heterozygosity to climatic environmental factors across Europe, we downloaded bioclimatic variables, altitude, and solar radiation ($kJ*m^{-2}*day^{-1}$) in 2.5 min resolution from WorldClim database v.2 (Fick & Hijmans, 2017) using the R package RASTER v.3.0–7 (Hijmans et al., 2020). GPS coordinates of study locations were imported into R, and environmental data were extracted location-wise using the R package RASTER. We standardized environmental variables, and removed autocorrelated variables ($r_{Sp} > 0.8$) since they can disturb modelling procedure (Dormann et al., 2013). We subsequently performed GLMMs between sexuality or heterozygosity and selected climatic environmental factors including specific interaction terms. Significant predictors were mostly in accordance with presented literature herein (potentially causal), and thus included in path analysis to get information about their relative significance in comparison to intrinsic factors.

To run the SEM, we applied the function psem within the R package PIECEWISESEM v.2.2.0 (Lefcheck, 2015, 2021), and employed the function glmmPQL within the R package MASS v.7.3–51.6 (Venables & Ripley, 2002; Ripley et al., 2020) for GLMMs. In the latest version,

standardized regression estimates and consequently effect sizes and directions were not simply accessible in the presence of non-normal distributed responses (binomial, poisson, and gaussian), and both numerical and categorical predictors (see also documentation Lefcheck, 2021). Therefore, nonstandardized estimates are hardly comparable (e.g., particularly when comparing the relationship between sexuality and isothermality in pre-GLMMs and in path analysis including intrinsic factors). We obtained effect direction (Fig. 4, Table 1) of numerical climatic predictors from pre-GLMMs and of categorical predictors from Wilcoxon (two groups) or Kruskal Wallis (three groups) tests accompanied by posthoc pairwise Wilcoxon rank-sum tests (with "fdr" correction). We used nonparametric tests because sexuality and heterozygosity values were non-normally distributed even after data transformation (arcsin) within-group levels and/or variances not homogenous among group levels. Categorical interaction terms were excluded in GLMMs of the SEM due to collinearities. We examined the influence of relevant predictor variable interactions on response variables by coding them as a single predictor variable and applying the same nonparametric tests as described above.

We downloaded satellite maps, created bubble plots where circle colour represents sexuality or heterozygosity values of a specific location, and added significant climatic environmental factors of the path analysis. We used the R packages DPLYR v.0.8.3 (Wickham et al., 2019), GGPLOT2 v.3.2.1 (Winston et al., 2019), GGMAP v.3.0.0 (Kahle et al., 2019), MAPS v.3.3.0 (Becker et al., 2018), and VIRIDIS v.0.5.1 (Garnier et al., 2018) for data visualization.

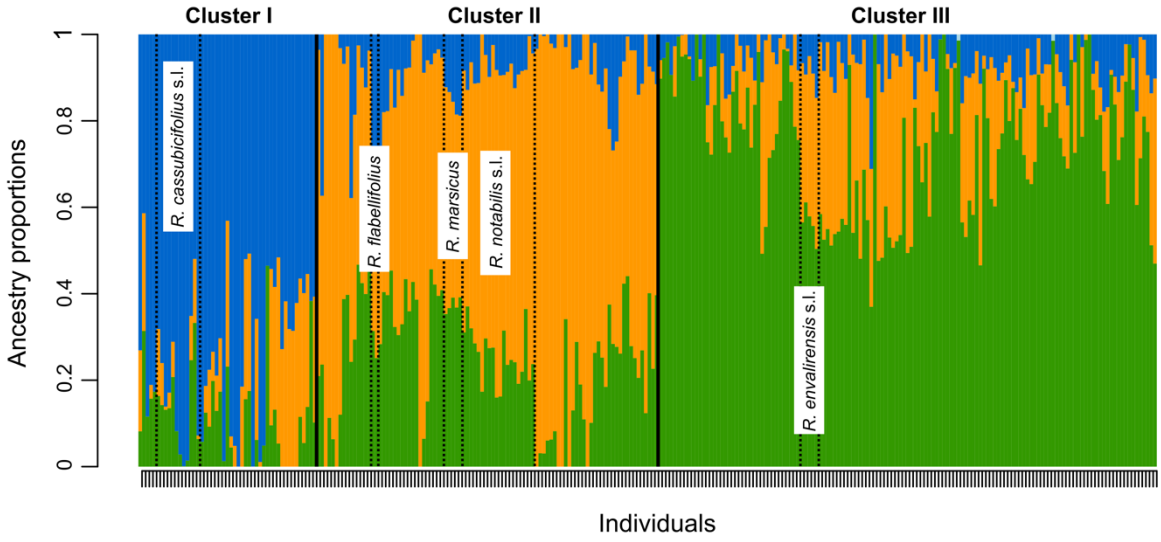


Figure 3. Bar graph of individual ancestry proportions using the likeliest K (K = 3 genetic partitions, shown in colours; see Fig. S10, Figure with individual labels on Figshare) based on a sNMF analysis of 280 individuals and the min30 RADseq alignment (1 SNP/locus; 33,165 loci). We highlighted sexual species by black dashed lines and separated genetic clusters (I–III) by black solid lines.

Results

Cytotpyes and Reproduction Mode

FC analyses revealed four main ploidy levels: Diploids (N = 13), triploids (N = 1), tetraploids (N = 197), and hexaploids (N = 12; Table S2). We also observed few populations of mixed ploidy levels: 3n+4n, 4n+5n, and 4n+6n. Diploids, triploids, and hexaploids regionally occur in Western to Eastern and Southern Europe, whereas tetraploids additionally occupy habitats in temperate Northern Europe (Fig. 1).

FCSS analyses unveiled different proportions of sexual/asexual seeds across individuals of different ploidy levels: Diploids 99.0%/1.0%, triploids 0.0%/100%, tetraploids 4.1%/95.9%, and hexaploids 6.3%/93.7% (Fig. S8). In total, we investigated 16 sexual (mean sexuality 98.3%), 62 facultative apomictic (mean sexuality 7.1%), and 143 obligate apomictic populations (Fig. S9). We detected only sexual populations at the diploid level (N = 13), and, at the tetra-/hexaploid level, we observed 1.5% (N = 3)/ 0% sexual, 28.6% (N = 56)/54.5% (N = 6) facultative apomictic, and 69.9% (N = 137)/45.5% (N = 5) obligate apomictic populations (Fig. 2; Table S2). We observed almost no population composed of sexual and apomictic individuals (except KK166), and almost no taxa of mixed sexual and apomictic populations

(except tetraploid *R. marsicus*). In addition, sexuals occupy a smaller area than apomicts ($1.23e^6$ km^2 and $3.66e^6$ km^2, respectively). Size of range area was comparable between facultative and obligate apomicts ($2.84e^6$ km^2 and $3.10e^6$ km^2, respectively; Fig. 1b).

Phylogenetic Relationships and Geographic Differentiation

The cross-entropy analysis based on sNMF results revealed three genetic partitions as the likeliest structure among *R. auricomus* individuals (Fig. S10). The individuals showed various degrees of admixture and fell according to their dominant genetic partition into three main genetic clusters (Fig. 3a): Cluster "I" contains the sexual species *R. cassubicifolius* s.l. and mainly tetra-to hexaploid apomicts, cluster "II" comprises the diploid sexual species *R. flabellifolius* and *R. notabilis* s.l., and the tetraploid sexual *R. marsicus*, and mainly tetraploid apomicts, and cluster "III" contains the diploid sexual species *R. envalirensis* s.l. and mainly tetraploid apomicts.

Results of the NeighborNet analysis indicated overall low genetic differentiation and reticulate relationships among main groups (Fig. S7). The three main genetic clusters (I–III) corresponded mostly to those of the sMNF analyses, with some incongruences between main clusters II and III. Accessions of apomictic taxa were widely scattered, also among main genetic groups, whereas accessions of sexual species clustered together.

IBD between genetic and geographic distances was significant for each reproduction mode, with strongest IBD for sexuals ($Z = 1.53e^4$, $p < .001$; $r_P = 0.43$, $R^2 = 0.18$), followed by obligate apomicts ($Z = 2.60e^5$, $p < .001$; $r_P = 0.28$, $R^2 = 0.08$), facultative apomicts ($Z = 4.01e^4$, $p < .01$; $r_P = 0.26$, $R^2 = 0.07$), and the total data set ($Z = 8.19e^5$, $p < .001$; $r_P = 0.21$, $R^2 = 0.04$; Fig. S11a–d). IBE between ecological and geographic distances was also significant for each reproduction mode, and was strongest for obligate apomicts ($Z = 4.71e^7$, $p < .001$; $r_P = 0.83$, $R^2 = 0.69$), followed by sexuals ($Z = 2.72e^6$, $p < .001$; $r_P = 0.73$, $R^2 = 0.53$), total data set ($Z = 1.57e^8$, $p < .001$; $r_P = 0.71$, $R^2 = 0.50$), and facultative apomicts ($Z = 7.35e^6$, $p < .001$; $r_P = 0.65$, $R^2 = 0.42$; Figs. S11e–h).

Path Analyses

The genetically-informed SEM represented the data very well (Fisher's $C = 51.23$, $p < .001$). We assessed widely significant, moderately to highly explained relationships between the

response variables sexuality (R^2_M = R_2^C = 0.58), petals (R^2_M = 0.15, R^2_C = 0.18), and heterozygosity (R^2_M = 0.43, R_2^C = 0.44) and their respective predictors.

Sexuality

In detail, we investigated significant relationships between sexuality and ploidy levels (2n: $p < .05$, 4n: $p < .01$, 6n: $p < .05$; Table 1, Fig. 4). Tetra- and hexaploid populations (means 3.9% and 5.7%, respectively) showed less sexuality compared with diploid populations (mean 99.0%; all $p < .001$; $Chi^2 = 95.2$, $df = 2$, $p < .001$; Fig. 5a). Hexaploids exhibited slightly higher sexuality than tetraploids ($p = .08$). Sexuality was marginally significantly related to habitat types (forest: $p = .06$, open habitats: $p = .07$; Table 1, Fig. 4), but forests did not differ from open habitats in post hoc comparisons (means 14.9% vs. 15.6%; $W = 6736$, $p = .75$; Fig. 5b).

With regard to the interaction between ploidy level and habitat type, we observed no differences between forests and open habitats at the diploid (means 99.2% and 98.7%, respectively; $p = .55$), tetraploid (means 2.7% and 5.4%, respectively; $p = .48$), and hexaploid level (means 4.1% and 7.1%, respectively; $p = .68$; $Chi^2 = 96.7$, $df = 5$, $p < .001$, Fig. S12). The maximum number of petals was significantly positively related to sexuality ($p < .001$, Table 1, Fig. 4 and Fig. S13). The pre-GLMM revealed an increase of sexuality with annual precipitation, temperature seasonality, and solar radiation (and their interactions), and a decrease with altitude (Table S4). In path analysis, only solar radiation ($p < .001$) and isothermality ($p < .01$) showed a significant effect on sexuality (Figs. 4 and 6, Figs. S14–16, Table S5).

Table 1. Genetically-informed SEM results based on GLMMs and a data set of 205 populations (see also Fig. 4). Sexuality (percent of sexual seeds), maximum number of petals, and heterozygosity (percent of heterozygous sites) of populations were included as response variables, and sexuality, ploidy levels (without triploids), reproduction modes, habitat types, and climatic environmental factors as predictor variables. Main genetic clusters found in sNMF analysis were used as random effect variable in GLMMs to control for potential phylogenetic dependence of included variables. For effect directions of climatic numerical and categorical predictors, estimates of pre-GLMMs (Tables S4 and S6) and mean values and significances (superscript letters) for within-group levels are given (separate relationships; see also Fig. S13 and legend). Next to it, nonstandardized estimates of the SEM are given (estimates of categorical predictors are model-estimated means). See Materials and Methods section for further details, and Figs. 5 and 7, S12–S22. Df, degrees of freedom; *4n–6n comparison marginal significant (p = .08). * p < .05, **p < .01, & *** p < .001.

response	predictor	level	mean / estimate (pre-GLMM)	est. mean / estimate (SEM)	std. error	df	critical value	p value
sexuality	**ploidy level**	2n	99.0$^{(a)}$	4.01	0.76	2	5.28	*
		4n	3.86$^{(c)}$	-3.62	0.19	2	-19.23	**
		6n	5.70$^{(b*)}$	-5.03	0.82	2	-6.13	*
sexuality	**habitat type**	forest	14.9$^{(a)}$	-1.51	0.38	2	-4.00	0.06
		open habitats	15.6$^{(a)}$	-1.58	0.45	2	-3.50	0.07
sexuality	**isothermality**		1.07	-1.06	0.36	192	-2.94	**
sexuality	**solar radiation**		0.87	1.24	0.33	192	3.80	***
petals	**sexuality**		-	0.70	0.11	201	6.59	***
heterozygosity	**sexuality**		-	-0.18	0.67	191	-0.27	0.79
heterozygosity	**ploidy level**	2n	0.61$^{(c)}$	9.75	0.29	2	33.69	***
		4n	1.28$^{(b)}$	10.10	0.16	2	63.40	***
		6n	1.72$^{(a)}$	10.26	0.21	2	49.13	***
heterozygosity	**reproduction mode**	sexual	0.64$^{(b)}$	9.88	0.57	2	17.28	**
		fac. apomictic	1.28$^{(a)}$	10.13	0.13	2	80.22	***
		obl. apomictic	1.36$^{(a)}$	10.10	0.14	2	69.96	***
heterozygosity	**habitat type**	forests	1.27$^{(a)}$	10.09	0.15	2	67.07	***
		open habitats	1.13$^{(b)}$	9.98	0.15	2	65.17	***
heterozygosity	**temperature seasonality**		0.01	0.08	0.03	191	2.56	*

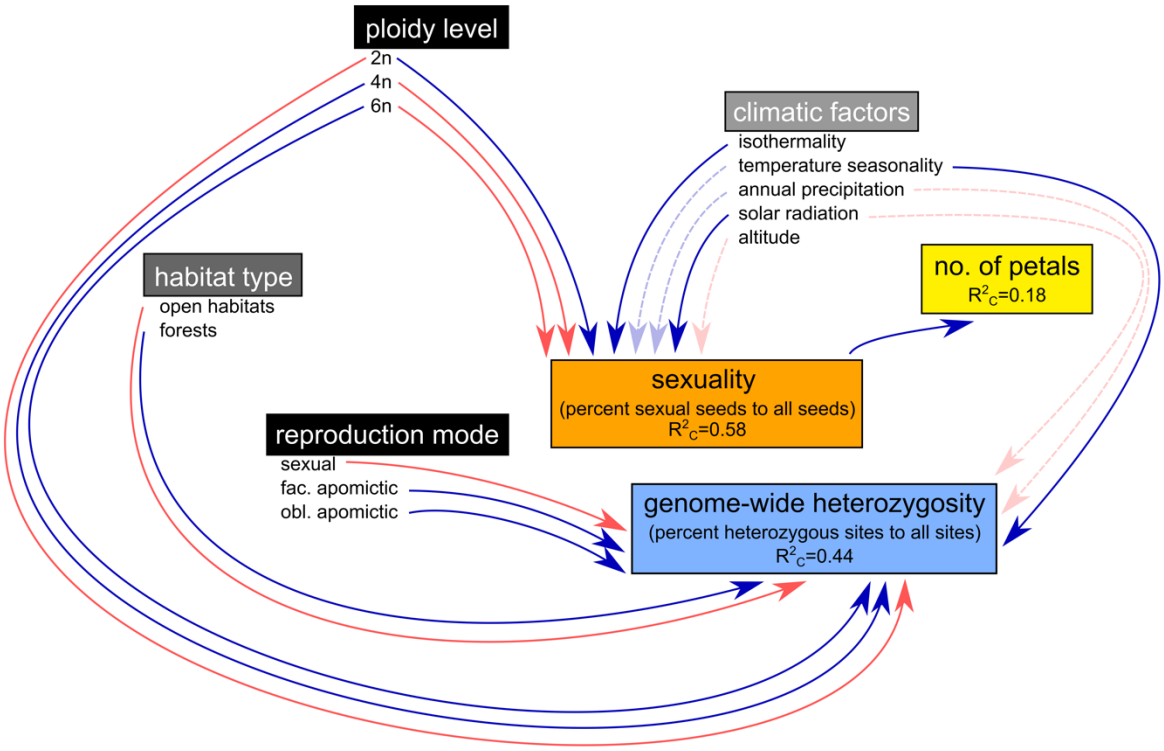


Figure 4. Path diagram based on results of the genetically-informed SEM and 205 populations, and further statistical analyses for effect direction (Figs. 5, 7, Table 1). Arrows indicate relationships among predictor and response variables (blue, positive, red, negative, solid, significant; dashed, climatic environmental significant in pre-GLMMs, Tables S4 and S6). Absent arrows indicate nonsignificant relationships, for example heterozygosity and sexuality ($p = .79$). See Materials and Methods and Results for more details. R^2_C, conditional R^2.

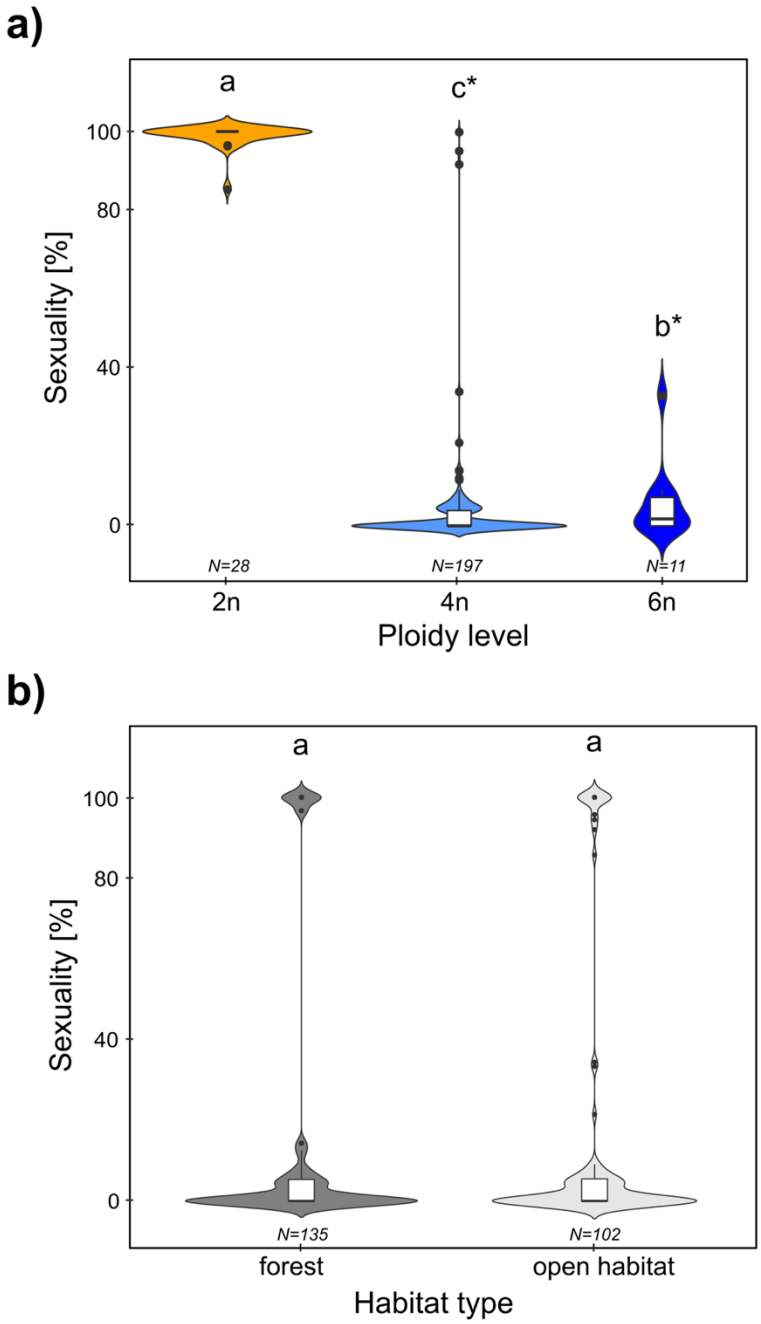


Figure 5. Violin plots showing sexuality of populations (percent of sexual seeds to all seeds) in relation to different (a) ploidy levels (without triploids) and (b) habitat types. Letters above violin plots indicate significant/nonsignificant differences between/among groups. *4n–6n comparison marginal significant (p = .08). N = number of populations.

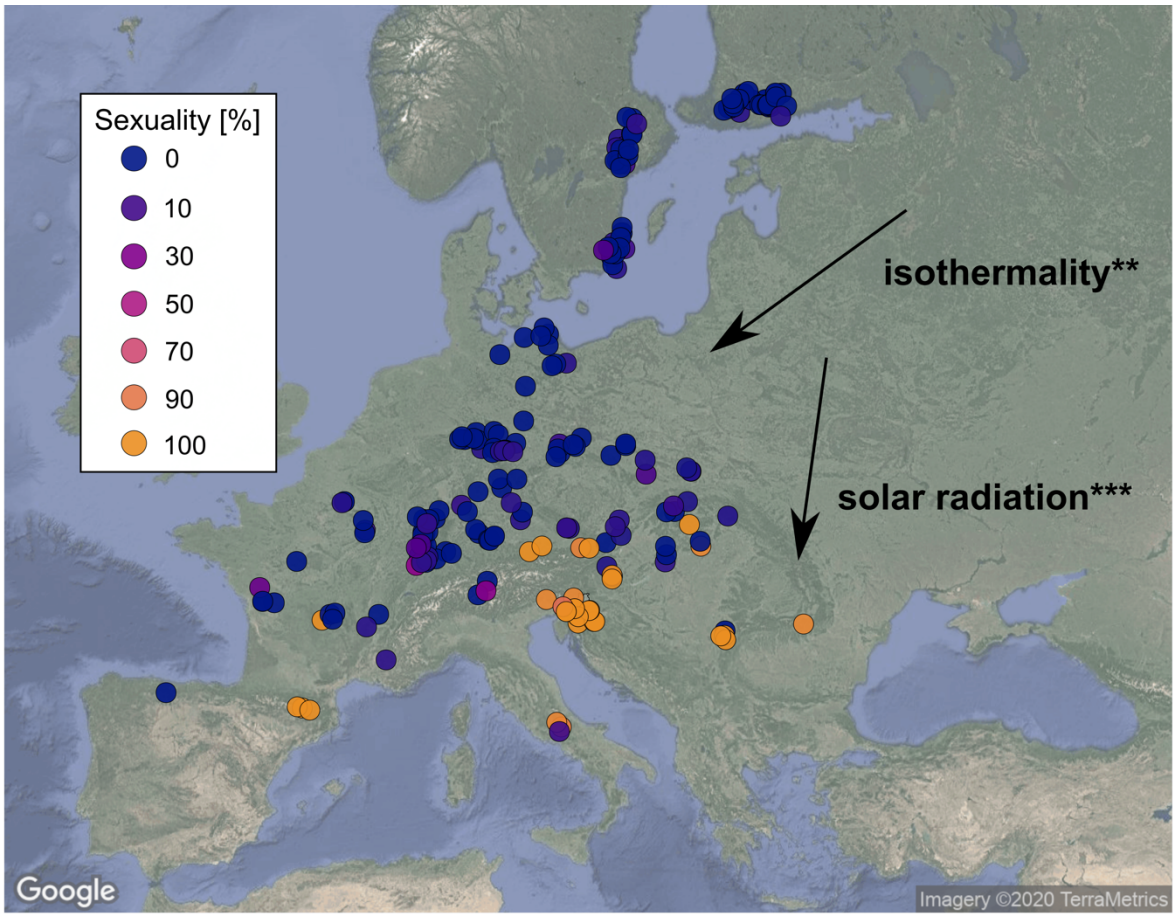


Figure 6. Geographic map illustrating sexuality of populations (percent of sexual seeds to all seeds) across Europe. Circle colour represents the value range of sexuality (see Legend, left of the centre). Climatic environmental factors significantly related to sexuality were drawn (retrieved from path analysis; *** $p < .001$, **$p < .01$). The arrow direction represents the main environmental gradient and was coarsely derived from Fig. S16. Solar radiation is given in $kJ*m^{-2}*day^{-1}$ and isothermality is the ratio of mean diurnal range to temperature annual range *100 (see Figs. S14–16). Image source: Google Imagery ©2020 TerraMetrics (retrieved on 14 September, 2020).

Heterozygosity

Genome-wide heterozygosity was significantly related to ploidy levels (all $p < .001$), habitat types (all $p < .001$), and reproduction modes (all $p < .01$), but not to sexuality ($p = .79$; Table 1, Fig. 4). Diploid cytotypes (mean 0.61%) were characterized by less heterozygosity compared with tetra- and hexaploid cytotypes exhibiting up to three-times higher heterozygosity (means 1.28% and 1.72%, respectively; all $p < .01$, respectively; $Chi^2 = 64.7$, $df = 2$, $p < .001$; Fig. 7a). Hexaploids possessed more heterozygosity than tetraploids ($p < .05$). We detected lower heterozygosity of sexuals (mean 0.64%) in comparison with facultative and obligate apomicts

(means 1.28% and 1.36%, respectively; all $p < .001$, respectively; $Chi^2 = 59.4$, $df = 2$, $p < .001$; Fig. 7b). Heterozygosity did not differ between apomictic reproduction modes ($p = .42$). Concerning the interaction ploidy level and reproduction mode, we detected no differences between sexuals, and facultative and obligate apomicts at the tetraploid (1.16%, 1.24%, and 1.30%, respectively; all $p > .05$), and hexaploid level (no sexual, 1.28%, and 2.00%, respectively; $p = .29$; $Chi^2 = 67.4$, $df = 5$, $p < .001$; Fig. S17). In addition, hexaploid facultative apomicts differed from facultative and obligate tetraploid apomicts (all $p < .05$). Populations in forest habitats were characterized by higher heterozygosity compared with populations in open habitats (means 1.27% and 1.13%, respectively; $W = 7451$, $p < .05$; Fig. 7c). Concerning the interaction ploidy level and habitat type, we detected no differences between forest and open habitats at the diploid (0.62% and 0.60%, respectively; $p = .69$), but at the tetraploid (1.33% and 1.20%, respectively; $p < .05$) and hexaploid level (2.11% and 1.12%, respectively; $p < .05$; $Chi^2 = 75.8$, $df = 5$, $p < .001$; Fig. S18). Besides, we observed the highest heterozygosity in hexaploid populations of forest habitats (all $p < .05$). Concerning the interaction habitat type and reproduction mode, we detected no differences between forest and open habitats for the sexual (0.60%, and 0.71%, respectively; $p = .35$), but for facultative (1.48% and 1.18%, respectively; $p = .05$) and obligate apomictic reproduction modes (1.35%, and 1.22%, respectively; $p < .05$; $Chi^2 = 67.1$, $df = 5$, $p < .001$; Fig. S19). The pre-GLMM revealed a significant increase in heterozygosity with interactions of solar radiation and annual precipitation or temperature seasonality (Table S6). In path analysis, we found only a significant increase of heterozygosity with temperature seasonality across Europe ($p < .05$; Figs. 4 and 8, Figs. S20 and S21, Table S5).

Discussion

We provide the first empirical evidence for associations of sexuality and genome-wide heterozygosity to intrinsic and extrinsic factors, which explain the GP pattern of the *R. auricomus* complex in temperate Europe. Ploidy level and environmental differences influenced sexuality, whereas heterozygosity was shaped by a combination of ploidy level, reproduction mode, habitat type, and environmental differences. Results also indicated that observed associations were independent from genetic structure (similar R^2_M and R^2_C). The here presented innovative approach allowed us to study the potential and complex relationships leading to GP.

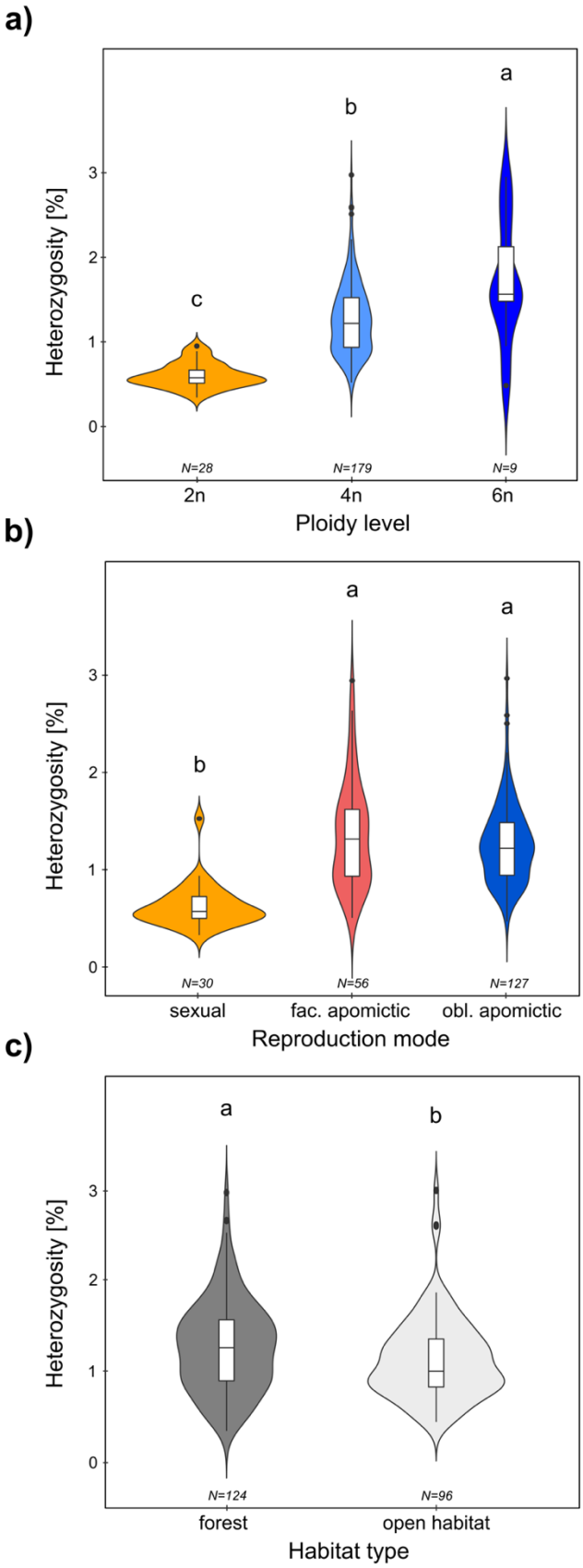


Figure 7. Violin plots showing population-wise heterozygosity (percent of heterozygous sites to all sites) in relation to different (a) ploidy levels (without triploids), (b) reproduction modes, and (c) habitat types. Letters above violin plots indicate significant/ nonsignificant differences between/among groups. N = number of populations.

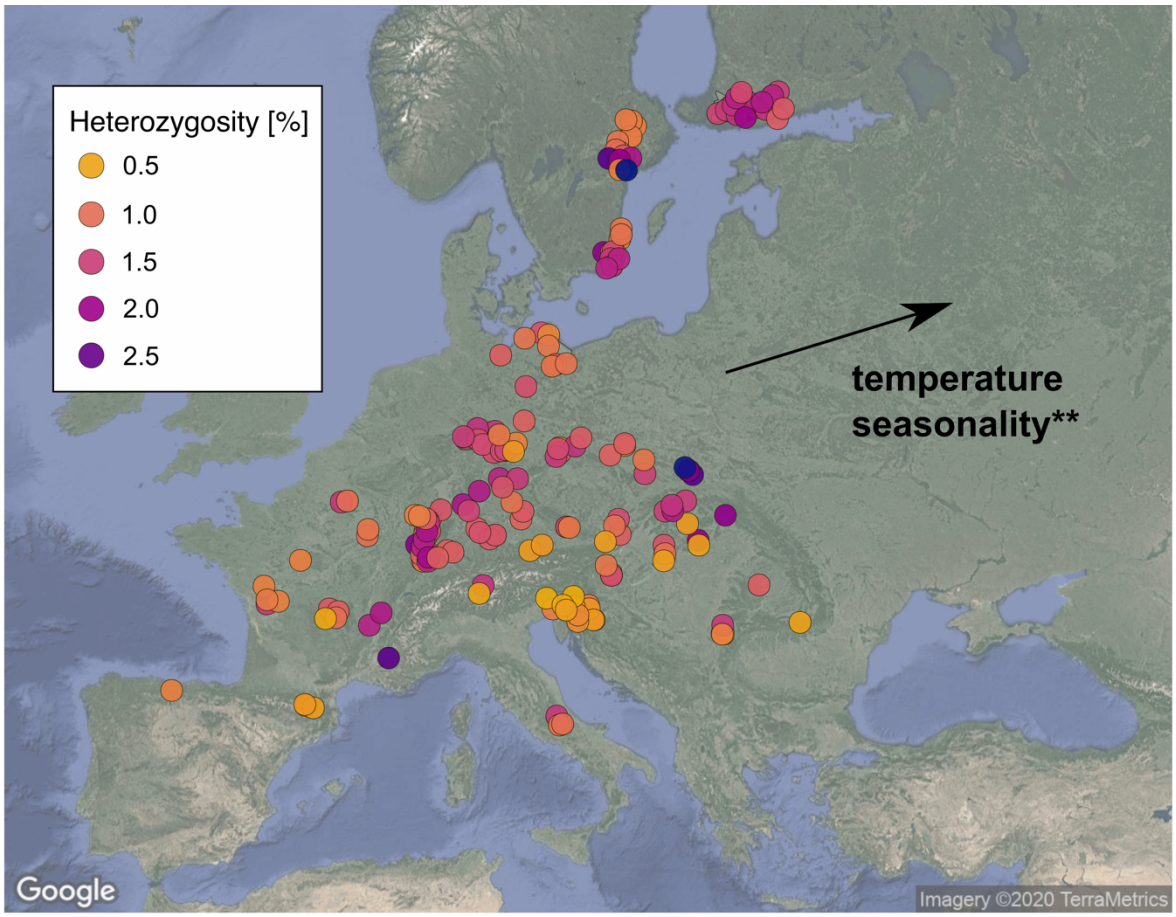


Figure 8. Geographic map illustrating population-wise heterozygosity (percent of heterozygous sites to all sites) across Europe. Circle colour represents the value range of heterozygosity (see Legend, left of the centre). Temperature seasonality (standard deviation*100) was significantly related to heterozygosity, and illustrated on the map (retrieved from path analysis; **$p < .01$; see Figs. S20 and S21). The arrow direction represents the main environmental gradient and was coarsely derived from Fig. S21. Image source: Google Imagery ©2020 TerraMetrics (retrieved on 14 September, 2020).

Peculiarities of the GP Pattern

We revealed a latitudinal GP pattern within the *R. auricomus* complex in the temperate zone of Europe. Apomictic populations were characterized by a larger and a more northern distribution range compared with sexual populations that inhabit a smaller range in more southern regions of Europe (non-centred distribution of sexuals; Fig. 1). Results are in line with other studies demonstrating larger distributions of apomicts towards higher latitudes than sexuals (e.g., *Crataegus*, Lo et al., 2013; *Hieracium alpinum*, Mráz et al., 2009; or *Potentilla crantzii*, Paule et al., 2015). We found a weak negative association between sexuality and altitude. Sexual species and facultative apomicts mainly inhabit lowlands, and only a few of them inhabit mid

to high altitudes of Central and Southern Europe mountain systems (Figs. 1 and 6). However, several obligate apomicts also occupy lower altitudes. Therefore, the classical GP pattern that apomicts tend to occupy higher altitudes than sexuals (Bierzychudek, 1985; Cosendai et al., 2013; Gregor, 2013; Schinkel et al., 2016) is only partly supported by our data.

The *R. auricomus* complex is structured into three main genetic clusters, with high levels of admixture in most individuals. The sexual species appear genetically coherent and distinct from each other as in previous phylogenomic studies (Karbstein et al., 2020a; Tomasello et al., 2020). Relationships of apomicts, however, were highly reticulate, and described morphospecies appeared mostly as polyphyletic, probably due to multiple allopolyploid origins (see Hörandl et al., 2009; Hodač et al., 2014, 2018; Karbstein et al., in prep.). We found no identical genotypes in apomictic polyploids, which does not fit predictions of clonality according to the classical GPG and FNV models (similar as in Cosendai et al., 2013).

Moreover, we detected IBD in sexuals and apomicts. Stronger IBD among sexual species can be attributed to geographical isolation in glacial refugia and to vicariance processes that triggered allopatric speciation (Tomasello et al., 2020). Apomicts are younger and probably rapidly colonized the temperate zone with many genotypes. Multiple colonization events and facultative sexuality allowing for gene flow probably reduced IBD (Cosendai et al., 2013; Lovell et al., 2014; Paun, Greilhuber, et al., 2006). Due to Late Pleistocene origins, the genetic divergence of apomictic lineages is generally low confirmed by low genetic distance values.

Sexuality in Relation to Ploidy Levels

Sexual *R. auricomus* populations were mainly diploid (except for tetraploid *R. cassubicifolius* and *R. marsicus* cytotypes). Apomicts were predominantly tetraploid (ca. 90%), followed by some hexaploid and a few penta- and triploid populations. The restriction of sexuality to lower, mainly di- and tetraploid levels is a commonly observed phenomenon in polyploid plant complexes (Hörandl, 2006; Lo et al., 2013; Mráz et al., 2009). The occurrence of apomictic polyploids in previously glaciated areas of the Alps and Scandinavia, and sexuals in Central and Southern Europe (Figs. 1 and 6) is also in line with previous studies (Bierzychudek, 1985; Hörandl, 2006; Hörandl et al., 2008). Natural diploid *R. auricomus* populations exhibited rare apomictic seed production, which occurred also in diploid F_1 *R. auricomus* crossings (Barke et al., 2018), and in other natural diploids (e.g., *Boechera holboellii*, Kantama et al., 2007; *Paspalum*, Ortiz et al., 2013 or *R. kuepferi*, Schinkel et al., 2016). Tetraploid *R. auricomus* populations reproduced to two-third obligate and one-third facultative-apomictically, whereas

hexaploid ones showed similar percentages of both reproduction modes (Fig. 2). Hexaploids also exhibited slightly higher sexuality compared with tetraploids (Fig. 5a). Results indicate ploidy-related differences, but they also reject the classical assumption of increasing obligate apomixis in higher ploidy levels (Grant, 1981). Gene dosage effects of the wild type alleles from sexual progenitors vs. apospory-controlling genetic factors are potentially relevant for the degree of apomixis (see Nogler, 2006). We only observed one strict asexual triploid population. Formation of triploids is rare because backcrossing of tetraploid apomictic to diploid sexual cytotypes is hampered by strong reproductive isolation via different ploidy levels, flowering time, or local habitat differentiation (Hörandl & Paun, 2007). Previous assumptions of predominant facultative sexuality of apomicts within the *R. auricomus* complex (e.g., Paun et al., 2006, Hörandl et al., 2009) are not confirmed by our results showing only 30% facultative apomictic populations with sexuality values <15% (outliers up to 34%, Fig. S9). However, experimental work on other perennial apomicts indicated that proportions of sexual seed formation of the same plant can vary between different years (Klatt et al., 2018), and also according to different pollination times (Espinoza, 2002). Hence, the mode of reproduction in wild populations might be more variable over larger timescales.

Sexuality in Relation to Habitat and Climatic Conditions

Forests and open habitats were similarly occupied by sexuals and apomicts (Fig. 5b; also within ploidy levels, Fig. S12). Our results are in contrast to previous research suggesting regional niche shifts of apomicts towards open habitats (Asker & Jerling, 1992; Hörandl & Paun, 2007). Our coarse classification could neglect habitat differences, although a more fine-scale classification also showed no significant differences (Fig. S22). Sexual *R. auricomus* species probably have evolved out of a common forest understory ancestor ca. 700,000 years ago (Tomasello et al., 2020), but nowadays some sexuals (except for *R. cassubicifolius* s.l. and *R. flabellifolius*) also inhabit anthropogenic and subalpine meadows, forest edges, and ditches. Obviously, some sexuals possessed enough genetic variation for selection processes to adapt to more open habitats during past climatic oscillations.

In contrast to habitat comparisons, we obtained significant relationships between sexuality and climatic conditions in the path analysis: Sexuality increased with solar radiation and isothermality across Europe (Fig. 6, Figs. S14–S16). Sexual and facultative apomictic populations were mainly found towards southern, solar radiation-rich habitats (Figs. S15, S16b), and they rather prefer isothermal habitats with a temperate climate in Central and

Southern Europe (Figs. S14, S16a). Isothermality describes the diurnal temperature range (monthly mean) in relation to the annual temperature range, and is mainly associated with seasonal variation or continentality (see Fig. S16a, and Rushworth et al., 2018). In the pre-GLMM, sexuality additionally increased with annual precipitation and temperature seasonality. However, in relation to other climatic and intrinsic factors, these factors were less important (Fig. 4).

The relationship between sexuality and solar radiation is surprising because previous GP studies in plants mainly revealed associations of apomicts to low temperature, temperature variation, high precipitation, precipitation seasonality, and high altitude (Alonso-Marcos et al., 2019; Coughlan et al., 2017; Kirchheimer et al., 2016; Klatt et al., 2018; Schinkel et al., 2016). High solar radiation besides low temperature variation and sufficient precipitation may promote sexuality in *R. auricomus* populations. Climate chamber experiments involving di-, tetra-, and hexaploid *R. auricomus* individuals showed that a prolonged photoperiod enhances sexual megaspore formation, with strongest effects in diploids, but did not affect sexual seed set (Klatt et al., 2016; Ulum et al., 2020). Effects of light intensity have not been tested experimentally so far. Under natural environmental conditions, solely solar radiation, which increases towards the south (Körner, 2003), is the main driver of sexuality in facultative apomicts, probably triggering sexual megaspore formation and seed development. Our results suggest that light conditions are an important, hitherto neglected factor for GP in plants. The negative effect of high temperature fluctuations within a year on sexuality (and less importance of daily temperature fluctuations), in general, is a less-observed aspect on plant sexual reproduction (see Zinn et al., 2010; Coughlan et al., 2017; Rushworth et al., 2018).

Apomixis in *R. auricomus* populations may be mainly promoted by high temperature fluctuations within a year and low light conditions as well as low levels of precipitation. In North American *Crataegus*, Coughlan et al. (2017) reported climatic niche shifts of polyploid apomictic species towards higher daily temperature fluctuation, more extreme cold or dry, high-temperature seasonality or -precipitation seasonality conditions compared to diploid sexuals. Rushworth et al. (2018) investigated that nonhybrid apomictic *Boechera* lineages settled in less isothermal habitats, supporting the findings presented herein. Lower temperatures down to freezing had triggered apomictic reproduction in alpine *R. kuepferi* (Klatt et al., 2018), but *R. auricomus* taxa mainly occupy temperate zones and thus might be more susceptible to freezing than *R. kuepferi* (Daubert, 2016). Under low-temperature and low-light conditions, uniparental reproduction via apomixis might ensure seed production and survival of *R. auricomus* in higher latitudes. Additionally, the slight increase of apomixis towards precipitation-poor habitats is

supported by Coughlan et al. (2017) for North American *Crataegus* apomicts, but not by Schinkel et al. (2016) for alpine, apomictic *R. kuepferi* populations. Different intra- and interspecific niche effects might play a pronounced role in this aspect. We found significant IBE regarding ecological and geographical distances (Fig. S11e–h). This supports the here found association between sexuality and certain bioclimatic factors across Europe. Our results also underline the findings of Paule et al. (2018; see also Rice et al., 2019) that *R. auricomus* polyploids occupy slightly colder and drier climates compared to diploids, potentially allowing them to spread more northwards.

Petal Formation in Relation to Sexuality

Sexual *R. auricomus* populations were characterized by more than 4-5 petals and apomictic ones by a variable number of petals (0–15), supporting previous investigations (Borchers-Kolb, 1983, 1985; Dunkel et al., 2018; Dunkel, 2014). Pollen transfer is probably mainly mediated via *Diptera*, *Hymenoptera*, and Coleoptera species (Steinbach & Gottsberger, 1994 for apomicts; K. Karbstein and E. Hörandl, personal observation). Self-incompatible sexuals with more petals may better attract pollinators, ensuring outcrossing. Highly facultative apomicts tend to possess more petals. However, some facultative-apomictic populations showed low levels of sexuality (<5%) and exhibited up to 15–30 petals. Obligate apomicts also showed a high petal variability (0–7). Apomicts are self-fertile and pollinator-independent, and therefore selection for petal formation to attract pollinators might be more relaxed than in sexuals. The functional and genetic background of petal variation, however, is still poorly understood (Text S1).

Heterozygosity in Relation to Ploidy, Reproduction Mode, and Environment

In general, genome-wide heterozygosity was significantly related to ploidy level, reproduction mode, and habitat types. Facultative- and obligate-apomictic, tetra- and hexaploids possessed up to three-times higher heterozygosity than sexual diploids, which is in line with previous studies using traditional genetic markers (Štorchová et al., 2002; Hörandl & Paun, 2007). Increased heterozygosity of apomicts is explainable by hybrid nature and fixed heterozygosity at duplicated loci (Beck et al., 2012; Gornall, 1999; Hodač et al., 2019; Hörandl & Paun, 2007; Pellino et al., 2013). Sexual, facultative-, and obligate-apomictic tetraploids, and facultative- and obligate-apomictic hexaploids showed similar heterozygosity, indicating a minor effect of

sexuality on heterozygosity within these ploidy levels. Significantly, more heterozygous sites in hexa- compared to tetraploids are potentially due to the Meselson effect, i.e., independent allelic sequence divergence over time in asexuals (Pellino et al., 2013; Welch & Meselson, 2000).

RADseq data represent a subset of the genome, including coding and non-coding regions (Davey et al., 2011). Non-coding DNA, if not involved in regulatory functions, is assumed to have higher heterozygosity than coding regions due to an absence of selective pressure against mutation accumulation. The herein presented heterozygosity values (0.35–2.98) based on RADseq loci were similar to those of self-compatible and annual sexual *Aethionema arabicum* lineages obtained from coding DNA regions (0.83–2.76), whereby tetraploid cytotypes had also twice as much heterozygosity as diploid cytotypes (Mohammadin et al., 2018). The estimated heterozygosity values of *R. auricomus* herein are probably representative for the nuclear genome because significant chloroplast fractions can be excluded, and "mean" orthologous RADseq locus assembly was ensured (see Materials and Methods, and Text S3). However, comparable whole-genome heterozygosity values of other apomicts are still missing.

Strikingly, we detected no relationship between heterozygosity and sexuality. On the one hand, heterozygosity is probably a consequence of hybrid origin of apomicts (Barke et al., 2018; Beck et al., 2012). On the other hand, facultative sexuality allows recombination, purging of deleterious mutations, and can reduce proportions of heterozygous individuals within populations (Cosendai et al., 2013; Gornall, 1999; Hörandl & Paun, 2007; Lushai et al., 2003). However, such effects might be better detectable with large-scale population studies.

Forest populations possessed more heterozygous individuals than populations in open habitats. Slightly higher heterozygosity of hexaploid facultative and obligate apomicts compared to tetraploid ones, and remarkably highly heterozygous hexaploid facultative apomictic populations from the Carpathian forests cause this effect (see Paun, Greilhuber, et al., 2006; Pellino et al., 2013). In contrast, hexaploid obligate-apomictic "*R. allemannii*" and hexaploid facultative-apomictic "*R. marsicus*" in subalpine meadows showed heterozygosity levels comparable to di- and tetraploids. Phylogenetic background and genetic bottlenecks instead of selection due to habitat conditions may better explain low heterozygosity in these geographically isolated taxa (see Tomasello et al., 2020).

We observed a hitherto unreported relationship between heterozygosity and temperature seasonality across Europe for GP plant complexes: Highly heterozygous apomictic populations rather occupy seasonally highly variable, mainly continental, areas of Northern, Eastern, and

Central Europe (Figs. 4 and 8). These areas are characterized by relatively dry, hot summers and cold winters. Higher levels of heterozygosity of apomictic polyploids might change the reaction norm and improve stress response, phenotypic plasticity, and the adaptive potential under more extreme conditions (Comai, 2005; Hörandl & Paun, 2007; Chen, 2007; Qiu et al., 2020). This way, heterozygosity indirectly promotes the GP pattern by enabling only polyploids to expand their range into areas with more extreme conditions.

Conclusions

We detected a latitudinal GP pattern within the *R. auricomus* complex. Apomicts showed a moderate to high admixture of genetic partitions and reticulate relationships, supporting the hypothesis of an allopolyploid origin (Barke et al., 2018; Hörandl et al., 2009; Hodač et al., 2014, 2018). Hybridization of divergent sexual progenitor species followed by polyploidization probably created numerous, highly heterozygous genotypes, which established populations under varying environmental conditions. Sexuality increased with solar radiation and isothermality, followed by minor positive precipitation and negative altitudinal effects, suggesting a hitherto undetected climatic niche differentiation. We conclude that niche differentiation in GP patterns is partly specific for species or apomictic complexes. The niche shift of polyploid apomicts towards more temperature variable and drier regions might be enabled by higher heterozygosity, increasing reaction norm, heterosis effects, stress resistance, and adaptive potential. Ploidy levels rather than environment explain differences in sexuality and heterozygosity. Apomicts can benefit from uniparental, pollinator-independent reproduction, which probably facilitated rapid colonization of deforested and/or deglaciated areas during range expansion (see Baker, 1967). In general, the range expansion of apomicts observed herein can be attributed to advantages of ploidy-dependent heterozygosity, facultative sexuality, uniparental reproduction, and climatic niche differentiation. We confirm previous hypotheses (Hörandl, 2006; Kirchheimer et al., 2018) that both genomic and environmental factors contribute to GP.

Acknowledgments

We thank the German Research Foundation for project funding (Ho4395/10-1 to E.H., within the priority program "Taxon-Omics: New Approaches for Discovering and Naming Biodiversity", SPP 1991), F. G. Dunkel for providing garden plants, B. H. Barke and J. Schmidt for technical help, S. Friedrichs for gardening. We also thank F. Hadacek, S. Hörandl, D. Metze, M. Karbstein, R. Karbstein, I. Schellin, and K. Spitzer for support during fieldwork.

Author Contributions

K.K. and E.H. designed research; K.K., L.H., and E.H. collected plant material; K.K., M.D., and E.L. performed research; K.K., E.L., and S.T. analysed data; K.K. wrote the paper with contributions from all authors.

Data Availability Statement

Basic data supporting the findings of this study are available within the manuscript or Supporting Information. R scripts are available on Github (https://github.com/KK260/ flow_cytom etric_and_RADseq_analyses). FC and FCSS data, and RADseq alignments are available on Figshare (https://doi.org/10.6084/m9.figsh are.13352429) upon publication. RADseq reads are stored in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA): BioProject ID PRJNA627796 (http://www.ncbi.nlm.nih.gov/bioproject/ 627796).

Orcid

Kevin Karbstein https://orcid.org/0000-0003-1424-6557
Salvatore Tomasello https://orcid.org/0000-0001-5201-1156
Ladislav Hodač https://orcid.org/0000-0002-6885-1317
Elvira Hörandl https://orcid.org/0000-0002-7600-1128

Chapter 5 – The Origin of the Polyploid Complex

Unraveling phylogenetic relationships, reticulate evolution, and genome composition of polyploid plant complexes by Rad-Seq and Hyb-Seq

Karbstein, K., Tomasello, S., Hodač, L., Wagner, N., Marinček, P.,
Barke, B. H., Pätzold, C., & Hörandl, E.

BioRxiv, 1–86 (2021b).
doi: 10.1101/2021.08.30.458250 (reuse allowed by author)

Complex genome evolution of young polyploid complexes is poorly understood. Besides challenges caused by hybridization, polyploidization, and incomplete lineage sorting, bioinformatic analyses are often exacerbated by missing information on progenitors, ploidy, and reproduction modes. By using a comprehensive, self-developed bioinformatic pipeline covering tree, structure, network, and SNP-origin analyses, we for the first time unraveled polyploid phylogenetic relationships and genome evolution within the large Eurasian *Ranunculus auricomus* species complex comprising more than 840 taxa. Our results rely on 97,312 genomic RADseq loci, target enrichment of 576 nuclear genes (48 phased), and 71 plastid regions (Hybseq; OMICS-data) derived from the 75 most widespread polyploid apomictic taxa and four di- and one tetraploid potential sexual progenitor species. Phylogenetic tree and structure analyses consistently showed 3–5 supported polyploid groups, each containing sexual progenitor species. In total, analyses revealed four diploid sexual progenitors and a one unknown, probably extinct progenitor, contributing to the genome composition of *R. auricomus* polyploids. Phylogenetic network, structure, and SNP-origin analyses based on RADseq loci and phased nuclear genes completed by plastid data demonstrated predominantly allopolyploid origins, each involving 2–3 different diploid sexual subgenomes. Allotetraploid genomes were characterized by subgenome dominance and large proportions of interspecific, non-

hybrid SNPs, indicating an enormous degree of post-origin evolution (i.e., Mendelian segregation of the diploid hybrid generations, back-crossings, and gene flow due to facultative sexuality of apomicts), but only low proportions of lineage-specific SNPs. The *R. auricomus* model system is the first large European polyploid species complex studied with reduced representation OMICS data. Our bioinformatic pipeline underlines the importance of combining different approaches and datasets to successfully unveil how reticulate evolution and post-origin processes shape the diversity of polyploid plant complexes.

Keywords

allopolyploidy, Europe, genome evolution, plastome data, apomixis, RADseq, *Ranunculus auricomus*, target enrichment

Introduction

Polyploidy, the presence of two or more full genomic complements (whole genome duplication), occurs across the tree of life (Otto & Whitton, 2000; Van De Peer et al., 2017; Rothfels, 2021). Whole-genome duplications have been observed in seed plants, and in several lineages of animals (mainly fish and amphibians), fungi, and protists (Mable et al., 2011; Van De Peer et al., 2017; Blischak et al., 2018). Polyploid cells and tissues occur throughout nature (also in humans) and are regarded as a cellular strategy for higher stress tolerance (Schoenfelder & Fox, 2015; Fox et al., 2020).

All flowering plants are ancient polyploids, as at least one polyploidization event occurred in their common ancestor, and several additional ones in various lineages (Soltis & Soltis, 2016; Van de Peer et al., 2017; Leebens-Mack et al., 2019). Neopolyploid formation for flowering plants is estimated to range between 30–70% of species and to cause upshifts of diversification rates in young polyploid complexes (Wood et al., 2009; Soltis et al., 2015; Landis et al., 2018). Key innovations in flowering plants have been hypothesized to be connected to polyploidy, for example, the carpel, double fertilization, and vessel elements (Soltis et al., 2015; Soltis & Soltis, 2016; Leebens-Mack et al., 2019). In addition to its evolutionary significance, important crop

plants are natural polyploids (e.g., wheat, potato, strawberry, coffee, cotton), and their evolution has often been exploited for agricultural purposes (Gordon et al., 2020).

The presence of multiple gene copies in polyploids allows for gene neo- and subfunctionalizations, epigenetic changes, and consequently a differential expression of homeologous genes (Comai, 2005; Blischak et al., 2018). Polyploidy provides larger physiological and phenotypic flexibility to respond to different environmental conditions (Hörandl, 2006; Blaine Marchant et al., 2016; Van De Peer et al., 2017; Karbstein et al., 2021), which facilitates colonization of various ecosystems (Te Beest et al., 2012; Rice et al., 2019; Fox et al., 2020; Meudt et al., 2021).

Genome evolution of polyploid lineages is complex and not only shaped by evolutionary origin and the genomic contributions of progenitors, but also by post-origin processes, resulting in a mosaic-like genome structure with parental, additive, and novel features (Soltis et al., 2015). Different polyploid formation types influence genome evolution: Autopolyploids arise within a species (tree-like evolution), whereas allopolyploids are formed by hybridization between different species/lineages followed by polyploidization (network-like evolution; Comai, 2005; Wendel, 2015; Blischak et al., 2018). Consequently, autopolyploids contain genetically similar subgenomes whereas genomes of allopolyploids are composed of previously diverged subgenomes. Allopolyploidization is considered particularly likely to create biotypes with novel genomic features (Abbott et al., 2013; Van de Peer et al., 2020; Rothfels, 2021). For instance, allopolyploids showed higher degrees of genomic, transcriptomic, and epigenetic changes than autopolyploids (Comai, 2005; Chen et al., 2007; Wendel, 2015; Soltis et al., 2015; Spoelhof et al., 2017).

After evolutionary origin, the genome structure of neopolyploids is fluid over evolutionary time scales, and genomes revert to a functionally diploid state (Soltis et al., 2015; Van De Peer et al., 2017). At the beginning of this process, various mechanisms influence polyploid genomes. Expression bias due to epigenetic changes and homeologous gene loss (biased fractionation) after polyploidization can cause subgenome dominance (Soltis et al., 2015; Wendel, 2015; Blischak et al., 2018; Alger & Edger, 2020). Moreover, Mendelian segregation in the first diploid hybrid generations before polyploidization, and/or backcrossing of polyploids to their sympatric progenitors might distort the original subgenome contributions (Barke et al., 2018; Hodač et al., 2018; Wagner et al., 2020). Gene flow between polyploid lineages further influences genome structure (Melichárková et al., 2020).

In plants, polyploidization and/or hybridization are frequently connected to apomixis, i.e., reproduction via asexually-formed seeds (Asker & Jerling, 1992; Brukhin et al., 2019; Hojsgaard & Hörandl, 2019). Noteworthy, not all neopolyploids are apomicts (e.g., Masci et al., 1994), and not all apomicts are polyploids (e.g., Brukhin et al., 2019). Apomixis is usually facultative, and residual sexuality allows for backcrossing to progenitors and intercrossing of polyploids, resulting in huge networks of hundreds to thousands hybridogenetic lineages (Fig. 1). Such complexes occur in many abundant plant genera, e.g., dandelions (*Taraxacum*), hawkweeds (*Hieracium* s.l.), brambles (*Rubus*), and *Citrus*. With higher ploidy levels and/or time, these lineages are expected to become fixed, and mutations remain as the only source of genetic variation (Grant, 1981; Coyne & Orr, 2004; Fig. 1). With reduced recombination, heterozygosity in allopolyploids is additionally increased by allelic sequence divergence in asexual lineages (Meselson effect; Welch & Meselson, 2000; Pellino et al., 2013). Studies using genome-wide data showed that heterozygosity significantly increased with higher ploidy levels (Mohammadin et al., 2018; Karbstein et al., 2021). In general, heterozygosity has several benefits, such as novel genetic combinations, heterosis, buffering effects of deleterious mutations, or changes in secondary metabolites (Comai, 2005; Qiu et al., 2020). Increased heterozygosity is considered an important factor for the spreading of polyploids towards more variable climatic conditions (Hörandl, 2006; Rice et al., 2019; Karbstein et al., 2021).

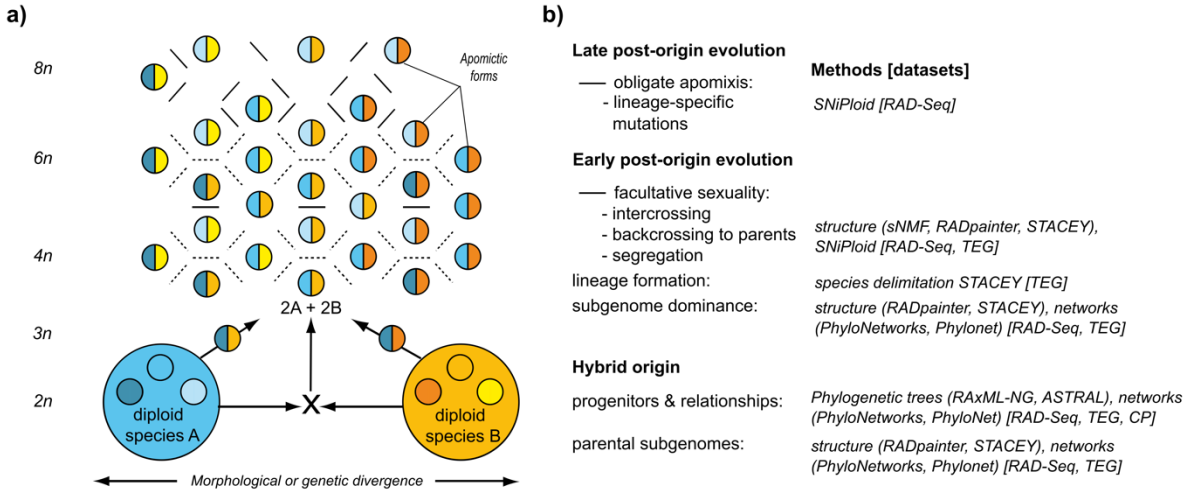


Fig. 1. Evolutionary processes in young polyploid species complexes and methods to address these processes (a) Evolution of an apomictic polyploid complex from two sexual progenitor species and evolution of lineages after origin (redrawn and modified after Babcock & Stebbins, 1938; see also Grant, 1981 and Coyne & Orr, 2004 for modern interpretations). (b) description of respective evolutionary processes and the corresponding analytical methods and pipelines applied here; CP = chloroplast regions, RAD-Seq = RADseq loci, TEG = target enriched nuclear genes. For a detailed scheme of bioinformatic pipelines see Fig. 3.

Despite the evolutionary, ecological, and economical importance of polyploidy, the understanding of phylogenetic relationships, genome diversity and evolution of fast-evolving, young polyploid species complexes remains limited. Traditional sequencing markers from organellar DNA were insufficient for reconstructing reticulate relationships in polyploid complexes because of uniparental inheritance (Rothfels, 2021). Nuclear markers from single regions (e.g., ribosomal DNA) are biased by a strong marker-specific evolution (e.g., Zarrei et al., 2014; Fehrer et al., 2021). Historically, polyploids were thus often avoided or dropped in phylogenetic studies (Freyman et al., 2020; Rothfels, 2021). Already the sexual progenitors of polyploid complexes are often characterized by low genetic divergence, incomplete lineage sorting (ILS), gene flow, and partial hybridogenic origins (Hörandl, 2018; Pease et al., 2018; Wagner et al., 2019; Karbstein et al., 2020a, 2020b). The use of 'OMICS'-data provides orders of magnitude more information compared with traditional genetic markers (Harrison and Kidner, 2011; Soltis et al., 2013). OMICS-data have proven effective at resolving diploid (and few polyploid) phylogenetic relationships of species that diversified more than 30 or even less than 0.1 Ma (Pellino et al., 2013; Hipp et al., 2014; Carter et al., 2019; Gordon et al., 2020; Karbstein et al., 2020b; Tomasello et al., 2020; Wagner et al., 2020).

Currently, two main approaches of reduced representation of OMICS-data are commonly used: restriction site-associated DNA sequencing (RADseq; and similar methods) and target enrichment (hybrid capture). RADseq covers a subset of anonymous, non-coding and coding regions across the entire genome and is mostly used for population genomics and phylogenomics of closely related species within genera, up to tens of million years of divergence (Davey et al., 2011, Ree & Hipp, 2015, McKain et al., 2018). RADseq is less costly and work-intense compared to target enrichment (McKain et al., 2018), and hence allows to process larger sample sets. Target enrichment usually addresses a subset of several hundreds of low-copy nuclear genes, and thus provides more conservative markers for resolving relationships within and among genera (Schmickl et al., 2016; McKain et al., 2018; Carter et al., 2019; Tomasello et al., 2020; Melichárková et al., 2020).

Although RADseq yields much more information (number of loci and SNPs) than target enrichment, locus dropout caused by mutation accumulation in cutting sites can become more and more problematic with increasing species divergence (but see in Eaton et al., 2017 for the influence of sequencing coverage). Moreover, the correct definition of loci and filtering of paralogs based on anonymous short sequence reads is a bioinformatic challenge (Ree & Hipp, 2015; O'Leary et al., 2018; McKain et al., 2018). Target enrichment loci are predefined from probe design and assembled loci are usually longer (McKain et al., 2018), allowing for gene tree estimation and allele phasing, and thus coalescent-based methods. Allelic information (segregating markers at a single locus) is particularly important for correct phylogenetic inferences in highly reticulate, young evolutionary relationships (Eriksson et al., 2018). Coalescent-based models can reconstruct correct species trees and estimate species boundaries while accounting for stochastic processes like ILS (Rannala & Yang, 2003; Jones et al., 2015; Rannala, 2015). In addition, plastid data can be easily gained from off-target reads of target enrichment (Hybseq; Weitemier et al., 2014; Folk et al., 2015; McKain et al., 2018). The incorporation of plastid data into network reconstruction to gain information on the maternal progenitor has been largely overlooked in the last few years. Nuclear-plastid discordances have been assessed on shallow to deep phylogenetic scales, and elucidated group-specific evolutionary processes (Huang et al., 2014; Stull et al., 2020).

Elucidating the evolution of allopolyploids is even more challenging due to reticulate evolution. Tree methods can give a first phylogenetic framework for polyploid reconstructions when no previous phylogenetic study exists. Particularly in evolutionary young species complexes containing genetically close taxa (see e.g. McDade, 1992 for tree stability in presence of hybrids from closely related taxa) and without any previous knowledge about auto- vs. allopolyploid

origins, trees and (quartet) support values can give valuable information on conflicting signals. For example, a non-conflicted, tree-like pattern rather hints at autopolyploids whereas trees with low (quartet) support values can indicate the presence of reticulations and/or ILS (Lo et al., 2010; Brandrud et al., 2020; Karbstein et al., 2020b, Tomasello et al., 2020). However, hybridogenic, network-like origins cannot be inferred by both standard and coalescent methods based on bifurcating models, leading to incongruences in tree reconstructions (McBreen & Lockhart, 2006; Rothfels, 2021). Consequently, phylogenetic relationships should not (only) be presented by bifurcating trees (McDade, 1992, 1995; Huson & Bryant, 2006; Rothfels, 2021).

Distance-based network methods like for example the popular NeighborNet algorithm can visualize reticulate relationships better than trees, but detailed information on ancestry or parentage in hybrid scenarios requires phylogenetic networks (Huson & Bryant, 2006; Oxelman et al., 2017). Recently developed software can model network-like evolution with maximum pseudolikelihood from gene trees or SNP-based multilocus sequence data under the coalescent model accommodating ILS (e.g., PhyloNet or PhyloNetworks; Than et al., 2008; Solís-Lemus et al., 2017; Wen et al., 2018; Olave & Meyer, 2020; Flouri et al., 2020).

Phylogenetic network inference requires information on ploidy level and diploid progenitors, allowing correct heterozygosity estimations and allele phasing in polyploids. Recently, an increasing number of studies focused on network estimations and technical improvements in polyploid reconstructions (e.g., phasing, allele sorting, subgenome assignment, or modeling the (allo)polyloidization process; Bertrand et al., 2015; Jones, 2017a; Oberprieler et al., 2017; Dauphin et al., 2018; Cao et al., 2019; Lautenschlager et al., 2020; Freyman et al., 2020, Yan et al., 2020; Šlenker et al., 2021; Tiley et al., 2021). Nevertheless, knowledge on putative progenitor species, number of contributing progenitors, ploidy levels, and formation types of the polyploids within large species complex are frequently missing. Moreover, only one resource-intensive program is currently capable to model the polyploidization process itself, i.e., that homeologues of an allotetraploid share demographic parameters or divergence times from their progenitors (Jones, 2017a). In addition, for current allele assignment methods (e.g., Lautenschlager et al., 2020; Šlenker et al., 2021), subgenomes should be well genetically differentiated, sequences of the diploid parents available/not extinct, and locus/gene datasets not too big. Therefore, more sophisticated methods (e.g., polyploid networks, multi-labeled subgenome trees) are often not applicable at that stage of research and/or currently still inappropriate for young polyploid species complexes.

Even with this information, reconstruction of relationships and discrimination between auto- and allopolyploid scenarios might be difficult. For example, when using maximum likelihood (or pseudo-likelihood) approaches, likelihood scores of networks are usually not directly comparable to those of trees. A-posteriori model comparisons need to be applied to discriminate among scenarios with different numbers of reticulations (e.g., Kamneva et al., 2017; Cai & Ané 2020). Moreover, a correct and unequivocal network inference is hard to reconstruct in young species complexes where progenitors exhibit high levels of genetic admixture and polyploids possess high levels of genome-wide heterozygosity. Therefore, relationships, reticulate evolutionary processes, genome composition, structure, and evolution within large polyploid species complexes remain largely uninvestigated.

In this study, we unravel for the first time the evolutionary processes shaping apomictic polyploid complexes on the model system *Ranunculus auricomus* by using reduced-representation genomic data. The complex ranges from Greenland, Europe to Western Siberia, and spans arctic, boreal, temperate, and Mediterranean climates (Jalas & Suominen, 1989). Linnaeus (1753) already described a species with dissected basal leaves, *R. auricomus*, and one with undivided basal leaves, *R. cassubicus*. Since then, more than 840 taxa (morphospecies) have been described, inhabiting stream- and riversides, and semi-dry to marshy meadows and forests (Karbstein et al., 2020b, 2021). Most of these taxa are tetra- to hexaploid and apomictic (Jalas & Suominen, 1989; Karbstein et al., 2021). Only four di- and one tetraploid, genetically and geographically distinct, sexual species were detected so far and originated 0.83–0.58 Mya (Karbstein et al., 2020a, 2020b; Tomasello et al., 2020): *R. cassubicifolius* s.l. (di- and autotetraploid) and *R. notabilis* s.l. (diploid) are most distantly related whereas *R. flabellifolius*, *R. envalirensis* s.l. (both diploid), and *R. marsicus* (tetraploid), are grouped in intermediate positions (Karbstein et al., 2020b). *R. cassubicifolius* s.l. and *R. flabellifolius* are characterized by non-dissected basal leaves whereas the other species show a strongly heterophyllous cycle with dissected basal leaves during anthesis (Karbstein et al., 2020b).

Vicariance processes probably triggered allopatric speciation during climatic deteriorations in the late Pleistocene from a widespread European ancestor (Tomasello et al., 2020). It has been hypothesized that the large number of asexual, mainly tetra- to hexaploid polyploids arose from hybridization of sexual progenitors (Hörandl et al., 2009; Hodač et al., 2014, 2018; Hojsgaard et al., 2014b; Barke et al., 2018). Some polyploid apomicts are probably less than 0.1 Mya (Paun et al., 2006a; Pellino et al., 2013). They occupy larger, more northern areas, possess higher levels of genome-wide heterozygosity, and are obligate apomictic or with low levels of facultative sexuality (Karbstein et al., 2021). Nevertheless, origin, relationships, and genomic

composition of the polyploid complex have never been analyzed due to genetic and bioinformatic limitations.

In this study, we compare a comprehensive taxon sampling based on genomic RADseq data (280 individuals, 80 taxa), (phased) nuclear genes (113 individuals, 50 taxa), and plastid regions (87 individuals, 45 taxa), to unravel phylogenetic relationships and genome composition of the large, evolutionary young, *R. auricomus* polyploid complex. We use a comprehensive, self-developed bioinformatic pipeline combining previous knowledge about sexual progenitors, ploidy and reproductive data with tree, structure, network, and SNP-origin methods across different datasets (Fig. 1) to answer the following questions: (i) Are the applied tree analyses able to give a first phylogenetic framework, and do well-supported (main) clades exist? (ii) Do genomic, nuclear-gene, and plastome data reveal congruent tree topologies or rather conflicting signals due to reticulate evolution? (iii) Do RADseq or phased nuclear gene data reflect any clear genetic and/or geographical structure? (iv) Are polyploid lineages of auto- or allopolyploid origin? (v) If the latter, how many progenitors contributed to their genomes? (vi) To which extent are polyploid genomes influenced by post-origin evolution? (vii) How can analyses of RADseq and Hybseq data be integrated for unraveling evolutionary processes in polyploid complexes?

Materials and Methods

Population Sampling

In the present study, we included four di- and one tetraploid sexual species, and 75 of the most widespread, tri- to hexaploid apomictic *R. auricomus* taxa. The new classification of sexual species is described in Karbstein et al. (2020b). Ploidy and reproduction mode measurements of *R. auricomus* individuals and populations (sexual, and facultative and obligate apomictic) needed for the performed analyses herein are published in Karbstein et al. (2021) (see also Table S1, Figs. S4, S5 in Karbstein et al., 2021, and data on FigShare). The diploids *R. sceleratus* and *R. pygmaeus* were used as outgroups. We collected silica-gel dried leaf material from living plants for all genetic analyses, and additionally, leaf material from herbarium specimens for target enrichment analyses. Finally, we used 280 samples originating from 235 collection sites (populations) across Europe for further analyses (Fig. 2, Table S1). Concerning genomic

analyses, the sexual progenitor species were treated as parental subgenomes and abbreviated according to the legend of Fig. 2.

Laboratory Work, Locus Assembly, and Parameter Optimization

DNA extraction of 280 *R. auricomus* and outgroup samples, adjustment of DNA concentration, DNA quality check, RAD lab workflow and sequencing with the cutting enzyme *PSTI* and single-end RAD sequencing of 100bp reads (Baird et al., 2008), raw read quality check, raw read demultiplexing, removal of adapter sequences and restriction overhang, and further quality filtering in IPYRAD (Eaton & Overcast, 2020) followed Karbstein et al. (2020b, 2021). Here, we used the already sequenced samples of Karbstein et al. (2020b, 2021).

PSTI is a methylation-sensitive enzyme and hence can considerably reduce the fraction of repetitive elements that is otherwise very high in plants. Therefore, the enzyme targets mostly nuclear genes and a few organelle sites (Fellers, 2008). This dataset of coding- and non-coding regions complements the markers derived from expressed genes (transcriptomes of flowering buds) and selected for target enrichment (see Tomasello et al., 2020 for baits design in *R. auricomus*) to get a comprehensive representation of the nuclear genome. Target enrichment further provides markers of the plastid genome. The self-developed, bioinformatic pipeline combining different datasets and analyses, and previously published *R. auricomus* studies (sexual progenitor species, reproduction modes) is illustrated in Fig. 3.

For target enrichment, we added 85 newly sequenced polyploid apomictic samples to the already existing 28 samples sequenced by Tomasello et al. (2020) (113 samples in total; Table S1). All plastome data (CP) from off-target reads is published here. We included almost the same samples as in the RADseq analyses and added as described above herbarium specimens (types or collections from type locations; Table S1). We used the bait set as described previously in Tomasello et al. (2020), consisting of 17,988 probes and capturing 736 target genomic regions. Library preparation and hybrid capture protocols are available as Text S1 in Tomasello et al. (2020). Libraries were sequenced in five different paired-end runs (24 samples each) with 2×250-bp on an Illumina MiSeq system at NGS Integrative Genomics Core Unit of the University of Göttingen (Germany).

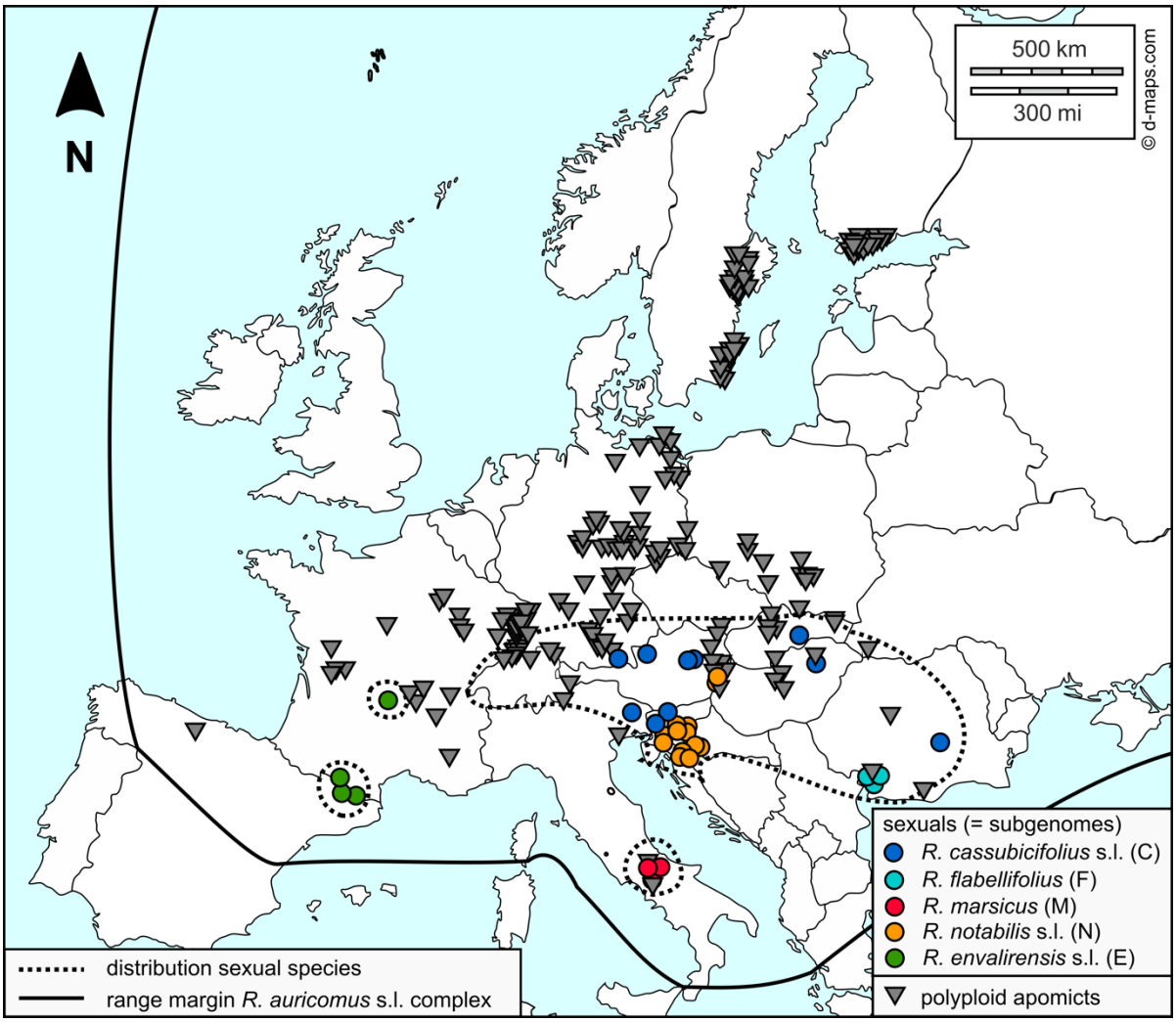


Figure 2. Locations of studied *R. auricomus* populations across Europe. We investigated 235 sexual and apomictic populations (see Table S1 for details). Symbols represent reproduction modes of populations (colored circles = sexuals, defined as subgenomes here for further data analyses), dark grey triangles = obligate or facultative apomictic, also in Karbstein et al. (2020b, 2021). Circles of sexual species were highlighted according to the color scheme of Fig. 4. The solid line shows the range margin of the *R. auricomus* complex, and the pointed lines highlight the distribution of sexual species. The original map was downloaded from https://d-maps.com/, created by Karbstein et al. (2021), and modified for this study.

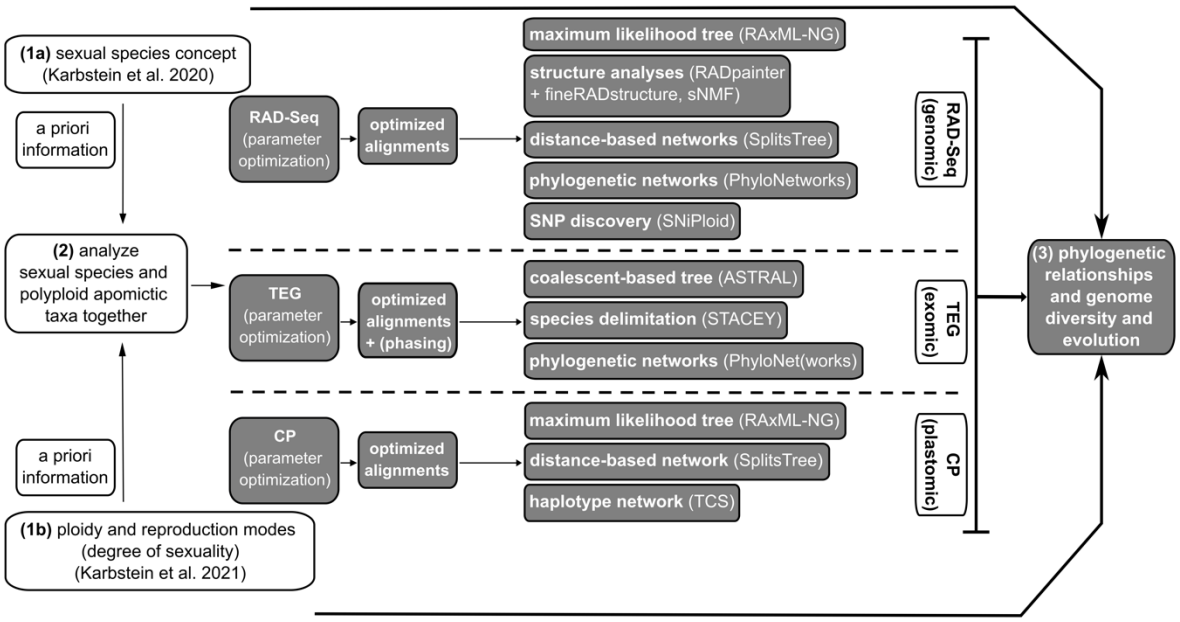


Figure 3. Bioinformatic pipeline to resolve polyploid species complexes. Here, we used *R. auricomus* as a model system and basically followed the concept of Hörandl (2018) to disentangle complicated species complexes. We analyzed (2) sexual species and apomictic taxa together, using a priori information about (1a) sexual species (Karbstein et al. 2020b) and (1b) ploidy levels and reproduction modes (Karbstein et al. 2020b, 2021). Analyses are based on the optimized alignments of three different datasets covering genomic parts (RAD-Seq), exomic nuclear regions (target enrichment, TEG), and plastome regions (chloroplast, CP). We used RADseq datasets to calculate maximum likelihood (ML) trees, genetic structure analyses, distance-based networks, maximum pseudolikelihood networks, and SNP discovery analyses. To study the robustness of RADseq results, we computed coalescent-based trees, species delimitation analysis, and maximum pseudolikelihood networks based on target enrichment datasets. A ML tree, and distance-based and haplotype networks of the CP dataset were also included to get further details about hybridogenic origins of polyploids.

For *de novo* assembly of RADseq loci and parameter optimization, we used IPYRAD v.0.9.14 (and v.0.9.52) on the local HPC-Cluster (GWDG, Göttingen, Germany). For parameter optimization, we applied an already established workflow accounting for different ploidy levels of *R. auricomus* individuals: The within-sample clustering similarity threshold was optimized for each ploidy level (2n–6n) balancing number of RADseq clusters, cluster depth, and clusters rejected due to high heterozygosity. Then, the among-sample clustering threshold was optimized for the merged assembly optimizing number of polymorphic loci, SNPs, loci filtered by maximum number of SNPs, removed duplicates, shared loci, and new polymorphic loci. Maximum number of SNPs per locus and of indels per locus were increased to 30% and 12, respectively, to account for greater genetic variation in polyploids as described in Karbstein et al. (2021).

For subsequent analyses, we created a 'without-outgroup' and a 'total' dataset. To assess effects of number of loci and missing data on phylogenetic analyses (Eaton et al., 2017; O'Leary et al., 2019; Karbstein et al., 2020b), we selected different minimum amounts of samples per locus and created 'min10' (10%), 'min30' (30%), and 'min50' (50%) alignments balancing the specific program requirements and informativeness of datasets (see below). The final sample size totals 282 individuals (incl. outgroup). For both datasets, sample filtering led to ca. 74% (min10), 55% (min30), and 44% (min50) missing data in the final sequence matrices.

For target enrichment data analysis, reads were processed with HYBPHYLOMAKER v1.6.4 (Fér & Schmickl, 2018) (Text S1), using the target regions (exons) selected for the bait design from transcriptomes as 'pseudoreference' for read mapping (Table S2 in Tomasello et al., 2020). Samples with more than 40% missing data were filtered out from each exon region. In addition, only loci including more than 90% of samples were further processed (579 genes). From those 579 genes, 50 loci were selected for the species delimitation and phylogenetic network analyses, to be informative, non-homoplasious, and free from paralogue sequences. To select these loci, we assessed four different parameters across the 579 alignments, scoring the respectively best performing 25% of loci with 1 and the remainder with 0. The parameters were the following (Herrando-Moraira et al. 2018): (i) the R^2 of mutational saturation regression curves (Philippe and Forterre 1999), (ii) the standard deviation of the sample-specific long-branch scores (LB scores; Struck et al. 2014), (iii) the clocklikeness, and (iv) average bootstrap (BT) support. This was done (i-iii) following the idea that in such a young species group like *R. auricomus* evolutionary rates will not change considerably among branches in orthologous regions. Finally, we selected the 50 loci with the highest overall score (Table S3).

For inference of polyploid origins, retrieving allelic information is crucial. We phased these 50 most informative loci using a similar approach as described in Eriksson et al. (2018). We processed the mapped BAM files of all samples with SAMTOOLS v.0.1.19 (Li et al., 2009) ('sort' and 'phase' commands). The polyploid samples (tri- to hexaploids) were phased further, looking at the phased BAM files in IGV v.2.8.9 (Robinson et al., 2011) (usually one of the BAM files was a consensus of alleles with some vowels corresponding to the heterozygous sites) and manually adding alleles in relation to the known ploidy level to the alignments using AliView vers. 1.26 (Larsson, 2014). For two of the 50 loci, it was not possible to unequivocally detect alleles in at least one of the polyploid samples. Therefore, we excluded these two loci from further analyses (Table S4).

Off-target reads were used to gain information on the plastid genome again performing HYBPHYLOMAKER. We used the *Ranunculus repens* plastid genome as reference (Table S5). Considering the low number of mapped reads and the resulting highly fragmented alignments, we excluded regions and samples with the highest amounts of missing data, to minimize phylogenetic inaccuracy in the subsequent analyses. First, we excluded samples with more than 50% missing data (from each plastome region separately), and regions containing sequence information for fewer than 50% of the samples. Second, we excluded from all regions all samples missing from more than 50% of the alignments. After filtering, we retrieved a subset of 71 regions including genes and intergenic spacers, for 87 samples from the original 113 (Table S6).

Maximum Likelihood Tree and Quartet Sampling (RADseq)

To infer phylogenetic relationships among sexuals and polyploid apomicts (Figs. 1, 3), maximum likelihood (ML) analyses were performed with RAXML-NG v.0.9.0 (Kozlov et al., 2019) on the final RADseq min10, min30, and min50 datasets. As input, we used the *.phy IPYRAD output files (each individual characterized by one sequence [majority-rule base calling], all loci concatenated into a supermatrix). Alignment patterns were compressed and stored in binary formats (RBA). Then, we inferred the respective tree under the GTR+GAMMA model with 10 random and 10 parsimony starting trees. Standard non-parametric BT were performed, and the MRE-based bootstopping test (cutoff: 0.05) applied. Felsenstein bootstrap proportion (FBP) and transfer bootstrap expectation (TBE) values were calculated by RAXML-NG. TBE is more appropriate for large phylogenies (>300 samples) and for phylogenies with conflicted (deep) branches compared to FBP (e.g., hybridization events; Lemoine et al., 2018). FBP and TBE values were mapped by RAXML-NG onto the best-scoring ML trees. The min10 compared with min30 and min50 alignments showed the highest mean BT support (FBP: 70/66/47, TBE: 85/82/64) and the largest number of monophyletic taxa (13/10/9) (Figs. S1–S3a–d). Moreover, the min10 tree topology was similar to min30 and min50 tree topologies (Figs. 4a, S1, S2). Therefore, we selected the min10 tree for final interpretations.

We addressed potentially inflated BT values in concatenated analyses with Quartet Sampling (QS) v.1.3.1 (Weisrock et al., 2012; Shen et al., 2017; Pease et al., 2018; and see also Karbstein et al., 2020b). The quartet concordance score (QC) is defined as the ratio of concordant to both discordant quartets (1: all concordant, > 0: more concordant patterns, < 0: more discordant patterns), the quartet differential score (QD) indicates the skewness of both discordant patterns

(1: equal, 0.3: skewed, 0: all topologies 1 or 2), and the quartet informativeness score (QI) describes the proportion of informative replicates (1: all informative, 0: none informative; see Pease et al., 2018). QD values around 1 indicate ILS (presence of both discordant topologies) whereas QD values towards 0 hint at directional introgression (presence of one alternative topology; Pease et al., 2018; see also Karbstein et al., 2020b). We set 100 replicates per branch and log-likelihood threshold cutoff to 2. The quartet concordance factor (QC), quartet differential (QD), and the quartet informativeness (QI) scores together with BT values were illustrated in Fig. 4a (detailed QS values in Fig. S3a–d).

Coalescent-Based Species Tree and Quartet Support (Targeted Genes)

As an equivalent to ML trees for RADseq alignments, we estimated a coalescent-based tree based on 576 target enriched genes. First, gene trees were inferred in RAXML v.8.2.12 (Stamatakis, 2014). Analyses were run with 100 standard BT replicates, setting the GTR+GAMMA model and partitioning by exons. Second, gene trees were rooted and combined into a single newick file in HYBPHYLOMAKER (Fér & Schmickl, 2018). Third, the species tree was inferred by applying the coalescent-based algorithm implemented in ASTRAL III v.5.6.3 (Zhang et al., 2018) with 100 multilocus BT replicates. To assess the amount of gene tree conflict on branches, we measured quartet support on the ASTRAL tree (Sayyari & Mirarab, 2016; Fig. 4b, Fig. S4a–c).

Plastome Phylogeny and Network Analysis (CP)

We used the 71 selected plastid regions to infer a ML tree with 100 BT replicates by RAXML-NG (Kozlov et al., 2019). Models of sequence evolution were assessed for each region separately using MODELTEST-NG v.0.1.6 (Darriba et al., 2020). Alignments were concatenated (80,461 base-pairs in total) and different regions were treated as different partitions, each with its respective sequence evolution model.

To gain additional information about haplotype evolution, the concatenated matrix was used to infer a haplotype network with TCS v.1.13 (Clement et al., 2000). We used the web-based software tcsBU (Múrias Dos Santos et al., 2016) to produce a graphical representation of TCS haplotype network. In addition to the TCS networks, we calculated neighbor-net networks as described in Karbstein et al. (2021).

Genetic Structure (RADseq)

To investigate genetic structure of the polyploid species complex (Figs. 1, 3), we first conducted analyses with RADPAINTER+FINERADSTRUCTURE v.0.3.2 (Malinsky et al., 2018) using the *.alleles.loci IPYRAD output files (each locus with a maximum of four allowed alleles (only two phases due to diploid SNP calls), with individual sequences). RADPAINTER is based on a coancestry matrix, uses all SNPs, allows a varying allele number, and tolerates moderate amounts of missing data (Malinsky et al., 2018). We ran RADPAINTER to calculate the coancestry matrix, and used FINERADSTRUCTURE to assign individuals to groups (1,000,000 burn-in and 1,000,000 sample iterations) including a simple tree building (MCMC; 100,000 burn-in). Finally, we plotted results using a modified R script of 'fineRADstructurePlot.R' (R v.4.0.3 (R Core Team, 2020) for all R analyses). We then compared results of min10, min30, and min50 alignments. With increasing number of loci and missing data, an increased number of groups and genetic dissimilarity among groups was detected (Fig. S5). Thus, we selected min10 for further interpretations.

Second, we carried out structure analyses applying SNMF within the R package 'LEA' v.3.0.0 (Frichot et al., 2014; Frichot & François, 2015). SNMF provides a fast and efficient estimation of individual coancestry, is robust to deviations from Hardy-Weinberg equilibrium, and can deal with moderate levels of missing data (Frichot & François, 2015). We used *.ugeno (each individual characterized by numbers indicating one randomly chosen per SNP locus) IPYRAD files, and set number of genetic clusters (K) from 1 to 80 (maximum number of included taxa), ploidy to 4 (as maximum), and repetitions to 7. To choose the number of ancestral Ks, we used the implemented cross-entropy criterion. Cross-entropies were plotted for all Ks. Across datasets, we found the optimal Ks between 3 and 5 (Fig. S6a–f). We plotted results of optimal Ks as bar graphs and across Europe (Figs. S7a–c, S8a–c). Additionally, we displayed ancestry coefficients (method 'max', i.e., at each point the cluster for which the ancestry coefficient is maximal; Figs. S9–11 without method 'max') of the respective best run of each K on geographical maps of Europe, using location coordinates and the R script POPSutilities.R (source: http://membres-timc.imag.fr/Olivier.Francois/POPSutilities.R; Fig. S8a–c). The min30 datasets balanced number of loci and amounts of missing data and revealed the most reasonable results (see explanation in legend of Fig. S8, and was therefore selected for further interpretations.

Genetic Structure (Targeted Genes)

To unravel the genetic structure of the polyploid complex based on nuclear genes (Figs. 1, 3), we utilized the coalescent-based species delimitation approach of STACEY v.1.2.1 (Jones, 2017b). Input files were prepared in BEAUTI v.2.6.1 (Bouckaert et al., 2014) using the 48 phased loci. For the analyses, each sample was treated as "minimal cluster" (i.e., alleles of the same individuals were represented by a single tip in the species tree/species delimitation results). Sequence substitution models were selected for each locus separately using the Bayesian Information Criterion (BIC) in MODELTEST-NG. Substitution models, clock models, and gene trees were treated as unlinked for all loci. To reduce the search space, parameters of the substitution models were fixed to those found in MODELTEST-NG. The strict clock was enforced for all loci fixed at an average rate of 1.0 in one random locus while estimating all other clock rates in relation to this locus. We set the 'collapse height' to 1×10^{-5}, which was estimated using a Beta prior with parameters $\alpha = 1.0$ and $\beta = 1.0$, and which represent a flat distribution between 0 and 1 (i.e., all possible species delimitation scenarios have an equal prior probability). Finally, we gave to the bdcGrowthRate prior a log-normal distribution (M = 4.6 and S = 1.5), a gamma shape ($\alpha = 0.1$ and $\beta = 3.0$) to the popPriorScale prior, and for the relativeDeathRate, we set a beta prior ($\alpha = 1.0$ and $\beta = 1.0$; optimized in Karbstein et al., 2020b).

The analyses were run for 2×10^9 iterations sampling every 200,000th generation in BEAST v.2.6.1 (Bouckaert et al., 2014). Two independent runs were performed and, after checking convergence between independent analyses and Effective Samples Size values (ESS > 100) in TRACER v.1.6 (Rambaut et al., 2018), we combined trees output files using LOGCOMBINER v.2.6.1 (Bouckaert et al., 2014) and discarding 10% of the analyses as burn-in (as described in Karbstein et al., 2020b). The obtained file was processed with the 'species delimitation analyser' (Jones et al., 2015). The similarity matrix was produced using a modified version of the R script (Jones et al., 2015).

Detecting Subgenome Contribution for Selected Polyploids (RADseq, Targeted Genes)

To investigate genome diversity, composition, and evolution of polyploids in more in detail (Figs. 1, 3), we selected 2–4 polyploid individuals with obvious reticulation (coancestry) signals per main genetic cluster and tested ten polyploids for subgenome contributions (H_1-H_{10}; Table 1, Table S1). We used hybrid binomials to clearly distinguish these taxa from sexual species (Hörandl et al., 2009). RADPAINTER+FINERADSTRUCTURE analyses include all SNPs per locus and varying ploidy levels (Malinsky et al., 2018), and is therefore here considered as

superior compared with SNMF and SPLITSTREE (1 SNP/locus). Therefore, we evaluated the RADPAINTER coancestry matrix, and calculated a median coancestry value of 343 (mean right-skewed). We took the median as the critical threshold to assess the potential subgenome contributions of polyploids. The same procedure was also applied to the STACEY posterior probability matrix (median = 0.000555). To ensure comparability among datasets, we aimed at selecting the same individuals. Only for '*R.* × *elatior*' (H_2), we selected another individual in STACEY analysis (*R.* × *elatior* is monophyletic, Figs. S1–S4a–c).

Phylogenetic Network Analyses (RADseq)

To corroborate the already gained information by appropriate network methods (Figs. 1, 3), we carried out analyses with PHYLONETWORKS v.0.12.0 (Solís-Lemus et al., 2017). PHYLONETWORKS allows network inference with maximum pseudolikelihood from multilocus sequences (SNAQ). SNAQ uses a multi-species network coalescent model (MSNC) that is capable of handling ILS (Solís-Lemus et al., 2017). Since SNPs were not intended as input for SNAQ, we used the recently published function SNPs2CF.R v.1.2 (Olave & Meyer, 2020) to transform SNP-based RADseq alignments into quartet concordance factors (CF). We selected the min30 dataset to avoid bias of network analyses by excessively high amounts of missing data. We converted the *.ustr (each individual characterized by numbers indicating one randomly chosen SNP per locus, two phases (maximum of four allowed alleles) per individual) IPYRAD file with a custom R script to an adequate input format for SNPs2CF. Network analyses are computationally intensive. Thus, we created ten subsets each containing one tetraploid accession (individual) and all available accessions (individuals) of diploid sexual progenitor species. The above-mentioned, preselected tetraploids were used for subset building (see Detecting Subgenome Contribution for selected Polyploids). We excluded the sexual tetraploid *R. marsicus* from network analyses because no significant subgenome contribution of this species was observed in previous genetic structure analyses. We used the converted *.ustr files and imap files containing individual-species associations as input for SNPs2CF. We specified 'between species only' comparisons, no maximum number of SNPs, maximum number of quartets of 1000, and 100 BTs.

We used the received quartet CF matrices and quartet-CF-based starting trees to run maximum pseudolikelihood (SNAQ) analyses with default settings. We initially allowed no hybridization event. Afterward, the output was used as a start network (net0) for the next analysis allowing one hybridization event (net1). Per polyploid, the likeliest network was commonly the one with

the polyploid as hybrid (seven out of ten). The polyploids H_4, H_5, and H_9, were not inferred as the likeliest hybrids. An explanation might be the low genetic divergence among polyploids and diploid progenitors causing problems in ILS and hybridization modeling. However, since SNAQ (PHYLONETWORKS) takes no hybrid constraint and polyploids cannot be the progenitor of diploid sexuals here, we had to select the less likely hybrid network in these cases for further polyploid analyses.

Phylogenetic Network Analyses (Targeted Genes)

To assess the validity of previous structure and RADseq-network results (Figs. 1, 3), we additionally performed phylogenetic network analyses using the 48 phased target genes. We also investigated $H_{1–10}$, taking gene trees as input and two different, separately performed, coalescent-based approaches: SNAQ implemented in PHYLONETWORKS (Solís-Lemus et al., 2017) as for RADseq data and the maximum pseudolikelihood (InferNetwork_MPL) approach implemented in PHYLONET v.3.8.2 (Than et al., 2008; Wen et al., 2018). PHYLONET is one of the most widely used and established programs for species tree/network reconstructions based on multilocus datasets. We told both programs that all alleles (phases in RAD-Seq-based networks) of diploid species are from their respective species, and alleles (pseudophases in RAD-Seq-based networks) of the polyploid are only from the single polyploid accession. Thus, we used network results based on phased nuclear genes for further validation of previous results.

For each polyploid testing, alignments were modified to include all diploid accessions (except for *R. cassubicifolius* s.l. LH006 and EH9126, and *R. flabellifolius* LH021; 22 samples in total) and the polyploid individuals. Models of sequence evolution were selected with MODELTEST-NG, and 100 BT gene trees were inferred with RAXML-NG for each of the 48 selected loci. Therefore, 100 gene trees per locus (4,800 trees in total) were used as input, to incorporate gene tree uncertainty while inferring species networks (and to ensure dataset comparability for SNAQ and PHYLONET). For the PHYLONETWORKS analyses, we used the gene trees and a mapping file (mapping alleles to species) to calculate a species-wise CF table. We continued the analyses as for the RADseq dataset, with the only exception that the starting tree was inferred using ASTRAL III. For the PHYLONET MPL analyses, the polyploid was always specified as the putative hybrid. We performed 10 runs per search, each returning five optimal networks. After the search, the returned species networks were optimized for their branch

lengths and inheritance probabilities under full likelihood (-po option in PhyloNet), using the default settings.

Subgenome Contributions and Polyploid Origin (RADseq, Target Genes, and CP)

We applied criteria for building consensus results on previously generated genetic structure and phylogenetic network results (details in legend of Table S7). We mainly assessed the parental subgenome contributions per polyploid individual as follows: (i) take the most abundant parent within a column; (ii) if there were two equally abundant parents (e.g., two-times sexual progenitor subgenomes C and F) within a column, both parental subgenome contributions were taken for the consensus result (e.g., C/F); (iii) if two parental subgenome contributions within a column existed, we included them with a value of '0.5' (instead of '1.0') in consensus calculations.

To validate the obtained consensus results and to infer genome evolution (tree-like, autopolyploid vs. network-like, allopolyploid), we submitted all previously generated results (before consensus results building) to the full likelihood approach implemented in PHYLONET. The CALCPROB function calculates the likelihood of gene trees under a given species network and thus the total likelihood of the same network. Thus, we employed the gene trees used in the network analyses based on the target enrichment dataset mentioned above.

To include RADPAINTER+FINERADSTRUCTURE and STACEY results, networks were manually constructed using the tree backbone topology in Karbstein et al. (2020b) and the first two putative progenitors identified by these methods (Table S7). The autopolyploid scenario was tested utilizing the ASTRAL III trees already used as starting tree for the PhyloNetworks analyses. We rooted all networks with *R. cassubicifolius* s.l. to make scenarios more comparable. To compare tree-like (autopolyploid) scenarios with network-like (allopolyploid) ones, we scored results using the Akaike Information Criterion (AIC), taking into account that the number of parameters in a tree/network is equal to the number of branch lengths plus (for the networks) the parental contributions (i.e., k = 8 and k = 13 for the tree and the networks, respectively).

We determined the final subgenome contribution(s) by correcting the consensus results by the previously generated full likelihood approach results of Phylonet and plastome (CP) results (Table S7). According to final results, we classified the origin of polyploids, and the number of subgenomes involved in polyploid formation.

SNP Discovery (RADseq)

To investigate post-origin evolution of allopolyploids in more detail (H_1–H_6, H_8, and H_{10}; Figs. 1, 3), we carried out SNIPLOID (v. 17th March 2016; Peralta et al., 2013) analyses mainly following the workflow of Wagner et al. (2020) (see bash-script on Github). SNIPLOID compares the genome of an allotetraploid and a diploid putative parental species (DIPLOID2) with a diploid parental reference (DIPLOID1). The resulting SNPs were categorized: cat 1&2 result from post-origin interspecific hybridization, e.g., backcrossing to the parental species; cat 3&4 represent post-origin lineage-specific SNPs (not present in the parents); cat 5 represents the homeo-SNPs from the hybrid origin from the two parents (Peralta et al., 2013; Wagner et al., 2020). For example, a first-generation hybrid is expected to have only homeo-SNPs inherited from the parental species (cat 5), and no interspecific SNPs (cat 1,2) or derived SNPs (cat 3/4).

We created references of diploids by merging all accessions of a single progenitor species into a single *.fastq file (all possible parental SNPs of genetically close individuals of a species have to be covered) and conducted within-sample clustering in IPYRAD (filtering and clustering settings identical to Karbstein et al., 2020b). Obtained consensus files were used as DIPLOID1 (reference) and merged *.fastq files as DIPLOID2. We specified a minimum read depth per position of 20 (default; majority of positions showed more than 100 reads coverage). First, we excluded the category ‘others’ (heterozygous positions of DIPLOID2) from final results. High percentages of this category (30–64%) are probably due to multi-sample accessions and high individual heterozygosity in natural diploid populations. To address the influence of ‘others’, we evaluated this category by splitting heterozygous positions (REF and ALT), and categorizing the remaining ALT SNPs of DIPLOID2 and REF SNPs of DIPLOID1 according to SNP categories of SNIPLOID. Moreover, we always observed a dominance of interspecific SNPs of cat 2 compared with cat 1 SNPs, independent of parental combination. This was probably due to neglection of natural genetic variation in the majority rule base call references. Therefore, we generally summarized both categories to ‘cat 1&2’ to avoid biases within interspecific SNP categorization.

Results

Phylogenetic Tree Analyses unraveled Five Main Clades and showed Large Congruence among Datasets

For sexual and asexual *R. auricomus* individuals across Europe, we generated genomic RADseq, nuclear target enrichment gene, and plastome (CP) data based on 97,312 loci (280 individuals), 576 genes (113 individuals), and 71 regions (87 individuals), respectively. Both ML (RADseq) and coalescent-based (nuclear genes) phylogenetic tree analyses revealed five main clades (I–V). BT support of tree 'backbones' was generally high (most FBP/TBE values 90–100, Fig. 4a,b). Clades were well supported (FBP/TBE values 70–100), but particularly FBP support of clades IV and V for the ML (39–40 vs. 98–99 TBE support) and of clade V for the coalescent-based tree (27) was very low. Within clades, BT support was very low or absent (most FBP/TBE values 0–80/40–90; Fig. S4a–c).). Each clade contained one sexual species and polyploid taxa of various geographical origins (ML and ASTRAL taxon names, respectively): (I) *R. cassubicifolius* (subgenome C) with 13 and 16 tetra- and hexaploid samples, (II) *R. flabellifolius* (subgenome F) and with 4 and 6 tri- to hexaploid, (III) *R. marsicus* (subgenome M) and one tetra- to hexaploid, (IV) *R. notabilis* (subgenome N) and with 10 and 17 tetraploid, and (V) *R. envalirensis* (subgenome E) and with 21 and 40 tetraploid taxa. Whereas clades I and II were predominantly characterized by undivided basal leaf types, clades III–IV exhibited only dissected ones.

ML tree nodes (RADseq) were highly informative (QI = 0.85–1, Fig. 4a). Quartet concordance metrics (QC) for RADseq and main topology for target genes (MT = Q_1) showed highly concordant patterns with almost no alternative topologies (QC = 1, QD=-/MT = 0.98, Q_2 = 0.01. Q_3 = 0.01, Fig. 4a,b) for the node splitting clade I and all remaining ones. All remaining nodes showed moderately to highly conflicting signals with varying distribution of alternative topologies (QC = -0.06–0.33, QD = 0.29–0.73/MT = 0.32–0.56, Q_2 = 0.23–0.38, Q_3 = 0.21–0.30), particularly the nodes splitting main clades IV and V (QC=-0.06)/M+N and E (MT = 0.42), supported by lowered BT values (72–98). The sample composition of clades was highly similar between the different approaches.

Plastome Phylogeny showed Incongruences with Nuclear Data

The ML tree based on plastome data (CP) revealed four well-supported main clades with FBP = 100 (i.e., haplotype groups; Fig. 5a, Fig. S12a–d). In general, within-clade (within-haplotype-

group) relationships were mainly low or not supported (FBP << 70). Clade I consists of haplotypes from *R. cassubicifolius*, *R. flabellifolius*, and various polyploids. Within the first clade, accessions of *R. cassubicifolius* and *R. flabellifolius* were completely intermingled, contrary to nuclear datasets (Fig. 4). Clade II contained only haplotypes from polyploid taxa. The remnant two haplotype clades III and IV consisted of *R. envalirensis* and few polyploid accessions, and *R. notabilis*, *R. marsicus*, and various polyploids, respectively. Interestingly, accessions of the diploid *R. notabilis* and the tetraploid *R. marsicus* are intermingled, indicating that they belong to the same haplotype group (Fig. S12b), contrary to nuclear data (Fig. 4). The splitsgraph of the neighbor-net analysis also exhibited four, weakly differentiated, clusters (Fig. 5b, Fig. S13). The same is true for the TCS haplotype network (Fig. S14).

Genetic Structure Analyses indicated Three to Five Clusters, Strong Reticulation, and a Geographical Pattern

Structure analyses based on RADPAINTER+FINERADSTRUCTURE revealed three supported main clusters (Fig. 6a). Sexual species were also clustered with polyploids: (I) *R. cassubicifolius* (C) and tetra- to hexaploid taxa, (II) *R. flabellifolius*, *R. marsicus*, and *R. notabilis* (F, M, and N) and tri-to hexaploid taxa, and (III) *R. envalirensis* (E) and tetraploid taxa. Commonly, polyploids showed high coancestry values, i.e., orange to red colors, with different clusters indicating reticulation events (see particularly polyploids H_1–H_{10}). In addition, highest values were found in relation to the sexual subgenomes occurring in the same cluster (Table S7). Polyploids of cluster I showed highest similarity values with C and lowest ones with N and F. In contrast, cluster II is genetically more heterogeneous (subclusters IIa, IIa). Polyploids shared high similarity values with N and low coancestry values with F, E, and C. In cluster III, polyploids only exhibited high similarity to E.

Structure analysis based on STACEY revealed similar results (Fig. 6b, Table S7). Sexuals are also surrounded by polyploids, and polyploids showed several reticulations and highest posterior probabilities with intra-cluster sexual subgenomes. There are few differences here: the former cluster II is divided into three distinct clusters each containing a single sexual species (II–IV), and many polyploids of the former cluster IIa are incorporated into cluster V. In addition, polyploids of cluster I also revealed significant posterior probabilities to E, and polyploids of cluster V to F, C, and N. In general, subgenome M shared no significant coancestry/posterior probability values with polyploids.

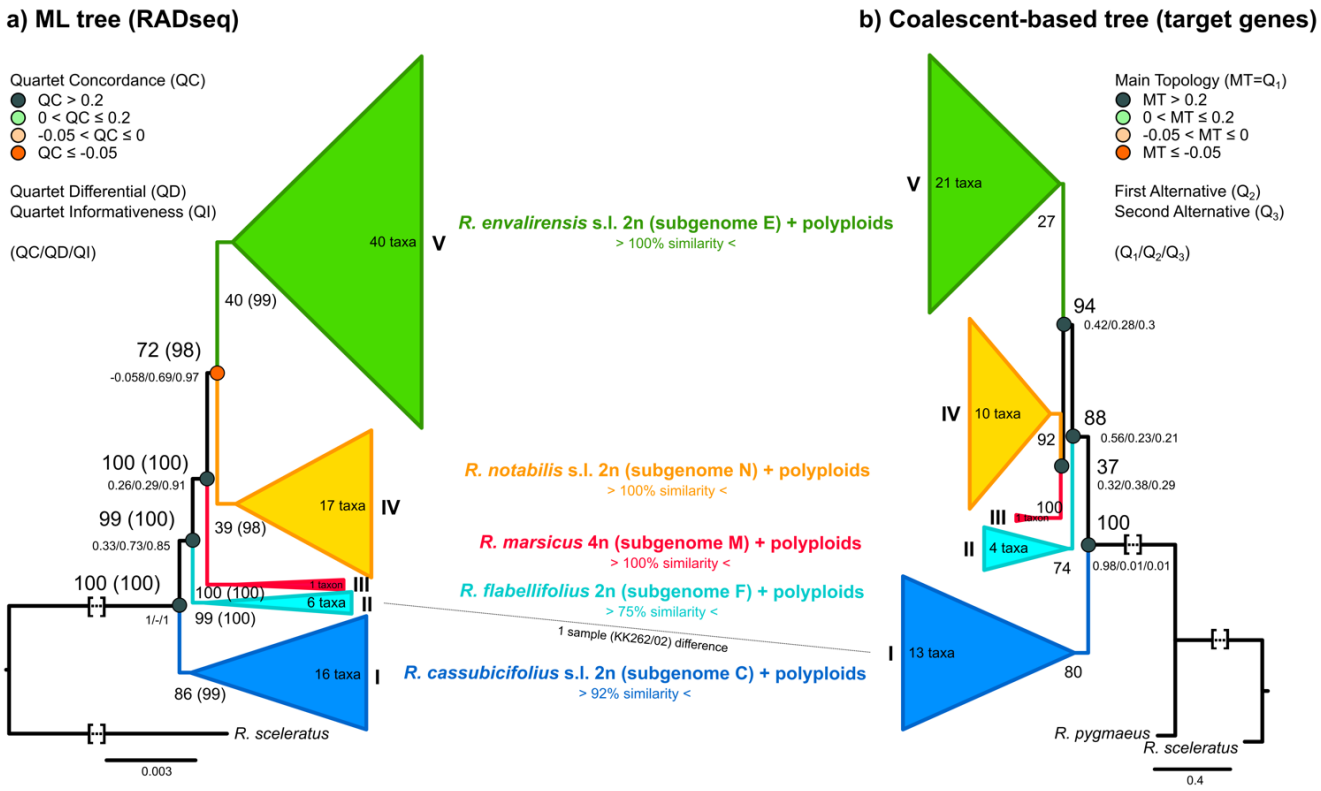


Figure 4. Phylogenetic trees based on RADseq and target enrichment data. (a) a ML tree based on RAXML-NG results and a min10 RADseq alignment (280 samples, 97,312 loci, 438,775 SNPs) and (b) a coalescent-based tree based on ASTRAL results and a target enrichment alignment (113 samples, 576 non-phased genes). For Fig. 4a, FBP, TBE, and quartet sampling scores (QC/QD/QI; see Fig. S3a-d and legend for explanations) are displayed per branch. Nodes are colored according to QC values (legend on the left). For Fig. 4b, FBP and quartet support scores (MT = $Q_1/Q_2/Q_3$) values are shown per branch. In general, only supported main clades (I-V, FBP/TBE>70; except Fig. 4b split between *R. notabilis* s.l. and *R. marsicus*) are illustrated because of mostly low/no BT support (BT<70) within clades. Each clade contains a sexual species and several polyploids. Number of polyploid taxa is given per main clade. We calculated sample composition between RADseq and target enrichment clades (only shared samples were evaluated due to different sample sizes), and illustrated values in the central part of the figure (0% = main clades are composed of completely different samples, 100% = main clades are composed of completely equal samples). Dotted lines show differences between clades of both datasets. See Figs. S1, S2, S4a-c and Figshare data repository for more details. Squared brackets: A part of the branch was cut for illustrative purposes.

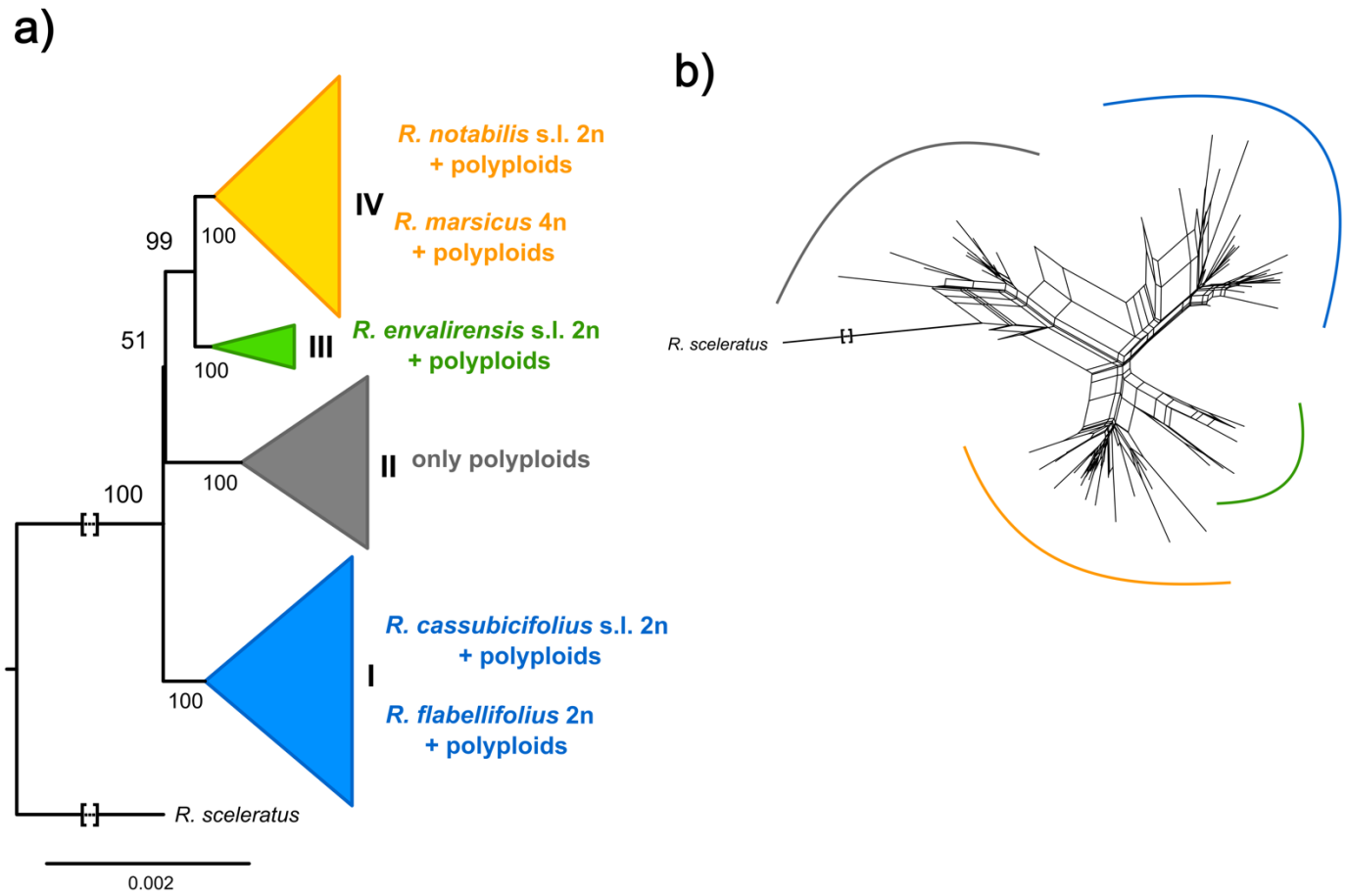
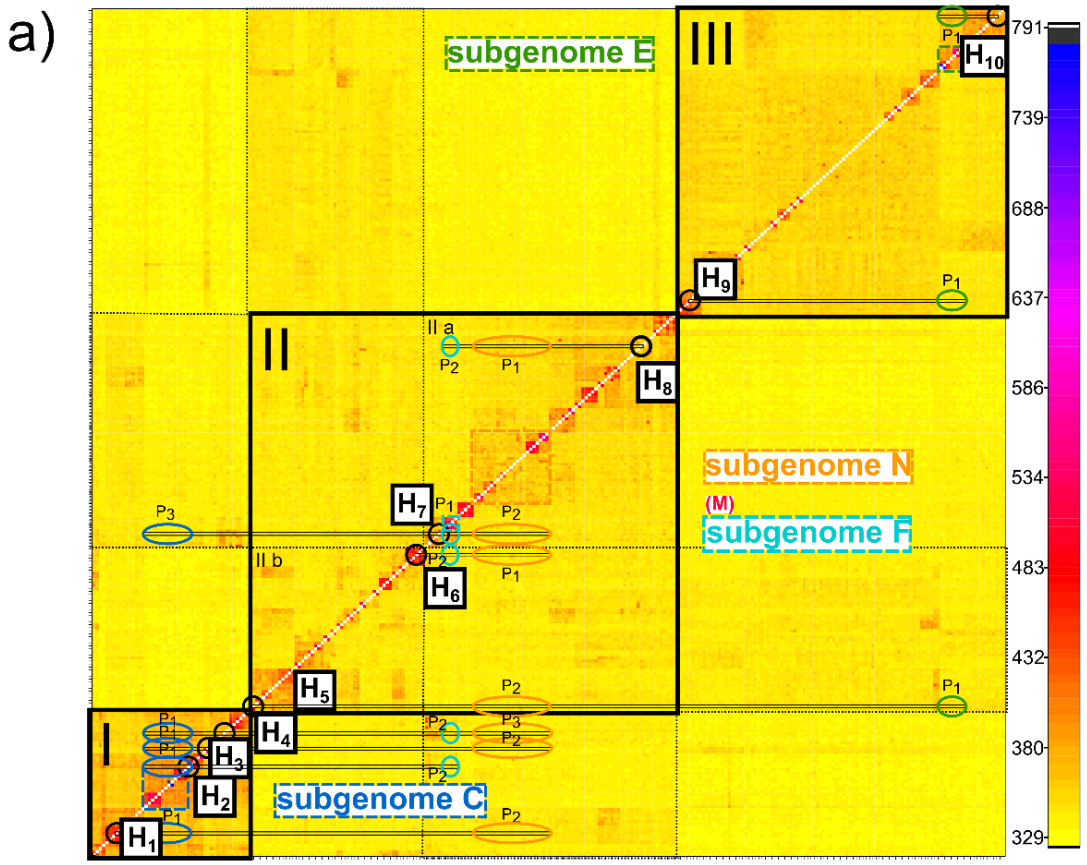


Figure 5. Phylogenetic tree and genetic structure based on plastome (CP) data. (a) ML tree (RAXML-NG) based on 87 samples and 71 plastid regions of the plastome (CP) dataset. Only main clades containing sexual species are shown (coloring according to Fig. 4). Concerning the clade in grey, plastid types of asexual polyploids were not found in any of the sexual species suggesting the former existence of a nowadays extinct sexual progenitor species. FBP values are given for each branch and clade (I-IV). (b) Neighbor-net analysis (SplitsTree) based on genetic distances (general time reversible [GTR] model with estimated site frequencies and ML), 87 samples, and 71 plastid regions of the plastome (CP) dataset. We colored main splits according to Fig. 4a. See Figs. S5, S6 for more details. Squared brackets: A part of the branch was cut for illustrative purposes.


a)
subgenome E
III
II
II a
II b
I
subgenome N
(M)
subgenome F
subgenome C
H1
H2
H3
H4
H5
H6
H7
H8
H9
H10
P1
P2
P3
791
739
688
637
586
534
483
432
380
329

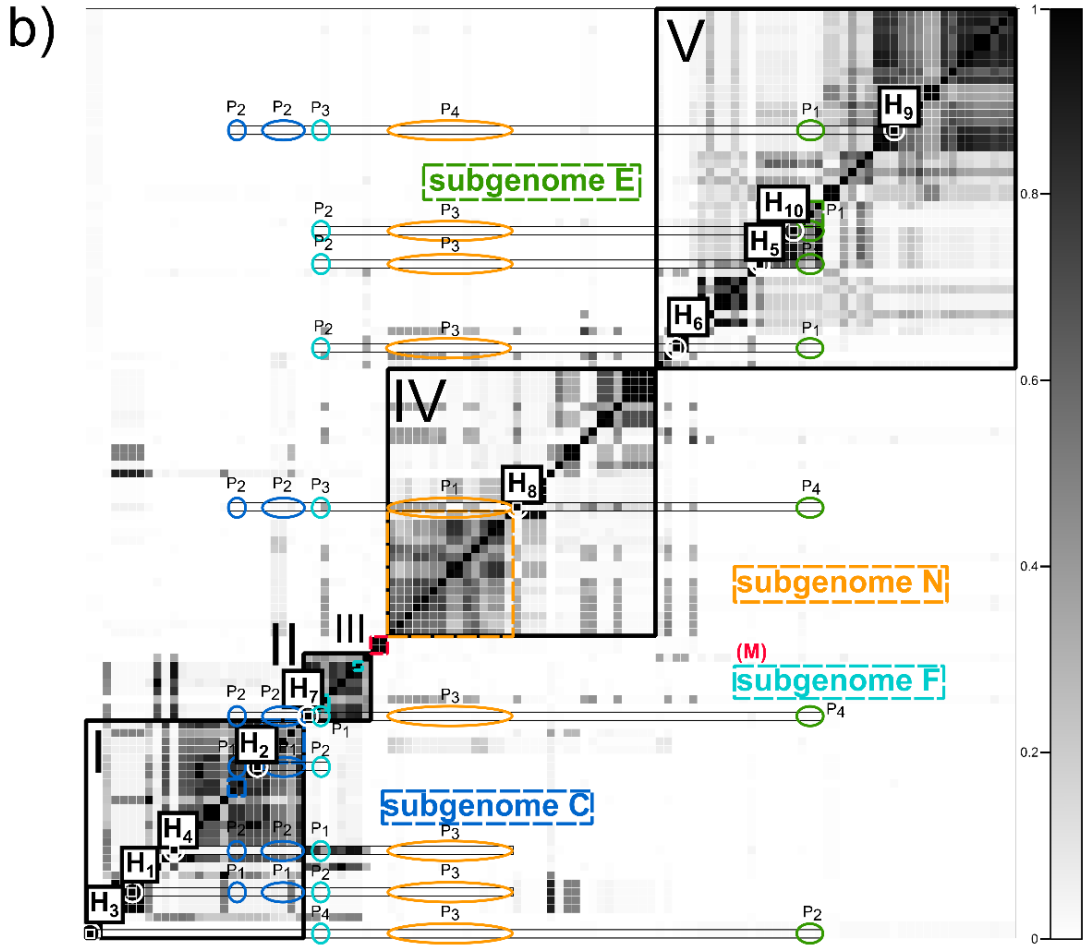

b)
V
IV
III
II
I
subgenome E
subgenome N
(M)
subgenome F
subgenome C
H1
H2
H3
H4
H5
H6
H7
H8
H9
H10
P1
P2
P3
P4
1
0.8
0.6
0.4
0.2
0

Figure 6. Genetic structure analyses based on RADseq and target enrichment data. (a) Clustered RADPAINTER+FINERADSTRUCTURE coancestry matrix of 280 sexual and polyploid apomictic individuals of the *R. auricomus* complex based on the 'min10' RADseq alignment (97,312 loci, 438,775 SNPs). The legend on the right shows the color-coding of genetic similarity (coancestry values): the darker the square, the higher the similarity between a pair of individuals. (b) Similarity matrix of STACEY species delimitation analyses of 113 individuals based on a phased target enrichment alignment (48 genes). Posterior probabilities for belonging to the same cluster (species) are shown for pairs of individuals in the legend on the right: black is for 1.0 posterior probability and white for 0.0. See Fig. 4b, Figs. S4a-c, S10 and high resolution figures on Figshare for clustering structure (a) tree with posterior probability group assignment probabilities and (b) coalescent-based tree. We indicated supported genetic clusters with solid lines, shared similarity among (sub)clusters with dotted lines, and sexual species with broad dashed colored squares (subgenomes C, F, M, N, and E). Small black squares ('H_n') indicate selected tetraploid apomicts, which were investigated for allo- vs autopolyploid origin (10 polyploids; a small square indicates the analyzed individual; see IDs in Table S1). Using lines and colored circles/ellipses, we highlighted the potential parental subgenome contributions for each polyploid (P_1, P_2, … and P_n with n = the n-th parental subgenome contribution. P_1 is always the parental subgenome contribution with the highest coancestry score/posterior probability (likeliest) followed by other parental/subgenome contributions with decreasing coancestry scores/posterior probabilities (minor parental/subgenome contributions not drawn, see Table S2 for more details).

Structure analyses based on SNMF also unraveled three to four (up to five) main clusters (Fig. 7a–d, Figs. S6–11, S15). Although polyploids were characterized by a dominant genetic partition, they also showed 1–3 minor genetic partitions (Fig. 7b,d, Fig. S7a–c). The likeliest number of K (clusters), K = 3, showed a west-east distribution of clusters across Europe (Fig. 7a,c). The clusters themselves are north-south distributed. *Ranunculus envalirensis* and related polyploids (E, green partition) mainly inhabit regions in southwestern, central, and northern Europe. *Ranunculus flabellifolius*, *R. marsicus*, and *R. notabilis* and related polyploids (F, M, and N, orange partition) predominantly occupy southern, central-eastern, and northern Europe. *Ranunculus cassubicifolius* and related polyploids (C, blue partition) range from southeastern to northern Europe, including a disjunct distribution in central Europe. When comparing results of K = 3 and K = 4, the only remarkable difference is the emergence of a genetic cluster in central and northern Europe without a sexual species (grey partition) out of the former green one (Fig. 7a,c). In general, SNMF results are comparable to all previous analyses (grey partition predominantly found in clade V (Fig. 4a,b)/cluster III (Fig. 6a)/cluster V (Fig. 6b), and the orange partition mostly situated in clade II–IV (Fig. 4a,b)/cluster III (Fig. 6a)/cluster II–IV (Fig. 6b). The splitsgraph of the neighbor-net analysis (RADseq) also exhibited three main genetic clusters weakly differentiated from each other (Figs. S16a and Fig. 6a).

Phylogenetic Networks supported by Genetic Structure revealed Allopolyploidy, Two to Three Contributing Subgenomes, and Subgenome Dominance

For most tested polyploids, phylogenetic networks based on RADseq and target enrichment datasets showed two different subgenome contributions (Figs. 8a–h, 9, Tables 1, Table S7). These polyploids were usually characterized by a dominant (P_1 = 51–99%, mean 74%) and a minor subgenome contribution (P_2 = 1–49%, mean 26%). Concerning PHYLONET likelihood+AIC calculations, reticulate evolution and thus allopolyploid origin was confirmed in most cases (H_1–H_6, H_8, and H_{10}). Within clade I and cluster I (Figs. 4, 6), polyploids H_1–H_4 possessed the dominant subgenome C whereas minor ones came from F followed by E and N. The blue haplotype C+F of these polyploids matched the dominant subgenome C. Final results indicated that '*R.* × *platycolpoides*' (H_1) is composed of subgenomes C and N, '*R.* × *elatior*' (H_2) of C and F, '*R.* × *pseudocassubicus*' (H_3) of C and E, and '*R.* × *hungaricus*' (H_4) of C and F.

Moreover, we inferred varying subgenome contributions for the polyploid '*R.* × *pilisiensis*' (H_7) of clade II (Fig. 4a,b) and cluster II/III (Fig. 6a/b), but consensus results supported by CP results revealed subgenome F as the dominant one. The likeliest scenario is a tree-like evolution (autopolyploid origin). The polyploid '*R.* × *indecorus*' (H_8), positioned in clade IV (Fig. 4a,b) and cluster II/IV (Fig. 6a/b), showed three subgenomes, whereas N was the slightly dominant one, and C and N the minor ones. '*Ranunculus* × *fissifolius*' (H_5) was characterized by the orange haplotype N and is also composed of three different subgenomes (E, F, and N).

The polyploids '*R.* × *glechomoides*' (H_6), '*R.* × *subglechomoides*' (H_9), and '*R.* × *leptomeris*' (H_{10}) exhibited subgenome E as the dominant contribution. In most tree and genetic structure analyses, these polyploids were also situated close to E. CP analyses showed the green haplotype E for H_6 and H_{10}, but not for H_9. Final results indicated E and F subgenome contributions for H_6 and H_{10}. '*Ranunculus* × *subglechomoides*' (H_9) exhibited the grey, unknown haplotype U and Phylonet AIC+likelihood calculations detected similarly-likely scenarios of reticulate (E and F) or tree-like evolution (E) (Table 1, Fig. 9).

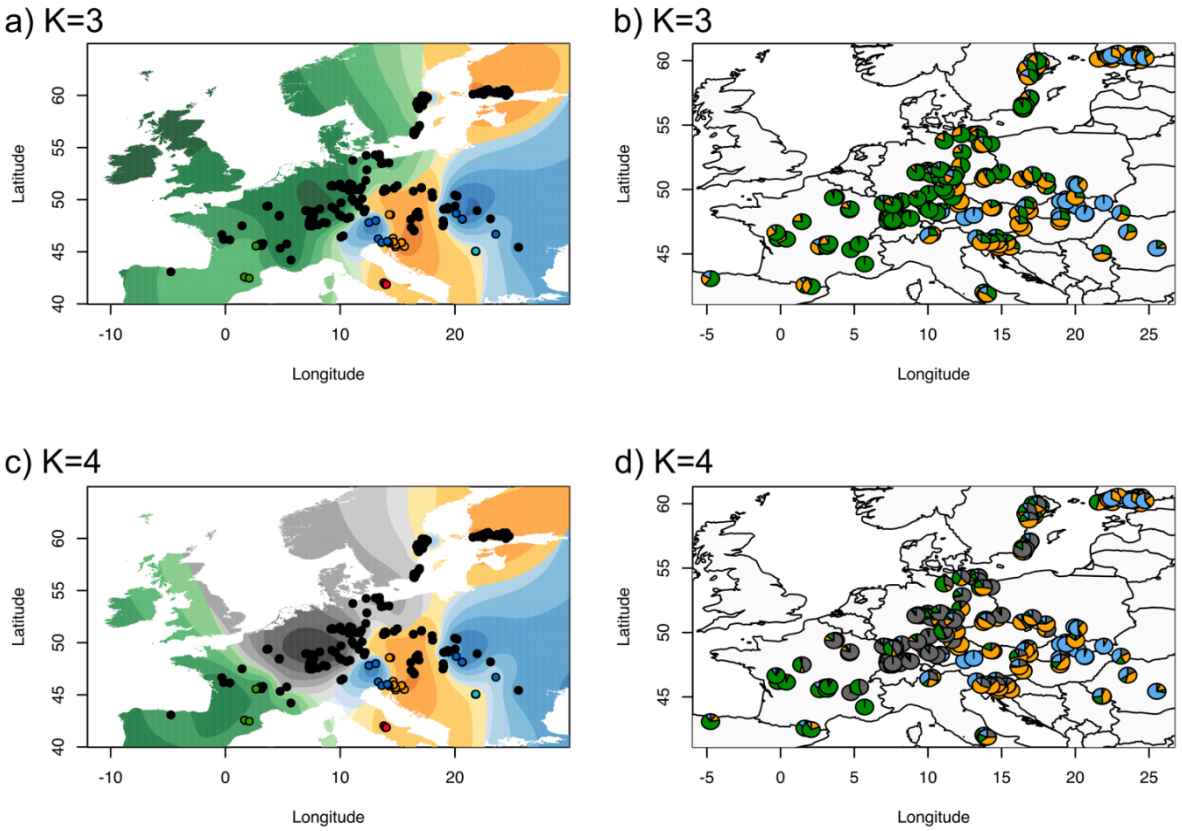


Figure 7. Geographic maps showing genetic clusters and ancestry coefficients across Europe. (a, c) interpolated values of ancestry coefficients (method 'max', i.e., at each point the cluster for which the ancestry coefficient is maximal) and (b, d) location-wise admixture estimate pie charts using K = 3 and 4 genetic clusters. Results are based on sNMF results of 280 sexual and apomictic *R. auricomus* individuals and the 'min30' unlinked-SNP RADseq alignment (33,165 loci). See Figs. S6-11, S15, and figures on Figshare. In (a) and (c), colored circles represent sexual species (coloring according to Fig. 4): blue = *R. cassubicifolius* s.l. (C), turquoise = *R. flabellifolius* (F), red = *R. marsicus* (M), green = *R. envalirensis* s.l. (E), and orange = *R. notabilis* s.l (N). We adopted the coloring also to pie charts in (b) and (d). Europe map source: https://maps.ngdc.noaa.gov.

Table 1. Genetic structure and phylogenetic network results of tested tetraploid *R. auricomus* accessions (H_1-H_{10}). Each row (H_1-H_{10}) represents a separately analyzed individual. Results are based on RADseq (RADPAINTER+FINERADSTRUCTURE, PHYLONETWORKS) and phased nuclear target enrichment gene (STACEY, PHYLONETWORKS, PHYLONET) datasets. Consensus results summarize all previously gained information. Final results indicate final subgenome contribution(s), i.e., consensus results corrected by the full likelihood approach+AIC calculations in Phylonet (likel+AIC, AIC = Akaike Information Criterion; more than one result if AIC network difference was less than 10 units) and plastome analysis results (CP type; C/F = plastid type shared by the diploid sexual species *R. cassubicifolius* and *R. flabellifolius*, * = not the same sample between CP and network analyses, # = haplotype from an unknown/extinct sexual progenitor species of Central Europe). Concerning the final results of H_9, we classified P_1 as "E(U)" because of the *R. envalirensis*-like U plastid type (see above). According to final results, we classified genome evolution of investigated polyploids (allo = allopolyploid, auto = autopolyploid), and the number of involved subgenomes in polyploid formation. See also Figs. 5, 6, 8, Tables S1, S7, and data on Figshare for sample IDs, genetic structure, and network results.

Analysis	**H_1** *"R. x platycolpoides"*			**H_2** *"R. x elatior"*				**H_3** *"R. x pseudocassubicus"*				**H_4** *"R. x hungaricus"*				**H_5** *"R. x fissifolius"*			**H_6** *"R. x glechomoides"*			**H_7** *"R. x pilisiensis"*				**H_8** *"R. x indecorus"*				**H_9** *"R. x subglechomoides"*				**H_{10}** *"R. x leptomeris"*		
	P_1	P_2	P_3	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4	P_1	P_2	P_3
consensus results	C	N	F	C	F			C	E	F	N	C	F	N		E	E F	N	E	F	N	F	F	C N	E	N	C	F	E	E	C F	F	N	E	E F	N
likel+AIC (PhyloNet)	reticulate (N, C)			reticulate (C, F)				reticulate C = E)				reticulate (F, C > C, C)				reticulate (3x E, F)			reticulate (E, F)			tree-like (F)				reticulate (E, C > N, C)				reticulate (E, F) > tree like (E)				reticulate (E, F)		
CP type	C		F	C		F		C		F		C		F		N*			E			C*		F*		N				U#				E		
final results	C	N		C	F			C	E			C	F			E	F	N	E	F		F				N	C	E		E	(F)			E	F	
genome evolution	allo			allo				allo				allo				allo			allo			auto				allo				allo vs. auto				allo		
no. sub-genome/s	2			2				2				2				3			2			1				3				1–2				2		

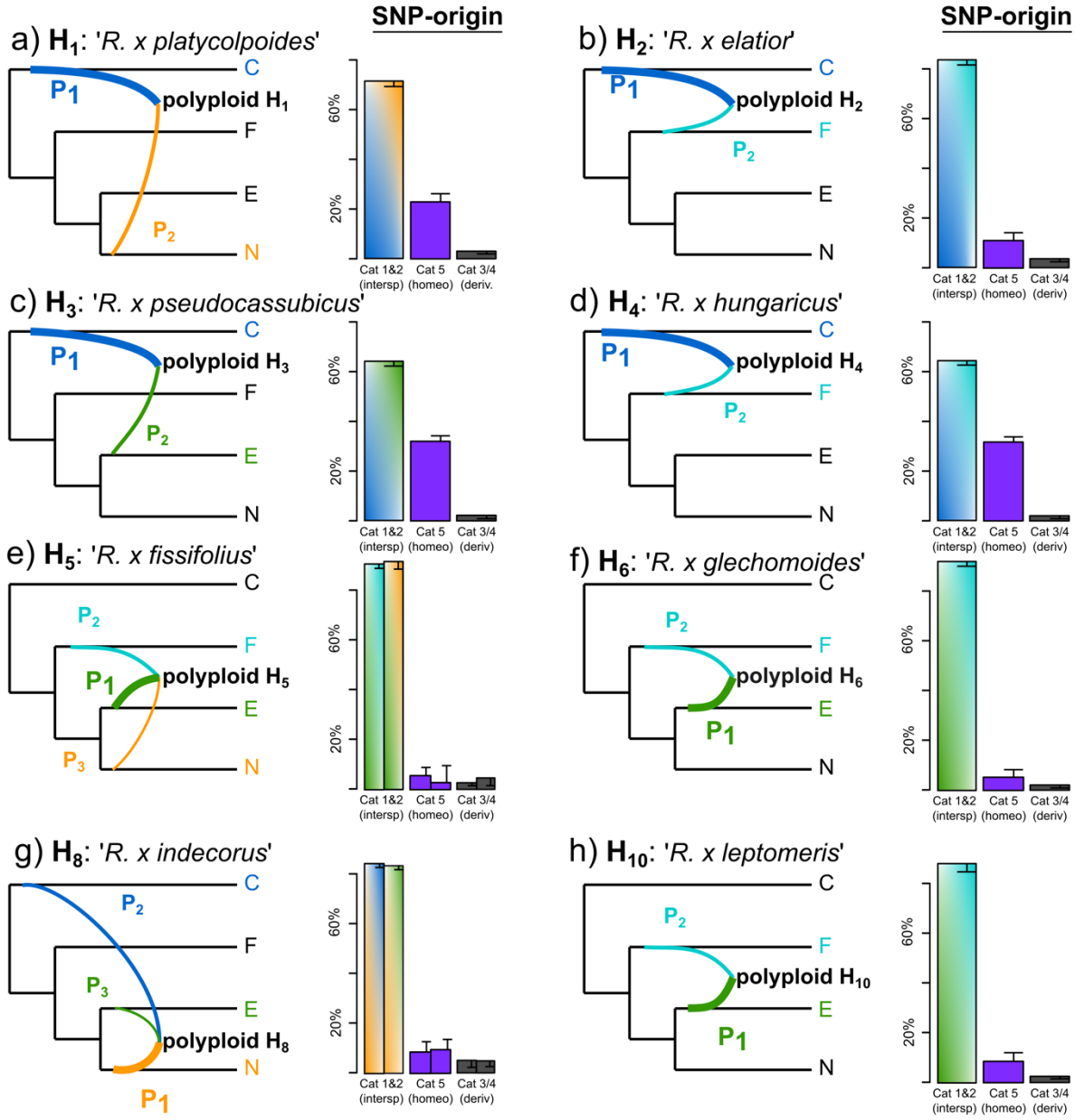


Figure 8. Reconstructed phylogenetic networks based on RADseq, target enrichment, and CP data. (a-h, left) Final networks of allopolyploids are based on genetic structure and phylogenetic network results (consensus results) corrected by the by the full likelihood approach+AIC calculations in Phylonet and CP data. P_1 defines the largest subgenome contribution, followed by P_2 and P_3. The network topology follows the published rooted phylogeny of *R. auricomus* sexuals (without tetraploid *R. marsicus*; Karbstein et al., 2020b). Curves indicate subgenome contributions (P_1-P_3). (a-h, right) Bar charts based on SNiPloid results are shown. Bar charts show SNP origins in percents (cat 1 = SNPs identical to DIPLOID2, cat 2 = SNPs identical to DIPLOID1/reference, cat ¾ = derived SNPs, cat 5 = homeo-SNPs. We highlighted SNPs percent concerning all SNPs (see Material and Methods for additional evaluation of SNP category 'others') with additional black T-bars. Concerning H_5 and H_8, we calculated two SNiPloid analyses because three parents have contributed to its origin. Coloring of sexual progenitor subgenomes is according to Fig. 4. Subgenomes of C = *R. cassubicifolius* s.l., F = *R. flabellifolius*, E = *R. envalirensis* s.l., and N = *R. notabilis* s.l.

SNP discovery (SNIPLOID) based on RADseq data supported allopolyploid hybrid origins with 3–33% (9–36% with evaluated others) homeo-SNPs of cat5. Whereas polyploids H_1, H_3, and H_4 showed relatively high percentages of homeo-SNPs (>20%), the polyploids H_2, H_4–H_{10} exhibited low amounts (<15%). The majority of SNPs, however, indicated considerable post-origin evolution of allopolyploids (Fig. 8a–h, Table S8). Across datasets, SNIPLOID assessed 64–93% (62–89% with evaluated 'others') interspecific SNPs of cat1&2, and 3–5% (2–3%) derived SNPs of cat3/4. Interspecific SNPs of cat1&2 were lowest for polyploids H_1, H_3, and H_4 and highest for H_2, H_4–H_{10}.

Discussion

Polyploid phylogenetics is an emerging and bioinformatically challenging field, with important consequences for understanding plant speciation and macroevolution (Soltis et al., 2015; Landis et al., 2018; Rothfels, 2021). Here, we used a comprehensive genomic, nuclear gene, and plastome dataset to unravel evolutionary processes of a less than 1 Mya polyploid species complex. Different kinds of evidence included in our self-developed bioinformatic pipeline confirmed that hybridization of sexual progenitors followed by polyploidization (allopolyploidy) is the dominant formation type in our model system (Table 1). Allopolyploidy also shaped the evolution of many other young polyploid complexes (Sochor, 2015; Spoelhof et al., 2017; Dauphin et al., 2018; Rothfels, 2021). In addition, we also demonstrated remarkable post-origin genome evolution of allopolyploids, mostly due to interspecific gene flow. The bioinformatic pipeline presented here disentangled the parental contributions and the genomic diversity and composition of different polyploid apomictic lineages that have evolved. The *Ranunculus auricomus* model system is the first well-studied, large polyploid European species complex using OMICS data. The major conceptual breakthrough presented here is the combination of several datasets and up-to-date NGS methods into a new pipeline, starting with the diploid progenitors and ending up with the polyploid derivatives, to receive, for the first time, a complete picture of the evolution of a large polyploid complex. The novel aspects particularly comprise network analyses, consensus result making of previous structure and network results, and auto- vs. allopolyploid testing.

The Phylogenetic Pattern

Phylogenetic trees based on RADseq and nuclear gene data surprisingly revealed only five well-supported main clades (Fig. 4a,b). Congruence between data sets hints at a strong evolutionary signal regardless of analyzing anonymous genomic regions (RADseq) or only coding nuclear genes, and regardless of applying different analytical approaches (concatenation and ML vs. coalescence). Since the two marker sets complete each other to some extent (see Materials & Methods), we infer a robust phylogenetic framework for all further analyses.

Each main clade contained a sexual species surrounded by asexual polyploid lineages, showing that polyploids largely derived from their clade-specific progenitors (Fig. 4). This pattern is rather unique compared with other polyploid complexes, where usually clades with several diploid and (allo)polyploid taxa, but also clades with only polyploids were found (Kirschner et al., 2015; Dauphin et al., 2018; Carter et al., 2019; Wagner et al., 2020). Despite well-supported tree backbones according to BT values and highly informative branches (Fig. 4a), quartet support is partly low and distribution of alternative topologies (QD, Q's; Fig. 4) hints at both inter-clade reticulation and ILS. Particularly within main clades, BT values are extremely low and quartet support metrics (QD, Q's) indicate high rates of introgression and ILS signals (Figs. S3, S4). Concerning the ML tree, TBE support was, especially at the backbone, higher than BT support probably due to less sensitivity of TBE to hybridization events and 'jumping taxa' (Lemoine et al., 2018). Reticulate evolution at the backbone of the tree is further indicated by the incongruent position of F polyploid derivatives in the plastid tree compared with the nuclear trees. The placement of diploid *R. flabellifolius* within the *R. cassubicifolius* plastid clade contrary to the nuclear trees suggests that *R. flabellifolius* could be an ancient homoploid hybrid species.

Within these main clades, the majority of described polyploid morphospecies are non-monophyletic (Figs. S1–S4a–c). No clear clade-/group-specific morphological trend is recognizable (Figs. 4–9), except that taxa with undivided basal leaves were mainly found in clades I+II. This incongruence with morphology also rejects the old Linnaean classification of two main morphotypes (previously also rejected for sexual progenitors in Karbstein et al., 2020b). Network-like evolution through hybridization (allopolyploidy) is well-known to cause severe conflicting signals in tree reconstructions (Lemoine et al., 2018; Pease et al., 2018; Rothfels, 2021). Moreover, cladogenetic speciation of *R. auricomus* allopolyploids from common polyploid ancestors is unlikely. Cladogenetic speciation would lead to bifurcating, tree-like post-origin evolution and thus only low conflicting signals (Jones, 2017a) at middle

and terminal branches. Here, these branches are extremely conflicting, suggesting extensive and repeated reticulate polyploid formation events. In contrast, coalescent-based phylogenetic analyses in the 20 Mya genus *Rubus* detected varying main clade positions, but highly resolved relationships within each clade (Carter et al., 2019). In brambles, ancient polyploidization events, strong geographical structure between continents, clade-specific and rather less post-origin evolution might explain this pattern. In *R. auricomus*, repeatedly ongoing hybridization and/or polyploidization of sexual progenitors during Pleistocene climate fluctuations in Europe and partly high facultative sexuality of polyploid apomicts in central and southern regions (Tomasello et al., 2020; Karbstein et al., 2021) potentially led to highly conflicting phylogenetic signals at middle and terminal branches.

Genetic Structure, Origin, and Parentage of Polyploids

Genetic structure also revealed 3–5 main clusters (Fig. 6), largely fitting the phylogenetic main clades. We detected more reticulation signals in RADseq data, probably favored by the incorporation of all SNPs of 97,312 loci, maximizing the information for analyzing young plant complexes. In contrast, analyses based on nuclear genes better inferred genetic boundaries by splitting each sexual progenitor species and accompanying polyploids into separate clusters. This advantage is probably related to the coalescent-based species delimitation using allelic information of the 48 phased loci (Jones, 2017b; Karbstein et al., 2020b; Tomasello et al., 2020). Interestingly, we observed a west-east distribution of clusters within Europe (Fig. 7a, see also Karbstein et al. 2020b, 2021): a western cluster related to *R. envalirensis* (E), a central cluster related to *R. flabellifolius*, *R. marsicus*, and *R. notabilis* (F, M, and N), and an eastern cluster related to *R. cassubicifolius* (C), each ranging from southern to northern Europe, respectively. West-east allopatric speciation of sexual progenitors in combination with both allopolyploidization events and north-south migration of populations due to past climatic changes may explain this pattern (Abbott et al., 2013; Tomasello et al., 2020). Moreover, we detected a subdivision of the western cluster (green, Fig. 7c,d). The grey, central European subcluster widely corresponds to the grey 'polyploid-only' haplotype of the ML plastid tree (Fig. 5a). The fact that ploidy levels of *R. auricomus* populations in central Europe are well-studied (Jalas & Sumoninen, 1989; Paule et al., 2018; Karbstein et al., 2021) makes it unlikely that an extant diploid was simply overlooked. Moreover, some polyploids possess a plastid type from an unknown diploid, suggesting an already extinct sexual progenitor related to subgenome E. Extinction of sexuals is a commonly considered or observed phenomenon in young polyploid

complexes shaped by past climatic deteriorations (Sochor et al., 2015; Rothfels, 2021) and is supported here by missing speciation events between 0.6 and 0.3 Ma (Million-years-ago; Tomasello et al., 2020). Alternatively, extinction could be due to past human activity.

Both genetic structure and phylogenetic network results based on RADseq and nuclear gene data revealed that the majority of polyploids were composed of two, surprisingly some of three, subgenome contributions (Figs. 6, 8, 9, Table 1). Plastome data underlined inferred subgenome contributions. We detected only one polyploid, that didn't show evidence of a reticulate evolutionary history: '*R.* × *pilisiensis*' (H_7). Here, varying subgenome contributions were found (Fig. 6, Table 1), but network analyses suggested autopolyploid origin from progenitor subgenome F. This lineage might also represent a segmental allopolyploid, as auto- and allopolyploidy are connected by transitions (Comai, 2005; Spoelhof et al., 2017). Autotetraploid cytotypes are also known from the otherwise diploid sexual species *R. cassubicifolius* (Hörandl & Greilhuber, 2002). However, autopolyploidy is currently only present in the *R. auricomus* complex in clades/clusters I and II, supporting its probably less frequent occurrence compared with allopolyploidy in nature (Spoelhof et al., 2017; but see many unnamed autopolyploids in Barker et al., 2016).

All diploid sexual progenitors were involved in allopolyploid formation (C, F, N, E, and U; Fig. 9). Polyploids were probably formed multiple times out of different progenitor combinations followed by considerable post-origin evolution. Close crosses result more easily in homoploids, whereas distant ones tend to become rather allopolyploid (Soltis & Soltis, 2009). However, we observed allopolyploids out of genetically distant (C+N, N+C+E), moderate (C+E, E+F+N), and close (C+F, E+F) crosses. Extant homoploid hybridization is probably inhibited by the allopatric distribution of sexual species (see Tomasello et al., 2020). Interestingly, allopolyploids showed more meiotic errors than homoploid hybrids in experimental crossings (distant C+N *R. auricomus* crosses; Barke et al., 2020). However, polyploids can escape hybrid sterility via apomixis, vegetative reproduction, and/or selfing (Hörandl, 2006; Soltis & Soltis, 2009; Barke et al., 2018, 2020).

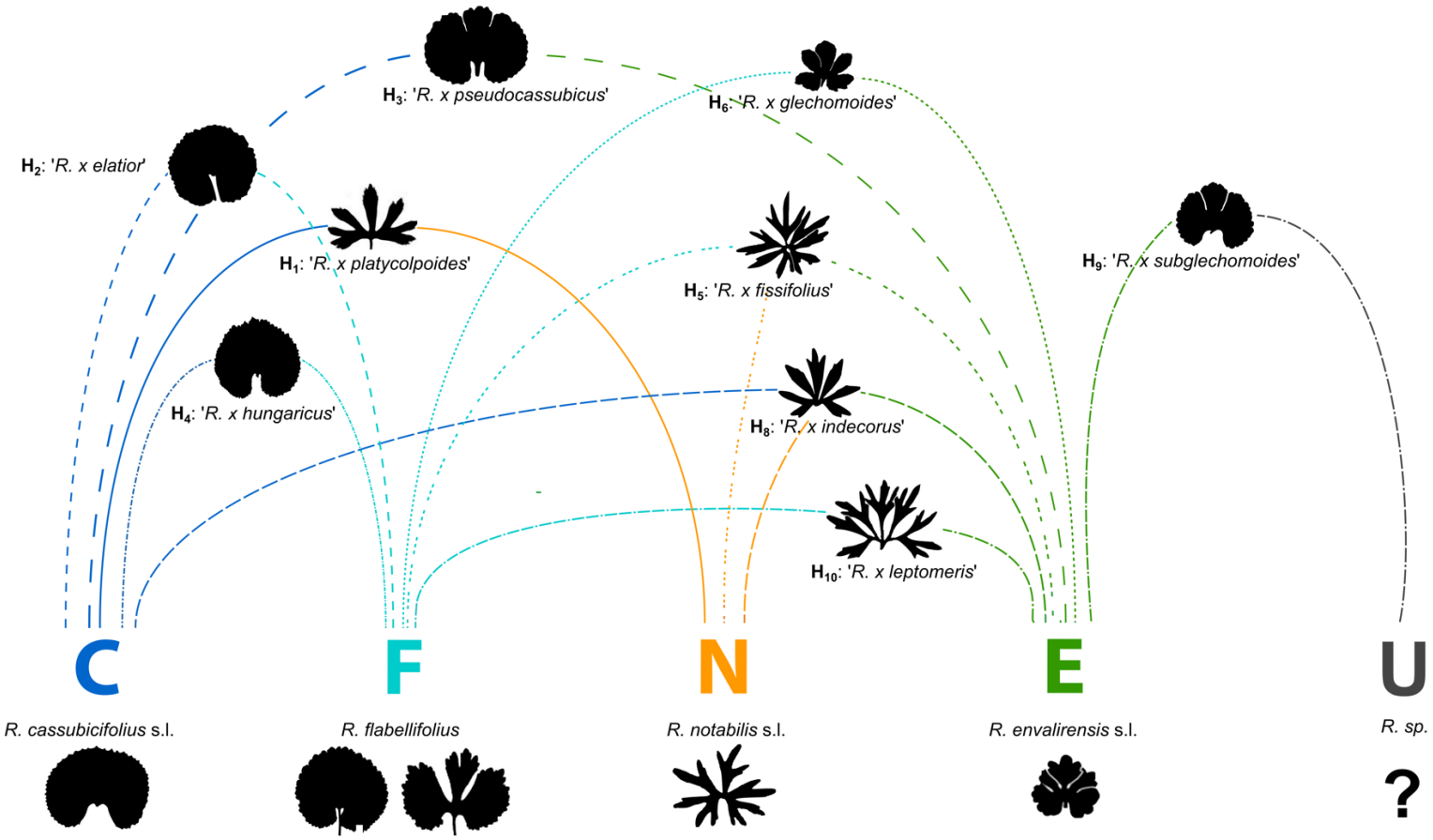


Figure 9. Hybrid scheme of sexual progenitors and selected polyploid *R. auricomus* derivatives. The diploid sexual progenitor species *R. cassubicifolius* s.l. (C), *R. flabellifolius* (F), *R. notabilis* s.l. (N), *R. envalirensis* s.l. (E), and a hypothetical unknown one (U) in different combinations gave rise to asexual polyploid derivatives (same polyploid individuals as in Fig. 8). Per allopolyploid, curves to the left and right indicate parental subgenome contributions. Subgenome dominance is shown by the relative position of the polyploid to the progenitors, for example '*R. x elatior*' is closer to subgenome C due to C subgenome dominance). We also illustrate characteristic basal leaf types of taxa during anthesis (variation not covered, two types illustrated for *R. flabellifolius* due to the frequent occurrence of undivided and divided types; source: herbarium type specimens). The hybrid scheme is based on phylogenetic network and genetic structure results of RADseq and target enrichment datasets supported by CP data (Table 1).

Extant sexuals have restricted ranges and are separated by thousands of kilometers across Europe (Fig. 2). Nevertheless, all main genetic clusters are present in central Europe (Fig. 7a, c). Sexual progenitors might have repeatedly met in this region during past interglacial times, giving rise to multiple allopolyploidization 'waves' with varying subgenome contributions. Interestingly, polyploids composed of three different subgenomes were only found in Central Europe and South Sweden (H_5, H_8), underlining the importance of secondary contact zones for the allopolyploid origin of the *R. auricomus* complex.

The genetic and phenotypic diversity of *R. auricomus* biotypes with more than 840 described morphospecies is probably formed by multiple allopolyploidization events from four extant and at least one probably extinct diploid sexual progenitor species. Other studies already demonstrated that few diploid progenitors were capable of producing a magnitude of allopolyploids, for example in *Botrychium* (Dauphin et al., 2018), *Rubus* (Sochor et al., 2015; Carter et al., 2019), or *Taraxacum* (Kirschner et al., 2015).

Post-Origin Evolution: Subgenome Dominance, Hybridization, and Mutation

Subgenome dominance was inferred in almost all tested allopolyploids and resulted in 74% mean inheritance probabilities from network analyses (Figs. 6, 8, 9, Table 1). Per main genetic cluster, allopolyploids were usually composed of dominant intra-cluster subgenomes and 1(-2) minor, varying inter-cluster subgenome(s), although trigenomic polyploids showed rather similar contributions (supported by SNP-origin analyses; Figs. 6, 8). Plastid data supported the dominant subgenome contribution and, in some cases, unraveled an additional or unknown/extinct subgenome (Table 1). Sequence subgenome dominance probably leads to the grouping of polyploids close to the dominant progenitor (Figs. 4, 6). Hence, polyploids of cluster I, II (II–IV), and III probably received at least one subgenome from C, N, and E(-like) progenitor lineages.

Subgenome dominance is a common feature of allopolyploids (Blischak et al., 2018; Alger & Edger, 2020). Here, segregation after hybridization, and after polyploidization, gene flow due to facultative sexuality of apomicts might cause the observed dominance. We consider biased fractionation of minor importance regarding the less than 1 Mya or even younger *R. auricomus* polyploids. For example, similar old *Brachypodium hybridum* showed only gradual gene loss and no subgenome expression dominance 1.4 million years after allopolyploidization (Gordon et al., 2020). The importance of diploid hybrid segregation and post-origin gene flow due to facultative sexuality is highlighted by SNiPloid analyses. We detected only a minority of SNPs

from hybridogenic origins (3–36%), but considerable proportions of interspecific, post-hybridization SNPs (62–93%). Since apomixis establishes only stepwise, the first diploid hybrid generations still exhibit predominant sexual reproduction (Barke et al., 2018), allowing for Mendelian segregation. The potential of Mendelian segregation in early hybrid generations for creating the nearly complete extant morphological diversity of the complex was demonstrated experimentally by Barke et al. (2018) and Hodač et al. (2018).

Moreover, facultative sexuality and maintenance of functional pollen probably allowed the newly formed hybrids for backcrossing with their parental species, and among each other. Extant polyploids under natural conditions exhibit usually low, but varying degrees of facultative sexuality (mean 2.15%, range 0–34%; Karbstein et al., 2021). Relatively high sexuality values of central and southern European asexual populations may indicate some ongoing gene flow. For example, in apomictic *Rubus* or *Pilosella* polyploids, which are also characterized by highly variable, facultative sexuality, post-origin scenarios by interlineage gene flow and backcrossing to sexual parents have also been considered (Sochor et al., 2015; Hörandl, 2018; Nardi et al., 2018; Carter et al., 2019). However, processes of fast segregation with backcrossing before vs. gene flow after polyploidization cannot be distinguished here, but we recognize here post-origin genome evolution as an important factor shaping genome structure of polyploids.

Only a few morphospecies appeared as monophyletic groups (e.g., H_2, H_6, H_7; see also Figs. S1–S3) and are evolving towards more stable lineages. The original hypothesis of Babcock and Stebbins (1938) predicted that such lineages would only form at higher ploidy levels (>6n, see Fig. 1). However, already Grant (1981) recognized that formation of stable apomictic lineages in a 'mature complex' is less dependent on cytotype but rather correlated to age, loss of sexuality, and extinction of sexual progenitors. According to Grant's (1981) definition, the *R. auricomus* complex is in an early mature stage of evolution, with extant diploid progenitors and a broad array of apomictic biotypes. This hypothesis is confirmed by the low proportion of lineage-specific SNPs (Fig. 8). SNP origin analyses revealed only 2–5% derived SNPs in allopolyploids attributable to mutations (Welch & Meselson, 2000; Pellino et al., 2013). Low degrees of post-origin sequence evolution are not surprising when comparing the evolutionary young *R. auricomus* polyploids to similar-old or older allopolyploids (Pellino et al., 2013; Gordon et al., 2020; Tomasello et al., 2020; Wagner et al., 2020). For example, several million years old *Salix* sexual allopolyploids exhibited 19–47% post-origin, species-specific SNPs (Wagner et al., 2020).

Integration of Datasets and Analyses for Unraveling Evolutionary Processes in Young Polyploid Complexes

Our case study demonstrates that even with OMICS approaches it is useful to rely on different complementary reduced-representation datasets to tackle polyploid complexes: genomic RADseq, nuclear genes, and plastomic regions. On the one hand, RADseq provided the highest number of (allelic) information (loci and SNPs) both from non-coding and coding regions. On the other hand, correct allele phasing and discrimination of homoeologues is desirable for polyploids (Eriksson et al., 2018; Freyman et al., 2020; Lautenschlager et al., 2020; Rothfels, 2021) but still a challenge for non-model plants. Both datasets together represent well the nuclear genome and can be much easier collected, cheaper sequenced, and easier analyzed for a large number of samples than entire transcriptomes or genomes (McKain et al., 2018; Johnsen et al., 2019). Moreover, target enrichment even allows the inclusion of young to old herbarium-type material (here, up to 74-years-old; up to 204 years in Brewer et al., 2019). This is particularly important in times of traveling restrictions and crucial for the correct application of taxon names in extremely morphologically diverse species complexes.

Here, we combined the advantages of three datasets to unravel evolutionary processes in polyploid complexes. We confirm previous approaches (e.g., Lo et al., 2010; Brandrud et al., 2020) demonstrating that a combination of tree building, structure, and network analyses is most useful to reconstruct non-hierarchical relationships in these complexes. The sexual progenitor species often diversify in a rather tree-like bifurcating manner that can be recognized with tree-building supported by genetic structure and/or morphometric methods (Burgess et al., 2015; Wagner et al., 2019; Karbstein et al., 2020b).

In our study, we demonstrate that the RADseq ML tree revealed a highly congruent topology compared with the target enrichment nuclear gene coalescent-based tree. These trees gave a first phylogenetic framework for the *R. auricomus* complex, as also shown in evolutionary young polyploid complexes of *Crataegus* (Lo et al., 2010) or *Dactylorhiza* (Brandrud et al., 2020). Here, the low interclade and extremely low intraclade (quartet) support values indicated the presence of reticulations and/or ILS. Nevertheless, allopolyploids originated by diverged progenitor species introduce errors in ordinary tree reconstructions due to network-like evolution and smushing of different evolutionary histories in consensus sequences (McDade, 1992, Oxelman et al., 2017, Rothfels, 2021). We regard this issue in our study as minor, since progenitors of polyploids are genetically less diverged (Karbstein et al., 2020b), tree and genetic

structure analyses show comparable results, and subgenome dominance of allopolyploids was probably also expressed in consensus sequences used for tree analyses.

The applied approaches provide a phylogenetic framework for recognizing sexual progenitor species and thus for unraveling the origins of (allo)polyploids. The detection of incongruences between plastid trees and nuclear datasets is a strong signal for hybridization events already on the diploid level (McKain et al., 2018; Dauphin et al., 2018). In this study, incongruences delivered valuable information for both sexual progenitors and polyploid derivatives (see e.g., parental contribution of a probably already extinct species in H_9). However, for the reconstruction of the reticulate relationships of allopolyploids, only genetic structure and network methods are able to unravel the correct evolutionary relationships.

Since neopolyploid complexes are evolutionary young and are characterized by reticulations and ILS, it is useful to employ genetic structure analyses that incorporate a maximum of allelic sequence diversity information. These analyses were previously often conducted with DNA fingerprinting markers (e.g., microsatellites, AFLPs). AFLP markers do not inform about heterozygosity, and in polyploids also microsatellites are usually scored as presence/absence, because allele dosages can often not be reliably assessed (e.g., Hodač et al., 2018; Karbstein et al., 2019; Melicharkova et al., 2020). RADseq covers a magnitude of markers, providing genome-wide sequence diversity per locus (SNPs) and thus robust results for genetic structure. RADpainter+fineRADstructure incorporates all SNPs, and varying allele numbers and amounts of missing data appropriate for young polyploid analyses (Malinsky et al., 2018; Wagner et al., 2021). In addition, the employed sNMF algorithm based on unlinked SNPs is not only faster, but also less sensitive to deviations from Hardy-Weinberg equilibrium (HWE) than the popular STRUCTURE software; it tolerates missing data, and is also applicable to different ploidy levels (Frichot et al., 2014; Frichot & François, 2020; Karbstein et al., 2021).

Whereas these methods impressively showed hybridity and a not yet recognized species gene pool of apomicts, coalescent-based STACEY species delimitation based on phased nuclear genes more clearly delimited the genetic structure of the polyploid complex (Fig. 6a,b). Allele phasing was demonstrated and is particularly considered as crucial for resolving young, reticulate relationships (Eriksson et al., 2018; Andermann et al., 2019; Freyman et al., 2020; Rothfels, 2021). Moreover, phylogenetic network analyses and subsequent tests mainly based on phased nuclear genes (see also Tiley et al. 2021) best unraveled subgenome contributions per polyploid and demonstrated predominant allopolyploid origins. Performing several network methods across different datasets informed by plastid information (i.e., consensus making) is

the most important part of our study to get a reliable picture about polyploid evolution in a such a young complex. Nevertheless, in general, potential limitations are the not realizable correct allele phasing of short RAD-Seq loci and relatively low number of nuclear genes, which were compensated by the combination of both datasets.

Disentangling genetic markers (SNPs) of polyploids for post-origin processes informs about divergence and stability of lineages. This information is crucial for classification and delimitation of species (Grant, 1981; Hörandl, 2018). Although incorporating ten thousands of RADseq loci, our SNiPloid analyses assigned only a minor fraction of RADseq-SNPs to homoeologous SNPs derived from hybrid origin. Here, a similar approach that incorporates homeologs (from various ploidy levels) derived from phased nuclear genes would be more favorable to assess polyploid (post-)origin evolution. A limitation of the SNiPloid pipeline is that the parental species must be defined for the input, only single samples can be analyzed, and that the algorithm is so far limited to tetraploids (i.e., not applicable to higher or lower ploidy levels; see also Wagner et al., 2020). However, the congruence of our results in eight independently analyzed hybrid lineages indicates two major trends in *R. auricomus*, namely considerable segregation of the diploid hybrid generation combined with gene flow after polyploidization, and so far only a low divergence via mutation in the more or less stable lineages.

Using the gained knowledge of this study, i.e., potential progenitor species of (allo)polyploids, ploidy levels and reproduction modes, and allo- vs. autopolyploid origins, subgenome assignments of allopolyploids and more appropriate phylogenetic allopolyploid networks (e.g., Jones, 2017a; Cao et al., 2019; Lautenschlager et al., 2020; Šlenker et al., 2021) are applicable for the polyploid complex. The combination of datasets and analytical pipelines gives a more comprehensive and complete picture of the evolution of young polyploid complexes.

Author Contributions

K.K., S.T., and E.H. designed research; K.K., L.H., and E.H. collected plant materials; K.K., S.T., P.M., B.B.H., and C.P. performed lab work; K.K., S.T., and N.W. analyzed data; K.K. wrote the paper with contributions of all authors.

Data Availability Statement

The authors declare that basic data supporting the findings are available within the manuscript and Supporting Information. RADseq, target enrichment, and CP alignments, and tables and figures supporting the results are deposited on FigShare (https://doi.org/10.6084/m9.figshare.-14046305). RADseq reads are deposited on the National Center for Biotechnology Information Sequence Read Archive (SRA): BioProject ID PRJNA627796 http://www.ncbi.nlm.nih.gov/-bioproject/627796). Flow cytometric (FC) and flow cytometric seed screening (FCSS) data are also stored in Figshare (https://doi.org/10.6084/m9.figshare.13352429).

Code Availability

We deposited custom bash, R, and Julia scripts on Github (https://github.com/KK260/-Ranunculus_auricomus_phylogenetic_network_scripts).

Funding

The work was supported by the German Research Foundation (DFG, grant number Ho4395/10-1).

Acknowledgments

We give thanks to the German Research Foundation for project funding to E.H., within the priority program “Taxon-Omics: New Approaches for Discovering and Naming Biodiversity” (SPP 1991). We also acknowledge Franz G. Dunkel for providing garden plants and herbarium specimens, Ena Lehtsaar and Julius Schmidt for technical help, and John Paul Bradican for suggestions on previous manuscript versions. We thank the herbaria of Jena (JE), Munich (M), Oslo (O), Uppsala (UPS), and the University of Vienna (WU) for loans of *R. auricomus* type species material.

III General Discussion

Extensively used, humid meadow close to Zakopane in the Tatra mountains (Lesser Poland, Poland). "*Ranunculus* × *subtatricus*", an undivided basal leaf, presumably polyploid asexual taxon described by A. Jasiewicz in 1954 (Jasiewicz, 1954; 1st June 2018), occurs in this region. Photograph: Kevin Karbstein

The Evolution of Young Polyploid Species Complexes

Darwin, and many of his contemporaries, recognized that specific evolutionary forces must have been responsible for the breathtaking, rapidly originating, and extraordinarily high species diversity of flowering plants (Darwin's "abominable mystery"; De Bodt et al., 2005; Friedman, 2009; Abbott et al., 2013; Buggs, 2021). Around 30 years after Darwin pondered this "abominable mystery", polyploidy was discovered at the beginning of the 20th century (Lutz, 1907; Gates, 1909; Strasburger, 1910; Buggs, 2021). Another 100 years later, in the current era of genomics, accumulating evidence indicates that polyploidy and hybridization are the major evolutionary forces connected to the rise of important developmental and regulatory genes (key innovations), large phenotypic diversity (morphology, biochemistry, etc.), and increases in species diversification rates (De Bodt et al., 2005; Marhold & Lihová, 2006; Abbott et al., 2013; Soltis et al., 2015; Landis et al., 2018; Van de Peer et al., 2020).

The study of young flowering plant species complexes is a crucial component in developing our understanding of evolutionary processes and speciation patterns accompanied by hybridization and polyploidy (Soltis & Soltis, 2009; Soltis et al., 2015; Pinheiro et al., 2018; Rothfels, 2021). In this thesis, I aimed at investigating the not yet comprehensively deciphered evolution and biogeography of young polyploid species complexes, using the large European apomictic *R. auricomus* complex as a model system. Previous technical, bioinformatic, and data limitations were overcome by applying a newly developed, multidisciplinary workflow based on up-to-date NGS methods: Genomic (RADseq), nuclear gene (target enrichment), and plastome datasets combined with current polyploid analysis methods, supported by flow cytometry, environmental data, geometric morphometrics, and crossing experiments (workflows summarized in Fig. 3, chapter 5: Karbstein et al., 2021b). The developed workflow represents a major conceptual breakthrough and reveals the first complete picture of both evolutionary processes and biogeographical relationships shaping a large polyploid species complex.

To enable the inference polyploid ancestry, sexual *R. auricomus* morphospecies were lumped into five well-defined progenitor species (chapters 1, 2; Karbstein et al., 2020b,c). Progenitors originated ca. 830–580,000 years ago from a Europe-wide distributed ancestor during Pleistocene climatic deteriorations (chapter 3; Tomasello et al., 2020). Besides advantages and success of polyploid apomicts in GP scenarios, the restricted distribution of sexual progenitors can also be explained by the negative effects of their breeding system in small and isolated populations (chapters 2: Karbstein et al., 2020c). Moreover, I demonstrated, for the first time,

that the here examined polyploid apomicts predominantly originated by allopolyploidization events involving five diploid sexual progenitor species (one of which is now likely extinct), and went through enormous post-origin genome evolution (chapter 5: Karbstein et al., 2021b). The larger distribution and success of polyploid apomicts is attributable to high levels of heterozygosity and uniparental reproduction, probably leading to increased environmental stress tolerance and an increased potential for rapid colonization of new habitats (chapter 4: Karbstein et al., 2021a).

By comparing the knowledge gained through my thesis with a comprehensive overview of related current research, I explain the evolutionary and biogeographical processes triggering young diploid sexual and polyploid apomictic plant speciation. I also show that in young polyploid species complexes microevolutionary processes (hybridization and polyploidization) are linked to macroevolutionary patterns in flowering plants: Few distinct sexual progenitor lineages (or species) generate novel, mainly polyploid, partly apomictic genotypes via allo- and autopolyploidization events. Backcrossing to progenitors and/or gene flow among derivatives further increase(s) genotypic diversity, leading to the development of enormous phenotypic richness. Arising polyploid or polyploid apomictic genotypes, compared to sexual progenitors, usually better cope with changing or more extreme environmental conditions, enabling their spread across large geographical regions. Genetically and morphologically unique, but also highly diverse lineages become fixed, for example, by geographical isolation or obligate apomixis, and are able to evolve as independent lineages in the long term.

The Sexual Progenitors

The delimitation of sexual progenitor species is the first step necessary to unravel polyploid phylogenetics and genome evolution in young species complexes (Grant, 1981; Burgess et al., 2015; Hörandl, 2018). In the past, elucidating young (phylo)genetic relationships was challenging, particularly if genetic divergence among taxa was low. Phylogenetic trees were usually reconstructed based on a few traditional or population genetic markers (e.g., internal transcribed spacers [ITS], microsatellites [SSR], amplified fragment length polymorphisms [AFLP]). These markers oftentimes did not provide sufficient genetic variation, or introduced false (non-homologous) or missing (null alleles) information among more diverged species and thus often produced unsupported phylogenies failing to reconstruct young relationships (e.g.,

R. auricomus: Hörandl et al., 2009; Freeland et al., 2011; *Hieracium*: Krak et al., 2013; *Rubus* subgen. *Rubus*: Kirschner et al., 2015).

In this thesis, phylogenomic approaches based on 86,782 parameter-optimized RADseq loci and 663 nuclear genes revealed the first fully resolved phylogeny of sexual *R. auricomus* morphospecies (BT ~90–100%; Figs. 3a,b, 5, chapter 1: Karbstein et al., 2020b). Despite the use of different methods and datasets, the concatenation-based ML tree (RADseq) and coalescent-based (nuclear gene) species tree showed highly congruent patterns: Five main clades that, in some cases, do not match with the already described morphospecies. Main clades originated ca. 830–580,000 years ago probably from a Europe-wide spread ancestor (Figs. 4, 5, chapter 3: Tomasello et al., 2020). RADseq provided an overall high resolution, particularly at the middle and terminal branches whereas nuclear genes delivered a better-resolved backbone. Outcomes are in line with previously published literature: The combination of highly variable, non-coding RADseq loci and less-variable, more conservative coding genes is useful for delimiting young species or population relationships (Baird et al., 2008; Davey et al., 2011; Cavender-Bares et al., 2015; Ree & Hipp, 2015; McKain et al., 2018).

Long reads generated using probes designed from pre-selected low-copy nuclear regions are easier to assemble into orthologous loci than RADseq reads, which is crucial for correct phylogenetic inferences particularly above the species and at the family level (Stephens et al., 2015; Eriksson et al., 2018; Vatanparast et al., 2018; Carter et al., 2019) but also at the population level (e.g., in plant pathogens, Bello Rodriguez et al., 2021). Highly congruent results also hint at an efficient orthologous locus assembly of RADseq loci. Without a reference genome, de novo sequence assembly of short, anonymous RADseq reads using only one minimum sequence clustering threshold within individuals (alleles to loci; ISCT) and across species (among loci; BSCT) easily leads to the formation of erroneously assembled clusters and paralogous loci. Currently, assembly pipelines are only capable of setting a minimum sequence clustering threshold within individuals and across species (e.g., IPYRAD: Eaton, 2014; Eaton & Overcast, 2020; STACKS: Catchen et al., 2013). However, genomes are collections of non-coding and coding loci, and each locus has its own evolutionary fate due to different selective pressures, resulting in specific genetic variabilities (Dawkins, 2008; Wolf et al., 2009).

To tackle this issue, recently, different criteria and workflows have been proposed, aiming at assembling orthologous loci and maximizing the "true phylogenetic signal" across the dataset for de novo sequence assemblies (Paris et al., 2017; McCartney et al., 2019; Pätzold et al., 2019; Wagner et al., 2020). In this thesis, I applied a combined approach built upon the ideas of Paris

et al. (2017) and Pätzold et al. (2019) as follows: Optimizing ISCT and BSCT separately, maximizing polymorphic loci and SNPs, and minimizing paralogs (chapter 1: Karbstein et al., 2020b), and adapting the complete procedure for different ploidy levels (chapters 4, 5: Karbstein et al., 2021a,b). This approach aims to yield an "optimal" RADseq dataset, where the true phylogenetic signal largely overpowers the assembly error. From a conceptual point of view, per-locus parameter optimization would be the best solution concerning de novo RADseq read assemblies. However, this has not yet been realized due to technical limitations. In addition, the methylation-sensitive cutting enzyme *PSTI* used here for RADseq analyses primarily targets nuclear genes (Fellers, 2008). This might also explain similar results obtained from both coding- and non-coding RADseq loci as well as nuclear target enrichment genes.

In general, ML tree, neighbor-net network, and STRUCTURE analyses based on RADseq data showed highly consistent results (Figs. 3, 4, chapter 1: Karbstein et al., 2020b). A deep split between the undivided basal leaf species *R. cassubicifolius* (first main clade) and the other undivided and dissected morphospecies (second main clade) was observed, supported by previous studies based on isoenzyme and ITS markers (Hörandl, 2004; Hörandl et al., 2009; Hodač et al., 2014). Findings clearly reject the old Linnean species classification distinguishing undivided (*R. cassubicus*) and divided basal leaf morphotypes (*R. auricomus*; Linnaeus, 1753; Kvist, 1987), at least for the sexual progenitor species. This rejection was already proposed based on the position of the undivided basal leaf apomict *R.* × *hungaricus* in relation to sexual progenitors using traditional and population genetic data (Hörandl et al., 2009). The position of the morphologically conspicuous species *R. flabellifolius* within the second main clade supports the rejection of the old classification in this thesis. Interestingly, genomic RADseq, nuclear gene, and plastome data conflicts hint at a homoploid hybrid origin (i.e., any kind of hybridization between sexual species without a shift in ploidy level; chapter 5: Karbstein et al., 2021b; see also Reichenbach, 1832; Nieto Feliner et al., 2017; Dunkel et al., 2018). Hybrid origin potentially also explains the rarely occurring three-lobed basal leaves and the unique fan-shaped stem leaf. *Ranunculus cassubicifolius* represents one progenitor of *R. flabellifolius* (same plastid type), and the second progenitor is potentially a species of the second main clade due to a similar nuclear genomic constitution.

Homoploid hybridization, in contrast to confirmed homoploid hybrid speciation events, is not rare in plants and can result in introgressive hybridization with the parental species (Nieto Feliner et al., 2017; Gramlich et al., 2018). For example, in the ox-eye daisy (*Leucanthemum*) polyploid complex, many homoploid hybridization events among diploid progenitor species have been observed (Wagner et al., 2019). It was previously thought that genetically close

crosses should result in homoploids whereas genetically distant ones would result in allopolyploids due to differing extents of genomic conflicts (see also "Darlington´s rule"; Darlington, 1937; Grant, 1981; Chapman & Burke, 2007; Soltis & Soltis, 2009). However, homoploid *R. flabellifolius* probably originated out of the most diverged progenitor subgenomes, and investigated allotetraploid apomicts within the *R. auricomus* complex showed a mixed pattern (chapter 5: Karbstein et al., 2021b). These findings support the conclusion of Buggs et al. (2011) that no general pattern is currently recognizable.

The only known polyploid sexual species without diploid cytotypes, *R. marsicus*, was placed in variable positions between trees (chapter 1: Karbstein et al., 2020b). In the ML RADseq tree, the branch splitting *R. marsicus* and Illyrian species received low BT support and was characterized by a high amount of discordant patterns indicating both ILS and reticulation events. Interestingly, *R. marsicus* and the diploid Illyrian species share the same plastid type (chapter 5: Karbstein et al., 2021b), an observation supported by their close position in trees and the neighbor-net network and high posterior probabilities of belonging to the same group in the coalescent-based species delimitation analysis (chapters 1, 3: Karbstein et al., 2020b; Tomasello et al., 2020). In addition, Masci et al. (1994) observed mainly two alleles per isoenzyme locus. The diploid behavior is typical for an allopolyploid origin. However, the majority of phylogenomic and genetic structure analyses (chapters 1, 5: Karbstein et al., 2020b, 2021b) indicate that at least one progenitor is an Illyrian morphospecies (e.g., *R. notabilis*). A second missing or extinct diploid species in the Central Apennines might be the potential second progenitor of *R. marsicus* (Dunkel et al., 2018). Taking *Taraxacum* (dandelions) for example, a genus which also contains mainly polyploid apomicts, only three sexual tetraploid out of ca. 2800 described taxa have been discovered thus far (Mogie & Ford, 1988; Záveská Drábkova et al., 2009; Trávníček et al., 2013). Concerning the *R. auricomus* complex, it´s feasible that subgenome divergence has to be low in the arising polyploid sexual such that subgenome conflicts are reduced, meiotic errors minimized, and the apomictic pathway not induced (see also autotetraploid sexual *R. cassubicifolius* cytotypes; Hörandl & Greilhuber, 2002). However, further (phylo)genomic research is needed to clarify the origin of *R. flabellifolius* and *R. marsicus*, to explain why the latter is the only observed sexual species with exclusively tetraploid cytotypes in the species complex, and why sexual polyploids occur so infrequently in young polyploid apomictic groups.

In general, all described sexual *R. auricomus* morphospecies are monophyletic. Nevertheless, highly conflicted middle and terminal branches in the ML RADseq tree, and shared posterior probability of belonging to the same cluster among some taxa in coalescent-based species

delimitation were identified (Figs. 3, 5, chapter 1: Karbstein et al., 2020b). The Central and Eastern European *R. cassubicifolius* and *R. carpaticola* form one genetic clade or cluster supported by previous research (Hörandl, 2004; Hörandl et al., 2009; Hodač et al., 2014). Conflicting signals indicated the presence of gene flow probably due to partial sympatry or young origin (split ca. 130,000 years ago, Fig. 4, chapter 3: Tomasello et al., 2020). Low genetic differentiation is supported morphologically by similar non-dissected basal leaves, broad lanceolate stem leaf segments, and narrow receptacles.

Moreover, the dwarf species *R. envalirensis* and the taller *R. cebennensis* of Southwestern Europe also split ca. 150,000 years ago, form one specific clade or cluster, and, as already mentioned by Dunkel et al. (2018), share similar dissected basal leaves, linear stem leaf segments with sinuses, and roundish receptacles. Habitats of the aforementioned species in the Pyrennees and Massif Central, respectively, are only separated by ca. 300 kilometers. In glacial times, populations might have been connected, and in interglacial times, isolation-by-distance (IBD) potentially led to weak genetic differentiation. IBD is a generally considered aspect in many evolutionarily young plants (Stebbins, 1984; Heywood, 1991; Taberlet et al., 1998; Brewer et al., 2002; Ali et al., 2012; Abbott et al., 2013). Intriguingly, all newly described Illyrian morphospecies and *R. notabilis* fall into a single clade, which is characterized by high levels of conflicting signals. Differentiation processes took place ca. 160,000–60,000 years ago (chapter 3: Tomasello et al., 2020), and totally missing geographical barriers (sometimes two to three morphospecies in a single meadow; Dunkel et al., 2018) probably enabled ongoing gene flow. Weak genetic differentiation is supported by a morphological continuum, ranging from basal leaves with three to five segments and linear stem leaves with and without sinuses. In experimental crossings, the Illyrian morphospecies revealed comparable inter- and intraspecific seed sets, indicating absent reproductive barriers (Fig. 3a, chapter 2: Karbstein et al., 2020c). Experiments also exhibited slight degrees of self-fertility (mean of 2.5%, up to 31%), which might explain local lineages and minorly differentiated morphotypes.

Finally, the lumping of the twelve described morphospecies, despite monophyly, into five progenitor species, is supported by multidisciplinary analyses based on genome-wide RADseq, nuclear gene, geometric morphometric, and crossing experiment data. Instead of “artificial” morphospecies characterized by missing genetic and morphological features, five clearly distinguishable genetic main lineages or clusters, that are well morphologically differentiated and geographically isolated were re-circumscribed (Fig. 5): *R. cassubicifolius* s.l. (incl. *R. carpaticola*), *R. envalirensis* s.l. (incl. *R. cebennensis*), *R. flabellifolius*, *R. marsicus*, and *R. notabilis* s.l. (incl. *R. austroslovenicus*, *R. calapius*, *R. mediocompositus*, *R. peracris*, and *R.*

subcarniolicus). This taxonomic treatment represents the combination of an evolutionary ancestor-descendant lineage concept with a cluster concept, including evolutionary role and circumscription (Mallet, 1995; De Queiroz, 2007; Freudenstein et al., 2017; Sukumaran & Knowles, 2017; Hörandl, 2018). The re-classified progenitor species present a robust phylogenetic framework for the following polyploid analyses in chapters 4 and 5 (Karbstein et al., 2021a,b).

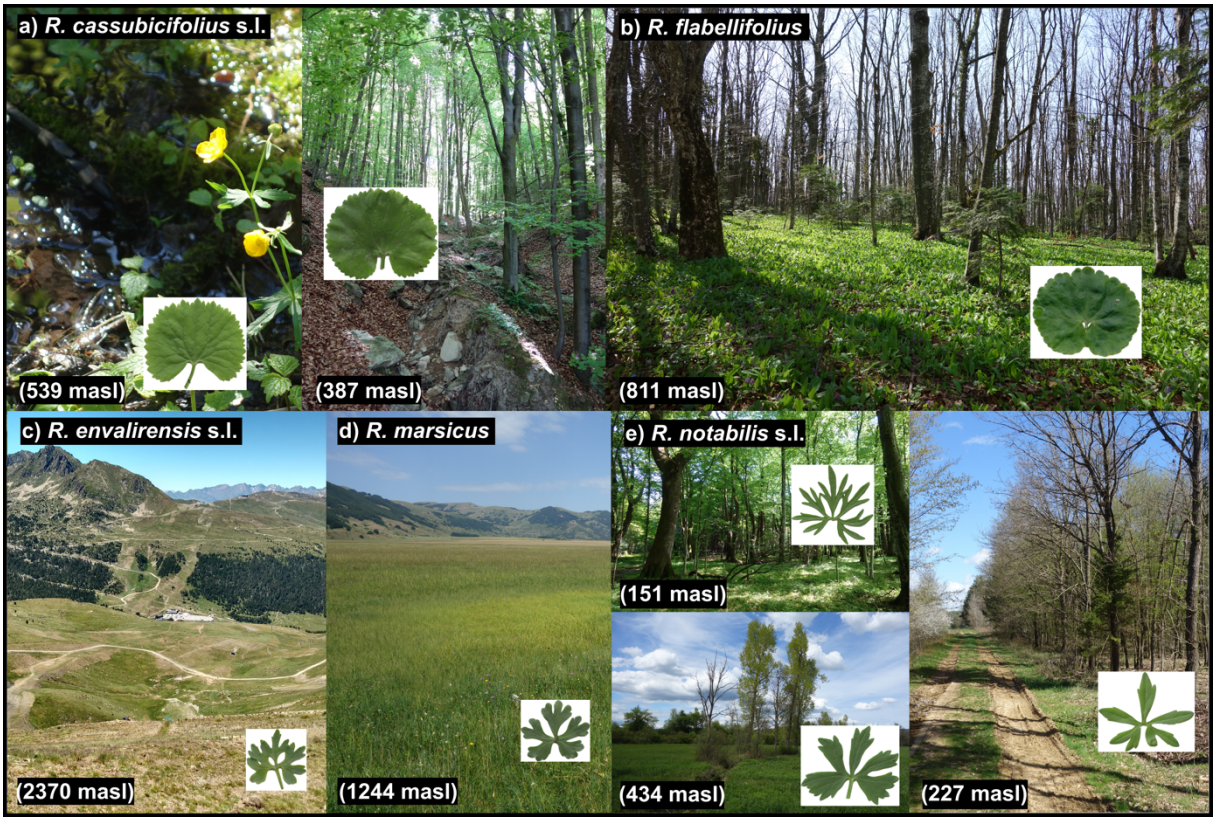


Figure 5. Habitats and accepted sexual species of the *R. auricomus* complex (basal leaf at anthesis shown, size of the image correlated to original leaf size; slightly modified figure from Karbstein et al., 2020b; published under Attribution 4.0 International, CC BY 4.0). (a) *R. cassubicifolius* s.l. occupies humid hornbeam and beech forests, and stream- and riversides; (b) *R. flabellifolius* inhabits a few forest locations in the Southern Carpathians; (c) *R. envalirensis* s.l. occurs in subalpine meadows in the Pyrennees and the Massif Central; (d) tetraploid *R. marsicus* inhabits humid subalpine meadows; (e) *R. notabilis* s.l. occupies forests and waysides, and humid to marshy meadows in the Illyrian region. See legend of Fig. S8 in Karbstein et al. (2020b) for more details.

Accepted sexual *R. auricomus* progenitors are distributed in restricted ranges across Central and Southern European mountain systems (Fig. 1, chapter 3: Tomasello et al., 2020). Despite the availability of suitable habitats, restricted ranges are a general feature of sexuals in GP scenarios, and are often interpreted as representing past glacial refugia during the Pleistocene (Bierzychudek, 1985; Hörandl, 2006; Hörandl et al., 2008; Cosendai et al., 2013; Kirchheimer et al., 2018; chapter 3: Tomasello et al., 2020). In this thesis, experiments demonstrated fewer seeds in intrapopulation crosses, and more seeds and seedlings in crosses involving genetically more distant and heterozygous lineages of *R. notabilis* s.l. (Figs. 5, 6, chapter 2: Karbstein et al., 2020c). Besides competition with conspecific apomicts and recent anthropogenic impacts (e.g., the decreasing area of extensively used meadows and forests), two- to three-times less genome-wide heterozygosity and negative inbreeding effects of the sexual reproductive system in small and/or isolated populations potentially explain their restricted ranges and lower propensity for colonization (Baker, 1967; Haag & Ebert, 2004; Hörandl, 2006; Freeland et al., 2011; Dunkel et al., 2018; chapter 1, 2, 4: Karbstein et al., 2020b, 2020c, 2021).

Recently, more has research has been conducted seeking to explain the restricted ranges of sexuals in GP scenarios: For example, in the polyploid apomictic *Potentilla puberula* complex in the Eastern European Alpes, the restricted distribution of diploid sexuals are explained by reproductive and competitive interactions with apomictic conspecifics and ecological differentiation towards drier, steeper and more southern slopes (Alonso-Marcos et al., 2019; Nardi et al., 2020); in the polyploid apomictic *Hieracium alpinum* complex, three-times lower seed set and four-times more variable seed reproduction due to pollen limitation in low-density sympatry, inbreeding effects, and consequently a lower colonizing ability are probably responsible for the restricted range of sexuals compared to obligate apomicts (Mráz et al., 2019; Pinc et al., 2020; Mráz & Mrázová, 2021).

The Polyploid Apomictic Derivatives

Species complexes are groups of newly arising, stable or unstable, closely related genetic lineages that are characterized by a magnitude of phenotypes with often missing boundaries among them (Freudenstein et al., 2017; Hörandl, 2018; Pinheiro et al., 2018). In plants, the artificial treatment of complexes as agglomerates often neglects their species richness (i.e., cryptic diversity), leading to flawed, underestimated, or unobserved biodiversity, which is

particularly distressing in the current period of global diversity crises (Great Britain: Richards, 2003; Hörandl, 2018; South America: Pinheiro et al., 2018; Żabicka et al., 2020). Nevertheless, to understand the first steps of species formation, the study of young complexes is most important (Pinheiro et al., 2018). Genomic data are mandatory to analyze the origin of polyploids, and in past times, polyploids were frequently dropped in phylogenetic analyses due to a lack of suitable markers (Freyman et al. 2020; Rothfels et al., 2017; Rothfels, 2021). Several authors recognized the need for unveiling these challenging evolutionary processes through a multidisciplinary approach incorporating enormous amounts of genomic, cytological, reproductive, ecological, and geographical data combined with taxonomic conclusions (Fig. 6; Hörandl, 2018; Pinheiro et al., 2018; Rothfels, 2021).

The complex evolutionary origin and the fascinating morphological diversity of the currently more than 800 polyploid apomict taxa within the *R. auricomus* complex have been a mystery and an object of not only a great deal of research, but also speculation for decades (Marklund, 1961, 1965; Ericsson, 1992; Hörandl, 1998; Paun et al., 2006a,b; Hörandl et al., 2009; Hodač et al., 2014, 2018; Hojsgaard et al., 2014b; Dunkel et al., 2018). Unraveling the evolution of young plant species complex is like untying a Gordian knot: the seemingly endless reticulate genetic and morphological relationships (the threads) are knotted so complexly that the starting point is not recognizable. The theoretical roadmap is already given, starting with the "sexuals-first principle" and ending with the classification of polyploid apomicts in relation to reproduction mode and lineage stability (Grant, 1981; Burgess et al., 2015; Hörandl, 2018), whereas its practical implementation is currently enabled by OMICS data and tools.

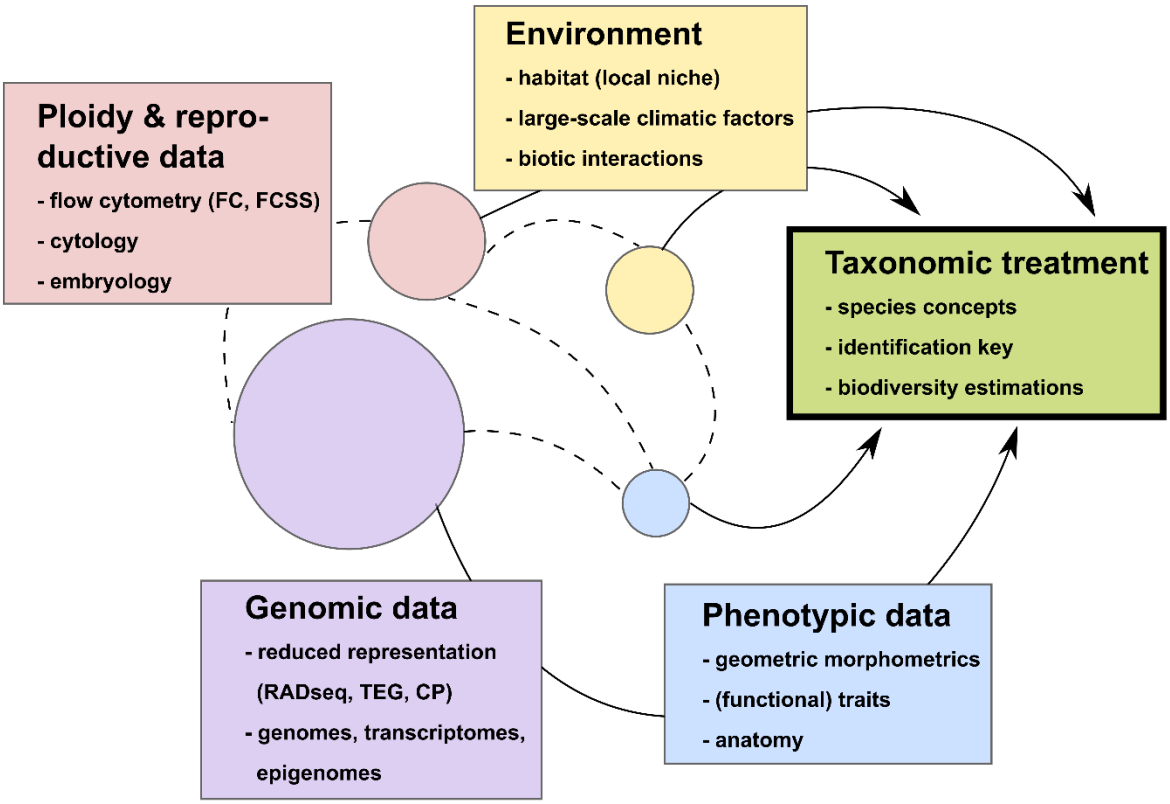


Figure 6. Multidisciplinary data needed to unravel evolutionary processes within young species complexes. Per data type, different analysis methods are given. Circle sizes represent the importance of each analysis type to decipher evolutionary processes. Results based on genomic data (informed by ploidy and reproductive data) should be considered as most important for final taxonomic treatments. The combination of ploidy and reproductive with environmental data can be used to understand biogeographical relationships and speciation processes. Selection processes on specific phenotypic traits under different environments are detectable by statistical modeling. Phenotypic data can be compared with ploidy and reproductive data, for example, to understand phenotypic richness in the light of polyploidization events. Finally, genetic clusters might be compared with phenotypic differentiation and used to support the final taxonomic treatment. CP = plastome data, FC = flow cytometry of leaves, FCSS = flow cytometric seed screening, TEG = target enrichment nuclear genes.

I developed a workflow based on multidisciplinary data and methods to decipher evolutionary processes within the less than 1.0 million-year-old *Ranunculus auricomus* polyploid model system (chapter 4, 5: Karbstein et al., 2021b). Genetic structure and network analyses, consensus making from previous results, and auto- versus allopolyploid testing mainly based on phased nuclear genes demonstrated predominant allopolyploid origins involving two to three progenitor subgenomes in each event. Findings are supported by data from several other young species complexes shaped by allopolyploidization events, for instance, apomictic polyploid

brambles (*Rubus*; Sochor et al., 2015), dandelions (*Taraxacum*; Šarhanová et al., 2012; Kirschner et al., 2015), and moonwort ferns (*Botrychium*; Dauphin et al. 2018), or polyploid sexual (though partly vegetatively reproducing) strawberries (*Fragaria*; Kamneva et al., 2017), cordgrasses (*Spartina*; Ainouche et al., 2004; Ayres et al., 2007), orchids (*Dactylorhiza*; Hedrén et al., 2008; Brandrud et al., 2020), ox-eye daisies (*Leucanthemum*, Wagner et al., 2017, 2019), and willows (*Salix*; Triest et al., 1999; Wagner et al., 2020). Moreover, the genetically informed path analysis based on GLMMs revealed empirical evidence for factors significantly shaping GP scenarios of large polyploid complexes in temperate Europe (Fig. 4, chapter 4: Karbstein et al., 2021a). The larger, more northern distribution of apomicts compared to sexuals within the *R. auricomus* complex (latitudinal GP) is explained by genomic features and climatic environmental factors. Latitudinal GP scenarios have been frequently observed in polyploid species groups of the Northern Hemisphere, for example, Alpine hawkweed (*Hieracium alpinum*, Mráz et al., 2009, 2019), hawthorns (*Crataegus*; Lo et al., 2013), or juneberries (*Amelanchier*; Burgess et al., 2014), but the combination of observed relationships between sexuality or genome-wide heterozygosity and specific environmental factors presents a novel aspect in GP research. The occurrence of apomicts towards previously glaciated areas is associated with solar radiation-poor, slightly drier habitats characterized by high-temperature fluctuations (compare Zinn et al., 2010; Schinkel et al., 2016; Coughlan et al., 2017; Rushworth et al., 2018). Uniparental reproduction and remarkably higher levels of genome-wide heterozygosity probably enable polyploid apomicts to colonize more northern, previously glaciated, and/or harsh habitats (Comai, 2005; Hörandl & Paun, 2007; chapter 4: Karbstein et al., 2021a). However, in general, findings support considerations that hybridization and polyploidization are likely to create novel genomic and phenotypic features and are (partly) responsible for the success of apomicts in GP scenarios and the great morphological diversity of polyploid apomicts in the *R. auricomus* complex (Liu & Wendel, 2003; Hörandl, 2006; Hörandl & Paun, 2007; Abbott et al., 2013; Coughlan et al., 2017; Van de Peer, et al., 2020; Rothfels, 2021).

In detail, the ML RADseq and coalescent-based nuclear gene trees provided a first and robust phylogenetic framework of diploids and polyploids based on genomic and nuclear gene data (Fig. 4a,b, chapter 5: Karbstein et al., 2021b). In contrast to expectations based on the large, continuous morphological diversity, the 75 included taxa were organized in only five main clades. Findings support the suitability for resolving both diploid and polyploid phylogenetic relationships of various evolutionary ages with RADseq (e.g., *Dactylorhiza*: Brandrud et al., 2020; *Salix*: Wagner et al., 2020) or nuclear gene analysis combined with plastome data (e.g.,

Rubus: Carter et al., 2019; Asteridae: Stull et al., 2020; *Prunus*: Hodel et al., 2021). However, this is the first study combining genomic RADseq, nuclear gene, and plastome data in hopes of resolving a large polyploid apomictic plant species complex. Surprisingly, despite different datasets and methods, the phylogenetic pattern is highly congruent regarding the five main clades and stable clade-wise sample composition: Only one out 113 shared samples was situated in different main clades between trees. Similar results between different trees were previously found in sexual progenitors (chapter 1, Fig. 3: Karbstein et al., 2020b), indicating a strong evolutionary signal, supporting the comparability of nuclear target enrichment genes and parameter-optimized non-coding and coding (anonymous) RADseq loci for investigating evolutionarily young relationships, and delivering a robust phylogenetic framework for subsequent polyploid analyses.

Per main clade, a single sexual species was surrounded by polyploid apomicts (chapter 5, Fig. 4a,b: Karbstein et al., 2021b). This pattern indicates that apomicts are primarily derived from their clade-specific progenitors, which is rather unique compared with other polyploid plant complexes. Typically, clades (or clusters) of both mixed-ploidy levels and only polyploids were observed, for example in brambles (*Rubus*: Carter et al., 2019), cuckoo flower complex (*Cardamine pratensis*: Melichárková et al., 2020), dandelions (*Taraxacum*: Kirschner et al., 2015), moonworts (*Botrychium*: Dauphin et al., 2018), or willows (*Salix*: Wagner et al., 2020). Despite partly well-resolved RADseq and nuclear gene tree backbones, quartet support values indicate the presence of reticulations and/or ILS. Allopolyploids are known to bias ordinary tree reconstructions due to reticulate evolution and mingling of different evolutionary histories in consensus sequences (McDade, 1992; Oxelman et al., 2017; Dauphin et al., 2018; Rothfels, 2021). Nevertheless, allopolyploids resulting from crosses of less diverged progenitor species are expected to not tremendously impact the reconstructed tree phylogeny (McDade, 1992; Soltis et al., 2008), and present DNA sequence subgenome dominance is probably also translated in consensus sequence based on majority base rule calling. The ML plastome tree phylogeny revealed remarkable incongruences with nuclear data (RADseq, nuclear genes) in the *R. auricomus* complex (Figs. 4, 5, chapter 5: Karbstein et al., 2021b). *Ranunculus flabellifolius* and polyploid apomicts (RADseq and nuclear gene main clade II) share the same plastid type with *R. cassubicifolius* s.l. (RADseq and nuclear gene main clade I), while *R. marsicus* and polyploid apomicts (RADseq and nuclear gene main clade III) share the same plastid type with *R. notabilis* s.l. (RADseq and nuclear gene main clade IV). Additionally, a new clade was observed containing only polyploid apomicts in Central Europe related to *R. envalirensis* s.l.

Nuclear-plastid (i.e., cytonuclear) discordance is a strong signal for hybridization events, as demonstrated at the diploid but also at the mixed-ploidy level (McKain et al., 2018; Dauphin et al., 2018; Carter et al., 2019; Stull et al., 2020), and as already explained herein for the probably homoploid *R. flabellifolius* and allopolyploid *R. marsicus* and their respective polyploid apomicts (chapters 1, 5: Karbstein et al., 2020b, 2021b). Interestingly, Hodel et al. (2021) also investigated a well-resolved nuclear-gene tree backbone despite cytonuclear discordance in the cherry and plum genus (*Prunus*), which is shaped by past hybridization and allopolyploidization events. The plastid type observed only in polyploids also hints at the existence of a sexual progenitor species not included in this analysis, one which is possibly already extinct. Unexpected new diploid sexuals have been found recently in less explored regions of the species range of young groups (e.g., in *Hieracium*, Eastern Carpathians; Mráz et al., 2019), and even in the *R. auricomus* complex (Illyrian region; Dunkel et al., 2018; Dunkel, 2019). Nevertheless, the possibility of an extinct progenitor is more likely, as ploidy levels of the *R. auricomus* complex in Central Europe are well-documented and sexuality is restricted to only rarely occurring diploid populations (Jalas & Sumoninen, 1989; Paule et al., 2018; chapter 4: Karbstein et al., 2021a). The existence of now-extinct sexual progenitors is a possibility often taken into account in research on young plant species groups (Stewart et al., 2010; Sochor et al., 2015; Rothfels, 2021). It is also considered a possibility in the *R. auricomus* complex, where a "speciation gap" was found between 0.6 and 0.3 million years ago (chapter 3: Tomasello et al., 2020). Surprisingly, I found a geographically isolated population composed of sexual and apomictic individuals ("KK166") in Eastern France, sharing the plastid type with the Central European apomicts characterized by an unknown sexual progenitor (chapter 5: Karbstein et al., 2021b). Further research is necessary in order to clarify whether the sexual tetraploid individuals of this population represent descendants of a former diploid progenitor species (Karbstein et al., in prep. a).

The application of appropriate and up-to-date polyploid genetic structure, network, and SNP-origin methods mainly enabled the unraveling of reticulate origins in this young neopolyploid complex. I found three to five, largely overlapping main genetic clusters using nearest neighbor coancestry (RADPAINTER), sparse nonnegative matrix factorization (SNMF), and coalescence-based species delimitation (STACEY) based on RADseq or nuclear gene datasets (Figs. 6, 7, chapter 5: Karbstein et al., 2021b). Main clusters largely overlap across methods and datasets, and are also similar to those found in tree reconstructions: A Western and Central European cluster comprising *R. envalirensis* s.l., the proposed extinct progenitor species, and related apomictic polyploids (split into two clusters in another sNMF analysis), a Central

European cluster comprising *R. notabilis* s.l., *R. marsicus*, and *R. flabellifolius* and related apomictic polyploids (split into three clusters in the STACEY analysis), and an Eastern European cluster comprising *R. cassubicifolius* s.l. and related apomicts. RADPAINTER, which uses all RADseq-SNPs of varying ploidy levels (Malinksy et al., 2018), was the most powerful tool in detecting significant reticulation signals between sexual progenitors and polyploid apomicts within the *R. auricomus* complex. SNMF based on RADseq-SNPs showed with remarkable resolution the west-east interclade and north-south intraclade distribution of apomicts. The program was specifically developed to tolerate moderate amounts of missing data usually present in RADseq matrices as well as deviations from the Hardy-Weinberg equilibrium as expected in apomicts (Freeland et al., 2011; Frichot et al., 2014; Eaton et al., 2017; Frichot & François, 2020). STACEY was best-suited to clearly delimit genetic structure, probably due to the advantages of using phased nuclear genes (Eriksson et al., 2018; Andermann et al., 2019; Rothfels, 2021), and revealed five clusters each containing a sexual progenitor species and related polyploid apomicts. Consequently, genetic structure results based on STACEY are most valuable for taxonomic conclusions, whereas RADPAINTER and SNMF give crucial information on hybridization signals and biogeographical relationships. Some of these patterns are also found in other polyploid species groups shaped by Pleistocene climatic deteriorations and ongoing gene flow: For example, taxa of *Dactylorhiza* orchids are also organized into large, weakly resolved RADseq-clusters characterized by highly reticulate relationships (Brandrud et al., 2020); taxa of brambles (*Rubus*) or dandelions (*Taraxacum*) showed a similar pattern based on nuclear genes (Carter et al., 2019); and a European west-east gradient was found in *Arabidopsis arenosa* (Schmickl et al., 2012) or *Cardamine pratensis* (Melichárková et al., 2020) species complexes. However, the combination of the observed patterns in a single species complex is rather unique and probably attributable to the young origin and enormous post-origin processes shaping apomicts within the *R. auricomus* group.

The most innovative part of my work is the combination of genetic structure, phylogenetic network, and plastid information via consensus result making, subsequent allo- versus autopolyploid origin statistical testing based on hundreds of nuclear genes, and final SNP-origin analyses of allopolyploids comprising nearly 100,000 RADseq loci (Table 1, Figs. 6, 8, chapter 5: Karbstein et al., 2021b). Analyses on various datasets (genomic RADseq, nuclear gene, and CP) and using different methods complement each other and illustrate a detailed picture of evolutionary processes within the complex: I predominantly found allopolyploids that are composed of two to three different subgenomes and largely derived from their clade- or cluster-specific progenitors; few autotetraploids (or segmental allopolyploids); and pronounced

subgenome dominance of allopolyploids mainly caused by tremendous post-origin evolution. Phylogenetic networks (PHYLONETWORKS: Solís-Lemus et al., 2017 and PHYLONET: Than et al., 2008; Wen et al., 2018) and subsequent statistical testing (i.e., combination of full likelihood and Akaike Information Criterion (AIC) calculations based on nuclear genes in PHYLONET) finally verified parental contributions and allopolyploid origins, and supported subgenome dominance by high inheritance probabilities of parent one (mean of ca. 74%). This approach is supported by simulations demonstrating the usefulness of modeling the correct network-like origins based on phased nuclear genes (Tiley et al., 2021), and by a similar statistical testing method based on haplotypes, which successfully shed light on the evolution of allopolyploid strawberries (*Fragaria*; Kamneva et al., 2017).

In general, subgenome dominance in young allopolyploids represents a hitherto less regarded feature, particularly for plant species complexes. Recent genomic studies demonstrate that many allopolyploid model species are characterized by subgenome dominance (Blischak et al., 2018; Edger et al., 2018; Alger & Edger, 2020). Two different, partly connected, timely overlapping processes exist: First, subgenome expression dominance and, second, subgenome sequence dominance (i.e., DNA expression or sequence dominance of one subgenome over the other). Differently methylated transposable elements (TEs) close to genes, retention of highly connected genes, or environmental influence causing adaption and selection of specific genes probably lead to expression and sequence dominance via gene loss (Schnable et al., 2011; Alger & Edger, 2020; Bird et al., 2021). Subgenome dominance has been documented for important allopolyploid crop plants, for example, 5–12 million years old maize cultivars (*Zea mays*; Schnable et al., 2011) or first-generation rape cultivars (*Brassica napus*; Bird et al., 2021), but also for less than 140-years-old, resynthesized monkeyflower allopolyploids (*Mimulus peregrinus*; Edger et al., 2017). In this thesis, I provide novel evidence for subgenome dominance in non-model allopolyploid species complexes. This aspect should be considered in future polyploid research (e.g., subgenome assignments or allopolyploid network modeling).

Subgenome sequence dominance can also be caused by various post-origin processes. In this thesis, RADseq-SNP-origin analyses (SNIPLOID: Peralta et al., 2013) showed low percentages of hybrid-SNPs, and high percentages of interspecific-SNPs solely derived either from the first or second progenitor. The observed sequence subgenome dominance in the less than 1.0 or potentially even less than 0.1 million years old *R. auricomus* allopolyploids (Pellino et al., 2013; chapter 3: Tomasello et al., 2020) could not only be explained by biased fractionation but also by segregation processes of the young hybrid generations, backcrossings to progenitors, and/or gene flow among facultative sexual apomicts after polyploidization (chapters 4, 5: Karbstein et

al., 2021a,b). Nevertheless, I regard the impact of biased fractionation as minor, as the majority of studies show no or relatively weak signs of biased fractionation in young allopolyploids, for example, in *Brachypodium hybridum* or wheat species (*Triticum*; Gordon et al., 2020; Schiavinato et al., 2021). Apomixis establishes step-wise in the *R. auricomus* complex (Barke et al., 2018), allowing Mendelian segregation of the early sexual hybrid generations and backcrossing of hybrids to progenitors (chapter 5: Karbstein et al., 2021b). In Central Europe, all main genetic clusters are present and clusters might have met during past interglacials, rendering Central Europe a potential center of origin for the *R. auricomus* complex. Nevertheless, further time-calibrated phylogenies and biogeographical analysis concerning the polyploid apomicts would shed light on this interesting point. In addition, populations are mainly obligate apomictic (ca. 65%) and the remaining facultative apomicts (ca. 28%) showed only low levels of residual sexuality within the *R. auricomus* complex (mean of 7.1%, up to 34%; chapter 4: Karbstein et al., 2021a). Sexuality is positively correlated to solar radiation and isothermality, and towards Central and Southern Europe, gene flow is enabled between sympatrically occurring facultative apomictic populations (chapter 4: Karbstein et al., 2021a). Gene flow in combination with allopolyploid origins based on less diverged progenitors is probably responsible for the low genetic divergence found in the RADPAINTER genetic structure analysis. Weak genetic structure is confirmed by negligible percentages of lineage-specific, post-origin RADseq-SNPs, suggesting slight mutation accumulation in young allotetraploids (Meselsson effect; Welch & Meselson, 2000; Pellino et al., 2013; chapter 5: Karbstein et al., 2021b). Low amounts of sexual reproduction in facultative apomicts might also purge (deleterious) mutations, as theoretically demonstrated by model calculations in Hodač et al. (2019). These findings are supported by a slow, gradual evolution of allopolyploid *B. hybridum* 1.4 million years after polyploidization (Gordon et al., 2020), and high percentages of lineage-specific mutations of several million years old willows (*Salix*; Wagner et al., 2020).

Despite the presence of main phylogenetic clades or genetic clusters, the polyploid apomictic *R. auricomus* species complex is genetically very heterogeneous. Different evolutionary processes are probably responsible: hybridization and subsequent polyploidization unifying two or three different progenitor subgenomes in a single individual, and significant post-origin evolution (i.e., hybrid segregation and backcrossing to parents, gene flow due to facultative sexuality of apomicts). Consequently, processes lead to an extremely high genomic diversity and thus the weak genetic differentiation regarding the entire species complex. Moreover, not only the inter-individual, but also the intra-individual genomic diversity is high. Path analysis revealed two to three times higher individual genome-wide heterozygosity in polyploid

apomicts compared to sexuals (chapter 4: Karbstein et al., 2021a). Hybrid origin, fixed heterozygosity at duplicated loci, and recombination due to facultative sexuality explain increased heterozygosity (Gornall, 1999; Hörandl & Paun, 2007; Cosendai et al., 2013; Hodač et al., 2019).

In general, polyploid populations or individuals are more heterozygous than their diploid counterparts (Soltis & Soltis, 2000), which is also confirmed by other population-genetic or genome-wide studies (e.g., Štorchová et al., 2002; Hörandl & Paun, 2007; Mohammadin et al., 2018). Inter- and intra-individual genomic diversity offer much genetic variation for natural selection to act, and is potentially responsible for several of the observed benefits of polyploids or polyploid apomicts: Enlarged reaction norm and phenotypic plasticity, improved stress tolerance, increased adaptive potential under more extreme conditions, and finally larger distribution and success in GP scenarios (Comai, 2005; Hörandl, 2006; Hörandl & Paun, 2007; Qiu et al., 2020; Van de Peer et al., 2020; chapter 4: Karbstein et al., 2021a). All these processes taken together might explain the extraordinarily high morphological diversity within the entire polyploid apomictic *R. auricomus* complex.

Microevolution linked with Macroevolutionary Patterns

Unraveling the evolutionary processes of polyploids and polyploid species complexes is one of the most challenging fields in evolutionary biology, as well as one of the most promising in furthering our understanding of the speciation process (Soltis et al., 2015; Pinheiro et al., 2018; Rothfels, 2021; Karbstein et al., 2021b). In this thesis, I show that the microevolutionary processes of polyploidy and hybridization have shaped the evolution of the *R. auricomus* and other various (apomictic) polyploid species complexes of different evolutionary ages. In natural flowering plant populations, polyploid formation is enforced by the frequent formation of unreduced gametes, providing a mechanism for plant speciation via allopolyploidy but also autopolyploidy (Ramsey & Schemske, 1998, 2002; Soltis et al., 2015; Spoelhof et al., 2017). Both formation types are considered to fundamentally contribute to the species diversity of flowering plants (Barker et al., 2016; Spoelhof et al., 2017; Van de Peer et al., 2020; Rothfels, 2021).

Newly arising polyploid lineages become stabilized by reproductive (e.g., due to obligate apomixis or ploidy-related crossing barriers), ecological (e.g., niche differentiation), and/or

geographical isolation (e.g., by geographical barriers or climatic fluctuations). Further evolutionary fate depends on their success in relation to abiotic and biotic environmental conditions. Polyploid lineages are not simply the sum of their diploid progenitors: Formation events and post-origin processes generate unique polyploids, each composed of a mosaic-like structure of novel, additive, and parental features (Soltis et al., 2015; Van de Peer et al., 2020), supported by (phylo)genomic results in this thesis (chapters 4, 5: Karbstein et al., 2021a,b). Genomic novelty, epi- and transcriptomic changes, increased heterozygosity and heterosis effects, and uniparental reproduction likely increase their colonizing ability and allow for better survival under extreme environmental conditions and in stressful periods of climate change (Bierzychudek, 1984; Hörandl, 2006; Van de Peer et al., 2020; chapter 4: Karbstein et al., 2021a). Polyploids and polyploid asexuals are frequently more successful than sexual progenitors in young polyploid species complexes, and in evolutionary time scales, these lineages might evolve into separate species.

Neopolyploids frequently occur in flowering plants, but in the long term, only a few of them seem to survive (Wood et al., 2009; Soltis et al., 2015; Van de Peer et al., 2017; Landis et al., 2018). In the past, and also in recent years, this observation and the theory of deleterious mutation accumulation in asexuals (Muller's Ratchet; Muller, 1964) has led many researchers to brand polyploids as dead-ends (Arrigo & Barker, 2012; Mayrose et al., 2015; Van de Peer et al., 2017). The expected burden of (deleterious) mutation accumulation has, however, not appeared in investigated polyploid genomes, probably due to purging of deleterious mutations by facultative sexuality and/or purifying selection (Pellino et al., 2013; Brandt et al., 2017; Hodač et al., 2019; Jaron et al., 2021). Again, contrary to expectations, polyploids do not inevitably become extinct: Novel genomic analyses impressively showed that all flowering plants share an ancient polyploid origin and that particularly young polyploidization events probably caused upshifts in species diversification rates (Landis et al., 2018; Leebens-Mack et al., 2019) due to the above-mentioned benefits of (allo)polyploidy. The paucity of ancient polyploidy in flowering plants is explained by the general rule of diploidization, i.e., the restructuring of polyploid genomes by i.a. chromosomal rearrangements, homeolog expression bias, and biased fractionation back to a functional diploid state over evolutionary time scales (Wolfe, 2001; Soltis et al., 2015; Blischak et al., 2018). In addition, I also expect that asexual reproduction switches back during the diploidization process to sexual reproduction because different subgenomes and the previously existing genomic conflicts no longer exist. Finally, after all these processes, the novel diploid contains a unique combination of genomic and consequently phenotypic features susceptible for further evolutionary processes (Soltis et al.,

2015; Edger et al., 2018). However, in the era of genomics, more and more studies will probably shed light on polyploid evolutionary processes and will discover currently hidden, ancient polyploidization events in the plant kingdom.

Taxonomic Implications

Famous and well-known species complexes exist in plants (e.g., *Salix* trees and shrubs: Wagner et al., 2018, 2020 or *Dactylorhiza* orchids: Brandrud et al., 2020), fungi (e.g., *Amanita muscari* fly agaric: Geml et al., 2006), and in animals (e.g., *Heliconius* butterflies: Kozak et al., 2021; or *Drosophila melanogaster* flies: Jagannathan et al., 2017). Most studied species complexes are from temperature regions, whereby tropical species complexes have been less well investigated, despite substantially contributing to local biodiversity (Pinheiro et al., 2018). Agglomerates often hide species richness of a complex whereas the treatment as (micro)species often lacks a genetic background (Richards, 2003; Hörandl, 2018; South America: Pinheiro et al., 2018; chapters 1, 4, 5: Karbstein et al., 2020b, 2021a,b; Żabicka et al., 2020). The impact on biodiversity estimations is tremendous: With classifying the agamospecies of some large apomictic polyploid complexes (*Hieracium*, *Taraxacum*, and *Rubus*), Richards (2003) calculated that the Flora of the British Islands would double (50% apomictic species). Haveman et al. (2002) also calculated an increase from 1450 to 1900 species for the Netherlands. Consequently, concerning the *R. auricomus* complex, taxonomic treatments of splitting, i.e., recognizing taxa as separate species, or lumping, i.e., recognizing only one agglomerate or few larger groups, would have a large impact on biodiversity estimations for Central (ca. 150 taxa) and Northern Europe (ca. 600 taxa; Fig. 7).

The majority of analyzed, rather widespread apomictic polyploid *R. auricomus* taxa were non-monophyletic in phylogenomic trees (e.g., 62–66 of 75 taxa in ML RADseq trees; chapter 5: Karbstein et al., 2021b). Nevertheless, as already expected from previous genetic and morphometric research, no clade- or group-specific morphological pattern was detected. Taxa characterized by mainly undivided basal leaves and broad lanceolate stem leaf segments were organized in main clades I and II (only a few exceptions). Dissected basal leaf taxa with narrow lanceolate stem leaf segments were predominantly found in main clades III–V. Consequently, as for sexual progenitors, the old Linnean classification of two main morphotypes (Linnaeus, 1753; Kvist, 1987) is also rejected for polyploid apomicts. Within main clades, BT values were

extremely low and quartet support values showed enormous amounts of discordant patterns (chapter 5: Karbstein et al., 2021b). Cladogenetic evolution from a few polyploid apomictic progenitors would lead to bifurcating, non-conflicted patterns in middle and shallow tree branches, which is not the case here. A more likely scenario explaining the extremely low support metrics is the repeatedly ongoing formation of polyploid apomicts and/or extensive post-origin evolution (e.g., hybrid segregation, backcrossings to parents, and gene flow of polyploid apomicts). In Central Europe, polyploid formation waves might have been frequently generated by Pleistocene climatic fluctuations, and, in Southern and Central Europe, high facultative sexuality of apomicts probably allowed for significant gene flow (chapters 3–5: Tomasello et al., 2020; Karbstein et al., 2021a,b). Interestingly, the degree of facultative sexuality decreases towards Northern Europe (chapter 4: Karbstein et al., 2021a), potentially associated with higher lineage stability and the presence of more monophyletic taxa. However, only three to five genetic clades or clusters could be recognized across the 75 included polyploid apomictic taxa (chapter 5: Karbstein et al., 2021b). Genetic differentiation among clusters is relatively weak but recognizable. Genomic composition within clades or clusters indicates various different contributions for the second or third parent but usually shows a clade-specific sexual progenitor. Three main clusters are distributed from Western to Eastern Europe, which largely do not overlap, restricting extant gene flow among these main clusters.

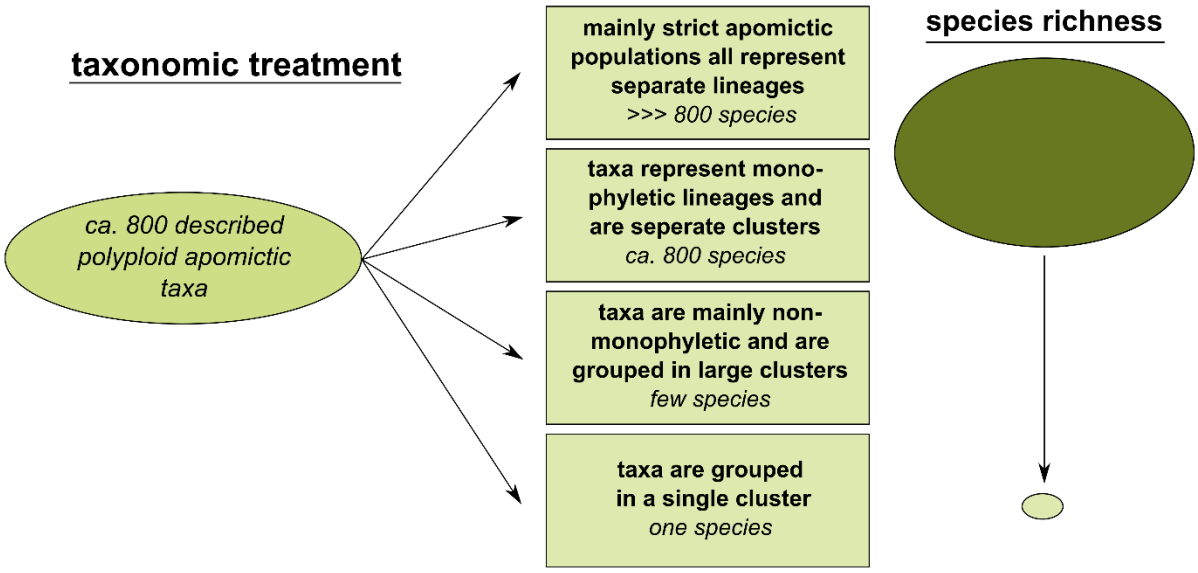


Figure 7. Different taxonomic treatments of the more than 800 described polyploid apomictic taxa within the *R. auricomus* complex. Treatments have different impacts on species richness of the complex. According to results presented in this thesis, the "few species" concept is currently the most likely scenario.

In Karbstein et al. (2020b), five sexual species were re-classified based on the combination of an evolutionary lineage (De Queiroz, 2007; Sukumaran & Knowles, 2017), a genetic, and a morphological cluster concept (Mallet, 1995), yielding a robust phylogenetic framework for polyploid analyses. However, findings of this thesis indicate that the application of an evolutionary lineage concept for apomictic polyploid *R. auricomus* taxa is not appropriate. The extremely low BT values and quartet support across middle and shallow tree branches (chapter 5: Karbstein et al., 2021b) suggest that the presence of widely distributed, stable or separate apomictic lineages evolving via ordinary speciation is unlikely. Multiple hybridization waves during the Pleistocene, enormous post-origin genome evolution, and the extremely young evolutionary age are probably responsible for the lack of widely distributed, stabilized lineages. Considering a cluster concept, three to five, partly geographically isolated, genetically heterogeneous main clusters might be accepted as species according to (phylo)genomic results (Fig. 7). An additional approach might consist of separating the few monophyletic taxa as agamospecies from the non-monophyletic ones within the main clusters. This is also problematic due to unknown morphological differentiation and the fact that not all monophyletic taxa are facultatively apomictic. Absent versus complete flowers help to distinguish apomicts from sexuals in most cases, but are less informative in discriminating apomictic taxa (Borchers-Kolb, 1983, 1985; Dunkel et al., 2018; chapter 4: Karbstein et al.,

2021a). Phenotypic differentiation is still regarded as an important criterion for species delimitation (e.g., Stuessy, 2009; Freudenstein et al., 2017). However, morphological differentiation is not easily recognizable among the potential species of the main cluster concept. For a final taxonomic treatment (Fig. 7), further research would need to be conducted, for example, geometric morphometrics to study morphological differentiation among main clusters (Karbstein et al., in prep. b) but also (phylo)genomics to clarify or confirm the presence of main clusters along the range margins of the European *R. auricomus* complex.

Conclusions and Outlook

The first steps towards understanding the speciation process of young species complexes accompanied by polyploidy and hybridization using OMICS data have been made. The thesis demonstrated that only an integrative approach with large amounts of data, from genomic, flow cytometry, morphometry to environmental data including final taxonomic treatments (Fig. 6), is able to decipher polyploid plant evolution, as already proposed by Hörandl (2018), Pinheiro et al. (2018), and Rothfels (2021). The theoretical approaches and practical implementations of this thesis allowed us to unravel the polyploid evolution and biogeography of the large *R. auricomus* species complex, a breakthrough in the context of more than 200 years of debate and research on this topic (Linnaeus, 1753; e.g., Reichenbach, 1832; Marklund, 1961, 1965; Ericsson, 1992; Hörandl, 1998; Paun et al., 2006a,b; Hörandl et al., 2009; Hodač et al., 2014; Dunkel et al., 2018). Five sexual progenitors generated predominantly allotetraploid apomicts characterized by subgenome dominance and tremendous post-origin evolution during Pleistocene deterioration (chapters 1, 3, 5; Karbstein et al., 2020b, 2021b; Tomasello et al., 2020). Sexuals have restricted ranges due to negative effects of their breeding system in small and/or isolated populations, whereas polyploid apomicts probably benefit from increased heterozygosity and uniparental reproduction, leading to higher environmental resistance and better colonizing abilities (chapters 2, 4: Karbstein et al., 2020c, 2021a).

The here presented phylogenomic relationships, genomic composition and diversity, and preliminary taxonomic treatments and implications give a first overview of the species complex evolution. However, steadily new software is developed in the field of polyploid phylogenetics and further detailed analyses are needed to study evolutionary processes of the complex across the entire European range. Consequently, the implementation of phased data sets has the highest

priority towards forwarding our understanding of the evolutionary processes within the *R. auricomus* complex. For instance, to enable the use of the best-suited polyploid (phylo)genomic methods, current phasing, allele sorting, and subgenome assignment methods (e.g., HOMOLOGIZER: Freyman et al., 2020; ALLCOPOL: Lautenschlager et al., 2020; Šlenker et al., 2021; PATE: Tiley et al., 2021) should be implemented based on the already available long-read nuclear gene data. Multilabeled trees, where each subgenome is represented by a tip, better illustrate the distinct evolutionary histories of subgenomes unified in allopolyploid individuals (Oxelman et al., 2017; Rothfels, 2021). The method would also enable the recognition of autopolyploids where two subgenomes would group close to one progenitor species. Afterward, more sophisticated network reconstructions appropriate for allopolyploid analyses (e.g., ALLOPPNET: Jones, 2017b) incorporating all sexual progenitor genomes and sorted subgenomes of allotetraploids can be applied to verify the progenitor-polyploid relationships presented in this thesis. If the only observed mixed sexual-apomictic tetraploid population (KK166) represents another sexual progenitor species (Karbstein et al., in prep. a), it should also be incorporated into network reconstructions.

Despite recent progress in transcriptome and genome sequencing, reduced representation methods (RADseq, target enrichment) represent the nuclear genome well, and are still easier to collect, cheaper to sequence, and easier to analyze for large sample sizes (McKain et al., 2018; Johnsen et al., 2019; chapter 5: Karbstein et al., 2021b). Nevertheless, deeply sequenced sexual (Karbstein et al., in prep. c) and apomictic *R. auricomus* genomes would be desirable, for instance, for reference-guided RADseq read assembly, inference of coding- and non-coding RADseq reads (including functionality), or exploration of asexual genome evolution. A promising novel approach is present in highly accurate long-read sequencing, yielding 10-25 kb read lengths with an extremely high accuracy of 99.5% (i.e., PacBio HiFi sequencing; Hon et al., 2020). In addition, CRISPR/Cas9 gene-editing methods are currently also applicable for model and non-model polyploid systems (e.g., Shan et al., 2018, 2020). These methods would allow unique insights into gene or cytonuclear interactions and phenotypic consequences in polyploids, for instance enabling us to understand the rise of highly morphologically diverse polyploid *R. auricomus* apomicts out of a few sexual progenitor species in detail.

Besides future (phylo)genomic research, the analysis of highly polyploid populations situated on the range margin of the *R. auricomus* complex has a potentially large impact on biogeographical relationships and taxonomic conclusions. For instance, associations found in Karbstein et al. (2021a) predict increased obligate apomixis in Northern Scandinavia, suggesting higher lineage stability due to missing gene flow and thus stable, separately evolving

lineages. However, observed percentages of facultative sexuality within the *R. auricomus* complex are usually low (0-34%, mean 2%; chapter 4: Karbstein et al., 2021a). It must be observed in small-scale population genetic studies (e.g., using RADseq data) how much facultative sexuality is needed to allow gene flow among apomictic populations in close sympatry. Finally, a general taxonomic treatment (Figs. 6, 7) should also include geometric morphometric analyses, potentially combined with nuclear-gene-based networks (Bastide et al., 2018a,b; Karbstein et al., in prep. b), in order to formulate a final taxonomic concept of the European *R. auricomus* complex.

IV References

General References

- General Introduction and General Discussion -

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*, 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Adams, K. L., & Wendel, J. F. (2005). Polyploidy and genome evolution in plants. *Current Opinion in Plant Biology*, *8*, 135–141. https://doi.org/10.1016/j.pbi.2005.01.001

Agapow, P., Bininda-Emonds, O. R. P., Crandall, K. A., Gittleman, J. L., Mace, G. M., Marshall, J. C., & Purvis, A. (2004). The impact of species concept on biodiversity studies. *The Quarterly Review of Biology*, *79*, 161–179. https://doi.org/10.1086/383542

Ainouche, M. L., Baumel, A., Salmon, A., & Yannic, G. (2004). Hybridization, polyploidy and speciation in *Spartina* (Poaceae). *New Phytologist*, *161*, 165–172. https://doi.org/10.1046/j.1469-8137.2003.00926.x

Alger, E. I., & Edger, P. P. (2020). One subgenome to rule them all: Underlying mechanisms of subgenome dominance. *Current Opinion in Plant Biology*, *54*, 108–113. https://doi.org/10.1016/j.pbi.2020.03.004

Ali, I. B. E. H., Guetat, A., & Boussaid, M. (2012). Genetic diversity and structure of wild Tunisian *Thymus capitatus* (L.) Hoffm. et Link. (Lamiaceae) assessed using isozyme markers. *African Journal of Ecology*, 50, 140–151. https://doi.org/10.1111/j.1365-2028.2011.01304.x

Alonso-Marcos, H., Hülber, K., Myllynen, T., Pérez Rodríguez, P., & Dobeš, C. (2018). Pollen precedence in sexual *Potentilla puberula* and its role as a protective reproductive barrier against apomictic cytotypes. *Taxon*, *67*, 1132–1142. https://doi.org/10.12705/676.9

Alonso-Marcos, H., Nardi, F. D., Scheffknecht, S., Tribsch, A., Hülber, K., & Dobeš, C. (2019). Difference in reproductive mode rather than ploidy explains niche differentiation in sympatric sexual and apomictic populations of *Potentilla puberula*. *Ecology and Evolution*, *9*, 3588–3598. https://doi.org/10.1002/ece3.4992

Andermann, T., Fernandes, A. M., Olsson, U., Töpel, M., Pfeil, B., Oxelman, B., ... Antonelli, A. (2019). Allele phasing greatly improves the phylogenetic utility of ultraconserved elements. *Systematic Biology*, *68*, 32–46. https://doi.org/10.1093/sysbio/syy039

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. Cambridge, UK: Babraham Bioinformatics, Babraham Institute. Retrieved from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Arrigo, N., & Barker, M. S. (2012). Rarely successful polyploids and their legacy in plant genomes. *Current Opinion in Plant Biology*, *15*, 140–146. https://doi.org/10.1016/j.pbi.2012.03.010

Asker, S., & Jerling, L. (1992). *Apomixis in plants*. CRC Press.

Ayres, D. R., Zaremba, K., Sloop, C. M., & Strong, D. R. (2007). Sexual reproduction of cordgrass hybrids (*Spartina foliosa* x *alterniflora*) invading tidal marshes in San Francisco Bay. *Diversity and Distributions*, *14*, 187–195. https://doi.org/10.1111/j.1472-4642.2007.00414.x

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, *3*, 1–7. https://doi.org/10.1371/journal.pone.0003376

Baker, H. G. (1967). Support for Baker's law –As a rule. *Evolution*, *21*, 853–856. https://doi.org/10.1017/CBO97 81107415324.004

Barke, B. H., Daubert, M., & Hörandl, E. (2018). Establishment of apomixis in diploid F_2 hybrids and inheritance of apospory from F_1 to F_2 hybrids of the *Ranunculus auricomus* complex. *Frontiers in Plant Science*, *9*, 1–12.

Barke, B. H., Karbstein, K., Daubert, M., & Hörandl, E. (2020). The relation of meiotic behaviour to hybridity, polyploidy and apomixis in the *Ranunculus auricomus* complex (Ranunculaceae). *BMC Plant Biology*, *20*, 523. https://doi.org/10.1186/s12870-020-02654-3

Barker, M. S., Arrigo, N., Baniaga, A. E., Li, Z., Levin, D. A. (2016). On the relative abundance of autopolyploids and allopolyploids. *New Phytologist*, *210*, 391–398. https://doi.org/10.1111/nph.13698

Bello Rodriguez, J. C., Hausbeck, M., & Sakalidis, M. L. (2021). Application of target enrichment sequencing for population genetic analyses of the obligate plant pathogens *Pseudoperonospora cubensis* and *P. humuli* in Michigan. *Molecular Plant-Microbe Interactions*, 1–65. https://doi.org/10.1094/MPMI-11-20-0329-TA

Bierzychudek, P. (1985). Patterns in plant parthenogenesis. *Experientia*, *41*, 1255–1264. https://doi.org/10.1007/BF019 52068

Bird, K. A., Niederhuth, C. E., Ou, S., Gehan, M., Pires, J. C., Xiong, Z., … Edger, P. P. (2021). Replaying the evolutionary tape to investigate subgenome dominance in allopolyploid *Brassica napus*. *New Phytologist*, *230*, 354–371. https://doi.org/10.1111/nph.17137

Blaine Marchant, D. B., Soltis, D. E., & Soltis, P. S. (2016). Patterns of abiotic niche shifts in allopolyploids relative to their progenitors. *New Phytologist*, *212*, 708–718. https://doi.org/10.1111/nph.14069

Blischak, P. D., Mabry, M. E., Conant, G. C., & Pires, J. C. (2018). Integrating networks,

phylogenomics, and population genomics for the study of polyploidy. *Annual Review of Ecology, Evolution, and Systematics*, *49*, 253–278. https://doi.org/10.1146/annurev-ecolsys-121415-032302

Borchers-Kolb, E. (1983). *Ranunculus* sect. *Auricomus* in Bayern und den angrenzenden Gebieten – I. Allgemeiner Teil. *Mitteilungen der Botanischen Staatssammlung München*, *19*, 363–429.

Borchers-Kolb, E. (1985). *Ranunculus* sect. *Auricomus* in Bayern und den angrenzenden Gebieten II. Spezieller Teil. *Mitteilungen der Botanischen Staatssammlung München*, *21*, 49–300.

Brandrud, M. K., Baar, J., Lorenzo, M. T., Athanasiadis, A., Bateman, R. M., Chase, M. W., ... Paun, O. (2020). Phylogenomic relationships of diploids and the origins of allotetraploids in *Dactylorhiza* (Orchidaceae). *Systematic Biology*, *69*, 91–109. https://doi.org/10.1093/sysbio/syz035

Brandt, A., Schaefer, I., Glanz, J., Schwander, T., Maraun, M., Scheu, S., & Bast, J. (2017). Effective purifying selection in ancient asexual oribatid mites. *Nature Communications*, *8*, 873. https://doi.org/10.1038/s41467-017-01002-8

Brewer, S., Cheddadi, R., de Beaulieu, J.-L., Reille, M., & Contributors, D. (2002). The spread of deciduous *Quercus* throughout Europe since the last glacial period. *Forest Ecology and Management*, *156*, 27–48. https://doi.org/10.1016/S0378-1127(01)00646-6

Brochmann, C., Brysting, A. K., Alsos, I. G., Borgen, L., Grundt, H. H., Scheen, A.-C., & Elven, R. (2004). Polyploidy in arctic plants. *Biological Journal of the Linnean Society*, *82*, 521–536. https://doi.org/10.1111/j.1095-8312.2004.00337.x

Brukhin, V., Osadtchiy, J. V., Florez-Rueda, A. M., Smetanin, D., Bakin, E., Nobre, M. S., Grossniklaus, U. (2019). The *Boechera* genus as a resource for apomixis research. *Frontiers in Plant Science*, *10*, 19. https://doi.org/10.3389/fpls.2019.00392

Buggs, R. J. A. (2021). The origin of Darwin's "abominable mystery." *American Journal of Botany*, *108*, 22–36. https://doi.org/10.1002/ajb2.1592

Buggs, R. J. A., Soltis, P. S., & Soltis, D. E. (2011). Biosystematic relationships and the formation of polyploids. *Taxon*, *60*, 324–332. https://doi.org/10.1002/tax.602003

Burgess, M. B., Cushman, K. R., Doucette, E. T., Frye, C. T., Campbell, C. S. (2015). Understanding diploid diversity: A first step in unraveling polyploid, apomictic complexity in *Amelanchier*. *American Journal of Botany*, *102*, 2041–2057. https://doi.org/10.3732/ajb.1500330

Burgess, M. B., Cushman, K. R., Doucette, E. T., Talent, N., Frye, C. T., & Campbell, C. S. (2014). Effects of apomixis and polyploidy on diversification and geographic distribution in *Amelanchier* (Rosaceae). *American Journal of Botany*, *101*, 1375–1387. https://doi.org/10.3732/ajb.1400113

Carman, J. G. (1997). Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony. *Biological Journal of the Linnean Society*, *61*, 51–94. https://doi.org/10.1006/bijl.1996.0118.

Carter, K. A., Liston, A., Bassil, N. V., Alice, L. A., Bushakra, J. M., Sutherland, B. L., … Hummer, K. E. (2019). Target capture sequencing unravels *Rubus* evolution. *Frontiers in Plant Science*, *10*, 1615. https://doi.org/10.3389/fpls.2019.01615

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). STACKS: an analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140. https://doi.org/10.1111/mec.12354

Cavender-Bares, J., González-Rodríguez, A., Eaton, D. A. R., Hipp, A. A. L., Beulke, A., & Manos, P. S. (2015). Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): A genomic and population genetics approach. *Molecular Ecology*, *24*, 3668–3687. https://doi.org/10.1111/mec.13269

Chapman, M. A., & Burke, J. M. (2007). Genetic divergence and hybrid speciation. *Evolution*, *61*, 1773–1780. https://doi.org/10.1111/j.1558-5646.2007.00134.x

Coates, D. J., Byrne, M., & Moritz, C. (2018). Genetic diversity and conservation units: Dealing with the species-population continuum in the age of genomics. *Frontiers in Ecology and Evolution*, *6*, 1–13. https://doi.org/10.3389/fevo.2018.00165

Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nature Reviews Genetics*, *6*, 836–846. https://doi.org/10.1038/nrg1711

Corlett, R. T. (2016). Plant diversity in a changing world: Status, trends, and conservation needs. *Plant Diversity*, *38*, 10–16. https://doi.org/10.1016/j.pld.2016.01.001

Cosendai, A. C., Wagner, J., Ladinig, U., Rosche, C., & Hörandl, E. (2013). Geographical parthenogenesis and population genetic structure in the alpine species *Ranunculus kuepferi* (Ranunculaceae). *Heredity*, *110*, 560–569. https://doi.org/10.1038/hdy.2013.1

Cosendai, A. C., & Hörandl, E. (2010). Cytotype stability, facultative apomixis and geographical parthenogenesis in *Ranunculus kuepferi* (Ranunculaceae). *Annals of Botany*, *105*, 457–470. https://doi.org/10.1093/aob/mcp304

Coughlan, J. M., Han, S., Stefanović, S., & Dickinson, T. A. (2017). Widespread generalist clones are associated with range and niche expansion in allopolyploids of Pacific Northwest Hawthorns (*Crataegus* L.). *Molecular Ecology*, *26*, 5484–5499. https://doi.org/10.1111/mec.14331

Darlington, C. D. (1937). *Recent advances in cytology* (sec ed). Philadelphia, USA: Blackiston.

Dauphin, B., Grant, J. R., Farrar, D. R., & Rothfels, C. J. (2018). Rapid allopolyploid radiation of moonwort ferns (*Botrychium*; Ophioglossaceae) revealed by PacBio sequencing of homologous and homeologous nuclear regions. *Molecular Phylogenetics and Evolution*, *120*, 342–353. https://doi.org/10.1016/j.ympev.2017.11.025

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Review Genetics*, *12*, 499–510. https://doi.org/10.1038/nrg3012

Dawkins, R. (2008). *Das egoistische Gen* (jubilee ed). https://doi.org/10.1007/978-3-642-55391-2

De Bodt, S., Maere, S., & Van de Peer, Y. (2005). Genome duplication and the origin of angiosperms. *Trends in Ecology and Evolution*, *20*, 591–597. https://doi.org/10.1016/j.tree.2005.07.008

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, *56*, 879–886. https://doi.org/10.1080/10635150701701083

Dobeš, C., Mitchell-Olds, T., & Koch, M. A. (2004). Intraspecific diversification in North American *Boechera stricta* (= *Arabis drummondii*), *Boechera x divaricarpa*, and *Boechera holboellii* (Brassicaceae) inferred from nuclear and chloroplast molecular markers – an integrative approach. *American Journal of Botany*, *91*, 2087–2101. https://doi.org/10.3732/ajb.91.12.2087

Dunkel, F. G. (2019). The *Ranunculus auricomus* L. complex (Ranunculaceae) in Slovenia. *Stapfia*, *111*, 33–91.

Dunkel, F. G., Gregor, T., & Paule, J. (2018). New diploid species in the *Ranunculus auricomus* complex (Ranunculaceae) from W and SE Europe. *Willdenowia*, *48*, 227–257. https://doi.org/10.3372/wi.48.48205

Dunkel, F. G. (2010). The *Ranunculus auricomus* L. complex (Ranunculaceae) in Northern Italy. *Webbia*, *65*, 179–227. https://doi.org/10.1080/00837792.2010.10670873

Dunkel, F. G. (2015). *Ranunculus pindicola* sp. nov., the only species of the *R. auricomus* complex (Ranunculaceae) in Greece. *Willdenowia*, *45*, 22–230. https://doi.org/10.3372/wi.45.45208

Dunkel, F. G. (2011). The *Ranunculus auricomus* L. complex (Ranunculaceae) in Central and Southern Italy, with additions to North Italian taxa. *Webbia*, *66*, 165–193. https://doi.org/10.1080/00837792.2011.10670895

Eaton, D. A. R. (2014). PYRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*, *30*, 1844–1849. https://doi.org/10.1093/bioinformatics/btu121

Eaton, D. A. R., Spriggs, E. L., Park, B., & Donoghue, M. J. (2017). Misconceptions on missing data in RADseq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology*, *66*, 399–412. https://doi.org/10.1093/sysbio/syw092

Eaton, D. A. R., & Overcast, I. (2020). IPYRAD: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, *36*, 2592–2594. https://doi.org/10.1093/bioinformatics/btz966

Ebdon, S., Laetsch, D. R., Dapporto, L., Hayward, A., Ritchie, M. G., Dincă, V., … Lohse, K. (2021). The Pleistocene species pump past its prime: Evidence from European butterfly sister species. *Molecular Ecology*, *30*, 3575–3589. https://doi.org/10.1111/mec.15981

Edger, P. P., McKain, M. R., Bird, K. A., & VanBuren, R. (2018). Subgenome assignment in allopolyploids: Challenges and future directions. *Current Opinion in Plant Biology*, *42*, 76–80. https://doi.org/10.1016/j.pbi.2018.03.006

Edger, P. P., Smith, R., McKain, M. R., Cooley, A. M., Vallejo-Marin, M., Yuan, Y., … Puzey, J. R. (2017). Subgenome dominance in an interspecific hybrid, synthetic allopolyploid, and a 140-year-old naturally established neo-allopolyploid monkeyflower. *The Plant Cell*, *29*, 2150–2167. https://doi.org/10.1105/tpc.17.00010

Emadzade, K., Gehrke, B., Linder, H. P., & Hörandl, E. (2011). The biogeographical history of the cosmopolitan genus *Ranunculus* L. (Ranunculaceae) in the temperate to meridional zones. *Molecular Phylogenetics and Evolution*, *58*, 4–21. https://doi.org/10.1016/j.ympev.2010.11.002

Emadzade, K., Lebmann, M., Hoffmann, M. H., Tkach, N., Lone, F., & Hörandl, E. (2015). Phylogenetic relationships and evolution of high mountain buttercups (*Ranunculus*) in North America and Central Asia. *Perspectives in Plant Ecology, Evolution and Systematics*, *17*, 131–141. https://doi.org/10.1016/j.ppees.2015.02.001

Ericsson, S. (1992). The microspecies of the *Ranunculus auricomus* complex treated at the species level. *Annales Botanici Fennici*, *29*, 123–158.

Eriksson, J. S., de Sousa, F., Bertrand, Y. J. C., Antonelli, A., Oxelman, B., & Pfeil, B. E. (2018). Allele phasing is critical to revealing a shared allopolyploid origin of *Medicago arborea* and *M. strasseri* (Fabaceae). *BMC Evolutionary Biology*, *18*, 9. https://doi.org/10.1186/s1286 2-018-1127-z

Ernst, A. (1918). *Bastardierung als Ursache der Apogamie im Pflanzenreich: eine Hypothese zur experimentellen Vererbungs- und Abstrammungslehre*. Jena, Germany: G. Fischer.

Fellers, J. P. (2008). Genome filtering using methylation-sensitive restriction enzymes with six base pair recognition sites. *The Plant Genome*, *1*, 146–152. https://doi.org/10.3835/plantgenome2008.05.0245

Folk, R. A., Mandel, J. R., & Freudenstein, J. V. (2015). A protocol for targeted enrichment of intron-containing sequence markers for recent radiations: A phylogenomic example from *Heuchera* (Saxifragaceae). *Applications in Plant Sciences*, *3*, 1500039. https://doi.org/10.3732/apps.1500039

Fox, D. T., Soltis, D. E., Soltis, P. S., Ashman, T. L., Van de Peer, Y. (2020). Polyploidy: a biological force from cells to ecosystems. *Trends in Cell Biology*, *30*, 688–694. https://doi.org/10.1016/j.tcb.2020.06.006

Freeland, J. R., Kirk, H., & Petersen, S. D. (2011). *Molecular Ecology* (sec ed). Chichester,

England: Wiley. https://doi.org/10.1002/9780470979365

Freudenstein, J. V., Broe, M. B., Folk, R. A., & Sinn, B. T. (2017). Biodiversity and the species concept – Lineages are not enough. *Systematic Biology*, *66*, 644–656. https://doi.org/10.1093/sysbio/syw098

Freyman, W. A., Johnson, M. G., & Rothfels, C. J. (2020). HOMOLOGIZER: Phylogenetic phasing of gene copies into polyploid subgenomes. *BioRxiv*, 1–55. https://doi.org/10.1101/2020.10.22.351486

Frichot E., & Francois O. (2020). LEA: An R package for landscape and ecological association studies. Retrieved from https://bioconductor.org/packages/release/bioc/html/LEA.html

Frichot E., & Francois O. (2020). LEA: An R package for landscape and ecological association studies. Retrieved from https://bioconductor.org/packages/release/bioc/html/LEA.html

Friedman, W. E. (2009). The meaning of Darwin's "abominable mystery." *American Journal of Botany*, *96*, 5–21. https://doi.org/10.3732/ajb.0800150

Hodel, R. G. J., Zimmer, E., & Wen, J. (2021). A phylogenomic approach resolves the backbone of *Prunus* (Rosaceae) and identifies signals of hybridization and allopolyploidy. *Molecular Phylogenetics and Evolution*, *160*, 107118. https://doi.org/10.1016/j.ympev.2021.107118

Gates, R. R. (1909). The behavior of the chromosomes in *Oenothera lata* x *O. gigas*. *Botanical Gazette*, *48*, 179–199. https://doi.org/10.1093/genetics/3.4.309

GBIF Secretariat (2017). *Ranunculus auricomus* L. https://doi.org/10.15468/39omei

Geml, J., Laursen, G. A., O'Neill, K., Nusbaum, H. C., & Taylor, D. L. (2006). Beringian origins and cryptic speciation events in the fly agaric (*Amanita muscaria*). *Molecular Ecology*, *15*, 225–239. https://doi.org/10.1111/j.1365-294X.2005.02799.x

Gordon, S. P., Contreras-Moreira, B., Levy, J. J., Djamei, A., Czedik-Eysenberg, A., Tartaglio, V. S., … Vogel, J. P. (2020). Gradual polyploid genome evolution revealed by pan-genomic analysis of *Brachypodium hybridum* and its diploid progenitors. *Nature Communications*, *11*, 1–16. https://doi.org/10.1038/s41467-020-17302-5

Gornall, R. J. (1999). Population genetic structure in agamospermous plants. In P. M. Hollingsworth, R. M. Bateman, & R. J. Gornall (Eds.), *Molecular Systematics and Plant Evolution* (pp. 118–138). London, UK: Taylor & Francis.

Gramlich, S., Wagner, N. D., & Hörandl, E. (2018). RAD-seq reveals genetic structure of the F2-generation of natural willow hybrids (*Salix* L.) and a great potential for interspecific introgression. *BMC Plant Biology*, *18*, 317. https://doi.org/10.1186/s12870-018-1552-6

Grant, V. (1981). *Plant speciation* (sec ed). New York, Columbia University Press.

Grau, J. (1984). Vorläufige Übersicht der iberischen Vertreter von *Ranunculus* sect. *Auricomus*. *Mitteilungen der Botanischen Staatssammlung München*, *20*, 11–28.

Grimanelli, D. (2012). Epigenetic regulation of reproductive development and the emergence of apomixis in angiosperms. *Current Opinion in Plant Biology*, *15*, 57–62. doi: https://doi.org/10.1016/j.pbi.2011.10.002

Haag, C. R., & Ebert, D. (2004). A new hypothesis to explain geographic parthenogenesis. *Annales Zoologici Fennici*, *41*, 539–544.

Häfliger, E. (1943). Zytologisch-embryologische Untersuchungen pseudogamer Ranunkeln der *Auricomus*-Gruppe. *Berichte der Schweizerischen Botanischen Gesellschaft*, *53*, 317–382.

Hand, M. L., & Koltunow, A. M. G. (2014). The genetic control of apomixis: asexual seed formation. *Genetics*, *197*, 441–450. https://doi.org/10.1534/genetics.114.163105

Hartmann, M., Štefánek, M., Zdvořák, P., Heřman, P., Chrtek, J., & Mráz, P. (2017). The Red Queen hypothesis and geographical parthenogenesis in the alpine hawkweed *Hieracium alpinum* (Asteraceae). *Biological Journal of the Linnean Society*, *122*, 681–696. https://doi.org/10.1093/biolinnean/blx105

Haveman, R., Schaminée, J. H. J., & Weeda, E. J. (2002). Apomicten: het belang van een genuanceerde taxonomie voor plantensociologisch onderzoek en natuurbeheer. *Stratiotes*, *25*, 3–25.

Hedrén, M., Nordström, S., & Ståhlberg, D. (2008). Polyploid evolution and plastid DNA variation in the *Dactylorhiza incarnata* / *maculata* complex (Orchidaceae) in Scandinavia. *Molecular Ecology*, *17*, 5075–5091. https://doi.org/10.1111/j.1365-294X.2008.03965.x

Heywood, J. (1991). Spatial analysis of genetic variation in plant populations. *Annual Reviews of Ecology and Systematics*, *22*, 335–355. https://doi.org/10.1146/annurev.ecolsys.22.1.335

Hodač, L., Barke, B. H., & Hörandl, E. (2018). Mendelian segregation of leaf phenotypes in experimental F2 hybrids elucidates origin of morphological diversity of the apomictic *Ranunculus auricomus* complex. *Taxon*, *67*, 1082–1092. https://doi.org/10.12705/676.6

Hodač, L., Klatt, S., Hojsgaard, D., Sharbel, T. F., & Hörandl, E. (2019). A little bit of sex prevents mutation accumulation even in apomictic polyploid plants. *BMC Evolutionary Biology*, *19*, 170. https://doi.org/10.1186/s12862-019-1495-z

Hodač, L., Scheben, A. P., Hojsgaard, D., Paun, O., & Hörand, E. (2014). ITS polymorphisms shed light on hybrid evolution in apomictic plants: A case study on the *Ranunculus auricomus* complex. *PLoS ONE*, *9*, 28–30. https://doi.org/10.1371/journal.pone.0103003

Hojsgaard, D. (2021). Molecular basis of apomixis in plants. *Genes*, *12*, 576. https://doi.org/10.3390/genes12040576

Hojsgaard, D., Greilhuber, J., Pellino, M., Paun, O., Sharbel, T. F., & Hörandl, E. (2014b). Emergence of apospory and bypass of meiosis via apomixis after sexual hybridisation and polyploidisation. *New Phytologist*, *204*, 1000–1012. https://doi.org/10.1111/nph.12954

Hojsgaard, D., & Hörandl, E. (2019). The rise of apomixis in natural plant populations. *Frontiers in Plant Science*, *10*, 358. https://doi.org/10.3389/fpls.2019.00358

Hojsgaard, D., Klatt, S., Baier, R., Carman, J. G., & Hörandl, E. (2014a). Taxonomy and biogeography of apomixis in angiosperms and associated biodiversity characteristics. *Critical Reviews in Plant Sciences*, *33*, 414–427. https://doi.org/10.1080/07352689.2014.898488

Hon, T., Mars, K., Young, G., Tsai, Y.-C., Karalius, J. W., Landolin, J. M., … Rank, D. R. (2020). Highly accurate long-read HiFi sequencing data for five complex genomes. *Scientific Data*, *7*, 399. https://doi.org/10.1038/s41597-020-00743-4

Hörandl, E. (1998). Species concepts in agamic complexes: Applications in the *Ranunculus auricomus* complex and general perspectives. *Folia Geobotanica*, *33*, 335–348. https://doi.org/10.1007/BF03216210

Hörandl, E. (2004). Comparative analysis of genetic divergence among sexual ancestors of apomictic complexes using isozyme data. *International Journal of Plant Sciences*, *165*, 615–622. https://doi.org/10.1086/386557

Hörandl, E. (2006). The complex causality of geographical parthenogenesis. *New Phytologist*, *171*, 525–538. https://doi.org/10.1111/j.1469-8137.2006.01769.x

Hörandl, E. (2010). The evolution of self-fertility in apomictic plants. *Sexual Plant Reproduction*, *23*, 73–86. https://doi.org/10.1007/s0049 7-009-0122-3

Hörandl, E. (2018). The classification of asexual organisms: Old myths, new facts, and a novel pluralistic approach. *Taxon*, *67*, 1066–1081. https://doi.org/10.12705/676.5

Hörandl, E., Cosendai, A. C., & Temsch, E. M. (2008). Understanding the geographic distributions of apomictic plants: A case for a pluralistic approach. *Plant Ecology and Diversity*, *1*, 309–320. https://doi.org/10.1080/17550 87080 2351175

Hörandl, E., Dobeš, C., & Lambrou, M. (1997). Chromosomen- und Pollenuntersuchungen an österreichischen Arten des apomiktischen *Ranunculus auricomus*-Komplexes. *Botanica Helvetica*, *107*, 195–209.

Hörandl, E., & Greilhuber, J. (2002). Diploid and autotetraploid sexuals and their relationships to apomicts in the *Ranunculus cassubicus* group: Insights from DNA content and isozyme variation. *Plant Systematics and Evolution*, *234*, 85–100.

Hörandl, E., Greilhuber, J., Klímová, K., Paun, O., Temsch, E., Emadzade, K., & Hodálová, I. (2009). Reticulate evolution and taxonomic concepts in the *Ranunculus auricomus* complex (Ranunculaceae): Insights from analysis of morphological, karyological and molecular data. *Taxon*, *58*, 1194–1216.

https://doi.org/10.1002/tax.584012

Hörandl, E., & Gutermann, W. (1995). Die Bearbeitung der *Ranunculus auricomus*-Gruppe für die 'Flora von Österreich' – ein Werkstattbericht. *Flora Austriacae Novitates*, *2*, 12–27. https://doi.org/10.1016/j.jmapro.2018.05.023; https://www.zobodat.at/pdf/Fl-Austr-Novit_2_0012-0027.pdf

Hörandl, E., & Gutermann, W. (1998a). Der *Ranunculus auricomus*-Komplex in Österreich 2. Die *R. cassubicus*-, *R. monophyllus*- und *R. fallax*-Sammelgruppe. *Botanische Jahrbücher fur Systematik, Pflanzengeschichte und Pflanzengeographie*, *120*, 545–598.

Hörandl, E., & Gutermann, W. (1998b). Zur Kenntnis des *Ranunculus auricomus*-Komplexes in Österreich: Die Arten der *R. phragmiteti*- und *R. indecorus*-Gruppe. *Phyton (Horn)*, *37*, 263–320.

Hörandl, E., & Gutermann, W. (1999). Der *Ranunculus auricomus*-Komplex in Österreich und benachbarten Gebieten 3. Die Arten der *R. latisectus*-, *R. puberulus*-, *R. stricticaulis*- und *R. argoviensis* (*R. auricomus*-Sammelgruppe). *Botanische Jahrbücher fur Systematik, Pflanzengeschichte und Pflanzengeographie*, *121*, 99–138.

Hörandl, E., & Paun, O. (2007). Patterns and sources of genetic diversity in apomictic plants: implications for evolutionary potentials and ecology. In: E. Hörandl, U. Grossniklaus, P. J. Van Dijk, & T. Sharbel (Eds.), *Apomixis: Evolution, Mechanisms and Perspective* (pp. 169–194). Ruggell, Liechtenstein: ARG-Gantner.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*, 254–267. https://doi.org/10.1093/molbev/msj030

Jagannathan, M., Warsinger-Pepe, N., Watase, G. J., & Yamashita, Y. M. (2017). Comparative analysis of satellite DNA in the *Drosophila melanogaster* species complex. *G3 Genes|Genomes|Genetics*, *7*, 693–704. https://doi.org/10.1534/g3.116.035352

Jalas, J., & Suominen, J. (1989). *Atlas florae europaeae*, vol. 8, Nymphaeaceae to Ranunculaceae. The committee for mapping the flora of Europe and Societas Biologica Fennica Vanamo.

Jaron, K. S., Bast, J., Nowell, R. W., Ranallo-Benavidez, T. R., Robinson-Rechavi, M., & Schwander, T. (2021). Genomic features of parthenogenetic animals. *Journal of Heredity*, *112*, 19–33. https://doi.org/10.1093/jhered/esaa031

Jasiewicz, A. (1954). Badania nad jaskrami z cyklu *Auricomi* Owcz. w okolicach Krakowa i w polnocnej czesci Karpat. - De *Ranunculis* e circulo *Auricomi* Owcz. in regione Cracoviensi nec non in Carpatorum parte boreali crescentibus. *Fragmenta Floristica et Geobotanica*, *2*, 62–110.

Johnson, M. G., Pokorny, L., Dodsworth, S., Botigue, L. R., Cowan, R. S., Devault, A., ... Baker, W. J. (2019). A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic Biology*, *68*,

594–606. https://doi.org/10.1093/sysbio/syy086

Jones, G. (2017a). Algorithmic improvements to species delimitation and phylogeny estimation under the multispecies coalescent. *Journal of Mathematical Biology*, *74*, 447–467. https://doi.org/10.1007/s00285-016-1034-0

Jones, G. (2017b). Bayesian phylogenetic analysis for diploid and allotetraploid species networks. *BioRxiv*, 1–15. https://doi.org/10.1101/129361

Julin, E. (1967). Der Formenkreis des *Ranunculus auricomus* L. in Schweden IV. Sippen von *R. auricomus* s. str. aus Oland. *Arkiv för Botanik*, *6*, 1–243.

Julin, E. (1980). *Ranunculus auricomus* L. in Södermanland, East-Central Sweden. *Opera Botanica*, *57*, 1–145.

Kamneva, O. K., Syring, J., Liston, A., & Rosenberg, N. A. (2017). Evaluating allopolyploid origins in strawberries (*Fragaria*) using haplotypes generated from target capture sequencing. *BMC Evolutionary Biology*, *17*, 180. https://doi.org/10.1186/s12862-017-1019-7

Kantama, L., Sharbel, T. F., Schranz, M. E., Mitchell-Olds, T., de Vries, S., & de Jong, H. (2007). Diploid apomicts of the *Boechera holboellii* complex display large-scale chromosome substitutions and aberrant chromosomes. *Proceedings of the National Academy of Sciences*, *104*, 14026–14031. https://doi.org/10.1073/pnas.0706647104

Karbstein, K., Barke, B.H., Pätzold, C., Hörandl, E. (in prep. c). A draft of the *Ranunculus auricomus* genome.

Karbstein, K., **Bradican, J.P., Hodač, L., Tomasello, S., Hörandl, E** (in prep. b). Geometric morphometrics combined with nuclear-gene data elucidates complex differentiation and evolution of diagnostic traits within the *Ranunculus auricomus* complex.

Karbstein, K., Prinz, K., Hellwig, F., & Römermann, C. (2020a). Plant intraspecific functional trait variation is related to within-habitat heterogeneity and genetic diversity in *Trifolium montanum* L. *Ecology and Evolution*, 10, 5015–5033. doi: https://doi.org/10.1002/ece3.6255

Karbstein, K., Rahmsdorf, E., Tomasello, S., Hodač, L., & Hörandl, E. (2020b). Breeding system of diploid sexuals within the *Ranunculus auricomus* complex and its role in a geographical parthenogenesis scenario. *Ecology and Evolution*, *10*, 14435–14450. https://doi.org/10.1002/ece3.7073

Karbstein, K., Tomasello, S., Bradican, J.P., Hodač, L., Hörandl, E (in prep. a). A new tetraploid, sexual-apomictic species of the *Ranunculus auricomus* complex.

Karbstein, K., Tomasello, S., Hodač, L., Dunkel, F. G., Daubert, M., & Hörandl, E. (2020a). Phylogenomics supported by geometric morphometrics reveals delimitation of sexual species within the polyploid apomictic *Ranunculus auricomus* complex (Ranunculaceae). *Taxon*, *69*, 1191–1220. https://doi.org/10.1002/tax.12365

Karbstein, K., Tomasello, S., Hodač, L., Lorberg, E., Daubert, M., & Hörandl, E. (2021a). Moving beyond assumptions: Polyploidy and environmental effects explain a geographical parthenogenesis scenario in European plants. *Molecular Ecology*, *30*, 2659–2675. https://doi.org/10.1111/mec.15919

Karbstein, K., Salvatore, T., Hodač, L., Wagner, N., Marinček, P., Barke, B. H., … Hörandl, E. (2021b). Unraveling the phylogenetic relationships, reticulate evolution, and genome composition of polyploid plant complexes by Rad-Seq and Hyb-Seq. *BioRxiv,* 1–86, https://doi.org/10.1101/2021.08.30.458250.

Karbstein, K., Tomasello, S., & Prinz, K. (2019). Desert-like badlands and surrounding (semi-)dry grasslands of Central Germany promote small-scale phenotypic and genetic differentiation in *Thymus praecox*. *Ecology and Evolution*, *9*, 14066–14084. *Ecology and Evolution*, https://doi.org/10.1002/ece3.5844

Karunarathne, P., Schedler, M., Martínez, E. J., Honfi, A. I., Novichkova, A., & Hojsgaard, D. (2018). Intraspecific ecological niche divergence and reproductive shifts foster cytotype displacement and provide ecological opportunity to polyploids. *Annals of Botany*, *121*, 1183–1196. https://doi.org/10.1093/aob/mcy004

Kearney, M. (2005). Hybridization, glaciation and geographical parthenogenesis. *Trends in Ecology and Evolution*, *20*, 495–502. https://doi.org/10.1016/j.tree.2005.06.005

Kirchheimer, B., Schinkel, C. C. F., Dellinger, A. S., Klatt, S., Moser, D., Winkler, M., … Dullinger, S. (2016). A matter of scale: Apparent niche differentiation of diploid and tetraploid plants may depend on extent and grain of analysis. *Journal of Biogeography*, *43*, 716–726. https://doi.org/10.1111/jbi.12663

Kirchheimer, B., Wessely, J., Gattringer, A., Hülber, K., Moser, D., Schinkel, C. C. F., … Dullinger, S. (2018). Reconstructing geographical parthenogenesis: Effects of niche differentiation and reproductive mode on Holocene range expansion of an alpine plant. *Ecology Letters*, *21*, 392–401. https://doi.org/10.1111/ele.12908

Kirschner, J., Záveská Drábková, L., Štěpánek, J., & Uhlemann, I. (2015). Towards a better understanding of the *Taraxacum* evolution (Compositae–Cichorieae) on the basis of nrDNA of sexually reproducing species. *Plant Systematics and Evolution*, *301*, 1135–1156. https://doi.org/10.1007/s00606-014-1139-0

Klatt, S., Hadacek, F., Hodač, L., Brinkmann, G., Eilerts, M., Hojsgaard, D., & Hörandl, E. (2016). Photoperiod extension enhances sexual megaspore formation and triggers metabolic reprogramming in facultative apomictic *Ranunculus auricomus*. *Frontiers in Plant Science*, *7*, 278. https://doi.org/10.3389/fpls.2016.00278

Koch, W. (1939). Zweiter Beitrag zur Kenntnis des Formenkreises von *Ranunculus auricomus* L.: Studien über kritische Schweizerpflanzen III. *Berichte der Schweizerischen Botanischen Gesellschaft*, *49*, 541–554.

Koltunow, A. M., & Grossniklaus, U. (2003). Apomixis: A developmental perspective. *Annual Review of Plant Biology*, *54*, 547–574. https://doi.org/10.1146/annurev.arplant.54.110901.160842

Kozak, K. M., Joron, M., McMillan, W. O., & Jiggins, C. D. (2021). Rampant genome-wide admixture across the *Heliconius* radiation. *Genome Biology and Evolution*, *13*, 1–17. https://doi.org/10.1093/gbe/evab099

Krak, K., Caklová, P., Chrtek, J., & Fehrer, J. (2013). Reconstruction of phylogenetic relationships in a highly reticulate group with deep coalescence and recent speciation (*Hieracium*, Asteraceae). *Heredity*, *110*, 138–151. https://doi.org/10.1038/hdy.2012.100

Kvist, G. (1987). Identity of the material of the *Ranunculus auricomus* group in the Linnaean Herbarium. *Annales Botanici Fennici*, *24*, 73–77.

Landis, J. B., Soltis, D. E., Zheng, L., Marx, H. E., Barker, M. S., Tank, D. C., & Soltis, P. S. (2018). Impact of whole-genome duplication events on diversification rates in angiosperms. *American Journal of Botany*, *105*, 348–363. https://doi.org/10.1002/ajb2.1060

Lautenschlager, U., Wagner, F., & Oberprieler, C. (2020). AllCoPol: Inferring allele co-ancestry in polyploids. *BMC Bioinformatics*, *21*, 444. https://doi.org/10.1186/s12859-020-03750-9

Leebens-Mack, J. H., Barker, M. S., Carpenter, E. J., Deyholos, M. K., Gitzendanner, M. A., Graham, S. W., … Wong, G. K. S. (2019). One thousand plant transcriptomes and the phylogenomics of green plants. *Nature*, *574*, 679–685. https://doi.org/10.1038/s41586-019-1693-2

Leslie, A. C. (2018). *Ranunculaceae*. *Section 2*. *Auricomus*. In P. Sell & G. Murrell (Eds.), Flora of Great Britain and Ireland, (pp. 671–677). Cambridge, UK: Cambridge University Press.

Linnaeus, C. (1753). *Species plantarum*. Holmiae [Stockholm]: impensis Laurentii Salvii. https://doi.org/10.5962/bhl.title.669

Liu, B., & Wendel, J. F. (2003). Epigenetic phenomena and the evolution of plant allopolyploids. *Molecular Phylogenetics and Evolution*, *29*, 365–379. https://doi.org/10.1016/S1055-7903(03)00213-6

Lo, E. Y. Y., Stefanovic, S., & Dickinson T. A. (2010). Reconstructing reticulation history in a phylogenetic framework and the potential of allopatric speciation driven by polyploidy in an agamic complex in *Crataegus* (Rosaceae). *Evolution*, *64*, 3593–3608. https://doi.org/10.1111/j.1558-5646.2010.01063.x

Lo, E. Y. Y., Stefanović, S., & Dickinson, T. A. (2013). Geographical parthenogenesis in Pacific Northwest hawthorns (*Crataegus*; Rosaceae). *Botany-Botanique*, *91*, 107–116. https://doi.org/10.1139/cjb-2012-0073

Lughadha, E. N., Govaerts, R., Belyaeva, I., Black, N., Lindon, H., Allkin, R., Magill, R., & Nicolson, N. (2016). Counting counts: Revised estimates of numbers of accepted species of flowering plants, seed plants, vascular plants and land plants with a review of other recent estimates. *Phytotaxa*, *272*, 082–088. https://doi.org/10.11646/phytotaxa.272.1.5

Lutz, A. M. (1907). A preliminary note on the chromosome of *Oenothera lamarckiana* and one of its mutants, *O. gigas*. *Science*, *26*, 151–152.

Mable, B. K., Alexandrou, M. A., & Taylor, M. I. (2011). Genome duplication in amphibians and fish: An extended synthesis. *Journal of Zoology*, *284*, 151–182. https://doi.org/10.1111/j.1469-7998.2011.00829.x

Magri, D. (2008). Patterns of post-glacial spread and the extent of glacial refugia of European beech (*Fagus sylvatica*). *Journal of Biogeography*, *35*, 450–463. https://doi.org/10.1111/j.1365-2699.2007.01803.x

Magri, D., Vendramin, G. G., Comps, B., Dupanloup, I., Geburek, T., Gomory, D., … de Beaulieu, J.-L. (2006). A new scenario for the Quaternary history of European beech populations: Palaeobotanical evidence and genetic consequences. *New Phytologist*, *171*, 199–221. https://doi.org/10.1111/j.1469-8137.2006.01740.x

Malinsky, M., Trucchi, E., Lawson, D. J., & Falush, D. (2018). RADpainter and fineRADstructure: Population inference from RADseq data. *Molecular Biology and Evolution*, *35*, 1284–1290. https://doi.org/10.1093/molbev/msy023

Mallet, J. (1995). A species definition for the Modern Synthesis. *Trends in Ecology and Evolution*, *10*, 294–299.

Marhold, K., & Lihová, J. (2006). Polyploidy, hybridization and reticulate evolution: lessons from the *Brassicaceae*. *Plant Systematics and Evolution*, *259*, 143–174. https://doi.org/10.1007/s00606-006-0417-x

Marklund, G. (1961). Der *Ranunculus auricomus*-Komplex in Finnland I. Diagnosen und Fundortslisten einiger Sippen des *R. auricomus* L. coll. (s. str.). *Flora Fennica*, *3*, 1–128.

Marklund, G. (1965). Der *Ranunculus auricomus*-Komplex in Finnland II: Diagnosen und Fundortslisten einiger Sippen von *R. fallax* (W. et Gr.) Schur, *R. monophyllus* Ovcz. und *R. cassubicus* L. *Flora Fennica*, *4*, 1–198.

Masci, S., Miho, A., & Marchi, P. (1994). *Ranunculus auricomus* L. aggr. (Ranunculaceae) in Italy. I. Sexual tetraploids on the Apennines. *Caryologia*, *47*, 97–108. https://doi.org/10.1080/00087114.1994.10797287

Matzk, F., Meister, A., & Schubert, I. (2000). An efficient screen for reproductive pathways using mature seeds of monocots and dicots. *The Plant Journal*, *21*, 97–108. https://doi.org/10.1046/j.1365-313X.2000.00647.x

Mayrose, I., Zhan, S. H., Rothfels, C. J., Arrigo, N., Barker, M. S., Rieseberg, L. H., & Otto, S. P. (2015). Methods for studying polyploid diversification and the dead-end

hypothesis: A reply to Soltis et al. (2014). *New Phytologist*, *206*, 27–35. https://doi.org/10.1111/nph.13192

McCartney-Melstad, E., Gidiş, M., & Shaffer, H. B. (2019). An empirical pipeline for choosing the optimal clustering threshold in RADseq studies. *Molecular Ecology Resources*, *19*, 1195–1204. https://doi.org/10.1111/1755-0998.13029

McDade, L. A. (1992). Hybrids and phylogenetic systematics. 1. The impact of hybrids on cladistic-analysis. *Evolution*, *46*, 1329–1346. https://doi.org/10.2307/2409940

McKain, M. R., Johnson, M. G., Uribe-Convers, S., Eaton, D., Yang, Y. (2018). Practical considerations for plant phylogenomics. *Applications in Plant Sciences*, *6*, 15. https://doi.org/10.1002/aps3.1038

Melichárková, A., Šlenker, M., Zozomová-Lihová, J., Skokanová, K., Šingliarová, B., Kačmárová T., ... Marhold, K. (2020). So closely related and yet so different: Strong contrasts between the evolutionary histories of species of the *Cardamine pratensis* polyploid complex in Central Europe. *Frontiers in Plant Science*, *11*, 1988. https://doi.org/10.3389/fpls.2020.588856

Meudt, H. M., Albach, D. C., Tanentzap, A. J., Igea, J., Newmarch, S. C., Brandt, A. J., ... Tate, J. A. (2021). Polyploidy on islands: Its emergence and importance for diversification. *Frontiers in Plant Science*, *12*, 637214. https://doi.org/10.3389/fpls.2021.637214

Mogie, M., & Ford, H. (1988). Sexual and asexual *Taraxacum* species. *Biological Journal of the Linnean Society*, *35*, 155–168. https://doi.org/10.1111/j.1095-8312.1988.tb00463.x

Mohammadin, S., Wang, W., Liu, T., Moazzeni, H., Ertugrul, K., Uysal, T., … Schranz, M. E. (2018). Genome-wide nucleotide diversity and associations with geography, ploidy level and glucosinolate profiles in *Aethionema arabicum* (Brassicaceae). *Plant Systematics and Evolution*, *304*, 619–630. https://doi.org/10.1007/s00606-018-1494-3

Mráz, P., Zdvořák, P., Hartmann, M., Štefánek, M., & Chrtek, J. (2018). Can obligate apomixis and more stable reproductive assurance explain the distributional successes of asexual triploids in *Hieracium alpinum* (Asteraceae)? *Plant Biology*, *21*, 227–236. https://doi.org/10.1111/plb.12930

Mráz, P., Chrtek, J., & Šingliarová, B. (2009). Geographical parthenogenesis, genome size variation and pollen production in the arctic-alpine species *Hieracium alpinum*. *Botanica Helvetica*, *119*, 41–51. https://doi.org/10.1007/s00035-009-0055-3

Mráz, P., Filipaş, L., Bărbos, M. I., Kadlecová, J., Paštová, L., Belyayev, A., & Fehrer, J. (2019). An unexpected new diploid *Hieracium* from Europe: Integrative taxonomic approach with a phylogeny of diploid *Hieracium* taxa. *Taxon*, *68*, 1258–1277. https://doi.org/10.1002/tax.12149

Mráz, P., & Mrázová, V. (2021). Greater reproductive assurance of asexual plant compared with sexual relative in a low-density sympatric population: Experimental evidence for

pollen limitation. *Journal of Evolutionary Biology*, *00*, 1–7. https://doi.org/10.1111/jeb.13910

Muller, H. J. (1964). The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *1*, 2–9.

Nardi, F. D., Hülber, K., Moser, D., Alonso-Marcos, H., Tribsch, A., & Dobeš, C. (2020). Occurrence of apomictic conspecifics and ecological preferences rather than colonization history govern the geographic distribution of sexual *Potentilla puberula*. *Ecology and Evolution*, *10*, 7306–7319. https://doi.org/10.1002/ece3.6455

Nieto Feliner, G., Álvarez, I., Fuertes-Aguilar, J., Heuertz, M., Marques, I., Moharrek, F., … Villa-Machío, I. (2017). Is homoploid hybrid speciation that rare? An empiricist's view. *Heredity*, *118*, 513–516. https://doi.org/10.1038/hdy.2017.7

Nogler, G. A. (1984). Genetics of apospory in apomictic *Ranunculus auricomus*. V. Conclusion. *Botanica Helvetica*, *94*, 411–422.

Olave, M., & Meyer, A. (2020). Implementing large genomic single nucleotide polymorphism data sets in phylogenetic reconstructions: A case study of particularly rapid radiations of cichlid fish. *Systematic Biology*, *69*, 848–862. https://doi.org/10.1093/sysbio/syaa005

Oxelman, B., Brysting, A. K., Jones, G. R., Marcussen, T., Oberprieler C., & Pfeil B. E. (2017). Phylogenetics of allopolyploids. *Annual Reviews of Ecology, Evolution, and Systematics*, *48*, 543–557. https://doi.org/10.1146/annurev-ecolsys-110316-022729

Ozias-Akins, P., & van Dijk, P. J. (2007). Mendelian genetics of apomixis in plants. *Annual Review of Genetics*, 41, 509–537. https://doi.org/10.1146/annurev.genet.40.110405.090511

Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: A road map for STACKS. *Methods in Ecology and Evolution*, *8*, 1360–1373. https://doi.org/10.1111/2041-210X.12775

Pätzold, C., Wood, K. R., Eaton, D. A. R., Wagner, W. L., & Appelhans, M. S. (2019). Phylogeny of Hawaiian Melicope (Rutaceae): RADseq resolves species relationships and reveals ancient introgression. *Frontiers in Plant Science*, *10*, 1–16. https://doi.org/10.3389/fpls.2019.01074

Paule, J., Dunkel, F. G., Schmidt, M., & Gregor, T. (2018). Climatic differentiation in polyploid apomictic *Ranunculus auricomus* complex in Europe. *BMC Ecology*, *18*, 1–12. https://doi.org/10.1186/s12898-018-0172-1

Paun, O., Greilhuber, J., Temsch, E. M., & Hörandl, E. (2006a). Patterns, sources and ecological implications of clonal diversity in apomictic *Ranunculus carpaticola* (*Ranunculus auricomus* complex, Ranunculaceae). *Molecular Ecology*, *15*, 897–910. https://doi.org/10.1111/j.1365-294X.2006.02800.x

Paun, O., Stuessy, T. F., & Hörandl, E. (2006b). The role of hybridization, polyploidization and glaciation in the origin and evolution of the apomictic *Ranunculus cassubicus* complex. *New Phytologist*, *171*, 223–236. https://doi.org/10.1111/j.1469-8137.2006.01738.x

Pease, J. B., Brown, J. W., Walker, J. F., Hinchliff, C. E., & Smith, S. A. (2018). Quartet Sampling distinguishes lack of support from conflicting support in the green plant tree of life. *American Journal of Botany*, *105*, 385–403. https://doi.org/10.1002/ajb2.1016

Pellino, M., Hojsgaard, D., Schmutzer, T., Scholz, U., Hörandl, E., Vogel, H., & Sharbel, T. F. (2013). Asexual genome evolution in the apomictic *Ranunculus auricomus* complex: Examining the effects of hybridization and mutation accumulation. *Molecular Ecology*, *22*, 5908–5921. https://doi.org/10.1111/mec.12533

Peralta, M., Combes, M., Cenci, A., Lashermes, P., & Dereeper, A. (2013). SNiPloid: A utility to exploit high-throughput SNP data derived from RNA-Seq in allopolyploid species. *International Journal of Plant Genomics*, *2013*, 1–6. https://doi.org/10.1155/2013/890123

Pinc, J., Chrtek, J., Latzel, V., & Mráz, P. (2020). Negative effect of inbreeding on fitness of an arctic–alpine *Hieracium alpinum* (Asteraceae), a species with a geographical parthenogenesis distribution pattern. *Plant Systematics and Evolution*, *306*, 62. https://doi.org/10.1007/s0060 6-020-01692-6

Pinheiro, F., Dantas-Queiroz, M. V., & Palma-Silva, C. (2018). Plant species complexes as models to understand speciation and evolution: A review of South American studies. *Critical Reviews in Plant Sciences*, *37*, 54–80. https://doi.org/10.1080/07352689.2018.1471565

Qiu, T., Liu, Z., & Liu, B. (2020). The effects of hybridization and genome doubling in plant evolution via allopolyploidy. *Molecular Biology Reports*, *47*, 5549–5558. https://doi.org/10.1007/s11033-020-05597-y

Raab-Straube, E., Hörandl, E., & Nardi, E. (2014). Ranunculaceae. http://ww2.bgbm.org/EuroPlusMed

Ramsey, J., & Schemske, D. W. (1998). Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics*, *29*, 467–501.

Ramsey, J., & Schemske, D. W. (2002). Neopolyploidy in flowering plants. *Annual Review of Ecology and Systematics*, *33*, 589–639. https://doi.org/10.1146/annurev.ecolsys.33.010802.150437

Rannala, B., & Yang, Z. (2003). Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*, *164*, 1645–1656.

Ree, R. H., & Hipp, A. L. (2015). *Inferring phylogenetic history from restriction site associated DNA (RADseq)*. In: Hörandl E., Appelhans M. (Eds.) Next-generation sequencing in plant systematics. Bratislava, IAPT, pp. 181–204.

Reichenbach, L. (1832) ("1830–1832"). *Flora Germanica excursoria*. Lipsiae [Leipzig]: apud Carolum Cnobloch. https://doi.org/10.5962/bhl.title.309

Rice, A., Smarda, P., Novosolov, M., Drori, M., Glick, L., Sabath, N., ... Mayrose, I. (2019). The global biogeography of polyploid plants. *Nature Ecology and Evolution*, *3*, 265–273. https://doi.org/10.1038/s41559-018-0787-9

Richards, A. J. (2003). Apomixis in flowering plants: an overview. *Philosophical Transactions of the Royal Society B*: *Biological Sciences*, *358*, 1085–1093. https://doi.org/10.1098/rstb.2003.1294

Richards, J. A. (1997). *Plant breeding systems* (sec ed). London, UK: Chapman and Hall.

Rothfels, C. J. (2021). Polyploid phylogenetics. *New Phytologist*, *230*, 66–72. https://doi.org/10.1111/nph.17105

Rothfels, C. J., Pryer, K., & Li, F.-W. (2017). Next-generation polyploid phylogenetics: rapid resolution of hybrid polyploid complexes using PacBio single-molecule sequencing. *New Phytologist*, *213*, 413–429. https://doi.org/10.1111/nph.14111

Rousi, A. (1956). Cytotaxonomy and reproduction in the apomictic *Ranunculus auricomus* group. *Annales Botanici Societatis Zoologicae-Botanicae Fennicae "Vanamo"*, *29*, 1–64.

Rushworth, C. A., Windham, M. D., Keith, R. A., & Mitchell-Olds, T. (2018). Ecological differentiation facilitates fine-scale coexistence of sexual and asexual *Boechera*. *American Journal of Botany*, *105*, 2051–2064. https://doi.org/10.1002/ajb2.1201

Šarhanová, P., Vašut, R. J., Dančák, M., Bureš, P., & Trávníček, B. (2012). New insights into the variability of reproduction modes in European populations of *Rubus* subgen. *Rubus*: How sexual are polyploid brambles*? Sexual Plant Reproduction*, *25*, 319–335. https://doi.org/10.1007/s00497-012-0200-9

Schiavinato, M., Bodrug-Schepers, A., Dohm, J. C., & Himmelbauer, H. (2021). Subgenome evolution in allotetraploid plants. *The Plant Journal*, *106*, 672–688. https://doi.org/10.1111/tpj.15190

Schinkel, C. C. F., Kirchheimer, B., Dellinger, A. S., Klatt, S., Winkler, M., Dullinger, S., & Hörandl, E. (2016). Correlations of polyploidy and apomixis with elevation and associated environmental gradients in an alpine plant. *AoB Plants*, *8*, plw064. https://doi.org/10.1093/aobpla/plw064

Schleuning, M., Niggemann, M., Becker, U., & Matthies, D. (2009). Negative effects of habitat degradation and fragmentation on the declining grassland plant *Trifolium montanum*. *Basic and Applied Ecology*, *10*, 61–69. https://doi.org/10.1016/j.baae.2007.12.002

Schmickl, R., Liston, A., Zeisek, V., Oberlander, K., Weitemier, K., Straub, S. C. K., ... Suda, J. (2016). Phylogenetic marker development for target enrichment from transcriptome and genome skim data: The pipeline and its application in southern

African *Oxalis* (Oxalidaceae). *Molecular Ecology Resources*, *16*, 1124–1135. https://doi.org/10.1111/1755-0998.12487

Schmickl, R., Paule, J., Klein, J., Marhold, K., & Koch, M. A. (2012). The evolutionary history of the *Arabidopsis arenosa* complex: Diverse tetraploids mask the Western Carpathian center of species and genetic diversity. *PLoS ONE*, *7*, e42691. https://doi.org/10.1371/journal.pone.0042691

Schnable, J. C., Springer, N. M., & Freeling, M. (2011). Differentiation of the maize subgenomes by genome dominance and both ancient and ongoing gene loss. *Proceedings of the National Academy of Sciences*, *108*, 4069–4074. https://doi.org/10.1073/pnas.1101368108

Schoenfelder, K. P., & Fox, D. T. (2015). The expanding implications of polyploidy. *Journal of Cell Biology*, *209*, 485–491. https://doi.org/10.1083/jcb.201502016

Schwarz, O. (1949). Beiträge zur Kenntnis kritischer Formenkreise im Gebiet der Flora von Thüringen. *Mitteilungen der Thüringischen Botanischen Gesellschaft*, *1*, 120–143.

Shan, S., Mavrodiev, E. V., Li, R., Zhang, Z., Hauser, B. A., Soltis, P. S., … Yang, B. (2018). Application of *CRISPR/Cas9* to *Tragopogon* (Asteraceae), an evolutionary model for the study of polyploidy. *Molecular Ecology Resources*, *18*, 1427–1443. doi: https://doi.org/10.1111/1755-0998.12935

Shan, S., Soltis, P. S., Soltis, D. E., & Yang, B. (2020). Considerations in adapting *CRISPR/Cas9* in nongenetic model plant systems. *Applications in Plant Sciences*, *8*, 1–17. https://doi.org/10.1002/aps3.11314

Šlenker, M., Kantor, A., Marhold, K., Schmickl, R., Mandáková, T., Lysak, M. A., … Zozomová-Lihová, J. (2021). Allele sorting as a novel approach to resolving the origin of allotetraploids using Hyb-Seq data: A case study of the Balkan Mountain endemic *Cardamine barbaraeoides*. *Frontiers in Plant Science*, *12*, 1–22. https://doi.org/10.3389/fpls.2021.659275

Sochor, M., Vašut, R. J., Sharbel, T. F., & Trávníček, B. (2015). How just a few makes a lot: Speciation via reticulation and apomixis on example of European brambles (*Rubus* subgen. *Rubus*, Rosaceae). *Molecular Phylogenetics and Evolution*, *89*, 13–27. https://doi.org/10.1016/j.ympev.2015.04.007

Solís-Lemus, C., Bastide, P., & Ané, C. (2017). PhyloNetworks: A package for phylogenetic networks. *Molecular Biology and Evolution*, *34*, 3292–3298. https://doi.org/10.1093/molbev/msx235

Soltis, D. E., Mavrodiev, E. V., Doyle, J. J., Rauscher, J., & Soltis, P. S. (2008). ITS and ETS sequence data and phylogeny reconstruction in allopolyploids and hybrids. *Systematic Botany*, *33*, 7–20. https://doi.org/10.1600/036364408783887401

Soltis, P. S., & Soltis, D. E. (2000). The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences*, *97*, 7051–7057. https://doi.org/10.1073/pnas.97.13.7051

Soltis, P. S., Marchant, D. B., Van de Peer, Y., & Soltis, D. E. (2015). Polyploidy and genome evolution in plants. *Current Opinion in Genetics and Development*, *35*, 119–125. https://doi.org/10.1016/j.gde.2015.11.003

Soltis, P. S., & Soltis, D. E. (2009). The role of hybridization in plant speciation. *Annual Review of Plant Biology*, *60*, 561–588. https://doi.org/10.1146/annurev.arplant.043008.092039

Soó, R. (1964). Die *Ranunculus auricomus* L. emend. Korsh. Artengruppe in der Flora Ungarns und der Karpaten I. *Acta Botanica Academiae Scientiarum Hungaricae*, *10*, 221–237.

Soó, R. (1965). Die *Ranunculus auricomus* L. emend. Korsh. Artengruppe in der Flora Ungarns und der Karpaten II. *Acta Botanica Academiae Scientiarum Hungaricae*, *11*, 395–404.

Spoelhof, J. P., Soltis, P. S., & Soltis, D. E. (2017). Pure polyploidy: Closing the gaps in autopolyploid research. *Journal of Systematics and Evolution*, *55*, 340–352. https://doi.org/10.1111/jse.12253

Stebbins, G. L. (1950). *Variation and evolution in plants*. New York, USA: Columbia University Press.

Stebbins, G. L. (1984). Polyploidy and the distribution of the arctic-alpine flora: New evidence and a new approach. *Botanica Helvetica*, *94*, 1–13.

Stephens, J. D., Rogers, W. L., Heyduk, K., Cruse-Sanders, J. M., Determann, R. O., Glenn, T. C., & Malmberg, R. L. (2015). Resolving phylogenetic relationships for the recently radiated carnivorous plant genus *Sarracenia* using target enrichment. *Molecular Phylogenetics and Evolution*, *85*, 76–87.

Stewart, J. R., Lister, A. M., Barnes, I., & Dalén, L. (2010). Refugia revisited: individualistic responses of species in space and time. *Proceedings of the Royal Society B*: *Biological Sciences*, *277*, 661–671. https://doi.org/10.1098/rspb.2009.1272

Štorchová, H., Chrtek, J., Bartish, I. V., Tetera, M., Kirschner, J., & Štěpánek, J. (2002). Genetic variation in agamospermous taxa of *Hieracium* sect. *Alpina* (Compositae) in the Tatry Mts. (Slovakia). *Plant Systematics and Evolution*, *235*, 1–17. https://doi.org/10.1007/s00606-002-0219-8

Strasburger, E. (1910). Chromosomenzahl. *Flora Oder Allgemeine Botanische Zeitung*, *100*, 398–446. https://doi.org/10.1016/S0367-1615(17)32797-0

Stuessy, T. F. (2009). *Plant taxonomy: The systematic evaluation of comparative data* (sec ed). New York, USA: Columbia University Press.

Stull, G. W., Soltis, P. S., Soltis, D. E., Gitzendanner, M. A., & Smith, S. A. (2020). Nuclear phylogenomic analyses of asterids conflict with plastome trees and support novel relationships among major lineages. *American Journal of Botany*, *107*, 790–805. https://doi.org/10.1002/ajb2.1468

Sukumaran, J., Holder, M. T., & Knowles, L. L. (2021). Incorporating the speciation process into species delimitation. *PLOS Computational Biology*, *17*, e1008924. https://doi.org/10.1371/journal.pcbi.1008924

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Acadamy of Science*, *114*, 1607–1612. https://doi.org/10.1073/pnas.1607921114

Taberlet, P., Fumagalli, L., Wust-Saucy, A. G., & Cosson, J.-F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, *7*, 453–464. https://doi.org/10.1046/j.1365-294x.1998.00289.x

Te Beest, M., Le Roux, J. J., Richardson, D. M., Brysting, A. K., Suda, J., Kubesova, M., Pysek, P. (2012). The more the better? The role of polyploidy in facilitating plant invasions. *Annnals of Botany*, *109*, 19–45.

Tenore, M. (1835–1836). Ad florae Neapolitanae syllogem, appendix quarta. In *Flora Napolitana ossia descrizione delle piante indigene del Regno di Napoli* (pp. III–XVI, vol. 5). Napoli [Naples]: dalla Stamperia e Cartiera del Fibreno. http://www.ortobotaniconapoli.it/paginadimenu.htm

Than, C., Ruths, D., & Nakhleh, L. (2008). PhyloNet: A software package for analyzing and reconstructing reticulate evolutionary relationships. *BMC Bioinformatics*, *9*, 322. https://doi.org/10.1186/1471-2105-9-322

Tiley, G. P., Crowl, A. A., Manos, P. S., Sessa, E. B., Solis-Lemus, C., Yoder, A. D., & Burleigh, J. G. (2021). Phasing alleles improves network inference with allopolyploids. *BioRxiv*, 1–41. https://doi.org/10.1101/2021.05.04.442457

Tilquin, A., & Kokko, H. (2016). What does the geography of parthenogenesis teach us about sex? *Philosophical Transactions of the Royal Society B*: *Biological Sciences*, *371*, 20150538. https://doi.org/10.1098/rstb.2015.0538

Tomasello, S., Karbstein, K., Hodač, L., Pätzold, C., & Hörandl, E. (2020). Phylogenomics unravels speciation patterns in temperate-montane plant species: A case study on the recently radiating *Ranunculus auricomus* species complex. *Molecular Ecology*, *29*, 2031–2049. https://doi.org/10.1111/mec.15458

Trávníček, P., Kirschner, J., Chudáčková, H., Rooks, F., & Štěpánek, J. (2013). Substantial genome size variation in *Taraxacum stenocephalum* (Asteraceae, Lactuceae). *Folia Geobotanica*, *48*, 271–284. https://doi.org/10.1007/s12224-013-9151-7

Triest, L., De Greef, B., Vermeersch, S., Van Slycken, J., & Coart, E. (1999). Genetic variation and putative hybridization in *Salix alba* and *S. fragilis* (Salicaceae): Evidence

from allozyme data. *Plant Systematics and Evolution*, *215*, 169–187. https://doi.org/10.1007/BF00984654

Ulum, F. B., Costa Castro, C., & Hörandl, E. (2020). Ploidy-dependent effects of light stress on the mode of reproduction in the *Ranunculus auricomus* complex (Ranunculaceae). *Frontiers in Plant Science*, *11*, 104. https://doi.org/10.3389/fpls.2020.00104

Van de Peer, Y., Ashman, T.-L., Soltis, P. S., & Soltis, D. E. (2020). Polyploidy: an evolutionary and ecological force in stressful times. *The Plant Cell*, *0*, 1–16. https://doi.org/10.1093/plcell/koaa015

Van De Peer, Y., Mizrachi, E., & Marchal, K. (2017). The evolutionary significance of polyploidy. *Nature Reviews Genetics*, *18*, 411–424. https://doi.org/10.1038/nrg.2017.26

Vandel, A. (1928). La parthénogenese geographique. Contribution a l'étude biologique et cytologique de la parthénogenese naturelle. *Bulletin Biologique de la France et de la Belgique*, *62*, 164–182.

Vašková, D., & Kolarčik, V. (2019). Breeding systems in diploid and polyploid hawthorns (*Crataegus*): Evidence from experimental pollinations of *C. monogyna*, *C. subsphaerica*, and natural hybrids. *Forests*, *10*, 1–18. https://doi.org/10.3390/F10121059

Vatanparast, M., Powell, A., Doyle, J. J., & Egan, A. N. (2018). Targeting legume loci: A comparison of three methods for target enrichment bait design in Leguminosae phylogenomics. *Applications in Plant Science*, *6*, e1036. https://doi.org/10.1002/aps3.1036

Wagner, F., Härtl, S., Vogt, R., & Oberprieler, C. (2017). "Fix Me Another Marguerite!": Species delimitation in a group of intensively hybridizing lineages of ox-eye daisies (*Leucanthemum* Mill., Compositae-Anthemideae). *Molecular Ecology*, *26*, 4260–4283. https://doi.org/10.1111/mec.14180

Wagner, F., Ott, T., Schall, M., Lautenschlager, U., Vogt, R., & Oberprieler, C. (2020). Taming the red bastards: Hybridisation and species delimitation in the *Rhodanthemum arundanum*-group (Compositae, Anthemideae). *Molecular Phylogenetics and Evolution*, *144*, 106702. https://doi.org/10.1016/j.ympev.2019.106702

Wagner, F., Ott, T., Zimmer, C., Reichhart, V., Vogt, R., & Oberprieler, C. (2019). 'At the crossroads towards polyploidy': Genomic divergence and extent of homoploid hybridization are drivers for the formation of the ox-eye daisy polyploid complex (*Leucanthemum*, Compositae-Anthemideae). *New Phytologist*, *223*, 2039–2053. https://doi.org/10.1111/nph.15784

Wagner, N. D., Gramlich, S., & Hörandl, E. (2018). RAD sequencing resolved phylogenetic relationships in European shrub willows (*Salix* L. subg. *Chamaetia* and subg. *Vetrix*) and revealed multiple evolution of dwarf shrubs. *Ecology and Evolution*, *17*, 8243–8255. doi: 10.1002/ece3.4360

Wagner, N. D., He, L., & Hörandl, E. (2020). Phylogenomic relationships and evolution of polyploid *Salix* species revealed by RAD sequencing data. *Frontiers in Plant Science*, *11*, 36–41. https://doi.org/10.3389/fpls.2020.01077

Weber, H. E. (1996). Former and modern taxonomic treatment of the apomictic *Rubus* complex. *Folia Geobotanica*, *31*, 373–380. https://doi.org/10.1007/BF02815381

Welch, D. M., & Meselson, M. (2000). Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. *Science*, *288*, 1211–1215. https://doi.org/10.1126/science.288.5469.1211

Wen, D., Yu, Y., Zhu, J., & Nakhleh, L. (2018). Inferring phylogenetic networks using PhyloNet. *Systematic Biology*, *67*, 735–740. https://doi.org/10.1093/sysbio/syy015

Wendel, J. F. (2015). The wondrous cycles of polyploidy in plants. *American Journal of Botany*, *102*, 1753–1756. https://doi.org/10.3732/ajb.1500320

Wolf, Y. I., Novichkov, P. S., Karev, G. P., Koonin, E. V., & Lipman, D. J. (2009). The universal distribution of evolutionary rates of genes and distinct characteristics of eukaryotic genes of different apparent ages. *Proceedings of the National Academy of Sciences*, *106*, 7273–7280. https://doi.org/10.1073/pnas.0901808106

Wolfe, K. H. (2001). Yesterday's polyploids and the mystery of diploidization. *Nature Reviews Genetics*, *2*, 333–341. https://doi.org/10.1038/35072009

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., Rieseberg, L. H. (2009). The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Science USA*, *106*, 13875–13879. https://doi.org/10.1073/pnas.0811575106

Yang, Z., & Rannala, B. (2014). Unguided species delimitation using DNA sequence data from multiple loci. *Molecular Biology and Evolution*, *31*, 3125–3135. https://doi.org/10.1093/molbev/msu279

Żabicka, J., Migdałek, G., Słomka, A., Sliwinska, E., Mackiewicz, L., Keczyński, A., & Kuta, E. (2020). Interspecific hybridization and introgression influence biodiversity - based on genetic diversity of Central European *Viola epipsila*-V. palustris complex. *Diversity*, *12*, 321. https://doi.org/10.3390/d12090321

Záveská Drábková, L., Kirschner, J., Štěpánek, J., Záveský, L., & Vlček, Č. (2009). Analysis of nrDNA polymorphism in closely related diploid sexual, tetraploid sexual and polyploid agamospermous species. *Plant Systematics and Evolution*, *278*, 67–85. https://doi.org/10.1007/s00606-008-0134-8

Zhang, Y., Wu, H., Hörandl, E., de Oliveira Franca, R., Wang, L., & Hao, J. (2021). Autonomous apomixis in *Praxelis clematidea* (Asteraceae: Eupatorieae), an invasive alien plant. *AoB PLANTS*, *13*, 1–12. https://doi.org/10.1093/aobpla/plab007

Zhao, M., Zhang, B., Lisch, D., & Ma, J. (2017). Patterns and consequences of subgenome differentiation provide insights into the nature of paleopolyploidy in plants. *The Plant Cell*, *29*, 2974–2994. https://doi.org/10.1105/tpc.17.00595

Zinn, K. E., Tunc-Ozdemir, M., & Harper, J. F. (2010). Temperature stress and plant sexual reproduction: Uncovering the weakest links. *Journal of Experimental Botany*, *61*, 1959–1968. https://doi.org/10.1093/jxb/erq053

Chapter 1

Ali, I. B. E. H., Guetat, A., & Boussaid, M. (2012). Genetic diversity and structure of wild Tunisian *Thymus capitatus* (L.) Hoffm. et Link. (Lamiaceae) assessed using isozyme markers. *African Journal of Ecology*, 50, 140–151. https://doi.org/10.1111/j.1365-2028.2011.01304.x

Alix, K., Gérard, P. R., Schwarzacher, T., & Heslop-Harrison, J. S. P. (2017). Polyploidy and interspecific hybridization: Partners for adaptation, speciation and evolution in plants. *Annals of Botany*, *120*, 183–194. https://doi.org/10.1093/aob/mcx079

Andermann, T., Fernandes, A. M., Olsson, U., Töpel, M., Pfeil, B., Oxelman, B., ... Antonelli, A. (2019). Allele phasing greatly improves the phylogenetic utility of ultraconserved elements. *Systematic Biology*, *68*, 32–46. https://doi.org/10.1093/sysbio/syy039

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. Cambridge, UK: Babraham Bioinformatics, Babraham Institute. Retrieved from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., ... Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, *3*, e3376. https://doi.org/10.1371/journal.pone.0003376

Barke, B. H., Daubert, M., & Hörandl, E. (2018). Establishment of apomixis in diploid F_2 hybrids and inheritance of apospory from F_1 to F_2 hybrids of the *Ranunculus auricomus* complex. *Frontiers of Plant Science*, *9*, 1111. https://doi.org/10.3389/fpls.2018.01111

Barke, B. H., Karbstein, K., Daubert, M., & Hörandl, E. (2020). The relation of meiotic behaviour to hybridity, polyploidy and apomixis in the *Ranunculus auricomus* complex (Ranunculaceae). *BMC Plant Biology*, *20*, 523. https://doi.org/10.1186/s12870-020-02654-3

Bookstein, F. L. (1997a). Landmark methods for forms without landmarks: Morphometrics of group differences in outline shape. *Medical Image Analysis*, *1*, 225–243. https://doi.org/10.1016/S1361-8415(97)85012-8

Bookstein, F. L. (1997b). *Morphometric tools for landmark data: Geometry and biology*. Cambridge, U.K.: Cambridge University Press.

Borchers-Kolb, E. (1983). *Ranunculus* sect. *Auricomus* in Bayern und den angrenzenden Gebieten – I. Allgemeiner Teil. *Mitteilungen der Botanischen Staatssammlung München*, *19*, 363–429.

Borchers-Kolb, E. (1985). *Ranunculus* sect. *Auricomus* in Bayern und den angrenzenden Gebieten II. Spezieller Teil. *Mitteilungen der Botanischen Staatssammlung München*, *21*, 49–300.

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., ... Drummond,

A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*, e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Burgess, M. B., Cushman, K. R., Doucette, E. T., Frye, C. T., & Campbell, C. S. (2015). Understanding diploid diversity: A first step in unraveling polyploid, apomictic complexity in amelanchier. *American Journal of Botany*, *102*, 2041–2057. https://doi.org/10.3732/ajb.1500330

Cavender-Bares, J., González-Rodríguez, A., Eaton, D. A. R., Hipp, A. A. L., Beulke, A., & Manos, P. S. (2015). Phylogeny and biogeography of the American live oaks (*Quercus* subsection *Virentes*): A genomic and population genetics approach. *Molecular Ecology*, *24*, 3668–3687. https://doi.org/10.1111/mec.13269

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods*, *9*, 772. https://doi.org/10.1038/nmeth.2109

De Queiroz, K. (2007). Species concepts and species delimitation. *Systematic Biology*, *56*, 879–886. https://doi.org/10.1080/10635150701701083

Dobeš, C., Mitchell-Olds, T., & Koch, M. A. (2004). Intraspecific diversification in North American *Boechera stricta* (= *Arabis drummondii*), *Boechera x divaricarpa*, and *Boechera holboellii* (Brassicaceae) inferred from nuclear and chloroplast molecular markers – an integrative approach. *American Journal of Botany*, *91*, 2087–2101. https://doi.org/10.3732/ajb.91.12.2087

Doležel, J., Greilhuber, J., & Suda, J. (2007). Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols*, *2*, 2233–2244. https://doi.org/10.1038/nprot.2007.310

Dunkel, F. G. (2010). The *Ranunculus auricomus* L. complex (Ranunculaceae) in Northern Italy. *Webbia*, *65*, 179–227. https://doi.org/10.1080/00837792.2010.10670873

Dunkel, F. G. (2011). The *Ranunculus auricomus* L. complex (Ranunculaceae) in Central and Southern Italy, with additions to North Italian taxa. *Webbia*, *66*, 165–193. https://doi.org/10.1080/00837792.2011.10670895

Dunkel, F. G. (2014). Le complexe de *Ranunculus auricomus* (Ranunculaceae) en Alsace. *Le Journal de Botanique de la Société de Botanique de France*, *66*, 3–53.

Dunkel, F. G. (2015). *Ranunculus pindicola* sp. nov., the only species of the *R. auricomus* complex (Ranunculaceae) in Greece. *Willdenowia*, *45*, 22–230. https://doi.org/10.3372/wi.45.45208

Dunkel, F. G., Gregor, T., & Paule, J. (2018). New diploid species in the *Ranunculus auricomus* complex (Ranunculaceae) from W and SE Europe. *Willdenowia*, *48*, 227–257. https://doi.org/10.3372/wi.48.48205

Earl, D. A., & Holdt, B. M. von (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method.

Conservation Genetics Resources, *4*, 359–361. https://doi.org/10.1007/s12686-011-9548-7

Eaton, D. A. R. (2014). PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*, *30*, 1844–1849. https://doi.org/10.1093/bioinformatics/btu121

Eaton, D. A. R., Spriggs, E. L., Park, B., & Donoghue, M. J. (2017). Misconceptions on missing data in RADseq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology*, *66*, 399–412. https://doi.org/10.1093/sysbio/syw092

Eaton, D. A. R., & Overcast, I. (2020). IPYRAD: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, *36*, 2592–2594. https://doi.org/10.1093/bioinformatics/btz966

Emadzade, K., Gehrke, B., Linder, H. P., & Hörandl, E. (2011). The biogeographical history of the cosmopolitan genus *Ranunculus* L. (Ranunculaceae) in the temperate to meridional zones. *Molecular Phylogenetics and Evolution*, *58*, 4–21. https://doi.org/10.1016/j.ympev.2010.11.002

Emadzade, K., Lebmann, M. J., Hoffmann, M. H., Tkach, N., Lone, F. A., & Hörandl, E. 2015. Phylogenetic relationships and evolution of high mountain buttercups (*Ranunculus*) in North America and Central Asia. *Perspectives in Plant Ecology, Evolution and Systematics*, *17*, 131–141. https://doi.org/10.1016/j.ppees.2015.02.001

Ericsson, S. (1992). The microspecies of the *Ranunculus auricomus* complex treated at the species level. *Annales Botanici Fennici*, *29*, 123–158.

Ericsson, S. (2001). Microspecies within the *Ranunculus auricomus* complex. In Jonsell, B. (ed.), *Flora Nordica* (pp. 382-397, vol. 2.). Stockholm: Bergius Foundation.

Eriksson, J. S., de Sousa, F., Bertrand, Y. J. K., Antonelli, A., Oxelman, B., & Pfeil, B. E. (2018). Allele phasing is critical to revealing a shared allopolyploid origin of *Medicago arborea* and *M. strasseri* (Fabaceae). *BMC Evolutionary Biology*, *18*, 9. https://doi.org/10.1186/s12862-018-1127-z

Feliner, G. N. (2011). Southern European glacial refugia: A tale of tales. *Taxon*, *60*, 365–372. https://doi.org/10.1002/tax.602007

Fér, T., & Schmickl, R. E. (2018). HybPhyloMaker: Target enrichment data analysis from raw reads to species trees. *Evolutionary Bioinformatics*, *14*, 1–9. https://doi.org/10.1177/1176934317742613

Folk, R. A., Mandel, J. R., & Freudenstein, J. V. (2015). A protocol for targeted enrichment of intron-containing sequence markers for recent radiations: A phylogenomic example from *Heuchera* (Saxifragaceae). *Applications in Plant Sciences*, *3*, 1500039. https://doi.org/10.3732/apps.1500039

Freeland, J. R., Kirk, H., & Petersen, S. D. (2011). *Molecular ecology* (sec ed). Chichester, England: Wiley. https://doi.org/10.1002/9780470979365

Freudenstein, J. V., Broe, M. B., Folk, R. A., & Sinn, B. T. (2017). Biodiversity and the species concept – Lineages are not enough. *Systematic Biology*, *66*, 644–656. https://doi.org/10.1093/sysbio/syw098

GBIF Secretariat (2017). *Ranunculus auricomus* L. GBIF Backbone Taxonomy. Checklist dataset https://doi.org/10.15468/39omei accessed via GBIF.org on 21 Nov 2018.

Grant, V. (1981). *Plant speciation*. New York: Columbia University Press.

Gratani, L. (2014). Plant phenotypic plasticity in response to environmental factors. *Advances in Botany*, *2014*, 208747. https://doi.org/10.1155/2014/208747

Grau, J. (1984). Vorläufige Übersicht der iberischen Vertreter von *Ranunculus* sect. *Auricomus*. *Mitteilungen der Botanischen Staatssammlung München*, *20*, 11–28.

Guillardín-Calvo, L., Mora-Márquez, F., Soto, Á., & López de Heredia, U. (2019). RADdesigner: A workflow to select the optimal sequencing methodology in genotyping experiments on woody plant species. *Tree Genetics & Genomes*, *15*, 64. https://doi.org/10.1007/s11295-019-1372-3

Gunz, P., & Mitteroecker, P. (2013). Semilandmarks: A method for quantifying curves and surfaces. *Hystrix*, *24*, 103–109. https://doi.org/10.4404/hystrix-24.1-6292

Häfliger, E. (1943). Zytologisch-embryologische Untersuchungen pseudogamer Ranunkeln der *Auricomus*-Gruppe. *Berichte der Schweizerischen Botanischen Gesellschaft*, *53*, 317–382.

Hammer, Ø., Harper, D. A. T., & Ryan, P. D. (2001). PAST: Paleontological statistics software package for education and data analysis. *Palaeontologia Electronica*, *4*, 4. https://palaeo-electronica.org/2001_1/past/past.pdf

Hassanpour, H., Zare-Maivan, H., Sonboli, A., Kazempour-Osaloo, S., Wagner, F., Tomasello, S., & Oberprieler, C. (2018). Phylogenetic species delimitation unravels a new species in the genus *Sclerorhachis* (Rech.f.) Rech.f. (Compositae, Anthemideae). *Plant Systematics and Evolution*, *304*, 185–203. https://doi.org/10.1007/s00606-017-1461-4

He, L., Wagner, N. D., & Hörandl, E. (2020). RAD sequencing data reveal a radiation of willow species (*Salix* L., Salicaceae) in the Hengduan Mountains and adjacent areas. *Journal of Systematics and Evolution*, *59*, 44-57. https://doi.org/10.1111/jse.12593

Heled, J., & Drummond, A. J. (2010). Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution*, *27*, 570–580. https://doi.org/10.1093/molbev/msp274

Hewitt, G. M. (2011). Quaternary phylogeography: The roots of hybrid zones. *Genetica*, *139*, 617–638. https://doi.org/10.1007/s10709-011-9547-3

Heywood, J. (1991). Spatial analysis of genetic variation in plant populations. *Annual Reviews of Ecology and Systematics*, *22*, 335–355.

https://doi.org/10.1146/annurev.ecolsys.22.1.335

Hipp, A. L., Eaton, D. A. R., Cavender-Bares, J., Fitzek, E., Nipper, R., & Manos, P. S. (2014). A framework phylogeny of the American oak clade based on sequenced RAD data. *PLoS ONE*, *9*, e102272. https://doi.org/10.1371/journal.pone.0093975

Hodač, L., Barke, B. H., & Hörandl, E. (2018). Mendelian segregation of leaf phenotypes in experimental F2 hybrids elucidates origin of morphological diversity of the apomictic *Ranunculus auricomus* complex. *Taxon*, *67*, 1082–1092. https://doi.org/10.12705/676.6

Hodač, L., Scheben, A. P., Hojsgaard, D., Paun, O., & Hörand, E. (2014). ITS polymorphisms shed light on hybrid evolution in apomictic plants: A case study on the *Ranunculus auricomus* complex. *PLoS ONE*, *9*, 28–30. https://doi.org/10.1371/journal.pone.0103003

Hojsgaard, D., Greilhuber, J., Pellino, M., Paun, O., Sharbel, T. F., & Hörandl, E. (2014). Emergence of apospory and bypass of meiosis via apomixis after sexual hybridisation and polyploidisation. *New Phytologist*, *204*, 1000–1012. https://doi.org/10.1111/nph.12954

Holliday, J. A., Zhou, L., Bawa, R., Zhang, M., & Oubida, R. W. (2016). Evidence for extensive parallelism but divergent genomic architecture of adaptation along altitudinal and latitudinal gradients in *Populus trichocarpa*. *New Phytologist*, *209*, 1240–1251. https://doi.org/10.1111/nph.13643

Hörandl, E. 1998. Species concepts in agamic complexes: Applications in the *Ranunculus auricomus* complex and general perspectives. *Folia Geobotanica*, *33*, 335–348. https://doi.org/10.1007/BF03216210

Hörandl, E. (2002). Morphological differentiation within the *Ranunculus cassubicus* group compared to variation of isozymes, ploidy levels, and reproductive systems: Implications for taxonomy. *Plant Systematics and Evolution*, 233, 65–78. https://doi.org/10.1007/s00606-002-0210-4

Hörandl, E. (2004). Comparative analysis of genetic divergence among sexual ancestors of apomictic complexes using isozyme data. *International Journal of Plant Sciences*, *165*, 615–622. https://doi.org/10.1086/386557

Hörandl, E. (2008). Evolutionary implications of self-compatibility and reproductive fitness in the apomictic *Ranunculus auricomus* polyploid complex (Ranunculaceae). *International Journal of Plant Sciences*, *169*, 1219–1228. https://doi.org/10.1086/591980

Hörandl, E. (2018). The classification of asexual organisms: Old myths, new facts, and a novel pluralistic approach. *Taxon*, *67*, 1066–1081. https://doi.org/10.12705/676.5

Hörandl, E., & Emadzade, K. (2012). Evolutionary classification: A case study on the diverse plant genus *Ranunculus* L. (Ranunculaceae). *Perspectives in Plant Ecology, Evolution and Systematics*, *14*, 310–324. https://doi.org/10.1016/j.ppees. 2012.04.001

Hörandl, E., & Greilhuber, J. (2002). Diploid and autotetraploid sexuals and their

relationships to apomicts in the *Ranunculus cassubicus* group: Insights from DNA content and isozyme variation. *Plant Systematics and Evolution*, *234*, 85–100. https://doi.org/10.1007/s00606-002-0209-x

Hörandl, E., & Gutermann, W. (1995). Die Bearbeitung der *Ranunculus auricomus*-Gruppe für die 'Flora von Österreich' – ein Werkstattbericht. *Flora Austriacae Novitates*, *2*, 12–27. https://doi.org/10.1016/j.jmapro.2018.05.023

Hörandl, E., & Gutermann, W. (1998a). Der *Ranunculus auricomus*-Komplex in Österreich. 1. Methodik. Gruppierung der mitteleuropäischen Sippen. *Botanische Jahrbücher fur Systematik, Pflanzengeschichte und Pflanzengeographie*, *120*, 1–44.

Hörandl, E., & Gutermann, W. (1998b). Der *Ranunculus auricomus*-Komplex in Österreich 2. Die *R. cassubicus*-, *R. monophyllus*- und *R. fallax*-Sammelgruppe. *Botanische Jahrbücher fur Systematik, Pflanzengeschichte und Pflanzengeographie*, *120*, 545–598.

Hörandl, E., & Gutermann, W. (1998c). Zur Kenntnis des *Ranunculus auricomus*-Komplexes in Österreich: Die Arten der *R. phragmiteti*- und *R. indecorus*-Gruppe. *Phyton (Horn)*, *37*, 263–320.

Hörandl, E., & Gutermann, W. (1999). Der *Ranunculus auricomus*-Komplex in Österreich und benachbarten Gebieten 3. Die Arten der *R. latisectus*-, *R. puberulus*-, *R. stricticaulis*- und *R. argoviensis* (*R. auricomus*-Sammelgruppe). *Botanische Jahrbücher fur Systematik, Pflanzengeschichte und Pflanzengeographie*, *121*, 99–138.

Hörandl, E., Dobeš, C., & Lambrou, M. (1997). Chromosomen- und Pollenuntersuchungen an österreichischen Arten des apomiktischen *Ranunculus auricomus*-Komplexes. *Botanica Helvetica*, *107*, 195–209.

Hörandl, E., Paun, O., Johansson, J. T., Lehnebach, C., Armstrong, T., Chen, L., & Lockhart, P. (2005). Phylogenetic relationships and evolutionary traits in *Ranunculus* s.l. (Ranunculaceae) inferred from ITS sequence analysis. *Molecular Phylogenetics and Evolution*, *36*, 305–327. https://doi.org/10.1016/j.ympev.2005.02.009

Hörandl, E., Greilhuber, J., Klímová, K., Paun, O., Temsch, E., Emadzade, K., & Hodálová, I. (2009). Reticulate evolution and taxonomic concepts in the *Ranunculus auricomus* complex (Ranunculaceae): Insights from analysis of morphological, karyological and molecular data. *Taxon*, *58*, 1194–1215. https://doi.org/10.1002/tax.584012

Hu, H., Al-Shehbaz, I. A., Sun, Y., Hao, G., Wang, Q., & Liu, J. (2015). Species delimitation in *Orychophragmus* (Brassicaceae) based on chloroplast and nuclear DNA barcodes. *Taxon*, *64*, 714–726. https://doi.org/10.12705/644.4

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, 23, 254–267. https://doi.org/10.1093/molbev/msj030

Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation

program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, *23*, 1801–1806. https://doi.org/10.1093/bioinformatics/btm233

Jalas, J., & Suominen, J. (1989). *Atlas florae Europaeae*, vol. 8, *Nymphaeaceae to Ranunculaceae*. Helsinki: The Committee for Mapping the Flora of Europe and Societas Biologica Fennica Vanamo.

Jensen, R. J., Ciofani, K. M., & Miramontes, L. C. (2002). Lines, outlines, and landmarks: Morphometric analyses of leaves of *Acer rubrum*, *Acer saccharinum* (Aceraceae) and their hybrid. *Taxon*, *51*, 475–492. https://doi.org/10.2307/1554860

Jiang, X., Hipp, A.L., Deng, M., Su, T., Zhou, Z., & Yan, M. (2019). East Asian origins of European holly oaks (*Quercus* section *Ilex* Loudon) via the Tibet-Himalaya. *Journal of Biogeography*, *46*, 2203–2214. https://doi.org/10.1111/jbi.13654

Jiao, Y., Wickett, N. J., Ayyampalayam, S., Chanderbali, A. S., Landherr, L., Ralph, P. E., ... Depamphilis, C. W. (2011). Ancestral polyploidy in seed plants and angiosperms. *Nature*, *473*, 97–100. https://doi.org/10.1038/nature09916

Jones, G. (2017). Algorithmic improvements to species delimitation and phylogeny estimation under the multispecies coalescent. *Journal of Mathematcial Biology*, *74*, 447–467. https://doi.org/10.1007/s00285-016-1034-0

Jones, G., Aydin, Z., & Oxelman, B. (2015). DISSECT: An assignment-free Bayesian discovery method for species delimitation under the multispecies coalescent. *Bioinformatics*, *31*, 991–998. https://doi.org/10.1093/bioinformatics/btu770

Julin, E. (1980). *Ranunculus auricomus* L. in Södermanland, East-Central Sweden. *Opera Botanica*, *57*, 1–145.

Junier, T., & Zdobnov, E. M. (2010). The Newick utilities: High-throughput phylogenetic tree processing in the UNIX shell. *Bioinformatics*, *26*, 1669–1670. https://doi.org/10.1093/bioinformatics/btq243

Kilian, N., Henning, T., Plitzner, P., Müller, A., Güntsch, A., Stöver, B. C., ... Borsch, T. (2015). Sample data processing in an additive and reproducible taxonomic workflow by using character data persistently linked to preserved individual specimens. *Database*, *2015*, bav094. https://doi.org/10.1093/database/bav094

Kirschner, J., Záveská Drábková, L., Štěpánek, J., & Uhlemann, I. (2015). Towards a better understanding of the *Taraxacum* evolution (Compositae–Cichorieae) on the basis of nrDNA of sexually reproducing species. *Plant Systematics and Evolution*, *301*, 1135–1156. https://doi.org/10.1007/s00606-014-1139-0

Klatt, S., Hadacek, F., Hodač, L., Brinkmann, G., Eilerts, M., Hojsgaard, D., & Hörandl, E. (2016). Photoperiod extension enhances sexual megaspore formation and triggers metabolic reprogramming in facultative apomictic *Ranunculus auricomus*. *Frontiers in Plant Science*, *7*, 278. https://doi.org/10.3389/fpls.2016.00278

Klatt, S., Schinkel, C. C. F., Kirchheimer, B., Dullinger, S., & Hörandl, E. (2018). Effects

of cold treatments on fitness and mode of reproduction in the diploid and polyploid alpine plant *Ranunculus kuepferi* (Ranunculaceae). *Annuals of Botany*, *121*, 1287–1298. https://doi.org/10.1093/aob/mcy017

Klingenberg, C. (2015). Analyzing fluctuating asymmetry with geometric morphometrics: Concepts, methods, and applications. *Symmetry*, *7*, 843–934. https://doi.org/10.3390/sym7020843

Klingenberg, C. P. (2011). MorphoJ: An integrated software package for geometric morphometrics. *Molecular Ecology Resources*, *11*, 353–357. https://doi.org/10.1111/j.1755-0998.2010.02924.x

Klingenberg, C. P., Barluenga, M., & Meyer, A. (2002). Shape analysis of symmetric structures: Quantifying variation among individuals and asymmetry. *Evolution*, *56*, 1909–1920. https://doi.org/10.1111/j.0014-3820.2002.tb00117.x

Klingenberg, C. P., Duttke, S., Whelan, S., & Kim, M. (2012). Developmental plasticity, morphological variation and evolvability: A multilevel analysis of morphometric integration in the shape of compound leaves. *Journal of Evolutionary Biology*, *25*, 115–129. https://doi.org/10.1111/j.1420-9101.2011.02410.x

Koch, W. (1939). Zweiter Beitrag zur Kenntnis des Formenkreises von *Ranunculus auricomus* L.: Studien über kritische Schweizerpflanzen III. *Berichte der Schweizerischen Botanischen Gesellschaft*, *49*, 541–554.

Körner, C. (2003). *Alpine plant life: Functional plant ecology of high mountain ecosystems*. Berlin & Heidelberg: Springer.

Košir, P., Čarni, A., & Di Pietro, R. (2008). Classification and phytogeographical differentiation of broad-leaved ravine forests in southeastern Europe. *Journal of Vegetation Science*, *19*, 331–342. https://doi.org/10.3170/2008-8-18372

Kozlov, A. M., Aberer, A. J., & Stamatakis, A. (2015). ExaML version 3: A tool for phylogenomic analyses on supercomputers. *Bioinformatics*, *31*, 2577–2579. https://doi.org/10.1093/bioinformatics/btv184

Krak, K., Caklová, P., Chrtek, J., & Fehrer, J. (2013). Reconstruction of phylogenetic relationships in a highly reticulate group with deep coalescence and recent speciation (*Hieracium*, Asteraceae). *Heredity*, *110*, 138–151. https://doi.org/10.1038/hdy.2012.100

Kubatko, L. S., Carstens, B. C., & Knowles, L. L. (2009). STEM: Species tree estimation using maximum likelihood for gene trees under coalescence. *Bioinformatics*, *25*, 971–973. https://doi.org/10.1093/bioinformatics/btp079

Kvist, G. (1987). Identity of the material of the *Ranunculus auricomus* group in the Linnaean Herbarium. *Annales Botanici Fennici*, *24*, 73–77.

Leliaert, F., Verbruggen, H., Wysor, B., & De Clerck, O. (2009). DNA taxonomy in morphologically plastic taxa: Algorithmic species delimitation in the *Boodlea* complex (Chlorophyta: Cladophorales). *Molecular Phylogenetics and Evolution*, *53*, 122–133.

https://doi.org/10.1016/j.ympev.2009.06.004

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Linnaeus, C. (1753). *Species plantarum*. Holmiae [Stockholm]: impensis Laurentii Salvii. https://doi.org/10.5962/bhl.title.669

Liu, L. (2008). BEST: Bayesian estimation of species trees under the coalescent model. *Bioinformatics*, *24*, 2542–2543. https://doi.org/10.1093/bioinformatics/btn484

MacLeod, N. (2008). Size and shape coordinates. *Newsletter Palaeontological Association*, *69*, 1-10.

Mallet, J. (1995). A species definition for the Modern Synthesis. *Trends in Ecology and Evolution*, *10*, 294–299.

Mandel, J. R., Dikow, R. B., Funk, V. A., Masalia, R. R., Staton, S. E., Kozik, A., ... Burke, J. M. (2014). A target enrichment method for gathering phylogenetic information from hundreds of loci: An example from the Compositae. *Applications in Plant Sciences*, *2*, 1300085. https://doi.org/10.3732/apps.1300085

Marklund, G. (1961). Der *Ranunculus auricomus*-Komplex in Finnland I. Diagnosen und Fundortslisten einiger Sippen des *R. auricomus* L. coll. (s. str.). *Flora Fennica*, *3*, 1–128.

Marklund, G. (1965). Der *Ranunculus auricomus*-Komplex in Finnland II: Diagnosen und Fundortslisten einiger Sippen von *R. fallax* (W. et Gr.) Schur, *R. monophyllus* Ovcz. und *R. cassubicus* L. *Flora Fennica*, *4*, 1–198.

Masci, S., Miho, A., & Marchi, P. (1994). *Ranunculus auricomus* L. aggr. (Ranunculaceae) in Italy. I. Sexual tetraploids on the Apennines. *Caryologia*, *47*, 97–108. https://doi.org/10.1080/00087114.1994.10797287

Matzk, F., Meister, A., & Schubert, I. (2000). An efficient screen for reproductive pathways using mature seeds of monocots and dicots. *The Plant Journal*, *21*, 97–108. https://doi.org/10.1046/j.1365-313X.2000.00647.x

Mayr, E. (1942). *Systematics and the origin of species*. New York: Columbia University Press.

McBreen, K., & Lockhart, P. J. (2006). Reconstructing reticulate evolutionary histories of plants. *Trends in Plant Science*, *11*, 398–404. https://doi.org/10.1016/j.tplants.2006.06.004

McCartney-Melstad, E., Gidiş, M., & Shaffer, H. B. (2019). An empirical pipeline for choosing the optimal clustering threshold in RADseq studies. *Molecular Ecolutionary Resources*, *19*, 1195–1204. https://doi.org/10.1111/1755-0998.13029

Mirarab, S., Reaz, R., Bayzid, M. S., Zimmermann, T., Swenson, M. S., & Warnow, T. (2014). ASTRAL: Genome-scale coalescent-based species tree estimation.

Bioinformatics, *30*, i541–i548. https://doi.org/10.1093/bioinformatics/btu462

Mráz, P., & Ronikier, M. (2016). Biogeography of the Carpathians: Evolutionary and spatial facets of biodiversity. *Biological Journal of the Linnean Society*, *119*, 528–559. https://doi.org/10.1111/bij.12918

Naciri, Y., Du Pasquier, P.-E., Lundberg, M., Jeanmonod, D., & Oxelman, B. (2017). A phylogenetic circumscription of *Silene* sect. *Siphonomorpha* (Caryophyllaceae) in the Mediterranean Basin. *Taxon*, *66*, 91–108. https://doi.org/10.12705/661.5

Nguyen, N., Mirarab, S., & Warnow, T. (2012). MRL and SuperFine+MRL: New supertree methods. *Algorithms for Molecular Biology*, *7*, 3. https://doi.org/10.1186/1748-7188-7-3

Nogler, G. A. (1984). Genetics of apospory in apomictic *Ranunculus auricomus*. V. Conclusion. *Botanica Helvetica*, *94*, 411–422.

Oberprieler, C., Wagner, F., Tomasello, S., & Konowalik, K. (2017). A permutation approach for inferring species networks from gene trees in polyploid complexes by minimising deep coalescences. *Methods in Ecology and Evolution*, *8*, 835–849. https://doi.org/10.1111/2041-210X.12694

Oliver, J. C. (2013). Microevolutionary processes generate phylogenomic discordance at ancient divergences. *Evolution*, *67*, 1823–1830. https://doi.org/10.1111/evo.12047

Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: A road map for STACKS. *Methods in Ecology and Evolution*, *8*, 1360–1373. https://doi.org/10.1111/2041-210X.12775

Pätzold, C., Wood, K. R., Eaton, D., Wagner, W. L., & Appelhans, M. S. (2019). Phylogeny of Hawaiian *Melicope* (Rutaceae): RADseq resolves species relationships and reveals ancient introgression. *Frontiers in Plant Science*, *10*, 1074. https://doi.org/10.3389/fpls.2019.01074

Paun, O., Greilhuber, J., Temsch, E. M., & Hörandl, E. (2006a). Patterns, sources and ecological implications of clonal diversity in apomictic *Ranunculus carpaticola* (*Ranunculus auricomus* complex, Ranunculaceae). *Molecular Ecology*, *15*, 897–910. https://doi.org/10.1111/j.1365-294X.2006.02800.x

Paun, O., Stuessy, T. F., & Hörandl, E. (2006b). The role of hybridization, polyploidization and glaciation in the origin and evolution of the apomictic *Ranunculus cassubicus* complex. *New Phytologist*, *171*, 223–236. https://doi.org/10.1111/j.1469-8137.2006.01738.x

Pease, J. B., Brown, J. W., Walker, J. F., Hinchliff, C. E., & Smith, S. A. (2018). Quartet Sampling distinguishes lack of support from conflicting support in the green plant tree of life. *American Journal of Botany*, *105*, 385–403. https://doi.org/10.1002/ajb2.1016

Pellino, M., Hojsgaard, D., Schmutzer, T., Scholz, U., Hörandl, E., Vogel, H., & Sharbel, T. F. (2013). Asexual genome evolution in the apomictic *Ranunculus auricomus* complex: Examining the effects of hybridization and mutation accumulation. *Molecular*

Ecology, *22*, 5908–5921. https://doi.org/10.1111/mec.12533

Perez, S. I., Bernal, V., & Gonzalez, P. N. (2006). Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. *Journal of Anatomy*, *208*, 769–784. https://doi.org/10.1111/j.1469-7580.2006.00576.x

Pirie, M. D., Humphreys, A. M., Barker, N. P., & Linder, H. P. (2009). Reticulation, data combination, and inferring evolutionary history: An example from Danthonioideae (Poaceae). *Systematic Biology*, *58*, 612–628. https://doi.org/10.1093/sysbio/syp068

Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., ... Vogler, A. P. (2006). Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology*, *55*, 595–609. https://doi.org/10.1080/10635150600852011

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959. https://doi.org/10.1111/j.1471-8286.2007.01758.x

R Core Team (2019). R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. http://www.r-project.org/

Raab-Straube, E., Hörandl, E., & Nardi, E. (2014). Ranunculaceae. http://ww2.bgbm.org/EuroPlusMed

Rahmsdorf, E. (2019). *Kreuzungsexperimente an sexuellen Arten des Ranunculus auricomus Komplexes (Ranunculaceae)*. University of Göttingen, Bachelor Thesis.

Rambaut, A. (2014). Figtree, a graphical viewer of phylogenetic trees, version v.1.4.3. http://tree.bio.ed.ac.uk/software/figtree

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using tracer 1.7. *Systematic Biology*, *67*, 901–904. https://doi.org/10.1093/sysbio/syy032

Rancilhac, L., Goudarzi, F., Gehara, M., Hemami, M. R., Elmer, K. R., Vences, M., & Steinfarz, S. (2019). Phylogeny and species delimitation of near Eastern *Neurergus* newts (Salamandridae) based on genome-wide RADseq data analysis. *Molecular Phylogenetics and Evolution*, *133*, 189–197. https://doi.org/10.1016/j.ympev.2019.01.003

Rannala, B. (2015). The art and science of species delimitation. *Current Zoology*, *61*, 846–853. https://doi.org/10.1093/czoolo/61.5.846

Reichenbach, L. (1832) ("1830–1832"). *Flora Germanica excursoria*. Lipsiae [Leipzig]: apud Carolum Cnobloch. https://doi.org/10.5962/bhl.title.309

Rohlf, F. J. (2015). The tps series of software. *Hystrix*, *26*, 9–12. https://doi.org/10.4404/hystrix-26.1-11264

Rosenberg, N. A. (2004). DISTRUCT: A program for the graphical display of population structure. *Molecular Ecology Notes*, *4*, 137–138. https://doi.org/10.1046/j.1471-8286.2003.00566.x

Rousi, A. (1956). Cytotaxonomy and reproduction in the apomictic *Ranunculus auricomus* group. *Annales Botanici Societatis Zoologicae-Botanicae Fennicae "Vanamo"*, *29*, 1–64.

Rubin, B. E. R., Ree, R. H., & Moreau, C. S. (2012). Inferring phylogenies from RAD sequence data. *PLoS ONE*, *7*, e33394. https://doi.org/10.1371/journal.pone.0033394

Ruiz-Sanchez, E. (2015). Parametric and non-parametric species delimitation methods result in the recognition of two new Neotropical woody bamboo species. *Molecular Phylogenetics and Evolution*, *93*, 261–273. https://doi.org/10.1016/j.ympev.2015.08.004

Sass, C., Iles, W. J. D., Barrett, C. F., Smith, S. Y., & Specht, C. D. (2016). Revisiting the Zingiberales: Using multiplexed exon capture to resolve ancient and recent phylogenetic splits in a charismatic plant lineage. *PeerJ*, *4*, e1584. https://doi.org/10.7717/peerj.1584

Schinkel, C. C. F., Kirchheimer, B., Dellinger, A. S., Klatt, S., Winkler, M., Dullinger, S., & Hörandl, E. (2016). Correlations of polyploidy and apomixis with elevation and associated environmental gradients in an alpine plant. *AoB Plants*, *8*, plw064. https://doi.org/10.1093/aobpla/plw064

Schmickl, R., Liston, A., Zeisek, V., Oberlander, K., Weitemier, K., Straub, S. C. K., ... Suda, J. (2016). Phylogenetic marker development for target enrichment from transcriptome and genome skim data: The pipeline and its application in southern African *Oxalis* (Oxalidaceae). *Molecular Ecology Resources*, *16*, 1124–1135. https://doi.org/10.1111/1755-0998.12487

Seo, T.-K. (2008). Calculating bootstrap probabilities of phylogeny using multilocus sequence data. *Molecular Biology and Evolution*, *25*, 960–971. https://doi.org/10.1093/molbev/msn043

Shen, X.-X., Hittinger, C. T., & Rokas, A. (2017). Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nature Ecology and Evolution*, *1*, 0126. https://doi.org/10.1038/s41559-017-0126

Solís-Lemus, C., Bastide, P., & Ané, C. (2017). PhyloNetworks: A package for phylogenetic networks. *Molecular Biology and Evolution*, *34*, 3292–3298. https://doi.org/10.1093/molbev/msx235

Soltis, P. S., & Soltis, D. E. (2000). The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Acadamy of Science USA*, *97*, 7051–7057. https://doi.org/10.1073/pnas.97.13.7051

Soó, R. (1964). Die *Ranunculus auricomus* L. emend. Korsh. Artengruppe in der Flora Ungarns und der Karpaten I. *Acta Botanica Academiae Scientiarum Hungaricae*, *10*, 221–237.

Soó, R. (1965). Die *Ranunculus auricomus* L. emend. Korsh. Artengruppe in der Flora Ungarns und der Karpaten II. *Acta Botanica Academiae Scientiarum Hungaricae*, *11*, 395–404.

Stace, C. A. (1998). Species recognition in agamosperms – The need for a pragmatic approach. *Folia Geobotanica*, *33*, 319–326. https://doi.org/10.1007/BF03216207

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stephens, J. D., Rogers, W. L., Heyduk, K., Cruse-Sanders, J. M., Determann, R. O., Glenn, T. C., & Malmberg, R. L. (2015). Resolving phylogenetic relationships of the recently radiated carnivorous plant genus *Sarracenia* using target enrichment. *Molecular Phylogenetics and Evolution*, *85*, 76–87. https://doi.org/10.1016/j.ympev.2015.01.015

Stevanović, V., Niketić, M., & Lakušić, D. (1991). Chorological additions to the flora of eastern Yugoslavia. *Flora Mediterranea*, *1*, 121–142.

Stuessy, T. F. (2009). *Plant taxonomy: The systematic evaluation of comparative data* (sec ed). New York, USA: Columbia University Press.

Sukumaran, J., & Knowles, L. L. (2017). Multispecies coalescent delimits structure, not species. *Proceedings of the National Acadamy of Science*, *114*, 1607–1612. https://doi.org/10.1073/pnas.1607921114

Tenore, M. (1835–1836). Ad florae Neapolitanae syllogem, appendix quarta. In *Flora Napolitana ossia descrizione delle piante indigene del Regno di Napoli* (pp. III–XVI, vol. 5). Napoli [Naples]: dalla Stamperia e Cartiera del Fibreno. http://www.ortobotaniconapoli.it/paginadimenu.htm

Than, C., Ruths, D., & Nakhleh, L. (2008). PhyloNet: A software package for analyzing and reconstructing reticulate evolutionary relationships. *BMC Bioinformatics*, *9*, 322. https://doi.org/10.1186/1471-2105-9-322

The GIMP Team (2018). GIMP 2 - GNU Image Manipulating Program. http://www.gimp.org/

Tomasello, S. (2018). How many names for a beloved genus? – Coalescent-based species delimitation in *Xanthium* L. (Ambrosiinae, Asteraceae). *Molecular Phylogenetics and Evolution*, *127*, 135–145. https://doi.org/10.1016/j.ympev.2018.05.024

Tomasello, S., Karbstein, K., Hodač, L., Pätzold, C., & Hörandl, E. (2020). Phylogenomics unravels speciation patterns in temperate-montane plant species: A case study on the recently radiating *Ranunculus auricomus* species complex. *Molecular Ecology*, *29*, 2031–2049. https://doi.org/10.1111/mec.15458

Toprak, Z., Pfeil, B. E., Jones, G., Marcussen, T., Ertekin, A. S., & Oxelman, B. (2016). Species delimitation without prior knowledge: DISSECT reveals extensive cryptic speciation in the *Silene aegyptiaca* complex (Caryophyllaceae). *Molecular Phylogenetics*

and Evolution, *102*, 1–8. https://doi.org/10.1016/j.ympev.2016.05.024

Tripp, E. A., Tsai, Y. H. E., Zhuang, Y., & Dexter, K. G. (2017). RADseq dataset with 90% missing data fully resolves recent radiation of *Petalidium* (Acanthaceae) in the ultra-arid deserts of Namibia. *Ecology and Evolution*, *7*, 7920–7936. https://doi.org/10.1002/ece3.3274

Turland, N. J., Wiersema, J. H., Barrie, F. R., Greuter, W., Hawksworth, D. L., Herendeen, P. S., ... Smith, G. F. (Eds.) (2018). *International Code of Nomenclature for algae, fungi, and plants (Shenzhen Code) adopted by the Nineteenth International Botanical Congress Shenzhen, China, July 2017*. Regnum Vegetabile 159. Glashütten: Koeltz Botanical Books. https://doi.org/10.12705/Code.2018

Tutin, T. G., & Akeroyd, J. R. (1993). *Ranunculus* subgen. *Ranunculus*. In Tutin, T. G., Burges, N. A., Edmondson, J. R., Chater, A. O., Heywood, V. H., Moore, D. M., ... Mill, R. R. (Eds.), *Flora Europaea* (sec ed, pp. 269–287, vol. 1). Cambridge, U.K.: Cambridge University Press.

Vachaspati, P., & Warnow, T. (2015). ASTRID: Accurate species trees from internode distances. *BMC Genomics*, *16*, S3. https://doi.org/10.1186/1471-2164-16-S10-S3

Vatanparast, M., Powell, A., Doyle, J. J., & Egan, A. N. (2018). Targeting legume loci: A comparison of three methods for target enrichment bait design in Leguminosae phylogenomics. *Applications in Plant Science*, *6*, e1036. https://doi.org/10.1002/aps3.1036

Volkova, P. A., Kasatskaya, S. A., Boiko, A. A., & Shipunov, A. B. (2010). Stability of leaf form and size during specimen preparation of herbarium specimens. *Feddes Repertorium*, *121*, 219–225. https://doi.org/10.1002/fedr.201000021

Wagner, F., Härtl, S., Vogt, R., & Oberprieler, C. (2017). "Fix Me Another Marguerite!": Species delimitation in a group of intensively hybridizing lineages of ox-eye daisies (*Leucanthemum* Mill., Compositae-Anthemideae). *Molecular Ecology*, *26*, 4260–4283. https://doi.org/10.1111/mec.14180

Wagner, N. D., Gramlich, S., & Hörandl, E. (2018). RAD sequencing resolved phylogenetic relationships in European shrub willows (*Salix* L. subg. *Chamaetia* and subg. *Vetrix*) and revealed multiple evolution of dwarf shrubs. *Ecology and* Evolution, *17*, 8243–8255. https://doi.org/10.1002/ece3.4360

Wanke, S., Granados Mendoza, C., Müller, S., Paizanni Guillén, A., Neinhuis, C., Lemmon, A. R., ... Samain, M.-S. (2017). Recalcitrant deep and shallow nodes in *Aristolochia* (Aristolochiaceae) illuminated using anchored hybrid enrichment. *Molecular Phylogenetics and Evolution*, *117*, 111–123. https://doi.org/10.1016/j.ympev.2017.05.014

Weisrock, D. W., Smith, S. D., Chan, L.M., Biebouw, K., Kappeler, P. M., & Yoder, A. D. (2012). Concatenation and concordance in the reconstruction of mouse lemur phylogeny: An empirical demonstration of the effect of allele sampling in phylogenetics.

Molecular Biology and Evolution, *29*, 1615–1630. https://doi.org/10.1093/molbev/mss008

Wen, D., & Nakhleh, L. (2018). Coestimating reticulate phylogenies and gene trees from multilocus sequence data. *Systematic Biology*, *67*, 439–457. https://doi.org/10.1093/sysbio/syx085

Yang, Z., & Rannala, B. (2014). Unguided species delimitation using DNA sequence data from multiple loci. *Molecular Biology and Evolution*, *31*, 3125–3135. https://doi.org/10.1093/molbev/msu279

Yu, Y., & Nakhleh, L. (2015). A maximum pseudo-likelihood approach for phylogenetic networks. *BMC Genomics*, *16*, S10. https://doi.org/10.1186/1471-2164-16-S10-S10

Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, *19*, 153. https://doi.org/10.1186/s12859-018-2129-y

Zhang, J., Kapli, P., Pavlidis, P., & Stamatakis, A. (2013). A general species delimitation method with applications to phylogenetic placements. Bioinformatics, *29*, 2869–2876. https://doi.org/10.1093/bioinformatics/btt499

Chapter 2

Alonso-Marcos, H., Hülber, K., Myllynen, T., Pérez Rodríguez, P., & Dobeš, C. (2018). Pollen precedence in sexual *Potentilla puberula* and its role as a protective reproductive barrier against apomictic cytotypes. *Taxon*, *67*, 1132–1142. https://doi.org/10.12705/676.9

Andrews, S. (2010). FastQC: A quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute. Asker, S., & Jerling, L. (1992). *Apomixis in plants*. CRC Press.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., ... Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, *3*, 1–7. https://doi.org/10.1371/journal.pone.0003376

Baker, H. G. (1967). Support for Baker's law – as a rule. *Evolution*, *21*, 853–856. https://doi.org/10.1017/CBO97 81107 415324.004

Barke, B. H., Daubert, M., & Hörandl, E. (2018). Establishment of apomixis in diploid F2 hybrids and inheritance of apospory from F1 to F2 hybrids of the *Ranunculus auricomus* complex. *Frontiers in Plant Science*, *9*, 1–12. https://doi.org/10.3389/fpls.2018.01111

Barrett, S. C. H., & Charlesworth, D. (1991). Effects of a change in the level of inbreeding on the genetic load. *Nature*, *352*, 522–524. https://doi.org/10.1038/352522a0

Bierzychudek, P. (1985). Patterns in plant parthenogenesis. *Experientia*, *41*, 1255–1264. https://doi.org/10.1007/BF019 52068

Brochmann, C., Brysting, A. K., Alsos, I. G., Borgen, L., Grundt, H. H., Scheen, A.-C., & Elven, R. (2004). Polyploidy in arctic plants. *Biological Journal of the Linnean Society*, *82*, 521–536. https://doi.org/10.1111/j.1095-8312.2004.00337.x

Brožová, V., Koutecký, P., & Doležal, J. (2019). Plant apomixis is rare in Himalayan high-alpine flora. *Scientific Reports*, *9*, 14386. https://doi.org/10.1038/s4159 8-019-50907 -5

Burt, A. (2000). Perspective: Sex, recombination, and the efficacy of selection - was Weismann right? *Evolution*, *54*, 337–351. https://doi.org/10.1111/j.0014-3820.2000.tb000 38.x

Charlesworth, D., & Charlesworth, B. (1987). Inbreeding depression and its evolutionary consequences. *Annual Review of Ecology and Systematics*, *18*, 237–268. https://doi.org/10.1146/annurev.es.18.110187.001321

Cosendai, A. C., & Hörandl, E. (2010). Cytotype stability, facultative apomixis and geographical parthenogenesis in *Ranunculus kuepferi* (Ranunculaceae). *Annals of Botany*, *105*, 457–470. https://doi.org/10.1093/aob/mcp304

Cosendai, A. C., Wagner, J., Ladinig, U., Rosche, C., & Hörandl, E. (2013). Geographical parthenogenesis and population genetic structure in the alpine species *Ranunculus kuepferi* (Ranunculaceae). *Heredity*, *110*, 560–569. https://doi.org/10.1038/hdy.2013.1

Crnokrak, P., & Barrett, S. C. H. (2002). Perspective: Purging the genetic load: A review of the experimental evidence. *Evolution*, *56*, 2347–2358. https://doi.org/10.1111/j.0014-3820.2002.tb00160.x

De Nettancourt, D. (2001). Incompatibility and incongruity in wild and cultivated plants (sec ed, vol. 3). Berlin, Germany: Springer.

Dunkel, F. G., Gregor, T., & Paule, J. (2018). New diploid species in the *Ranunculus auricomus* complex (Ranunculaceae) from W and SE Europe. *Willdenowia*, *48*, 227–257. https://doi.org/10.3372/wi.48.48205

Eaton, D. A. R. (2014). PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*, *30*, 1844–1849. https://doi.org/10.1093/bioinformatics/btu121

Eaton, D. A. R., & Overcast, I. (2020). IPYRAD: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, *36*, 2592–2594. https://doi.org/10.1093/bioinformatics/btz966

Freeland, J. R., Kirk, H., & Petersen, S. D. (2011). *Molecular Ecology* (sec ed). Chichester, England: Wiley. https://doi.org/10.1002/9780470979365

Gai, J., & Lu, J. (2013). Outbreeding. In S. Maloy, & K. Hughes (Eds.), *Brenner's Encyclopedia of Genetics* (sec ed, pp. 197–198). Amsterdam, Netherlands: Elsevier. https://doi.org/10.1016/B978-0-12-37498 4-0.01106 -2

Gibbs, P. E. (2014). Late-acting self-incompatibility – the pariah breeding system in flowering plants. *New Phytologist*, *203*, 717–734. https://doi.org/10.1111/nph.12874

Goldberg, E. E., Kohn, J. R., Lande, R., Robertson, K. A., Smith, S. A., & Igić, B. (2010). Species selection maintains self-incompatibility. *Science*, *330*, 493–496. https://doi.org/10.5040/9780755621101.0007

Haag, C. R., & Ebert, D. (2004). A new hypothesis to explain geographic parthenogenesis. *Annales Zoologici Fennici*, *41*, 539–544.

Hijmans, R. J., Williams, E., & Vennes, C. (2019). *GEOSPHERE: Spherical trigonometry*. Retrieved from https://cran.r-project.org/

Hojsgaard, D., Greilhuber, J., Pellino, M., Paun, O., Sharbel, T. F., & Hörandl, E. (2014). Emergence of apospory and bypass of meiosis via apomixis after sexual hybridisation and polyploidisation. *New Phytologist*, *204*, 1000–1012. https://doi.org/10.1111/nph.12954

Hörandl, E. (2006). The complex causality of geographical parthenogenesis. *New Phytologist*, *171*, 525–538. https://doi.org/10.1111/j.1469-8137.2006.01769.x

Hörandl, E. (2008). Evolutionary implications of self-compatibility and reproductive fitness in the apomictic *Ranunculus auricomus* polyploid complex (Ranunculaceae). *International Journal of Plant Science*, *169*, 1219–1228. https://doi.org/10.1086/591980

Hörandl, E. (2009). Geographical parthenogenesis: Opportunities for asexuality. In I. Schön, K. Martens, & P. Dijk (Eds.), *Lost Sex: The evolutionary biology of parthenogenesis* (pp. 1–615). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-90-481-2770-2

Hörandl, E. (2010). The evolution of self-fertility in apomictic plants. *Sexual Plant Reproduction*, *23*, 73–86. https://doi.org/10.1007/s0049 7-009-0122-3

Hörandl, E., Cosendai, A. C., & Temsch, E. M. (2008). Understanding the geographic distributions of apomictic plants: A case for a pluralistic approach. *Plant Ecology and Diversity*, *1*, 309–320. https://doi.org/10.1080/17550 87080 2351175

Hörandl, E., Dobeš, C., & Lambrou, M. (1997). Chromosomen- und Pollenuntersuchungen an österreichischen Arten des apomiktischen *Ranunculus auricomus*-Komplexes. *Botanica Helvetica*, *107*, 195–209.

Hörandl, E., Dobeš, C., Suda, J., Vít, P., Urfus, T., Temsch, E. M., ... Ladinig, U. (2011). Apomixis is not prevalent in subnival to nival plants of the European Alps. *Annals of Botany*, *108*, 381–390. https://doi.org/10.1093/aob/mcr142

Hörandl, E., Greilhuber, J., & Dobeš, C. (2000). Isozyme variation and ploidy levels within the apomictic *Ranunculus auricomus* complex: Evidence for a sexual progenitor species in Southeastern Austria. *Plant Biology*, *2*, 53–62. https://doi.org/10.1055/s-2000-9148

Hörandl, E., & Gutermann, W. (1998). The *Ranunculus auricomus* complex in Austria: The species of the *R. phragmiteti* and *R. indecorus* groups. *Phyton* (*Annales Rei Botanicae*), *37*, 263–320.

Hörandl, E., Jakubowsky, G., & Dobeš, C. (2001). Isozyme and morphological diversity within apomictic and sexual taxa of the *Ranunculus auricomus* complex. *Plant Systematics and Evolution*, *226*, 165–185. https://doi.org/10.1007/s0060 60170064

Hörandl, E., & Temsch, E. (2009). Introgression of apomixis into sexual species is in the *Ranunculus auricomus* complex inhibited by mentor effects and ploidy barriers. *Annals of Botany*, *104*, 81–89.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*, 254–267. https://doi.org/10.1093/molbev/msj030

Karbstein, K., Tomasello, S., Hodač, L., Dunkel, F. G., Daubert, M., & Hörandl, E. (2020). Phylogenomics supported by geometric morphometrics reveals delimitation of sexual species within the polyploid apomictic *Ranunculus auricomus* complex (Ranunculaceae). *Taxon*, 00, 1–30. https://doi.org/10.1002/tax.12365

Karunarathne, P., Schedler, M., Martínez, E. J., Honfi, A. I., Novichkova, A., & Hojsgaard, D. (2018). Intraspecific ecological niche divergence and reproductive shifts

foster cytotype displacement and provide ecological opportunity to polyploids. *Annals of Botany*, *121*, 1183–1196. https://doi.org/10.1093/aob/mcy004

Kearney, M. (2005). Hybridization, glaciation and geographical parthenogenesis. *Trends in Ecology and Evolution*, *20*, 495–502. https://doi.org/10.1016/j.tree.2005.06.005

Kirchheimer, B., Wessely, J., Gattringer, A., Hülber, K., Moser, D., Schinkel, C. C. F., ... & Dullinger, S. (2018). Reconstructing geographical parthenogenesis: Effects of niche differentiation and reproductive mode on Holocene range expansion of an alpine plant. *Ecology Letters*, *21*, 392–401. https://doi.org/10.1111/ele.12908

Levin, D. A. (1975). Minority cytotype exclusion in local plant populations. *Taxon*, *24*, 35–43. https://doi.org/10.2307/1218997

Lo, E. Y. Y., Stefanović, S., & Dickinson, T. A. (2009). Population genetic structure of diploid sexual and polyploid apomictic hawthorns (*Crataegus*; Rosaceae) in the Pacific Northwest. *Molecular Ecology*, *18*, 1145–1160. https://doi.org/10.1111/j.1365-294X.2009.04091.x

Lo, E. Y. Y., Stefanović, S., & Dickinson, T. A. (2013). Geographical parthenogenesis in Pacific Northwest hawthorns (*Crataegus*; Rosaceae). *Botany-Botanique*, *91*, 107–116. https://doi.org/10.1139/cjb-2012-0073

Lohwasser, U. (2001). Biosystematische Untersuchungen an *Ranunculus auricomus* L. (Ranunculaceae) in Deutschland. *Dissertationes Botanicae*, *343*, 1–220.

Lovell, J. T., Grogan, K., Sharbel, T. F., & McKay, J. K. (2014). Mating system and environmental variation drive patterns of adaptation in *Boechera spatifolia* (Brassicaceae). *Molecular Ecology*, *23*, 4486–4497. https://doi.org/10.1111/mec.12879

Lundqvist, A. (1990). The complex S-gene system for control of self-incompatibility in the buttercup genus *Ranunculus*. *Hereditas*, *113*, 29–46. https://doi.org/10.1111/j.1601-5223.1990.tb006 95.x

Lundqvist, A. (1994). The self-incompatibility system in *Ranunculus repens* (Ranunculaceae). *Hereditas*, *120*, 151–157. https://doi.org/10.1111/j.1601-5223.1994.00151.x

Lundqvist, A. (1998). Disomic control of self-incompatibility in the tetraploid *Ranunculus repens* (Ranunculaceae). *Hereditas*, *128*, 181–183. https://doi.org/10.1111/j.1601-5223.1998.00181.x

Mau, M., Lovell, J. T., Corral, J. M., Kiefer, C., Koch, M. A., Aliyu, O. M., & Sharbel, T. F. (2015). Hybrid apomicts trapped in the ecological niches of their sexual ancestors. *Proceedings of the National Academy of Sciences*, *112*, E2357–E2365. https://doi.org/10.1073/pnas.1423447112

Maynard-Smith, J. (1978). *The evolution of sex*. Cambridge University Press.

Mráz, P., Zdvořák, P., Hartmann, M., Štefánek, M., & Chrtek, J. (2018). Can obligate apomixis and more stable reproductive assurance explain the distributional successes of asexual triploids in *Hieracium alpinum* (Asteraceae)? *Plant Biology*, *21*, 227–236. https://doi.org/10.1111/plb.12930

Müller-Schneider, P. (1986). *Verbreitungsbiologie der Blütenpflanzen Graubündens*. Veröffentlichungen Des Geobotanischen Instituts Der ETH, Stiftung Rübel, Zürich, *85*, 1–263.

Nardi, F. D., Hülber, K., Moser, D., Alonso-Marcos, H., Tribsch, A., & Dobeš, C. (2020). Occurrence of apomictic conspecifics and ecological preferences rather than colonization history govern the geographic distribution of sexual *Potentilla puberula*. *Ecology and Evolution*, *10*, 7306–7319. https://doi.org/10.1002/ece3.6455

Pannell, J. R. (2015). Evolution of the mating system in colonizing plants. *Molecular Ecology*, 24(9), 2018–2037. https://doi.org/10.1111/mec.13087

Paradis, E., Blomberg, S., Bolker, B., Brown, J., Claude, J., Cuong, H. S., & de Vienne, D. (2019). *ape: Analyses of phylogenetics and evolution*. Retrieved from http://ape-package.ird.fr/

Paris, J. R., Stevens, J. R., & Catchen, J. M. (2017). Lost in parameter space: A road map for STACKS. *Methods in Ecology and Evolution*, *8*, 1360–1373. https://doi.org/10.1111/2041-210X.12775

Park, B., & Donoghue, M. J. (2019). Phylogeography of a widespread eastern North American shrub, *Viburnum lantanoides*. *American Journal of Botany*, *106*, 389–401. https://doi.org/10.1002/ajb2.1248

Pätzold, C., Wood, K. R., Eaton, D. A. R., Wagner, W. L., & Appelhans, M. S. (2019). Phylogeny of Hawaiian Melicope (Rutaceae): RADseq resolves species relationships and reveals ancient introgression. *Frontiers in Plant Science*, *10*, 1–16. https://doi.org/10.3389/fpls.2019.01074

Paule, J., Dunkel, F. G., Schmidt, M., & Gregor, T. (2018). Climatic differentiation in polyploid apomictic *Ranunculus auricomus* complex in Europe. *BMC Ecology*, *18*, 1–12. https://doi.org/10.1186/s12898-018-0172-1

Pickering, C. M. (1997). Breeding systems of Australian *Ranunculus* in the alpine region. *Nordic Journal of Botany*, *17*, 613–620. https://doi.org/10.1111/j.1756-1051.1997.tb00357.x

Pinc, J., Chrtek, J., Latzel, V., & Mráz, P. (2020). Negative effect of inbreeding on fitness of an arctic–alpine *Hieracium alpinum* (Asteraceae), a species with a geographical parthenogenesis distribution pattern. *Plant Systematics and Evolution*, *306*, 62. https://doi.org/10.1007/s0060 6-020-01692-6

Quilichini, A., Debussche, M., & Thompson, J. D. (2001). Evidence for local outbreeding depression in the Mediterranean island endemic *Anchusa crispa* Viv. (Boraginaceae). *Heredity*, *87*, 190–197. https://doi.org/10.1046/j.1365-2540.2001.00897.x

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Renner, S. S. (2014). The relative and absolute frequencies of angiosperm sexual systems: Dioecy, monoecy, gynodioecy, and an updated online database. *American Journal of Botany*, *101*, 1588–1596. https://doi.org/10.3732/ajb.1400196

Richards, J. A. (1997). *Plant breeding systems* (sec. ed). London, UK: Chapman and Hall.

Steinbach, K., & Gottsberger, G. (1995). Phenology and pollination biology of five *Ranunculus* species in Central Germany. *Plant Systematics and Evolution*, *9*, 319–323.

Tomasello, S., Karbstein, K., Hodač, L., Pätzold, C., & Hörandl, E. (2020). Phylogenomics unravels Quaternary vicariance and allopatric speciation patterns in temperate-montane plant species: A case study on the *Ranunculus auricomus* species complex. *Molecular Ecology*, *29*, 2031–2049. https://doi.org/10.1111/mec.15458

Vandel, A. (1928). La parthénogenese geographique. Contribution a l'étude biologique et cytologique de la parthénogenese naturelle. *Bulletin Biologique de la France et de la Belgique*, *62*, 164–182.

Vandelook, F. (2009). *Seed germination ecology of temperate woodland herbs*. PhD Thesis. Katholieke Universiteit Leuven.

Vašková, D., & Kolarčik, V. (2019). Breeding systems in diploid and polyploid hawthorns (*Crataegus*): Evidence from experimental pollinations of *C. monogyna*, *C. subsphaerica*, and natural hybrids. *Forests*, *10*, 1–18. https://doi.org/10.3390/F10121059

Vrijenhoek, R. C., & Parker, E. D. (2009). Geographical parthenogenesis: General purpose genotypes and frozen niche variation. In I. Schön, K. Martens, & P. Dijk, (Eds.), *Lost sex: The evolutionary biology of parthenogenesis* (pp. 99–131). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-90-481-2770-2

Wirth, L. R., Waser, N. M., Graf, R., Gugerli, F., Landergott, U., Erhardt, A., ... Holderegger, R. (2011). Effects of floral neighborhood on seed set and degree of outbreeding in a high-alpine cushion plant. *Oecologia*, *167*, 427–434. https://doi.org/10.1007/s00442-011-1985-1

Chapter 3

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*, 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Agudo, A. (2017). *Evolución en Anacyclus L. (Anthemideae, Asteraceae). Análisis de la zona de contacto entre A. clavatus (Desf.) Pers. y A. valentinus L.*. Universidad Autónoma de Madrid, PhD thesis.

Alroy, J., Koch, P. L., & Zachos, J. C. (2000). Global climate change and North American mammalian evolution. *Paleobiology*, *26*, 259–288. https://doi.org/10.1017/S0094837300026968

Alsos, I. G. (2005). Impact of the ice ages on circumpolar molecular diversity: Insights from an ecological key species. *Molecular Ecology*, *14*, 2739–2753.

Andermann, T., Fernandes, A. M., Olsson, U., Töpel, M., Pfeil, B., Oxelman, B., … Antonelli, A. (2019). Allele phasing greatly improves the phylogenetic utility of ultraconserved elements. *Systematic Biology*, *68*, 32–46. https://doi.org/10.1093/sysbio/syy039

Bahr, A., Kaboth, S., Hodell, D., Zeeden, C., Fiebig, J., & Friedrich, O. (2018). Oceanic heat pulses fueling moisture transport towards continental Europe across the mid-Pleistocene transition. *Quaternary Science Reviews*, *179*, 48–58. https://doi.org/10.1016/j.quasc irev.2017.11.009

Baker, H. G. (1967). Support for Baker's law – As a rule. *Evolution*, *21*, 853–856. https://doi.org/10.1111/j.1558-5646.1967.tb03440.x

Bierzychudek, P. (1985). Patterns in plant parthenogenesis. *Experientia*, *41*, 1255–1264. https://doi.org/10.1007/BF019 52068

Bolger, A. M., Lohse, M., & Usade, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Borowiec, M. L. (2016). AMAS: A fast tool for alignment manipulation and computing of summary statistics. *PeerJ*, *4*, e1660. https://doi.org/10.7717/peerj.1660

Bouckaert, R., Vaughan, T. G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A., … Drummond, A. J. (2019). BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *15*, e1006650.

Brewer, S., Cheddadi, R., de Beaulieu, J.-L., Reille, M., & Contributors, D. (2002). The spread of deciduous *Quercus* throughout Europe since the last glacial period. *Forest Ecology and Management*, *156*, 27–48. https://doi.org/10.1016/S0378-1127(01)00646-6

Capella-Gutiérrez, S., Silla-Martínez, J. M., & Gabaldón, T. (2009). trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics*, *25*, 1972–1973. https://doi.org/10.1093/bioinformatics/btp348

Chamala, S., García, N., Godden, G. T., Krishnakumar, V., Jordon-Thaden, I. E., Smet, R. D., … Soltis, P. S. (2015). MarkerMiner 1.0: A new application for phylogenetic marker development using angiosperm transcriptomes. *Applications in Plant Sciences*, *3*, 1400115. https://doi.org/10.3732/apps.1400115

Chao, Z., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, *19*, 153. https://doi.org/10.1186/s1285 9-018-2129-y

Chen, L. Y., Zhao, S. Y., Wang, Q. F., & Moody, M. L. (2015). Transcriptome sequencing of three *Ranunculus* species (Ranunculaceae) reveals candidate genes in adaptation from terrestrial to aquatic habitats. *Scientific Reports*, *5*, e10098. https://doi.org/10.1038/srep10098

Clark, P. U., Archer, D., Pollard, D., Blum, J. D., Rial, J. A., Brovkin, V., … Roy, M. (2006). The middle Pleistocene transition: Characteristics, mechanisms, and implications for long-term changes in atmospheric pCO_2. *Quaternary Science Reviews*, *25*, 3150–3184. https://doi.org/10.1016/j.quascirev.2006.07.008

Comes, H. P., & Kadereit, J. W. (1998). The effect of Quaternary climatic changes on plant distribution and evolution. *Trends in Plant Science*, *3*, 432–438. https://doi.org/10.1016/S1360-1385(98)01327-2

R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org

Cosendai, A. C., Wagner, J., Ladinig, U., Rosche, C., & Hörandl, E. (2013). Geographical parthenogenesis and population genetic structure in the alpine species *Ranunculus kuepferi* (Ranunculaceae). *Heredity*, *110*, 560–569.

Couvreur, T. L. P., Pirie, M. D., Chatrou, L. W., Saunders, R. M. K., Su, Y. C. F., Richardson, J. E., & Erkens, R. H. J. (2011). Early evolutionary history of the flowering plant family Annonaceae: Steady diversification and boreotropical geodispersal. *Journal of Biogeography*, *38*, 664–680. https://doi.org/10.1111/j.1365-2699.2010.02434.x

Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: More models, new heuristics and parallel computing. *Nature Methods*, *9*, 772. https://doi.org/10.1038/nmeth.2109

Després, L., Loriot, S., & Gaudeul, M. (2002). Geographic pattern of genetic variation in the European globeflower *Trollius europaeus* L. (Ranunculaceae) inferred from amplified fragment length polymorphism markers. *Molecular Ecology*, *11*, 2337–2347. https://doi.org/10.1046/j.1365-294X.2002.01618.x

Drummond, A. J., & Bouckaert, R. R. (2015). *Bayesian evolutionary analysis with BEAST*. Cambridge, UK: Cambridge University Press.

Dunkel, F. G. (2011). The *Ranunculus auricomus* L. complex (Ranunculaceae) in Central and Southern Italy, with additions to North Italian taxa. *Webbia*, *66*, 165–193.

Dunkel, F. G., Gregor, T., & Paule, J. (2018). New diploid species in the *Ranunculus auricomus* complex (Ranunculaceae) from W and SE Europe. *Willdenowia*, *48*, 227–257. https://doi.org/10.3372/wi.48.48205

Elderfield, H., Ferretti, P., Greaves, M., Crowhurst, S., McCave, I. N., Hodell, D., & Piotrowski, A. M. (2012). Evolution of ocean temperature and ice volume through the Mid-Pleistocene Climate Transition. *Science*, *337*, 704–709. https://doi.org/10.1126/science.1221294

Emadzade, K., Gehrke, B., Linder, H. P., & Hörandl, E. (2011). The biogeographical history of the cosmopolitan genus *Ranunculus* L. (Ranunculaceae) in the temperate to meridional zones. *Molecular Phylogenetics and Evolution*, *58*, 4–21. https://doi.org/10.1016/j.ympev.2010.11.002

Emadzade, K., & Hörandl, E. (2011). Northern Hemisphere origin, transoceanic dispersal, and diversification of *Ranunculeae* DC. (Ranunculaceae) in the Cenozoic. *Journal of Biogeography*, *38*, 517–530. https://doi.org/10.1111/j.1365-2699.2010.02404.x

Emadzade, K., Lebmann, M., Hoffmann, M. H., Tkach, N., Lone, F., & Hörandl, E. (2015). Phylogenetic relationships and evolution of high mountain buttercups (*Ranunculus*) in North America and Central Asia. *Perspectives in Plant Ecology, Evolution and Systematics*, *17*, 131–141. https://doi.org/10.1016/j.ppees.2015.02.001

Eriksson, J. S., de Sousa, F., Bertrand, Y. J. C., Antonelli, A., Oxelman, B., & Pfeil, B. E. (2018). Allele phasing is critical to revealing a shared allopolyploid origin of *Medicago arborea* and *M. strasseri* (Fabaceae). *BMC Evolutionary Biology*, *18*, 9. https://doi.org/10.1186/s1286 2-018-1127-z

Fér, T., & Schmickl, R. E. (2018). Hybphylomaker: Target enrichment data analysis from raw reads to species trees. *Evolutionary Bioinformatics*, *14*, 1–9. https://doi.org/10.1177/1176934317 742613

Folk, R. A., Mandel, J. R., & Freudenstein, J. V. (2015). A protocol for targeted enrichment of intron-containing sequence markers for recent radiations: A phylogenomic example with genomic resources from *Heuchera* (Saxifragaceae). *Applications in Plant Sciences*, *3*, 1500039.

Fragoso-Martínez, I., Salazar, G. A., Martínez-Gordillo, M., Magallón, S., Sánchez-Reyes, L., Moriarty Lemmon, E., … Granados Mendoza, C. (2017). A pilot study applying the plant Anchored Hybrid Enrichment method to New World sages (*Salvia* subgenus *Calosphace*; Lamiaceae). *Molecular Phylogenetics and Evolution*, *117*, 124–134. https://doi.org/10.1016/j.ympev.2017.02.006

Gaboardi, M., Deng, T., & Wang, Y. (2005). Middle Pleistocene climate and habitat change at Zhoukoudian, China, from the carbon and oxygen isotopic record from herbivore tooth enamel. *Quaternary Research*, *63*, 329–338. https://doi.org/10.1016/j.yqres.2005.02.006

Giarla, T. C., & Esselstyn, J. (2015). The challenges of resolving a rapid, recent radiation: Empirical and simulated phylogenomics of Philippine shrews. *Systematic Biology*, *64*, 727–740. https://doi.org/10.1093/sysbio/syv029

Grummer, J. A., Morando, M. M., Avila, L. J., Sites, J. W. Jr, & Leaché, A. D. (2018). Phylogenomic evidence for a recent and rapid radiation of lizards in the Patagonian *Liolaemus ftzingerii* species group. *Molecular Phylogenetics and Evolution*, *125*, 243–254.

Harvey, M. G., Smith, B. T., Glenn, T. C., Faircloth, C. B., & Brumfield, R. T. (2016). Sequence capture versus restriction site associated DNA sequencing for shallow systematics. *Systematic Biology*, *65*, 910–924. https://doi.org/10.1093/sysbio/syw036

Head, M. J., & Gibbard, P. L. (2005). Early-Middle Pleistocene transitions: An overview and recommendation for the defining boundary. In M. L. Head, & E. L. Gibbard (Eds.), *Early-middle pleistocene transitions: The land-ocean evidence* (pp. 1–18). London, UK: Geological Society.

Herrando-Moraira, S., & The Cardueae Radiations Group (2018). Exploring data processing strategies in NGS target enrichment to disentangle radiations in the tribe Cardueae (Compositae). *Molecular Phylogenetics and Evolution*, *128*, 69–87. https://doi.org/10.1016/j.ympev.2018.07.012

Hewitt, G. M. (1999). Post-glacial re-colonization of European biota. *Biological Journal of the Linnean Society*, *68*, 87–112. https://doi.org/10.1111/j.1095-8312.1999.tb01160.x

Hojsgaard, D., Greilhuber, J., Pellino, M., Paun, O., Sharbel, T. F., & Hörandl, E. (2014). Emergence of apospory and bypass of meiosis via apomixis after sexual hybridisation and polyploidisation. *New Phytologist*, *204*, 1000–1012. https://doi.org/10.1111/nph.12954

Hörandl, E. (2004). Comparative analysis of genetic divergence among sexual ancestor of apomictic complexes using isozyme data. *International Journal of Plant Sciences*, *165*, 615–622.

Hörandl, E. (2006). The complex causality of geographical parthenogenesis. *New Phytologist*, *171*, 525–538. https://doi.org/10.1111/j.1469-8137.2006.01769.x

Hörandl, E. (2008). Evolutionary implications of self-compatibility and reproductive fitness in the apomictic *Ranunculus auricomus* polyploid complex (Ranunculaceae). *International Journal of Plant Sciences*, *169*, 1219–1228.

Hörandl, E., & Greilhuber, J. (2002). Diploid and autotetraploid sexuals and their relationships to apomicts in the *Ranunculus cassubicus* group: Insights from DNA content and isozyme variation. *Plant Systematics and Evolution*, *234*, 85–100.

Hörandl, E., Greilhuber, J., & Dobes, C. (2000). Isozyme variation and ploidy levels within the apomictic *Ranunculus auricomus* complex: Evidence for a sexual progenitor species in southeastern Austria. *Plant Biology*, *2*, 53–62. https://doi.org/10.1055/s-2000-9148

Hörandl, E., Greilhuber, J., Klímová, K., Paun, O., Temsch, E., Emadzade, K., & Hodálová, I. (2009). Reticulate evolution and taxonomic concepts in the *Ranunculus auricomus* complex (Ranunculaceae): Insights from analysis of morphological, karyological and molecular data. *Taxon*, *58*, 1194–1216. https://doi.org/10.1002/tax.584012

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*, 254–267. https://doi.org/10.1093/molbev/msj030

Junier, T., & Zdobnov, E. M. (2010). The Newick utilities: Highthroughput phylogenetic tree processing in the UNIX shell. *Bioinformatics*, *26*, 1669–1670. https://doi.org/10.1093/bioinformatics/btq243

Kadereit, J. W., Griebeler, E. M., & Comes, H. P. (2004). Quaternary diversification in European Alpine plants: Pattern and process. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *359*, 265–274.

Karbstein, K., Tomasello, S., Hodač, L., Dunkel, F. G., Daubert, M., & Hörandl, E. (2020). Phylogenomics supported by geometric morphometrics reveals delimitation of sexual species within the polyploid apomictic *Ranunculus auricomus* complex (Ranunculaceae). *BioRxiv*, 1–58. https://doi.org/10.1101/2020.01.07.896902

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780. https://doi.org/10.1093/molbev/mst010

Kent, W. J. (2002). BLAT - The BLAST-like alignment tool. *Genome Research*, *12*, 656–664. https://doi.org/10.1101/gr.229202

Kirchheimer, B., Wessely, J., Gattringer, A., Hülber, K., Moser, D., Schinkel, C. C. F., … Dullinger, S. (2018). Reconstructing geographical parthenogenesis: Effects of niche differentiation and reproductive mode on Holocene range expansion of an alpine plant. *Ecology Letters*, *21*, 392–401. https://doi.org/10.1111/ele.12908

Knapp, M., Stöckler, K., Havell, D., Delsuc, F., Sebastiani, F., & Lockhart, P. (2005). Relaxed molecular clock provides evidence for long-distance dispersal of *Nothofagus* (southern beech). *PLoS Biology*, *3*, e14. https://doi.org/10.1371/journal.pbio.0030014

Knowles, L. L. (2001). Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Molecular Ecology*, *10*, 691–701.

Knowles, L. L., & Chan, Y.-H. (2008). Resolving species phylogenies of recent evolutionary radiations. *Annals of the Missouri Botanical Garden*, *95*, 224–231.

Kramp, K., Huck, S., Niketić, M., Tomović, G., & Schmitt, T. (2008). Multiple glacial refugia and complex postglacial range shifts of the obligatory woodland plant *Polygonatum verticillatum* (Convallariaceae). *Plant Biology*, *11*, 392–404.

Kropf, M., Kadereit, J. W., & Comes, H. P. (2003). Differential cycles of range contraction and expansion in European high mountain plants during the Late Quaternary: Insights from *Pritzelago alpina* (L.) O. Kuntze (Brassicaceae). *Molecular Ecology*, *12*, 931–949. https://doi.org/10.1046/j.1365-294X.2003.01781.x

Kubatko, L. S., & Degnan, J. H. (2007). Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology*, *56*, 17–24. https://doi.org/10.1080/10635150601146041

Kumar, S., Filipski, A. J., Battistuzzi, F. U., Pond, S. L. K., & Tamura, K. (2012). Statistics and truth in phylogenomics. *Molecular Biology and Evolution*, *29*, 457–472. https://doi.org/10.1093/molbev/msr202

Levin, D. (1975). Minority cytotype exclusion in local plant populations. *Taxon*, *24*, 35–43. https://doi.org/10.2307/1218997

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., & Homer, N., …1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Lisiecki, L. L., & Raymo, M. E. (2005). A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. *Paleoceanography*, *20*, PA1003. https://doi.org/10.1029/2004PA001071

López-Giráldez, F., & Townsend, J. P. (2011). PhyDesign: An online application for profiling phylogenetic informativeness. *BMC Evolutionary Biology*, *11*, 152. https://doi.org/10.1186/1471-2148-11-152

Magri, D. (2008). Patterns of post-glacial spread and the extent of glacial refugia of European beech (*Fagus sylvatica*). *Journal of Biogeography*, *35*, 450–463. https://doi.org/10.1111/j.1365-2699.2007.01803.x

Magri, D., & Palombo, M. R. (2013). Early to Middle Pleistocene dynamics of plant and mammal communities in South West Europe. *Quaternary International*, *288*, 63–72. https://doi.org/10.1016/j.quaint.2012.02.028

Magri, D., Vendramin, G. G., Comps, B., Dupanloup, I., Geburek, T., Gomory, D., … de Beaulieu, J.-L. (2006). A new scenario for the Quaternary history of European beech

populations: Palaeobotanical evidence and genetic consequences. *New Phytologist*, *171*, 199–221. https://doi.org/10.1111/j.1469-8137.2006.01740.x

Masci, S., Miho, A., & Marchi, P. (1994). *Ranunculus auricomus* L. aggr. (Ranunculaceae) in Italy. I. Sexual tetraploids on the Apennines. *Caryologia*, *47*, 97–108.

Matzke, N. J. (2013). Probabilistic historical biogeography: New models for founder-event speciation, imperfect detection, and fossils allow improved accuracy and model-testing. *Frontiers of Biogeography*, *5*, 242–248. https://doi.org/10.21425/F55419694

Matzke, N. J. (2018). *BioGeoBEARS: BioGeography with Bayesian (and likelihood) evolutionary analysis with R scripts*. Retrieved from https://doi.org/10.5281/zenodo.1478250

McCormack, J. E., Harvey, M. G., Faircloth, B. C., Crawford, N. G., Glenn, T. C., & Brumfield, R. T. (2013). A phylogeny of birds based on over 1,500 loci collected by target enrichment and high-throughput sequencing. *PLoS ONE*, *8*, e54848. https://doi.org/10.1371/journal.pone.0054848

McKain, M. R., Johnson, M. G., Uribe-Convers, S., Eaton, D., & Yang, Y. (2018). Practical considerations for plant phylogenomics. *Applications in Plant Sciences*, *6*, e1038. https://doi.org/10.1002/aps3.1038

Naydenov, K., Senneville, S., Beaulieu, J., Tremblay, F., & Bousquet, J. (2007). Glacial vicariance in Eurasia: Mitochondrial DNA evidence from Scots pine for a complex heritage involving genetically distinct refugia at mid-northern latitudes and in Asia Minor. *BMC Evolutionary Biology*, *7*, 233. https://doi.org/10.1186/1471-2148-7-233

Nei, M. (1975). Molecular population genetics and evolution. *Frontiers in Biology*, *40*, 1–288.

Paradis, E., & Schliep, K. (2018). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526–528. https://doi.org/10.1093/bioinformatics/bty633

Pastre, J.-F., Gauthier, A., Nomade, S., Orth, P., Andrieu, A., Goupille, F., … Renne, P. R. (2007). The Alleret maar (Massif Central, France): A new lacustrine sequence of the early Middle Pleistocene in western Europe. *Comptes Rendus Geoscience*, *339*, 987–997. https://doi.org/10.1016/j.crte.2007.09.019

Paun, O., Stuessy, T. F., & Hörandl, E. (2006). The role of hybridization, polyploidization and glaciation in the origin and evolution of the apomictic *Ranunculus cassubicus* complex. *New Phytologist*, *171*, 223–236. https://doi.org/10.1111/j.1469-8137.2006.01738.x

Pease, J. B., Brown, J. W., Walker, J. F., Hinchliff, C. E., & Smith, S. A. (2018). Quartet Sampling distinguishes lack of support from conflicting support in the green plant tree of life. *American Journal of Botany*, *105*, 385–403. https://doi.org/10.1002/ajb2.1016

Pellino, M., Hojsgaard, D., Schmutzer, T., Scholz, U., Hörandl, E., Vogel, H., & Sharbel, T. F. (2013). Asexual genome evolution in the apomictic *Ranunculus auricomus* complex: Examining the effects of hybridization and mutation accumulation. *Molecular Ecology*, *22*, 5908–5921.

Pond, S. L. K., Frost, S. D. W., & Muse, S. V. (2005). HyPhy: Hypothesis testing using phylogenies. *Bioinformatics*, *21*, 676–679. https://doi.org/10.1093/bioinformatics/bti079

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2 – Approximately maximum-likelihood trees for large alignments. *PLoS ONE*, *5*, e9490. https://doi.org/10.1371/journal.pone.0009490

Qiu, Y.-X., Guan, B.-C., Fu, C.-X., & Comes, H. P. (2009). Did glacials and/or interglacials promote allopatric incipient speciation in East Asian temperate plants? Phylogeographic and coalescent analyses on refugial isolation and divergence in *Dysosma versipellis*. *Molecular Phylogenetics and Evolution*, *51*, 281–293. https://doi.org/10.1016/j.ympev.2009.01.016

Rambaut, A., & Drummond, A. J. (2007). *Tracer v1.4: MCMC trace analyses tool*. Retrieved from http://beast.bio.ed.ac.uk/Tracer

Raymo, M. E., Lisiecki, L. E., & Nisancioglu, K. H. (2006). Plio-Pleistocene ice volume, antarctic climate, and the global δ18O record. *Science*, *313*, 492–495.

Reisch, C. (2008). Glacial history of *Saxifraga paniculata* (Saxifragaceae): Molecular biogeography of a disjunct arcticalpine species from Europe and North America. *Biological Journal of the Linnean Society*, *93*, 385–398. https://doi.org/10.1111/j.1095-8312.2007.00933.x

Ronquist, F., & Sanmartín, I. (2011). Phylogenetic methods in biogeography. *Annual Review of Ecology, Evolution, and Systematics*, *42*, 441–464. https://doi.org/10.1146/annurev-ecolsys-102209-144710

Rull, V. (2009). Microrefugia. *Journal of Biogeography*, *36*, 481–484. https://doi.org/10.1111/j.1365-2699.2008.02023.x

Sánchez Goñi, M. F., Rodrigues, T., Hodell, D. A., Polanco-Martínez, J. M., Alonso-García, M., Hernández-Almeida, I., … Ferretti, P. (2016). Tropically-driven climate shifts in southwestern Europe during MIS 19, a low eccentricity interglacial. *Earth and Planetary Science Letters*, *448*, 81–93. https://doi.org/10.1016/j.epsl.2016.05.018

Sanderson, M. J. (2002). Estimating absolute rates of molecular evolution and divergence times: A penalized likelihood approach. *Molecular Biology Evolution*, *19*, 101–109. https://doi.org/10.1093/oxfordjournals.molbev.a003974

Sanmartín, I. (2009). Dispersal versus vicariance. In McGraw-Hill (ed.), *Yearbook of Science and Technology* (pp. 85–88). New York, NY: McGraw-Hill.

Sanmartín, I. (2012). Historical biogeography: Evolution in time and space. *Evolution: Education and Outreach*, *5*, 555–568. https://doi.org/10.1007/s12052-012-0421-2

Sanz, M., Schönswetter, P., Vallés, J., Schneeweiss, G. M., & Vilartesana, R. (2014). Southern isolation and northern long distance dispersal shaped the phylogeography of the widespread, but highly disjunct, European high mountain plant *Artemisia eriantha* (Asteraceae). *Botanical Journal of the Linnean Society*, *174*, 214–226.

Sayyari, E., & Mirarab, S. (2016). Fast coalescent-based computation of local branch support from quartet frequencies. *Molecular Biology and Evolution*, *33*, 1654–1668. https://doi.org/10.1093/molbev/msw079

Schönswetter, P., Stehlik, I., Holderegger, R., & Tribsch, A. (2005). Molecular evidence for glacial refugia of mountain plants in the European Alps. *Molecular Ecology*, *14*, 3547–3555. https://doi.org/10.1111/j.1365-294X.2005.02683.x

Schönswetter, P., & Tribsch, A. (2005). Vicariance and dispersal in the Alpine perennial, *Bupleurum stellatum* L. (Apiaceae). *Taxon*, *54*, 725–732.

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Stebbins, G. L. (1984). Polyploidy and the distribution of the Arctic-Alpine flora: New evidence and a new approach. *Botanica Helvetica*, *94*, 1–13.

Stephens, J. D., Rogers, W. L., Heyduk, K., Cruse-Sanders, J. M., Determann, R. O., Glenn, T. C., & Malmberg, R. L. (2015). Resolving phylogenetic relationships for the recently radiated carnivorous plant genus *Sarracenia* using target enrichment. *Molecular Phylogenetics and Evolution*, *85*, 76–87.

Stuchlik, L., & Wójcik, A. (2001). Pollen analysis of Malopolanian Interglacial deposits at Lowisko (Kolbuszowa Upland, southern Poland). *Acta Palaeobotanica*, *41*, 15–26.

Taberlet, P., Fumagalli, L., Wust-Saucy, A. G., & Cosson, J.-F. (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Molecular Ecology*, *7*, 453–464. https://doi.org/10.1046/j.1365-294x.1998.00289.x

Tomasello, S., & Oberprieler, C. (2017). Frozen ploidies: A phylogeographical analysis of the Leucanthemopsis alpina polyploid complex (Asteraceae, Anthemideae). *Botanical Journal of the Linnean Society*, *183*, 211–235. https://doi.org/10.1093/botlinnean/bow009

Townsend, J. P. (2007). Profiling phylogenetic informativeness. *Systematic Biology*, 56, 222–231. https://doi.org/10.1080/10635150701311362

Tribsch, A., Schönswetter, P., & Stuessy, T. F. (2002). *Saponaria pumila* (Caryophyllaceae) and the ice age in the European Alps. *American Journal of Botany*, *89*, 2024–2033. https://doi.org/10.3732/ajb.89.12.2024

Van Dam, M., & Matzke, N. J. (2016). Evaluating the influence of connectivity and distance on biogeographic patterns in the south-western deserts of North America. *Journal of Biogeography*, *43*, 1514–1532.

Vargas, P. (2003). Molecular evidence for multiple diversification patterns of Alpine plants in Mediterranean Europe. *Taxon*, *52*, 463–476. https://doi.org/10.2307/3647383

Weisrock, D. W., Smith, S. D., Chan, L. M., Biebouw, K., Kappeler, P. M., & Yoder, A. D. (2012). Concatenation and concordance in the reconstruction of mouse lemur phylogeny: An empirical demonstration of the effect of allele sampling in phylogenetics. *Molecular Biology and Evolution*, *29*, 1615–1630. https://doi.org/10.1093/molbev/mss008

Weitemier, K., Straub, S. C., Cronn, R. C., Fishbein, M., Schmickl, R., McDonnell, A., & Liston, A. (2014). Hybseq: Combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Sciences*, *2*, 1400042. https://doi.org/10.3732/apps.1400042

Widmer, A., & Lexer, C. (2001). Glacial refugia: Sanctuaries for allelic richness, but not for gene diversity. *Trends in Ecology and Evolution*, *16*, 267–269. https://doi.org/10.1016/S0169-5347(01)02163-2

Wolfe, K. H., Gouy, M. L., Yang, Y. W., Sharp, P. M., & Li, W. H. (1989). Date of the monocot dicot divergence estimated from chloroplast DNA-sequence data. *Proceedings of the National Academy of Sciences*, *86*, 6201–6205. https://doi.org/10.1073/pnas.86.16.6201

Xu, H., Luo, X., Qian, J., Pang, X., Song, J., Qian, G., … Chen, S. (2012). FastUniq: A fast de novo duplicates removal tool for paired short reads. *PLoS ONE*, *7*, e52249. https://doi.org/10.1371/journal.pone.0052249

Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, *24*, 1586–1591. https://doi.org/10.1093/molbev/msm088

Zhao, C., Wang, C. B., Ma, X. G., Liang, Q. L., & He, X. J. (2013). Phylogeographic analysis of a temperate-deciduous forest restricted plant (*Bupleurum longiradiatum* Turcz.) reveals two refuge areas in China with subsequent refugial isolation promoting speciation. *Molecular Phylogenetics and Evolution*, *68*, 628–643. https://doi.org/10.1016/j.ympev.2013.04.007

Zhao, W., Tarasov, P. E., Lozhkin, A. V., Anderson, P. M., Andreev, A. A., Korzun, J. A., … Wennrich, V. (2017). High-latitude vegetation and climate changes during the Mid-Pleistocene Transition inferred from apalynological record from Lake El'gygytgyn, NE Russian Arctic. *Boreas*, *47*, 137–149. https://doi.org/10.1111/bor.12262

Zhou, X., Yang, J., Wang, S., Xiao, G., Zhao, K., Zheng, Y., … Li, X. (2018). Vegetation change and evolutionary response of large mammal fauna during the Mid-Pleistocene Transition in temperate northern East Asia. *Palaeogeography Palaeoclimatology Palaeoecology*, *505*, 287–292. https://doi.org/10.1016/j.palaeo.2018.06.007

Chapter 4

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*, 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Alix, K., Gérard, P. R., Schwarzacher, T., & Heslop-Harrison, J. S. P. (2017). Polyploidy and interspecific hybridization: Partners for adaptation, speciation and evolution in plants. *Annals of Botany*, *120*, 183–194. https://doi.org/10.1093/aob/mcx079

Alonso-Marcos, H., Nardi, F. D., Scheffknecht, S., Tribsch, A., Hülber, K., & Dobeš, C. (2019). Difference in reproductive mode rather than ploidy explains niche differentiation in sympatric sexual and apomictic populations of *Potentilla puberula*. *Ecology and Evolution*, *9*, 3588–3598. https://doi.org/10.1002/ece3.4992

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. Retrieved from https://www.bioin formatics.babraham.ac.uk/proje cts/fastqc/.

Asker, S., & Jerling, L. (1992). *Apomixis in plants*. CRC Press.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A., & Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, *3*, 1–7. https://doi.org/10.1371/journal.pone.0003376

Baker, H. G. (1965). Characteristics and modes of origin of weeds. In H. G. Baker, & G. L. Stebbins (Eds.), *The genetics of colonizing species* (pp. 147–172). New York, USA: Academic Press.

Baker, H. G. (1967). Support for Baker's law –As a rule. *Evolution*, *21*, 853–856. https://doi.org/10.1017/CBO97 81107415324.004

Barke, B. H., Daubert, M., & Hörandl, E. (2018). Establishment of apomixis in diploid F2 hybrids and inheritance of apospory from F1 to F2 hybrids of the *Ranunculus auricomus* complex. *Frontiers in Plant Science*, *9*, 1–12.

Beck, J. B., Alexander, P. J., Allphin, L., Al-Shehbaz, I. A., Rushworth, C., Bailey, C. D., & Windham, M. D. (2012). Does hybridization drive the transition to asexuality in diploid *Boechera*? *Evolution*, *66*, 985–995. https://doi.org/10.1111/j.1558-5646.2011.01507.x.

Becker, D., Verheul, J., Zickel, M., & Willmes, C. (2015). LGM paleoenvironment of Europe –map (CRC806-Database). https://doi.org/10.5880/SFB806.15

Becker, R. A., Wilks, A. R., Brownrigg, R., Minka, T. P., & Deckmyn, A. (2018). *MAPS: Draw geographical maps*. Retrieved from https://cran.r-project.org/.

Bierzychudek, P. (1985). Patterns in plant parthenogenesis. *Experientia*, *41*, 1255–1264.

Borchers-Kolb, E. (1983). *Ranunculus* sect. *Auricomus* in Bayern und den angrenzenden Gebieten -I. Allgemeiner Teil. *Mitteilungen der Botanischen Staatssammlung München*, *19*, 363–429

Borchers-Kolb, E. (1985). Ranunculus sect. Auricomus in Bayern und den angrenzenden Gebieten II Spezieller Teil. *Mitteilungen der Botanischen Staatssammlung München*, *21*, 49–300.

Carman, J. G. (1997). Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony. *Biological Journal of the Linnean Society*, *61*, 51–94. https://doi.org/10.1006/bijl.1996.0118.

Chen, Z. J. (2007). Genetic and epigenetic mechanisms for gene expression and phenotypic variation in plant polyploids. *Annual Review of Plant Biology*, *58*, 377–406.

Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nature Reviews Genetics*, *6*, 836–846. https://doi.org/10.1038/nrg1711

Cosendai, A. C., Wagner, J., Ladinig, U., Rosche, C., & Hörandl, E. (2013). Geographical parthenogenesis and population genetic structure in the alpine species *Ranunculus kuepferi* (Ranunculaceae). *Heredity*, *110*, 560–569. https://doi.org/10.1038/hdy.2013.1

Coughlan, J. M., Han, S., Stefanović, S., & Dickinson, T. A. (2017). Widespread generalist clones are associated with range and niche expansion in allopolyploids of Pacific Northwest Hawthorns (*Crataegus* L.). *Molecular Ecology*, *26*, 5484–5499. https://doi.org/10.1111/mec.14331

Daubert, M. (2016). Effects of cold stress on the reproduction mode of synthetic hybrids from the *Ranunculus auricomus* complex. University of Göttingen, Bachelor thesis.

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., & Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, *12*, 499–510. https://doi.org/10.1038/nrg3012.

Dobeš, C. H., Mitchell-Olds, T., & Koch, M. A. (2004). Extensive chloroplast haplotype variation indicates Pleistocene hybridization and radiation of North American *Arabis drummondii*, *A. x divaricarpa*, and A. *holboellii* (Brassicaceae). *Molecular Ecology*, *13*, 349–370. https://doi.org/10.1046/j.1365-294X.2003.02064.x.

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D. & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*, 027–046. https://doi.org/10.1111/j.1600-0587.2012.07348.x

Dunkel, F. G., Gregor, T., & Paule, J. (2018). New diploid species in the *Ranunculus auricomus* complex (Ranunculaceae) from W and SE Europe. *Willdenowia*, *48*, 227–257. https://doi.org/10.3372/wi.48.48205

Dunkel, F. G. (2014). Le complexe de *Ranunculus auricomus* (Ranunculaceae) en Alsace. *Journal de Botanique de la Société Botanique de France*, *66*, 3–53.

Eaton, D. A. R. (2014). PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics*, *30*, 1844–1849. https://doi.org/10.1093/bioinformatics/btu121

Eaton, D. A. R., & Overcast, I. (2020). IPYRAD: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, *36*, 2592–2594. https://doi.org/10.1093/bioinformatics/btz966.

Espinoza, F. (2002). Effect of pollination timing on the rate of apomictic reproduction revealed by RAPD markers in *Paspalum notatum*. *Annals of Botany*, *89*, 165–170. https://doi.org/10.1093/aob/mcf024

Felsenstein, J. (1974). The evolutionary advantage of recombination. *Genetics*, *78*, 737–756.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*, 4302–4315. https://doi.org/10.1002/joc.5086

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, *196*, 973–983. https://doi.org/10.1534/genetics.113.160572

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, *6*, 925–929. https://doi.org/10.1111/2041-210X.12382.

Frichot, E., & Francois, O. (2020). *LEA: An R package for landscape and ecological association studies*. Retrieved from https://bioconductor.org/. https://doi.org/10.18129/B9.bioc.LEA

Garnier, S., Ross, N., Rudis, B., Sciaini, M., & Scherer, C. (2018). *viridis: Defaut color maps from "matplotlib"*. Retrieved from https://github.com/sjmga rnier/viridis

GBIF Secretariat (2017). *Ranunculus auricomus* L. https://doi.org/10.15468/39omei

Goslee, S., & Urban, D. (2020). *ECODIST*. Retrieved from https://cran.r-project.org/.

Gornall, R. J. (1999). Population genetic structure in agamospermous plants. In P. M. Hollingsworth, R. M. Bateman, & R. J. Gornall (Eds.), *Molecular Systematics and Plant Evolution* (pp. 118–138). London, UK: Taylor & Francis.

Grant, V. (1981). *Plant speciation*. Ney York, USA: Columbia University Press.

Gregor, T. (2013). Apomicts in the vegetation of Central Europe. *Tuexenia*, *33*, 233–257.

Hand, M. L., & Koltunow, A. M. G. (2014). The genetic control of apomixis: Asexual seed formation. *Genetics*, *197*, 441–450. https://doi.org/10.1534/genetics.114.163105

Hartmann, M., Štefánek, M., Zdvořák, P., Heřman, P., Chrtek, J., & Mráz, P. (2017). The Red Queen hypothesis and geographical parthenogenesis in the alpine hawkweed

Hieracium alpinum (Asteraceae). *Biological Journal of the Linnean Society*, *122*, 681–696. https://doi.org/10.1093/biolinnean/blx105

Hijmans, R. J., Sumner, M., Cheng, J., Bevan, A., Bivand, R., Busetto, L., … Wueest, R. (2020). *RASTER: Geographic data analysis and modeling*. Retrieved from https://cran.r-project.org

Hijmans, R. J., Williams, E., & Vennes, C. (2019). *GEOSPHERE: Spherical trigonometry*. Retrieved from https://cran.r-project.org/.

Hodač, L., Barke, B. H., & Hörandl, E. (2018). Mendelian segregation of leaf phenotypes in experimental F_2 hybrids elucidates origin of morphological diversity of the apomictic *Ranunculus auricomus* complex. *Taxon*, *67*, 1082–1092. https://doi.org/10.12705/676.6

Hodač, L., Klatt, S., Hojsgaard, D., Sharbel, T. F., & Hörandl, E. (2019). A little bit of sex prevents mutation accumulation even in apomictic polyploid plants. *BMC Evolutionary Biology*, *19*, 170. https://doi.org/10.1186/s12862-019-1495-z

Hodač, L., Scheben, A. P., Hojsgaard, D., Paun, O., & Hörand, E. (2014). ITS polymorphisms shed light on hybrid evolution in apomictic plants: A case study on the *Ranunculus auricomus* complex. *PLoS ONE*, *9*, 28–30. https://doi.org/10.1371/journal.pone.0103003

Hojsgaard, D., Greilhuber, J., Pellino, M., Paun, O., Sharbel, T. F., & Hörandl, E. (2014). Emergence of apospory and bypass of meiosis via apomixis after sexual hybridisation and polyploidisation. *New Phytologist*, *204*, 1000–1012. https://doi.org/10.1111/nph.12954

Hojsgaard, D., & Hörandl, E. (2019). The rise of apomixis in natural plant populations. *Frontiers in Plant Science*, *10*, 1–13. https://doi.org/10.3389/fpls.2019.00358

Hörandl, E., & Raab-Straube, E. (2015). *Ranunculaceae*. Retrieved from Euro+Med Plantbase – The information resource for Euro-Mediterranean plant diversity. Retrieved from http://ww2.bgbm.org/EuroPlusMed

Hörandl, E. (2008). Evolutionary implications of self-compatibility and reproductive fitness in the apomictic *Ranunculus auricomus* polyploid complex (Ranunculaceae). *International Journal of Plant Science*, *169*, 1219–1228. https://doi.org/10.1086/591980

Hörandl, E. (2004). Comparative analysis of genetic divergence among sexual ancestors of apomictic complexes using isozyme data. *International Journal of Plant Sciences*, *165*, 615–622. https://doi.org/10.1086/386557

Hörandl, E. (2006). The complex causality of geographical parthenogenesis. *New Phytologist*, *171*, 525–538. https://doi.org/10.1111/j.1469-8137.2006.01769.x

Hörandl, E. (2010). The evolution of self-fertility in apomictic plants. *Sexual Plant Reproduction*, *23*, 73–86. https://doi.org/10.1007/s00497-009-0122-3

Hörandl, E., Cosendai, A.-C., & Temsch, E. M. (2008). Understanding the geographic distributions of apomictic plants: A case for a pluralistic approach. *Plant Ecology and Diversity*, *1*, 309–320. https://doi.org/10.1080/17550870802351175

Hörandl, E., & Greilhuber, J. (2002). Diploid and autotetraploid sexuals and their relationships to apomicts in the *Ranunculus cassubicus* group: Insights from DNA content and isozyme variation. *Plant Systematics and Evolution*, *234*, 85–100. https://doi.org/10.1007/s00606-002-0209-x

Hörandl, E., Greilhuber, J., Klímová, K., Paun, O., Temsch, E., Emadzade, K., & Hodálová, I. (2009). Reticulate evolution and taxonomic concepts in the *Ranunculus auricomus* complex (Ranunculaceae): Insights from analysis of morphological, karyological and molecular data. *Taxon*, *58*, 1194–1215.

Hörandl, E., & Paun, O. (2007). Patterns and sources of genetic diversity in apomictic plants: implications for evolutionary potentials and ecology. In: E. Hörandl, U. Grossniklaus, P. J. Van Dijk, & T. Sharbel (Eds.), *Apomixis: Evolution, Mechanisms and Perspective* (pp. 169–194). Ruggell, Liechtenstein: ARG-Gantner.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*, 254–267. https://doi.org/10.1093/molbev/msj030

Jalas, J., & Suominen, J. (1989). *Atlas florae europaeae*, vol. 8, Nymphaeaceae to Ranunculaceae. The committee for mapping the flora of Europe and Societas Biologica Fennica Vanamo.

Kahle, D., Wickham, D., Jackson, S., & Korpela, M. (2019). *ggmap: Spatial visualization with ggplot2*. Retrieved from https://github.com/dkahle/ggmap.

Kantama, L., Sharbel, T. F., Schranz, M. E., Mitchell-Olds, T., de Vries, S., & de Jong, H. (2007). Diploid apomicts of the *Boechera holboellii* complex display large-scale chromosome substitutions and aberrant chromosomes. *Proceedings of the National Academy of Sciences*, *104*, 14026–14031. https://doi.org/10.1073/pnas.0706647104

Karbstein, K., Tomasello, S., Hodač, L., Dunkel, F. G., Daubert, M., & Hörandl, E. (2020a). Phylogenomics supported by geometric morphometrics reveals delimitation of sexual species within the polyploid apomictic *Ranunculus auricomus* complex (Ranunculaceae). *Taxon*, *69*, 1191–1220. https://doi.org/10.1002/tax.12365

Karbstein, K., Rahmsdorf, E., Tomasello, S., Hodač, L., & Hörandl, E. (2020b). Breeding system of diploid sexuals within the *Ranunculus auricomus* complex and its role in a geographical parthenogenesis scenario. *Ecology and Evolution*, *10*, 14435–14450. https://doi.org/10.1002/ece3.7073

Karbstein, K., Tomasello, S., Hodač, L., Lorberg, E., Daubert, M., & Hörandl, E. (2020). The biodiversity of apomictic polyploid plants: The *Ranunculus auricomus* complex, National Center for Biotechnology Information Sequence Read Archive (SRA)

BioProject ID PRJNA627796. Retrived from http://www.ncbi.nlm.nih.gov/bioproject/627796

Karbstein, K., Tomasello, S., Hodač, L., Lorberg, E., Daubert, M., & Hörandl, E. (2021). The biodiversity of apomictic polyploid plants: The *Ranunculus auricomus* complex - FC and FCSS data. Figshare, https://doi.org/10.6084/m9.figshare.13352429

Karunarathne, P., Schedler, M., Martínez, E. J., Honfi, A. I., Novichkova, A., & Hojsgaard, D. (2018). Intraspecific ecological niche divergence and reproductive shifts foster cytotype displacement and provide ecological opportunity to polyploids. *Annals of Botany*, *121*, 1183–1196. https://doi.org/10.1093/aob/mcy004

Kearney, M. (2005). Hybridization, glaciation and geographical parthenogenesis. *Trends in Ecology and Evolution*, *20*, 495–502. https://doi.org/10.1016/j.tree.2005.06.005

Kirchheimer, B., Schinkel, C. C. F., Dellinger, A. S., Klatt, S., Moser, D., Winkler, M., … Dullinger, S. (2016). A matter of scale: Apparent niche differentiation of diploid and tetraploid plants may depend on extent and grain of analysis. *Journal of Biogeography*, *43*, 716–726. https://doi.org/10.1111/jbi.12663

Kirchheimer, B., Wessely, J., Gattringer, A., Hülber, K., Moser, D., Schinkel, C. C. F., … Dullinger, S. (2018). Reconstructing geographical parthenogenesis: Effects of niche differentiation and reproductive mode on Holocene range expansion of an alpine plant. *Ecology Letters*, *21*, 392–401. https://doi.org/10.1111/ele.12908

Klatt, S., Hadacek, F., Hodač, L., Brinkmann, G., Eilerts, M., Hojsgaard, D., & Hörandl, E. (2016). Photoperiod extension enhances sexual megaspore formation and triggers metabolic reprogramming in facultative apomictic *Ranunculus auricomus*. *Frontiers in Plant Science*, *7*, 1–13. https://doi.org/10.3389/fpls.2016.00278

Klatt, S., Schinkel, C. C. F., Kirchheimer, B., Dullinger, S., & Hörandl, E. (2018). Effects of cold treatments on fitness and mode of reproduction in the diploid and polyploid alpine plant *Ranunculus kuepferi* (Ranunculaceae). *Annals of Botany*, *121*, 1287–1298. https://doi.org/10.1093/aob/mcy017.

Körner, C. (2003). *Alpine plant life: Functional plant ecology of high mountain ecosystems*. Springer Verlag.

Kumar, S. (2017). Epigenetic control of apomixis: A new perspective of an old enigma. *Advances in Plants & Agriculture Research*, *7*, 1–8. https://doi.org/10.15406/apar.2017.07.00243

Leebens-Mack, J. H., Barker, M. S., Carpenter, E. J., Deyholos, M. K., Gitzendanner, M. A., Graham, S. W., … Wong, G. K. S. (2019). One thousand plant transcriptomes and the phylogenomics of green plants. *Nature*, *574*, 679–685. https://doi.org/10.1038/s41586-019-1693-2.

Lefcheck, J. S. (2015). PiecewiseSEM: Piecewise structural equation modelling in R for ecology, evolution, and systematics. *Methods in Ecology and Evolution*, *7*, 573–579. https://doi.org/10.1111/2041-210X.12512.

Lefcheck, J. S. (2021). *Piecewise structural equation modeling*. Retrieved from https://cran.r-project.org.

Lo, E. Y. Y., Stefanović, S., & Dickinson, T. A. (2013). Geographical parthenogenesis in Pacific Northwest hawthorns (*Crataegus*; Rosaceae). *Botany*, *91*, 107–116. https://doi.org/10.1139/cjb-2012-0073.

Lovell, J. T., Grogan, K., Sharbel, T. F., & McKay, J. K. (2014). Mating system and environmental variation drive patterns of adaptation in *Boechera spatifolia* (Brassicaceae). *Molecular Ecology*, *23*, 4486–4497. https://doi.org/10.1111/mec.12879

Lushai, G., Loxdale, H. D., & Allen, J. A. (2003). The dynamic clonal genome and its adaptive potential. *Biological Journal of the Linnean Society*, *79*, 193–208. https://doi.org/10.1046/j.1095-8312.2003.00189.x

Lughadha, E. N., Govaerts, R., Belyaeva, I., Black, N., Lindon, H., Allkin, R., Magill, R., & Nicolson, N. (2016). Counting counts: Revised estimates of numbers of accepted species of flowering plants, seed plants, vascular plants and land plants with a review of other recent estimates. *Phytotaxa*, *272*, 082–088. https://doi.org/10.11646/phytotaxa.272.1.5

Lynch, M. (1984). Destabilizing hybridization, general-purpose genotypes and geographic parthenogenesis. *Quarterly Review of Biology*, *59*, 257–290.

Matzk, F., Meister, A., & Schubert, I. (2000). An efficient screen for reproductive pathways using mature seeds of monocots and dicots. *Plant Journal*, *21*, 97–108. https://doi.org/10.1046/j.1365-313X.2000.00647.x

Meusel, H., Jäger, E., Rauschert, S., & Weinert, E. (1978). *Vergleichende Chorologie der Zentraleuropaischen Flora*, Vol. 2. Gustav Fischer Verlag.

Mohammadin, S., Wang, W., Liu, T., Moazzeni, H., Ertugrul, K., Uysal, T., Christodoulou, C. S., Edger, P. P., Pires, J. C., Wright, S. I., & Schranz, M. E. (2018). Genome-wide nucleotide diversity and associations with geography, ploidy level and glucosinolate profiles in *Aethionema arabicum* (Brassicaceae). *Plant Systematics and Evolution*, *304*, 619–630. https://doi.org/10.1007/s00606-018-1494-3

Mráz, P., Chrtek, J., & Šingliarová, B. (2009). Geographical parthenogenesis, genome size variation and pollen production in the arctic-alpine species *Hieracium alpinum*. *Botanica Helvetica*, *119*, 41–51. https://doi.org/10.1007/s00035-009-0055-3

Muller, H. J. (1964). The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *1*, 2–9.

Naumova, T. N., Hayward, M. D., & Wagenvoort, M. (1999). Apomixis and sexuality in diploid and tetraploid accessions of *Brachiaria decumbens*. *Sexual Plant Reproduction*, *12*, 43–52. https://doi.org/10.1007/s004970050170

Nogler, G. A. (1984). Genetics of apospory in apomictic *Ranunculus auricomus*. V. Conclusion. *Botanica Helvetica*, *94*, 411–422.

Nogler, G. A. (2006). The lesser-known Mendel: His experiments on *Hieracium*. *Genetics*, *172*(1), 1–6.

One Thousand Plant Transcriptomes Initiative (2019). One thousand plant transcriptomes and the phylogenomics of green plants. *Nature*, *574*, 679–685. https://doi.org/10.1038/s41586-019-1693-2

Ortiz, J. P. A., Quarin, C. L., Pessino, S. C., Acuña, C., Martínez, E. J., Espinoza, F., Hojsgaard, D. H., Sartor, M. E., Cáceres, M. E., & Pupilli, F. (2013). Harnessing apomictic reproduction in grasses: What we have learned from *Paspalum*. *Annals of Botany*, *112*, 767–787. https://doi.org/10.1093/aob/mct152

Ozias-Akins, P., & van Dijk, P. J. (2007). Mendelian genetics of apomixis in plants. *Annual Review of Genetics*, *41*, 509–537. https://doi.org/10.1146/annurev.genet.40.110405.090511

Paradis, E., Blomberg, S., Bolker, B., Brown, J., Claude, J., Cuong, H. S., de Vienne, D. (2020). *APE: Analyses of phylogenetics and evolution*. Retrieved from http://ape-package.ird.fr/

Paule, J., Dunkel, F. G., Schmidt, M., & Gregor, T. (2018). Climatic differentiation in polyploid apomictic *Ranunculus auricomus* complex in Europe. *BMC Ecology*, *18*, 1–12. https://doi.org/10.1186/s12898-018-0172-1

Paule, J., Kolář, F., & Dobeš, C. (2015). Arctic-alpine and serpentine differentiation in polyploid *Potentilla crantzii*. *Preslia*, *87*, 195–215.

Paun, O., Greilhuber, J., Temsch, E. M., & Hörandl, E. (2006). Patterns, sources and ecological implications of clonal diversity in apomictic *Ranunculus carpaticola* (*Ranunculus auricomus* complex, Ranunculaceae). *Molecular Ecology*, *15*, 897–910. https://doi.org/10.1111/j.1365-294X.2006.02800.x

Paun, O., Stuessy, T. F., & Hörandl, E. (2006). The role of hybridization, polyploidization and glaciation in the origin and evolution of the apomictic *Ranunculus cassubicus* complex. *New Phytologist*, *171*, 223–236. https://doi.org/10.1111/j.1469-8137.2006.01738.x

Pegoraro, L., Baker, E. C., Aesschimann, D., Balant, M., Douzet, R., Garnatje, T., … Pellicer, J. (2020). The correlation of phylogenetics, elevation, and ploidy on the incidence of apomixis in Asteraceae in the European Alps. *Botanical Journal of the Linnean Society*, *194*, 410–422. https://doi.org/10.1093/aobpla/plw064

Pellino, M., Hojsgaard, D., Schmutzer, T., Scholz, U., Hörandl, E., Vogel, H., & Sharbel, T. F. (2013). Asexual genome evolution in the apomictic *Ranunculus auricomus* complex: Examining the effects of hybridization and mutation accumulation. *Molecular Ecology*, *22*, 5908–5921. https://doi.org/10.1111/mec.12533

Qiu, T., Liu, Z., & Liu, B. (2020). The effects of hybridization and genome doubling in plant evolution via allopolyploidy. *Molecular Biology Reports*, *47*, 5549–5558. https://doi.org/10.1007/s11033-020-05597-y

R Core Team (2020). *R: A language and environment for statistical computing*. Retrieved from http://www.r-project.org/.

Rice, A., Šmarda, P., Novosolov, M., Drori, M., Glick, L., Sabath, N., … Mayrose, I. (2019). The global biogeography of polyploid plants. *Nature Ecology and Evolution*, *3*, 265–273. https://doi.org/10.1038/s41559-018-0787-9

Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., & Firth, D. (2020). *MASS: Support functions and datasets for Venables and Ripley's MASS*. Retrieved from https://cran.r-project.org.

Rushworth, C. A., Windham, M. D., Keith, R. A., & Mitchell-Olds, T. (2018). Ecological differentiation facilitates fine-scale coexistence of sexual and asexual *Boechera*. *American Journal of Botany*, *105*, 2051–2064. https://doi.org/10.1002/ajb2.1201

Schinkel, C. C. F., Kirchheimer, B., Dellinger, A. S., Klatt, S., Winkler, M., Dullinger, S., & Hörandl, E. (2016). Correlations of polyploidy and apomixis with elevation and associated environmental gradients in an alpine plant. *AoB Plants*, *8*, plw064. https://doi.org/10.1093/aobpla/plw064

Soltis, P. S., & Soltis, D. E. (2000). The role of genetic and genomic attributes in the success of polyploids. *Proceedings of the National Academy of Sciences*, *97*, 7051–7057. https://doi.org/10.1073/pnas.97.13.7051

Stebbins, G. L. (1950). *Variation and evolution in plants*. New York, USA: Columbia University Press.

Steinbach, K., & Gottsberger, G. (1994). Phenology and pollination biology of five *Ranunculus* species in Giessen, Central Germany. *Phyton* (Horn, Austria), *34*, 203–218.

Štorchová, H., Chrtek, J., Bartish, I. V., Tetera, M., Kirschner, J., & Štěpánek, J. (2002). Genetic variation in agamospermous taxa of *Hieracium* sect. *Alpina* (Compositae) in the Tatry Mts. (Slovakia). *Plant Systematics and Evolution*, *235*, 1–17. https://doi.org/10.1007/s00606-002-0219-8

Tomasello, S., Karbstein, K., Hodač, L., Pätzold, C., & Hörandl, E. (2020). Phylogenomics unravels Quaternary vicariance and allopatric speciation patterns in temperate-montane plant species: A case study on the *Ranunculus auricomus* species complex. *Molecular Ecology*, *29*, 2031–2049. https://doi.org/10.1111/mec.15458

Ulum, F. B., Costa Castro, C., & Hörandl, E. (2020). Ploidy-dependent effects of light stress on the mode of reproduction in the *Ranunculus auricomus* complex (Ranunculaceae). *Frontiers in Plant Science*, *11*, 104. https://doi.org/10.3389/fpls.2020.00104

Van de Peer, Y., Mizrachi, E., & Marchal, K. (2017). The evolutionary significance of polyploidy. *Nature Reviews Genetics*, *18*, 411. https://doi.org/10.1038/nrg.2017.26

van Dijk, P. J. (2003). Ecological and evolutionary opportunities of apomixis: Insights from *Taraxacum* and *Chondrilla*. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*, 1113–1121. https://doi.org/10.1098/rstb.2003.1302

Vandel, A. (1928). La parthénogenese geographique. Contribution a l'étude biologique et cytologique de la parthénogenese naturelle. *Bulletin Biologique de La France et de La Belgique*, *62*, 164–182.

Venables, W. N., & Ripley, B. D. (2002). Modern applied statistics with S (4th ed.). Springer.

Vrijenhoek, R. C. (1979). Factors affecting clonal diversity and coexistence. *American Zoologist*, *19*, 787–797.

Vrijenhoek, R. C., & Parker, E. D. (2009). Geographical parthenogenesis: General purpose genotypes and frozen niche variation. In I. Schön, K. Martens, & P. Dijk, (Eds.), *Lost sex: The evolutionary biology of parthenogenesis* (pp. 99–131). Dordrecht, Netherlands: Springer. https://doi.org/10.1007/978-90-481-2770-2

Welch, D. M., & Meselson, M. (2000). Evidence for the evolution of Bdelloid rotifers without sexual reproduction or genetic exchange. *Science*, *288*, 1211–1215. https://doi.org/10.1126/science.288.5469.1211

Wickham, H., Francois, R., Henry, L., & Müller, K. (2019). *dplyr: A grammar of data manipulation*. Retrieved from https://github.com/tidyverse/dplyr.

Winston, C., Lionel, H., Thomas Lin, P., Kohske, T., Claus, W., Kara, W., & Hiroaki, Y. (2019). *GGPLOT2: Create elegant data visualisations using the grammar of graphics*. Retrieved from https://cran.r-project.org.

Zinn, K. E., Tunc-Ozdemir, M., & Harper, J. F. (2010). Temperature stress and plant sexual reproduction: Uncovering the weakest links. *Journal of Experimental Botany*, *61*(7), 1959–1968. https://doi.org/10.1093/jxb/erq053

Chapter 5

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., … Zinner, D. (2013). Hybridization and speciation. *Journal of Evolutionary Biology*, *26*, 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Alger, E. I., & Edger, P. P. (2020). One subgenome to rule them all: Underlying mechanisms of subgenome dominance. *Current Opinion in Plant Biology*, *54*, 108–113. https://doi.org/10.1016/j.pbi.2020.03.004

Andermann, T., Fernandes, A. M., Olsson, U., Töpel, M., Pfeil, B., Oxelman, B., ... Antonelli, A. (2019). Allele phasing greatly improves the phylogenetic utility of ultraconserved elements. *Systematic Biology*, *68*, 32–46. https://doi.org/10.1093/sysbio/syy039

Asker, S., & Jerling, L. (1992). *Apomixis in plants*. Boca Raton, CRC press.

Babcock, G.T., & Stebbins, G.L. (1938). *The American species of Crepis. Their interrelationships and distribution as affected by polyploidy and apomixis*. Carnegie Inst. Washington 504.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., … Johnson, E. A. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, *3*, 1–7. https://doi.org/10.1371/journal.pone.0003376

Barke, B. H., Daubert, M., & Hörandl, E. (2018). Establishment of apomixis in diploid F_2 hybrids and inheritance of apospory from F_1 to F_2 hybrids of the *Ranunculus auricomus* complex. *Frontiers in Plant Science*, *9*, 1–12. https://doi.org/10.3389/fpls.2018.01111

Barke, B. H., Karbstein, K., Daubert, M., & Hörandl, E. (2020). The relation of meiotic behaviour to hybridity, polyploidy and apomixis in the *Ranunculus auricomus* complex (Ranunculaceae). *BMC Plant Biology*, *20*, 523. https://doi.org/10.1186/s12870-020-02654-3

Barker, M. S., Arrigo, N., Baniaga, A. E., Li, Z., Levin, D. A. (2016). On the relative abundance of autopolyploids and allopolyploids. *New Phytologist*, *210*, 391–398. https://doi.org/10.1111/nph.13698

Te Beest, M., Le Roux, J. J., Richardson, D. M., Brysting, A. K., Suda, J., Kubesova, M., Pysek, P. (2012). The more the better? The role of polyploidy in facilitating plant invasions. *Annnals of Botany*, *109*, 19–45.

Bertrand, Y. J., Scheen, A. C., Marcussen, T., Pfeil, B. E., de Sousa, F., Oxelman, B. (2015). Assignment of homoeologs to parental genomes in allopolyploids for species tree inference, with an example from *Fumaria* (Papaveraceae). *Systematic Biology*, *64*, 448–471. https://doi.org/10.1093/sysbio/syv004

Blischak, P. D., Mabry, M. E., Conant, G. C., & Pires, J. C. (2018). Integrating networks, phylogenomics, and population genomics for the study of polyploidy. *Annual Review of*

Ecology, Evolution, and Systematics, *49*, 253–278. https://doi.org/10.1146/annurev-ecolsys-121415-032302

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., … Drummond, A. J. (2014). BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*, e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Brandrud, M. K., Baar, J., Lorenzo, M. T., Athanasiadis, A., Bateman, R. M., Chase, M. W., ... Paun, O. (2020). Phylogenomic relationships of diploids and the origins of allotetraploids in *Dactylorhiza* (Orchidaceae). *Systematic Biology*, *69*, 91–109. https://doi.org/10.1093/sysbio/syz035

Brewer, G. E., Clarkson, J. J., Maurin, O., Zuntini, A. R., Barber, V., Bellot, S., ... Baker W. J. (2019). Factors affecting targeted sequencing of 353 nuclear genes from herbarium specimens spanning the diversity of angiosperms. *Frontiers in Plant Science*, *10*, 1102. https://doi.org/10.3389/fpls.2019.01102

Brukhin, V., Osadtchiy, J. V., Florez-Rueda, A. M., Smetanin, D., Bakin, E., Nobre, M. S., Grossniklaus, U. (2019). The *Boechera* genus as a resource for apomixis research. *Frontiers in Plant Science*, *10*, 19. https://doi.org/10.3389/fpls.2019.00392

Burgess, M. B., Cushman, K. R., Doucette, E. T., Frye, C. T., Campbell, C. S. (2015). Understanding diploid diversity: A first step in unraveling polyploid, apomictic complexity in *Amelanchier*. *American Journal of Botany*, *102*, 2041–2057. https://doi.org/10.3732/ajb.1500330

Cai, R., & Ané, C. (2020). Assessing the fit of the multi-species network coalescent to multi-locus data. *Bioinformatics*, 37, 634–641. https://doi.org/10.1093/bioinformatics/btaa863

Cao, Z., Liu, X., Ogilvie, H. A., Yan, Z., & Nakhleh L. (2019). Practical aspects of phylogenetic network analysis using Phylonet. *BioRxiv*, 1–39. https://doi.org/10.1101/746362

Carter, K. A., Liston, A., Bassil, N. V., Alice, L. A., Bushakra, J. M., Sutherland, B. L., … Hummer, K. E. (2019). Target capture sequencing unravels *Rubus* evolution. *Frontiers in Plant Science*, *10*, 1615. https://doi.org/10.3389/fpls.2019.01615

Clement, M., Posada, D., & Crandall, K. A. (2000). TCS: A computer program to estimate gene genealogies. *Molecular Ecology*, *9*, 1657–1659. https://doi.org/10.1046/j.1365-294x.2000.01020.x

Comai, L. (2005). The advantages and disadvantages of being polyploid. *Nature Reviews Genetics*, *6*, 836–846. https://doi.org/10.1038/nrg1711

Coyne J., & Orr H. (2004). *Speciation*. Sunderland, Massachusetts, U.S.A., Sinauer Ass.

Fehrer, J., Slavikova, R., Pastova, L., Josefiova, J., Mráz, P., Chrtek, J., & Bertrand, Y. J. K. (2021). Molecular evolution and organization of ribosomal DNA in the hawkweed tribe Hieraciinae (Cichorieae, Asteraceae). *Frontiers in Plant Science*, *12*, 23. https://doi.org/10.3389/fpls.2021.647375

Darriba, D., Posada, D., Kozlov, A. M., Stamatakis, A., Morel, B., & Flouri, T. (2020). ModelTest-NG: A new and scalable tool for the selection of DNA and protein evolutionary models. *Molecular Biology and Evolution*, *37*, 291–294. https://doi.org/10.1093/molbev/msz189

Dauphin, B., Grant, J. R., Farrar, D. R., & Rothfels, C. J. (2018). Rapid allopolyploid radiation of moonwort ferns (*Botrychium*; Ophioglossaceae) revealed by PacBio sequencing of homologous and homeologous nuclear regions. *Molecular Phylogenetics and Evolution*, *120*, 342–353. https://doi.org/10.1016/j.ympev.2017.11.025

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Review Genetics*, *12*, 499–510. https://doi.org/10.1038/nrg3012

Eaton, D. A. R., & Overcast, I. (2020). IPYRAD: Interactive assembly and analysis of RADseq datasets. *Bioinformatics*, *36*, 2592–2594. https://doi.org/10.1093/bioinformatics/btz966

Eaton, D. A. R., Spriggs, E. L., Park, B., & Donoghue, M. J. (2017). Misconceptions on missing data in RADseq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology*, *66*, 399–412. https://doi.org/10.1093/sysbio/syw092

Eriksson, J. S., de Sousa, F., Bertrand, Y. J. K., Antonelli, A., Oxelman, B., & Pfeil, B. E. (2018). Allele phasing is critical to revealing a shared allopolyploid origin of *Medicago arborea* and *M. strasseri* (Fabaceae). *BMC Evolutionary Biology*, *18*, 9. https://doi.org/10.1186/s12862-018-1127-z

Fér, T., & Schmickl, R. E. (2018). HybPhyloMaker: Target enrichment data analysis from raw reads to species trees. *Evolutionary Bioinformatics*, *14*, 1–9. https://doi.org/10.1177/1176934317742613

Fellers, J. P. (2008). Genome filtering using methylation-sensitive restriction enzymes with six base pair recognition sites. *Plant Genome*, *1*, 146–152. https://doi.org/10.3835/plantgenome2008.05.0245

Flouri, T., Jiao, X., Rannala, B., Yang, Z. (2020). A Bayesian implementation of the multispecies coalescent model with introgression for phylogenomic analysis. *Molecular Biology and Evolution*, *37*, 1211–1223. https://doi.org/10.1093/molbev/msz296

Folk, R. A., Mandel, J. R., & Freudenstein, J. V. (2015). A protocol for targeted enrichment of intron-containing sequence markers for recent radiations: A phylogenomic example from *Heuchera* (Saxifragaceae). *Applications in Plant Sciences*, *3*, 1500039. https://doi.org/10.3732/apps.1500039

Fox, D. T., Soltis, D. E., Soltis, P. S., Ashman, T. L., Van de Peer, Y. (2020). Polyploidy: a biological force from cells to ecosystems. *Trends in Cell Biology*, *30*, 688–694. https://doi.org/10.1016/j.tcb.2020.06.006

Freyman, W. A., Johnson, M. G., & Rothfels, C. J. (2020). HOMOLOGIZER:

Phylogenetic phasing of gene copies into polyploid subgenomes. *BioRxiv*, 1–55. https://doi.org/10.1101/2020.10.22.351486

Frichot, E., & François, O. (2015). LEA: An R package for landscape and ecological association studies. *Methods in Ecology and Evolution*, *6*, 925–929. https://doi.org/10.1111/2041-210X.12382

Frichot E., & Francois O. (2020). LEA: An R package for landscape and ecological association studies. Retrieved from https://bioconductor.org/packages/release/bioc/html/LEA.html

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & Francois, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, *196*, 973–983. https://doi.org/10.1534/genetics.113.160572

Gordon, S. P., Contreras-Moreira, B., Levy, J. J., Djamei, A., Czedik-Eysenberg, A., Tartaglio, V. S., … Vogel, J. P. (2020). Gradual polyploid genome evolution revealed by pan-genomic analysis of *Brachypodium hybridum* and its diploid progenitors. *Nature Communications*, *11*, 1–16. https://doi.org/10.1038/s41467-020-17302-5

Gornall, R. J. (1999). Population genetic structure in agamospermous plants. In P. M. Hollingsworth, R. M. Bateman, & R. J. Gornall (Eds.), *Molecular Systematics and Plant Evolution* (pp. 118–138). London, UK: Taylor & Francis.

Grant, V. (1981). *Plant speciation* (sec ed). New York, USA: Columbia University Press.

Herrando-Moraira, S., Calleja, J. A., Carnicero, P., Fujikawa, K., Galbany-Casals, M., Garcia-Jacas, N., … Vilatersana, R. (2018). Exploring data processing strategies in NGS target enrichment to disentangle radiations in the tribe Cardueae (Compositae). *Molecular Phylogenetics and Evolution*, *128*, 69–87. https://doi.org/10.1016/j.ympev.2018.07.012

Hipp, A. L., Eaton, D. A. R., Cavender-Bares, J., Fitzek, E., Nipper, R., & Manos, P. S. (2014). A framework phylogeny of the American oak clade based on sequenced RAD data. *PLoS ONE*, *9*, e93975. https://doi.org/10.1371/journal.pone.0093975

Hodač, L., Barke, B. H., & Hörandl, E. (2018). Mendelian segregation of leaf phenotypes in experimental F2 hybrids elucidates origin of morphological diversity of the apomictic *Ranunculus auricomus* complex. *Taxon*, *67*, 1082–1092. https://doi.org/10.12705/676.6

Hodač, L., Klatt, S., Hojsgaard, D., Sharbel, T. F., & Hörandl, E. (2019). A little bit of sex prevents mutation accumulation even in apomictic polyploid plants. *BMC Evolutionary Biology*, *19*, 170. https://doi.org/10.1186/s12862-019-1495-z

Hodač, L., Scheben, A. P., Hojsgaard, D., Paun, O., & Hörand, E. (2014). ITS polymorphisms shed light on hybrid evolution in apomictic plants: A case study on the *Ranunculus auricomus* complex. *PLoS ONE*, *9*, 28–30. https://doi.org/10.1371/journal.pone.0103003

Harrison, N., & Kidner, C. A. (2011). Next-generation sequencing and systematics: What

can a billion base pairs of DNA sequence data do for you? *Taxon*, *60*, 1552–1566. https://doi.org/10.1002/tax.606002

Hojsgaard, D., Greilhuber, J., Pellino, M., Paun, O., Sharbel, T.F., & Hörandl, E. (2014). Emergence of apospory and bypass of meiosis via apomixis after sexual hybridisation and polyploidisation. *New Phytologist*, *204*, 1000–1012. https://doi.org/10.1111/nph.12954

Hojsgaard, D., & Hörandl, E. (2019). The rise of apomixis in natural plant populations. *Frontiers in Plant Science*, *10*, 358. https://doi.org/10.3389/fpls.2019.00358

Hörandl, E. (2006). The complex causality of geographical parthenogenesis. *New Phytologist*, *171*, 525–538. https://doi.org/10.1111/j.1469-8137.2006.01769.x

Hörandl, E. (2018). The classification of asexual organisms: Old myths, new facts, and a novel pluralistic approach. *Taxon*, *67*, 1066–1081. https://doi.org/10.12705/676.5

Hörandl, E., Greilhuber, J., Klímová, K., Paun, O., Temsch, E., Emadzade, K., & Hodálová, I. (2009). Reticulate evolution and taxonomic concepts in the *Ranunculus auricomus* complex (Ranunculaceae): Insights from analysis of morphological, karyological and molecular data. *Taxon*, *58*, 1194–1215.

Hörandl, E., & Paun, O. (2007). Patterns and sources of genetic diversity in apomictic plants: Implications for evolutionary potentials. In: E. Hörandl, U. Grossniklaus, P. J. Van Dijk, & T. Sharbel (Eds.), *Apomixis: Evolution, Mechanisms and Perspective* (pp. 169–194). Ruggell, Liechtenstein: ARG-Gantner.

Hörandl, E., & Greilhuber, J. (2002). Diploid and autotetraploid sexuals and their relationships to apomicts in the *Ranunculus cassubicus* group: insights from DNA content and isozyme variation. *Plant Systematics and Evolution*, *234*, 85–100. https://doi.org/10.1007/s00606-002-0209-x

Huang, D. I., Hefer, C. A., Kolosova, N., Douglas, C. J., & Cronk, Q. C. B. (2014). Whole plastome sequencing reveals deep plastid divergence and cytonuclear discordance between closely related balsam poplars, *Populus balsamifera* and *P. trichocarpa* (Salicaceae). *New Phytologist*, *204*, 693–703. https://doi.org/10.1111/nph.12956

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*, 254–267. https://doi.org/10.1093/molbev/msj030

Jalas, J., & Suominen, J. (1989). *Atlas Florae Europaeae. Distribution of vascular plants in Europe, vol. 8, Nymphaeaceae to Ranunculaceae*. Helsinki, The Committee for Mapping the Flora of Europe. Societas Biologica Fennica. Vanamo.

Jaron, K. S., Bast, J., Nowell, R. W., Ranallo-Benavidez, T. R., Robinson-Rechavi, M., & Schwander, T. (2020). Genomic features of parthenogenetic animals. *Journal of Heredity*, *112*, 19–33. https://doi.org/10.1093/jhered/esaa031

Johnson, M. G., Pokorny, L., Dodsworth, S., Botigue, L. R., Cowan, R. S., Devault, A., ...

Baker, W. J. (2019). A universal probe set for targeted sequencing of 353 nuclear genes from any flowering plant designed using k-medoids clustering. *Systematic Biology*, *68*, 594–606. https://doi.org/10.1093/sysbio/syy086

Jones, G. (2017a). Algorithmic improvements to species delimitation and phylogeny estimation under the multispecies coalescent. *Journal of Mathematical Biology*, *74*, 447–467. https://doi.org/10.1007/s00285-016-1034-0

Jones, G. (2017b). Bayesian phylogenetic analysis for diploid and allotetraploid species networks. *BioRxiv*, 1–15. https://doi.org/10.1101/129361

Jones, G., Aydin, Z., & Oxelman, B. (2015). DISSECT: An assignment-free Bayesian discovery method for species delimitation under the multispecies coalescent. *Bioinformatics*, *31*, 991–998. https://doi.org/10.1093/bioinformatics/btu770

Kamneva, O. K., Syring, J., Liston, A., & Rosenberg, N. A. (2017). Evaluating allopolyploid origins in strawberries (*Fragaria*) using haplotypes generated from target capture sequencing. *BMC Evolutionary Biology*, *17*, 180. https://doi.org/10.1186/s12862-017-1019-7

Karbstein, K., Rahmsdorf, E., Tomasello, S., Hodač, L., & Hörandl, E. (2020). Breeding system of diploid sexuals within the *Ranunculus auricomus* complex and its role in a geographical parthenogenesis scenario. *Ecology and Evolution*, *10*, 14435–14450. https://doi.org/10.1002/ece3.7073

Karbstein, K., Tomasello, S., Hodač, L., Dunkel, F. G., Daubert, M., & Hörandl, E. (2020). Phylogenomics supported by geometric morphometrics reveals delimitation of sexual species within the polyploid apomictic *Ranunculus auricomus* complex (Ranunculaceae). *Taxon*, *69*, 1191–1220. https://doi.org/10.1002/tax.12365

Karbstein, K., Tomasello, S., Hodač, L., Lorberg, E., Daubert, M., & Hörandl, E. (2021). Moving beyond assumptions: Polyploidy and environmental effects explain a geographical parthenogenesis scenario in European plants. *Molecular Ecology*, *00*, 1–17. https://doi.org/10.1111/mec.15919

Karbstein, K., Tomasello, S., Hodač, L., Lorberg, E., Daubert, M., & Hörandl, E. (2020). The biodiversity of apomictic polyploid plants: The *Ranunculus auricomus* complex, National Center for Biotechnology Information Sequence Read Archive (SRA) BioProject ID PRJNA627796, http://www.ncbi.nlm.nih.gov/bioproject/627796.

Karbstein, K., Tomasello, S., Hodač, L., Lorberg, E., Daubert, M., & Hörandl, E. (2021). The biodiversity of apomictic polyploid plants: The *Ranunculus auricomus* complex - FC and FCSS data, Figshare, https://doi.org/10.6084/m9.figshare.13352429.

Karbstein, K., Tomasello, S., Hodač, L., Wagner, W., Marinček, P., Barke, B. H., Pätzold, C., & Hörandl, E. (2021). The biodiversity of apomictic polyploid plants: The *Ranunculus auricomus* complex - RADseq, TE, and CP (genomic) data, Figshare, https://doi.org/10.6084/m9.figshare.14046305

Karbstein, K., Tomasello, S., Hodač, L., Wagner, W., Marinček, P., Barke, B. H., Pätzold, C., & Hörandl, E. (2021). Ranunculus_auricomus_scripts_phylogenetic_networks_sniploid, Github, https://github.com/KK260/Ranunculus_auricomus_phylogenetic_network_scripts.

Karbstein, K., Tomasello, S., & Prinz, K. (2019). Desert-like badlands and surrounding (semi-)dry grasslands of Central Germany promote small-scale phenotypic and genetic differentiation in *Thymus praecox*. *Ecology and Evolution*, *9*, 14066–14084. https://doi.org/10.1002/ece3.5844

Kirschner, J., Záveská Drábková, L., Štěpánek, J., & Uhlemann, I. (2015). Towards a better understanding of the *Taraxacum* evolution (Compositae–Cichorieae) on the basis of nrDNA of sexually reproducing species. *Plant Systematics and Evolution*, *301*, 1135–1156. https://doi.org/10.1007/s00606-014-1139-0

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, *35*, 4453–4555. https://doi.org/10.1093/bioinformatics/btz305

Landis, J. B., Soltis, D. E., Zheng, L., Marx, H. E., Barker, M. S., Tank, D. C., & Soltis, P. S. (2018). Impact of whole-genome duplication events on diversification rates in angiosperms. *American Journal of Botany*, *105*, 348–363. https://doi.org/10.1002/ajb2.1060

Lautenschlager, U., Wagner, F., & Oberprieler, C. (2020). AllCoPol: Inferring allele co-ancestry in polyploids. *BMC Bioinformatics*, *21*, 444. https://doi.org/10.1186/s12859-020-03750-9

Leebens-Mack, J. H., Barker, M. S., Carpenter, E. J., Deyholos, M. K., Gitzendanner, M. A., Graham, S. W., … Wong, G. K. S. (2019). One thousand plant transcriptomes and the phylogenomics of green plants. *Nature*, *574*, 679–685. https://doi.org/10.1038/s41586-019-1693-2

Lemoine, F., Domelevo Entfellner, J.-B., Wilkinson, E., Correia, D., Dávila Felipe, M., De Oliveira, T., & Gascuel, O. (2018). Renewing Felsenstein's phylogenetic bootstrap in the era of big data. *Nature*, *556*, 452–456. https://doi.org/10.1038/s41586-018-0043-0

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., & Homer, ... 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lo, E. Y. Y., Stefanovic, S., & Dickinson T. A. (2010). Reconstructing reticulation history in a phylogenetic framework and the potential of allopatric speciation driven by polyploidy in an agamic complex in *Crataegus* (Rosaceae). *Evolution*, *64*, 3593–3608. https://doi.org/10.1111/j.1558-5646.2010.01063.x

Mable, B. K., Alexandrou, M. A., & Taylor, M. I. (2011). Genome duplication in amphibians and fish: An extended synthesis. *Journal of Zoology*, *284*, 151–182.

https://doi.org/10.1111/j.1469-7998.2011.00829.x

Malinsky, M., Trucchi, E., Lawson, D. J., & Falush, D. (2018). RADpainter and fineRADstructure: Population inference from RADseq data. *Molecular Biology and Evolution*, *35*, 1284–1290. https://doi.org/10.1093/molbev/msy023

Marchant, D. B., Soltis, D. E., & Soltis, P. S. (2016). Patterns of abiotic niche shifts in allopolyploids relative to their progenitors. *New Phytologist*, *212*, 708–718. https://doi.org/10.1111/nph.14069

Mayrose, I., Zhan, S. H., Rothfels, C. J., Arrigo, N., Barker, M. S., Rieseberg, L. H., & Otto, S. P. (2015). Methods for studying polyploid diversification and the dead-end hypothesis: A reply to Soltis et al. (2014). *New Phytologist*, *206*, 27–35. https://doi.org/10.1111/nph.13192

McBreen, K., & Lockhart, P. J. (2006). Reconstructing reticulate evolutionary histories of plants. *Trends in Plant Science*, *11*, 398–404. https://doi.org/10.1016/j.tplants.2006.06.004

McDade, L. A. (1992). Hybrids and phylogenetic systematics. 1. The impact of hybrids on cladistic-analysis. *Evolution*, *46*, 1329–1346. https://doi.org/10.2307/2409940

McDade, L. A. (1995). *Hybridization and phylogenetics*. In: Hoch, P. C., Stephenson, A. (Eds.), Experimental and Molecular Approaches to Plant Biosystematics (pp. 305–331). St. Louis, USA: Missouri Botanical Garden.

McKain, M. R., Johnson, M. G., Uribe-Convers, S., Eaton, D., Yang, Y. (2018). Practical considerations for plant phylogenomics. *Applications in Plant Sciences*, *6*, 15. https://doi.org/10.1002/aps3.1038

Melicharkova, A., Slenker, M., Zozomova-Lihova, J., Skokanova, K., Singliarova, B., Kacmarova T., ... Marhold, K. (2020). So closely related and yet so different: Strong contrasts between the evolutionary histories of species of the *Cardamine pratensis* polyploid complex in Central Europe. *Frontiers in Plant Science*, *11*, 1988. https://doi.org/10.3389/fpls.2020.588856

Meudt, H. M., Albach, D. C., Tanentzap, A. J., Igea, J., Newmarch, S. C., Brandt, A. J., ... Tate, J. A. (2021). Polyploidy on islands: Its emergence and importance for diversification. *Frontiers in Plant Science*, *12*, 637214. https://doi.org/10.3389/fpls.2021.637214

Mohammadin, S., Wang, W., Liu, T., Moazzeni, H., Ertugrul, K., Uysal, T., … Schranz, M. E. (2018). Genome-wide nucleotide diversity and associations with geography, ploidy level and glucosinolate profiles in *Aethionema arabicum* (Brassicaceae). *Plant Systematics and Evolution*, *304*, 619–630. https://doi.org/10.1007/s00606-018-1494-3

Mráz, P., Filipaş, L., Bărbos, M. I., Kadlecová, J., Paštová, L., Belyayev, A., & Fehrer, J. (2019). An unexpected new diploid *Hieracium* from Europe: Integrative taxonomic approach with a phylogeny of diploid *Hieracium* taxa. *Taxon*, *68*, 1258–1277.

https://doi.org/10.1002/tax.12149

Múrias Dos Santos, A., Cabezas, M. P., Tavares, A. I., Xavier, R., & Branco, M. (2016). TcsBU: A tool to extend TCS network layout and visualization. *Bioinformatics*, *32*, 627–628. https://doi.org/10.1093/bioinformatics/btv636

Nardi, F. D., Dobes, C., Muller, D., Grasegger, T., Myllynen, T., Alonso-Marcos, H., Tribsch, A. (2018). Sexual intraspecific recombination but not de novo origin governs the genesis of new apomictic genotypes in *Potentilla puberula* (Rosaceae). *Taxon*, 67, 1108–1131. https://doi.org/10.12705/676.8

Near, T. J., MacGuigan, D. J., Parker, E., Struthers, C. D., Jones, C. D., & Dornburg, A. (2018). Phylogenetic analysis of Antarctic Notothenioids illuminates the utility of RADseq for resolving Cenozoic adaptive radiations. *Molecular Phylogenetics and Evolution*, *129*, 268–279. https://doi.org/10.1016/j.ympev.2018.09.001

Oberprieler, C., Wagner, F., Tomasello, S., & Konowalik, K. (2017). A permutation approach for inferring species networks from gene trees in polyploid complexes by minimising deep coalescences. *Methods in Ecology and Evolution*, *8*, 835–849. https://doi.org/10.1111/2041-210X.12694

Olave, M., & Meyer, A. (2020). Implementing large genomic single nucleotide polymorphism data sets in phylogenetic reconstructions: A case study of particularly rapid radiations of cichlid fish. *Systematic Biology*, *69*, 848–862. https://doi.org/10.1093/sysbio/syaa005

Otto, S. P., & Whitton, J. (2000). Polyploid incidence and evolution. *Annual Reviews of Genetics*, 34, 401–437. https://doi.org/10.1146/annurev.genet.34.1.401

Oxelman, B., Brysting, A. K., Jones, G. R., Marcussen, T., Oberprieler C., & Pfeil B. E. (2017). Phylogenetics of allopolyploids. *Annual Reviews of Ecology, Evolution, and Systematics*, *48*, 543–557. https://doi.org/10.1146/annurev-ecolsys-110316-022729

Paun, O., Stuessy, T. F., Hörandl, E. (2006). The role of hybridization, polyploidization and glaciation in the origin and evolution of the apomictic *Ranunculus cassubicus* complex. *New Phytologist*, *171*, 223–236. https://doi.org/10.1111/j.1469-8137.2006.01738.x

Paule, J., Dunkel, F. G., Schmidt, M., & Gregor, T. (2018). Climatic differentiation in polyploid apomictic *Ranunculus auricomus* complex in Europe. *BMC Ecology*, *18*, 16. https://doi.org/10.1186/s12898-018-0172-1

Pease, J. B., Brown, J. W., Walker, J. F., Hinchliff, C. E., & Smith, S. A. (2018). Quartet Sampling distinguishes lack of support from conflicting support in the green plant tree of life. *American Journal of Botany*, *105*, 385–403. https://doi.org/10.1002/ajb2.1016

Pellino, M., Hojsgaard, D., Schmutzer, T., Scholz, U., Hörandl, E., Vogel, H., & Sharbel, T. F. (2013). Asexual genome evolution in the apomictic *Ranunculus auricomus* complex: Examining the effects of hybridization and mutation accumulation. *Molecular Ecology*, *22*, 5908–5921. https://doi.org/10.1111/mec.12533

Peralta, M., Combes, M., Cenci, A., Lashermes, P., & Dereeper, A. (2013). SNiPloid: A utility to exploit high-throughput SNP data derived from RNA-Seq in allopolyploid species. *International Journal of Plant Genomics*, *2013*, 1–6. https://doi.org/10.1155/2013/890123

Philippe, H., & Forterre, P. (1999). The rooting of the universal tree of life is not reliable. *Journal of Molecular Evolution*, *49*, 509–523. https://doi.org/10.1007/pl00006573.

Qiu, T., Liu, Z., & Liu, B. (2020). The effects of hybridization and genome doubling in plant evolution via allopolyploidy. *Molecular Biology Reports*, *47*, 5549–5558. https://doi.org/10.1007/s11033-020-05597-y

R Core Team (2020). *R: A language and environment for statistical computing*. Retrieved from http://www.r-project.org/

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, *67*, 901–904. https://doi.org/10.1093/sysbio/syy032

Rannala, B. (2015). The art and science of species delimitation. *Current Zoology*, *61*, 846–853. https://doi.org/10.1093/czoolo/61.5.846

Rannala, B., & Yang, Z. (2003). Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics*, *164*, 1645–1656.

Ree, R. H., & Hipp, A. L. (2015). *Inferring phylogenetic history from restriction site associated DNA (RADseq)*. In: Hörandl E., Appelhans M. (Eds.) Next-generation sequencing in plant systematics. Bratislava, IAPT, pp. 181–204.

Rice, A., Smarda, P., Novosolov, M., Drori, M., Glick, L., Sabath, N., ... Mayrose, I. (2019). The global biogeography of polyploid plants. *Nature Ecology and Evolution*, *3*, 265–273. https://doi.org/10.1038/s41559-018-0787-9

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature Biotechnology*, *29*, 24–26. https://doi.org/10.1038/nbt.1754

Rothfels, C. J. (2021). Polyploid phylogenetics. *New Phytologist*, *230*, 66-72. https://doi.org/10.1111/nph.17105

Rothfels, C. J., Pryer, K., & Li, F.-W. (2017). Next-generation polyploid phylogenetics: rapid resolution of hybrid polyploid complexes using PacBio single-molecule sequencing. *New Phytologist*, *213*, 413–429. https://doi.org/10.1111/nph.14111

Sayyari, E., & Mirarab, S. (2016). Fast coalescent-based computation of local branch support from quartet frequencies. *Molecular Biology and Evolution*, *33*, 1654–1668. https://doi.org/10.1093/molbev/msw079

Schmickl, R., Liston, A., Zeisek, V., Oberlander, K., Weitemier, K., Straub, S. C. K., … Suda, J. (2016). Phylogenetic marker development for target enrichment from

transcriptome and genome skim data: The pipeline and its application in southern African *Oxalis* (Oxalidaceae). *Molecular Ecology Resources*, *16*, 1124–1135. https://doi.org/10.1111/1755-0998.12487

Schoenfelder, K. P., & Fox, D. T. (2015). The expanding implications of polyploidy. *Journal of Cell Biology*, *209*, 485–491. https://doi.org/10.1083/jcb.201502016

Shen, X.-X., Hittinger, C.T., & Rokas A. (2017). Contentious relationships in phylogenomic studies can be driven by a handful of genes. *Nature Ecology and Evolution*, *1*, 0126. https://doi.org/10.1038/s41559-017-0126

Šlenker, M., Kantor, A., Marhold, K., Schmickl, R., Mandáková, T., Lysak, M. A., ... Zozomová-Lihová, J. (2021). Allele sorting as a novel approach to resolving the origin of allotetraploids using Hyb-Seq data: A case study of the Balkan mountain endemic *Cardamine barbaraeoides*. *Frontiers in Plant Science*, *12*, 659275. https://doi.org/10.3389/fpls.2021.659275

Sochor, M., Vašut, R. J., Sharbel, T. F., & Trávníček, B. (2015). How just a few makes a lot: Speciation via reticulation and apomixis on example of European brambles (*Rubus* subgen. *Rubus*, Rosaceae). *Molecular Phylogenetics and Evolution*, *89*, 13–27. https://doi.org/10.1016/j.ympev.2015.04.007

Solís-Lemus, C., Bastide, P., & Ané, C. (2017). PhyloNetworks: A package for phylogenetic networks. *Molecular Biology and Evolution*, *34*, 3292–3298. https://doi.org/10.1093/molbev/msx235

Soltis, D. E., Gitzendanner, M. A., Stull, G., Chester, M., Chanderbali, A., Chamala S., ... Barbazuk W. B. (2013). The potential of genomics in plant systematics. *Taxon*, *62*, 886–898. https://doi.org/10.12705/625.13

Soltis, P. S., Marchant, D. B., Van de Peer, Y., & Soltis, D. E. (2015). Polyploidy and genome evolution in plants. *Current Opinion in Genetics and Development*, *35*, 119–125. https://doi.org/10.1016/j.gde.2015.11.003

Soltis, P. S., & Soltis, D. E. (2009). The role of hybridization in plant speciation. *Annual Review of Plant Biology*, *60*, 561–588. https://doi.org/10.1146/annurev.arplant.043008.092039

Soltis, P. S., & Soltis, D. E. (2016). Ancient WGD events as drivers of key innovations in angiosperms. *Current Opinion in Plant Biology*, *30*, 159–165. https://doi.org/10.1016/j.pbi.2016.03.015.

Spoelhof, J. P., Soltis, P. S., & Soltis, D. E. (2017). Pure polyploidy: Closing the gaps in autopolyploid research. *Journal of Systematics and Evolution*, *55*, 340–352. https://doi.org/10.1111/jse.12253

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30*, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033

Struck, T. H., Wey-Fabrizius, A. R., Golombek, A., Hering, L., Weigert, A., Bleidorn, C., … Hankeln, T. (2014). Platyzoan paraphyly based on phylogenomic data supports a noncoelomate ancestry of *Spiralia*. *Molecular Biology and Evolution*, *31*, 1833–1849. https://doi.org/10.1093/molbev/msu143

Stull, G. W., Soltis, P. S., Soltis, D. E., Gitzendanner, M. A., & Smith, S. A. (2020). Nuclear phylogenomic analyses of asterids conflict with plastome trees and support novel relationships among major lineages. *American Journal of Botany*, *107*, 790–805. https://doi.org/10.1002/ajb2.1468

Tiley, G. P., Crowl, A. A., Manos, P. S., Sessa, E. B., Solis-Lemus, C., Yoder, A. D., & Burleigh, J. G. (2021). Phasing alleles improves network inference with allopolyploids. *BioRxiv*, 1–41. https://doi.org/10.1101/2021.05.04.442457

Than, C., Ruths, D., & Nakhleh, L. (2008). PhyloNet: A software package for analyzing and reconstructing reticulate evolutionary relationships. *BMC Bioinformatics*, *9*, 322. https://doi.org/10.1186/1471-2105-9-322

Tomasello, S., Karbstein, K., Hodač, L., Pätzold, C., & Hörandl, E. (2020). Phylogenomics unravels Quaternary vicariance and allopatric speciation patterns in temperate-montane plant species: A case study on the *Ranunculus auricomus* species complex. *Molecular Ecology*, *29*, 2031–2049. https://doi.org/10.1111/mec.15458

Van de Peer, Y., Ashman, T.-L., Soltis, P. S., & Soltis, D. E. (2020). Polyploidy: an evolutionary and ecological force in stressful times. *The Plant Cell*, *0*, 1–16. https://doi.org/10.1093/plcell/koaa015

Van De Peer, Y., Mizrachi, E., & Marchal, K. (2017). The evolutionary significance of polyploidy. *Nature Reviews Genetics*, *18*, 411–424. https://doi.org/10.1038/nrg.2017.26

Wagner, F., Ott, T., Zimmer, C., Reichhart, V., Vogt, R., & Oberprieler, C. (2019). 'At the crossroads towards polyploidy': Genomic divergence and extent of homoploid hybridization are drivers for the formation of the ox-eye daisy polyploid complex (*Leucanthemum*, Compositae-Anthemideae). *New Phytologist*, *223*, 2039–2053. https://doi.org/10.1111/nph.15784

Wagner, N. D., He, L., & Hörandl, E. (2020). Phylogenomic relationships and evolution of polyploid *Salix* species revealed by RAD sequencing data. *Frontiers in Plant Science*, *11*, 36–41. https://doi.org/10.3389/fpls.2020.01077

Wagner, N. D., Clements, M. A., Simpson, L., & Nargar, K. (2021). Conservation in the face of hybridisation: genome-wide study to evaluate taxonomic delimitation and conservation status of a threatened orchid species. *Conservation Genetics*, *22*, 151–168. https://doi.org/10.1007/s10592-020-01325-y

Weisrock, D. W., Smith, S. D., Chan, L. M., Biebouw, K., Kappeler, P. M., & Yoder, A. D. (2012). Concatenation and concordance in the reconstruction of mouse lemur phylogeny: An empirical demonstration of the effect of allele sampling in phylogenetics. *Molecular Biology and Evolution*, *29*, 1615–1630.

https://doi.org/10.1093/molbev/mss008

Weitemier, K., Straub, S. C. K., Cronn, R. C., Fishbein, M., Schmickl, R., McDonnell, A., & Liston, A. (2014). Hybseq: Combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Sciences*, *2*, 1400042. https://doi.org/10.3732/apps.1400042

Welch, D. M., & Meselson, M. (2000). Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. *Science*, *288*, 1211–1215. https://doi.org/10.1126/science.288.5469.1211

Wen, D., Yu, Y., Zhu, J., & Nakhleh, L. (2018). Inferring phylogenetic networks using PhyloNet. *Systematic Biology*, *67*, 735–740. https://doi.org/10.1093/sysbio/syy015

Wendel, J. F. (2015). The wondrous cycles of polyploidy in plants. *American Journal of Botany*, *102*, 1753–1756. https://doi.org/10.3732/ajb.1500320

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., Rieseberg, L. H. (2009). The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Science USA*, *106*, 13875–13879. https://doi.org/10.1073/pnas.0811575106

Yan, Z., Cao, Z., Liu, Y., & Nakhleh, L. (2020). Maximum parsimony inference of phylogenetic networks in the presence of polyploid complexes. *BioRxiv*, 1–22. https://doi.org/10.1101/2020.09.28.317651

Yang, Z., & Rannala, B. (2014). Unguided species delimitation using DNA sequence data from multiple loci. *Molecular Biology and Evolution*, *31*, 3125–3135. https://doi.org/10.1093/molbev/msu279

Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, *19*, 153. https://doi.org/10.1186/s12859-018-2129-y

V Supplementaries

Chapter 1

Table S1. Description of the basic flow cytometric procedure. Steps mainly followed Klatt et al. (2016) and Barke et al. (2018).

Step	Description
1	We used Otto buffers to first isolate and then stain nuclei (Otto, 1990; Doležel & Bartoš, 2005; Doležel et al., 2007). The composition of buffers followed Otto (1990) with slight modifications: 100 ml Otto I buffer are composed of 99 ml ddH_2O, 2.1014 g citric acid, and 0.5 ml Tween 20, and 100 ml Otto II contained 99 ml ddH_2O, 5.6784 g dibasic sodium phosphate, and 1 ml (300 µg/ml) 4′,6-Diamidin-2-phenylindol (DAPI) solution.
2	We put 200 µl extraction buffer Otto I to the ground leaf material and immediately inverted samples for 30 s to isolate the nuclei from the plant cells.
3	Afterward, we filtered samples through CellTrics filters with a mesh of 30 µm into flow cytometric sample tubes.
4	We stained samples with 800 µl Otto II containing DAPI.

Table S2. Flow cytometry (FC) of leaves and single-seed flow cytometric seed screening (ssFCSS) of diploid and tetraploid populations within the *R. auricomus* complex. We only measured populations with yet unknown DNA ploidy and reproductive pathway. Peak indices are mean values (with minima and maxima in brackets). See Dunkel et al. (2018) for already published DNA ploidy and reproductive pathway measurements and references given in the manuscript. N_{ind} = number of individuals, N_{seeds} = number of seeds.

		FC		ssFCSS					
Pop. ID	**Taxon**	**N_{ind}**	**DNA ploidy**	**N_{ind}**	**N_{seeds}**	**Ploidy embryo**	**Ploidy endosperm**	**Peak index (endosperm : embryo)**	**Reproductive pathway**
LH012	*R. austroslovenicus*	11	2C	3	15	2C	3C	1.53 (1.47–1.62)	Sexual
Du-34889	*R. calapius*	1	2C						
Du-35351	*R. calapius*	1	2C						
LH040	*R. carpaticola*	11	2C	3	14	2C	3C	1.50 (1.38–1.62)	Sexual
LH006	*R. cassubicifolius*	15	2C	3	15	2C	3C	1.54 (1.38–1.67)	Sexual
LH007	*R. cassubicifolius*	13	2C	3	15	2C	3C	1.49 (1.43–1.55)	Sexual
LH016	*R. cassubicifolius*	9	2C	3	12 2	2C 2C	3C 6C	1.54 (1.38–1.78) 2.98 (2.86–3.10)	Sexual Asexual
Du-33354	*R. cebennensis*	1	2C						
EH10316	*R. envalirensis*	36	2C	3	15	2C	3C	1.50 (1.44–1.70)	Sexual
LH023	*R. flabellifolius*	6	2C	1	5	2C	3C	1.48 (1.36–1.53)	Sexual
LH025	*R. flabellifolius*	3	2C	2	6	2C	3C	1.52 (1.43–1.55)	Sexual
LH017	*R. marsicus*	9	4C, 6C	3	1 8	4C 6C	6C 18C	1.53 2.90 (2.71–3.15)	Sexual Asexual
LH018	*R. marsicus*	14	4C	3	23	2C	3C	1.47 (1.36–1.72)	Sexual
LH014	*R. mediocompositus*	9	2C	3	15	2C	3C	1.48 (1.44–1.53)	Sexual
LH015	*R. mediocompositus*	9	2C	3	15	2C	3C	1.49 (1.38–1.6)	Sexual
10137	*R. notabilis*	2	2C						
LH028	*R. notabilis*	6	2C						
LH010	*R. peracris*	11	2C	3	15	2C	3C	1.61 (1.45–1.83)	Sexual
LH011	*R. peracris*	15	2C	3	15	2C	3C	1.50 (1.44–1.54)	Sexual
LH013	*R. subcarniolicus*	9	2C	3	14	2C	3C	1.52 (1.42–1.64)	Sexual

Table S3. Quartet sampling (QS; Pease et al., 2018) results concerning the quartet fidelity (QF) score.

Taxon (sample ID)	QF
R. austroslovenicus Du-30441/01 SI	0.82
R. austroslovenicus Du-30442/01 SI	0.87
R. austroslovenicus LH012/01 SI	0.81
R. austroslovenicus LH012/08 SI	0.84
R. calapius Du/34889/01 HR	0.77
R. calapius Du/34889/11 HR	0.69
R. calapius Du/35351/15 HR	0.76
R. carpaticola 9126/02 RO	0.86
R. carpaticola 9126/06 RO	0.86
R. carpaticola LH040/04 SK	0.80
R. cassubicifolius Du-15980/01 IT	0.87
R. carpaticola Du-21047/01 HU	0.79
R. cassubicifolius Du-28673/01 SI	0.85
R. cassubicifolius LH006/05 DE	0.90
R. cassubicifolius LH006/17 DE	0.90
R. cassubicifolius LH007/06 AT	0.92
R. cassubicifolius LH007/11 AT	0.88
R. cassubicifolius LH016/01 SI	0.88
R. cassubicifolius LH016/14 SI	0.86
R. cebennensis Du-33354/02 FR	0.89
R. cebennensis Du-33354/21 FR	0.86
R. envalirensis 10316/14 AD	0.88
R. envalirensis Du-29983/01 AD	0.88
R. envalirensis Du-29988/01 FR	0.91
R. flabellifolius Du-25795 RO	0.80
R. flabellifolius LH023/04 RO	0.82
R. flabellifolius LH025/03 RO	0.80
R. marsicus Du-23722/01 IT	0.82
R. marsicus LH017/01 IT	0.68
R. marsicus LH018/02 IT	0.76
R. marsicus LH018/18 IT	0.71
R. mediocompositus LH014/03 SI	0.72

R. mediocompositus LH014 10 SI	0.76
R. mediocompositus LH015/03 SI	0.70
R. mediocompositus LH015 10 SI	0.72
R. notabilis 10137/3 AT	0.80
R. notabilis 10137/8 AT	0.82
R. notabilis LH028/02 AT	0.80
R. peracris LH010/01 SI	0.85
R. peracris LH010/09 SI	0.87
R. peracris LH011/01 SI	0.81
R. peracris LH011/14 SI	0.82
R. sceleratus 10426/1 DE	0.93
R. sceleratus 10426/2 DE	0.92

Table S4. Results of the Evanno test concerning the total dataset, the subset consisting of *R. carpaticola* and *R. cassubicifolius*, and the subset consisting of *R. flabellifolius*, *R. envalirensis*, …, and the Illyrian species. We highlighted the likeliest K value. See the documentation of STRUCTURE HARVESTER (Earl & Holdt, 2012) for details.

	K value	Replicates	Mean LnP(K)	Stdev LnP(K)	Ln′(K)	|Ln″(K)|	Delta K
Total dataset	1	10	−780507.69	487.17	NA	NA	NA
	2	**10**	**−667325.97**	**573.46**	**113181.72**	**1481068.94**	**2582.68**
	3	10	−2035213.19	4311638.14	−1367887.22	5198509.93	1.21
	4	10	−8601610.34	16689733.94	−6566397.15	99773314.35	5.98
	5	9	−114941321.8	106168812.4	−106339711.5	75751986.4	0.71
	6	9	−145529046.9	190254761.3	−30587725.1	137716322.8	0.72
	7	9	−313833094.8	306028640.8	−168304047.9	221593315.9	0.72
	8	9	−260543826.8	117831677.5	53289268.04	108450238.5	0.92
	9	8	−315704797.2	229997413.9	−55160970.43	157638700.7	0.69
	10	8	−528504468.4	484253936.3	−212799671.2	258841562.8	0.53
	11	5	−482462576.7	339927580.2	46041891.64	386930428.7	1.14
	12	5	−823351113.8	572349546.3	−340888537.1	NA	NA
Subset (*carp.*/cass.)	1	10	−302819.26	213	NA	NA	NA
	2	**10**	**−4130631.99**	**1809055.97**	**−3827812.73**	**44595959.68**	**24.65**
	3	10	−52554404.4	33638282.3	−48423772.41	19747862.78	0.59
	4	10	−120726039.6	67724624.68	−68171635.19	18384174.67	0.27
	5	10	−207281849.5	128589521.6	−86555809.86	191557765.2	1.49
	6	10	−485395424.5	182136946.1	−278113575	216404468.8	1.19
	7	6	−547104530.7	370028569.7	−61709106.19	NA	NA
Subset (*flab.*/.../*per.*)	1	10	−672385.94	274.02	NA	NA	NA
	2	10	−1531770.84	1810423.94	−859384.9	1038639.82	0.57
	3	**10**	**−1352515.92**	**1194659.54**	**179254.92**	**35218205.49**	**29.48**
	4	10	−36391466.49	44189080.16	−35038950.57	105634921.9	2.39
	5	10	−177065338.9	78338755.9	−140673872.4	237468525.6	3.03
	6	10	−555207736.9	237071810.9	−378142398	239094479.8	1.01
	7	10	−694255655.1	456645146.8	−139047918.2	406841039.9	0.89
	8	10	−1240144613	507374741.5	−545888958.2	296794947.5	0.58
	9	10	−1489238624	596294321.3	−249094010.6	474518235.4	0.8
	10	5	−2212850870	741454925.4	−723612246	NA	NA

Table S5. Herbarium specimens or data from the literature (in PDF format) incorporated in the geometric morphometric analyses. If a herbarium specimen or data from the literature (PDF) from the locus classicus were included, we marked them with "x". A particular morphological trait is also marked with "x" if it was included in data analysis. See also Table 1 for most location details.

Herbarium specimen / PDF	Herbarium ID	Publication	Taxon	Locus classicus	Basal leave(s)	Stem leave(s)	Recep-tacle(s)
PDF	Du-25795	(Dunkel et al., 2018)	*R. flabellifolius*	x	x	x	
Herbarium specimen	Du-30441		*R. austroslovenicus*	x	x	x	x
Herbarium specimen / PDF	Du-28639	(Dunkel et al., 2018)	*R. mediocompositus*		x	x	x
PDF	Du-29983	(Dunkel et al., 2018)	*R. envalirensis*		x	x	
Herbarium specimen	GOET019892		*R. envalirensis*		x	x	x
Herbarium specimen	M-0008124		*R. envalirensis*		x	x	
Herbarium specimen	M-0008125		*R. envalirensis*		x	x	
Herbarium specimen / PDF	Du-30446	(Dunkel et al., 2018)	*R. peracris*		x	x	x
Herbarium specimen / PDF	Du-33354	(Dunkel et al., 2018)	*R. cebennensis*	x	x	x	x
Herbarium specimen	Du-34772		*R. subcarniolicus*		x	x	x
PDF	Du-34773	(Dunkel et al., 2018)	*R. subcarniolicus*		x	x	
Herbarium specimen / PDF	Du-34889	(Dunkel et al., 2018)	*R. calapius*	x	x	x	x
Herbarium specimen	Du-35342		*R. calapius*		**x**	**x**	**x**
Herbarium specimen	Du-35351		*R. calapius*		**x**	**x**	**x**

Table S6. Test statistics for (**A**) described taxa and (**B**) accepted species regarding mean trichome density (per 0.25 mm² transect) of receptacles. Mean trichome density per 0.25 mm² transect of taxa and pairwise *p* values based on pairwise comparisons of the Wilcoxon rank-sum test with Holm correction are given between taxa.

A

	R. carpaticola	***R. cassubicifolius***	***R. flabellifolius***	***R. envalirensis***	***R. marsicus***	***R. mediocompositus***	***R. austrosloveniсus***	***R. subcarniolicus***	***R. notabilis***	***R. peracris***
R. cassubicifolius	0.18									
R. flabellifolius	0.01	< 0.001								
R. envalirensis	< 0.001	< 0.001	0.90							
R. marsicus	0.11	< 0.01	1.00	0.89						
R. mediocompositus	<0.01	< 0.001	1.00	1.00	1.00					
R. austrosloveniсus	< 0.01	< 0.001	1.00	1.00	1.00	1.00				
R. subcarniolicus	0.05	< 0.001	1.00	1.00	1.00	1.00	1.00			
R. notabilis	1.00	0.90	< 0.01	< 0.001	0.06	< 0.001	< 0.001	0.02		
R. peracris	< 0.01	< 0.001	0.42	1.00	0.47	1.0	1.00	1.00	< 0.01	
mean	8.14	12.67	0.22	1.01	0.00	0.67	1.22	0.71	9.50	1.27

B

	***R. cassubicifolius* s.l.**	***R. flabellifolius***	***R. envalirensis* s.str.**	***R. marsicus***	***R. notabilis* s.l.**
R. flabellifolius	< 0.001				
***R. envalirensis* s.str.**	< 0.001	0.17			
R. marsicus	< 0.01	0.55	0.17		
***R. notabilis* s.l.**	< 0.001	0.10	0.17	0.14	
mean	12.12	0.22	1.01	0.00	2.25

Appendix S1 Target enrichment lab workflow and initial bioinformatic analyses.

(A) Library preparation and hybrid capture protocol for target enrichment.

Genomic DNA was extracted from ~1.5 cm² leaf material of silica-dried samples or herbarium collection. For the scope, we used the Qiagen DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions, except for the sample incubation time in the lysis buffer that was increased to one hour. DNA quality and fragments length were checked by gel electrophoresis in a 1.5 Agarose gel and using the Roti-Load DNAstain 3 (Carl Roth, Karlsruhe, Germany), particularly for herbarium specimens. Extract concentration was estimated using the Qubit fluorometer and the Qubit dsDNA HS Assay Kit (ThermoFisher Scientific, Waltham, Massachusetts, U.S.A.). Sequencing libraries were prepared using either the NEBNext Ultra II DNA Library Prep Kit for Illumina (E7645) or the NEBNext Ultra II FS DNA Library Prep Kit for Illumina (E7805) (New England BioLabs, Ipswich, Massachusetts, U.S.A.). In the former case, we sheared DNA with a Bioruptor Pico (Diagenode, Seraing, Belgium) before library preparation. Extracts were diluted to 10 ng/µl and sonicated for eight cycles of 15 s sonication and 90 s break in order to obtain fragments of approximately 300–500 bp. In the latter case, enzymatic shearing is combined with the first steps of the library preparation. Fragmentation was carried out for 12 min at 37°C in order to obtain DNA fragments of the same length of the sonicated samples. For the herbarium collections Hoe5615 and Du33351-15, shearing incubation was shorter, 3 and 10 min, respectively. In both cases, we followed the manufacturer's instructions. At the end of the library preparation procedure, samples were PCR amplified for 14 cycles during which sample-specific dual indices (NEBNext Multiplex Oligos for Illumina, E7600; New England BioLabs) were added to the fragments. Indexed samples were pooled in equal quantities (four samples per 500 ng pool), dehydrated in a Concentrator Plus (Eppendorf, Hamburg, Germany), and diluted in 7 µl of ddH_2O. Each pool was enriched using the custom baits kit following the manufacturer's protocol. Hybridization took place for 21 h at 65°C.

Enriched products were PCR-amplified for 14 cycles using the 2X KAPA HiFi HotStart Mix (KAPA Biosystems, Wilmington, Massachusetts, U.S.A.) and the P7 and P5 adapters as primers. Amplified enriched libraries were purified with 50 µl of AMPure XP Beads (New England BioLabs) following the manufacturer's protocol. Concentrations were measured with the Qubit fluorometer and fragment length distributions were checked with a Bioanalyzer (Agilent, Santa Clara, California, U.S.A.). In the few cases in which fragment length was different from the desired one (and especially in cases short fragments were present), pools

were undergone to size selection with the BluePippin (Sage Science, Beverly, Massachusetts, U.S.A.).

Sequencing took place on an Illumina MiSeq System (Illumina, San Diego, California, U.S.A.) at the Transcriptome and Genome Analysis Laboratory (Georg-August-Universität, Göttingen, Germany). Pools were mixed equimolarly and sequenced in two different paired-end runs (6 pools, 24 samples each) with a 2 × 250 bp (500 cycles) v2 kit.

(B) Read processing and alignments

We checked the quality of raw reads with FASTQC v.0.11.8 (Andrews, 2010; available at: http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc). Further processing of the raw reads was done using the pipeline HYBPHYLOMAKER (all scripts available at: https://github.com/tomas-fer/HybPhyloMaker/; Fér & Schmickl, 2018). This pipeline offers a set of bash scripts assisting the procedure of assembly of captured sequences from read quality-trimming to the reconstruction of phylogenetic trees. In HYBPHYLOMAKER, quality-trimmed individual raw reads are mapped to a reference sequence and then merged into contigs that are aligned for each gene separately. As pseudo-reference for read mapping, we used a sequence consisting of the concatenation of the target exonic sequences separated by stretches of 800 Ns. Sequence adapters were removed and reads were quality-trimmed using TRIMMOMATIC v.0.32 (Bolger et al., 2014) with the default settings used in HYBPHYLOMAKER. Duplicated reads were removed with FASTUNIQ v.1.1 (Xu et al., 2012). Mapping to the pseudo-reference genome was done with BWA v.0.7.16 (Li & Durbin, 2010). In order to avoid the loss of allelic information during the process of allele mapping and consensus sequence production, we took out of the HYBPHYLOMAKER pipeline the *.bam files produced after mapping and phased them with SAMTOOLS v.0.1.19 (Li et al., 2009) using the commands “samtools calmd” and “samtools phase”. The respective *.bai files were also duplicated and named consequently. The phased *.bam and *.bai files were then placed back in a the HYBPHYLOMAKER working directory for further processing within the pipeline workflow. A new sample file with duplicate sample names was placed in the “/10rawreads/”. The pipeline was therefore resumed for the computing of the consensus sequences (allele-wise consensus sequences) by calling the script “HybPhyloMaker2_readmapping.sh” but specifying “mapping = no” in the setting file. Consensus sequences were produced with CONSENSUSFIXER (available at: https://github.com/cbg-ethz/ConsensusFixer) since this is the only of the approaches available in HYBPHYLOMAKER able to call ambiguity DNA codes in case of multiple bases per site in the mapped reads. For CONSENSUSFIXER, we used

the following settings: Minimum relative abundance of the alternative base ("plurality" in the setting file of HYBPHYLOMAKER) of 0.2 and a minimum read coverage for ambiguity calling ("mincov") of 5.

Consensus sequences were matched to sequences of the target exons to produce PSLX files using BLAT v.35.1 (Kent, 2002). Afterward, they were combined to produce exon-wise matrices with "assembled_exons_to_fastas.py" (Weitemier et al., 2014). These were aligned with MAFFT v.7.029 (Katoh & Standley, 2013) using the default program settings. HYBPHYLOMAKER performs two consecutive steps to check the alignments for missing data and filter out those that exceed certain levels. First, sequences with more than a certain percentage of Ns in an alignment ("MISSINGPERCENT" in the settings file) are deleted. We set this option to 40. Secondly, alignments with less than a certain percent of sequences ("SPECIESPRESENCE" in the settings file) are filtered out. We set this value to 75 so that alignments with more than 25% of missing sequences were excluded.

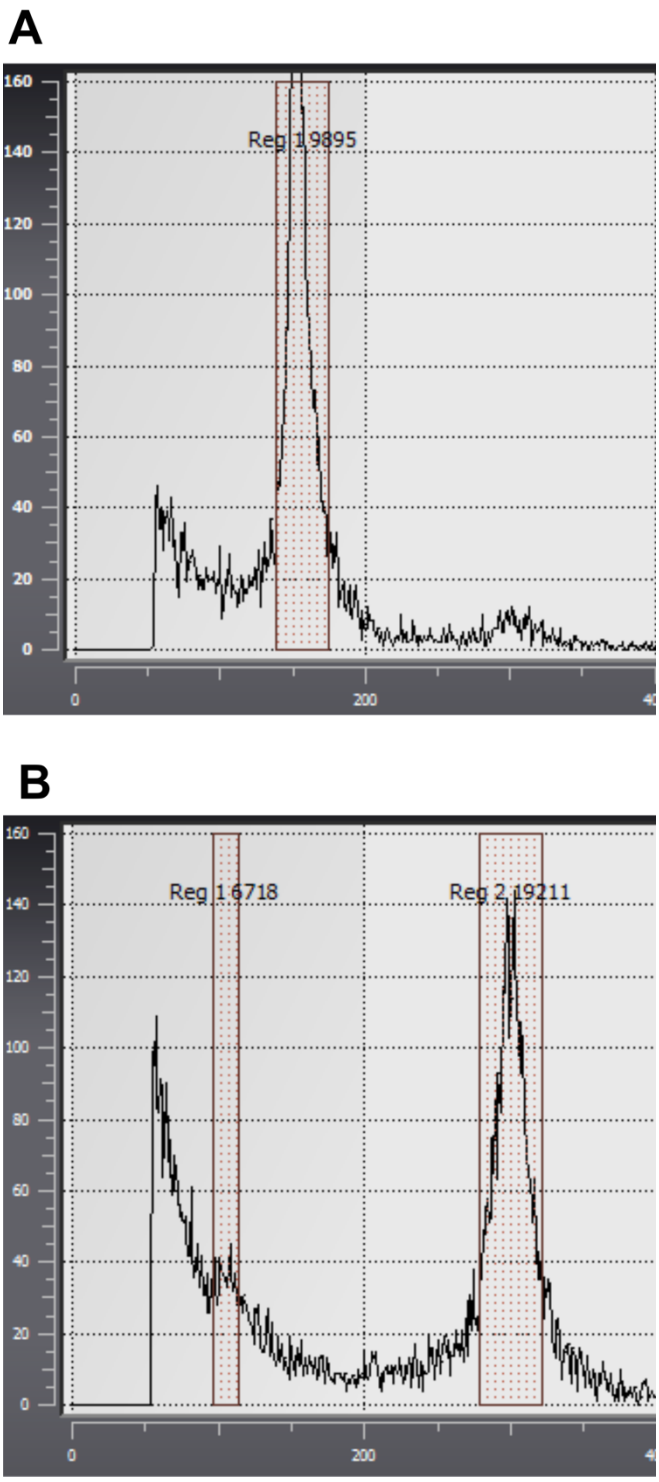


Figure S1. A, Exemplary single-seed flow cytometric seed screening (ssFCSS) measurement of seed "LH007_08_seed2" (*Ranunculus cassubicifolius* s.l.). We detected a clearly visible endosperm peak but no embryo peak. This was the case in ca. 30% of our ssFCSS measurements (see also section Flow cytometry and single seed flow cytometric seed screening in Materials and Methods). **B**, Exemplary ssFCSS measurement of an asexually formed seed (LH016_07_seed1) with embryo and endosperm peak, by a diploid mother plant (*R. cassubicifolius* s.l.).

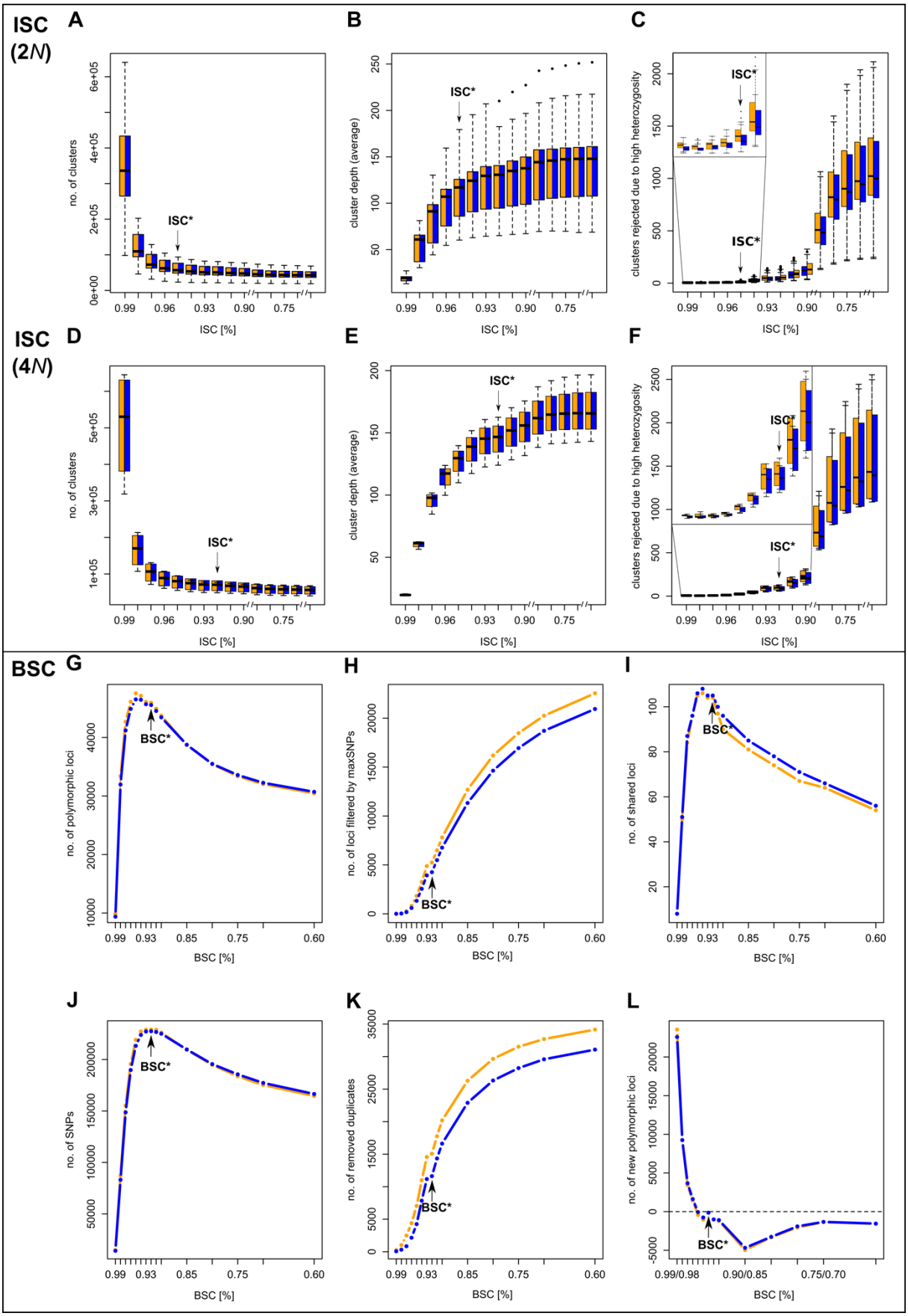


Figure S2. Evaluation of in-sample clustering threshold (ISCT) concerning diploid and tetraploid (and one hexaploid) sexual individuals within the *R. auricomus* complex, and between-sample clustering threshold (BSCT) concerning the total dataset. We illustrated

boxplots of (**A**, **D**) number of clusters, (**B**, **E**) cluster depth (average), and (**C**, **F**) clusters rejected due to high heterozygosity in relation to ISCT. Due to huge differences within clusters rejected due to high heterozygosity across ISCTs, we added an extended plot of ISCTs ranging from 94% to 99%. We illustrated boxplots of (**G**) number of polymorphic loci, (**H**) no. of loci filtered by maxSNPs, (**I**) no. of shared loci across samples no. of loci filtered by maxSNPs, (**J**) no. of SNPs, (**K**) no. of removed duplicates and (**L**) no. of new polymorphic loci in relation to BSCT. Orange boxplots represent "mindepth6" setting, and blue boxplots refer to "mindepth12" setting. ISCT* = selected ISCT for cluster optimization.

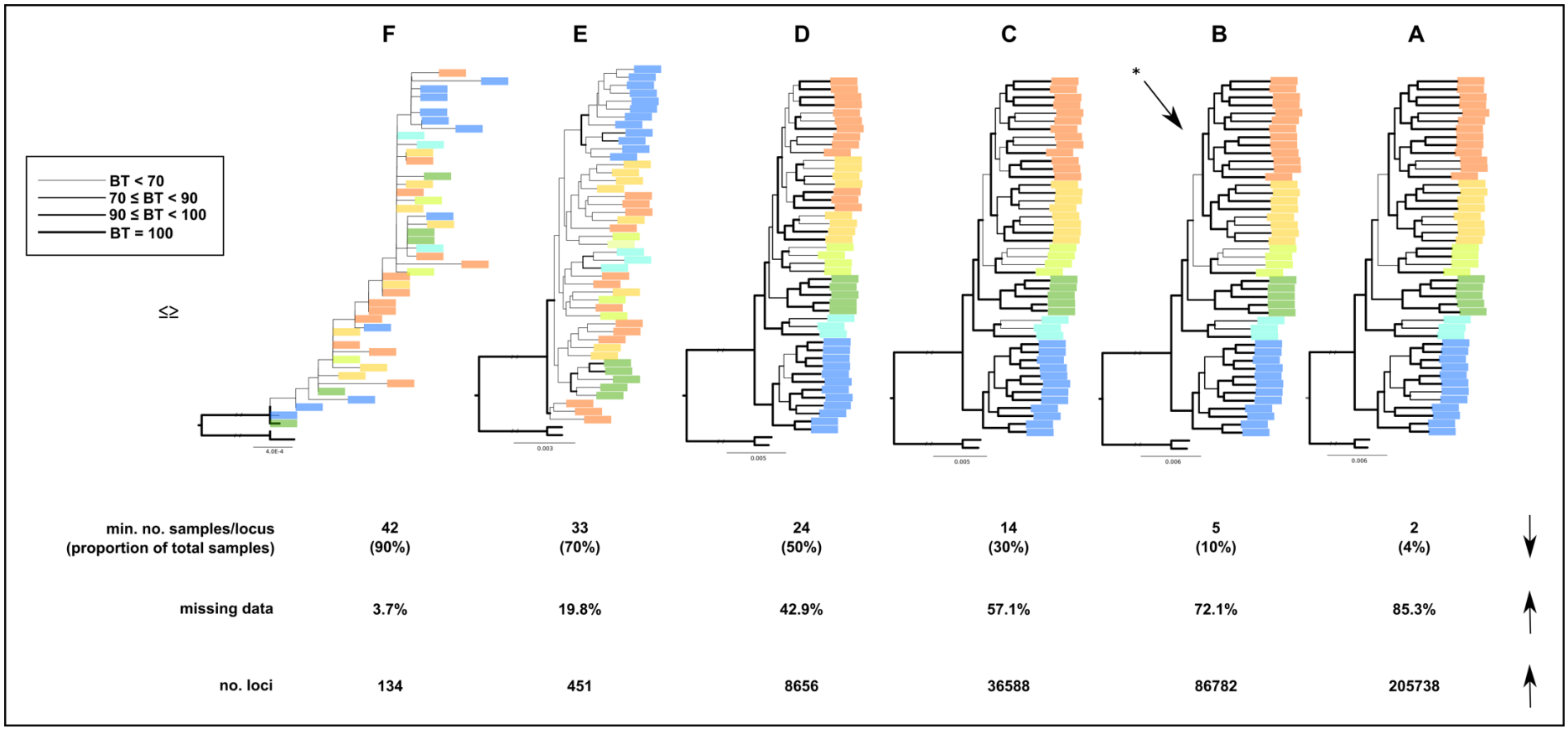


Figure S3. Maximum likelihood (ML) trees of different minimum number of samples per locus (2, 4%; 5, 10%; 14, 30%; 24, 50%; 33, 70%; and 42, 90%) (**A–F**). Line strengths represent the magnitude of the bootstrap value concerning the previous node (see legend within the figure). Taxa are colored according to accepted species (see Fig. 3). Below the ML trees, we put details on minimum number of samples per locus (and proportion of total samples), percent of missing data, and number of loci. Arrows specify the trend (either positive or negative). * (B): selected alignment for further data analyses (see Materials and Methods).

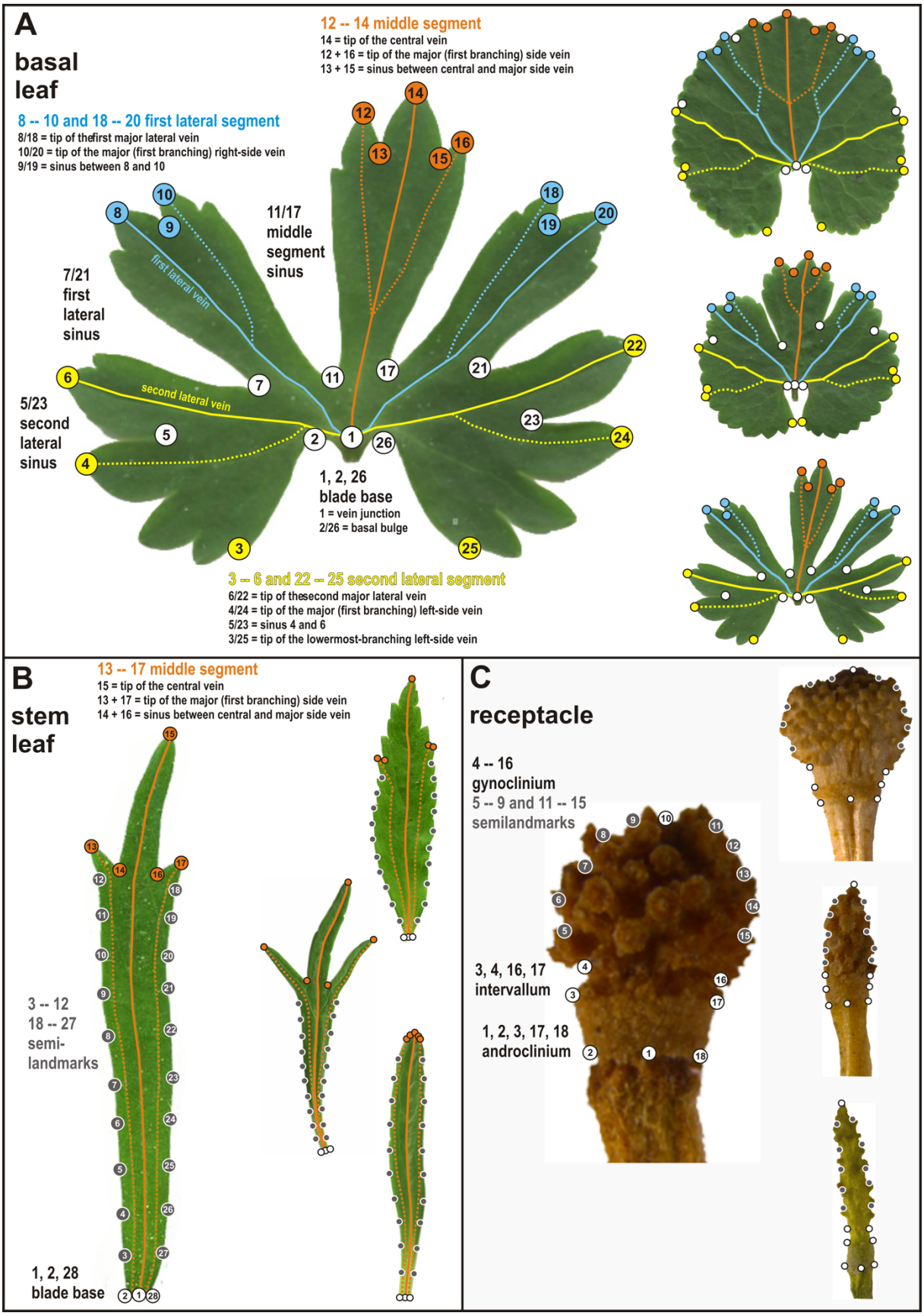


Figure S4. Landmark digitization. **A**, Delimitation of 26 landmarks (yellow, white, blue, and orange dots) on basal-leaf outline following homologous leaf venation patterns, segments, and sinuses. **B**, Delimitation of 8 landmarks (white and orange dots) and 20 semilandmarks (grey dots) on the stem leaf outline. **C**, Delimitation of 8 landmarks (white dots) and 10

semilandmarks (grey dots) on the receptacle outline. Images on the right side of each figure illustrate the morphological variability of species within the *R. auricomus* complex. The symmetrization of the object's left and right halves happens in three steps: (1) Exchange of left vs. right object halves, (2) Procrustes superimposition (Zelditch et al., 2012) of original and mirrored objects, (3) averaging of original and corresponding mirrored objects. In the first step, we put original landmark configurations together with their mirrored counterparts into one dataset. The mirrored landmark configurations was acquired by exchanging x- and y-coordinates within each pair of symmetrically corresponding landmarks. In basal leaves, for example, we reciprocally exchanged the positions of landmarks 2 and 26 with respect to the axis of bilateral symmetry as defined by landmarks 1 and 14. In the final mirrored configurations (after exchanging all corresponding landmarks), we have multiplied all x-coordinates by −1. In the next step, both original and mirrored configurations were subjected to the Procrustes superimposition (Zelditch et al., 2012). Procrustes fit method translates centroids of all configurations to the same point [0, 0] and rotates the configurations so that distances among homologous landmarks are minimized, and rescales the configurations to the unit size. Images consisting of landmark configurations without semilandmarks (i.e., basal leaves) were superimposed in the program MORPHOJ v.1.06d (Klingenberg, 2011), and those with semilandmarks (i.e., stem leaves and receptacles) in TPSRELW v.1.70 (Rohlf, 2015) allowing semilandmarks (without fixed position) to slide along curves between adjacent (semi-)landmarks. Superimposed landmark configurations, i.e., original configurations and their corresponding mirrored counterparts, were subsequently averaged resulting in configurations perfectly symmetric with respect to the axis of object bilateral symmetry. Symmetrized landmark configurations representing original leaf and receptacle shapes were averaged over each plant using the software MORPHOJ v.1.06d (Klingenberg, 2011).

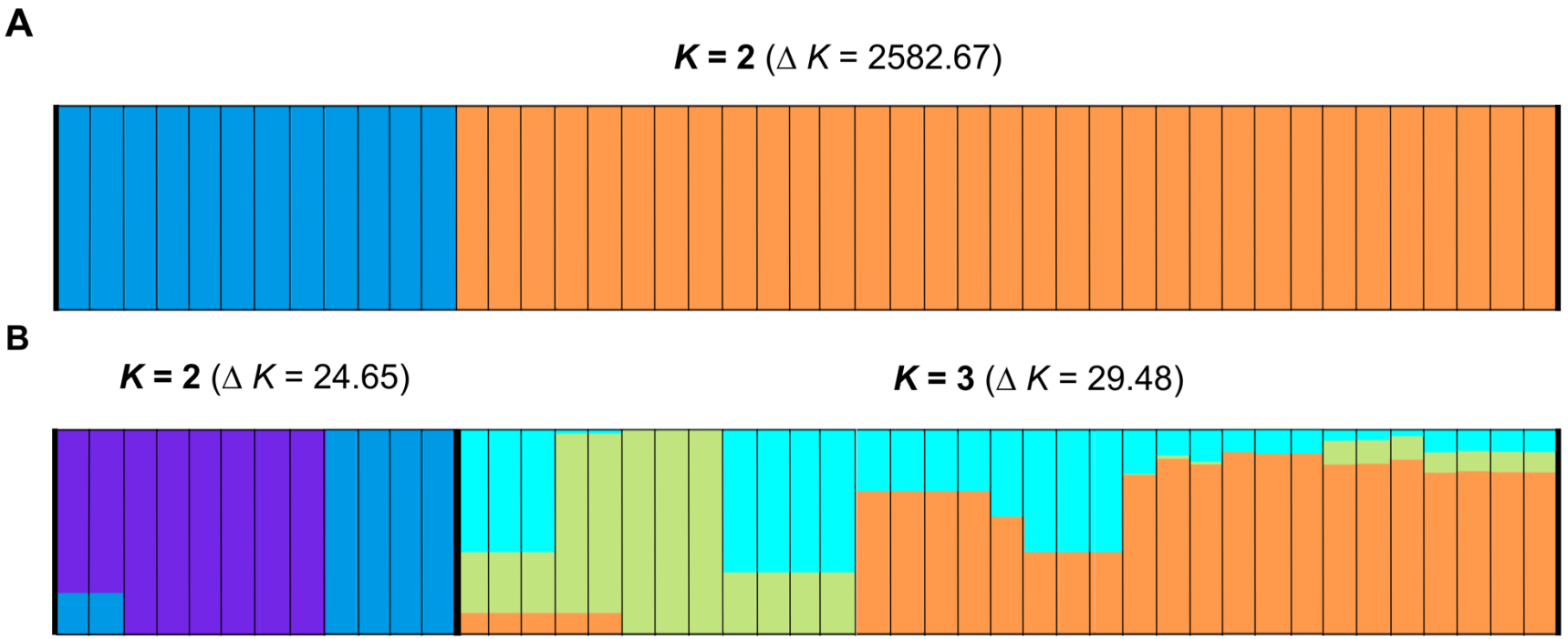


Figure S5. Bar graphs of STRUCTURE analyses are drawn for (**A**) the total dataset and (**B**) for each of the two subsets (see Table S3 for results of the Evanno tests). We chose the dataset with the likeliest *K* value.

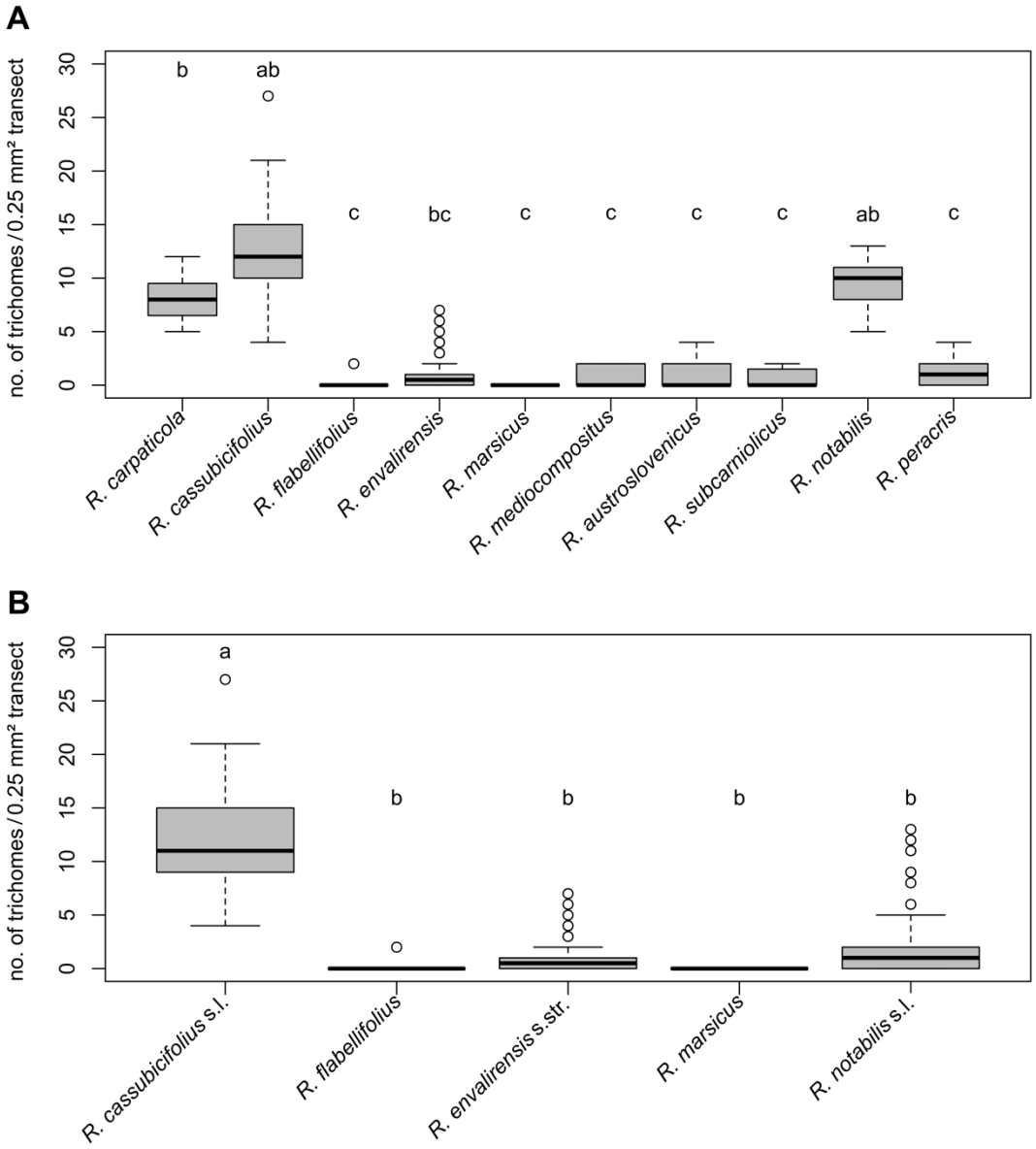


Figure S6. Boxplots of receptacle trichome density for (**A**) all described taxa and (**B**) all accepted taxa. The number of trichomes on a 0.25 mm² transect. In total, we assessed 219 receptacles. Significant differences in trichome density (**A**) among species (among all Chi^2 = 151.69, df = 9, $p < .001$) and (**B**) accepted species among all (Chi^2 = 129.97, df = 4, $p < .001$). Letters a–c (see Table S6) indicate pairwise differences.

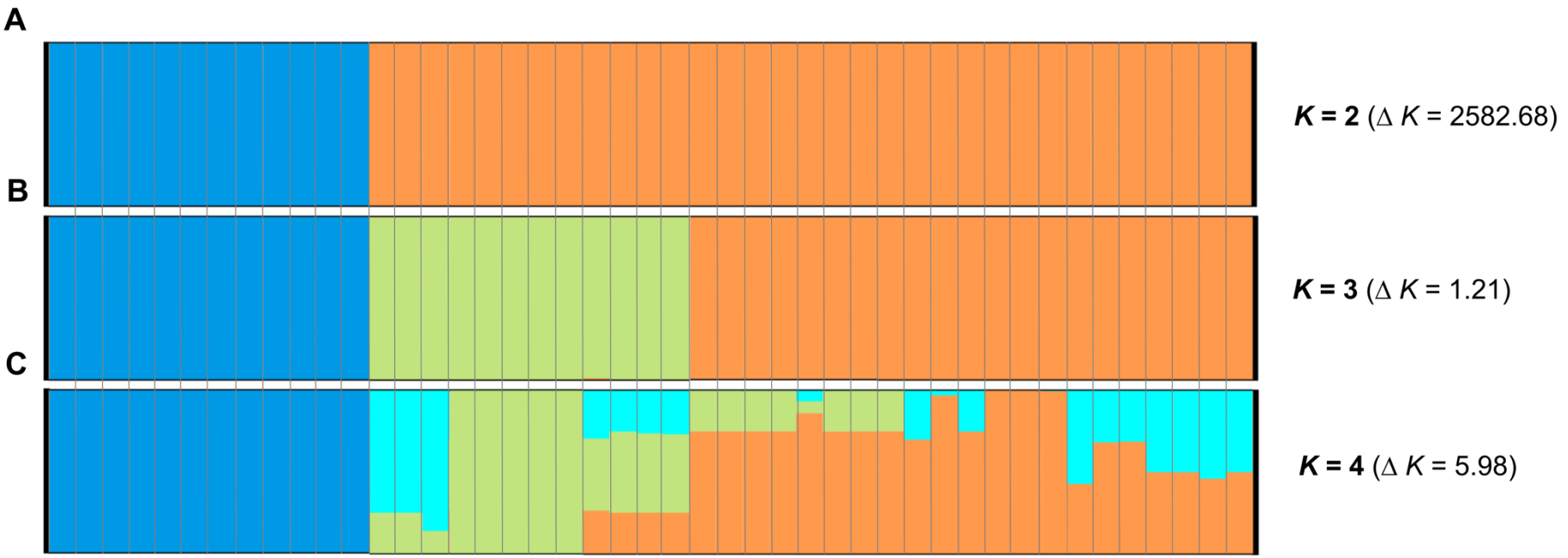


Figure S7. Bar graphs of STRUCTURE analyses are drawn for the total dataset and K = 2 to K = 4 (**A–C**; see Table S3 for results of the Evanno tests). We illustrated the dataset with the likeliest K value (*K* = 2), and also the datasets K = 3 and K = 4.

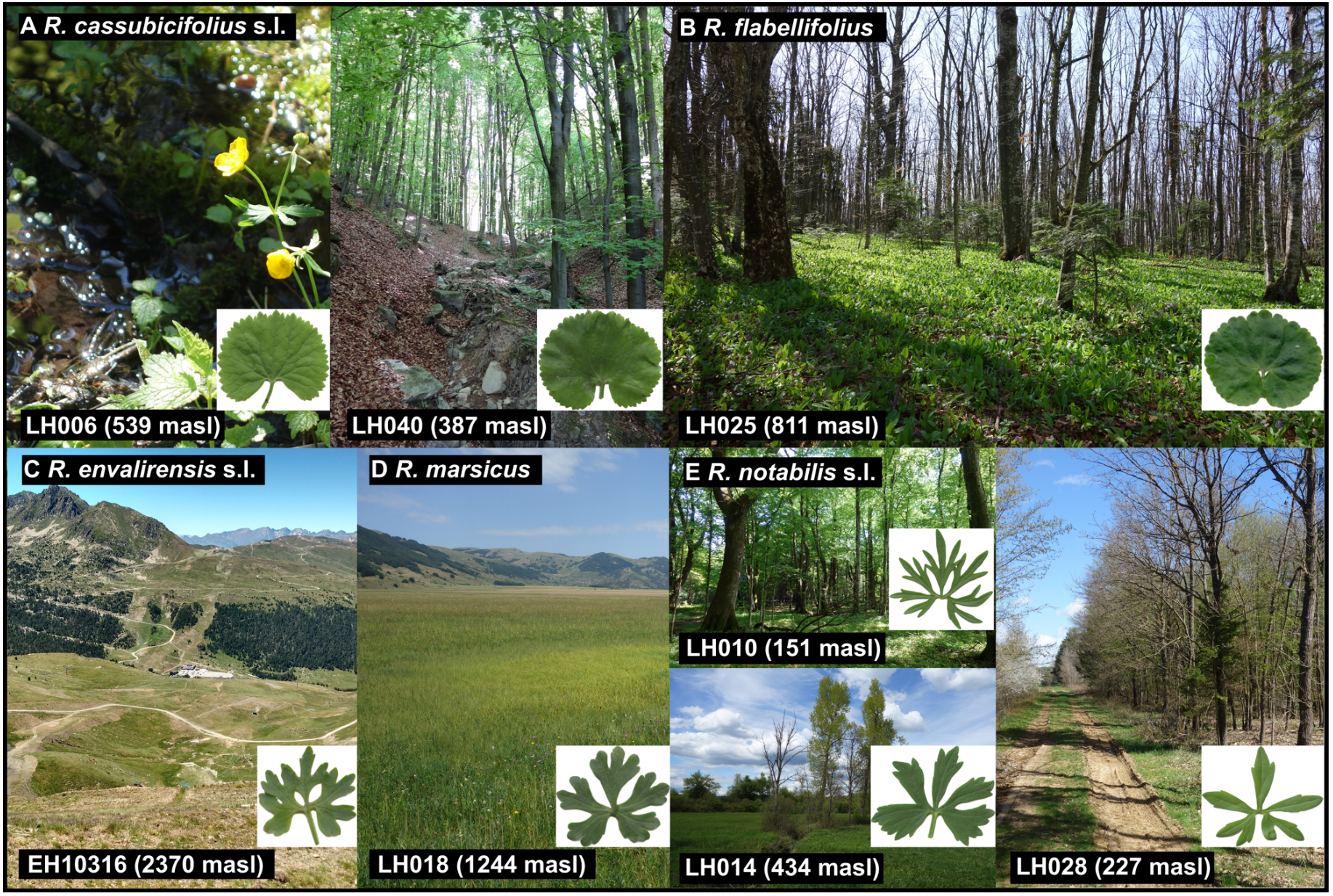

A R. cassubicifolius s.l.
LH006 (539 masl)
LH040 (387 masl)
B R. flabellifolius
LH025 (811 masl)
C R. envalirensis s.l.
EH10316 (2370 masl)
D R. marsicus
LH018 (1244 masl)
E R. notabilis s.l.
LH010 (151 masl)
LH014 (434 masl)
LH028 (227 masl)

Figure S8. Habitats, and basal spring leaf of accepted sexual species within the *R. auricomus* complex (former names in brackets): **A**, *R. cassubicifolius* s.l. inhabits humid beech, hornbeam forests, and streamsides (LH006 = *R. cassubicifolius*, LH040 = *R. carpaticola*); **B**, *R. flabellifolius* (LH025 = *R. flabellifolius*) was found in forests in a restricted area of the South Carpathians; **C**, *R. envalirensis* s.l. (EH10316 = *R. envalirensis*) inhabits (subalpine) meadows in the Massif Central and the Pyrennees; **D**, *R. marsicus* (LH018 = *R. marsicus*) occupies humid subalpine meadows (3-lobed leaf); **E**, *R. notabilis* s.l. (LH010 = *R. peracris*, LH014 = *R. mediocompositus*, and LH028 = *R. notabilis*; *R. austroslovenicus*, *R. subcarniolicus*, and *R. calapius* not illustrated) occurs in deciduous forest, forest and waysides, and humid to marshy meadows. Usually, five corolla leaves characterize the sexual species. We illustrated the most dissected-leaf morphotypes (taxonomically most informative). Image sources: A, B, D, and E (Ladislav Hodač); C (Elvira Hörandl).

References

Andrews, S. (2010). *FastQC: A quality control tool for high throughput sequence data*. Cambridge, UK: Babraham Bioinformatics, Babraham Institute. Retrieved from https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Barke, B. H., Daubert, M. & Hörandl, E. (2018). Establishment of apomixis in diploid F_2 hybrids and inheritance of apospory from F_1 to F_2 hybrids of the *Ranunculus auricomus* complex. *Frontiers in Plant Science*, *9*, 1111. https://doi.org/10.3389/fpls.2018.01111

Bolger, A. M., Lohse, M. & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics, 30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Doležel, J. & Bartoš, J. (2005). Plant DNA flow cytometry and estimation of nuclear genome size. *Annals of Botany*, *95*, 99–110. https://doi.org/10.1093/aob/mci005

Doležel, J., Greilhuber, J. & Suda, J. (2007). Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols*, *2*, 2233–2244. https://doi.org/10.1038/nprot.2007.310

Dunkel, F. G., Gregor, T. & Paule, J. (2018). New diploid species in the *Ranunculus auricomus* complex (Ranunculaceae) from W and SE Europe. *Willdenowia*, 48, 227–257. https://doi.org/10.3372/wi.48.48205

Earl, D. A. & Holdt, B. M. von (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetic Resources*, *4*, 359–361. https://doi.org/10.1007/s12686-011-9548-7

Fér, T. & Schmickl, R. E. (2018). HybPhyloMaker: Target enrichment data analysis from raw reads to species trees. *Evolutionary Bioinformatics*, *14*, 1–9. https://doi.org/10.1177/1176934317742613

Katoh, K. & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780. https://doi.org/10.1093/molbev/mst010

Kent, W. J. (2002). BLAT - the BLAST-like alignment tool. *Genome Resources*, *12*, 656–664. https://doi.org/10.1101/gr.229202

Klatt, S., Hadacek, F., Hodač, L., Brinkmann, G., Eilerts, M., Hojsgaard, D. & Hörandl, E. (2016). Photoperiod extension enhances sexual megaspore formation and triggers metabolic reprogramming in facultative apomictic *Ranunculus auricomus*. *Frontiers in Plant Science*, *7*, 278. https://doi.org/10.3389/fpls.2016.00278

Klingenberg, C. P. (2011). MorphoJ: An integrated software package for geometric morphometrics. *Molecular Ecology Resources*, *11*, 353–357. https://doi.org/10.1111/j.1755-0998.2010.02924.x

Li, H. & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, *26*, 589–595. https://doi.org/10.1093/bioinformatics/btp698

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Otto, F. (1990). DAPI staining of fixed cells for high-resolution flow cytometry of nuclear DNA. In Darzynkiewicz, Z. & Crissman, H.A. (Eds.), *Methods in Cell Biology* (pp. 105-110), vol. 33. San Diego, USA: Academic Press.

Pease, J. B., Brown, J. W., Walker, J. F., Hinchliff, C. E. & Smith, S. A. (2018). Quartet Sampling distinguishes lack of support from conflicting support in the green plant tree of life. *American Journal of Botany*, *105*, 385–403. https://doi.org/10.1002/ajb2.1016

Rohlf, F. J. (2015). The tps series of software. *Hystrix*, *26*, 9–12. https://doi.org/10.4404/hystrix-26.1-11264

Weitemier, K., Straub, S. C. K., Cronn, R. C., Fishbein, M., Schmickl, R., McDonnell, A. & Liston, A. (2014). Hybseq: Combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Science*, *2*, 1400042. https://doi.org/10.3732/apps.1400042

Xu, H., Luo, X., Qian, J., Pang, X., Song, J., Qian, G., ... Chen, S. (2012). FastUniq: A fast de novo duplicates removal tool for paired short reads. *PLoS ONE*, *7*, e52249. https://doi.org/10.1371/journal.pone.0052249

Zelditch, M. L., Swiderski, D. L., Sheets, H. D. & Fink, W. L. (2012). *Geometric morphometrics for biologists: A primer*, vol. 95. London: Academic Press. https://doi.org/10.1016/B978-0-12-778460-1.X5000-5

Chapter 2

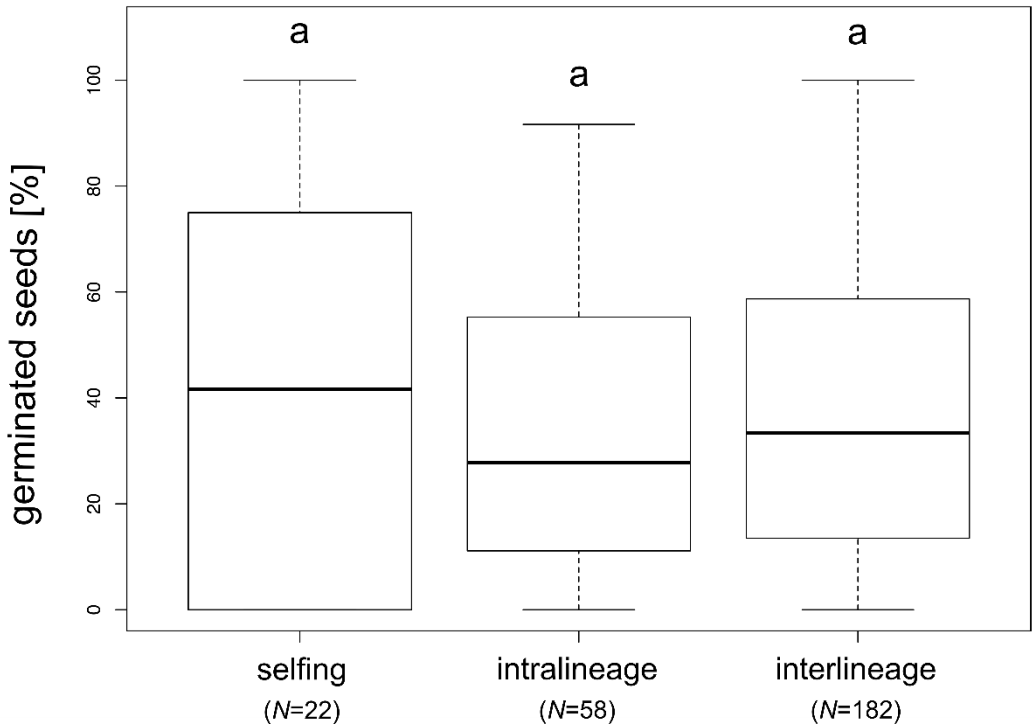


Figure S1. Boxplots showing germination rates of different treatments (selfing, intra-, and interlineage). Germination rate = percent of germinated seeds in relation to viable seeds. Letters above boxplots indicate non-significant differences between groups. See Table 2 for details. N = sample size (number of different crossings/crossing IDs used to determine germination rates).

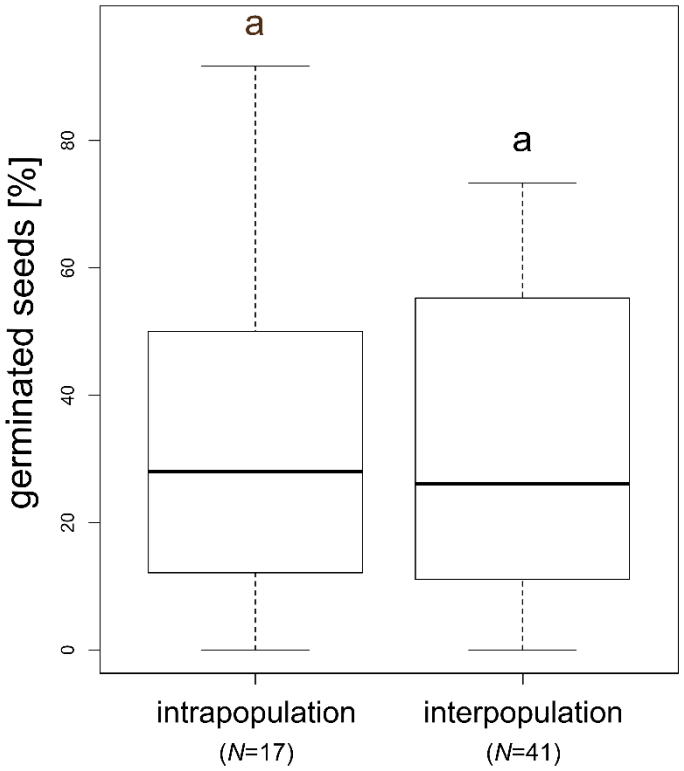


Figure S2. Boxplots showing germination rates of intra- und interpopulation crossings (see also Fig. 3b for seed set of intra- and interpopulation crossings). Germination rate = percent of germinated seeds in relation to viable seeds. Letters above boxplots indicate non-significant differences between groups. See Table 2 for details. N = sample size (number of different crossings/crossing IDs used to determine germination rates).

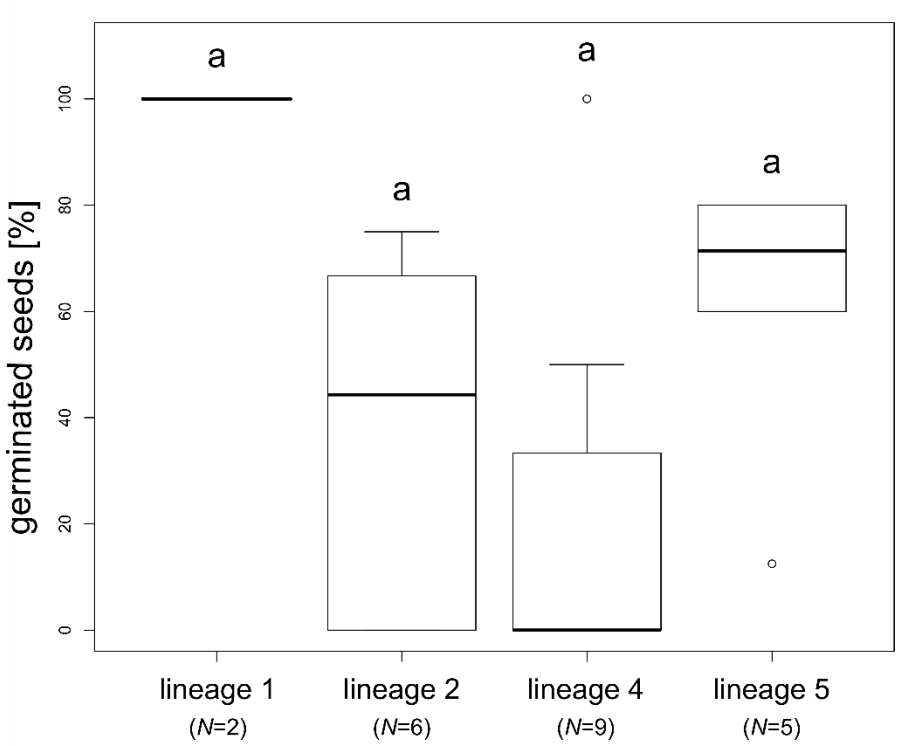


Figure S3. Boxplots showing germination rates of different lineages (selfings). Germination rate = percent of germinated seeds in relation to viable seeds. Letters above boxplots indicate non-significant differences between groups. See Table 2 for details. N = sample size (number of different crossings/crossing IDs used to determine germination rates).

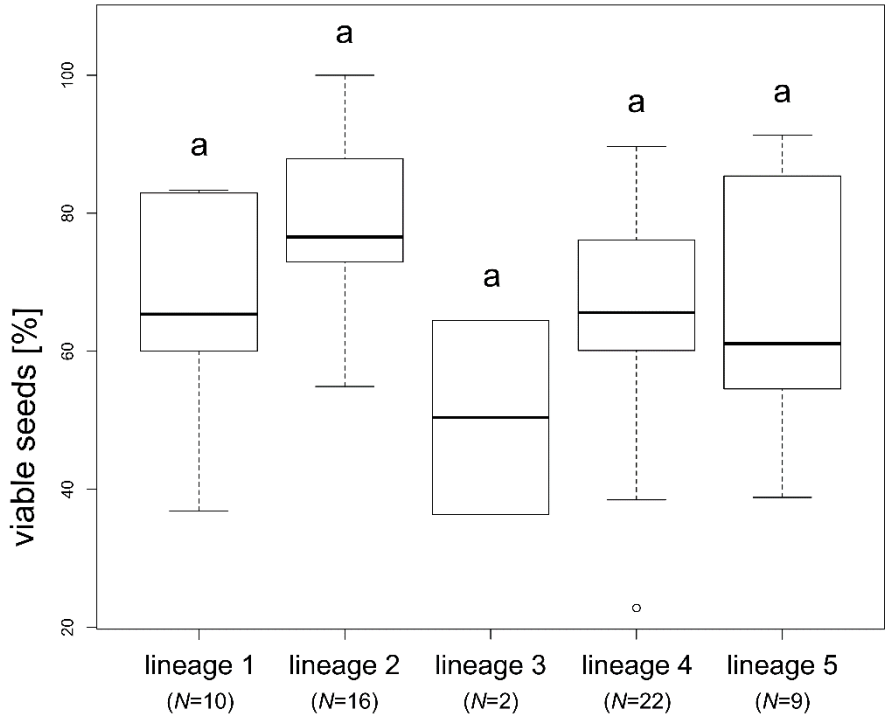


Figure S4. Boxplots showing seed set of different lineages (intralineage crossings). Seed set = percent of viable seeds in relation to all seeds. Letters above boxplots indicate non-significant differences between groups. See Table 2 for details. N = sample size (number of crossings/crossing IDs used to determine seed set).

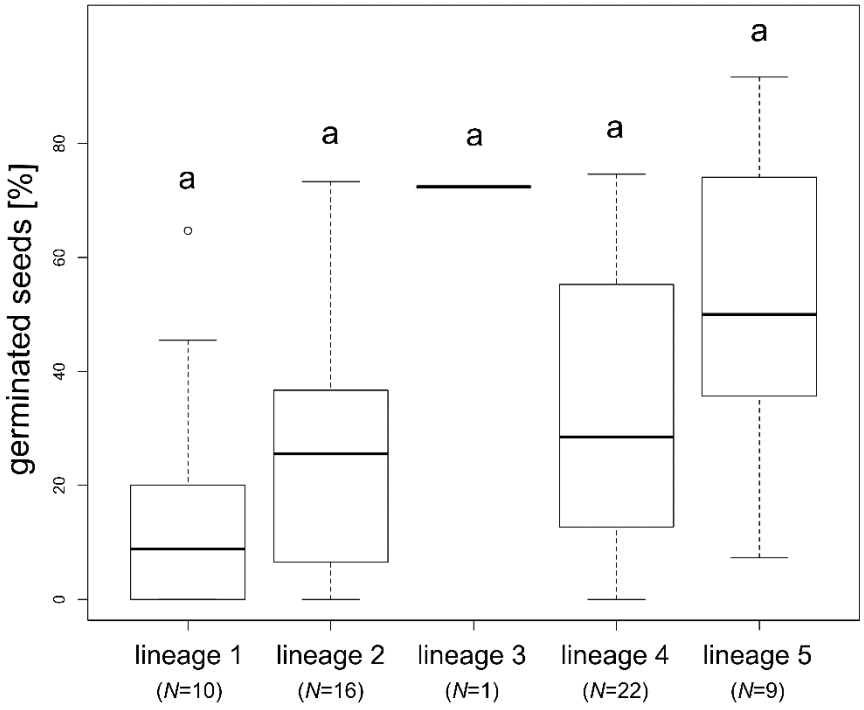


Figure S5. Boxplots showing germination rates of different lineages (intralineage crossings). Germination rate = percent of germinated seeds in relation to viable seeds. Letters above boxplots indicate non-significant differences between groups. See Table 2 for details. N = sample size (number of different crossings/crossing IDs used to determine germination rates).

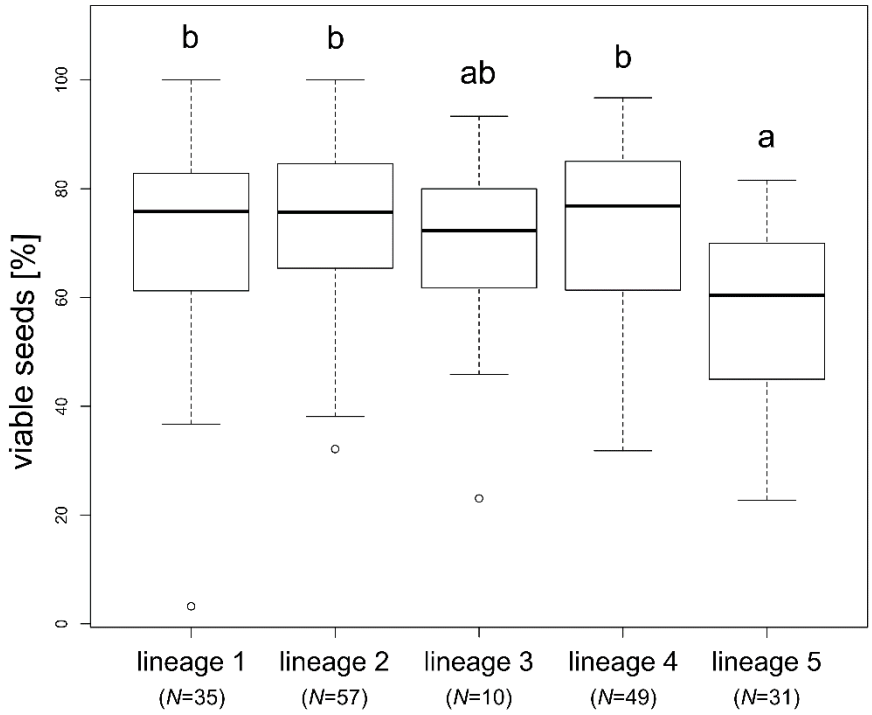


Figure S6. Boxplots showing seed set of different lineages (interlineage crossings). Seed set = percent of viable seeds in relation to all seeds. Letters above boxplots indicate significant/non-significant differences between groups. See Table 2 for details. N = sample size (number of crossings/crossing IDs used to determine seed set).

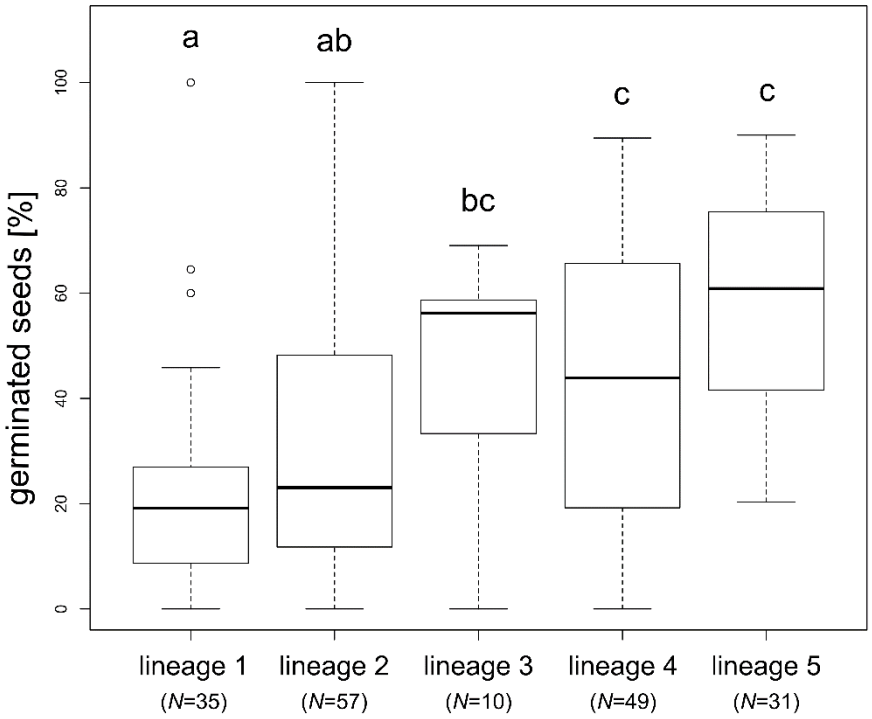


Figure S7. Boxplots showing germination rates of different lineages (interlineage crossings). Germination rate = percent of germinated seeds in relation to viable seeds. Letters above boxplots indicate significant/non-significant differences between groups. See Table 2 for details. N = sample size (number of different crossings/crossing IDs used to determine germination rates).

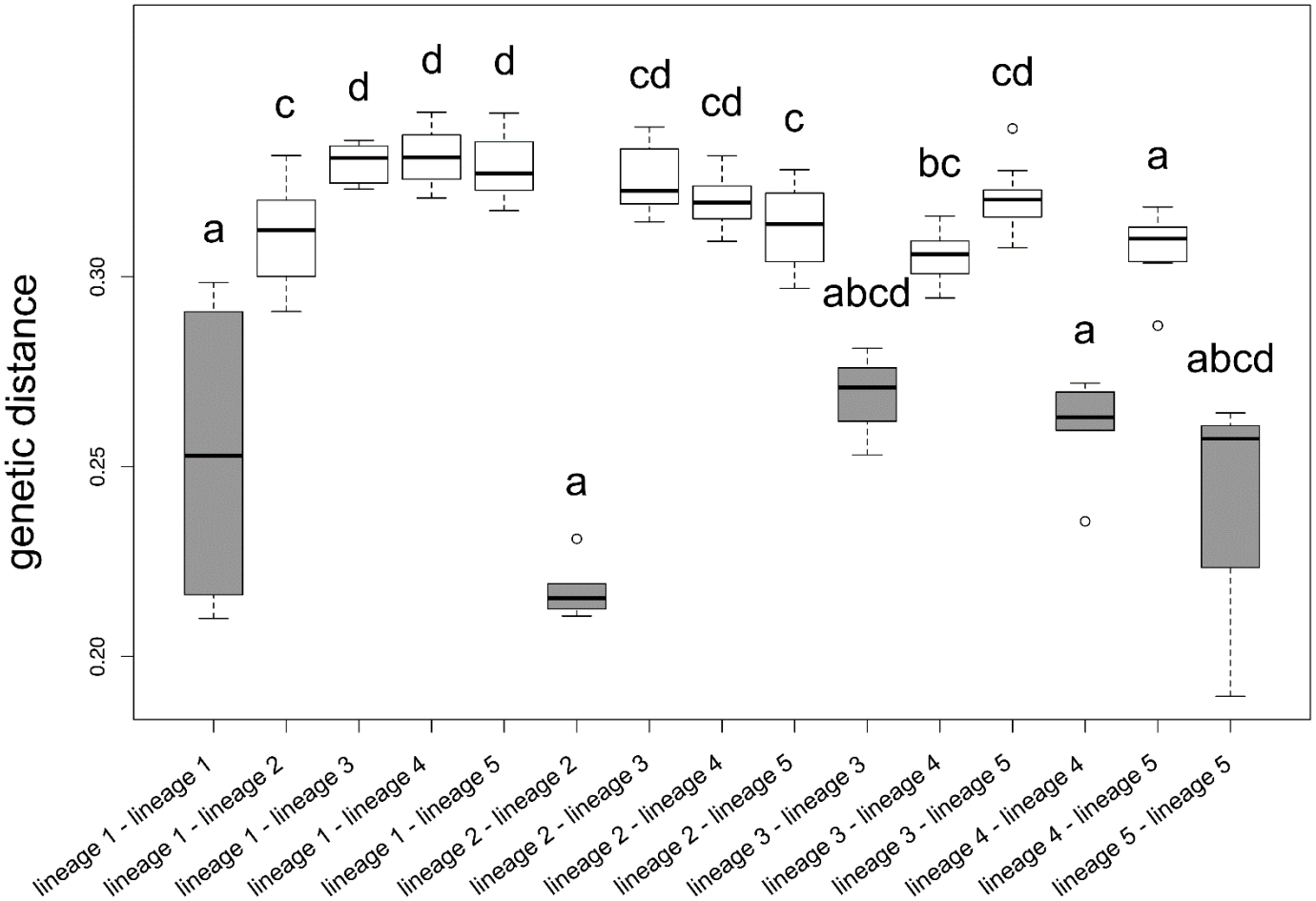


Figure S8. Boxplots showing genetic distances of different lineage combinations (intra- (e.g., "lineage 1 - lineage 1"; boxplots with grey filling) and interlineage (e.g., "lineage 2 - lineage 4"; boxplots with white filling)); lineages indicated by numbers). Letters above boxplots indicate significant/non-significant differences between groups. See Tables 1, 3 for details.

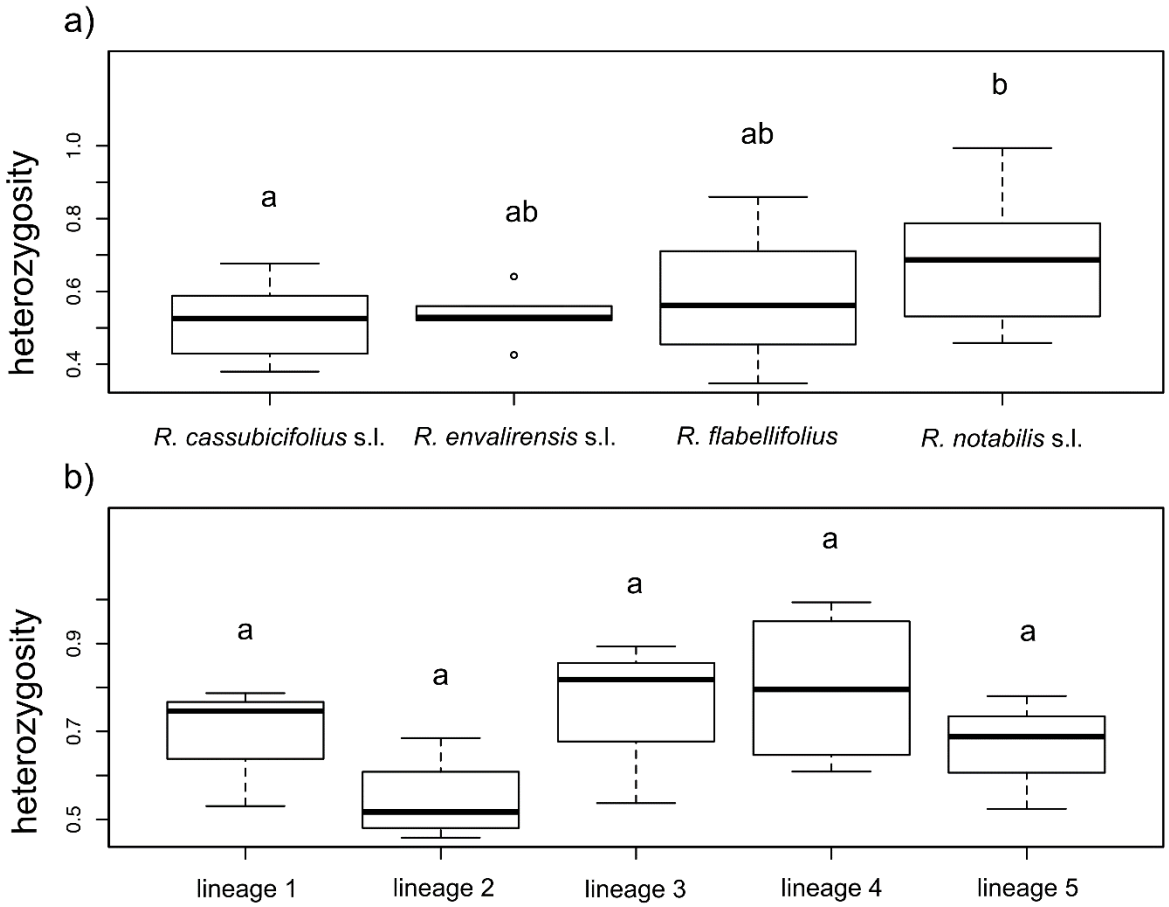


Figure S9. Boxplots showing (a) individual heterozygosity values among four sexual species of the *R. auricomus* complex and (b) individual heterozygosity values among *R. notabilis* lineages (data from Karbstein et al., subm.). Letters above boxplots indicate significant/non-significant differences between groups.

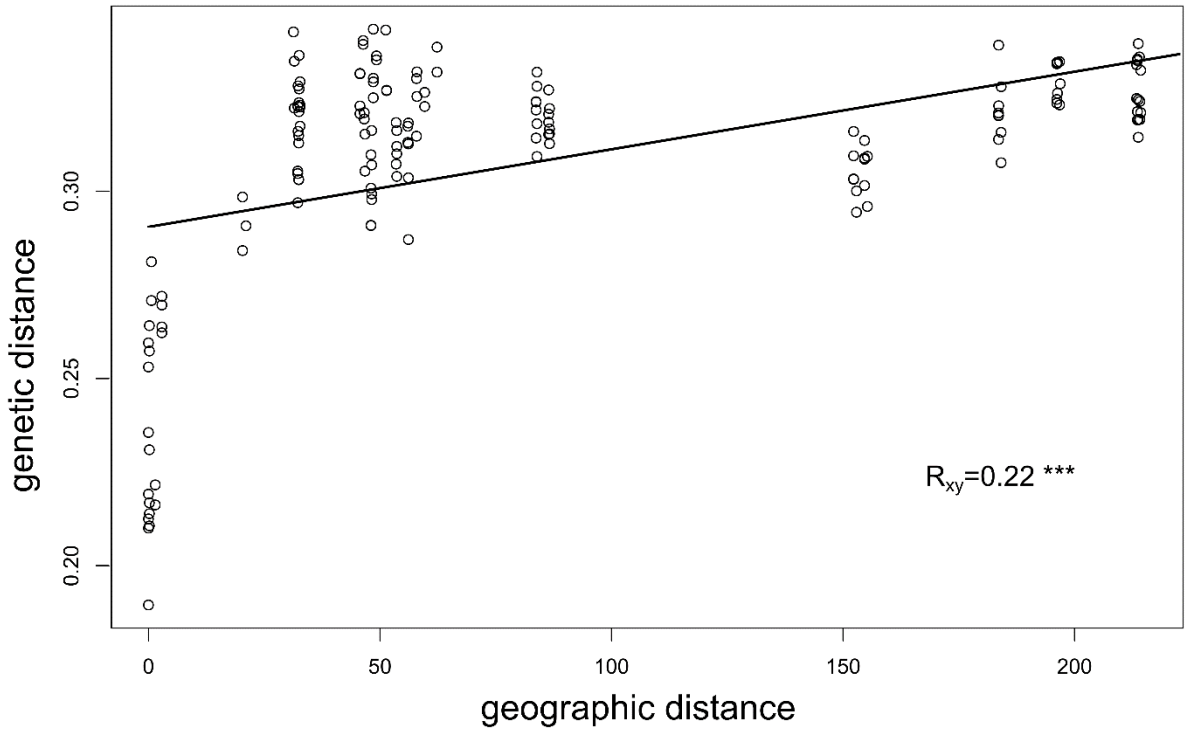


Figure S10. Scatter plot with a regression line based on Mantel test results (Spearman; 9999 permutations) of genetic distance as a function of geographical distances among individuals of *R. notabilis* s.l. lineages.

Chapter 3

Table S1 (electronical supplement). List of targeted loci and information about their function in *Arabidopsis thaliana* (source: The Arabidopsis Information Resource (TAIR); https://www.arabidopsis.org/tools/bulk/genes/index.jsp).

The electronical supplement is available online:
https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fmec.15458&file=mec15458-sup-0001-Supinfo.pdf

Supplementary Methods

Library Preparation

Genomic DNA was extracted from ~1.5 cm² leaf material of silica-dried samples or herbarium collections using the Qiagen DNeasy Plant Mini Kit® (Qiagen, Hilden, Germany). We followed the manufacturer's instructions, except for the sample incubation time in the lysis buffer, which was increased to one hour. DNA quality and fragment length were checked by gel electrophoresis in a 1.5% Agarose gel and using the Roti®-Load DNAstain 3 (Carl Roth, Karlsruhe, Germany). Extract concentration was estimated using the Qubit® fluorometer and the Qubit® dsDNA HS Assay Kit (ThermoFisher Scientific, Waltham, USA). Sequencing libraries were prepared using either the "NEBNext Ultra II DNA Library Prep Kit for Illumina®" (E7645) or the "NEBNext Ultra II FS DNA Library Prep Kit for Illumina®" (E7805) (New England BioLabs, Ipswich, USA). In the former case, we sheared DNA with a Bioruptor® Pico (Diagenode, Seraing, Belgium) prior to library preparation. Extracts were diluted to 10 ng/µL and sonicated for eight cycles of 15" sonication and 90" break to obtain fragments of approximately 300-500 bp. In the latter case, enzymatic shearing is combined with the first steps of the library preparation. Fragmentation was carried out for 12' at 37°C to obtain DNA fragments of the same length as for the sonicated samples. For the herbarium collections Hoe5615 and Du33351-15, shearing incubation was shorter, 3' and 10', respectively. In both cases, we followed the manufacturer's instructions. At the end of the library preparation procedure, samples were PCR-amplified for 14 cycles, during which sample-specific dual indices ("NEBNext Multiplex Oligos for Illumina®", E7600; New England BioLabs, Ipswich, USA) were added to the fragments. Indexed samples were pooled in equal quantities (four samples per 500 ng pool), dehydrated in a Concentrator Plus (Eppendorf, Hamburg, Germany), and diluted in 7 µL of ddH_2O. Each pool was enriched using the custom baits kit and following the manufacturer's protocol. Hybridization took place for 21h at 65 °C.

Enriched products were PCR-amplified for 14 cycles using the 2X KAPA HiFi HotStart Mix (KAPA Biosystems, Wilmington, USA) and the P7 and P5 adapters as primers. Amplified enriched libraries were purified with 50 µL of AMPure XP Beads (New England BioLabs, Ipswich, USA), following the manufacturer's protocol. Concentrations were measured with the Qubit® fluorometer and fragment length distributions were checked with a Bioanalyzer (Agilent, Santa Clara, USA). In the few cases in which fragment length was different from the desired one (and especially in case short fragments were present), pools were subjected to size selection with the BluePippin (Sage Science, Beverly, USA). All sequencing took place on an Illumina MiSeq System (Illumina Inc., San Diego, USA) at the Transcriptome and Genome Analysis Laboratory

(Georg-August-Universität, Gottingen, Germany). Pools were mixed equimolarly and sequenced in two different paired-end runs (6 pools, 24 samples each) with a 2x 250 bp (500 cycles) v2 kit.

Table S2. Detailed information about positions (along with their estimated substitution rates) excluded after position filtering for the different datasets.

a) Consensus alignments

Locus	position	rate
Assembly_AT1G09280	1522	84.1606
Assembly_AT1G09280	1494	84.1606
Assembly_AT1G09280	1492	84.1606
Assembly_AT1G09280	1479	84.1606
Assembly_AT1G09280	1470	84.1606
Assembly_AT1G09280	1468	84.1606
Assembly_AT1G09280	1458	84.1606
Assembly_AT1G09280	1457	84.1606
Assembly_AT1G09280	1456	84.1606
Assembly_AT1G09280	1454	84.1606
Assembly_AT1G09280	1453	84.1606
Assembly_AT1G09280	1452	84.1606
Assembly_AT1G09280	1451	84.1606
Assembly_AT1G09280	1449	84.1606
Assembly_AT1G09280	1447	84.1606
Assembly_AT1G09280	1446	84.1606
Assembly_AT1G09280	1445	84.1606
Assembly_AT1G09280	1444	84.1606
Assembly_AT1G09850	344	230.3317
Assembly_AT1G21370	374	716.5574
Assembly_AT1G31970	279	84.1606
Assembly_AT1G31970	267	84.1606
Assembly_AT1G75210	662	84.1606
Assembly_AT1G75210	13	55.6806
Assembly_AT1G78280	2444	190.5341
Assembly_AT2G03430	394	55.0661
Assembly_AT2G38060	587	178.2413
Assembly_AT2G38060	584	440.6491
Assembly_AT2G38060	564	91.2023
Assembly_AT2G38060	563	81.385
Assembly_AT2G38060	476	1966.7312
Assembly_AT2G38060	475	1724.6578
Assembly_AT2G38060	466	220.3121
Assembly_AT3G11830	788	1217.649
Assembly_AT3G12940	202	84.1606
Assembly_AT3G55260	523	84.1606
Assembly_AT3G55260	462	84.1606
Assembly_AT3G55260	461	84.1606
Assembly_AT3G55260	446	178.3747
Assembly_AT3G55260	440	84.1606
Assembly_AT3G55260	414	84.1606
Assembly_AT3G55260	407	84.1606
Assembly_AT3G55260	404	84.1606
Assembly_AT4G13590	594	236.9334
Assembly_AT4G13590	593	265.9859
Assembly_AT4G13590	592	236.9334
Assembly_AT4G13590	591	236.9334
Assembly_AT4G13590	590	197.9026
Assembly_AT4G13590	589	236.9334

Assembly_AT4G13590	588	197.9026
Assembly_AT4G13590	587	265.9859
Assembly_AT4G13590	586	236.9334
Assembly_AT4G13590	584	197.9026
Assembly_AT4G13590	583	197.9026
Assembly_AT4G13590	582	236.9334
Assembly_AT4G13590	581	265.9859
Assembly_AT4G13590	580	265.9859
Assembly_AT4G20325	378	229.9387
Assembly_AT4G20325	377	223.273
Assembly_AT4G30310	882	101.4457
Assembly_AT5G08720	289	84.1606
Assembly_AT5G08720	244	84.1606
Assembly_AT5G08720	234	84.1606
Assembly_AT5G20040	522	1762.6633
Assembly_AT5G21070	1152	84.1606
Assembly_AT5G21070	1134	84.1606
Assembly_AT5G48520	737	84.1606
Assembly_AT5G48520	736	84.1606
Assembly_AT5G48520	735	84.1606
Assembly_AT5G48520	734	84.1606
Assembly_AT5G48520	733	84.1606
Assembly_AT5G48520	731	84.1606
Assembly_AT5G48520	730	84.1606
Assembly_AT5G64940	828	84.1606
Assembly_AT5G64940	795	84.1606
Assembly_AT5G64940	637	84.1606
Assembly_AT5G66120	405	884.0018

b) Consensus alignments, paralog filtering

Locus	**position**	**rate**
Assembly_AT1G31970	279	84.1606
Assembly_AT1G31970	267	84.1606
Assembly_AT1G78280	2444	190.5341
Assembly_AT3G55260	523	84.1606
Assembly_AT3G55260	462	84.1606
Assembly_AT3G55260	461	84.1606
Assembly_AT3G55260	446	178.3747
Assembly_AT3G55260	440	84.1606
Assembly_AT3G55260	414	84.1606
Assembly_AT3G55260	407	84.1606
Assembly_AT3G55260	404	84.1606
Assembly_AT4G20325	378	229.9387
Assembly_AT4G20325	377	223.273
Assembly_AT5G21070	1152	84.1606
Assembly_AT5G21070	1134	84.1606
Assembly_AT5G66120	405	884.0018

c) Allele alignments

Exons	position	rate
Assembly_AT1G21370_13	78	110.5288
Assembly_AT2G18940_1	409	62.854
Assembly_AT2G18940_1	349	86.7073
Assembly_AT2G18940_49	409	62.854
Assembly_AT2G18940_49	349	86.7073
Assembly_AT3G04870_2376	6	48.9575
Assembly_AT4G17300_3349	293	1572.8704
Assembly_AT4G17300_3349	284	786.4385
Assembly_AT4G33060_1179	4	75.0819
Assembly_AT5G01010_17	93	221.1661
Assembly_AT5G01010_17	89	263.8743
Assembly_AT5G14600_85	36	98.6839
Assembly_AT5G64050_4	287	101.2217

d) Allele alignments, paralog filtering

Exons	position	rate
Assembly_AT2G18940_1	409	262.9397
Assembly_AT2G18940_1	349	443.5075
Assembly_AT2G18940_49	409	262.9397
Assembly_AT2G18940_49	349	443.5075
Assembly_AT3G04870_2376	6	76.6957
Assembly_AT4G17300_3349	293	179.3766
Assembly_AT4G17300_3349	284	229.6578
Assembly_AT4G33060_1179	4	118.3271
Assembly_AT5G19850_953	37	22.0063

Table S3. Number of stop codons registered by HYBPHYLOMAKER in the three possible reading frames, lowest number of stop codons and possible reading frames for all exons of the genes selected for the *BEAST analyses. Regions in bold are those for which no more than one possible reading frame was found for all exons.

exon	n. stop codons			min. n stop codons	possible r. frames
	1	2	3		
AT1G01090_1427	375	114	235	114	0
AT1G01090_556	0	419	812	0	1
AT1G02910_1170	0	60	213	0	1
AT1G02910_1483	59	93	0	0	1
AT1G02910_1723	0	98	122	0	1
AT1G02910_27	32	25	53	25	0
AT1G02910_951	159	0	58	0	1
AT1G06560_1278	0	60	60	0	1
AT1G06560_1816	30	58	0	0	1
AT1G06560_182	81	111	0	0	1
AT1G06560_2281	30	118	0	0	1
AT1G06560_2572	0	213	187	0	1
AT1G06560_3128	0	119	157	0	1
AT1G06560_4288	146	116	60	60	0
AT1G06560_600	435	0	151	0	1
AT1G08660_1345	360	0	60	0	1
AT1G08660_2125	116	31	0	0	1
AT1G08660_2562	416	31	236	31	0
AT1G08660_68	208	116	122	116	0
AT1G23400_1221	0	132	89	0	1
AT1G23400_1595	0	60	118	0	1
AT1G23400_1998	0	60	87	0	1
AT1G23400_2290	43	281	528	43	0
AT1G23400_903	30	204	0	0	1
AT1G23400_93	334	573	0	0	1
AT1G28680_1007	120	0	120	0	1
AT1G28680_1385	60	150	88	60	0
AT1G28680_75	539	0	325	0	1
AT1G30360_1235	0	89	148	0	1
AT1G30360_1608	0	667	509	0	1
AT1G30360_2545	152	0	297	0	1
AT1G30360_2889	131	203	121	121	0
AT1G30360_983	0	121	0	0	2
AT1G31480_1235	0	90	209	0	1
AT1G31480_1758	0	90	4	0	1
AT1G31480_2243	0	90	90	0	1
AT1G31480_2508	0	151	117	0	1
AT1G31480_3008	0	30	114	0	1
AT1G31480_3247	88	316	0	0	1
AT1G31480_5465	376	232	326	232	0
AT1G31480_893	0	41	150	0	1
AT1G32080_1652	0	237	0	0	2
AT1G32080_1899	182	87	0	0	1
AT1G32080_2127	28	59	150	28	0
AT1G32080_23	130	30	49	30	0
AT1G32080_441	0	153	134	0	1
AT1G32080_896	103	89	0	0	1
AT1G57770_1467	149	0	90	0	1
AT1G57770_1791	0	30	260	0	1
AT1G57770_2073	0	57	119	0	1
AT1G57770_2325	0	174	202	0	1
AT1G57770_2647	50	349	302	50	0
AT1G57770_510	113	0	89	0	1
AT1G68100_1262	0	0	0	0	3

AT1G68100_129	238	60	105	60	0
AT1G68100_1471	388	0	177	0	1
AT1G68100_2034	0	120	93	0	1
AT1G68100_2850	89	149	88	88	0
AT1G69500_1164	343	0	239	0	1
AT1G69500_1631	0	90	121	0	1
AT1G69500_1903	179	353	109	109	0
AT1G69500_39	150	60	270	60	0
AT1G69500_480	0	90	119	0	1
AT1G69500_767	0	319	191	0	1
AT1G74640_1220	178	30	0	0	1
AT1G74640_1459	225	21	187	21	0
AT1G74640_544	117	491	0	0	1
AT1G74640_91	45	133	0	0	1
AT1G78280_1042	237	0	60	0	1
AT1G78280_1355	0	59	88	0	1
AT1G78280_1831	0	355	472	0	1
AT1G78280_2532	0	60	120	0	1
AT1G78280_2913	0	153	301	0	1
AT1G78280_3337	90	178	0	0	1
AT1G78280_3815	0	31	323	0	1
AT1G78280_4056	59	179	0	0	1
AT1G78280_4332	143	118	186	118	0
AT1G78280_813	60	91	0	0	1
AT1G78280_96	60	105	0	0	1
AT2G13100_1158	30	349	415	30	0
AT2G13100_291	550	0	518	0	1
AT2G21370_1137	**0**	**1**	**89**	**0**	**1**
AT2G21370_1578	**0**	**150**	**270**	**0**	**1**
AT2G21370_222	**60**	**0**	**59**	**0**	**1**
AT2G21370_503	**56**	**0**	**30**	**0**	**1**
AT2G21370_772	**0**	**56**	**133**	**0**	**1**
AT2G21720_1245	**60**	**0**	**30**	**0**	**1**
AT2G21720_1455	**89**	**154**	**0**	**0**	**1**
AT2G21720_1720	**503**	**0**	**421**	**0**	**1**
AT2G21720_2568	**121**	**0**	**150**	**0**	**1**
AT2G21720_487	**216**	**0**	**92**	**0**	**1**
AT2G21720_751	**29**	**0**	**60**	**0**	**1**
AT2G21720_966	0	173	174	0	1
AT2G28070_1355	30	0	119	0	1
AT2G28070_1833	0	59	180	0	1
AT2G28070_205	59	339	26	26	0
AT2G28070_2064	0	171	29	0	1
AT2G28070_2508	0	296	89	0	1
AT2G28070_3343	47	180	236	47	0
AT2G28070_975	0	180	119	0	1
AT2G34640_1307	180	0	149	0	1
AT2G34640_1649	30	100	0	0	1
AT2G34640_230	242	686	0	0	1
AT2G34640_2333	202	104	61	61	0
AT2G40190_1080	90	0	30	0	1
AT2G40190_1536	0	105	258	0	1
AT2G40190_2023	114	171	17	17	0
AT2G40190_242	0	165	402	0	1
AT2G40190_696	0	173	229	0	1
AT2G45500_121	329	0	145	0	1
AT2G45500_1937	0	120	225	0	1
AT2G45500_2327	0	85	139	0	1
AT2G45500_2719	142	171	233	142	0
AT2G45500_966	0	56	327	0	1
AT3G04480_1764	0	76	370	0	1
AT3G04480_2168	179	237	0	0	1

AT3G04480_2583	0	132	234	0	1
AT3G04480_3003	0	209	292	0	1
AT3G04480_3460	158	158	262	158	0
AT3G04480_982	140	0	120	0	1
AT3G07720_826	**0**	**332**	**767**	**0**	**1**
AT3G07720_874	**0**	**329**	**767**	**0**	**1**
AT3G23940_1182	120	0	30	0	1
AT3G23940_1504	119	0	157	0	1
AT3G23940_2296	60	0	0	0	2
AT3G23940_279	203	332	0	0	1
AT3G23940_3399	0	89	137	0	1
AT3G23940_3654	68	76	139	68	0
AT3G27530_1430	0	148	119	0	1
AT3G27530_3449	0	118	149	0	1
AT3G27530_4296	318	555	0	0	1
AT3G27530_5491	30	238	710	30	0
AT3G27530_656	115	0	89	0	1
AT3G44880_1106	76	0	30	0	1
AT3G44880_1322	0	100	123	0	1
AT3G44880_1541	90	0	59	0	1
AT3G44880_1774	0	120	209	0	1
AT3G44880_2152	0	66	155	0	1
AT3G44880_777	0	0	246	0	2
AT3G44880_82	319	0	59	0	1
AT3G54090_1178	144	386	688	144	0
AT3G54090_81	347	485	56	56	0
AT3G54510_1032	179	145	0	0	1
AT3G54510_1269	0	120	148	0	1
AT3G54510_1673	0	204	147	0	1
AT3G54510_2050	0	210	1	0	1
AT3G54510_2398	0	148	120	0	1
AT3G54510_2720	89	166	229	89	0
AT3G54510_358	161	0	72	0	1
AT3G54510_738	0	60	178	0	1
AT3G58470_1334	30	140	54	30	0
AT3G58470_527	0	30	207	0	1
AT3G58470_64	209	130	0	0	1
AT3G59040_12	405	16	229	16	0
AT3G59040_1512	0	90	64	0	1
AT3G59040_1903	741	18	317	18	0
AT3G59040_852	0	408	498	0	1
AT3G59770_1123	0	0	0	0	3
AT3G59770_186	0	138	448	0	1
AT3G59770_2805	45	183	0	0	1
AT3G59770_3091	120	90	0	0	1
AT3G59770_3474	0	167	396	0	1
AT3G59770_4092	210	691	0	0	1
AT3G59770_5039	0	61	169	0	1
AT3G59770_5345	0	169	152	0	1
AT3G59770_880	0	120	30	0	1
AT3G61320_1089	0	86	89	0	1
AT3G61320_1385	39	65	218	39	0
AT3G61320_63	29	423	383	29	0
AT3G61320_835	0	89	166	0	1
AT4G13250_1704	266	0	150	0	1
AT4G13250_2352	120	88	0	0	1
AT4G13250_2750	125	0	60	0	1
AT4G13250_3191	0	90	60	0	1
AT4G13250_3405	0	150	4	0	1
AT4G13250_3603	179	0	90	0	1
AT4G13250_3890	178	149	161	149	0
AT4G15850_1270	0	196	90	0	1

AT4G15850_1554	150	0	115	0	1
AT4G15850_1840	0	116	59	0	1
AT4G15850_2128	150	0	175	0	1
AT4G15850_3036	202	134	173	134	0
AT4G15850_612	147	0	59	0	1
AT4G15850_935	0	142	149	0	1
AT4G19860_1391	0	150	60	0	1
AT4G19860_1684	0	148	179	0	1
AT4G19860_1	0	134	219	0	1
AT4G19860_2279	131	196	451	131	0
AT4G19860_833	118	0	59	0	1
AT4G27340_1478	0	148	152	0	1
AT4G27340_2301	90	118	0	0	1
AT4G27340_2946	58	88	60	58	0
AT4G27340_63	1014	0	418	0	1
AT4G29310_1487	226	240	336	226	0
AT4G29310_35	418	29	181	29	0
AT4G29310_758	377	0	209	0	1
AT4G33440_1242	0	124	37	0	1
AT4G33440_1675	61	34	0	0	1
AT4G33440_2117	632	535	63	63	0
AT5G06580_16	0	58	27	0	1
AT5G06580_4289	59	0	60	0	1
AT5G06580_4949	119	423	122	119	0
AT5G06830_1146	**60**	**0**	**30**	**0**	**1**
AT5G06830_1362	**0**	**90**	**107**	**0**	**1**
AT5G06830_1605	656	0	246	0	1
AT5G06830_2255	240	80	0	0	1
AT5G06830_2544	89	0	181	0	1
AT5G10910_1171	0	120	90	0	1
AT5G10910_1654	88	328	472	88	0
AT5G10910_178	2	224	0	0	1
AT5G10910_661	151	0	235	0	1
AT5G10920_106	1	292	0	0	1
AT5G10920_1144	0	210	61	0	1
AT5G10920_1404	294	0	140	0	1
AT5G10920_1947	3	176	292	3	0
AT5G10920_546	0	180	538	0	1
AT5G27950_1125	0	61	89	0	1
AT5G27950_1379	150	0	117	0	1
AT5G27950_1624	237	0	150	0	1
AT5G27950_2136	60	30	0	0	1
AT5G27950_2339	440	472	0	0	1
AT5G27950_2897	148	174	118	118	0
AT5G27950_662	32	187	0	0	1
AT5G41150_1058	180	0	137	0	1
AT5G41150_1654	0	118	62	0	1
AT5G41150_1932	0	416	740	0	1
AT5G41150_3149	0	360	60	0	1
AT5G41150_3508	0	120	302	0	1
AT5G41150_3822	224	488	149	149	0
AT5G41150_50	771	0	613	0	1
AT5G45900_1305	117	103	0	0	1
AT5G45900_1686	0	0	151	0	2
AT5G45900_1917	0	150	352	0	1
AT5G45900_2353	0	11	238	0	1
AT5G45900_2826	0	60	177	0	1
AT5G45900_3063	146	86	192	86	0
AT5G45900_497	60	166	0	0	1
AT5G45900_793	0	90	156	0	1
AT5G47090_521	441	885	67	67	0
AT5G47090_71	0	17	0	0	2

AT5G54910_1226	0	150	328	0	1
AT5G54910_1667	0	89	178	0	1
AT5G54910_1941	0	60	60	0	1
AT5G54910_2180	0	356	246	0	1
AT5G54910_2678	21	84	0	0	1
AT5G54910_2932	454	30	148	30	0
AT5G54910_66	12	113	213	12	0
AT5G54910_689	0	205	186	0	1
AT5G57590_1200	200	0	119	0	1
AT5G57590_1775	0	60	89	0	1
AT5G57590_1969	1	210	0	0	1
AT5G57590_206	325	575	0	0	1
AT5G57590_2876	60	143	0	0	1
AT5G57590_3180	57	58	0	0	1
AT5G57590_3400	119	31	0	0	1
AT5G57590_3614	0	2	82	0	1
AT5G57590_4104	170	204	369	170	0
AT5G57590_922	150	0	204	0	1
AT5G64050_2652	91	0	30	0	1
AT5G64050_2904	210	248	281	210	0
AT5G64050_4	144	205	84	84	0
AT5G64050_753	30	88	0	0	1
AT5G64050_960	0	89	89	0	1
AT5G64370_1236	0	119	149	0	1
AT5G64370_1554	56	284	411	56	0
AT5G64370_647	0	60	35	0	1
AT5G64370_930	1	143	125	1	0

Table S4. Gene-wise dN/dS estimates (ω) for all the loci used for the BEAST divergence time estimates.

Locus	omega (dN/dS)
AT1G01090	0.00266
AT1G02910	0.08252
AT1G06560	0.31212
AT1G08660	0.02384
AT1G23400	0.04956
AT1G28680	0.09499
AT1G30360	0.30462
AT1G31480	0.32739
AT1G32080	0.13297
AT1G57770	0.0463
AT1G68100	0.24461
AT1G69500	0.03692
AT1G74640	0.10571
AT1G78280	0.21776
AT2G13100	0.07836
AT2G21370	0.19494
AT2G21720	0.16757
AT2G28070	0.06293
AT2G34640	0.15694
AT2G40190	0.13775
AT2G45500	0.14595
AT3G04480	0.21084
AT3G07720	0.09081
AT3G23940	0.04214
AT3G27530	0.08172
AT3G44880	0.23239
AT3G54510	0.13805
AT3G58470	0.25813
AT3G59040	0.14138
AT3G59770	0.14236
AT3G61320	0.01141
AT4G13250	0.05148

Locus	omega (dN/dS)
AT4G15850	0.39361
AT4G19860	0.16078
AT4G27340	0.1433
AT4G29310	0.01389
AT4G33440	0.04804
AT5G06580	0.14231
AT5G06830	0.18238
AT5G10910	0.55319
AT5G10920	0.09421
AT5G27950	0.17565
AT5G41150	0.13296
AT5G45900	0.2288
AT5G54910	0.19479
AT5G57590	0.21234
AT5G64050	0.07354
AT5G64370	0.09457

Table S5. The 50 loci with the highest number of parsimony informative sites along with their best-fitting substitution model and characteristics.

		Model	f(a)	f(c)	f(g)	f(t)	kappa	titv	Ra	Rb	Rc	Rd	Re	Rf	pInv	gamma
AT1G01090	BIC	HKY+I	0.26	0.14	0.28	0.32	10.75	5.1	1	10.745	1	1	10.745	1	0.87	N/A
AT1G02910	BIC	HKY+I	0.28	0.22	0.21	0.29	4.33	2.12	1	4.329	1	1	4.329	1	0.9	N/A
AT1G06560	BIC	TrN+I	0.29	0.19	0.26	0.26	0	0	1	1.819	1	1	4.161	1	0.82	N/A
AT1G08660	BIC	TrN+I	0.3	0.2	0.22	0.28	0	0	1	1.564	1	1	3.843	1	0.86	N/A
AT1G23400	BIC	HKY+I+G	0.31	0.21	0.21	0.28	3.77	1.83	1	3.772	1	1	3.772	1	0.81	0.8
AT1G28680	BIC	K80+I+G	0.25	0.25	0.25	0.25	4.26	2.13	1	4.26	1	1	4.26	1	0.85	0.65
AT1G30360	BIC	HKY+I	0.28	0.21	0.21	0.3	2.43	1.18	1	2.425	1	1	2.425	1	0.84	N/A
AT1G31480	BIC	HKY+I	0.3	0.17	0.24	0.29	4.03	1.96	1	4.027	1	1	4.027	1	0.9	N/A
AT1G32080	BIC	TPM3uf+I	0.25	0.22	0.21	0.31	0	0	4.094	5.241	1	4.094	5.241	1	0.86	N/A
AT1G57770	BIC	HKY+I	0.27	0.19	0.23	0.31	4.94	2.4	1	4.943	1	1	4.943	1	0.94	N/A
AT1G68100	BIC	TrN+G	0.29	0.2	0.21	0.3	0	0	1	2.58	1	1	5.425	1	N/A	0.02
AT1G69500	BIC	HKY+I	0.32	0.2	0.2	0.28	3.66	1.75	1	3.659	1	1	3.659	1	0.84	N/A
AT1G74640	BIC	TPM2uf+I	0.29	0.21	0.22	0.28	0	0	4.181	10.045	4.181	1	10.045	1	0.87	N/A
AT1G78280	BIC	HKY+I	0.3	0.19	0.23	0.29	2.7	1.32	1	2.699	1	1	2.699	1	0.85	N/A
AT2G13100	BIC	HKY+G	0.23	0.14	0.27	0.36	2.44	1.1	1	2.438	1	1	2.438	1	N/A	0.02
AT2G21370	BIC	K80+I	0.25	0.25	0.25	0.25	4.27	2.14	1	4.273	1	1	4.273	1	0.84	N/A
AT2G21720	BIC	HKY+I	0.32	0.18	0.23	0.27	4.58	2.26	1	4.583	1	1	4.583	1	0.84	N/A
AT2G28070	BIC	HKY+I	0.27	0.21	0.21	0.31	3.05	1.49	1	3.051	1	1	3.051	1	0.89	N/A
AT2G34640	BIC	HKY+I+G	0.31	0.2	0.25	0.24	2.24	1.14	1	2.242	1	1	2.242	1	0.7	0.9
AT2G40190	BIC	HKY+I	0.3	0.19	0.22	0.3	3.83	1.85	1	3.83	1	1	3.83	1	0.9	N/A
AT2G45500	BIC	HKY+I	0.33	0.19	0.19	0.28	3.47	1.65	1	3.466	1	1	3.466	1	0.87	N/A
AT3G04480	BIC	HKY+I	0.29	0.21	0.19	0.31	5.28	2.53	1	5.278	1	1	5.278	1	0.86	N/A
AT3G07720	BIC	TrNef+I	0.25	0.25	0.25	0.25	0	0	1	1.995	1	1	5.547	1	0.67	N/A
AT3G23940	BIC	TrN+I+G	0.31	0.17	0.24	0.28	0	0	1	1.439	1	1	5.935	1	0.75	0.6
AT3G27530	BIC	HKY+I	0.32	0.18	0.24	0.27	3.75	1.85	1	3.746	1	1	3.746	1	0.89	N/A
AT3G44880	BIC	TPM3+I	0.25	0.25	0.25	0.25	0	0	2.556	5.89	1	2.556	5.89	1	0.88	N/A
AT3G54090	BIC	HKY+I	0.3	0.2	0.22	0.28	3.13	1.53	1	3.129	1	1	3.129	1	0.86	N/A
AT3G54510	BIC	HKY	0.27	0.19	0.21	0.34	4.07	1.94	1	4.074	1	1	4.074	1	N/A	N/A
AT3G58470	BIC	HKY+I	0.3	0.18	0.24	0.28	5.47	2.69	1	5.472	1	1	5.472	1	0.88	N/A
AT3G59040	BIC	HKY+I	0.33	0.17	0.23	0.28	3.69	1.81	1	3.69	1	1	3.69	1	0.87	N/A
AT3G59770	BIC	HKY+G	0.26	0.21	0.23	0.29	3.89	1.92	1	3.892	1	1	3.892	1	N/A	0.02
AT3G61320	BIC	K80+I+G	0.25	0.25	0.25	0.25	2.19	1.09	1	2.187	1	1	2.187	1	0.83	0.88
AT4G13250	BIC	HKY+I	0.31	0.19	0.23	0.27	3.56	1.76	1	3.564	1	1	3.564	1	0.93	N/A
AT4G15850	BIC	HKY	0.27	0.19	0.23	0.31	2.51	1.22	1	2.51	1	1	2.51	1	N/A	N/A
AT4G19860	BIC	HKY+I	0.31	0.17	0.24	0.28	3.2	1.58	1	3.201	1	1	3.201	1	0.79	N/A
AT4G27340	BIC	TPM2uf+I	0.29	0.14	0.27	0.3	0	0	0.463	1.697	0.463	1	1.697	1	0.86	N/A
AT4G29310	BIC	HKY+I	0.24	0.22	0.23	0.31	2.39	1.18	1	2.394	1	1	2.394	1	0.85	N/A
AT4G33440	BIC	TrN+I	0.31	0.17	0.25	0.27	0	0	1	1.726	1	1	8.945	1	0.91	N/A
AT5G06580	BIC	TPM3uf+I	0.28	0.27	0.15	0.29	0	0	3.294	6.95	1	3.294	6.95	1	0.85	N/A
AT5G06830	BIC	HKY+I	0.32	0.18	0.23	0.27	5.45	2.69	1	5.445	1	1	5.445	1	0.82	N/A
AT5G10910	BIC	TPM3uf+I	0.3	0.19	0.24	0.26	0	0	2.735	4.181	1	2.735	4.181	1	0.85	N/A
AT5G10920	BIC	TIM2+I	0.3	0.2	0.22	0.28	0	0	3.081	1.847	3.081	1	5.983	1	0.89	N/A
AT5G27950	BIC	HKY+I	0.32	0.18	0.24	0.27	6.11	3.03	1	6.114	1	1	6.114	1	0.8	N/A
AT5G41150	BIC	HKY+I	0.29	0.2	0.23	0.28	3.74	1.84	1	3.743	1	1	3.743	1	0.84	N/A
AT5G45900	BIC	TPM2uf+I	0.25	0.21	0.24	0.3	0	0	0.489	1.367	0.489	1	1.367	1	0.91	N/A
AT5G47090	BIC	TPM1uf+I+G	0.33	0.16	0.28	0.23	0	0	1	7.505	2.516	2.516	7.505	1	0.71	0.64
AT5G54910	BIC	HKY+G	0.33	0.19	0.22	0.27	2.25	1.11	1	2.254	1	1	2.254	1	N/A	0.02
AT5G57590	BIC	HKY+I	0.28	0.17	0.25	0.3	3.47	1.7	1	3.473	1	1	3.473	1	0.89	N/A
AT5G64050	BIC	HKY+I+G	0.28	0.2	0.21	0.31	2.51	1.21	1	2.511	1	1	2.511	1	0.79	0.89
AT5G64370	BIC	HKY+I	0.28	0.21	0.23	0.28	3.93	1.93	1	3.925	1	1	3.925	1	0.85	N/A

Table S6. Number of reads per sample, along with information about quality trimming and duplicate removal.

Sample no.	Taxon	Nr. of pairs	Nr. of reads	Both surviving	Forward only surviving	Reverse only surviving	Nr. reads after quality trimming	% quality trimmed reads	Nr. reads without duplicates	% duplicates
DU003-1	*Ranunculus austroslovenicus*	434604	869208	418577	9204	4216	850574	2.15	788804	7.27
LH012-8	*Ranunculus austroslovenicus*	614059	1228118	593017	11436	6398	1203868	1.98	1088796	9.56
2	*Ranunculus calapius*	591569	1183138	547175	33091	4235	1131676	4.35	1122814	0.79
Du35351-15	*Ranunculus calapius*	878989	1757978	843735	20017	7997	1715484	2.42	1630888	4.94
9126-2	*Ranunculus carpaticola*	878231	1756462	836843	19697	9823	1703206	3.04	1545974	9.24
DU013-1	*Ranunculus carpaticola*	791056	1582112	742839	36202	5318	1527198	3.48	1370336	10.28
LH40-4	*Ranunculus carpaticola*	590893	1181786	552794	28544	4068	1138200	3.69	1070354	5.97
LH006-17	*Ranunculus cassubicifolius*	715201	1430402	674038	29304	5430	1382810	3.33	1284166	7.14
LH016-14	*Ranunculus cassubicifolius*	1089476	2178952	1051712	20883	10570	2134877	2.03	1928887	9.65
LH008-7	*Ranunculus cassubicifolius*	742197	1484394	712605	17458	7015	1449683	2.34	1318205	9.07
LH009-3	*Ranunculus cassubicifolius*	677510	1355020	589344	77486	2341	1258515	7.13	1193781	5.15
Du33354-2	*Ranunculus cebennensis*	734001	1468002	694081	26409	6366	1420937	3.21	1360769	4.24
DU018-1	*Ranunculus envalirensis*	471152	942304	452010	11863	4279	920162	2.35	834448	9.32
Du019	*Ranunculus envalirensis*	618223	1236446	525997	81023	2217	1135234	8.19	1120850	1.27
DU021-1	*Ranunculus flabellifolius*	796503	1593006	770210	13476	8247	1562143	1.94	1370079	12.3
DU054-1	*Ranunculus flabellifolius*	563973	1127946	541083	14506	4726	1101398	2.36	979444	11.08
LH2501	*Ranunculus flabellifolius*	533925	1067850	505220	20250	3864	1034554	3.12	956330	7.57
LH017-1	*Ranunculus marsicus*	549225	1098450	527303	12906	5484	1072996	2.32	967954	9.79
LH018-2	*Ranunculus marsicus*	715224	1430448	680873	22667	6395	1390808	2.78	1275636	8.29
LH014-3	*Ranunculus mediocompositus*	690788	1381576	650077	30227	4615	1334996	3.38	1226172	8.16
LH015-10	*Ranunculus mediocompositus*	758808	1517616	729505	17696	6762	1483468	2.26	1332162	10.2
10137-3	*Ranunculus notabilis*	684406	1368812	647306	26682	4787	1326081	3.13	1226859	7.49
Hoe5615	*Ranunculus notabilis*	788655	1577310	653152	121175	2219	1429698	9.36	1359838	4.89
LH010-7	*Ranunculus peracris*	433500	867000	416171	10395	4167	846904	2.32	794542	6.19

LH011-1	*Ranunculus peracris*	680873	1361746	638035	32232	4542	1312844	3.6	1193156	9.12
LH011-14	*Ranunculus peracris*	482678	965356	461588	13306	4441	940923	2.54	883035	6.16
LG09	*Ranunculus pygmaeus*	908098	1816196	875316	19488	8040	1778160	2.1	1619494	8.93
104263	*Ranunculus sceleratus*	1047365	2094730	991745	40081	7439	2031010	3.05	1886966	7.1
DU049-1	*Ranunculus subcarniolicus*	656379	1312758	631484	14967	6148	1284083	2.19	1183105	7.87
LH013-3	*Ranunculus subcarniolicus*	479816	959632	463766	8596	4762	940890	1.96	876464	6.85

Table S7. Number and percent of mapped read per sample.

Sample no.	Taxon	Total nr. reads	Nr. paired reads	Nr. forward unpaired reads	Nr. reverse unpaired reads	Nr. mapped reads	Percent of mapped reads
DU003-1	*Ranunculus austroslovenicus*	828673	387692	9204	4216	662966	80.003
LH012-8	*Ranunculus austroslovenicus*	1144964	535481	11436	6398	935564	81.711
2	*Ranunculus calapius*	1136354	542744	33091	4235	305569	26.89
Du35351-15	*Ranunculus calapius*	1658283	801437	20017	7997	974768	58.781
9126-2	*Ranunculus carpaticola*	1555793	758227	19697	9823	857218	55.098
DU013-1	*Ranunculus carpaticola*	1449882	664408	36202	5318	1079453	74.451
LH40-4	*Ranunculus carpaticola*	1100804	518871	28544	4068	717409	65.171
LH006-17	*Ranunculus cassubicifolius*	1334773	624716	29304	5430	910910	68.244
LH016-14	*Ranunculus cassubicifolius*	2015353	948717	20883	10570	1416897	70.305
LH008-7	*Ranunculus cassubicifolius*	1363349	646866	17458	7015	927282	68.015
LH009-3	*Ranunculus cassubicifolius*	1230345	556977	77486	2341	830803	67.526
Du33354-2	*Ranunculus cebennensis*	1384000	663997	26409	6366	839598	60.664
DU018-1	*Ranunculus envalirensis*	880508	409153	11863	4279	615685	69.923
Du019	*Ranunculus envalirensis*	1147489	518805	81023	2217	611987	53.332
DU021-1	*Ranunculus flabellifolius*	1424710	674178	13476	8247	1015231	71.258
DU054-1	*Ranunculus flabellifolius*	1032953	480106	14506	4726	768506	74.398
LH2501	*Ranunculus flabellifolius*	984472	466108	20250	3864	618623	62.838
LH017-1	*Ranunculus marsicus*	1015923	474782	12906	5484	760511	74.859
LH018-2	*Ranunculus marsicus*	1342779	623287	22667	6395	946497	70.487
LH014-3	*Ranunculus mediocompositus*	1290181	595665	30227	4615	1041981	80.762
LH015-10	*Ranunculus mediocompositus*	1410906	653852	17696	6762	1141346	80.894
10137-3	*Ranunculus notabilis*	1283419	597695	26682	4787	1008654	78.591
Hoe5615	*Ranunculus notabilis*	1379116	618222	121175	2219	863888	62.64
LH010-7	*Ranunculus peracris*	825825	389990	10395	4167	619412	75.005
LH011-1	*Ranunculus peracris*	1254824	578191	32232	4542	935436	74.547
LH011-14	*Ranunculus peracris*	929053	432644	13306	4441	702164	75.578
LG09	*Ranunculus pygmaeus*	1715802	795983	19488	8040	1415286	82.485

104263	*Ranunculus sceleratus*	1986542	919723	40081	7439	1571110	79.087
DU049-1	*Ranunculus subcarniolicus*	1254050	580995	14967	6148	994326	79.289
LH013-3	*Ranunculus subcarniolicus*	920035	431553	8596	4762	747517	81.248

Table S8. Tukey multiple comparisons of means. Significant p values are marked in bolt.

	diff	lwr	upr	p adj
AlleleAlign.ParalogFilt.PositionUnfilt.-AlleleAlign.ParalogFilt.PositionFilt.	0.00098571	-0.0178477	0.01981912	0.9999999
AlleleAlign.ParalogUnfilt.PositionFilt.-AlleleAlign.ParalogFilt.PositionFilt.	0.03172143	0.01288802	0.05055483	**0.0000090**
AlleleAlign.ParalogUnfilt.PositionUnfilt.-AlleleAlign.ParalogFilt.PositionFilt.	0.03446429	0.01563088	0.05329769	**0.0000006**
ConsensusAlign.ParalogFilt.PositionFilt.-AlleleAlign.ParalogFilt.PositionFilt.	-0.0065917	-0.0255987	0.01241532	0.9664137
ConsensusAlign.ParalogFilt.PositionUnfilt.-AlleleAlign.ParalogFilt.PositionFilt.	-0.0074324	-0.0264394	0.01157458	0.9363370
ConsensusAlign.ParalogUnfilt.PositionFilt.-AlleleAlign.ParalogFilt.PositionFilt.	0.04046389	0.0214569	0.05947088	**0.0000000**
ConsensusAlign.ParalogUnfilt.PositionUnfilt.-AlleleAlign.ParalogFilt.PositionFilt.	0.04115648	0.02214949	0.06016347	**0.0000000**
AlleleAlign.ParalogUnfilt.PositionFilt.-AlleleAlign.ParalogFilt.PositionUnfilt.	0.03073571	0.01190231	0.04956912	0.0000207
AlleleAlign.ParalogUnfilt.PositionUnfilt.-AlleleAlign.ParalogFilt.PositionUnfilt.	0.03347857	0.01464517	0.05231198	0.0000018
ConsensusAlign.ParalogFilt.PositionFilt.-AlleleAlign.ParalogFilt.PositionUnfilt.	-0.0075774	-0.0265844	0.01142961	0.9297382
ConsensusAlign.ParalogFilt.PositionUnfilt.-AlleleAlign.ParalogFilt.PositionUnfilt.	-0.0084181	-0.0274251	0.01058887	0.8825819
ConsensusAlign.ParalogUnfilt.PositionFilt.-AlleleAlign.ParalogFilt.PositionUnfilt.	0.03947817	0.02047119	0.05848516	**0.0000000**
ConsensusAlign.ParalogUnfilt.PositionUnfilt.-AlleleAlign.ParalogFilt.PositionUnfilt.	0.04017077	0.02116378	0.05917776	**0.0000000**
AlleleAlign.ParalogUnfilt.PositionUnfilt.-AlleleAlign.ParalogUnfilt.PositionFilt.	0.00274286	-0.0160905	0.02157626	0.9998539
ConsensusAlign.ParalogFilt.PositionFilt.-AlleleAlign.ParalogUnfilt.PositionFilt.	-0.0383131	-0.0573201	-0.0193061	**0.0000000**
ConsensusAlign.ParalogFilt.PositionUnfilt.-AlleleAlign.ParalogUnfilt.PositionFilt.	-0.0391538	-0.0581608	-0.0201469	**0.0000000**
ConsensusAlign.ParalogUnfilt.PositionFilt.-AlleleAlign.ParalogUnfilt.PositionFilt.	0.00874246	-0.0102645	0.02774945	0.8602767
ConsensusAlign.ParalogUnfilt.PositionUnfilt.-AlleleAlign.ParalogUnfilt.PositionFilt.	0.00943505	-0.0095719	0.02844204	0.8052769
ConsensusAlign.ParalogFilt.PositionFilt.-AlleleAlign.ParalogUnfilt.PositionUnfilt.	-0.041056	-0.0600629	-0.022049	**0.0000000**
ConsensusAlign.ParalogFilt.PositionUnfilt.-AlleleAlign.ParalogUnfilt.PositionUnfilt.	-0.0418967	-0.0609037	-0.0228897	**0.0000000**
ConsensusAlign.ParalogUnfilt.PositionFilt.-AlleleAlign.ParalogUnfilt.PositionUnfilt.	0.0059996	-0.0130074	0.02500659	0.9802237
ConsensusAlign.ParalogUnfilt.PositionUnfilt.-AlleleAlign.ParalogUnfilt.PositionUnfilt.	0.0066922	-0.0123148	0.02569918	0.9635074
ConsensusAlign.ParalogFilt.PositionUnfilt.-ConsensusAlign.ParalogFilt.PositionFilt.	-0.0008407	-0.0200197	0.01833826	1.0000000
ConsensusAlign.ParalogUnfilt.PositionFilt.-ConsensusAlign.ParalogFilt.PositionFilt.	0.04705556	0.02787656	0.06623456	**0.0000000**
ConsensusAlign.ParalogUnfilt.PositionUnfilt.-ConsensusAlign.ParalogFilt.PositionFilt.	0.04774815	0.02856915	0.06692715	**0.0000000**
ConsensusAlign.ParalogUnfilt.PositionFilt.-ConsensusAlign.ParalogFilt.PositionUnfilt.	0.0478963	0.0287173	0.0670753	**0.0000000**
ConsensusAlign.ParalogUnfilt.PositionUnfilt.-ConsensusAlign.ParalogFilt.PositionUnfilt.	0.04858889	0.02940989	0.06776789	**0.0000000**
ConsensusAlign.ParalogUnfilt.PositionUnfilt.-ConsensusAlign.ParalogUnfilt.PositionFilt.	0.00069259	-0.0184864	0.01987159	1.0000000

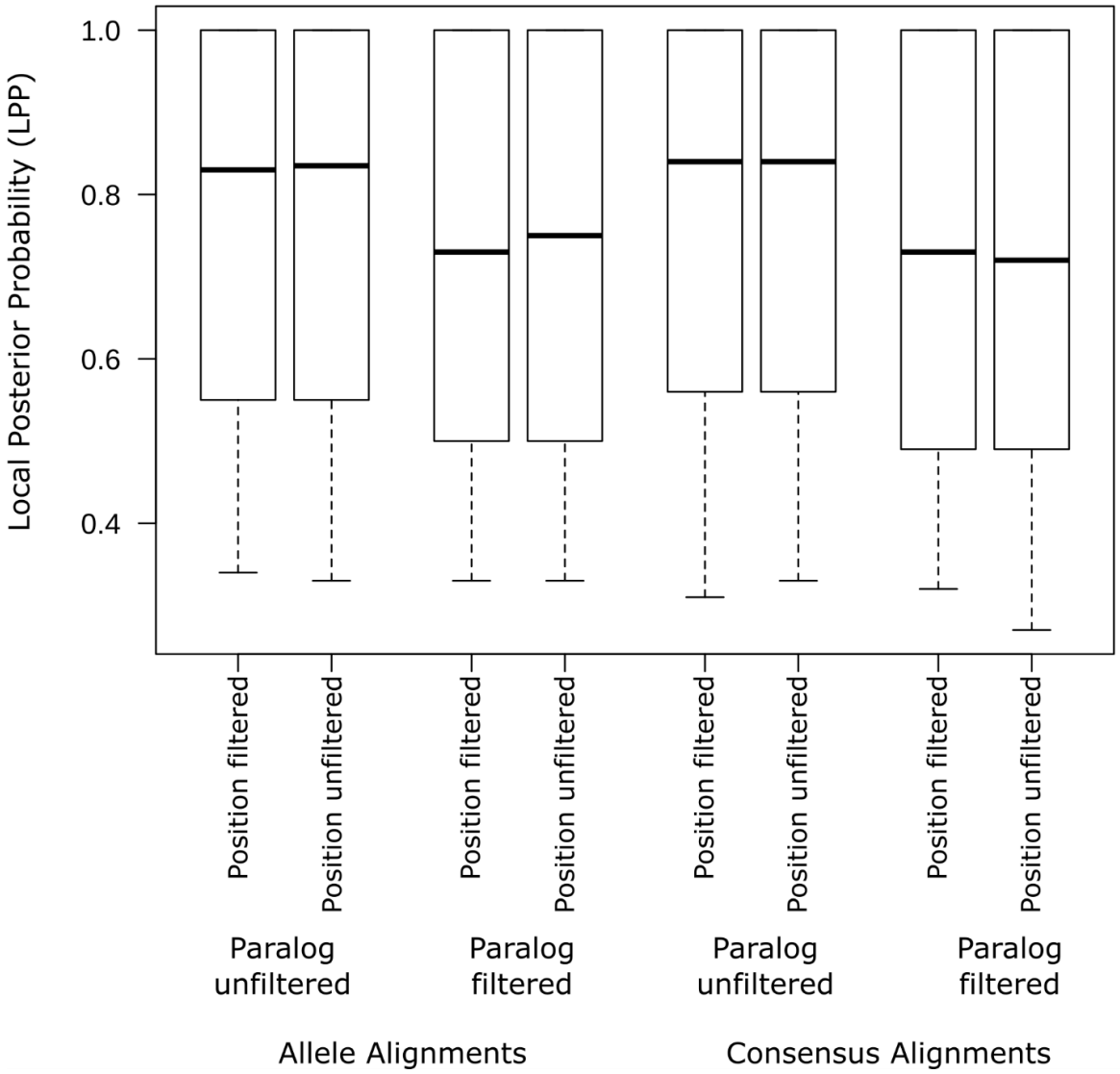


Figure S1. Boxplot showing performance (mean Local Posterior Probability of the bootstrap ASTRAL trees) of the different phasing and filtering schemes.

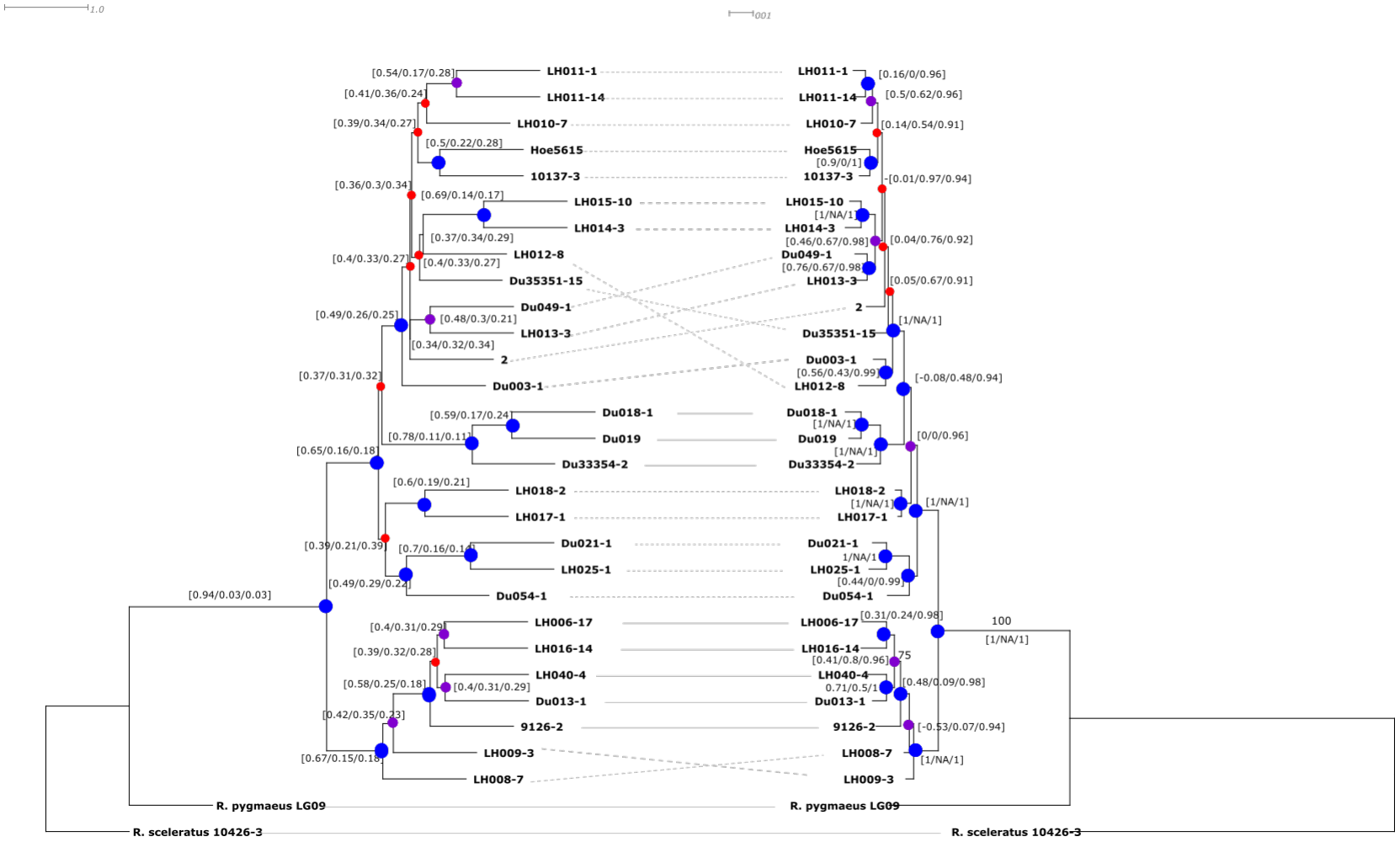


Figure S2. Phylogenetic trees drawn opposite to each other (tanglegram) from the coalescent-based Astral species tree (left) and concatenated maximum likelihood analysis (right), based on the consensus dataset after applying position and paralog filtering. Continuous lines indicate accession with congruent position, whereas dashed lines indicate incongruence. Circles on the nodes of the trees refer to the bootstrap support values. Blue bigger circles are for support values above 95; purples medium-sized circles are for values between 75 and 95; small red circles are for bootstraps below 75. Numbers in proximity of nodes indicate ASTRAL quartet supports (left) and results from quartet sampling on the maximum likelihood tree (right).

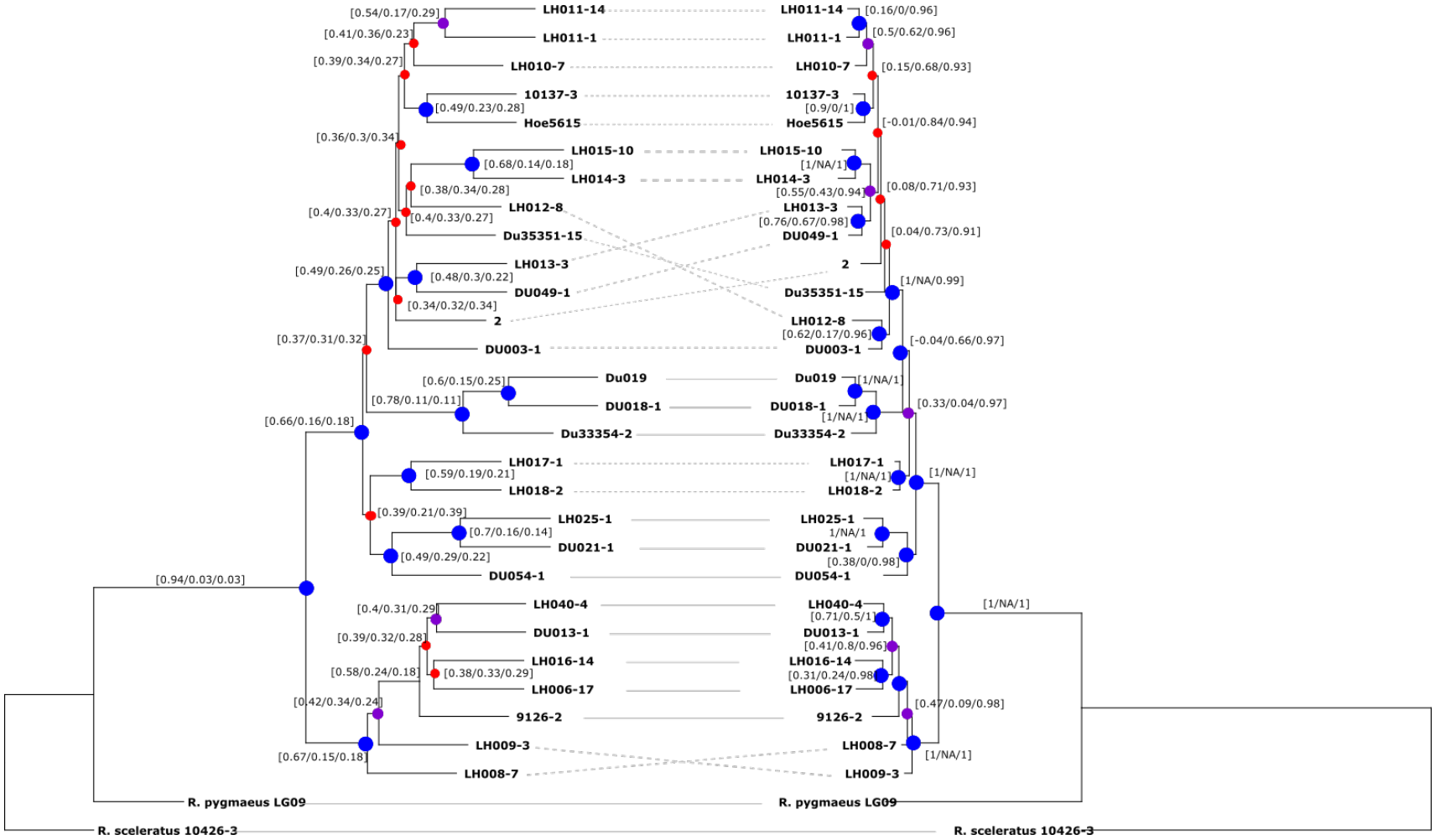


Figure S3. Phylogenetic trees drawn opposite to each other (tanglegram) from the coalescent-based Astral species tree (left) and concatenated maximum likelihood analysis (right), based on the consensus dataset after applying and paralog filtering. Continuous lines indicate accession with congruent position, whereas dashed lines indicate incongruence. Circles on the nodes of the trees refer to the bootstrap support values. Blue bigger circles are for support values above 95; purples medium-sized circles are for values between 75 and 95; small red circles are for bootstraps below 75. Numbers in proximity of nodes indicate ASTRAL quartet supports (left) and results from quartet sampling on the maximum likelihood tree (right).

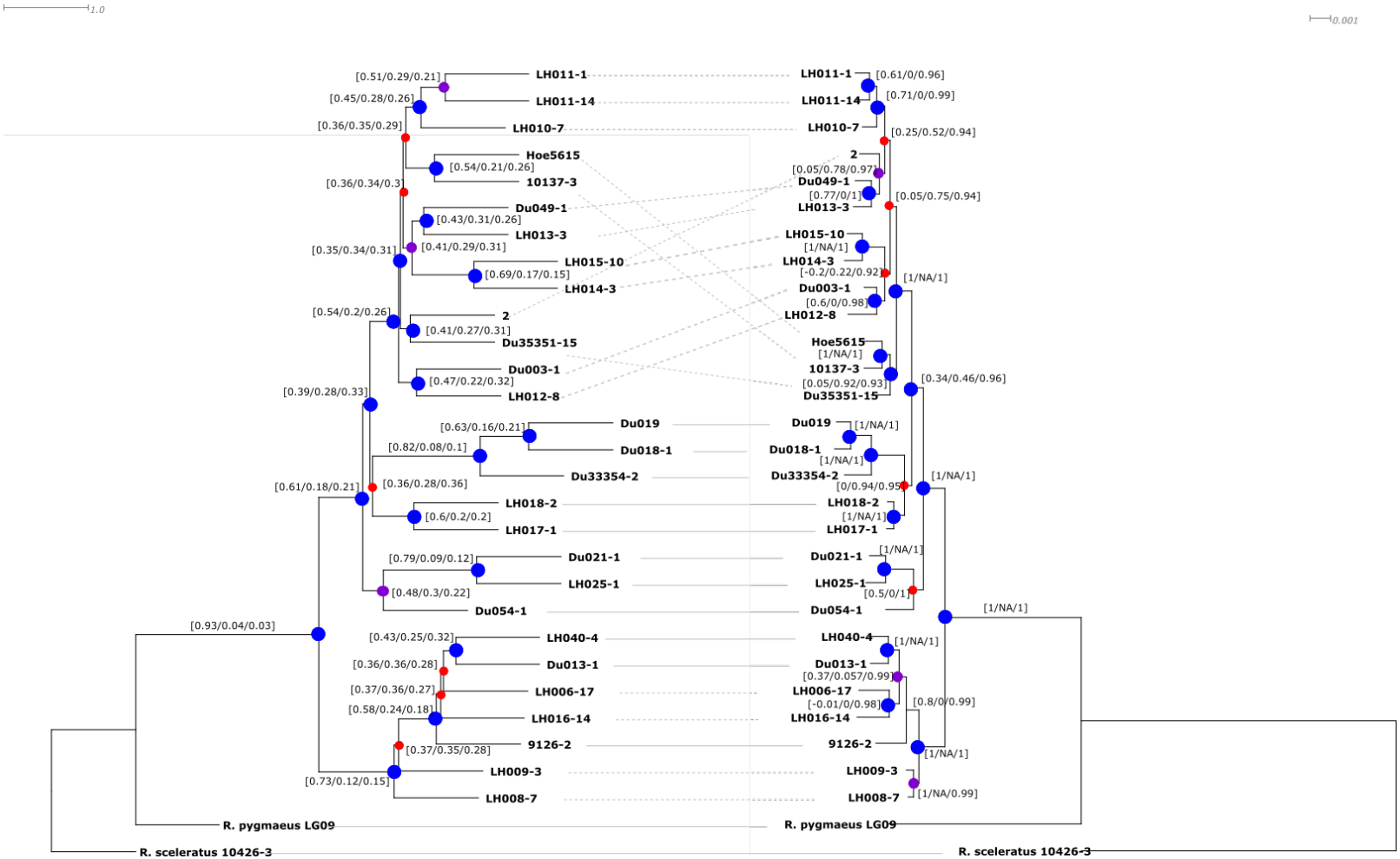


Figure S4. Phylogenetic trees drawn opposite to each other (tanglegram) from the coalescent-based Astral species tree (left) and concatenated maximum likelihood analysis (right), based on the consensus dataset after applying position filtering and without paralog filtering. Continuous lines indicate accession with congruent position, whereas dashed lines indicate incongruence. Circles on the nodes of the trees refer to the bootstrap support values. Blue bigger circles are for support values above 95; purples medium-sized circles are for values between 75 and 95; small red circles are for bootstraps below 75. Numbers in proximity of nodes indicate ASTRAL quartet supports (left) and results from quartet sampling on the maximum likelihood tree (right).

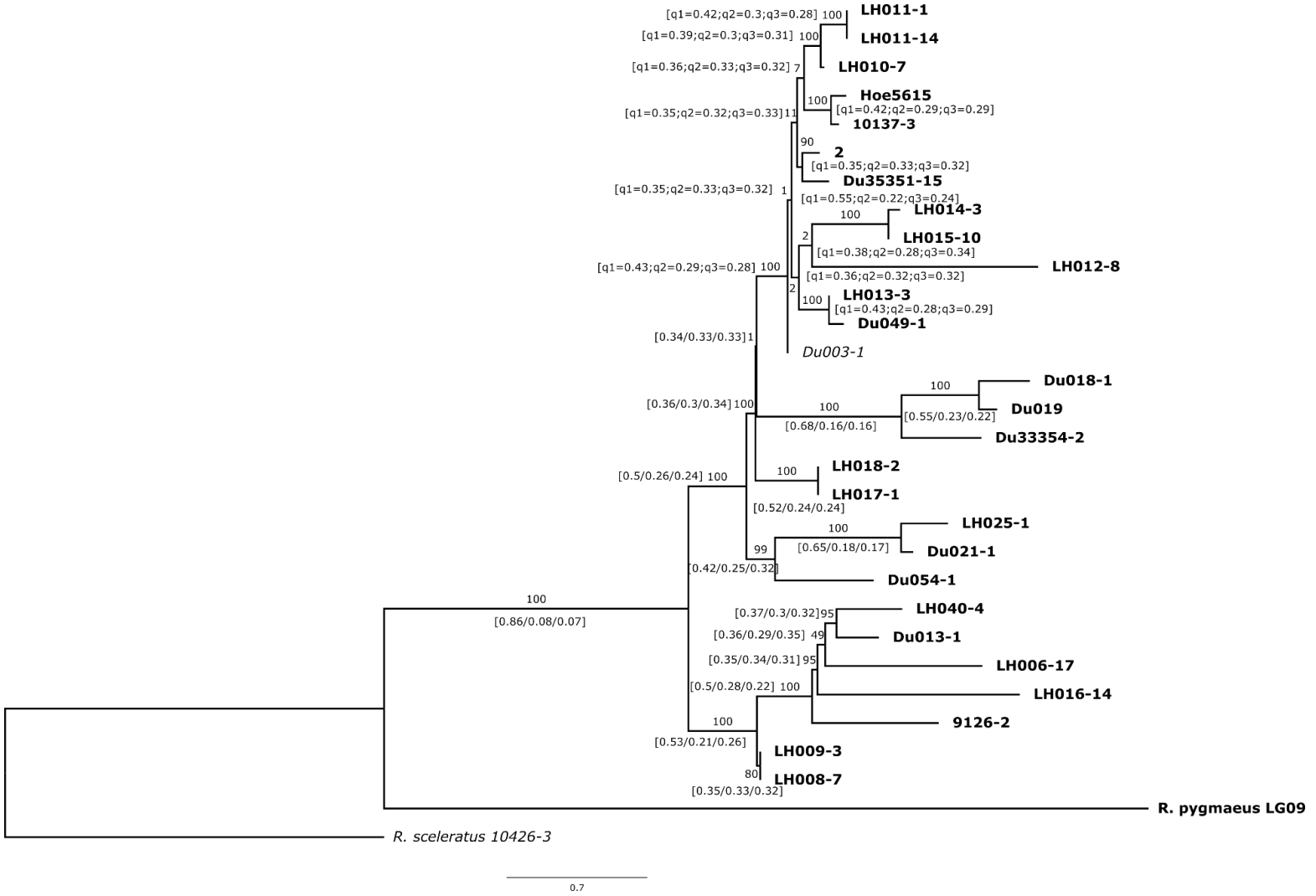


Figure S5. Astral species tree based on the allele dataset after applying position and paralog filtering. Numbers above branches are bootstrap values (only above 75 are shown). Numbers in square brackets in proximity of nodes indicate ASTRAL quartet supports for the main/first alternative/second alternative topologies.

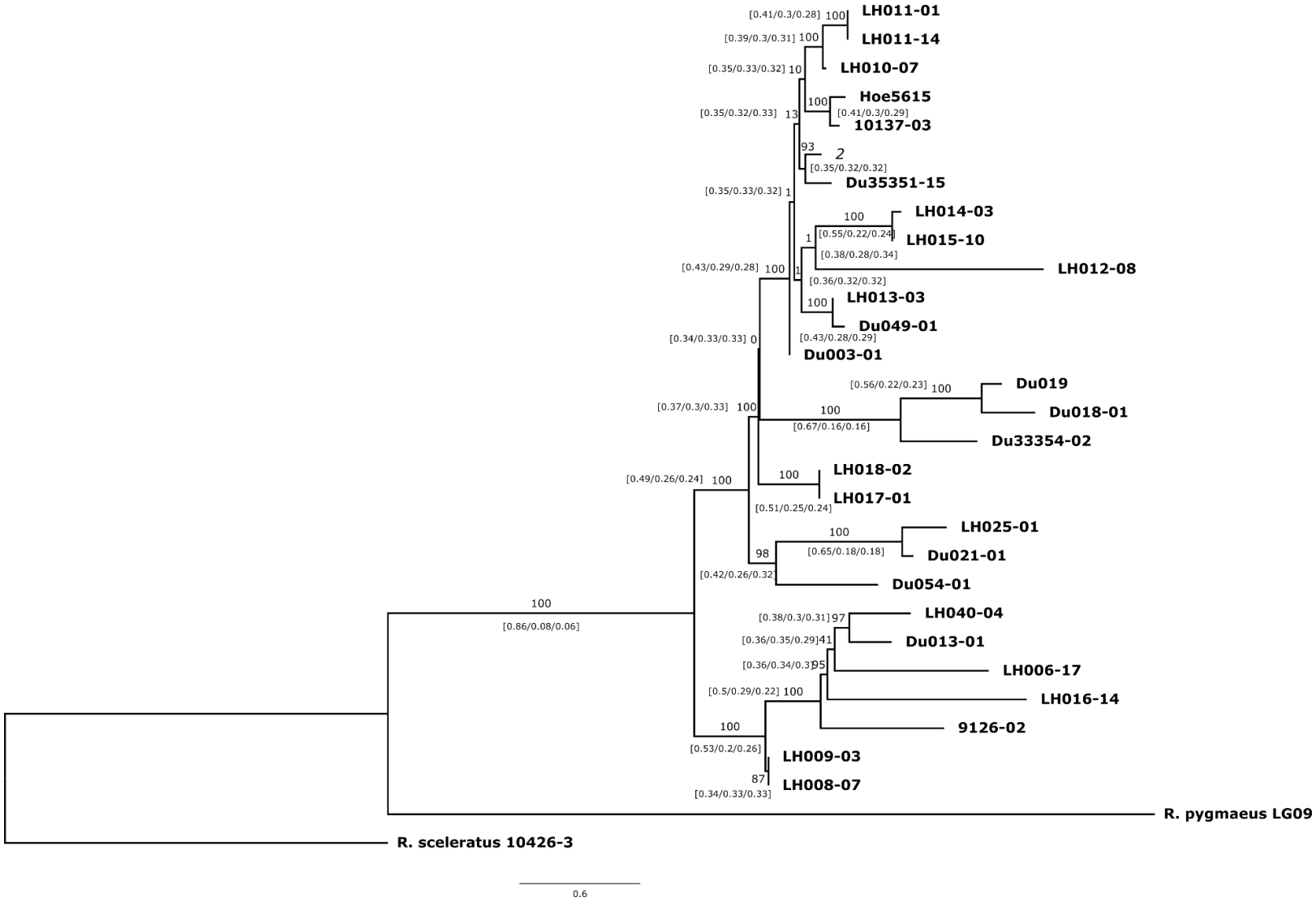


Figure S6. Astral species tree based on the allele dataset after applying paralog filtering. Numbers above branches are bootstrap values (only above 75 are shown). Numbers in square brackets in proximity of nodes indicate ASTRAL quartet supports for the main/first alternative/second alternative topologies.

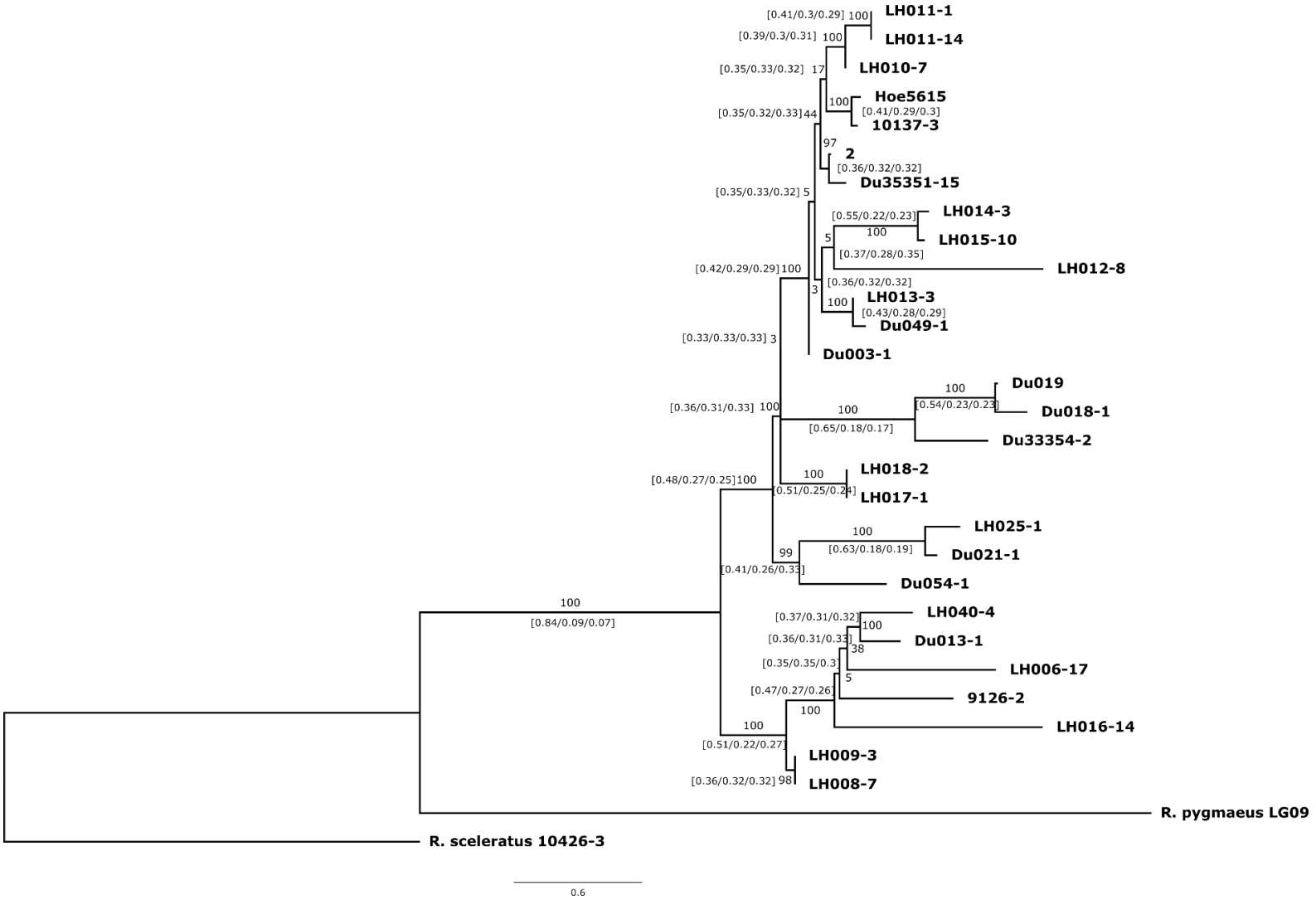
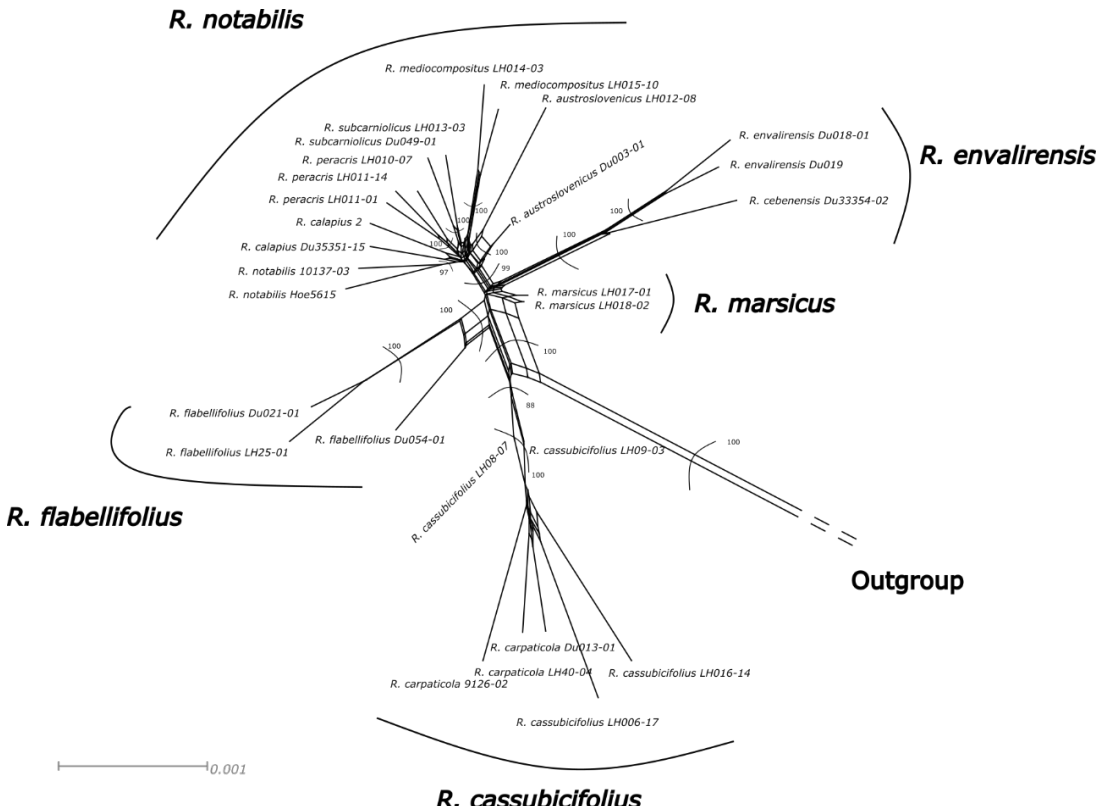
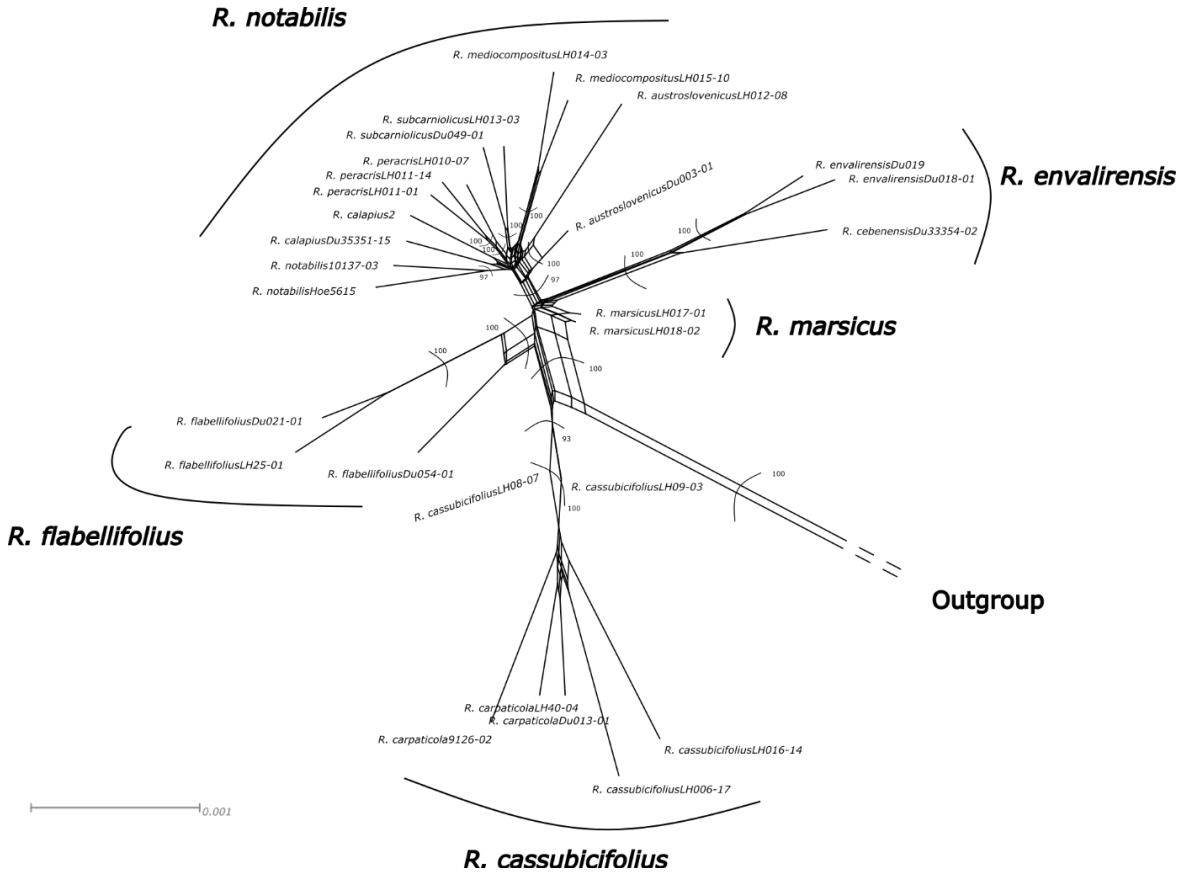


Figure S7. Astral species tree based on the allele dataset after applying position filtering and without paralog filtering. Numbers above branches are bootstrap values (only above 75 are shown). Numbers in square brackets in proximity of nodes indicate ASTRAL quartet supports for the main/first alternative/second alternative topologies.

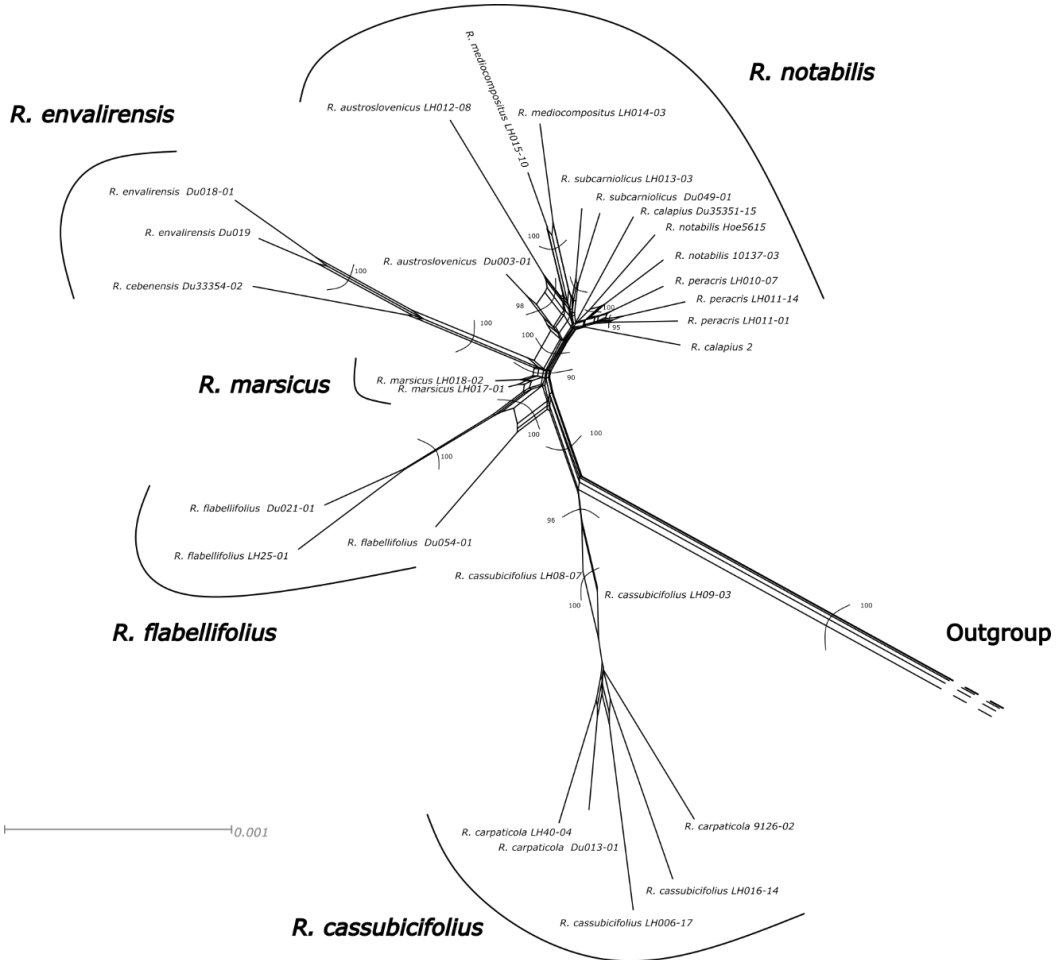
a)


R. notabilis
R. mediocompositus LH014-03
R. mediocompositus LH015-10
R. austroslovenicus LH012-08
R. subcarniolicus LH013-03
R. subcarniolicus Du049-01
R. peracris LH010-07
R. peracris LH011-14
R. peracris LH011-01
R. calapius 2
R. calapius Du35351-15
R. notabilis 10137-03
R. notabilis Hoe5615
R. austroslovenicus Du003-01
R. envalirensis Du018-01
R. envalirensis Du019
R. cebenensis Du33354-02
R. envalirensis
R. marsicus LH017-01
R. marsicus LH018-02
R. marsicus
R. flabellifolius Du021-01
R. flabellifolius Du054-01
R. flabellifolius LH25-01
R. cassubicifolius LH08-07
R. cassubicifolius LH09-03
R. flabellifolius
Outgroup
R. carpaticola Du013-01
R. carpaticola LH40-04
R. carpaticola 9126-02
R. cassubicifolius LH016-14
R. cassubicifolius LH006-17
0.001
R. cassubicifolius

b)


R. notabilis
R. mediocompositusLH014-03
R. mediocompositusLH015-10
R. austroslovenicusLH012-08
R. subcarniolicusLH013-03
R. subcarniolicusDu049-01
R. peracrisLH010-07
R. peracrisLH011-14
R. peracrisLH011-01
R. calapius2
R. calapiusDu35351-15
R. notabilis10137-03
R. notabilisHoe5615
R. austroslovenicusDu003-01
R. envalirensisDu019
R. envalirensisDu018-01
R. cebenensisDu33354-02
R. envalirensis
R. marsicusLH017-01
R. marsicusLH018-02
R. marsicus
R. flabellifoliusDu021-01
R. flabellifoliusLH25-01
R. flabellifoliusDu054-01
R. cassubicifoliusLH08-07
R. cassubicifoliusLH09-03
R. flabellifolius
Outgroup
R. carpaticolaLH40-04
R. carpaticolaDu013-01
R. carpaticola9126-02
R. cassubicifoliusLH016-14
R. cassubicifoliusLH006-17
0.001
R. cassubicifolius

c)


R. notabilis
R. envalirensis
R. mediocompositus LH015-10
R. austrosloveniicus LH012-08
R. mediocompositus LH014-03
R. subcarniolicus LH013-03
R. subcarniolicus Du049-01
R. calapius Du35351-15
R. notabilis Hoe5615
R. notabilis 10137-03
R. peracris LH010-07
R. peracris LH011-14
R. peracris LH011-01
R. calapius 2
R. envalirensis Du018-01
R. envalirensis Du019
R. cebenensis Du33354-02
R. austrosloveniicus Du003-01
R. marsicus
R. marsicus LH018-02
R. marsicus LH017-01
R. flabellifolius Du021-01
R. flabellifolius LH25-01
R. flabellifolius Du054-01
R. cassubicifolius LH08-07
R. cassubicifolius LH09-03
R. flabellifolius
Outgroup
0.001
R. carpaticola LH40-04
R. carpaticola Du013-01
R. carpaticola 9126-02
R. cassubicifolius LH016-14
R. cassubicifolius LH006-17
R. cassubicifolius

d)

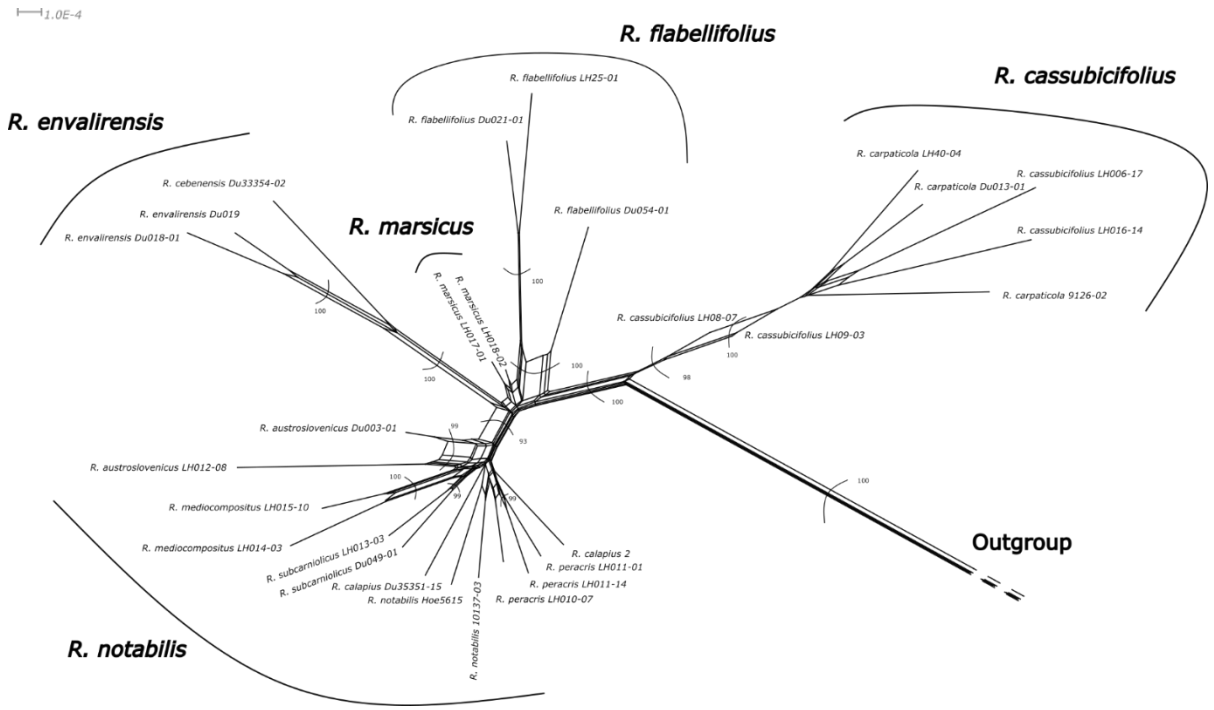


Figure S8. Neighbor-net analysis based on distances of a General Time Reversible (GTR) model with estimated site frequencies and ML. Bootstrap values above 90 for major clusters are shown. a) Dataset unfiltered for paralogs and position filtered; b) unfiltered for both paralogs and positions; c) filtered for both paralogs and positions; d) paralog filtered and position unfiltered.

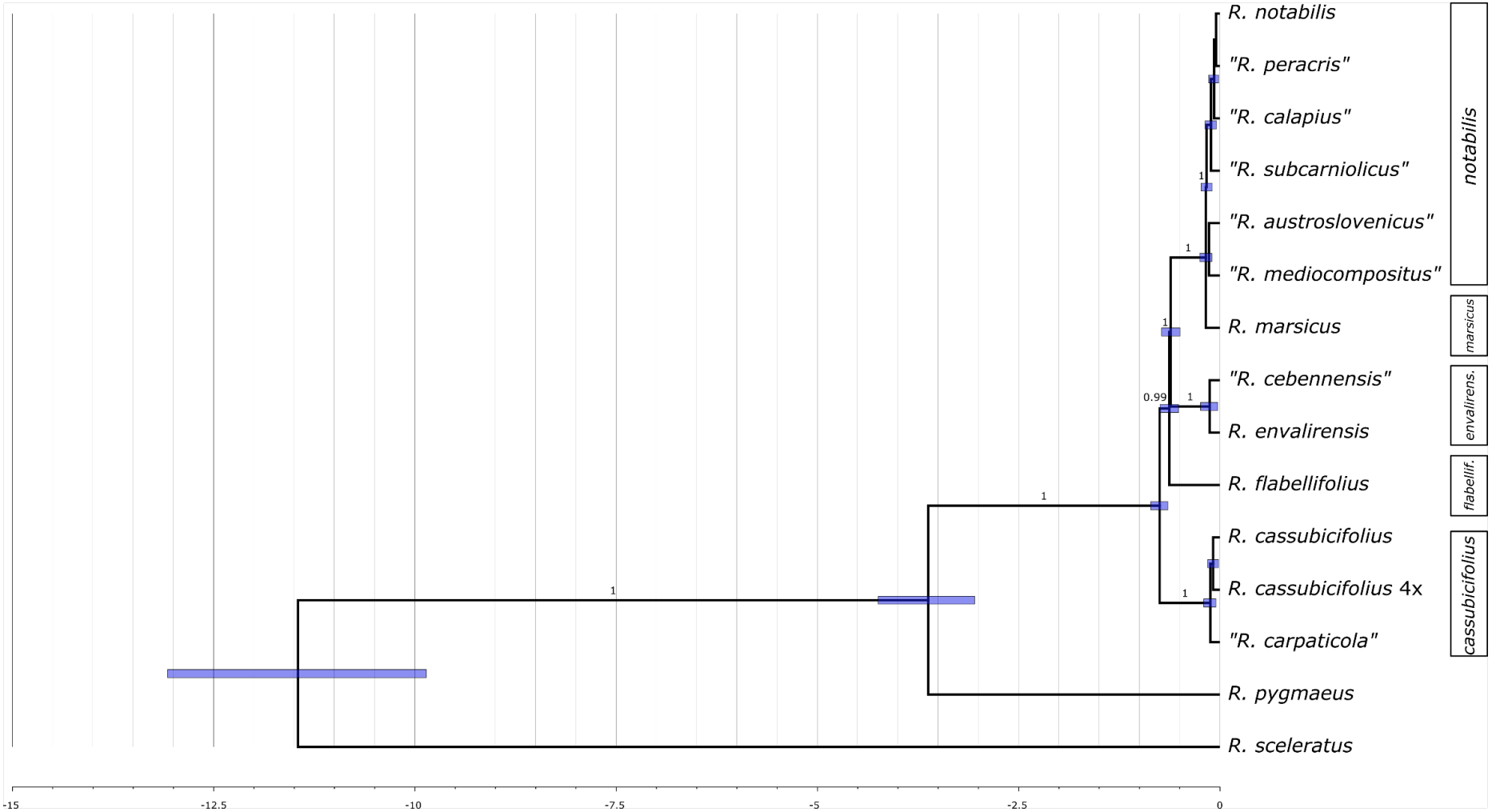


Figure S9. Maximum clade credibility tree estimated in *BEAST using the 50 most informative loci from the consensus dataset and including outgroup samples. Blue bars indicate 95% highest posterior density (HPD) intervals of the age estimate. Numbers above branches indicate Bayesian posterior probabilities. Only values above 0.7 are shown. On the right, bars indicate the current nomenclature following Karbstein et al. (2020).

Table S9. Results of the BEAST analyses when using separately calibration point(s) or clock rate, for both the datasets including and excluding outgroup accessions. In bold are estimates of the main analyses presented in Fig. 4 (without outgroup) and Fig. S9 (including outgroup). Estimates are expressed in ka and number in brackets refers to 95% HPD (hight posterior density).

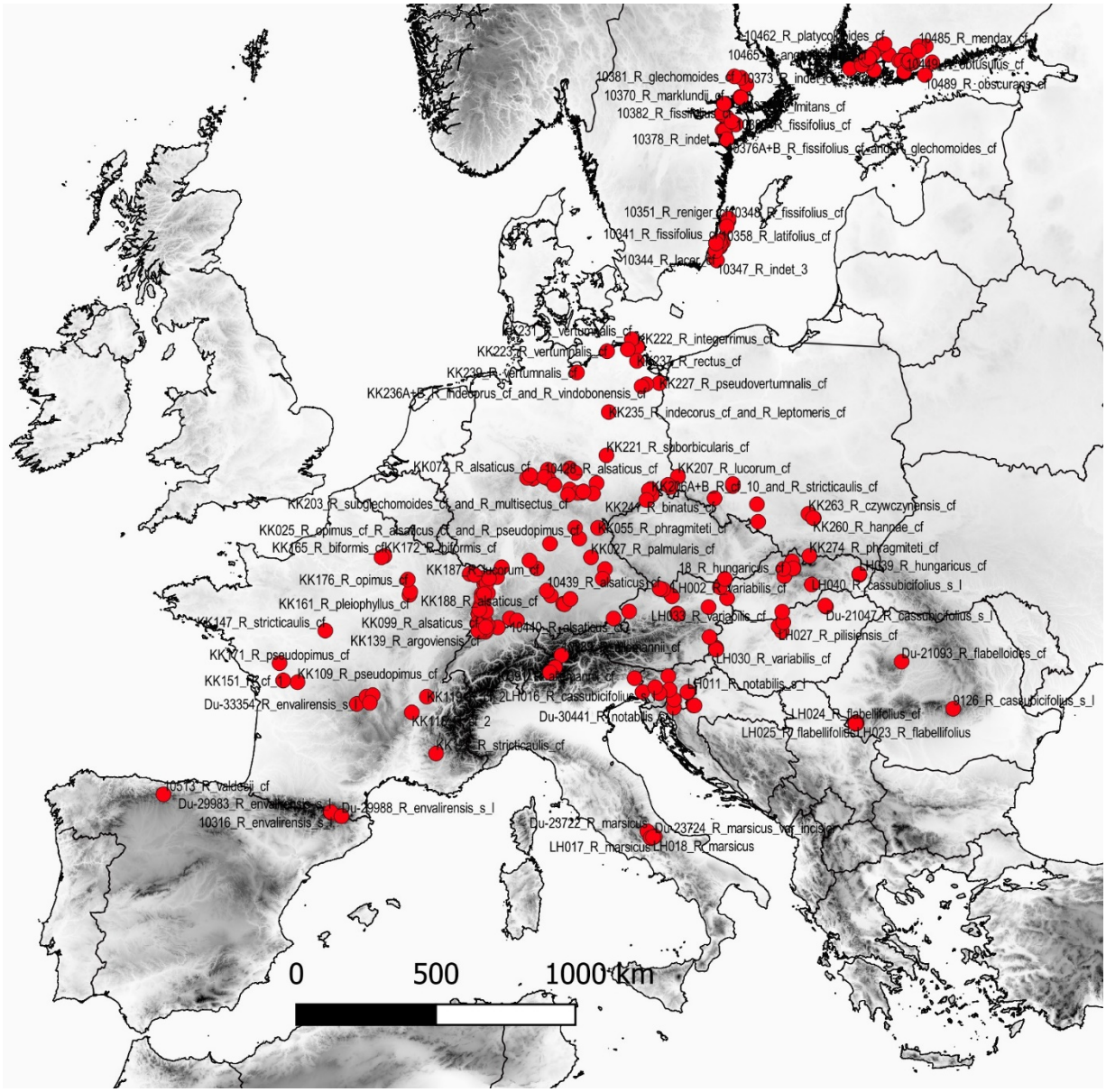
		Crown age *R. auricomus*	**Stem age *R. envalirensis***	**Crown age *R. notabilis***
Only Ingroup	**Calibr. and clock rate**	**733.52 (639.23-829.12)**	**683.01 (577.10-780.56)**	**175.88 (103.14-257.52)**
	Only calibration	727.58 (708.71-746.71)	505.90 (380.14-634.43)	154.56 (96.84-219.36)
	Only clock rate	738.91 (625.38-854.98)	688.73 (576.96-796.12)	170.54 (106.49-239.32)
With Outgroup	**Calibr. and clock rate**	**749.7 (645.06-862.43)**	**611.26 (494.93-725.23)**	**166.51 (95.33-234.2)**
	Only calibrations	732.63 (713.78-751.53)	698.24 (647.08-742.05)	286.82 (178.33-417.58)
	Only clock rate	732.96 (713.43-750.50)	701.11 (647.4-740.38)	303.27 (200.52-461.7)

Table S10. Results from the Likelihood Ratio Test (LRT) of different nested models.

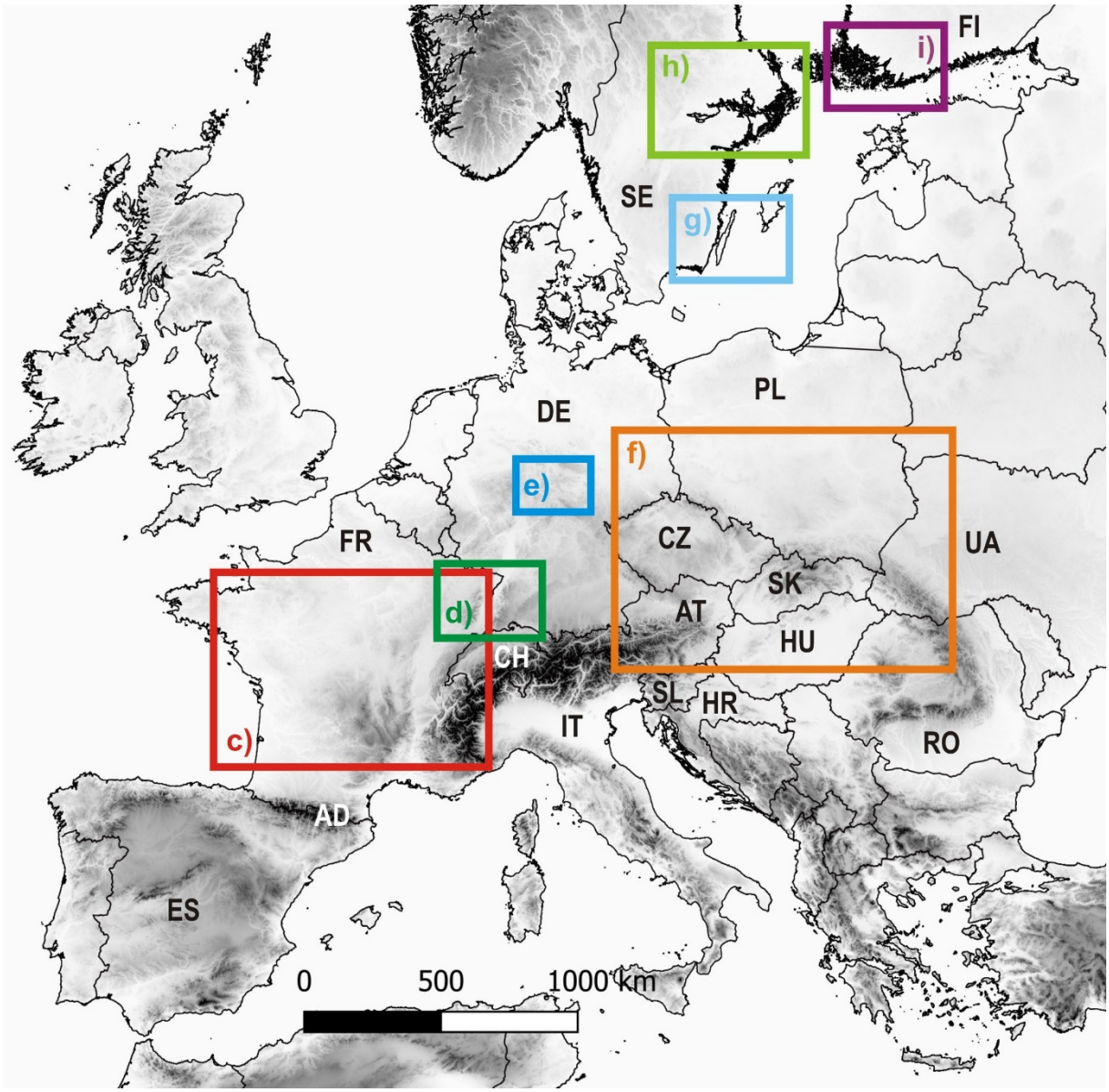
alt	null	LnLalt	LnLnull	DFalt	DFnull	DF	Dstatistic	pval	test	tail	AIC1	AIC2	AICwt1	AICwt2	AICweight ratio model1	AICweight ratio model2
DEC+x	DEC	-22.756713	-26.228416	3	2	1	6.94340487	0.00841288	chi-squared	one-tailed	51.51343	56.45683	0.9221341	0.07786591	11.84259	0.08444098
DEC+j	DEC	-24.273943	-26.228416	3	2	1	3.90894538	0.04802973	chi-squared	one-tailed	54.54789	56.45683	0.7220138	0.2779862	2.597301	0.3850151
DEC+x+j	DEC+x	-22.597258	-22.756713	4	3	1	0.31891002	0.57226342	chi-squared	one-tailed	53.19452	51.51343	0.30142	0.69858	0.4314753	2.31763
DIVA+x	DIVA	-18.306908	-21.475135	3	2	1	6.3364542	0.01182811	chi-squared	one-tailed	42.61382	46.95027	0.8973598	0.1026402	8.74277	0.1143802
DIVA+j	DIVA	-21.473241	-21.475135	3	2	1	0.00378824	0.95092226	chi-squared	one-tailed	48.94648	46.95027	0.269314	0.730686	0.3685769	2.713138
DIVA+x+j	DIVA+x	-17.718453	-18.306908	4	3	1	1.17690952	0.27798605	chi-squared	one-tailed	43.43691	42.61382	0.3985417	0.6014583	0.6626255	1.509148
BAYAREA+x	BAYAREA	-29.506118	-31.871805	3	2	1	4.73137481	0.0296172	chi-squared	one-tailed	65.01224	67.74361	0.7966825	0.2033175	3.918416	0.2552052
BAYAREA+j	BAYAREA	-26.981695	-31.871805	3	2	1	9.78021918	0.00176399	chi-squared	one-tailed	59.96339	67.74361	0.9799664	0.02003356	48.91625	0.02044311
BAYAREA+x+j	BAYAREA+x	-25.354431	-29.506118	4	3	1	8.30337279	0.00395715	chi-squared	one-tailed	58.70886	65.01224	0.9589751	0.04102488	23.37545	0.04277992

Chapter 4

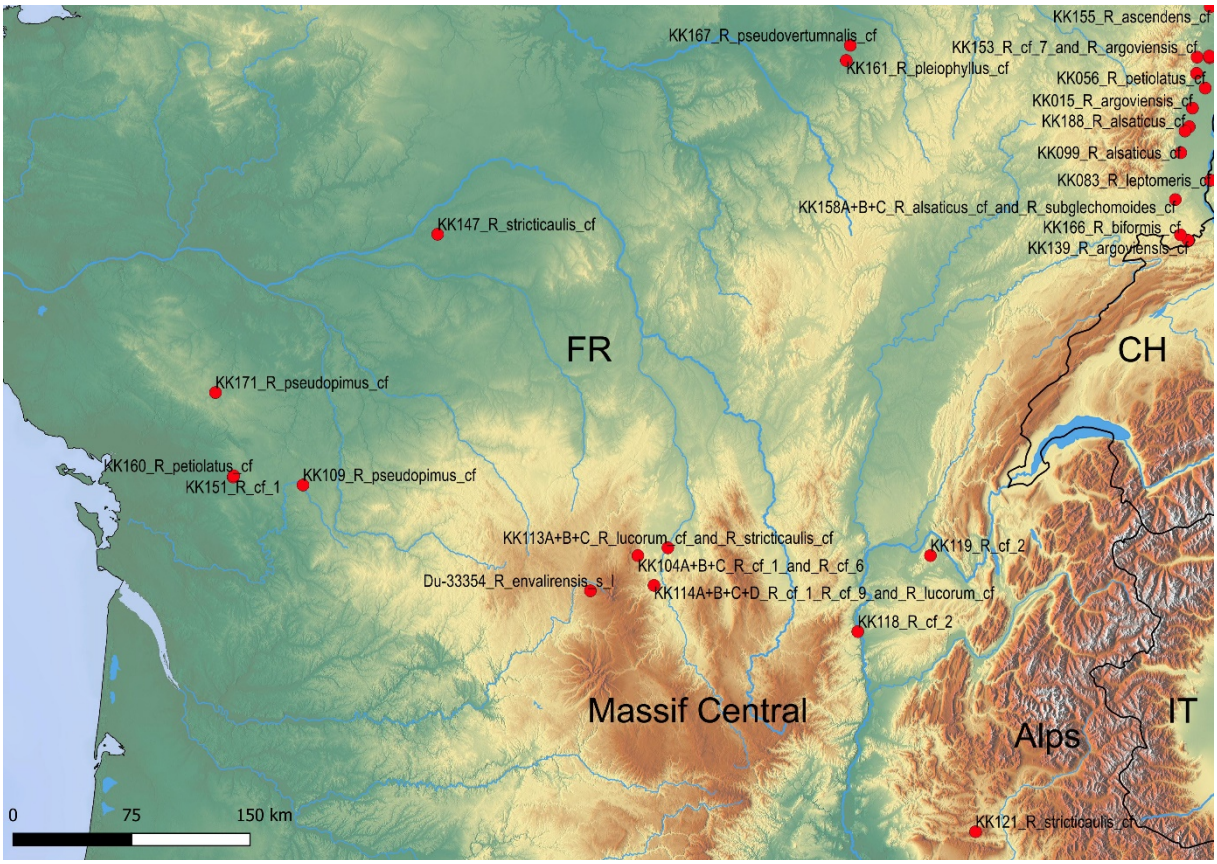
a)


10462_R_platycoloides_cf
10485_R_mendax_cf
10381_R_glechomoides_cf
10370_R_marklundii_cf
10382_R_fissifolius_cf
10378_R_indet
10489_R_obscurans_cf
10376A+B_R_fissifolius_cf_and_R_glechomoides_cf
10351_R_reniger_cf
10348_R_fissifolius_cf
10341_R_fissifolius_cf
10358_R_latifolius_cf
10344_R_lacer_cf
10347_R_indet_3
KK222_R_integerrimus_cf
KK237_R_rectus_cf
KK227_R_pseudovertumnalis_cf
KK235_R_indecorus_cf_and_R_leptomeris_cf
KK221_R_sphorbicularis_cf
KK207_R_lucorum_cf
KK205A+B_R_cf_10_and_R_stricticaulis_cf
KK263_R_czywczynensis_cf
KK260_R_hannae_cf
KK274_R_phragmiteti_cf
KK039_R_hungaricus_cf
LH040_R_cassubicifolius_s_l
Du-21047_R_cassubicifolius_s_l
LH027_R_pilisiensis_cf
Du-21093_R_flabelloides_cf
LH030_R_variabilis_cf
LH011_R_notabilis_s_l
9126_R_cassubicifolius_s_l
LH024_R_flabellifolius_cf
LH025_R_flabellifolius
LH023_R_flabellifolius
LH016_R_cassubicifolius_s_l
Du-30441_R_notabilis_cf
10513_R_valdesii_cf
Du-29983_R_envalirensis_s_l
Du-29988_R_envalirensis_s_l
10316_R_envalirensis_s_l
Du-23722_R_marsicus
Du-23724_R_marsicus_var_incisus
LH017_R_marsicus
LH018_R_marsicus
0
500
1000 km

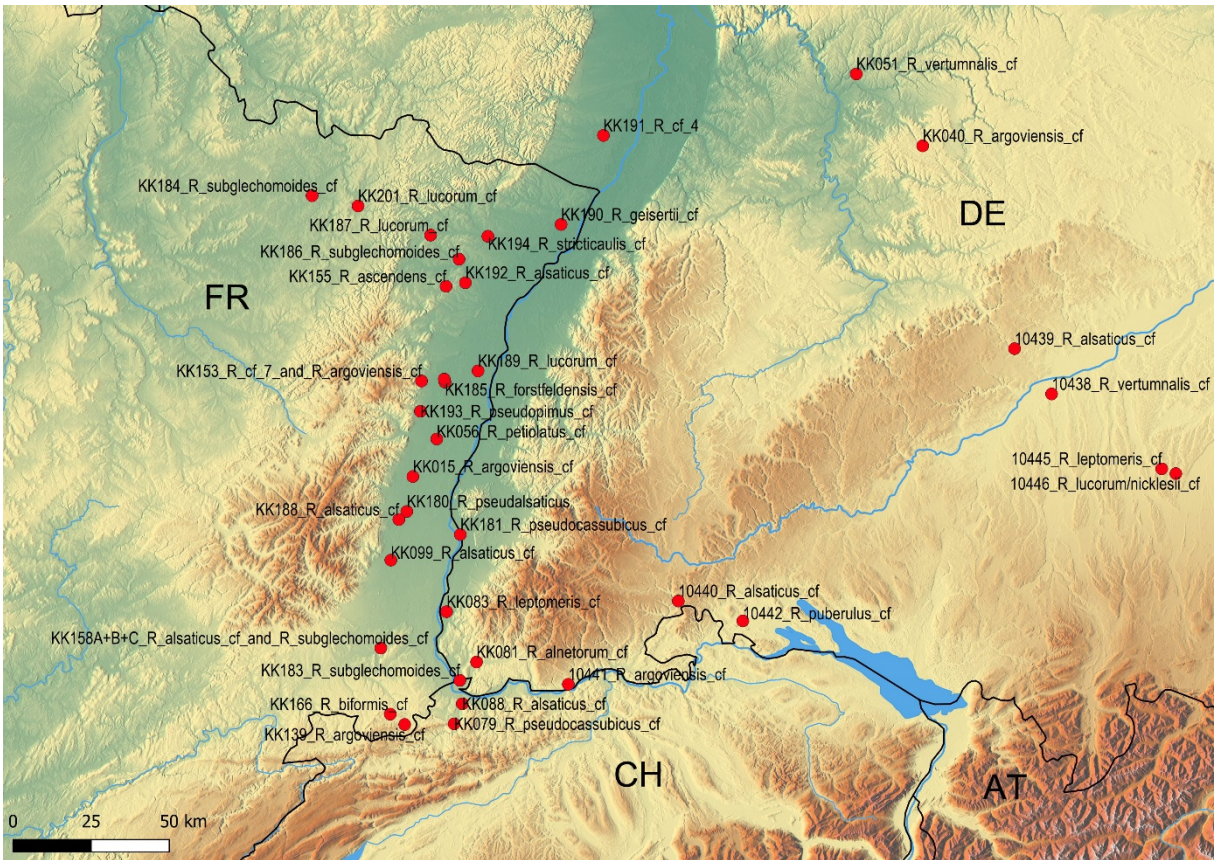
b)


FI
i)
h)
SE
g)
PL
DE
f)
e)
FR
CZ
UA
SK
AT
d)
HU
CH
SL
HR
c)
IT
RO
AD
ES
0
500
1000 km

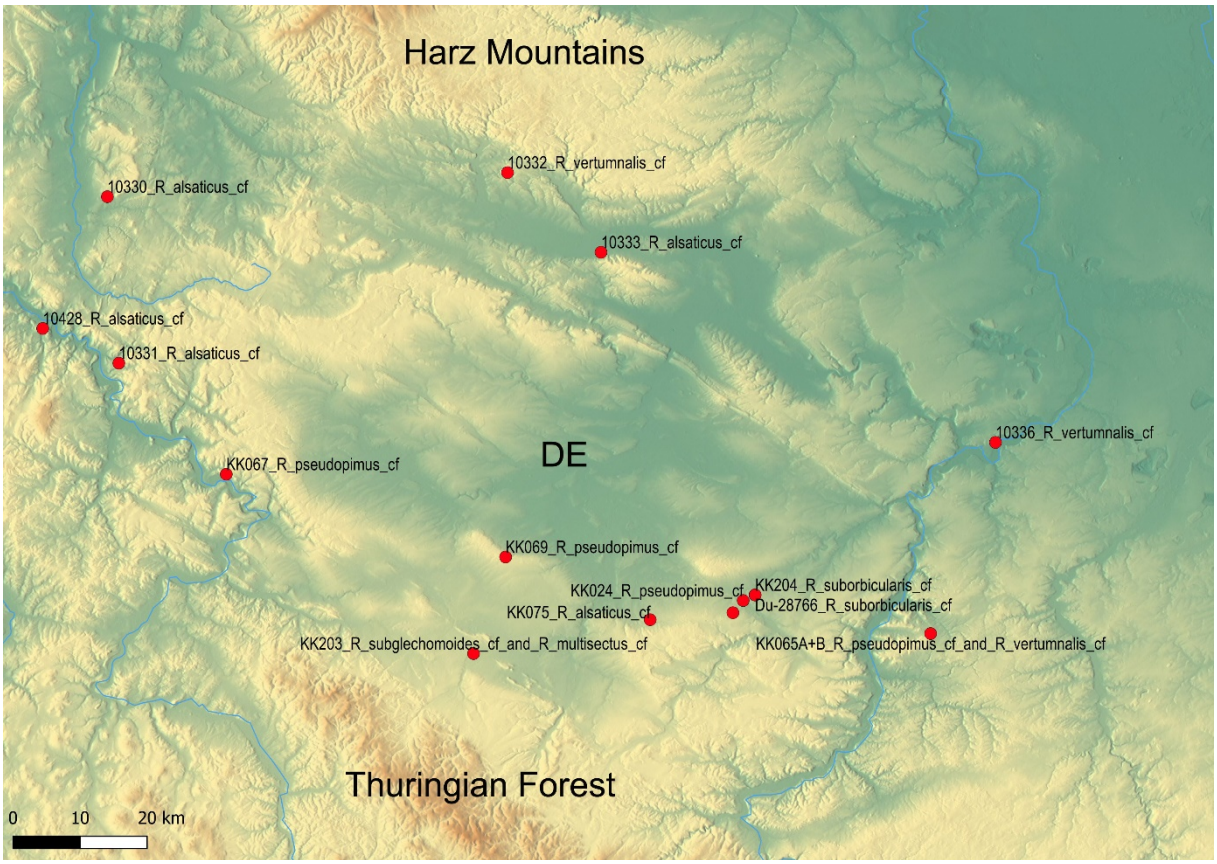
c)


KK155_R_ascendens_cf
KK167_R_pseudovertumnalis_cf
KK153_R_cf_7_and_R_argoviensis_cf
KK161_R_pleiophyllus_cf
KK056_R_petiolatus_cf
KK015_R_argoviensis_cf
KK188_R_alsaticus_cf
KK099_R_alsaticus_cf
KK083_R_leptomeris_cf
KK158A+B+C_R_alsaticus_cf_and_R_subglechomoides_cf
KK166_R_biformis_cf
KK139_R_argoviensis_cf
KK147_R_stricticaulis_cf
FR
CH
KK171_R_pseudopimus_cf
KK160_R_petiolatus_cf
KK151_R_cf_1
KK109_R_pseudopimus_cf
KK113A+B+C_R_lucorum_cf_and_R_stricticaulis_cf
KK119_R_cf_2
KK104A+B+C_R_cf_1_and_R_cf_6
Du-33354_R_envalirensis_s_l
KK114A+B+C+D_R_cf_1_R_cf_9_and_R_lucorum_cf
KK118_R_cf_2
Massif Central
Alps
IT
0
75
150 km
KK121_R_stricticaulis_cf

d)


KK051_R_vertumnalis_cf
KK191_R_cf_4
KK040_R_argoviensis_cf
KK184_R_subglechomoides_cf
KK201_R_lucorum_cf
KK187_R_lucorum_cf
KK190_R_geisertii_cf
KK194_R_stricticaulis_cf
KK186_R_subglechomoides_cf
KK155_R_ascendens_cf
KK192_R_alsaticus_cf
DE
FR
10439_R_alsaticus_cf
KK153_R_cf_7_and_R_argoviensis_cf
KK189_R_lucorum_cf
KK185_R_forstfeldensis_cf
10438_R_vertumnalis_cf
KK193_R_pseudopinus_cf
KK056_R_petiolatus_cf
10445_R_leptomeris_cf
10446_R_lucorum/nicklesii_cf
KK015_R_argoviensis_cf
KK188_R_alsaticus_cf
KK180_R_pseudalsaticus
KK181_R_pseudocassubicus_cf
KK099_R_alsaticus_cf
10440_R_alsaticus_cf
KK083_R_leptomeris_cf
10442_R_puberulus_cf
KK158A+B+C_R_alsaticus_cf_and_R_subglechomoides_cf
KK081_R_alnetorum_cf
KK183_R_subglechomoides_cf
10441_R_argoviensis_cf
KK166_R_biformis_cf
KK088_R_alsaticus_cf
KK079_R_pseudocassubicus_cf
KK139_R_argoviensis_cf
CH
AT
0
25
50 km

e)


Harz Mountains
10332_R_vertumnalis_cf
10330_R_alsaticus_cf
10333_R_alsaticus_cf
10428_R_alsaticus_cf
10331_R_alsaticus_cf
10336_R_vertumnalis_cf
DE
KK067_R_pseudopimus_cf
KK069_R_pseudopimus_cf
KK024_R_pseudopimus_cf
KK204_R_suborbicularis_cf
Du-28766_R_suborbicularis_cf
KK075_R_alsaticus_cf
KK203_R_subglechomoides_cf_and_R_multisectus_cf
KK065A+B_R_pseudopimus_cf_and_R_vertumnalis_cf
Thuringian Forest
0
10
20 km

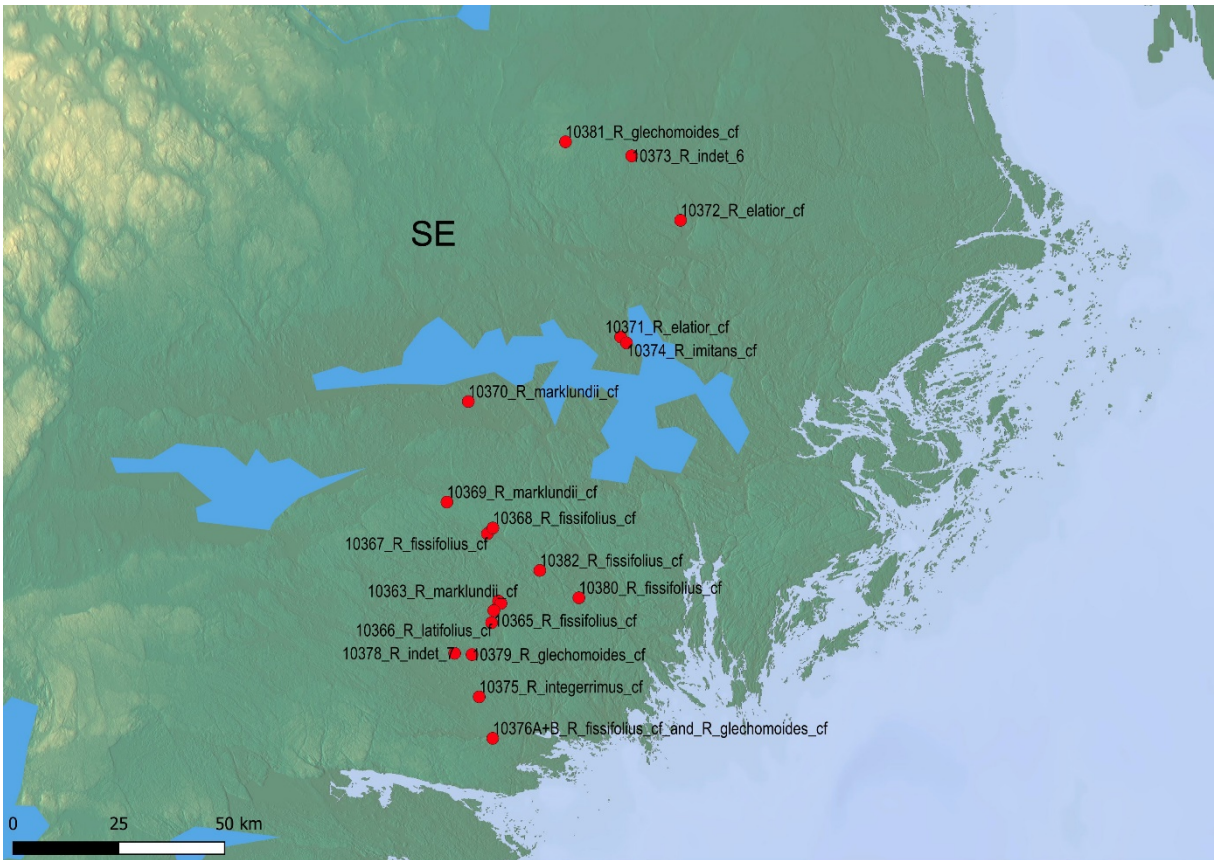
f)


DE
PL
CZ
Bohemian Massif
UA
Carpathians
AT
SK
HU
RO
Alps
KK207_R_lucorum_cf
KK261_R_argoviensis_cf
KK205_R_cf_5
KK208_R_aemulans_cf
KK255_R_vertumnalis_cf
KK252_R_phragmiteti_cf
KK208A+B_R_cf_10_and_R_stricticaulis_cf
KK211_R_phragmiteti_cf
KK241_R_binatus_cf
KK266_R_bayerae_cf_and_R_palmularis_cf
KK263_R_czywczynensis_cf
KK259_R_marginicolus_cf
KK267A_R_binatus_cf
KK260_R_hannae_cf
KK273_R_pawlowksi_cf_and_R_kalinensis_cf
KK274_R_phragmiteti_cf
KK275_R_carpaticola_cf
LH042_R_carpaticola_cf
42_R_hungaricus_cf
18_R_hungaricus_cf
LH039_R_hungaricus_cf
10_R_hungaricus_cf
LH045_R_pannonicus_cf
LH040_R_cassubicifolius_s_l
LH044_R_pilisiensis_cf
LH032_R_phragmiteti_cf
LH001_R_variabilis_cf
LH002_R_variabilis_cf
LH043_R_integerrimus_cf
LH007_R_cassubicifolius_s_l
LH033_R_variabilis_cf
LH038_R_hungaricus_cf
LH037_R_cf_11
Du-21047_R_cassubicifolius_s_l
LH009_R_cassubicifolius_s_l
LH036_R_pilisiensis_cf
LH027_R_pilisiensis_cf
LH026_R_pilisiensis_cf
LH034_R_variabilis_cf
10137_R_notabilis_s_l
LH029_R_variabilis_cf
10136_R_variabilis_cf
LH030_R_variabilis_cf
Du-21093_R_flabelloides_cf
0
75
150 km

g)


SE
Øland
10352_R_latifolius_cf
10351_R_reniger_cf
10348_R_fissifolius_cf
10353_R_fissifolius_cf
10342_R_coleophorus_cf
10359_R_latifolius_cf
10343_R_indet_2
10362_R_fissifolius_cf
10358_R_latifolius_cf
10357_R_indet_4
10354_R_reniger_cf
10344_R_lacer_cf
10360_R_reniger_cf
10345_R_litoricola_cf
10346_R_parcipilosus_cf
10347_R_indet_3
0
25
50 km

h)


10381_R_glechomoides_cf
10373_R_indet_6
10372_R_elatior_cf
SE
10371_R_elatior_cf
10374_R_imitans_cf
10370_R_marklundii_cf
10369_R_marklundii_cf
10368_R_fissifolius_cf
10367_R_fissifolius_cf
10382_R_fissifolius_cf
10363_R_marklundii_cf
10380_R_fissifolius_cf
10365_R_fissifolius_cf
10366_R_latifolius_cf
10378_R_indet_7
10379_R_glechomoides_cf
10375_R_integerrimus_cf
10376A+B_R_fissifolius_cf_and_R_glechomoides_cf
0
25
50 km

i)

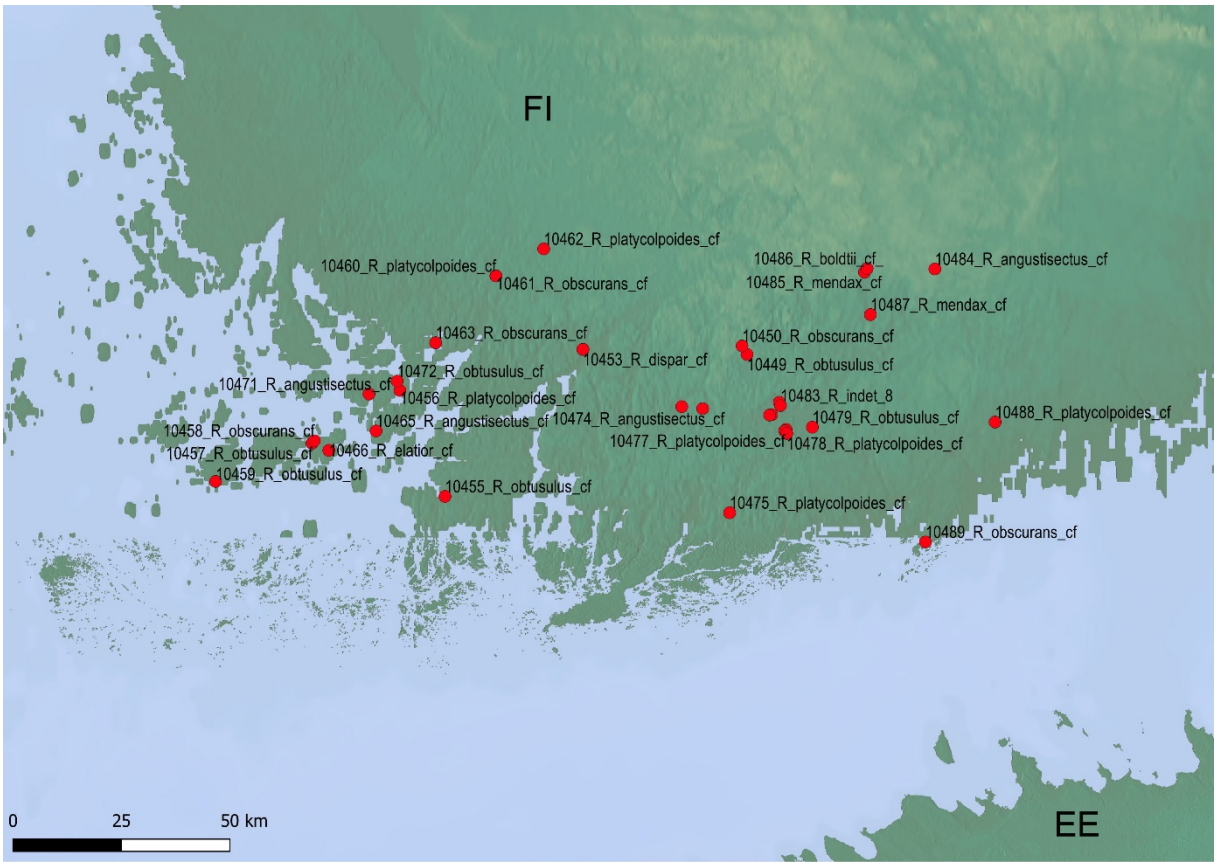


Figure S1. Geographical maps (a-i) of *R. auricomus* populations included in the present study. Sampling locations in (a) Europe, (b) subregions, (c) Western Europe, (d) Alsace, (e) Central Germany, (f) Eastern Europe, (g) Oeland (Northern Europe), (h) Uppland (Northern Europe), and (i) Finnland (Northern Europe). Names of taxa are tentative (please see Tables S1, S2 for latest identifications and taxon names; *R. angustisectus* = *R. cosmophyllus*, *R. petiolatus* = *R. bochers-kolbiae*). Maps were created with QGIS v.3.10.2 (QGIS Development Team, 2020).

Table S1 (electronical supplement). Detailed information about populations included in the present study. For each population, the provisional taxon name, ploidy level, reproduction mode, locality (incl. ISO code 3166-2), collection data, altitude (in meter above sea level), latitude (north), longitude (east), habitat, collector, and deposition of herbarium specimens are given. See Fig. S1a-i for geographic maps of locations via population IDs. * = locality is at locus classicus or nearby, # = holotype, + = private herbarium F. G. Dunkel.

The electronical supplement is available online:
https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fmec.15919&file=mec15919-sup-0002-TableS1-S3.xlsx

Text S1. Cross-pollination, flower isolation, seed harvesting, further garden work, and comparability of facultative apomictic sexuality and of number of petals.

If populations were met in flowering stage, we cross-pollinated individuals and isolated the complete flowering structure immediately with porous plastic bags (Baumann Saatzuchtbedarf GmbH, Waldenburg, Germany) to avoid cross-pollination with other taxa during transport or in the garden. Bags were sealed with tape to prevent seed loss during harvesting. Some populations in early spring 2018 were met before early bud stage, and seeds developed under garden conditions. If populations were already in fruiting stage, we repeated the same cross-pollination procedure in the garden in 2018 and 2019. All plastic bags containing mature seeds were collected in summer 2018 and 2019. Because some populations were already met in fruiting stage or petals were not recorded during sampling, we again counted petals of all populations in the garden in 2019. Seeds of sexual species (already under garden culture) were mainly harvested in 2017 and 2018 (Karbstein et al., 2020a). Moreover, we collected herbarium specimens for all populations and deposited them in the herbarium GOET.

To assess whether treatments (in situ, garden) influenced facultative sexuality (apomicts) and petal number, we checked group differences with Wilcoxon rank-sum test due to non-normally distributed data (Shapiro test, and QQ-plot). We collected plants in a stage where the sexual or apomictic pathway has already been fixed (before the bud stage). We tested whether seed development was affected by ripening in situ or under garden conditions. We found no significant differences in sexuality (percent sexual seeds to all seeds) between in situ (7.5%) and garden (6.9%) seed development across facultative apomictic populations ($W = 409$, $p = .55$). Therefore, sexuality of populations is comparable across the dataset. We also found no significant difference in mean number of petals between the treatments in situ (2018) and garden (2019; $W = 11128$, $p = .78$).

Table S2 (electronical supplement). Flow cytometry of leaves (FC) and single-seed flow cytometric seed screening (FCSS) of sexual diploid and polyploid populations within the *R. auricomus* complex. For each population, i.a. the (provisional) taxon names, number of examined individuals or seeds (FC, FCSS), ploidy levels (leaf, embryo), references (for added data from already published DNA ploidy and reproduction mode measurements), peak indices (ratio of endosperm ploidy to embryo ploidy), sexuality (percent of sexual seeds to all seeds), reproduction mode, petals, habitat types, seed development (y = in situ, n = in garden), latitude (north), longitude (east), altitude, and scaled bioclimatic variables (bio1-19) and solar radiation values (derived from WorldClim database v.2, Fick & Hijmans, 2017), genome-wide heterozygosity (percent heterozygous sites to all sites), and the genetic clusters based on the sNMF analysis are given. We merged very close or identical locations (labels "A", "B", "C", etc.), which explains high sample sizes of some populations in FC and FCSS measurements. See Materials and Methods section for more details.

The electronical supplement is available online:
https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fmec.15919&file=mec15919-sup-0002-TableS1-S3.xlsx

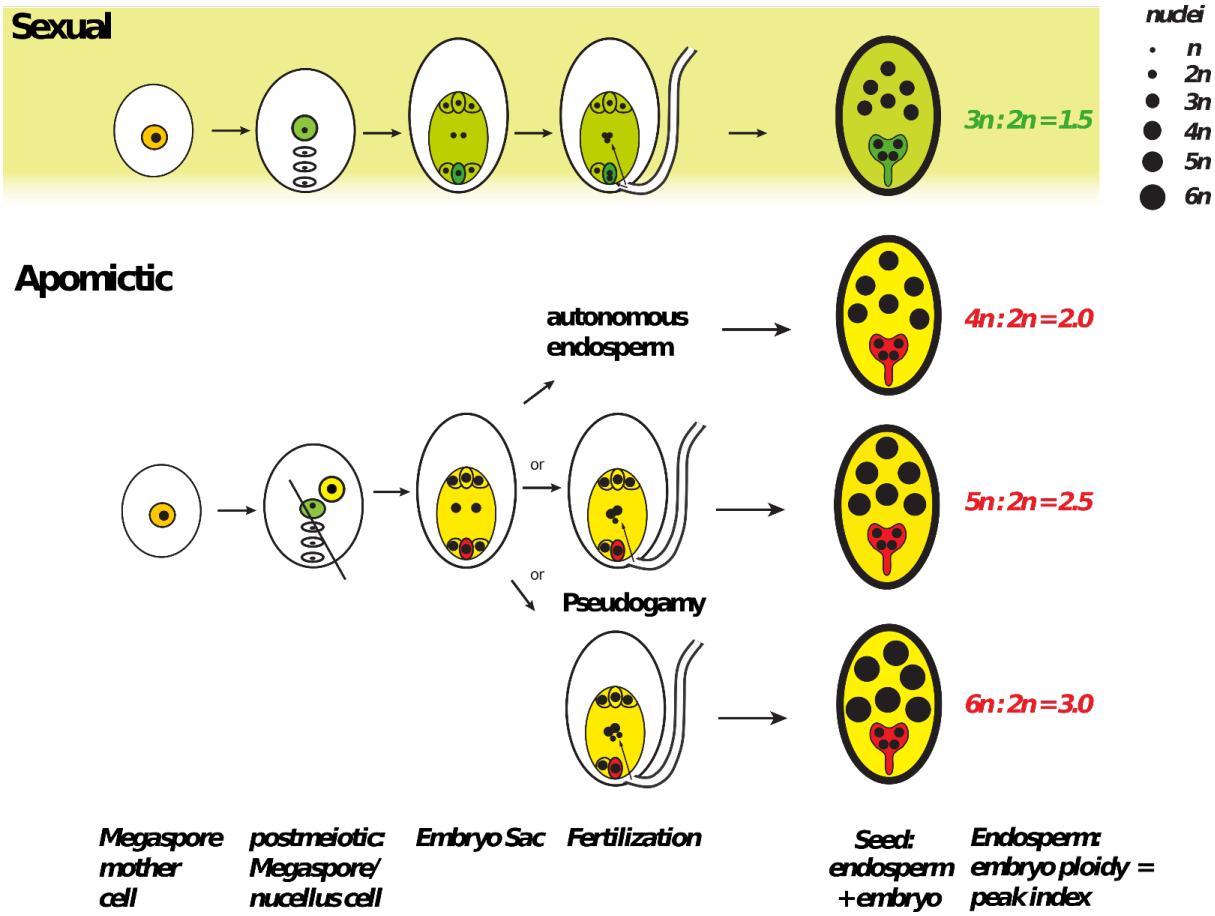


Figure S2. Developmental pathways of sexual vs. apomictic mode of reproduction in the *R. auricomus* complex (edited figure from Hojsgaard & Hörandl 2019; Creative Commons Attribution Licence CC-BY). Ploidy levels of nuclei (n) are symbolized by size of black dots. Sexual development results after meiosis in a reduced megaspore, which develops into a reduced embryo sac, which forms after double fertilization of egg cell and polar nuclei a diploid embryo and a triploid endosperm. Apomictic development starts from a somatic (unreduced) cell and results in an unreduced embryo sac, in which the unreduced egg cell develops parthenogenetically into an embryo. Polar nuclei are either unfertilized (autonomous apomixis), or fertilized by one or both sperm nuclei (pseudogamy). The resulting endosperm is tetra-, penta- or hexaploid, respectively. Flow cytometric seed screening measures the respective ratios of endosperm:embryo ploidies in each seed. These ratios can be also given as peak index (PI, last column).

Text S2. Classification of reproduction modes and filtering.

Within the apomictic pathway, the unreduced egg cell develops parthenogenetically, and the endosperm is formed without fertilization or with contribution of one or two reduced/unreduced sperm nuclei (PI = 2, 2.5, 3.0, and 4.0, respectively; Matzk et al., 2000; Doležel et al., 2007; Klatt et al., 2016). If the embryo was still in development and hence the peak too small, we used individual- or population-wise (taking mixed ploidy in populations into account) embryo mean values to calculate the ratios. We classified all PI values below 1.7 as sexual reproduction mode (excluding also the possibility of paternal BIII hybrids (PI = 1.3), see e.g., Barke et al., 2018) and all PI values above 1.9 as apomictic pathways. Additionally, we checked for maternal BIII hybrids (PI = 1.7; see Barke et al., 2018) but we didn't find any. We set the following thresholds to classify the different apomictic reproduction modes: 1.9-2.1 as autonomous endosperm development (excluding parthenogenetic development of the embryo; see also Barke et al., 2018); 2.4-2.6 as endosperm development with one reduced pollen; 2.8-3.2 as endosperm development with two reduced or one unreduced pollen, and 3.7-4.3 as endosperm development with two unreduced pollen. By comparing maternal leaf with embryo ploidy, we observed nine embryo polyploidizations (4n mother plants with mostly 8n embryos), 10 diploidized embryos from 4n mother plants, and six triploidized embryos from 4n mother plants (haploid parthenogenesis, unequal chromosome segregation), and excluded them from further analyses. In addition, we sometimes observed slight differences between maternal leaf and embryo peak potentially due to experimental error or physiological differences (see also Lo, Stefanović, & Dickinson, 2013).

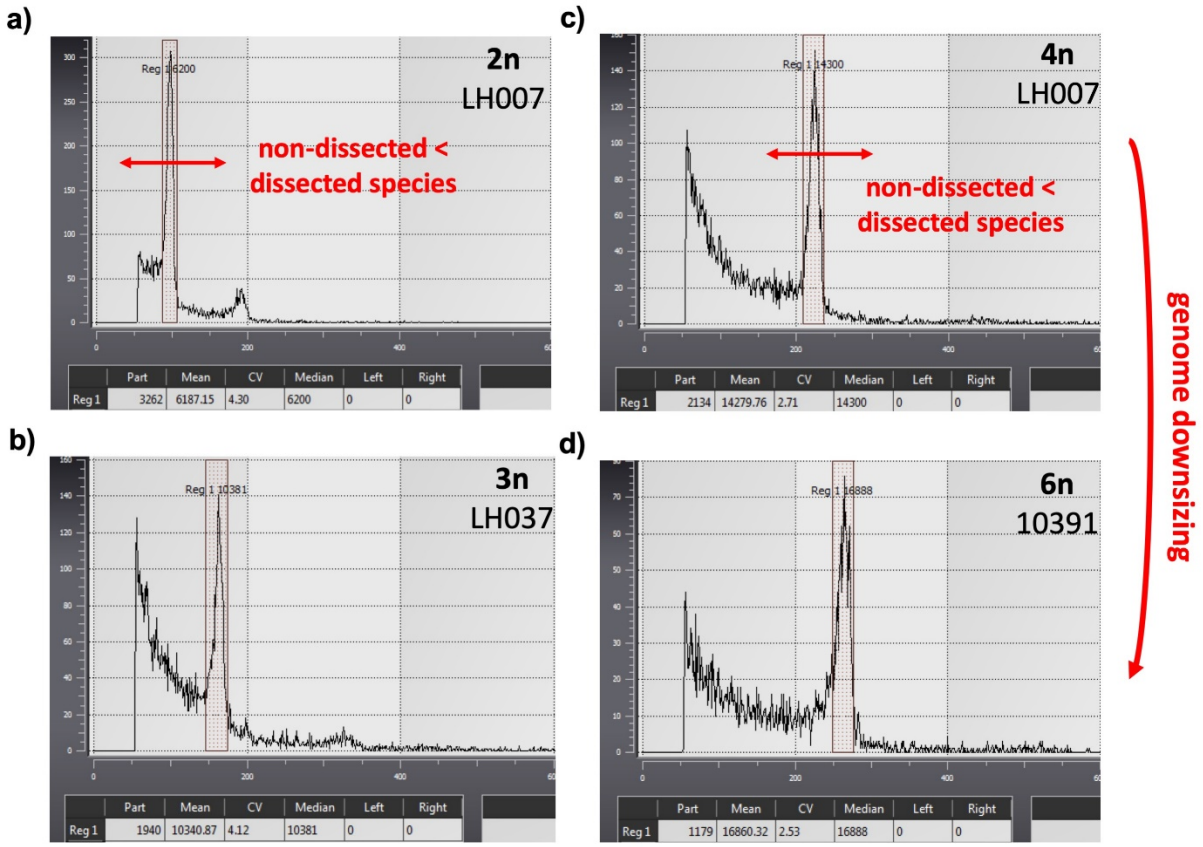


Figure S3. Examples of different ploidy levels in (a) diploid (2n), (b) triploid (3n), (c) tetraploid (4n), and (d) hexaploid (6n) *R. auricomus* individuals assessed by flow cytometry (FC). Population IDs are given for each example. Diploid non-dissected leaf species, in general, have lower relative DNA content than dissected ones (see Paule et al., 2018). This peak range variation also affects the tetraploid individuals that arose from diploid progenitor species. We also observed a remarkable genome downsizing effect from the tetra- to hexaploid ploidy level.

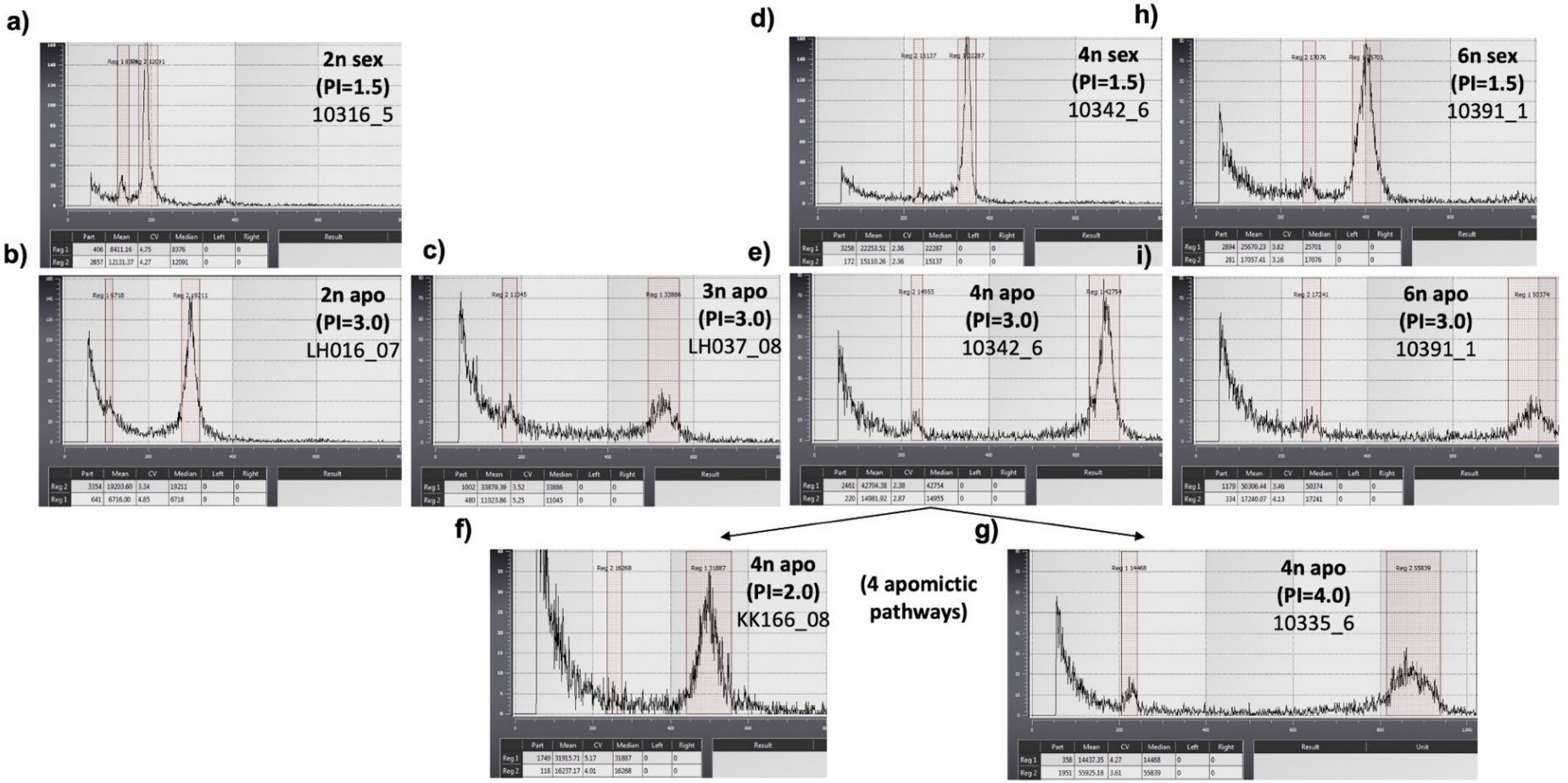


Figure S4. Examples of different reproduction modes in (a, b) diploid (2C), (c) triploid (3C), (d-g) tetraploid (4C), and (h, i) hexaploid (6C) *R. auricomus* individuals assessed by flow cytometric seed screening (FCSS). The first peak represents the embryo peak, and the second one the endosperm peak. Peak index (PI) was calculated as the ratio of endosperm to embryo peak. Individual IDs are given for each example. sex = sexual, apo = apomictic.

Text S3. Within-sample (ISC) and among-sample clustering (BSC), assurance of orthologous locus assembly, and potential degree of chloroplast sequences in the final assembly.

First, we separated di-, tri-, tetra-, and hexaploid taxa into four datasets; second, we separately optimized within-sample clustering thresholds (ISCT, within individuals) of di-, tri-, tetra-, and hexaploid taxa; third, we merged all assemblies and optimized among-sample clustering threshold (BSCT, between individuals). For each dataset (2n, 3n, 4n, and 6n), we kept all parameters constant (default) except clustering threshold (minimal sequence similarity in percent) for de novo assembly. We adapted the maximum number of alleles settings according to ploidy levels (2, 3, 4, and 6), and slightly increased maximum SNPs per locus to 30% and maximum indels per locus to 12 accounting for increased genetic variation in polyploids within the last filtering step (IPYRAD step 7, BSC). We specified a minimum depth of 12 because fewer SNPs and duplicates were filtered compared to minimum depth 6 setting from previous analyses (see Karbstein et al., 2020a).

We selected the ISCTs: 95% (2n), 95% (3n), 96% (4n), and 95 (6n, Fig. S5). Selected ISCT thresholds have already a stabilized number of clusters and cluster depth (by 5% allowed SNPs per locus and per sample) and fewer clusters rejected due to high heterozygosity, indicating the transition between under- and overmerging of reads and thus avoidance of oversplitting of loci. The optimal BSCT value of the merged assembly was 97% (Fig. S6). Settings of maximum SNPs per locus (30%) and maximum indels per locus (12) filtered less than 0.1% and 0.01% of all assessed 754,033 loci after determining the optimal BSCT, and therefore these settings did not remarkably influence the filtering procedure. The selected BSC threshold maximizes SNPs and balances over- and undermerging of loci.

To assess the potential degree of chloroplast sequences in the final assembly, we used the published *Ranunculus repens* plastome (GenBank accession: NC 036976.1) and the alignment program ALIVIEW v.1.18 (Larsson, 2014) to search for the recognition site of *PST*I ("TGCAG"). We detected 162 recognition sites (loci), i.e., 0.021% of all retained RAD loci (754,033). Followingly, the applied restriction enzyme PstI has only few cutting sites in the *R. repens* plastome, and the number of chloroplast loci in the final assembly is probably lower than 1% (14,489 final loci) and thus negligible.

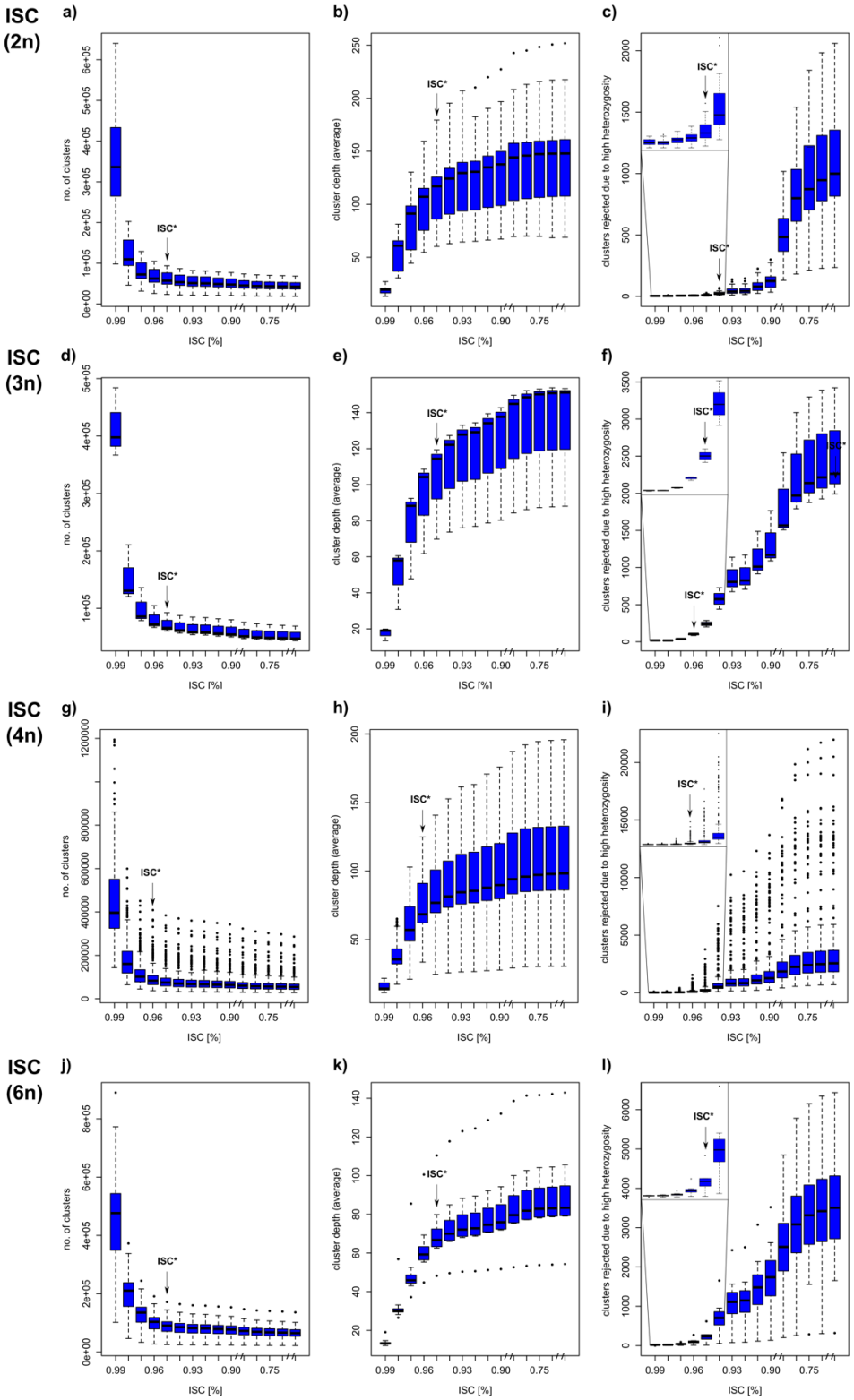


Figure S5. Evaluation of within-sample clustering threshold (ISCT) concerning di- (Karbstein et al., 2020a), tri, tetra, and hexaploid individuals of the *R. auricomus* complex. Boxplots illustrate (a, d, g, j) number of clusters, (b, e, h, k) average cluster depth (coverage), and (c, f, i, l) clusters rejected due to high heterozygosity in relation to within-sample clustering. „Mindepth“ was set to 12 for all analyses.

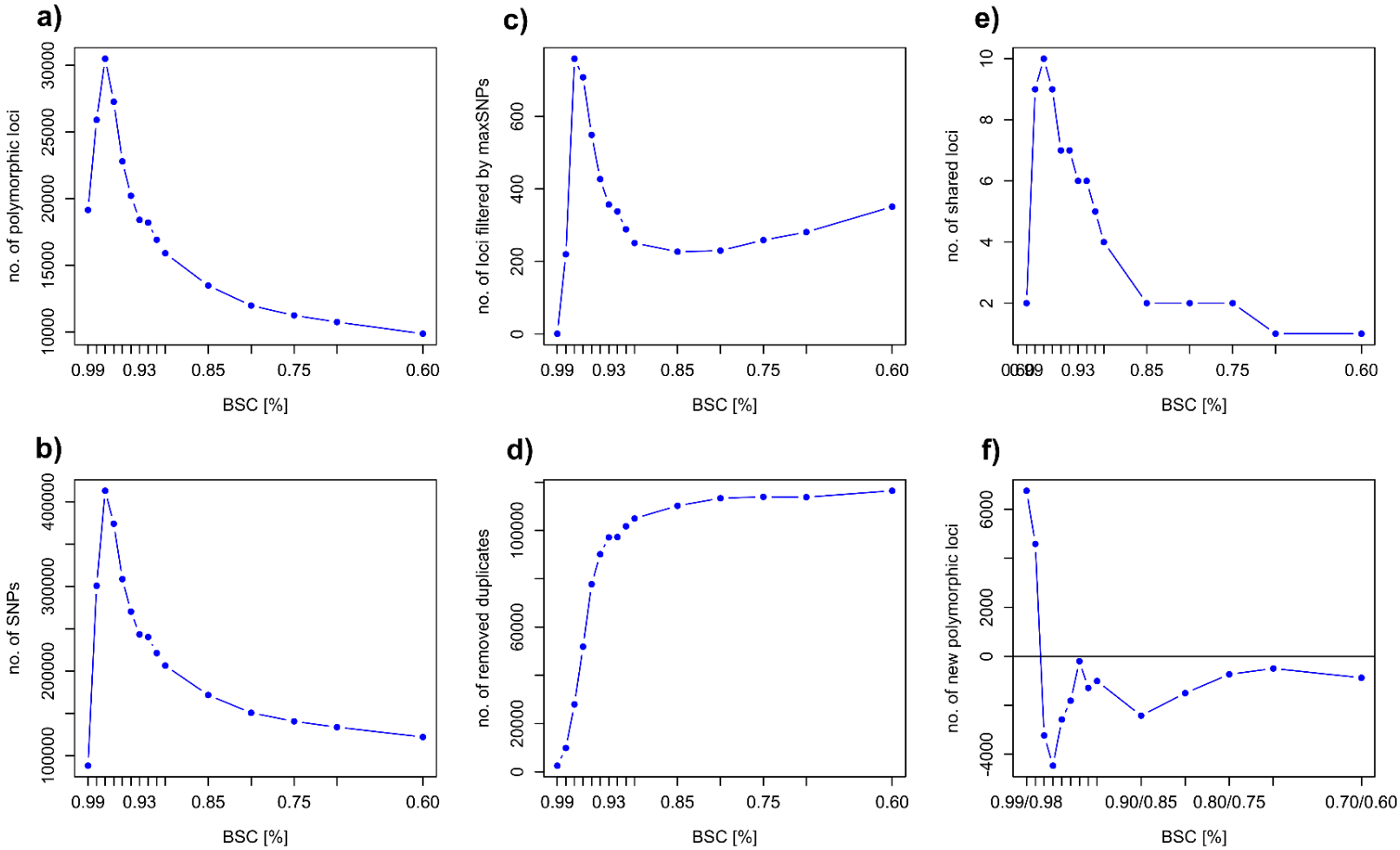


Figure S6. Evaluation of among-sample clustering threshold (BSCT) of the merged assembly (di- (Karbstein et al., 2020a), tri-, tetra-, and hexaploid individuals of the *R. auricomus* complex). Boxplots illustrate number of (a) polymorphic loci, (b) number of SNPs, (c) number of loci filtered by maxSNPs, (d) number of removed duplicates, (e) number of shared loci across samples, and (f) number of new polymorphic loci in relation to among-sample clustering. "Mindepth" was set to 12 for all analyses.

Table S3 (electronical supplement). Inferred heterozygosity (percent of heterozygous sites to all sites) of *R. auricomus* individuals. We averaged these values per population (see Table S2). See Materials and Methods section for further details.

The electronical supplement is available online:
https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fmec.15919&file=mec15919-sup-0002-TableS1-S3.xlsx

Text S4. Heterozygosity in relation to reproduction modes within the *R. auricomus* complex.

According to theory and previous population genetic studies (Freeland, Kirk, & Petersen, 2011; Hörandl, Jakubowsky, & Dobeš, 2001; Paun, et al., 2006), we expect Hardy-Weinberg-equilibrium and genetic cohesion in sexual outcrossing populations and predominant clonality within apomictic populations. We found similar heterozygosity values among individuals (2-6) of sexual and apomictic populations (Table S3), additionally confirmed by significant differences between two sampled individuals of a population, $W = 1931$, $p = .31$). Results indicate comparable intrapopulation heterozygosity, and main differences in individual heterozygosity are probably due to ploidy effects (see also Results and Discussion sections). Thus, in apomicts, we regard one (up to four) individuals as representative for heterozygosity, while at least two individuals per population were targeted for sexuals (except "Du" populations, close to other sampled populations).

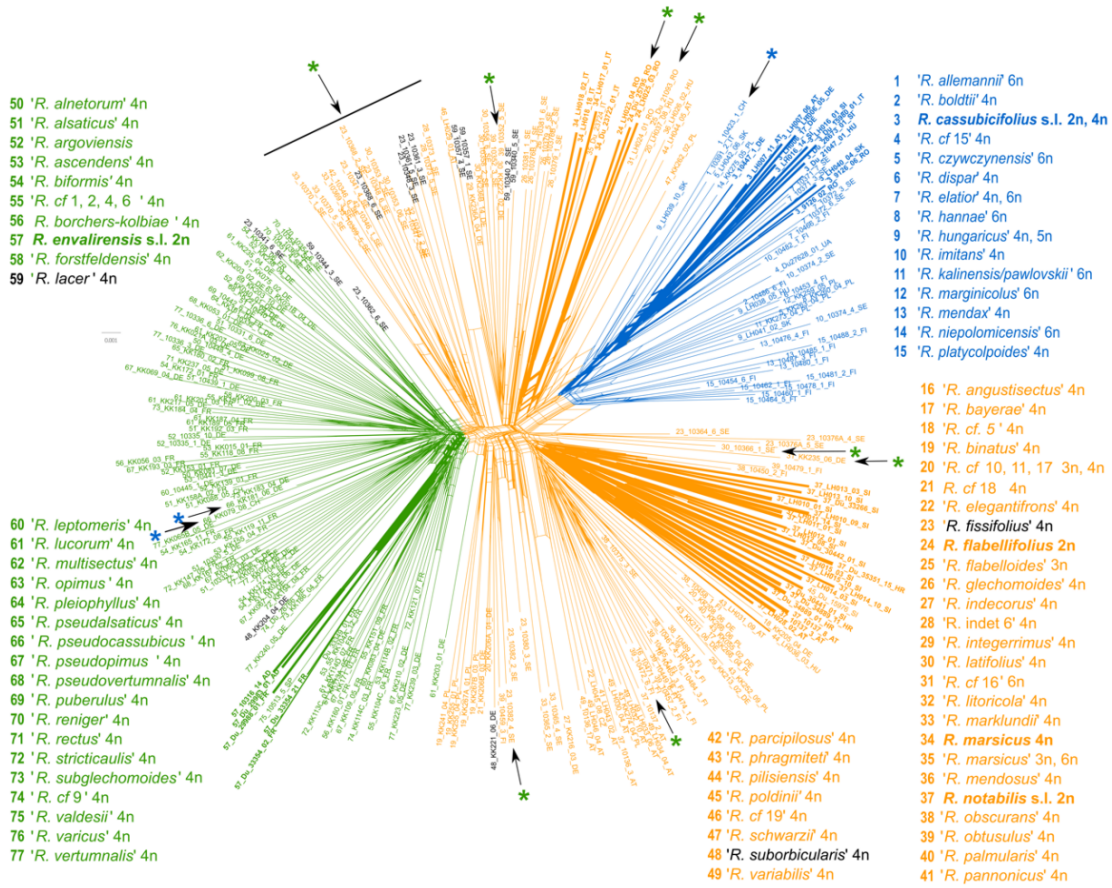


Figure S7. Splitsgraph of Neighbor-net analysis of 280 sexual and polyploid apomictic individuals based on genetic distances (GTR model with estimated site frequencies and ML) and the "min30" alignment (unlinked SNPs, i.e., 1 SNP/locus). The figure illustrates highly reticulate relationships and a lack of genetic coherence of samples within described apomictic morphospecies. The accessions are labeled with the respective species number, followed by the sample ID and a country code (see Table S1, S2). The five accepted sexual species (Karbstein et al., 2020a) are marked in bold. The three main genetic groups are indicated in colors (note that the 3D NeighborNet is plotted here as 2D graph, thereby artificially drawing the orange cluster on two sides). Species belonging to two groups are labeled in black. Arrows indicate a different cluster position compared with the sNMF analysis (Fig. 3, Figure with individual labels on Figshare) and asterisks colors show the position in the sNMF analysis (blue = cluster I, orange = cluster II, green = cluster III).

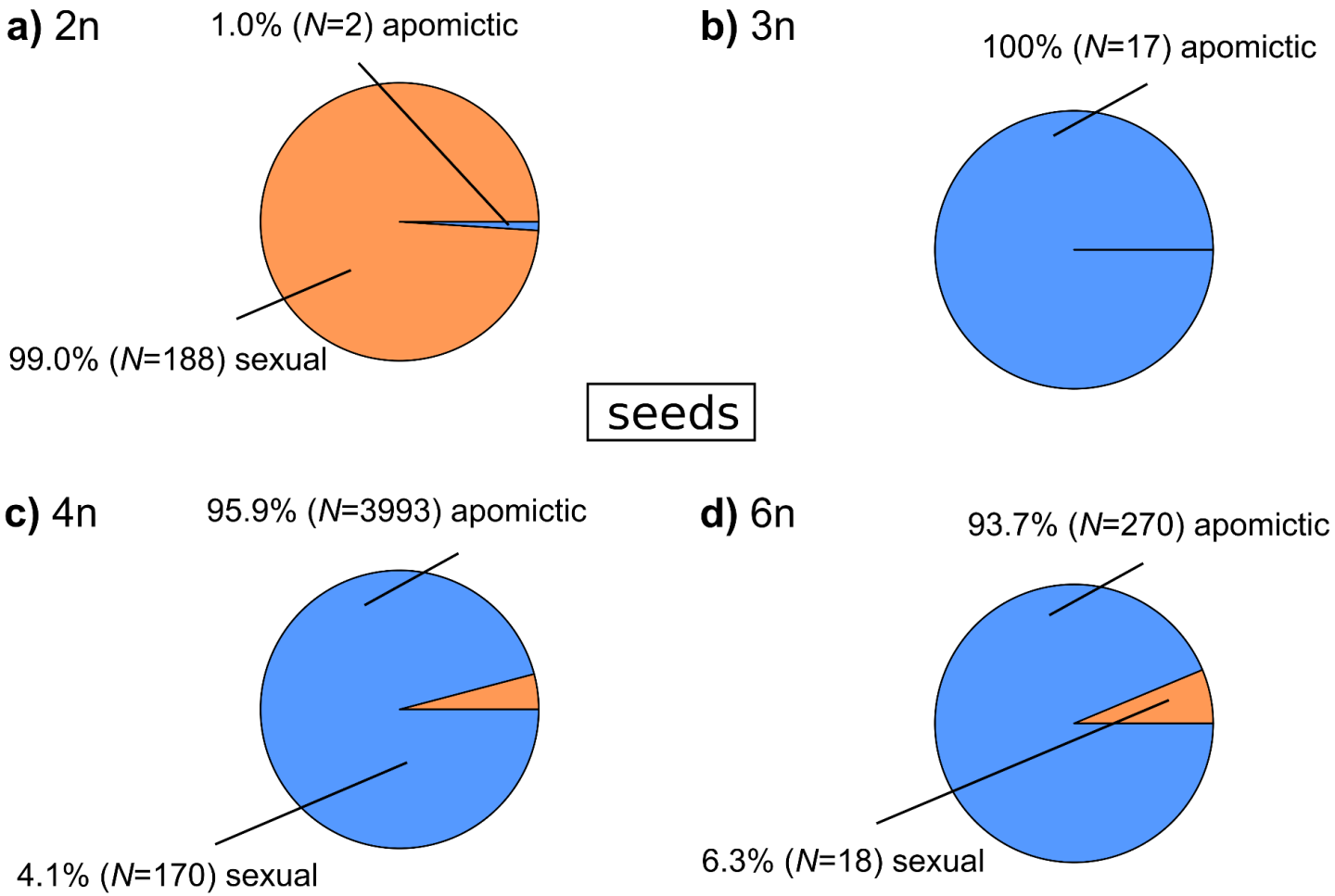


Figure S8. Proportions of observed sexual and apomictic *R. auricomus* seeds for each ploidy level (2n, a; 3n, b, 4n, c; 6n, d) illustrated in pie charts. N = number of seeds.

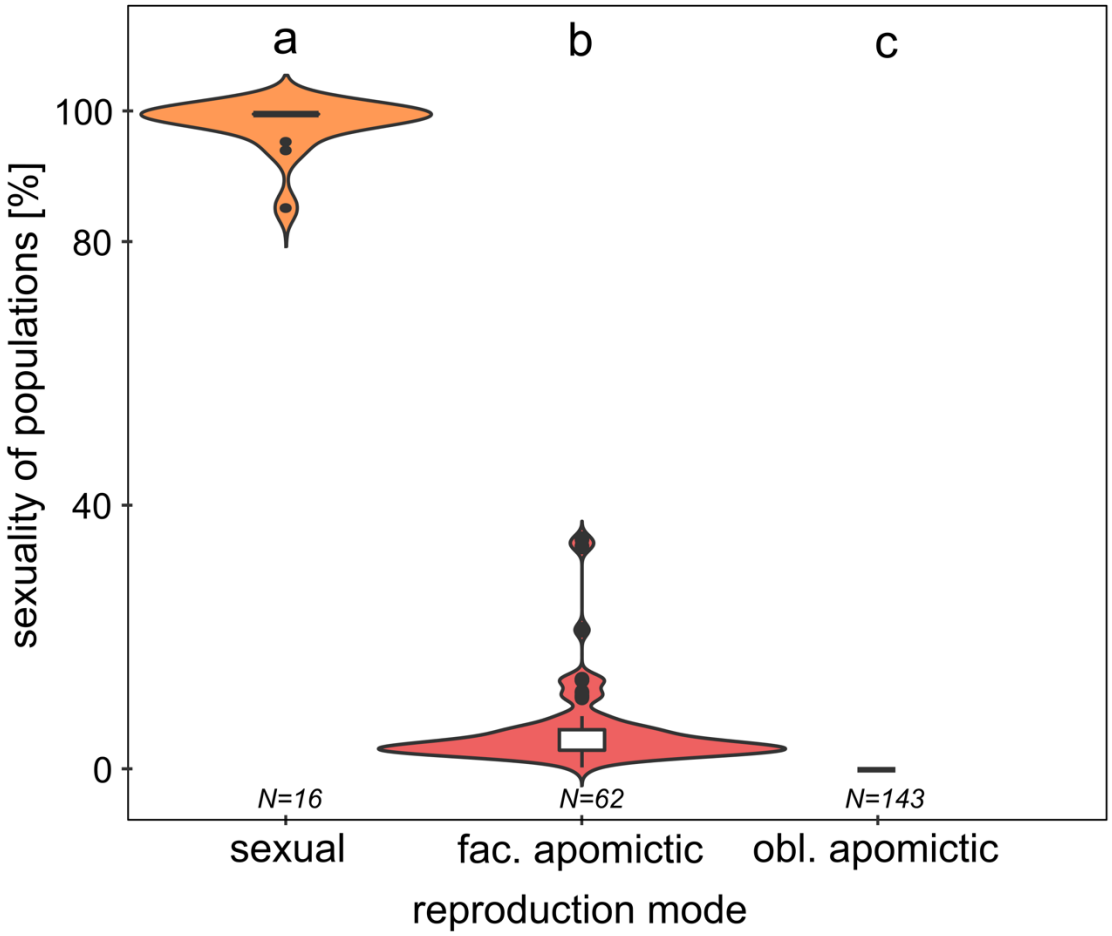


Figure S9. Violin plot of population sexuality (percent sexual seeds to all seeds) in relation to different reproduction modes (without added populations from literature). Letters above indicate significant/non-significant differences between/among groups. We checked differences among groups with Kruskal Wallis test ($Chi^2 = 230.2$, $df = 2$, $p < .001$) followed by a pairwise Wilcox rank sum-test (sexual-fac. apo.: $p < .001$, sexual-obl. apo.: $p < .001$, fac. apo.-obl. apo: $p < .001$). N = number of populations per group level.

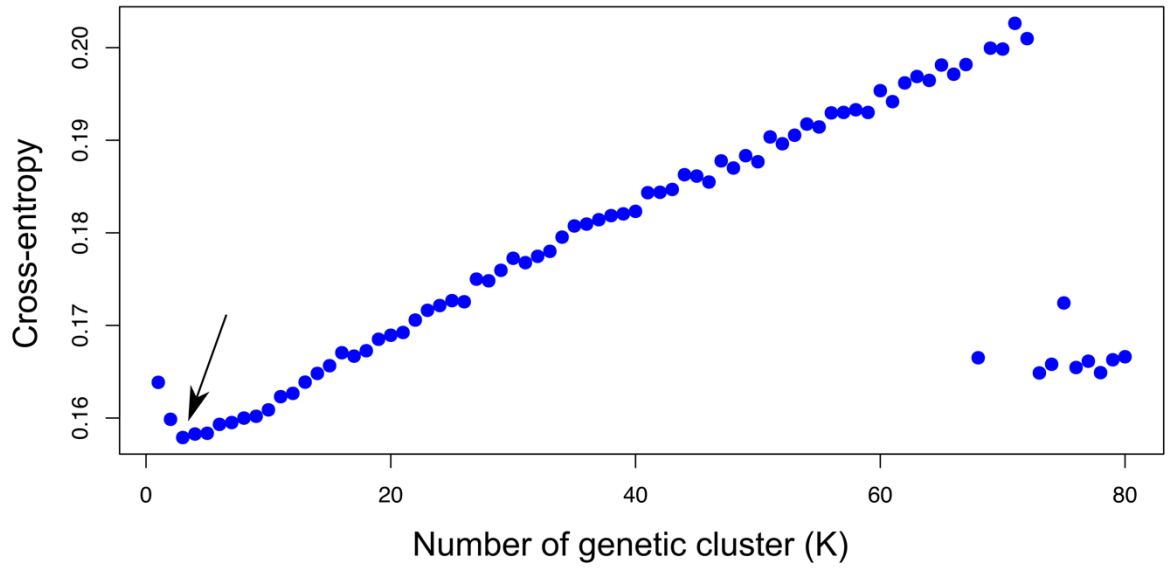


Figure S10. Cross-entropy calculations based on different numbers of genetic clusters (K = 1-80) and the "min30" RADseq alignment (1 SNP/locus; 33,165 loci). The criterion examines the prediction of masked genotypes to evaluate the fit of a model with K populations and, thus, the lower the value, the better the specific model. The arrow indicates the global cross-entropy minimum.

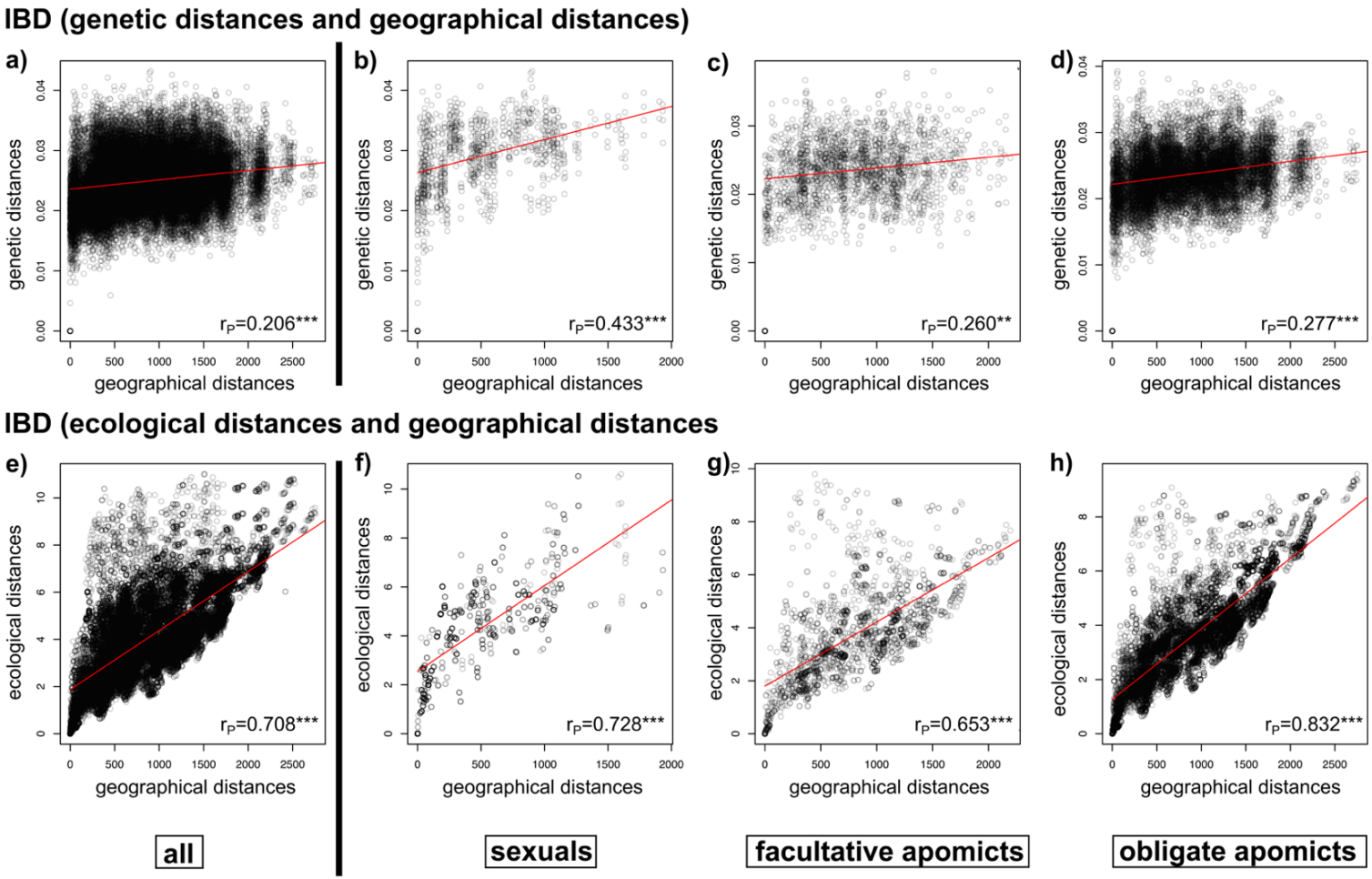


Figure S11. Patterns of isolation-by-distance (IBD, a-d) or isolation-by-ecology (IBE, e-h) for (a) all populations (N = 269), (b) sexual populations (N = 45), (c) facultative-apomictic populations (N = 65), and (d) obligate-apomictic populations (N = 159). Regression lines are based on linear regression models. Total sample size (N = 269) differs from sNMF and SplitsTree analysis (N = 280) because some individuals were removed if the exact reproduction mode was not available (see also missing reproduction modes in Table 2). N = number of individuals, r_S = Pearson's correlation coefficient *** = p < .001, ** = p < .01.

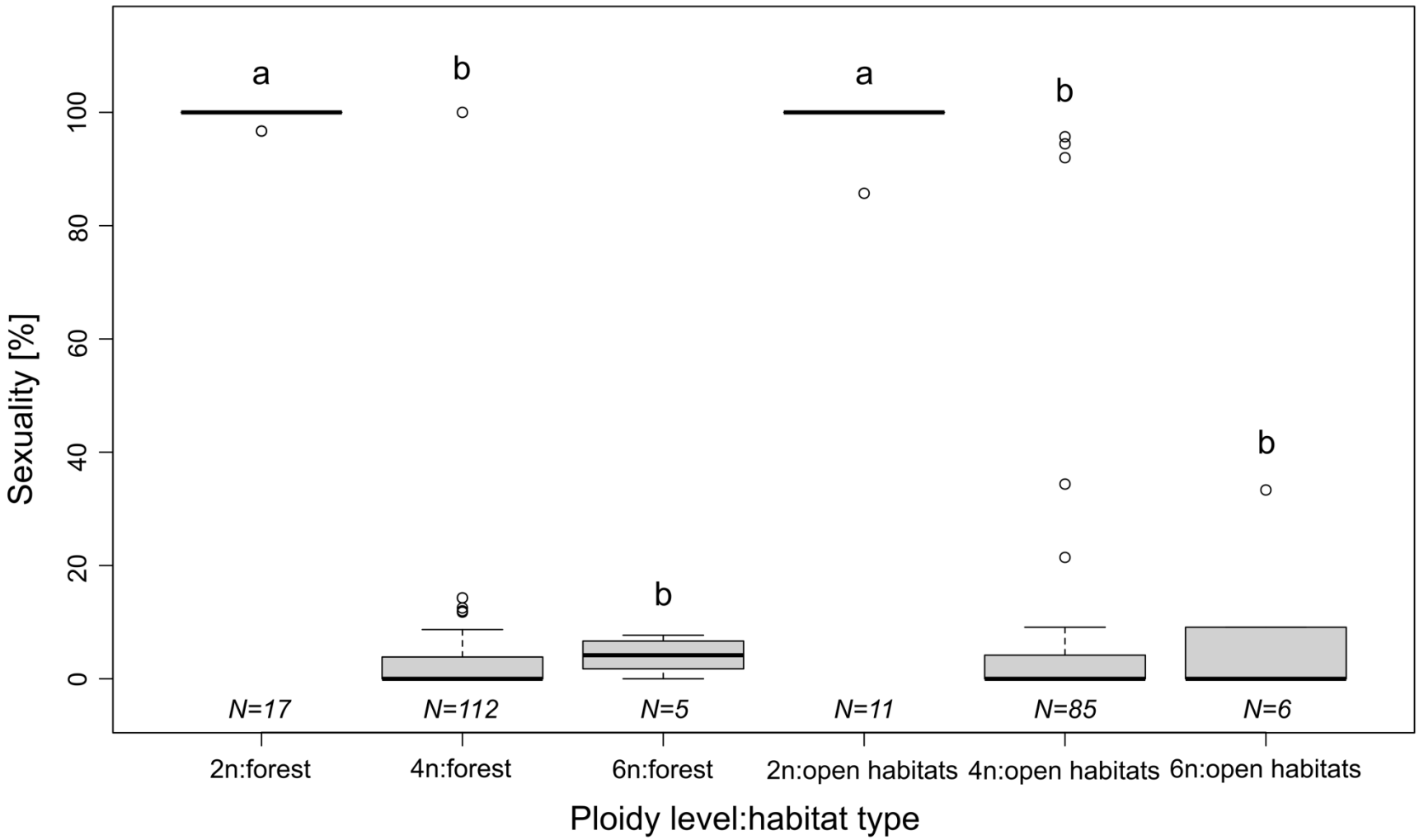


Figure S12. Boxplots of population sexuality (percent of sexual seeds to all seeds) in relation to different ploidy level-habitat type levels (interaction). Letters above boxplots indicate significant/non-significant differences between/among groups. We checked differences among groups with Kruskal Wallis test followed by a pairwise Wilcox rank sum-test. N = number of populations per group level.

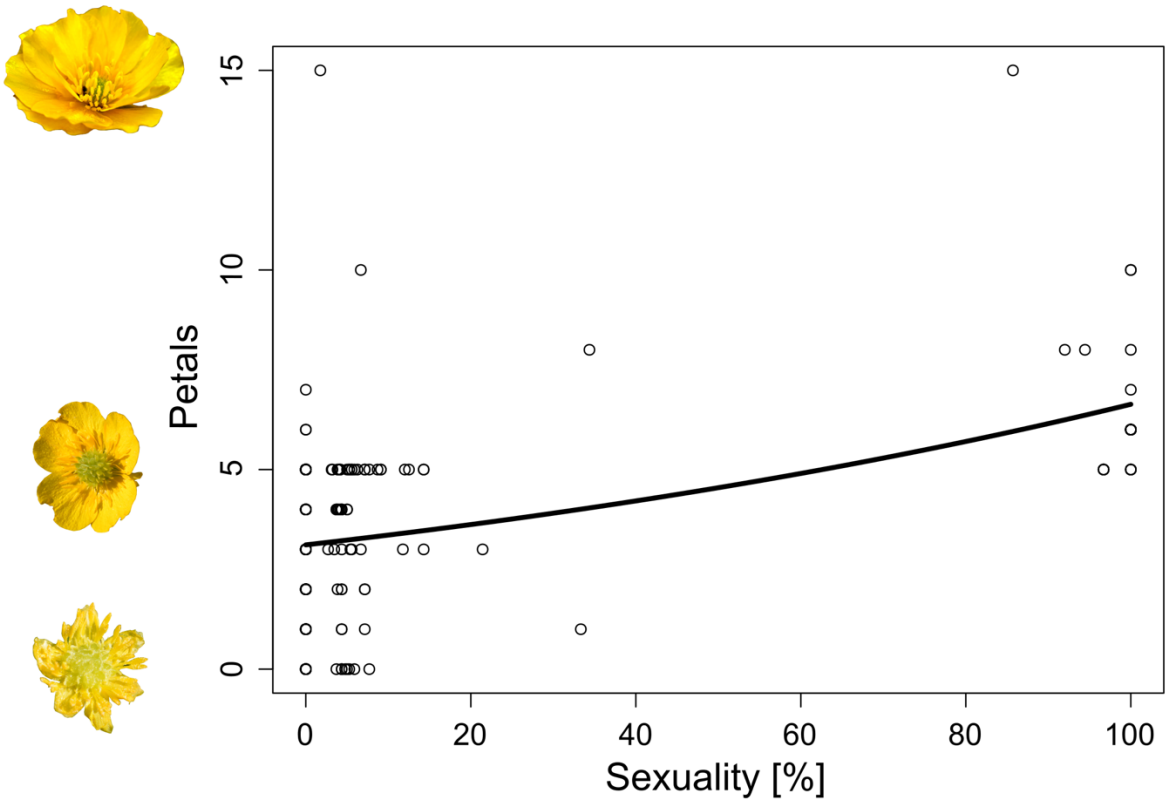


Figure S13. Regression plot illustrating maximum number of petals (response) in relation to population sexuality (predictor) of 229 *R. auricomus* populations. The solid curve indicates a significant relationship (path analysis, $p < .001$, Table 1, S5). We removed an extreme outlier (population KK263 "*R. czywczynensis* cf" containing only a few individuals with 30 petals/individual). The regression line was drawn based on a separate GLM (poisson) analysis (estimate = 0.76, $p < .001$), and thus the separate relationship was comparable to the SEM result (Table 1). In addition, the non-significant (and not illustrated) SEM relationship between heterozygosity and sexuality is significant in a separate GLM (estimate = -0.71, $p < .001$; compare to Table 1).

Table S4. Results of the generalized linear mixed effect model (GLMM) between sexuality (percent of sexual seeds to all seeds) and climatic environmental factors (and their interactions) including 214 *R. auricomus* populations. P values of significant relationships are highlighted in bold. Isothermality = bio3, temperature seasonality = bio4, annual precipitation = bio12, df = degrees of freedom.

predictor	estimate	std. error	df	t value	p value
isothermality	1.07	0.43	204	2.47	**<0.05**
temperature seasonality	1.19	0.34	204	3.53	**<0.001**
temperature seasonality : solar radiation	0.94	0.29	204	3.24	**<0.01**
annual precipitation	1.85	0.26	204	6.97	**<0.001**
solar radiation	0.87	0.35	204	2.46	**<0.05**
altitude	-1.49	0.50	204	-2.97	**<0.01**
altitude : solar radiation	1.04	0.33	204	3.15	**<0.01**

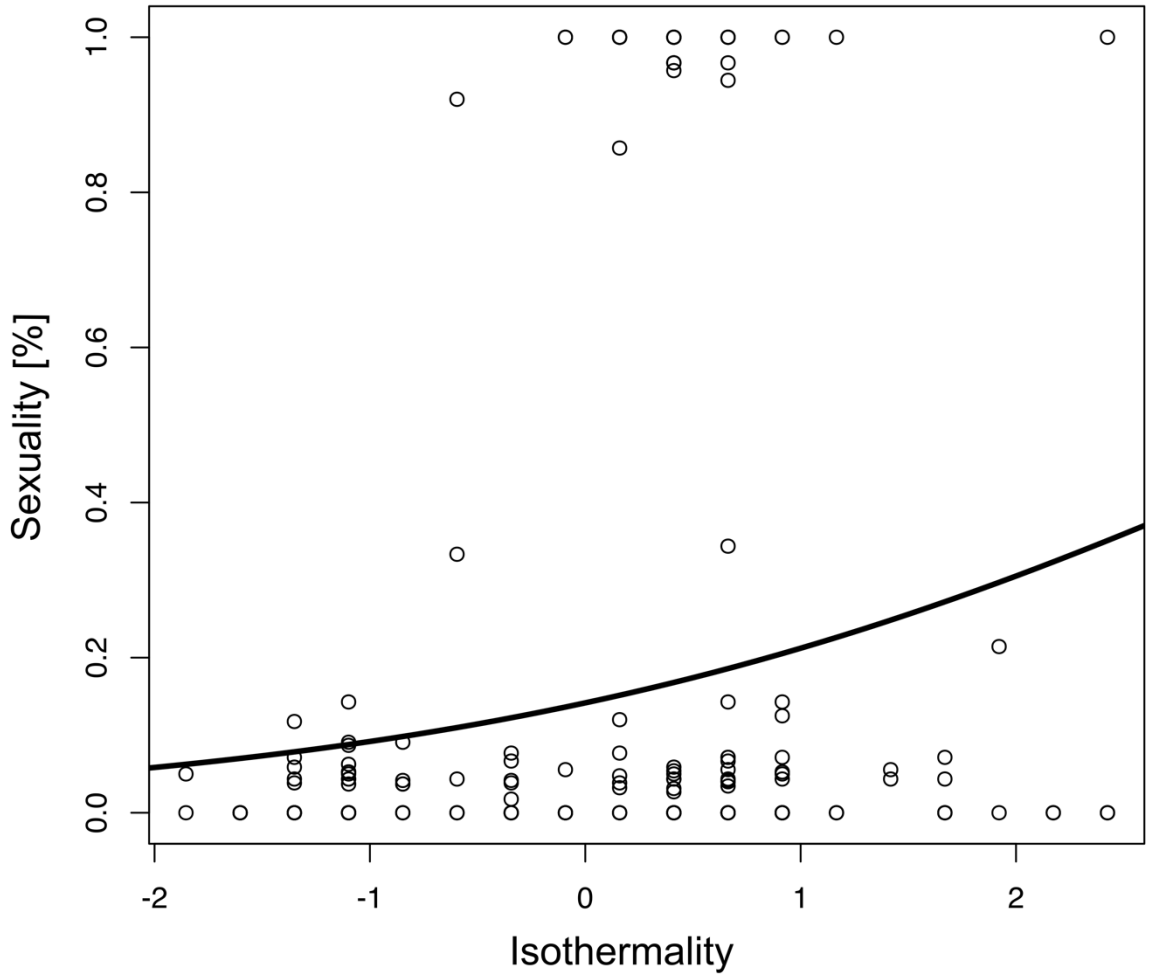


Figure S14. Regression plot illustrating the relationship between population sexuality (percent of heterozygous sites to all sites; response) and isothermality (bio3; predictor) of 237 *R. auricomus* populations. Isothermality is the ratio of mean diurnal range to temperature annual range*100. The solid curve indicates a significant relationship (path analysis, $p < .01$, Table 1, S5). The regression line was drawn based on a separate GLM (binomial) analysis.

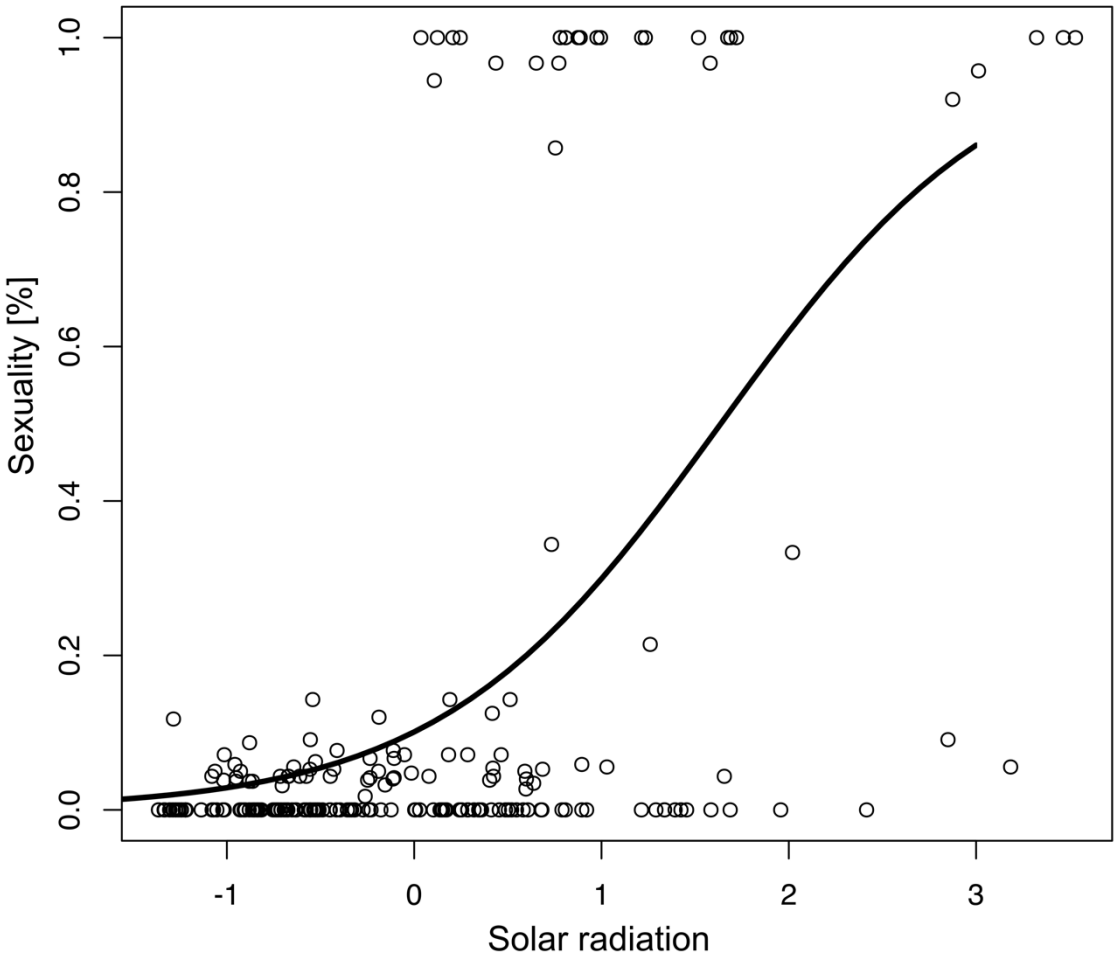


Figure S15. Regression plot illustrating the relationship between population sexuality (percent of heterozygous sites to all sites; response) and solar radiation (predictor) of 237 *R. auricomus* populations. The solid curve indicates a significant relationship (path analysis, $p < .001$, Table 1, S5). The regression line was drawn based on a separate GLM (binomial) analysis.

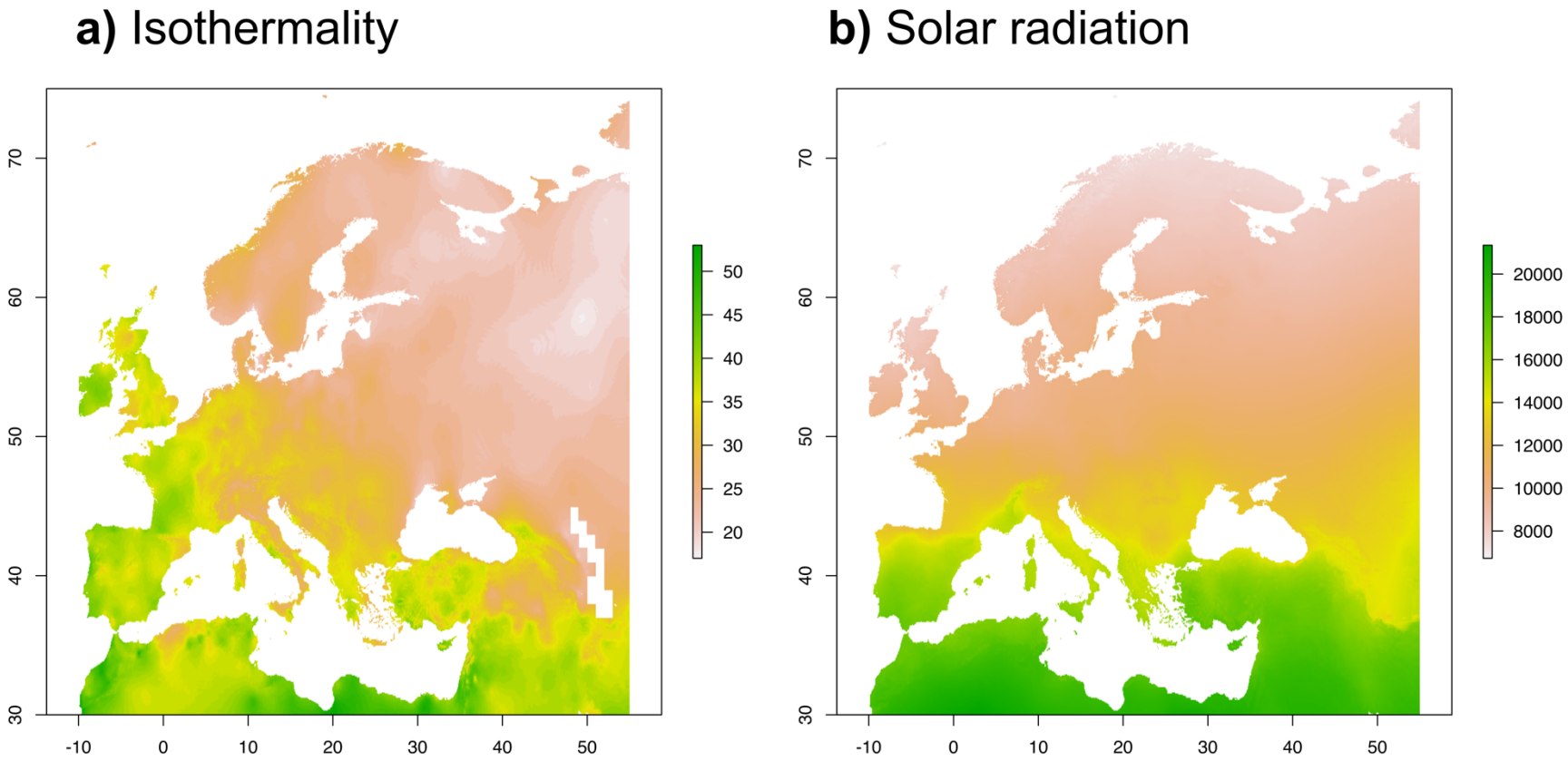


Figure S16. Europe maps of (a) isothermality (i.e., the ratio of mean diurnal range to temperature annual range (*100)) and (b) solar radiation (kJ*m^{-2}*day^{-1}) in 2.5 minutes resolution retrieved from WorldClim database v.2 (Fick & Hijmans, 2017).

Table S5. Results of the structural equation model (SEM; or confirmatory path analysis) based on generalized linear mixed-effect models (GLMMs) and a dataset without missing information of 205 *R. auricomus* populations (see also Fig. 4). We included sexuality (percent sexual seeds to all seeds), maximum number of petals, and heterozygosity (percent of heterozygous sites to all sites) of populations as response (endogenous) variables, and sexuality, ploidy levels (without 3n), reproduction modes, habitat types, and climatic environmental factors as predictor (exogenous) variables. Main genetic clusters found in sNMF analysis were used as random effect variable in GLMMs to account for potential phylogenetic non-independence of included variables. Estimates are not standardized. In contrast to numeric predictors, estimates of categorical predictors are model-estimated mean values in the SEM. For categorical predictors, mean values and significances (superscript letters) are given for within-group levels. The SEM also includes an independence claim "sexuality - reproductive pathway" ($p < .001$) and 11 error terms. P values of significant relationships are highlighted in bold. See Materials and Methods section for further details, and see Tables S4-S6, Figs. 5, 7, S12-21). df = degrees of freedom, * = 4n -6n comparison marginal significant ($p < .1$).

response	predictor	level	mean [%]	estimate	std. error	df	crit. value	p value
sexuality	**ploidy level**	2n	99.0$^{(a)}$	4.01	0.76	2	5.28	**< 0.05**
		4n	3.86$^{(b)}$	-3.62	0.19	2	-19.23	**< 0.01**
		6n	5.70$^{(c*)}$	-5.03	0.82	2	-6.13	**< 0.05**
sexuality	**habitat type**	forest	14.9$^{(a)}$	-1.51	0.38	2	-4.00	0.06
		open habitats	15.6$^{(a)}$	-1.58	0.45	2	-3.50	0.07
sexuality	**isothermality**			-1.06	0.36	192	-2.94	**< 0.01**
sexuality	**temperature seasonality**			-0.34	0.23	192	-1.49	0.14
sexuality	**temperature seasonality : solar rad**			-0.12	0.21	192	-0.59	0.56
sexuality	**annual precipitation**			0.04	0.32	192	0.11	0.91
sexuality	**solar radiation**			1.24	0.33	192	3.80	**< 0.001**
sexuality	**solar radiation : altitude**			-0.33	0.27	192	-1.22	0.22
sexuality	**altitude**			0.53	0.40	192	1.31	0.19
petals	**sexuality**			0.70	0.11	201	6.59	**< 0.001**
heterozygosity	**sexuality**			-0.18	0.67	191	-0.27	0.79
heterozygosity	**ploidy level**	2n	0.61$^{(c)}$	9.75	0.29	2	33.69	**< 0.001**
		4n	1.28$^{(b)}$	10.10	0.16	2	63.40	**< 0.001**
		6n	1.72$^{(a)}$	10.26	0.21	2	49.13	**< 0.001**
heterozygosity	**reproduction mode**	sexual	0.64$^{(a)}$	9.88	0.57	2	17.28	**< 0.01**
		fac. apomictic	1.28$^{(b)}$	10.13	0.13	2	80.22	**< 0.001**
		obl. apomictic	1.36$^{(b)}$	10.10	0.14	2	69.96	**< 0.001**
heterozygosity	**habitat type**	forests	1.27$^{(a)}$	10.09	0.15	2	67.07	**< 0.001**
		open habitats	1.13$^{(b)}$	9.98	0.15	2	65.17	**< 0.001**
heterozygosity	**temperature seasonality**			0.08	0.03	191	2.56	**< 0.05**

heterozygosity	**temperature seasonality : solar radiation**	-0.03	0.02	191	-1.26	0.21
heterozygosity	**annual precipitation**	0.05	0.034	191	1.34	0.18
heterozygosity	**annual precipitation : solar radiation**	-0.03	0.03	191	-0.85	0.40
heterozygosity	**solar radiation**	0.01	0.04	191	0.28	0.78

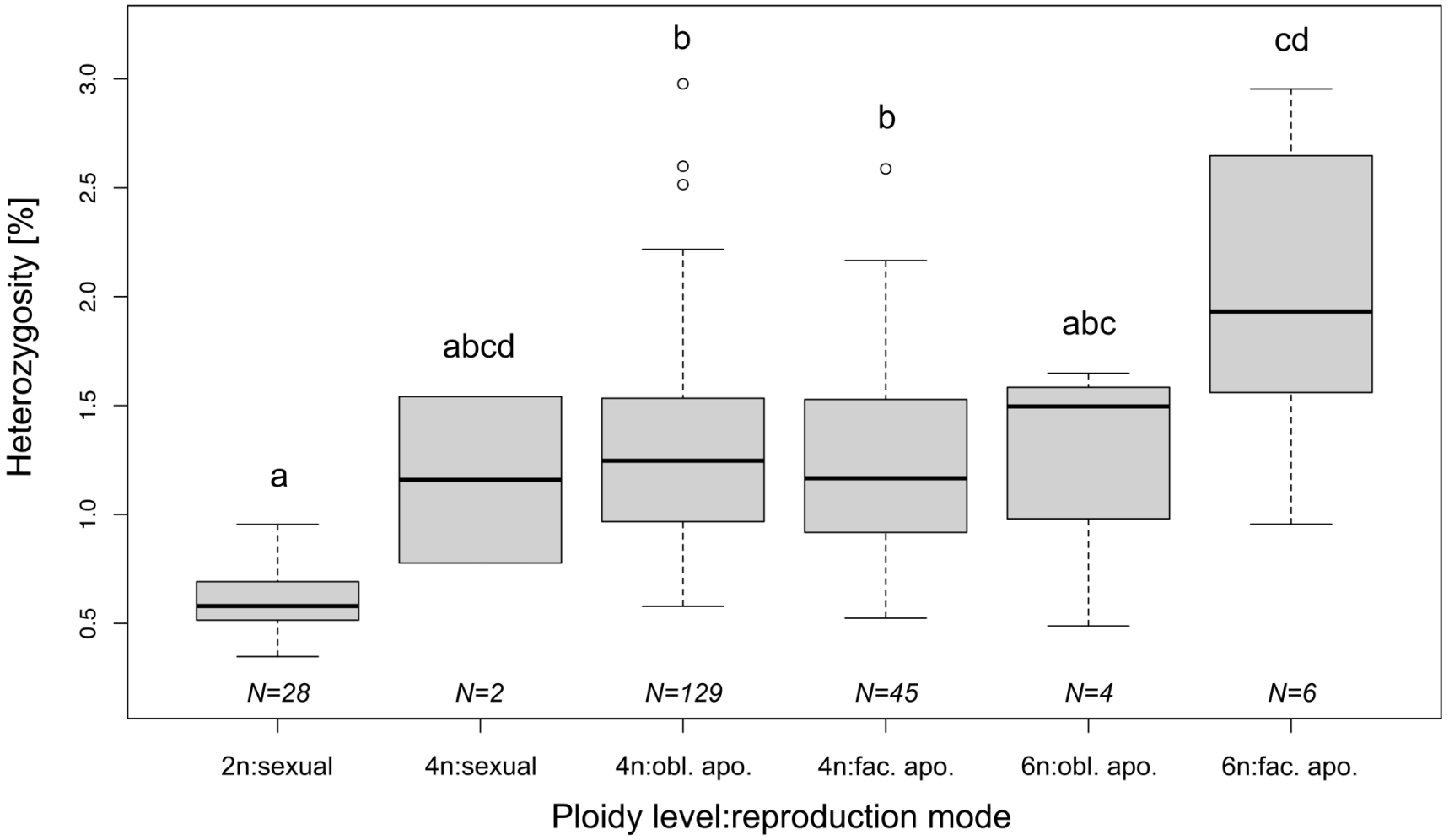


Figure S17. Boxplots of (mean) heterozygosity per population (percent heterozygous sites to all sites) in relation to different ploidy level-reproduction mode levels (interaction). Letters above indicate significant/non-significant differences between/among groups. We checked differences among groups with Kruskal Wallis test followed by a pairwise Wilcox rank sum-test. N = number of populations per group level.

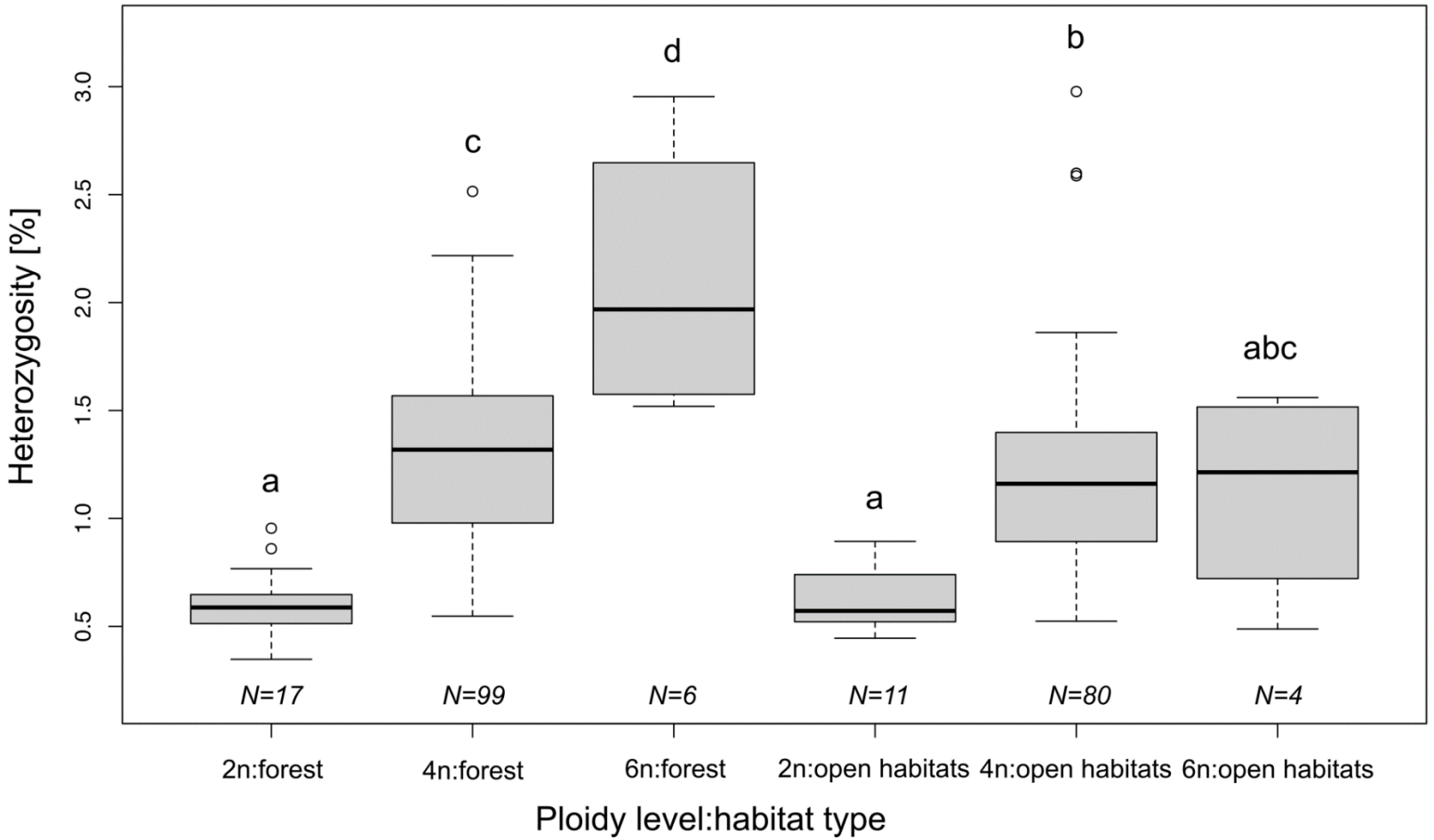


Figure S18. Boxplots of (mean) heterozygosity per population (percent heterozygous sites to all sites) in relation to different ploidy level-habitat type levels (interaction). Letters above indicate significant/non-significant differences between/among groups. We checked differences among groups with Kruskal Wallis test followed by a pairwise Wilcox rank sum-test. N = number of populations per group level.

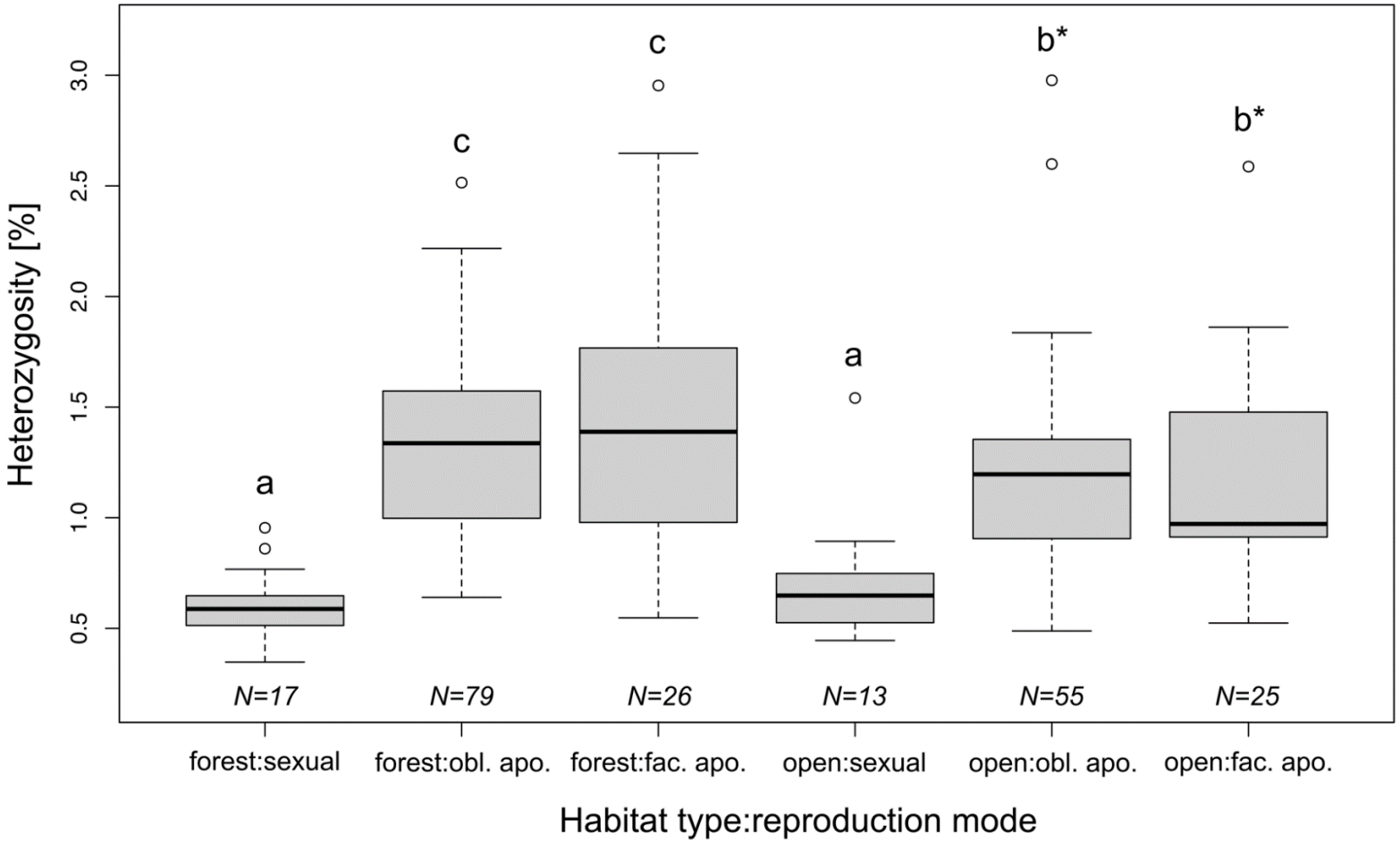


Figure S19. Boxplots of (mean) heterozygosity per population (percent heterozygous sites to all sites) in relation to different habitat type-reproduction mode levels (interaction). Letters above indicate significant/non-significant differences between/among groups. We checked differences among groups with Kruskal Wallis test followed by a pairwise Wilcox rank sum-test (* = b marginal significant different to c, $p < .1$). N = number of populations per group level.

Table S6. Results of the generalized linear mixed effect models (GLMMs) between heterozygosity (percent of heterozygous sites to all sites) and environmental factors (and their interactions) including 220 *R. auricomus* populations. Heterozygosity was obtained from RADseq data (1-4 individuals per population, see Table S3). P values of significant relationships are highlighted in bold. bio4 = temperature seasonality, bio12 = annual precipitation, df = degrees of freedom.

predictor	estimate	std. error	df	t value	p value
temperature seasonality	0.01	0.03	212	0.38	0.71
temperature seasonality : solar radiation	-0.10	0.02	212	-4.29	**< 0.001**
annual precipitation	-0.06	0.03	212	-1.84	0.07
annual precipitation : solar radiation	-0.12	0.03	212	-3.91	**< 0.001**
solar radiation	-0.07	0.03	212	-2.03	**< 0.05**

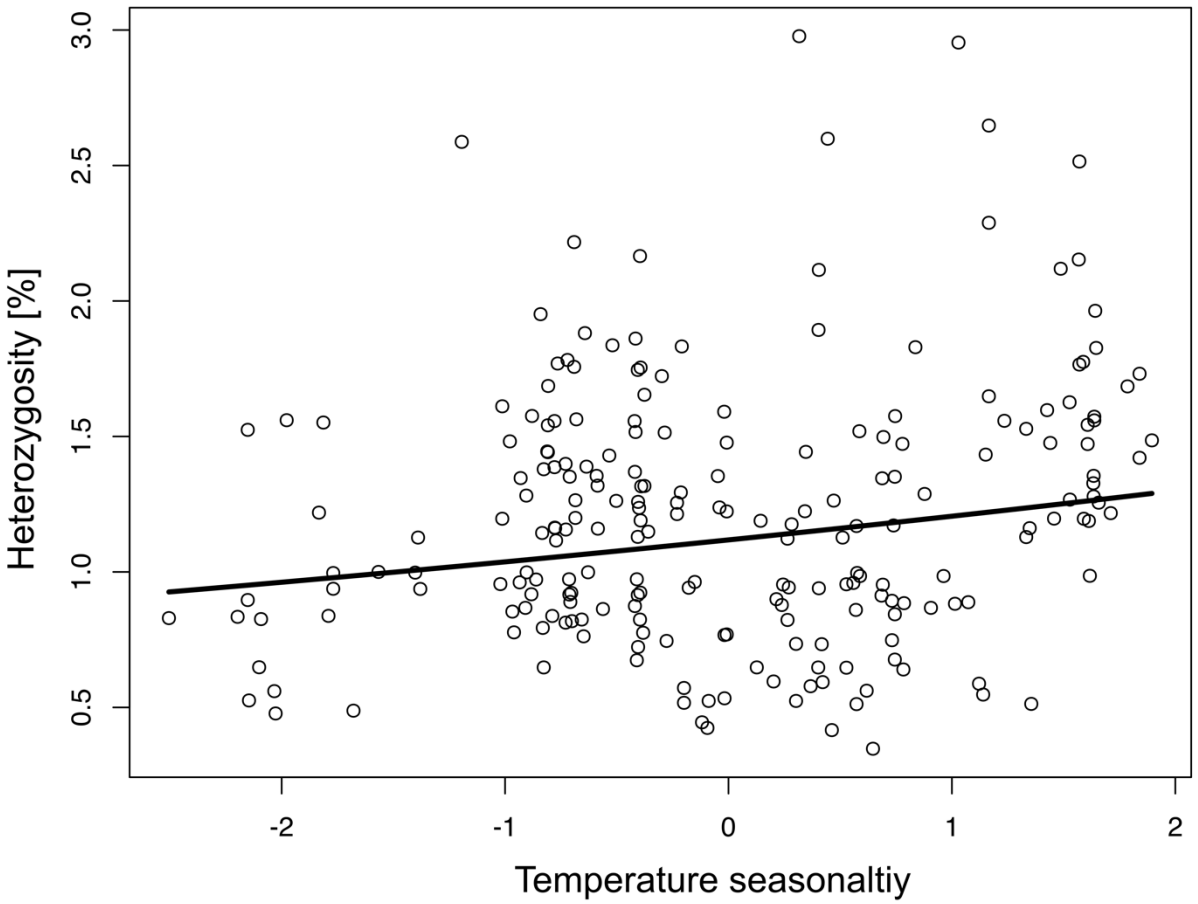


Figure S20. Regression plot illustrating the relationship between heterozygosity (percent of heterozygous sites to all sites; response) and temperature seasonality (bio4; predictor) of 220 *R. auricomus* populations. The solid curve indicates a significant relationship (path analysis, $p < .05$, Table 1, S5). The regression line was drawn based on a separate GLM (gaussian) analysis.

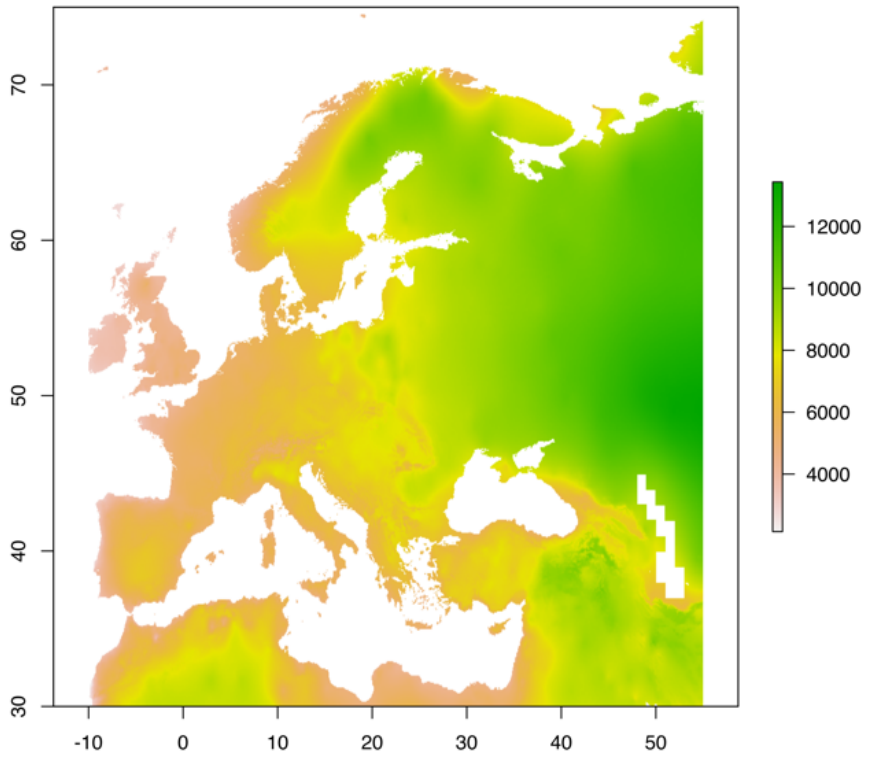


Figure S21. Europe map of temperature seasonality (standard deviation*100) in 2.5 minutes resolution retrieved from WorldClim database v.2 (Fick & Hijmans, 2017).

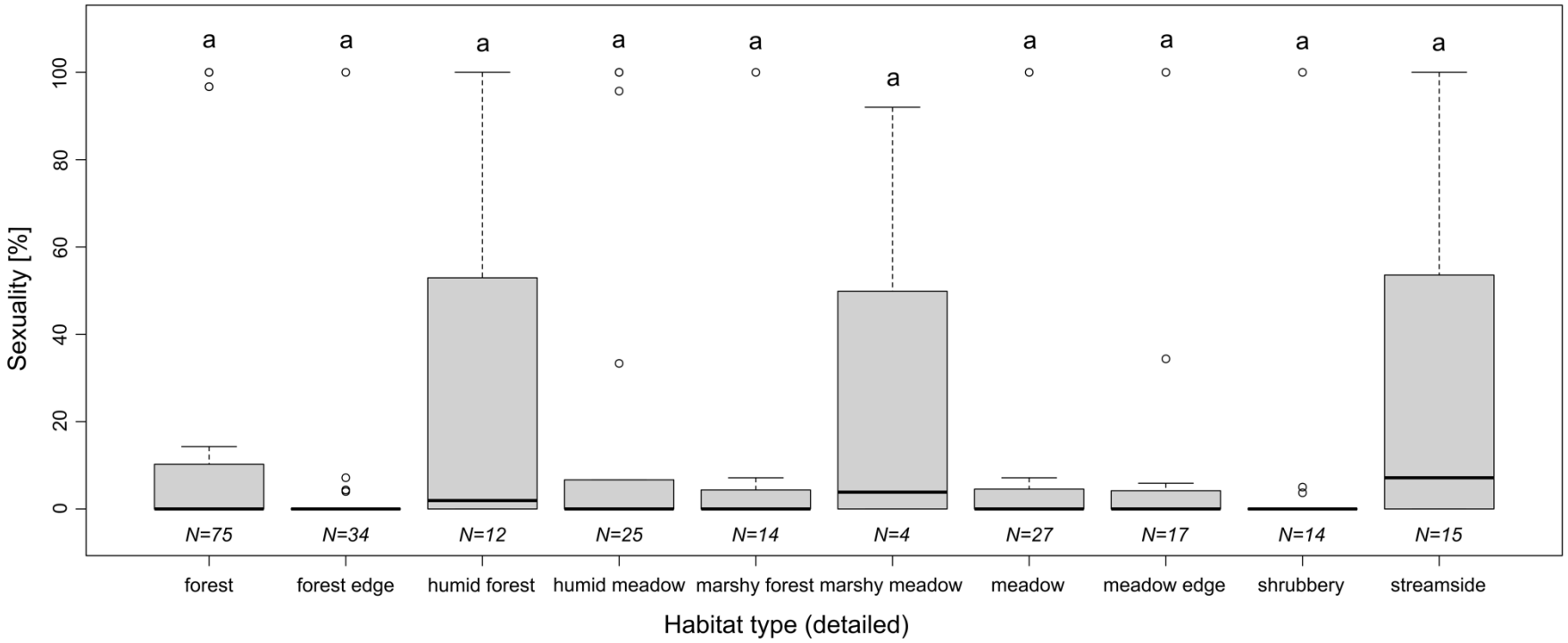


Figure S22. Boxplots of population sexuality (percent of sexual seeds to all seeds) in relation to different habitat types (detailed analysis). Letters above boxplots indicate non-significant differences between/among groups. The Kruskal-Wallis test revealed marginal significant differences among habitat types ($Chi^2 = 15.51$, $df = 9$, $p = .08$), but posthoc comparisons found no significant ($p < .1$) between-level differences.

References

Barke, B. H., Daubert, M., & Hörandl, E. (2018). Establishment of apomixis in diploid F_2 hybrids and inheritance of apospory from F_1 to F_2 hybrids of the *Ranunculus auricomus* complex. *Frontiers in Plant Science*, *9*, 1–12.

Doležel, J., Greilhuber, J., & Suda, J. (2007). Estimation of nuclear DNA content in plants using flow cytometry. *Nature Protocols*, *2*, 2233–2244. https://doi.org/10.1038/nprot.2007.310

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*, 4302–4315. https://doi.org/10.1002/joc.5086

Freeland, J. R., Kirk, H., & Petersen, S. D. (2011). *Molecular Ecology* (sec ed). Chichester, England: Wiley. https://doi.org/10.1002/9780470979365

Hojsgaard, D., & Hörandl, E. (2019). The rise of apomixis in natural plant populations. *Frontiers in Plant Science*, *10*, 1–13. https://doi.org/10.3389/fpls.2019.00358

Hörandl, E., Jakubowsky, G., & Dobeš, C. (2001). Isozyme and morphological diversity within apomictic and sexual taxa of the *Ranunculus auricomus* complex. *Plant Systematics and Evolution*, *226*, 165–185. https://doi.org/10.1007/s006060170064

Karbstein, K., Tomasello, S., Hodač, L., Dunkel, F. G., Daubert, M., & Hörandl, E. (2020a). Phylogenomics supported by geometric morphometrics reveals delimitation of sexual species within the polyploid apomictic *Ranunculus auricomus* complex (Ranunculaceae). *Taxon*, *69*, 1191–1220. https://doi.org/10.1002/tax.12365.

Klatt, S., Hadacek, F., Hodač, L., Brinkmann, G., Eilerts, M., Hojsgaard, D., & Hörandl, E. (2016). Photoperiod extension enhances sexual megaspore formation and triggers metabolic reprogramming in facultative apomictic *Ranunculus auricomus*. *Frontiers in Plant Science*, *7*, 1–13. https://doi.org/10.3389/fpls.2016.00278

Larsson, A. (2014). AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*, *30*, 3276–3278. https://doi.org/10.1093/bioinformatics/btu531

Lo, E. Y. Y., Stefanović, S., & Dickinson, T. A. (2013). Geographical parthenogenesis in Pacific Northwest hawthorns (*Crataegus*; Rosaceae). *Botany*, *91*, 107–116. https://doi.org/10.1139/cjb-2012-0073

Matzk, F., Meister, A., & Schubert, I. (2000). An efficient screen for reproductive pathways using mature seeds of monocots and dicots. *Plant Journal*, *21*, 97–108. https://doi.org/10.1046/j.1365-313X.2000.00647.x

Paule, J., Dunkel, F. G., Schmidt, M., & Gregor, T. (2018). Climatic differentiation in polyploid apomictic *Ranunculus auricomus* complex in Europe. *BMC Ecology*, *18*, 1–12. https://doi.org/https://doi.org/10.1186/s12898-018-0172-1

Paun, O., Greilhuber, J., Temsch, E. M., & Hörandl, E. (2006). Patterns, sources and ecological implications of clonal diversity in apomictic *Ranunculus carpaticola* (*Ranunculus auricomus* complex, Ranunculaceae). *Molecular Ecology*, *15*, 897–910. https://doi.org/10.1111/j.1365-294X.2006.02800.x

Chapter 5

Table S1 (electronical supplement). Location details of sampled *R. auricomus* populations across Europe. We investigated 235 sexual and apomictic populations (see also Fig. 2). Population ID, taxon name, ploidy, main reproduction mode (see also Karbstein et al., 2021, assumed reproduction modes due to missing data in brackets), locality by country and ISO code 3166-2, collection date, altitude (in meter above sea level, m.a.s.l.), latitude (N, decimal), longitude (E, decimal), habitat, collector, herbarium voucher specimens ($^{\#}$ = holotype, $^{+}$ = private herbarium F.G.Dunkel), and available samples per dataset (RADseq, TE = target enrichment dataset, and CP = plastome dataset) and individuals in network analyses (tested polyploids) are given.

The electronical supplement is available online: https://doi.org/10.1101/2021.08.30.458250

Text S1. HYBPHYLOMAKER settings (target enrichment analysis).

Sequence adapters were removed, and reads were quality-trimmed using TRIMMOMATIC v.0.32 (Bolger et al., 2014), with the default settings implemented in HYBPHYLOMAKER v.1.6.4. Duplicated reads were removed with FASTUNIQ v.1.1 (Xu et al., 2012). We used the concatenated sequences of target exons separated by stretches of 800 Ns as "pseudo-reference" for read mapping with BWA v.0.7.12 (Li & Durbin, 2010). Consensus sequences of mapped reads were produced with CONSENSUSFIXER v.0.4 (available at: https://github.com/cbg-ethz/ConsensusFixer), as this is the only approach available in HYBPHYLOMAKER for calling ambiguity DNA codes in case of multiple bases per site in the mapped reads. For CONSENSUSFIXER, we used the following settings: Minimum relative abundance of the alternative base ("plurality" in the settings file of HYBPHYLOMAKER) of 0.2 and a minimum read coverage for ambiguity calling ("mincov") of 5.

Consensus sequences were matched to sequences of the target exons to produce *.pslx files using BLAT v.35.1 (Kent, 2002), and combined across samples to produce exon-wise matrices with "assembled_exons_to_fastas.py" (Weitemier et al., 2014). Matrices were aligned with MAFFT v.7.029 (Katoh & Standley, 2013), using the default program settings, and then gene-wise concatenated with AMAS v.1.0 (Borowiec, 2016). We excluded samples with more than 40% of missing data from each exon region, and subsequently, retained only exons including more than 90% of samples.

Table S2 (electronical supplement). Quality trimming and read mapping (target enrichment analysis). Detailed results of quality and duplicate trimming (Trimming) and read mapping to the "pseudo-reference" consisting of concatenated target regions (Mapping).

The electronical supplement is available online: https://doi.org/10.1101/2021.08.30.458250

Table S3 (electronical supplement). Loci selection (target enrichment analysis). We scored exon regions to select the 50 best loci to be phased and used in network analyses (see Materials for details on how the scoring was performed). "R_squared" defines the R^2 of mutational saturation regression curves, and "LBscoresSD" describes the standard deviation of the sample-specific long-branch scores (LB scores). The last column gives information on the amount of parsimony informative sites in each exon region. The selected loci are highlighted in pink.

The electronical supplement is available online: https://doi.org/10.1101/2021.08.30.458250

Table S4 (electronical supplement). Results of the phasing procedure (target enrichment analysis). Information is given on the number of alleles found for each sample in the 50 loci selected for the species delimitation and network analyses. The last column gives information on the best-fitting model of sequence evolution found with ModelTest-NG.

The electronical supplement is available online: https://doi.org/10.1101/2021.08.30.458250

Table S5 (electronical supplement). Read mapping (plastome analysis). Detailed results of the read mapping to the chloroplast reference genome. We used the *Ranunculus repens* plastid genome (GenBank accession number NC_036976; Dann et al., 2017) as assembly reference.

The electronical supplement is available online: https://doi.org/10.1101/2021.08.30.458250

Table S6 (electronical supplement). List of regions used for the chloroplast phylogeny (plastome analysis). The alignments were selected excluding first samples with more than 50% missing data and second alignments including less than 50% of samples. From those selected alignments (71 in total), we excluded samples missing from more than 50% of the alignments (marked in pink). Numbers indicate percentages of missing data (0 = no missing data; empty cell = completely missing).

The electronical supplement is available online: https://doi.org/10.1101/2021.08.30.458250

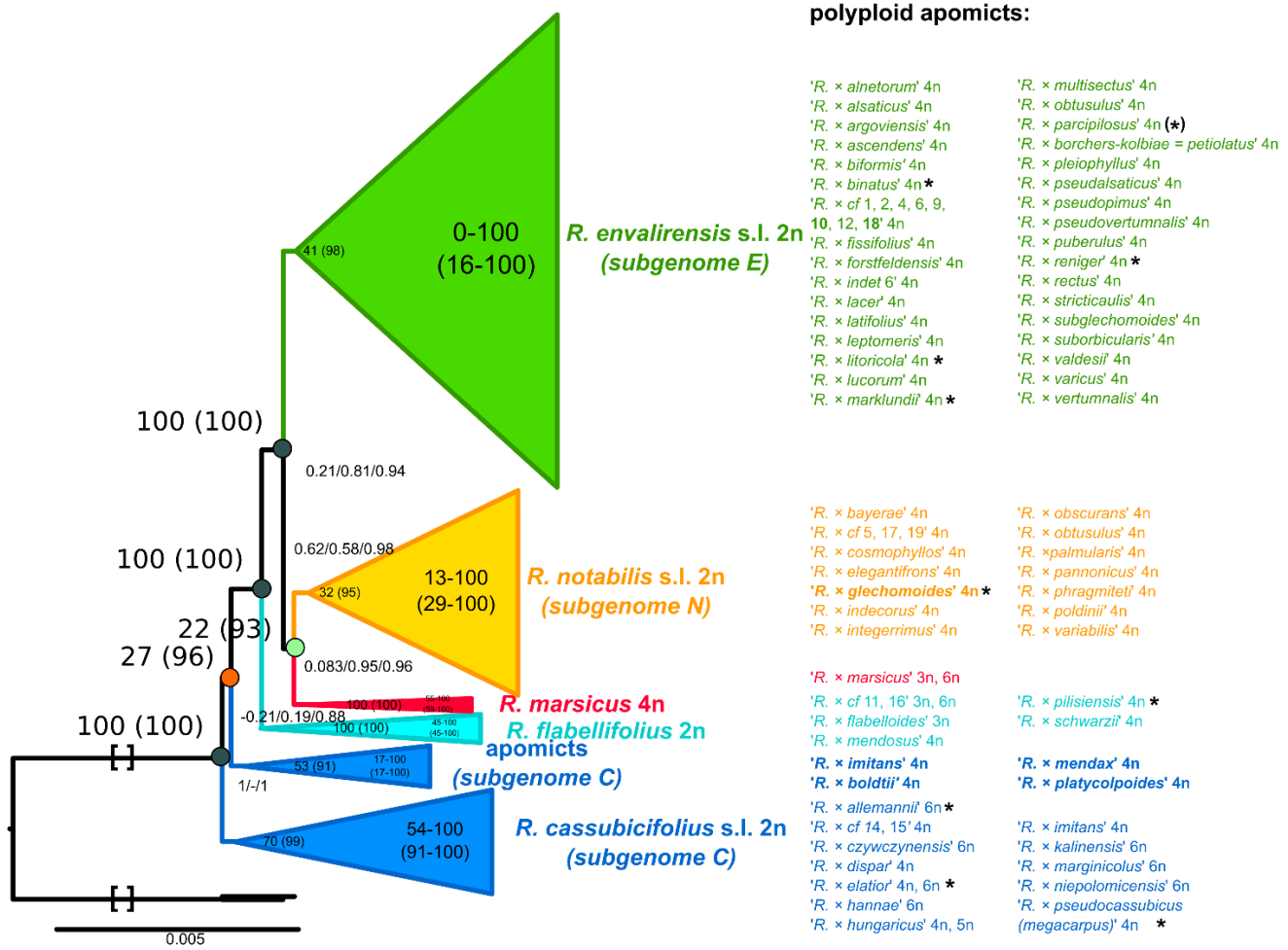


Figure S1. Concatenated ML tree with summarized main clades (min30). Results are based on RAxML-NG results and the "min30" RADseq alignment (280 samples, 33,165 loci, 194,083 SNPs). Felsenstein Bootstrap Proportions (FBP) and Transfer Bootstrap Expectation (TBE, in brackets) values are displayed on branches. Range of BT values (i.e., FBP (TBE) value range of all internal nodes) is shown for each main clade. Names of sexual species and apomictic taxa resolved in the main clades are listed to the right in corresponding colors, with samples resolved differently compared to the "min10" ML tree in bold. Ploidy levels are listed to the right of species/taxon names. See Figshare data repository for the complete ML tree. * = monophyletic polyploid apomictic taxon. Squared brackets: A part of the branch was cut for illustrative purposes.

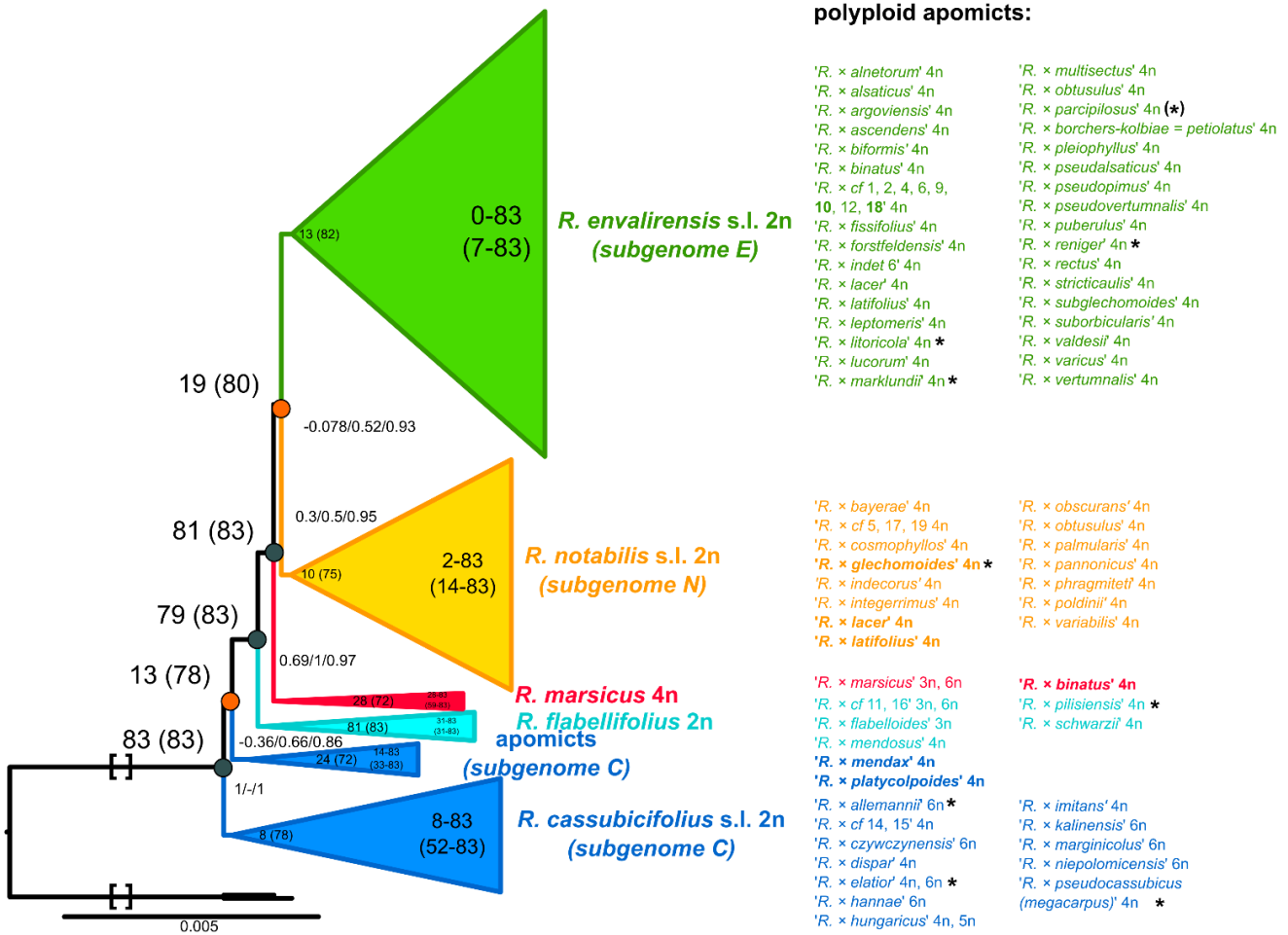
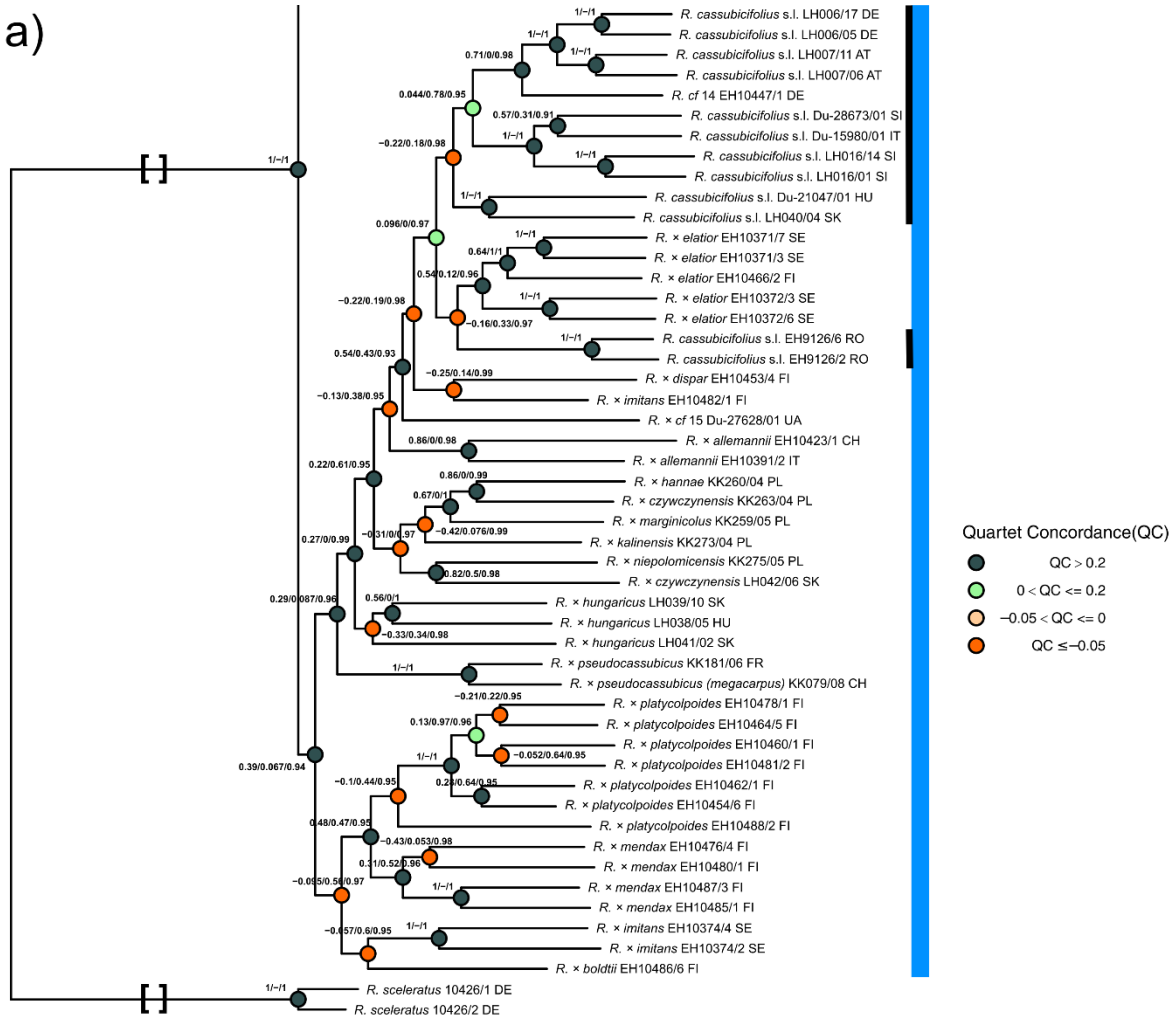
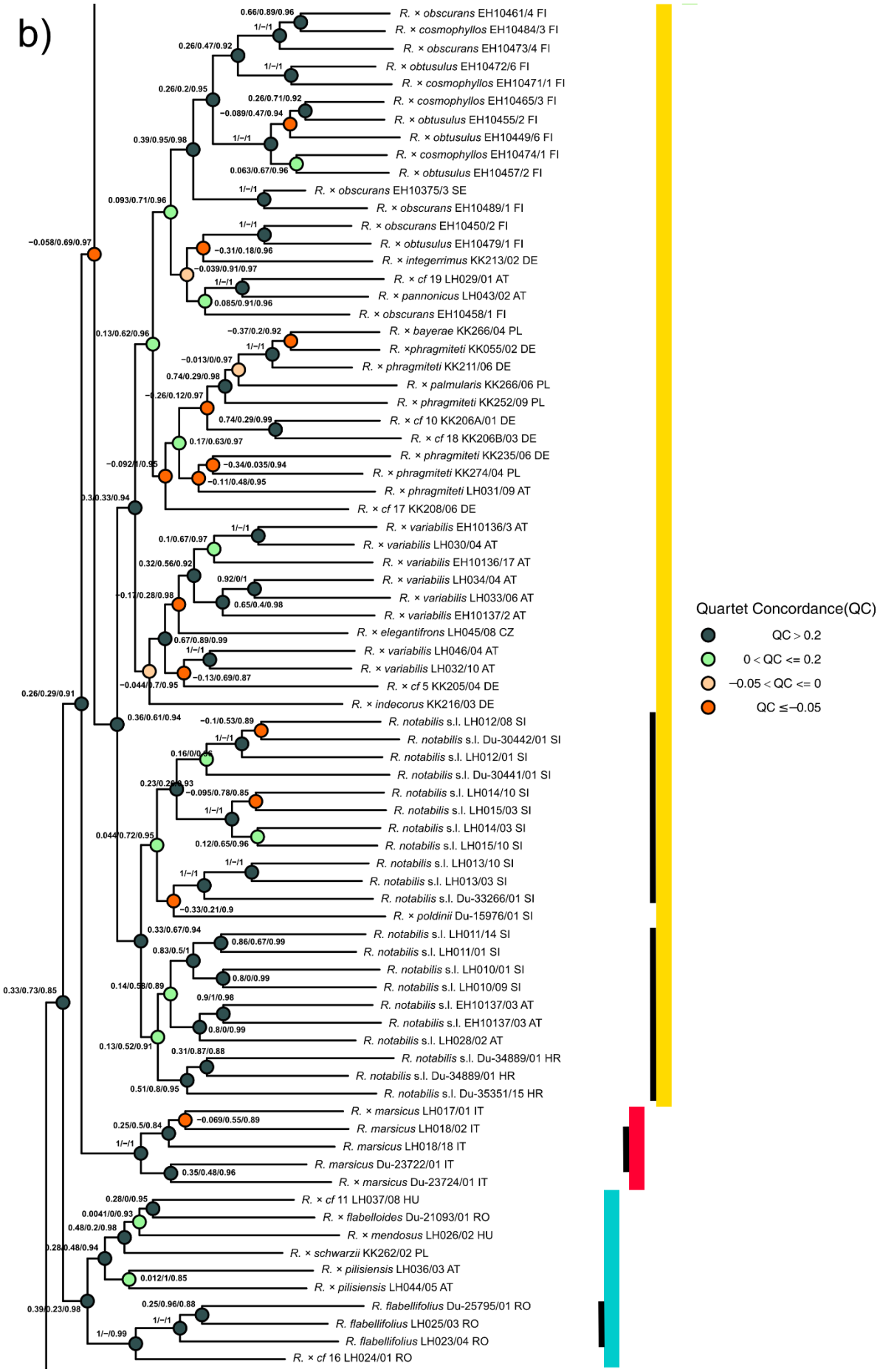
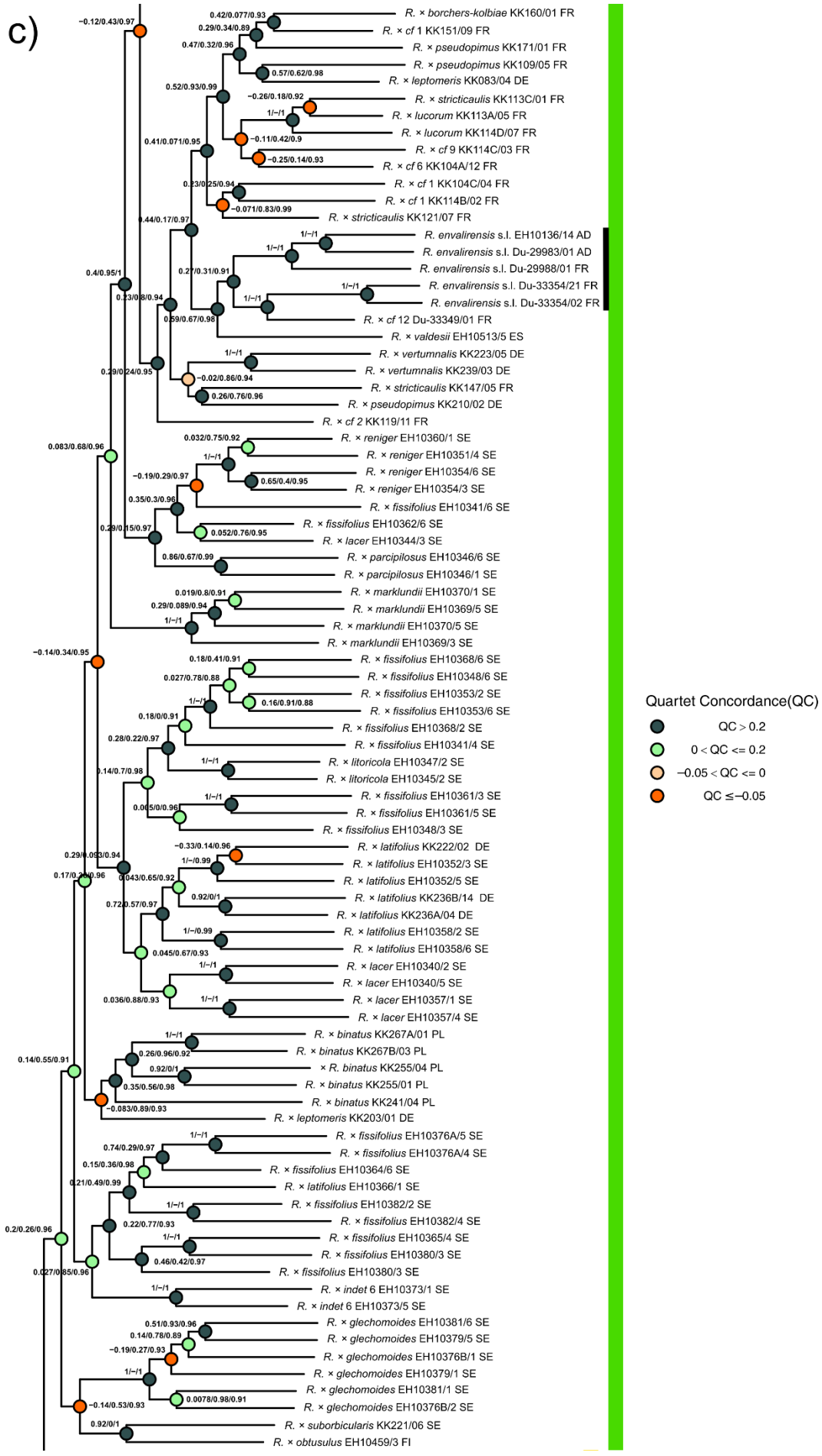


Figure S2. Concatenated ML tree with summarized main clades (min50). Results are based on RAxML-NG results and the "min50" RADseq alignment (280 samples, 11,196 loci, 64,554 SNPs). Felsenstein Bootstrap Proportions (FBP) and Transfer Bootstrap Expectation (TBE, in brackets) values are displayed on branches. Range of BT values (i.e., FBP (TBE) value range of all internal nodes) is shown for each main clade. Names of sexual species and apomictic taxa resolved in the main clades are listed to the right in corresponding colors, with samples resolved differently compared to the "min10" ML tree in bold. Ploidy levels are listed to the right of species/taxon names. See Figshare data repository for the complete ML tree. * = monophyletic polyploid apomictic taxon. Squared brackets: A part of the branch was cut for illustrative purposes.


a)
R. cassubicifolius s.l. LH006/17 DE
R. cassubicifolius s.l. LH006/05 DE
R. cassubicifolius s.l. LH007/11 AT
R. cassubicifolius s.l. LH007/06 AT
R. cf 14 EH10447/1 DE
R. cassubicifolius s.l. Du-28673/01 SI
R. cassubicifolius s.l. Du-15980/01 IT
R. cassubicifolius s.l. LH016/14 SI
R. cassubicifolius s.l. LH016/01 SI
R. cassubicifolius s.l. Du-21047/01 HU
R. cassubicifolius s.l. LH040/04 SK
R. × elatior EH10371/7 SE
R. × elatior EH10371/3 SE
R. × elatior EH10466/2 FI
R. × elatior EH10372/3 SE
R. × elatior EH10372/6 SE
R. cassubicifolius s.l. EH9126/6 RO
R. cassubicifolius s.l. EH9126/2 RO
R. × dispar EH10453/4 FI
R. × imitans EH10482/1 FI
R. × cf 15 Du-27628/01 UA
R. × allemannii EH10423/1 CH
R. × allemannii EH10391/2 IT
R. × hannae KK260/04 PL
R. × czywczynensis KK263/04 PL
R. × marginicolus KK259/05 PL
R. × kalinensis KK273/04 PL
R. × niepolomicensis KK275/05 PL
R. × czywczynensis LH042/06 SK
R. × hungaricus LH039/10 SK
R. × hungaricus LH038/05 HU
R. × hungaricus LH041/02 SK
R. × pseudocassubicus KK181/06 FR
R. × pseudocassubicus (megacarpus) KK079/08 CH
R. × platycolpoides EH10478/1 FI
R. × platycolpoides EH10464/5 FI
R. × platycolpoides EH10460/1 FI
R. × platycolpoides EH10481/2 FI
R. × platycolpoides EH10462/1 FI
R. × platycolpoides EH10454/6 FI
R. × platycolpoides EH10488/2 FI
R. × mendax EH10476/4 FI
R. × mendax EH10480/1 FI
R. × mendax EH10487/3 FI
R. × mendax EH10485/1 FI
R. × imitans EH10374/4 SE
R. × imitans EH10374/2 SE
R. × boldtii EH10486/6 FI
R. sceleratus 10426/1 DE
R. sceleratus 10426/2 DE
Quartet Concordance(QC)
QC > 0.2
0 < QC <= 0.2
−0.05 < QC <= 0
QC ≤ −0.05


b)
R. × obscurans EH10461/4 FI
R. × cosmophyllos EH10484/3 FI
R. × obscurans EH10473/4 FI
R. × obtusulus EH10472/6 FI
R. × cosmophyllos EH10471/1 FI
R. × cosmophyllos EH10465/3 FI
R. × obtusulus EH10455/2 FI
R. × obtusulus EH10449/6 FI
R. × cosmophyllos EH10474/1 FI
R. × obtusulus EH10457/2 FI
R. × obscurans EH10375/3 SE
R. × obscurans EH10489/1 FI
R. × obscurans EH10450/2 FI
R. × obtusulus EH10479/1 FI
R. × integerrimus KK213/02 DE
R. × cf 19 LH029/01 AT
R. × pannonicus LH043/02 AT
R. × obscurans EH10458/1 FI
R. × bayerae KK266/04 PL
R. ×phragmiteti KK055/02 DE
R. × phragmiteti KK211/06 DE
R. × palmularis KK266/06 PL
R. × phragmiteti KK252/09 PL
R. × cf 10 KK206A/01 DE
R. × cf 18 KK206B/03 DE
R. × phragmiteti KK235/06 DE
R. × phragmiteti KK274/04 PL
R. × phragmiteti LH031/09 AT
R. × cf 17 KK208/06 DE
R. × variabilis EH10136/3 AT
R. × variabilis LH030/04 AT
R. × variabilis EH10136/17 AT
R. × variabilis LH034/04 AT
R. × variabilis LH033/06 AT
R. × variabilis EH10137/2 AT
R. × elegantifrons LH045/08 CZ
R. × variabilis LH046/04 AT
R. × variabilis LH032/10 AT
R. × cf 5 KK205/04 DE
R. × indecorus KK216/03 DE
R. notabilis s.l. LH012/08 SI
R. notabilis s.l. Du-30442/01 SI
R. notabilis s.l. LH012/01 SI
R. notabilis s.l. Du-30441/01 SI
R. notabilis s.l. LH014/10 SI
R. notabilis s.l. LH015/03 SI
R. notabilis s.l. LH014/03 SI
R. notabilis s.l. LH015/10 SI
R. notabilis s.l. LH013/10 SI
R. notabilis s.l. LH013/03 SI
R. notabilis s.l. Du-33266/01 SI
R. × poldinii Du-15976/01 SI
R. notabilis s.l. LH011/14 SI
R. notabilis s.l. LH011/01 SI
R. notabilis s.l. LH010/01 SI
R. notabilis s.l. LH010/09 SI
R. notabilis s.l. EH10137/03 AT
R. notabilis s.l. EH10137/03 AT
R. notabilis s.l. LH028/02 AT
R. notabilis s.l. Du-34889/01 HR
R. notabilis s.l. Du-34889/01 HR
R. notabilis s.l. Du-35351/15 HR
R. × marsicus LH017/01 IT
R. marsicus LH018/02 IT
R. marsicus LH018/18 IT
R. marsicus Du-23722/01 IT
R. × marsicus Du-23724/01 IT
R. × cf 11 LH037/08 HU
R. × flabelloides Du-21093/01 RO
R. × mendosus LH026/02 HU
R. × schwarzii KK262/02 PL
R. × pilisiensis LH036/03 AT
R. × pilisiensis LH044/05 AT
R. flabellifolius Du-25795/01 RO
R. flabellifolius LH025/03 RO
R. flabellifolius LH023/04 RO
R. × cf 16 LH024/01 RO
Quartet Concordance(QC)
QC > 0.2
0 < QC <= 0.2
−0.05 < QC <= 0
QC ≤−0.05


c)
R. × borchers-kolbiae KK160/01 FR
R. × cf 1 KK151/09 FR
R. × pseudopimus KK171/01 FR
R. × pseudopimus KK109/05 FR
R. × leptomeris KK083/04 DE
R. × stricticaulis KK113C/01 FR
R. × lucorum KK113A/05 FR
R. × lucorum KK114D/07 FR
R. × cf 9 KK114C/03 FR
R. × cf 6 KK104A/12 FR
R. × cf 1 KK104C/04 FR
R. × cf 1 KK114B/02 FR
R. × stricticaulis KK121/07 FR
R. envalirensis s.l. EH10136/14 AD
R. envalirensis s.l. Du-29983/01 AD
R. envalirensis s.l. Du-29988/01 FR
R. envalirensis s.l. Du-33354/21 FR
R. envalirensis s.l. Du-33354/02 FR
R. × cf 12 Du-33349/01 FR
R. × valdesii EH10513/5 ES
R. × vertumnalis KK223/05 DE
R. × vertumnalis KK239/03 DE
R. × stricticaulis KK147/05 FR
R. × pseudopimus KK210/02 DE
R. × cf 2 KK119/11 FR
R. × reniger EH10360/1 SE
R. × reniger EH10351/4 SE
R. × reniger EH10354/6 SE
R. × reniger EH10354/3 SE
R. × fissifolius EH10341/6 SE
R. × fissifolius EH10362/6 SE
R. × lacer EH10344/3 SE
R. × parcipilosus EH10346/6 SE
R. × parcipilosus EH10346/1 SE
R. × marklundii EH10370/1 SE
R. × marklundii EH10369/5 SE
R. × marklundii EH10370/5 SE
R. × marklundii EH10369/3 SE
R. × fissifolius EH10368/6 SE
R. × fissifolius EH10348/6 SE
R. × fissifolius EH10353/2 SE
R. × fissifolius EH10353/6 SE
R. × fissifolius EH10368/2 SE
R. × fissifolius EH10341/4 SE
R. × litoricola EH10347/2 SE
R. × litoricola EH10345/2 SE
R. × fissifolius EH10361/3 SE
R. × fissifolius EH10361/5 SE
R. × fissifolius EH10348/3 SE
R. × latifolius KK222/02 DE
R. × latifolius EH10352/3 SE
R. × latifolius EH10352/5 SE
R. × latifolius KK236B/14 DE
R. × latifolius KK236A/04 DE
R. × latifolius EH10358/2 SE
R. × latifolius EH10358/6 SE
R. × lacer EH10340/2 SE
R. × lacer EH10340/5 SE
R. × lacer EH10357/1 SE
R. × lacer EH10357/4 SE
R. × binatus KK267A/01 PL
R. × binatus KK267B/03 PL
× R. binatus KK255/04 PL
R. × binatus KK255/01 PL
R. × binatus KK241/04 PL
R. × leptomeris KK203/01 DE
R. × fissifolius EH10376A/5 SE
R. × fissifolius EH10376A/4 SE
R. × fissifolius EH10364/6 SE
R. × latifolius EH10366/1 SE
R. × fissifolius EH10382/2 SE
R. × fissifolius EH10382/4 SE
R. × fissifolius EH10365/4 SE
R. × fissifolius EH10380/3 SE
R. × fissifolius EH10380/3 SE
R. × indet 6 EH10373/1 SE
R. × indet 6 EH10373/5 SE
R. × glechomoides EH10381/6 SE
R. × glechomoides EH10379/5 SE
R. × glechomoides EH10376B/1 SE
R. × glechomoides EH10379/1 SE
R. × glechomoides EH10381/1 SE
R. × glechomoides EH10376B/2 SE
R. × suborbicularis KK221/06 SE
R. × obtusulus EH10459/3 FI
Quartet Concordance(QC)
QC > 0.2
0 < QC <= 0.2
−0.05 < QC <= 0
QC ≤ −0.05

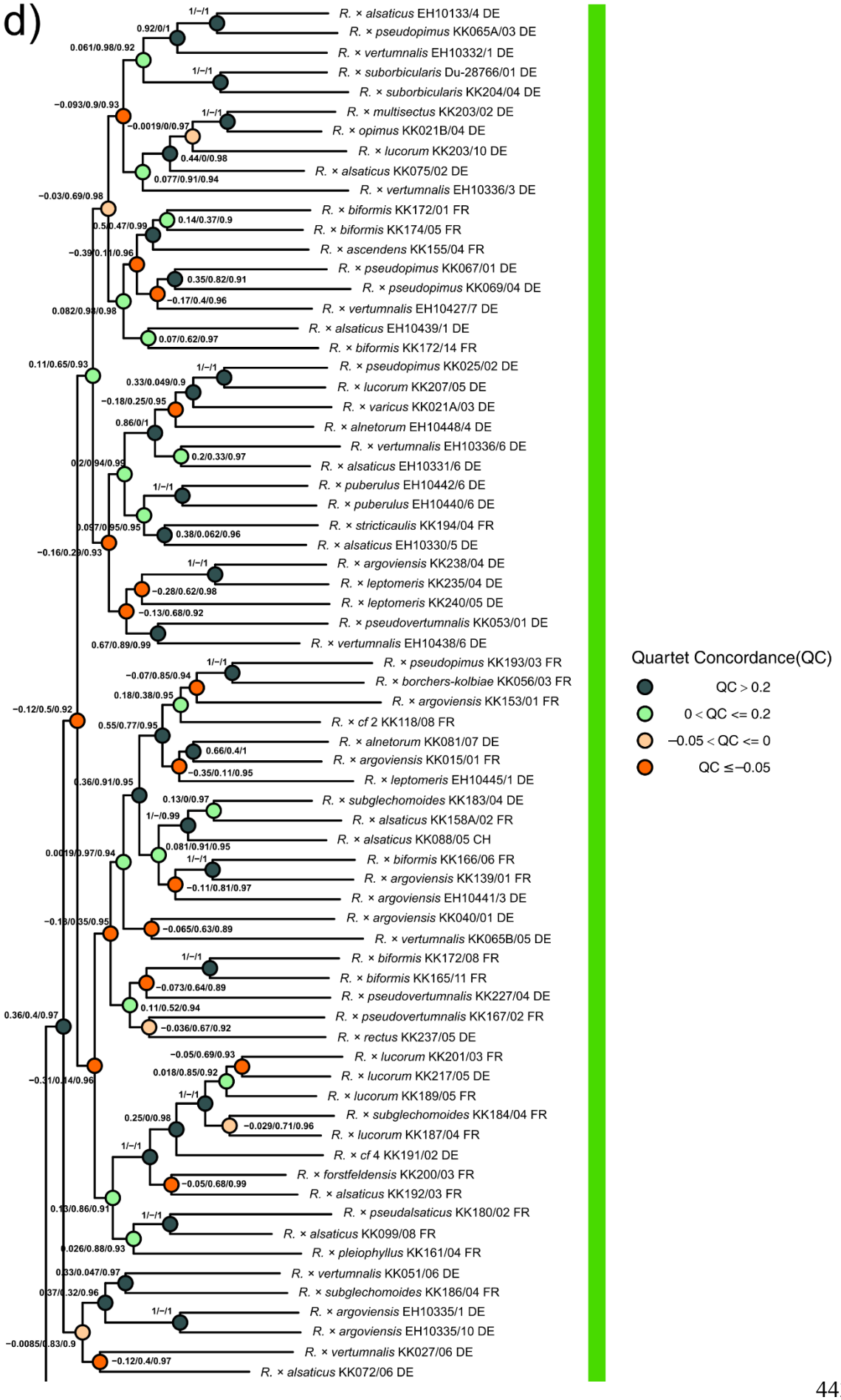
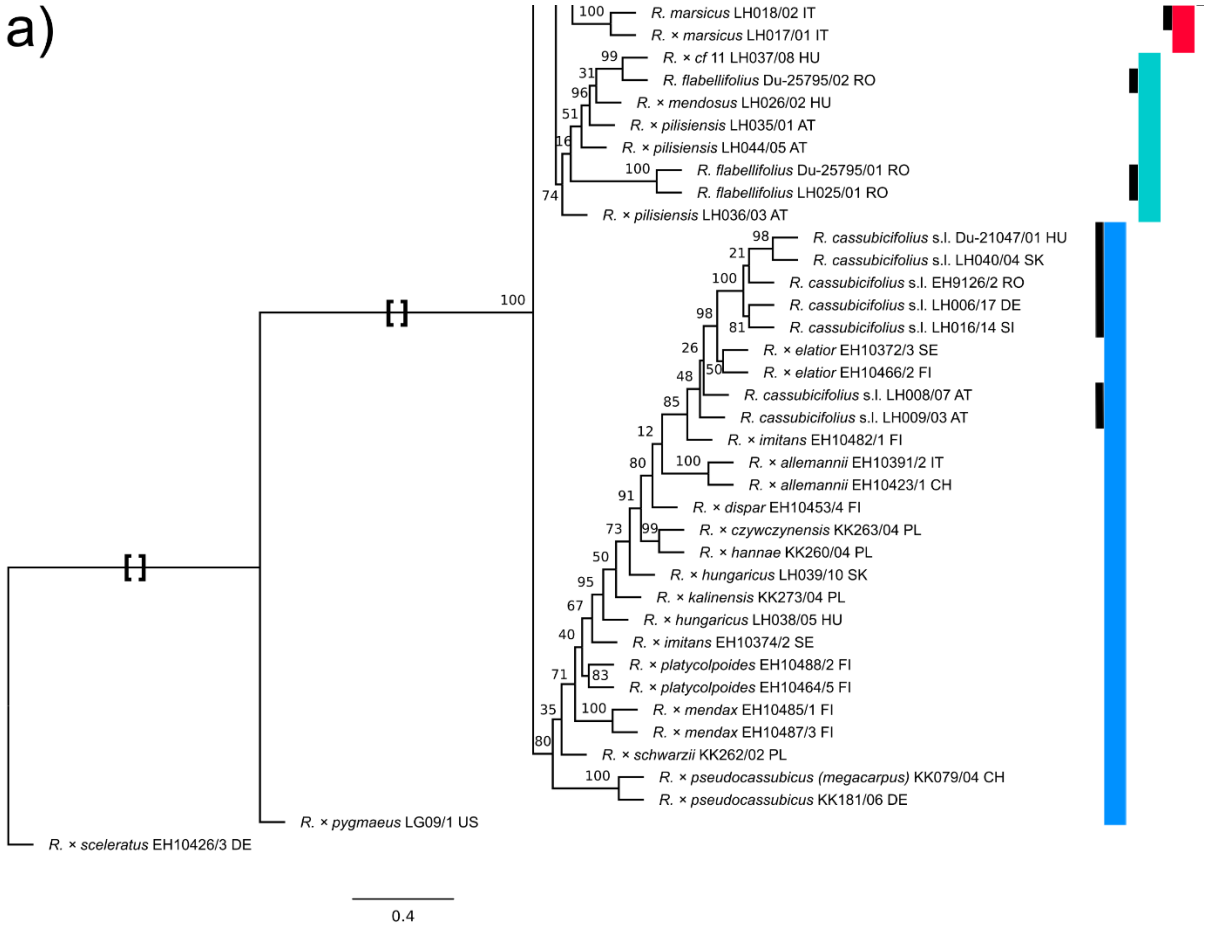
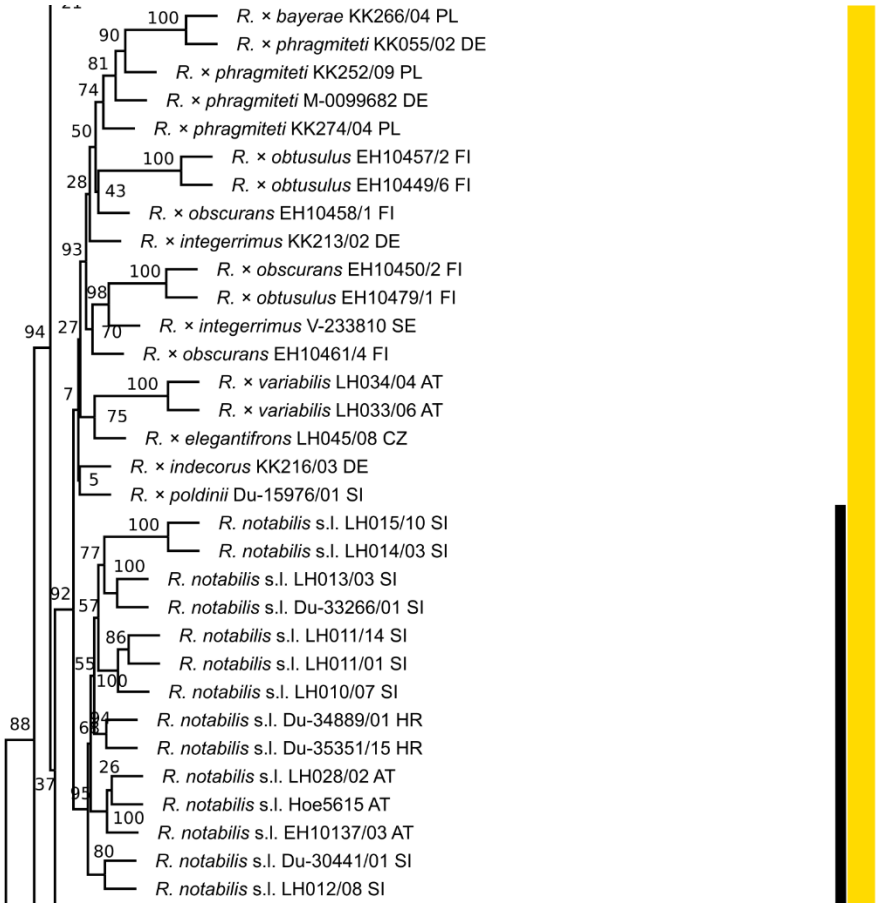

d)
R. × alsaticus EH10133/4 DE
R. × pseudopimus KK065A/03 DE
R. × vertumnalis EH10332/1 DE
R. × suborbicularis Du-28766/01 DE
R. × suborbicularis KK204/04 DE
R. × multisectus KK203/02 DE
R. × opimus KK021B/04 DE
R. × lucorum KK203/10 DE
R. × alsaticus KK075/02 DE
R. × vertumnalis EH10336/3 DE
R. × biformis KK172/01 FR
R. × biformis KK174/05 FR
R. × ascendens KK155/04 FR
R. × pseudopimus KK067/01 DE
R. × pseudopimus KK069/04 DE
R. × vertumnalis EH10427/7 DE
R. × alsaticus EH10439/1 DE
R. × biformis KK172/14 FR
R. × pseudopimus KK025/02 DE
R. × lucorum KK207/05 DE
R. × varicus KK021A/03 DE
R. × alnetorum EH10448/4 DE
R. × vertumnalis EH10336/6 DE
R. × alsaticus EH10331/6 DE
R. × puberulus EH10442/6 DE
R. × puberulus EH10440/6 DE
R. × stricticaulis KK194/04 FR
R. × alsaticus EH10330/5 DE
R. × argoviensis KK238/04 DE
R. × leptomeris KK235/04 DE
R. × leptomeris KK240/05 DE
R. × pseudovertumnalis KK053/01 DE
R. × vertumnalis EH10438/6 DE
R. × pseudopimus KK193/03 FR
R. × borchers-kolbiae KK056/03 FR
R. × argoviensis KK153/01 FR
R. × cf 2 KK118/08 FR
R. × alnetorum KK081/07 DE
R. × argoviensis KK015/01 FR
R. × leptomeris EH10445/1 DE
R. × subglechomoides KK183/04 DE
R. × alsaticus KK158A/02 FR
R. × alsaticus KK088/05 CH
R. × biformis KK166/06 FR
R. × argoviensis KK139/01 FR
R. × argoviensis EH10441/3 DE
R. × argoviensis KK040/01 DE
R. × vertumnalis KK065B/05 DE
R. × biformis KK172/08 FR
R. × biformis KK165/11 FR
R. × pseudovertumnalis KK227/04 DE
R. × pseudovertumnalis KK167/02 FR
R. × rectus KK237/05 DE
R. × lucorum KK201/03 FR
R. × lucorum KK217/05 DE
R. × lucorum KK189/05 FR
R. × subglechomoides KK184/04 FR
R. × lucorum KK187/04 FR
R. × cf 4 KK191/02 DE
R. × forstfeldensis KK200/03 FR
R. × alsaticus KK192/03 FR
R. × pseudalsaticus KK180/02 FR
R. × alsaticus KK099/08 FR
R. × pleiophyllus KK161/04 FR
R. × vertumnalis KK051/06 DE
R. × subglechomoides KK186/04 FR
R. × argoviensis EH10335/1 DE
R. × argoviensis EH10335/10 DE
R. × vertumnalis KK027/06 DE
R. × alsaticus KK072/06 DE
Quartet Concordance(QC)
QC > 0.2
0 < QC <= 0.2
−0.05 < QC <= 0
QC ≤ −0.05

Figure S3. Concatenated ML tree with quartet sampling scores (QS). (a–d) Results are based on RAxML-NG results and the "min10" RADseq alignment (a–d, see also Fig. 4; 280 samples, 97,312 loci, 435,678 SNPs). The quartet concordance score (QC) is the ratio of concordant to both discordant quartets (1: all concordant, > 0: more concordant patterns, < 0: more discordant patterns), the quartet differential score (QD) describes the skewness of both discordant patterns (1: equal, 0.3: skewed, 0: all topologies 1 or 2), and the quartet informativeness score (QI) shows the proportion of informative replicates (1: all informative, 0: none informative; see also Pease et al., 2018). Incomplete lineage sorting (ILS) is indicated by QD values around 1 and means the presence of both discordant topologies (random pattern). QD values towards 0 hint at directional introgression, i.e., the presence of only one particular alternative topology (Pease et al., 2018; Karbstein et al., 2020). We used the R script plot_QC_ggtree.R (available at: https://github.com/-ShuiyinLIU/QS_visualization) and the R packages GGTREE v.2.2.4 (Yu et al., 2017), Treeio vers 1.12.0 (Wang et al., 2020), GGPLOT2 v.3.3.1 (Winston et al., 2020), and APE v.5.4 (Paradis & Schliep, 2019; Paradis et al., 2020) for visualization. Colored bars highlight main clades of Fig. 4a. Black bars indicate sexual species. Names of polyploid apomictic taxa are tentative. Squared brackets: A part of the branch was cut for illustrative purposes.


a)
R. marsicus LH018/02 IT
R. × marsicus LH017/01 IT
R. × cf 11 LH037/08 HU
R. flabellifolius Du-25795/02 RO
R. × mendosus LH026/02 HU
R. × pilisiensis LH035/01 AT
R. × pilisiensis LH044/05 AT
R. flabellifolius Du-25795/01 RO
R. flabellifolius LH025/01 RO
R. × pilisiensis LH036/03 AT
R. cassubicifolius s.l. Du-21047/01 HU
R. cassubicifolius s.l. LH040/04 SK
R. cassubicifolius s.l. EH9126/2 RO
R. cassubicifolius s.l. LH006/17 DE
R. cassubicifolius s.l. LH016/14 SI
R. × elatior EH10372/3 SE
R. × elatior EH10466/2 FI
R. cassubicifolius s.l. LH008/07 AT
R. cassubicifolius s.l. LH009/03 AT
R. × imitans EH10482/1 FI
R. × allemannii EH10391/2 IT
R. × allemannii EH10423/1 CH
R. × dispar EH10453/4 FI
R. × czywczynensis KK263/04 PL
R. × hannae KK260/04 PL
R. × hungaricus LH039/10 SK
R. × kalinensis KK273/04 PL
R. × hungaricus LH038/05 HU
R. × imitans EH10374/2 SE
R. × platycolpoides EH10488/2 FI
R. × platycolpoides EH10464/5 FI
R. × mendax EH10485/1 FI
R. × mendax EH10487/3 FI
R. × schwarzii KK262/02 PL
R. × pseudocassubicus (megacarpus) KK079/04 CH
R. × pseudocassubicus KK181/06 DE
R. × pygmaeus LG09/1 US
R. × sceleratus EH10426/3 DE
0.4

b)


R. × bayerae KK266/04 PL
R. × phragmiteti KK055/02 DE
R. × phragmiteti KK252/09 PL
R. × phragmiteti M-0099682 DE
R. × phragmiteti KK274/04 PL
R. × obtusulus EH10457/2 FI
R. × obtusulus EH10449/6 FI
R. × obscurans EH10458/1 FI
R. × integerrimus KK213/02 DE
R. × obscurans EH10450/2 FI
R. × obtusulus EH10479/1 FI
R. × integerrimus V-233810 SE
R. × obscurans EH10461/4 FI
R. × variabilis LH034/04 AT
R. × variabilis LH033/06 AT
R. × elegantifrons LH045/08 CZ
R. × indecorus KK216/03 DE
R. × poldinii Du-15976/01 SI
R. notabilis s.l. LH015/10 SI
R. notabilis s.l. LH014/03 SI
R. notabilis s.l. LH013/03 SI
R. notabilis s.l. Du-33266/01 SI
R. notabilis s.l. LH011/14 SI
R. notabilis s.l. LH011/01 SI
R. notabilis s.l. LH010/07 SI
R. notabilis s.l. Du-34889/01 HR
R. notabilis s.l. Du-35351/15 HR
R. notabilis s.l. LH028/02 AT
R. notabilis s.l. Hoe5615 AT
R. notabilis s.l. EH10137/03 AT
R. notabilis s.l. Du-30441/01 SI
R. notabilis s.l. LH012/08 SI

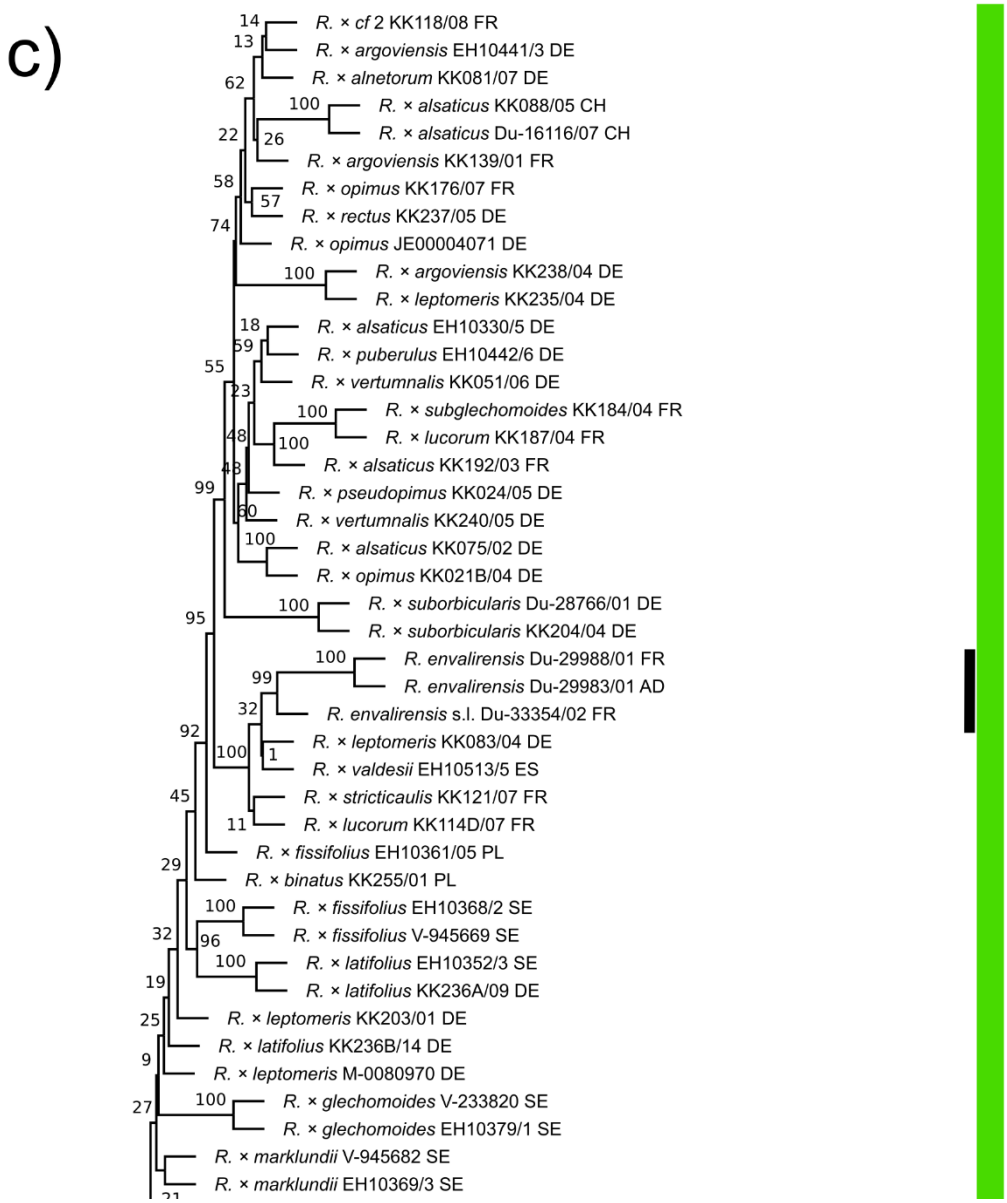
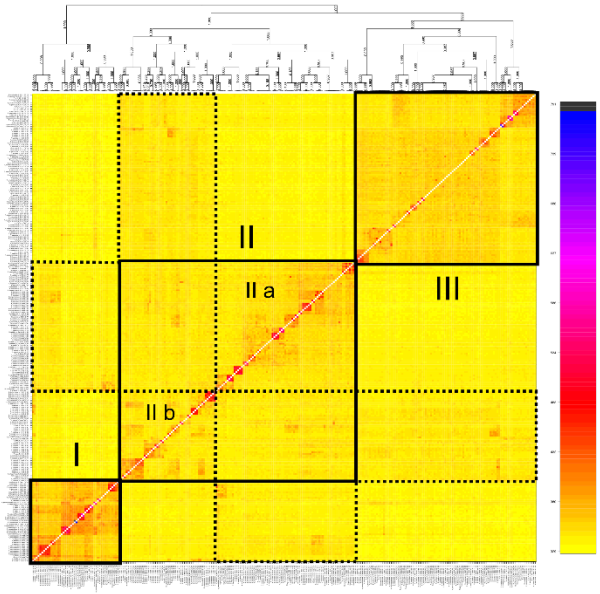
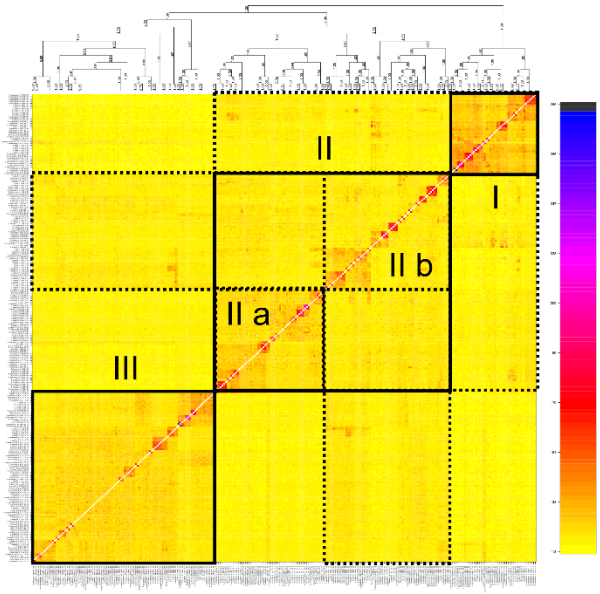


Figure S4. Coalescent-based tree ASTRAL tree. (a–c) Results are based on a TE alignment of 113 samples and 576 non-phased genes. Felsenstein Bootstrap Proportions (FBP) values are displayed on branches. Black bars indicate sexual species. Colored bars highlight main clades of Fig. 4b. Names of apomictic polyploid taxa are tentative. Squared brackets: A part of the branch was cut for illustrative purposes.

a) min10



b) min30



c) min50

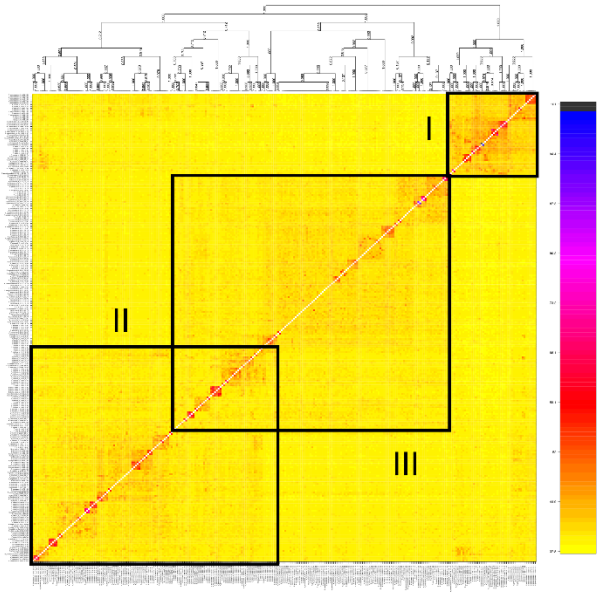


Figure S5. Clustered fineRADstructure coancestry matrices. Results are based on 280 sexual and polyploid apomictic individuals of the *R. auricomus* complex and (a) the "min10" (97,312 loci, 438,775 SNPs), (b) the "min30" (33,165 loci, 194,083 SNPs), or (c) the "min50" (11,196 loci, 64,554 SNPs) RADseq alignments. The legend (on the right) shows the color-coding of pairwise genetic similarity: The darker the color, the higher the similarity between a pair of individuals. We highlighted supported genetic clusters with solid lines, shared similarity among (sub)clusters with dotted lines, and sexual species with broad dashed lines (subgenome C, F, M, N, and E). Above the coancestry matrix, the clustering structure is given (posterior probability group assignment probabilities). See Fig. 5 and Figshare data repository for different datasets and more detailed results.

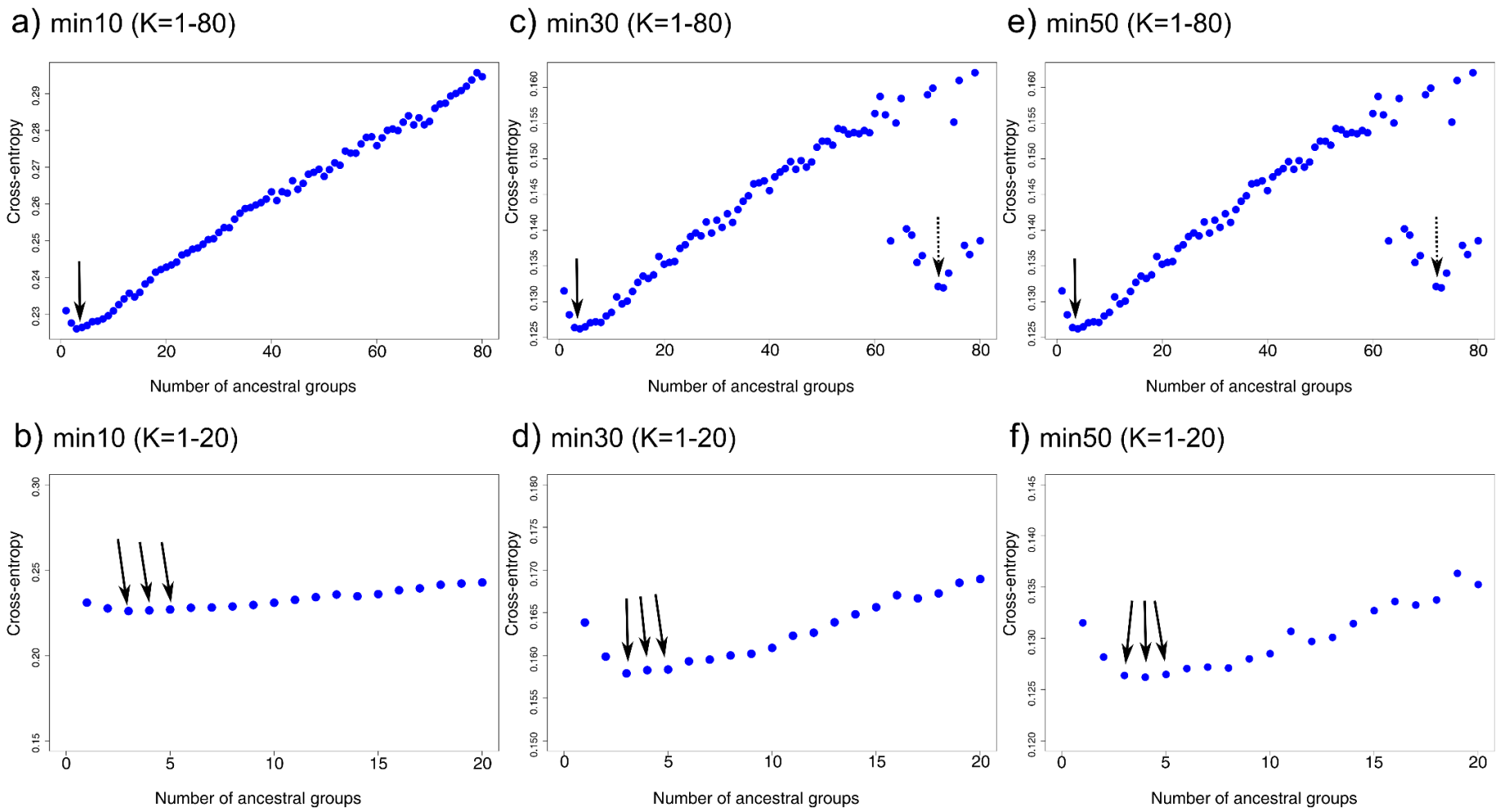


Figure S6. Cross-entropies based on sNMF results. Results are based on different unlinked-SNP-RADseq datasets of K = 80 assumed ancestral genetic groups incl. smaller sections (K = 20), (a, b) "min10" (97,312 loci), (c, d) "min30" (33,165 loci), and (e, f) "min50" (11,196 loci). The cross-entropy criterion is based on the prediction of masked genotypes to evaluate the fit of a model with K populations (the lower, the better; Frichot et al., 2014). Solid arrows indicate global cross-entropy minima (K = 3–5 across datasets), and dashed arrows show local minima.

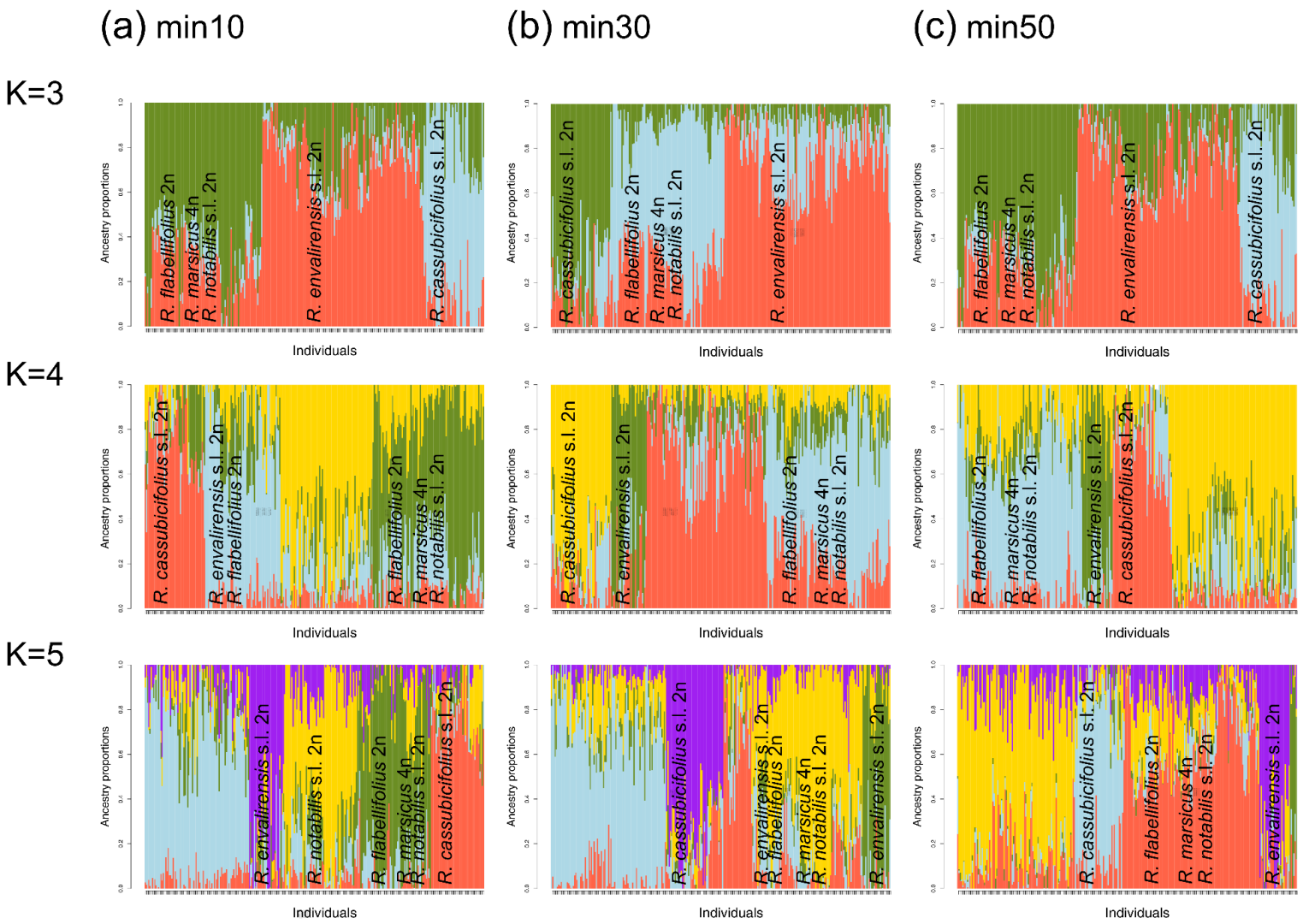


Figure S7. Ancestry proportions based on sNMF analyses. Results are based on 280 sexual and apomictic *R. auricomus* individuals, and different unlinked-SNP-RADseq datasets and numbers of ancestral genetic groups (K = 3–5), i.e. (a) "min10" (97,312 loci), (b) "min30" (33,165 loci), and (c) "min50" (11,196 loci). See Figshare data repository for detailed results (including sample names).

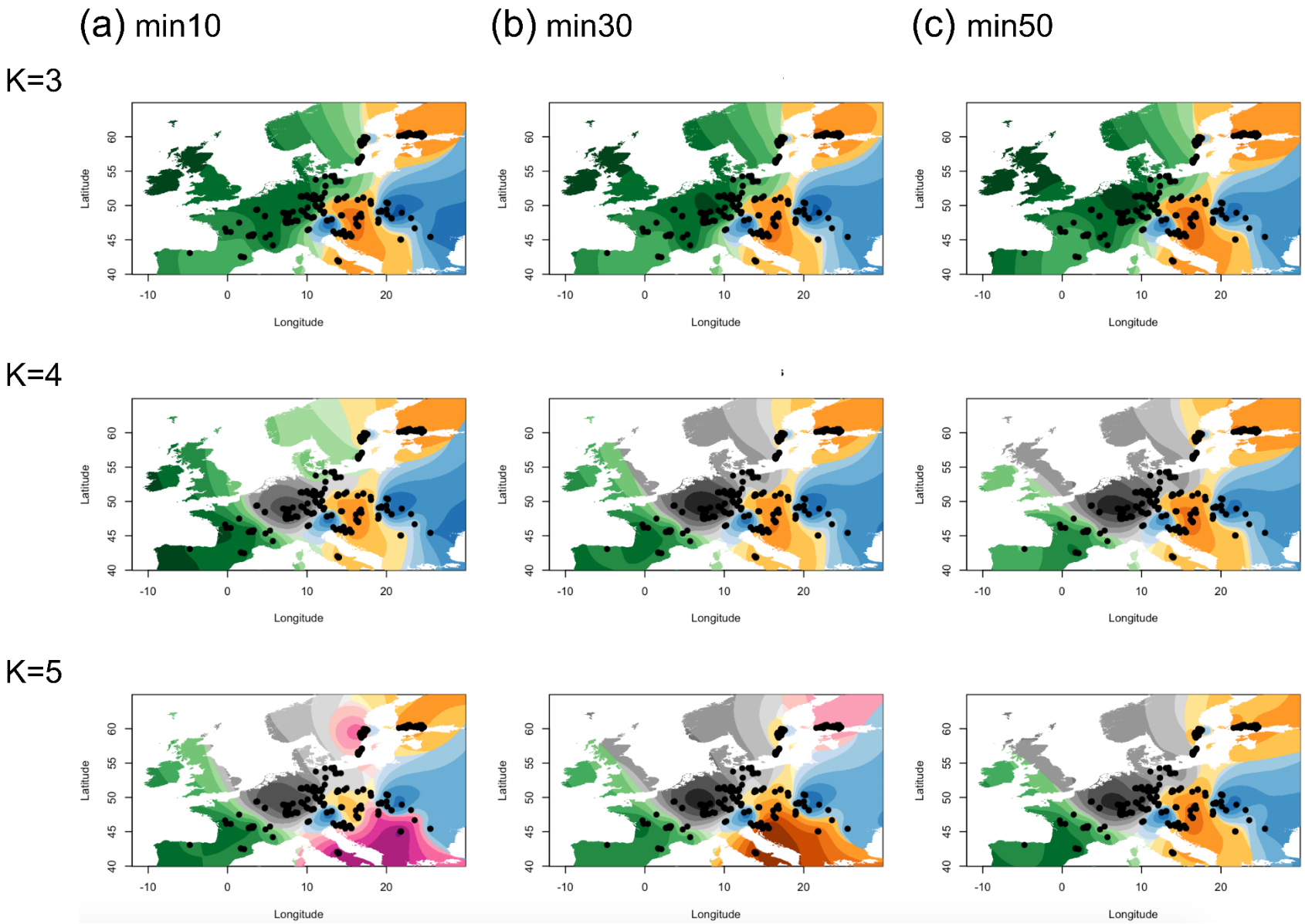
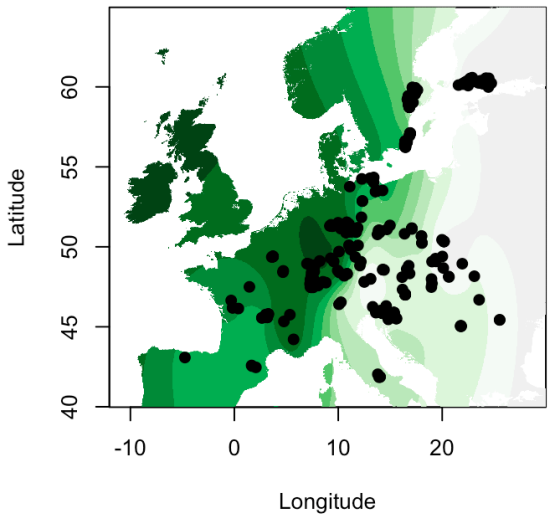
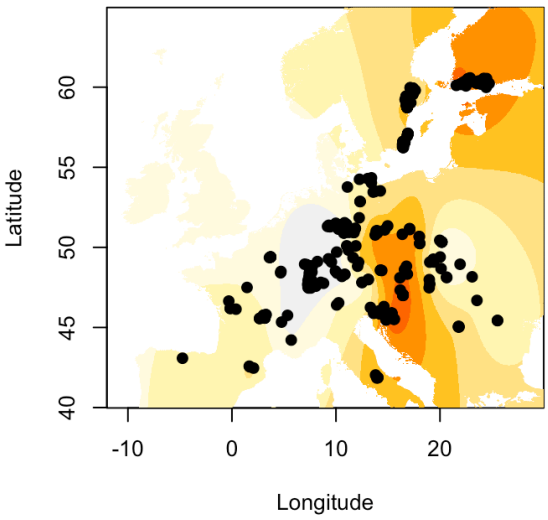

(a) min10
(b) min30
(c) min50
K=3
K=4
K=5
Latitude
Longitude
40
45
50
55
60
-10
0
10
20

Figure S8. Geographic maps showing ancestry coefficients across Europe (K = 3–5). Ancestry coefficients were illustrated with method "max", i.e., at each point the cluster for which the ancestry coefficient is maximal. Results are based on sNMF analyses of 280 sexual and apomictic *R. auricomus* individuals, and different unlinked-SNP-RADseq datasets and numbers of ancestral genetic groups (K = 3–5), (a) min10 (97,312 loci), (b) min30 (33,165 loci), and (c) min50 (11,196 loci). We colored maps according to Fig. 4 (except grey = Central European apomicts related to main clade E, pink = either (a) *R. flabellifolius*, *R. marsicus*, and Southern Scandinavian apomicts, (b) a separate apomictic cluster containing many Finnish taxa (related to the orange one (subgenome N, most reasonable according to other results), or (c) a separate apomictic cluster containing a few Finnish taxa). Results were similar among K = 3 dataset. K = 4 datasets slightly differed from each other (e.g., min10 unlikely because of disjunct grouping the subgenome "E" group). K = 5 datasets remarkably differed from each other (e.g., min10 unlikely because of disjunct distribution of *R. flabellifolius*, *R. marsicus*, and Southern Scandinavian apomicts or min50 unlikely because of a separate apomictic cluster containing a few Finnish taxa). Europe map source: https://maps.ngdc.noaa.gov.

a)


Latitude
40
45
50
55
60
-10
0
10
20
Longitude

b)


Latitude
40
45
50
55
60
-10
0
10
20
Longitude

c)

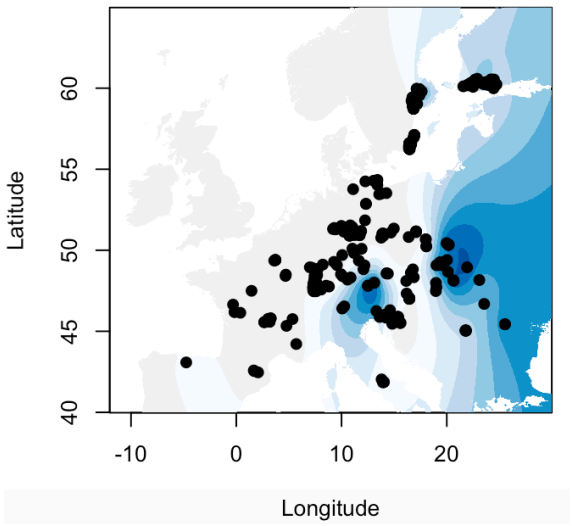
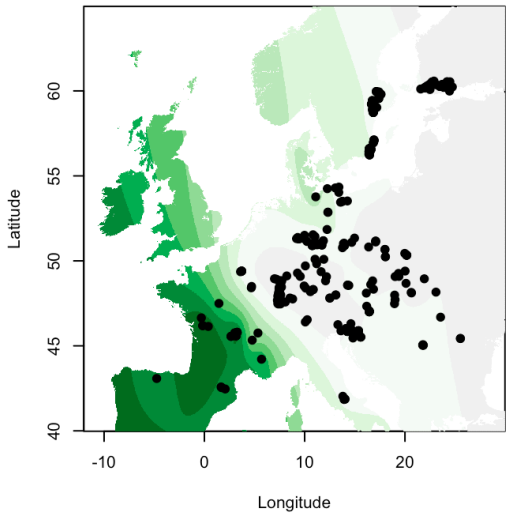
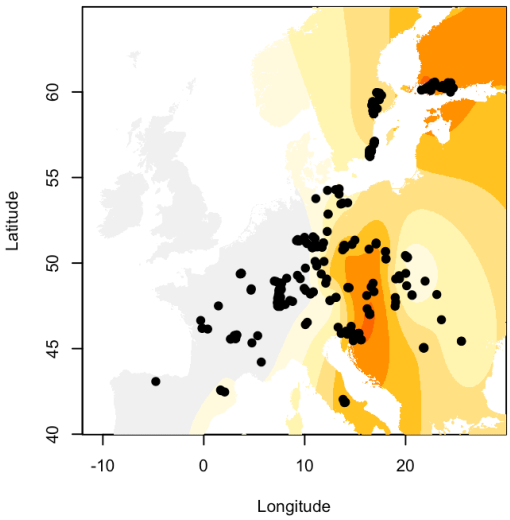


Figure S9. Geographic maps showing separate ancestry coefficients across Europe (K = 3). Ancestry coefficients are illustrated. Results are based on sNMF analyses of 280 sexual and apomictic *R. auricomus* individuals and the "min30" unlinked-SNP-RADseq dataset (33,165 loci) for K = 3 genetic clusters (a–c; see Fig. 7, S12, S13 and Results for details). Europe map source: https://maps.ngdc.noaa.gov.

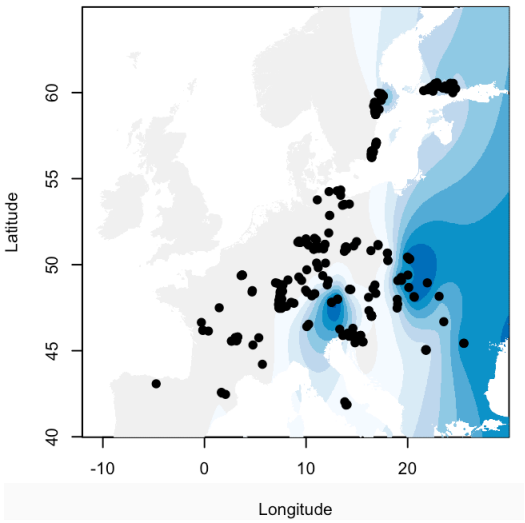
a)


Latitude
60
55
50
45
40
-10
0
10
20
Longitude

b)


Latitude
60
55
50
45
40
-10
0
10
20
Longitude

c)



d)

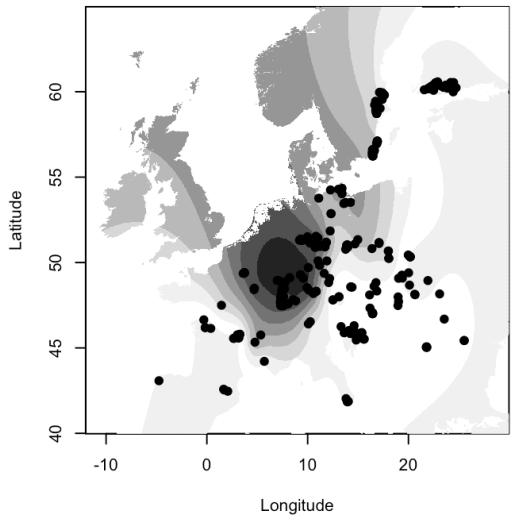
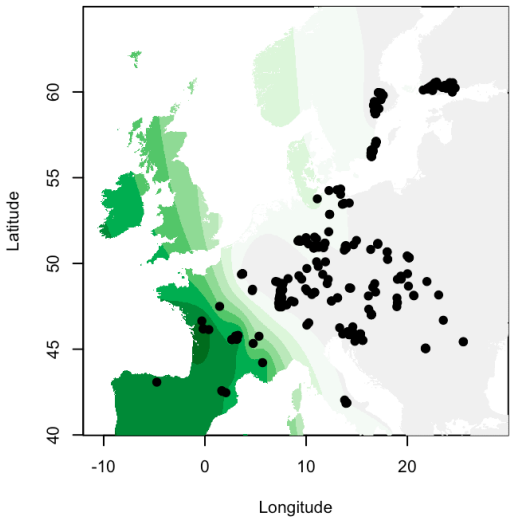
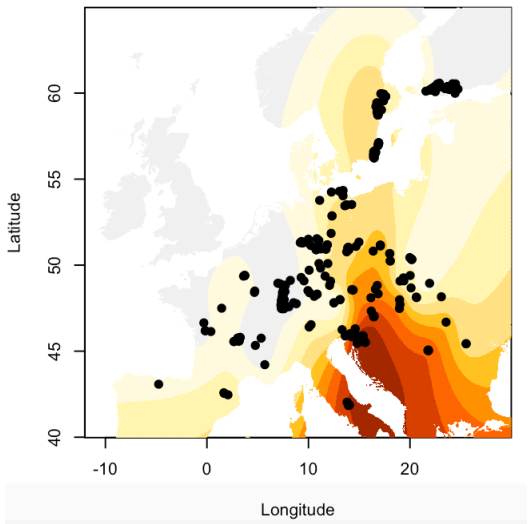


Figure S10. Geographic maps showing separate ancestry coefficients across Europe (K = 4). Ancestry coefficients are illustrated. Results are based on sNMF analyses of 280 sexual and apomictic *R. auricomus* individuals and the "min30" unlinked-SNP-RADseq dataset (33,165 loci) for K = 4 genetic clusters (a–d; see Figs. 7, S12, S13 and Results for details). Europe map source: https://maps.ngdc.noaa.gov.

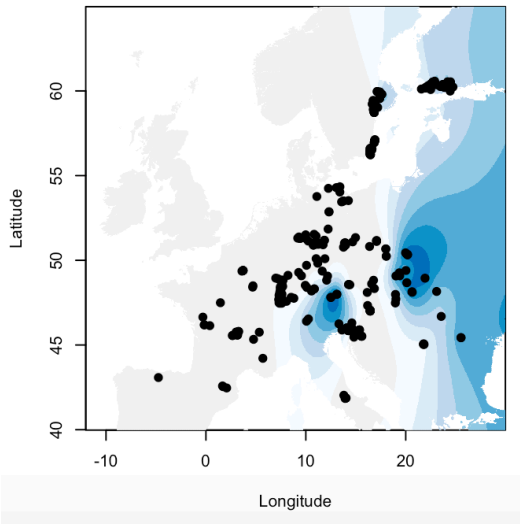
a)


Latitude
60
55
50
45
40
-10
0
10
20
Longitude

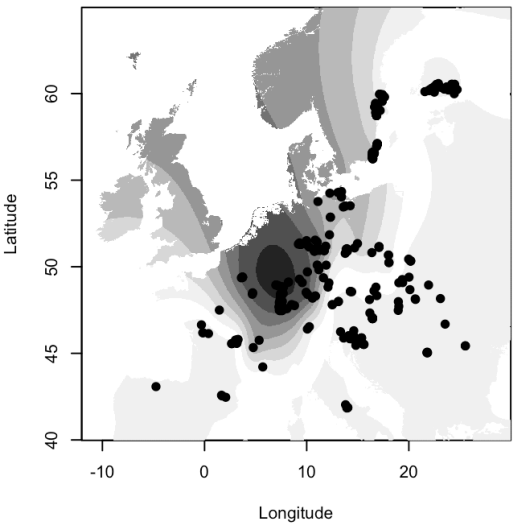
b)


Latitude
60
55
50
45
40
-10
0
10
20
Longitude

c)


Latitude
60
55
50
45
40
-10
0
10
20
Longitude

d)


Latitude
60
55
50
45
40
-10
0
10
20
Longitude

e)

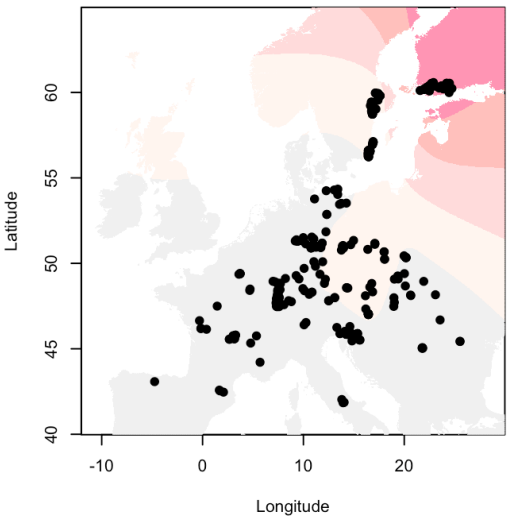
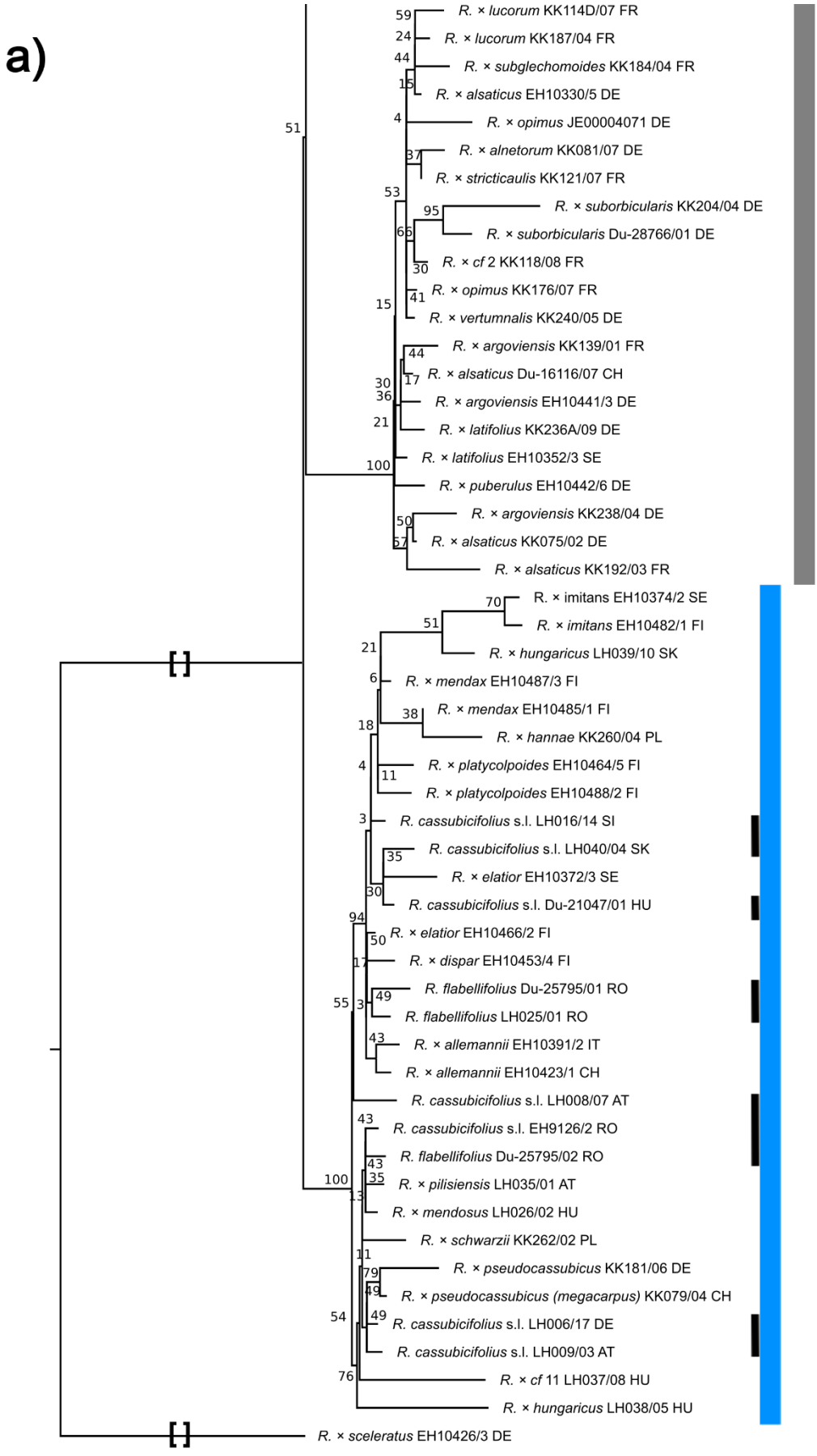


Figure S11. Geographic maps showing separate ancestry coefficients across Europe (K = 5). Ancestry coefficients are illustrated. Results are based on sNMF analyses of 280 sexual and apomictic *R. auricomus* individuals and the "min30" unlinked-SNP-RADseq dataset (33,165 loci) for K = 5 genetic clusters (a–e; see Figs. 7, S12, S13 and Results for details). Europe map source: https://maps.ngdc.noaa.gov.


a)
R. × lucorum KK114D/07 FR
R. × lucorum KK187/04 FR
R. × subglechomoides KK184/04 FR
R. × alsaticus EH10330/5 DE
R. × opimus JE00004071 DE
R. × alnetorum KK081/07 DE
R. × stricticaulis KK121/07 FR
R. × suborbicularis KK204/04 DE
R. × suborbicularis Du-28766/01 DE
R. × cf 2 KK118/08 FR
R. × opimus KK176/07 FR
R. × vertumnalis KK240/05 DE
R. × argoviensis KK139/01 FR
R. × alsaticus Du-16116/07 CH
R. × argoviensis EH10441/3 DE
R. × latifolius KK236A/09 DE
R. × latifolius EH10352/3 SE
R. × puberulus EH10442/6 DE
R. × argoviensis KK238/04 DE
R. × alsaticus KK075/02 DE
R. × alsaticus KK192/03 FR
R. × imitans EH10374/2 SE
R. × imitans EH10482/1 FI
R. × hungaricus LH039/10 SK
R. × mendax EH10487/3 FI
R. × mendax EH10485/1 FI
R. × hannae KK260/04 PL
R. × platycolpoides EH10464/5 FI
R. × platycolpoides EH10488/2 FI
R. cassubicifolius s.l. LH016/14 SI
R. cassubicifolius s.l. LH040/04 SK
R. × elatior EH10372/3 SE
R. cassubicifolius s.l. Du-21047/01 HU
R. × elatior EH10466/2 FI
R. × dispar EH10453/4 FI
R. flabellifolius Du-25795/01 RO
R. flabellifolius LH025/01 RO
R. × allemannii EH10391/2 IT
R. × allemannii EH10423/1 CH
R. cassubicifolius s.l. LH008/07 AT
R. cassubicifolius s.l. EH9126/2 RO
R. flabellifolius Du-25795/02 RO
R. × pilisiensis LH035/01 AT
R. × mendosus LH026/02 HU
R. × schwarzii KK262/02 PL
R. × pseudocassubicus KK181/06 DE
R. × pseudocassubicus (megacarpus) KK079/04 CH
R. cassubicifolius s.l. LH006/17 DE
R. cassubicifolius s.l. LH009/03 AT
R. × cf 11 LH037/08 HU
R. × hungaricus LH038/05 HU
R. × sceleratus EH10426/3 DE
0.002

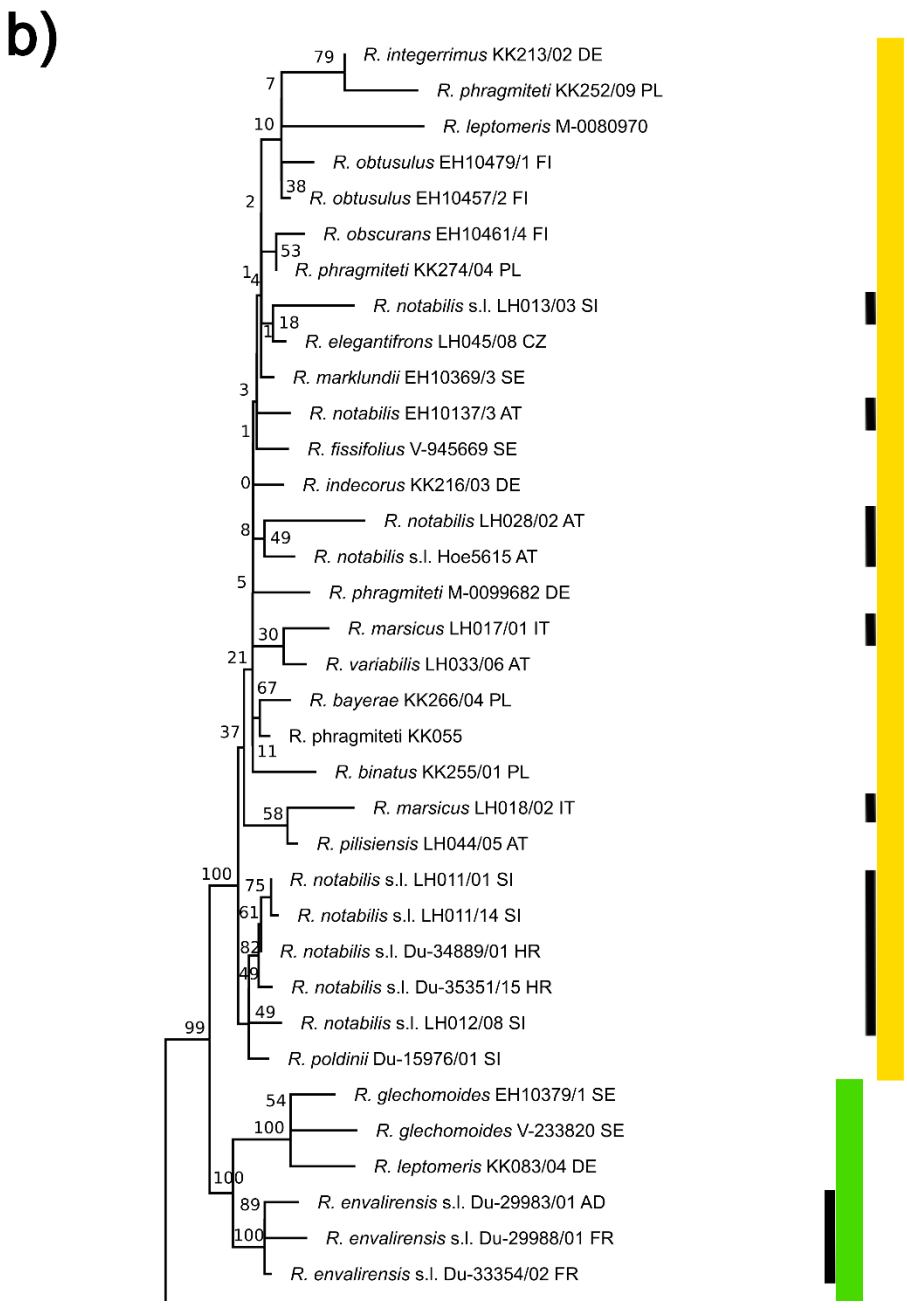


Figure S12. Concatenated ML tree inferred from plastid regions (CP). Results are based on a CP alignment of 87 samples and 71 plastid regions. Per branch, Felsenstein Bootstrap Proportions (FBP) values are displayed. Black bars indicate sexual species. Colored bars highlight main clades of Fig. 5a. Names of apomictic polyploid taxa are tentative. Squared brackets: A part of the branch was cut for illustrative purposes.

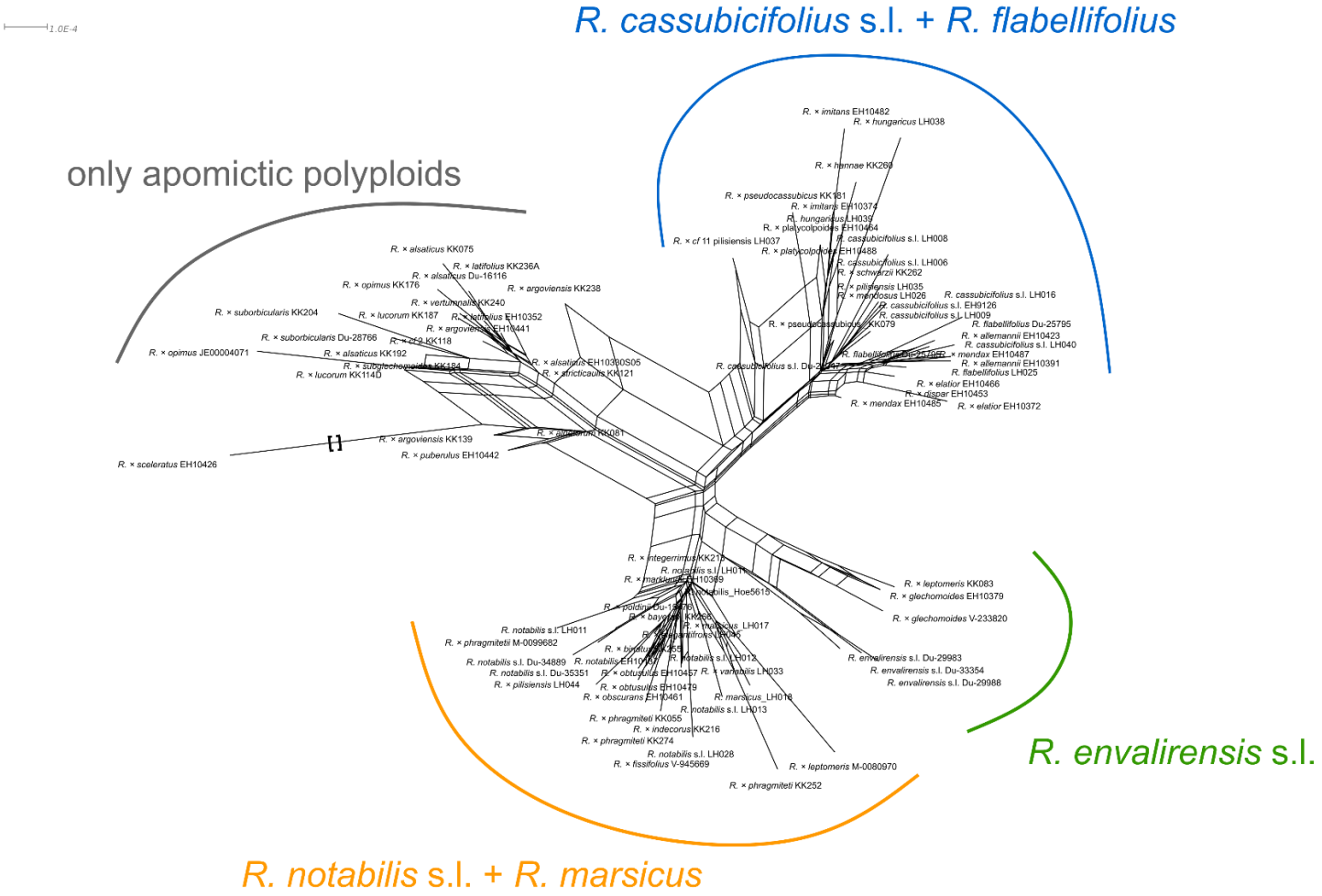


Figure S13. Neighbor-net analysis (SplitsTree) of CP data. Results are based on 87 samples and 71 plastid regions of the plastome dataset and genetic distances (general time reversible [GTR] model with estimated site frequencies and ML). We colored main splits according to Fig. 5a. No BT values are given for reasons of clarity and comprehensibility. Squared brackets: A part of the branch was cut for illustrative purposes.

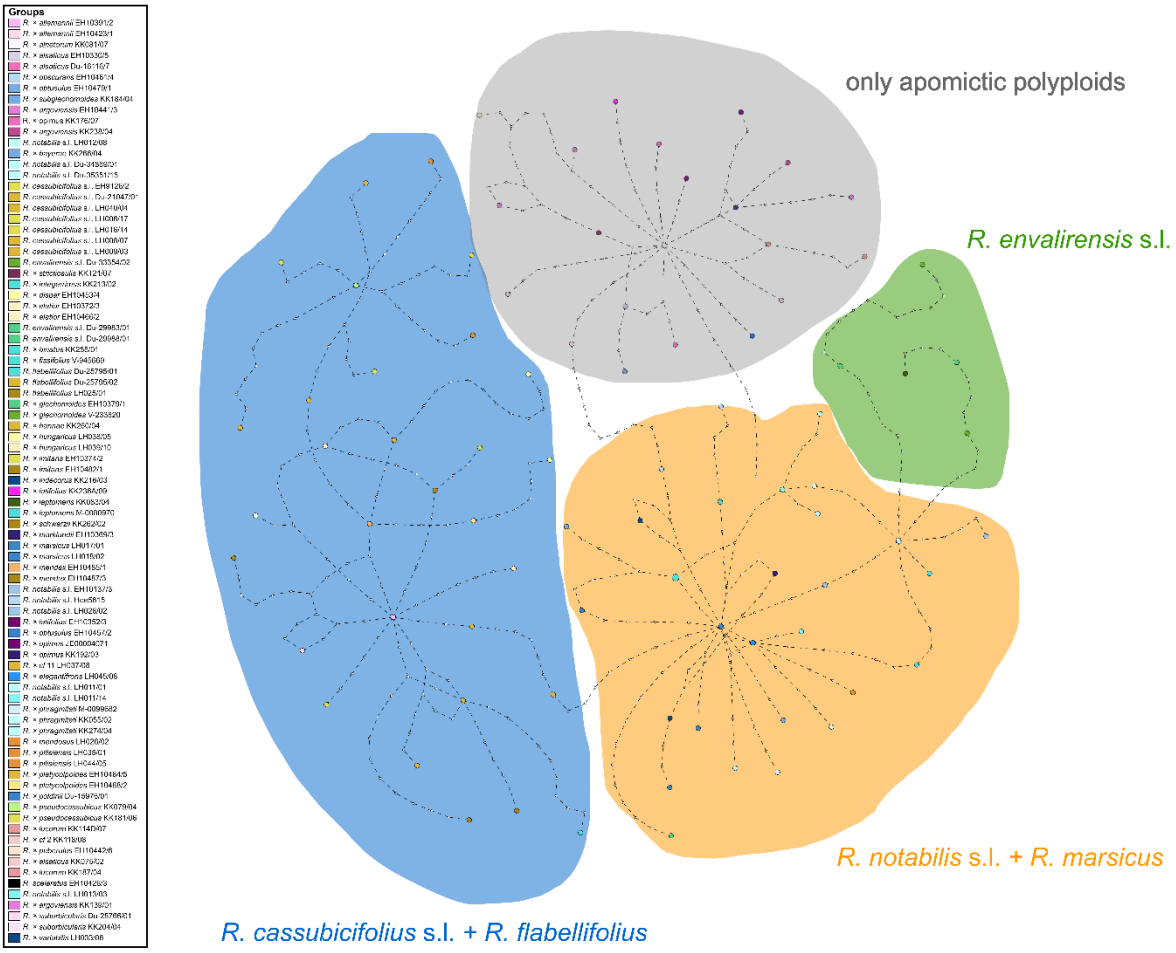


Figure S14. Haplotype network (TCS) of CP data. Results are based on 87 samples and 71 plastid regions. We colored main clusters according to Fig. 5a.

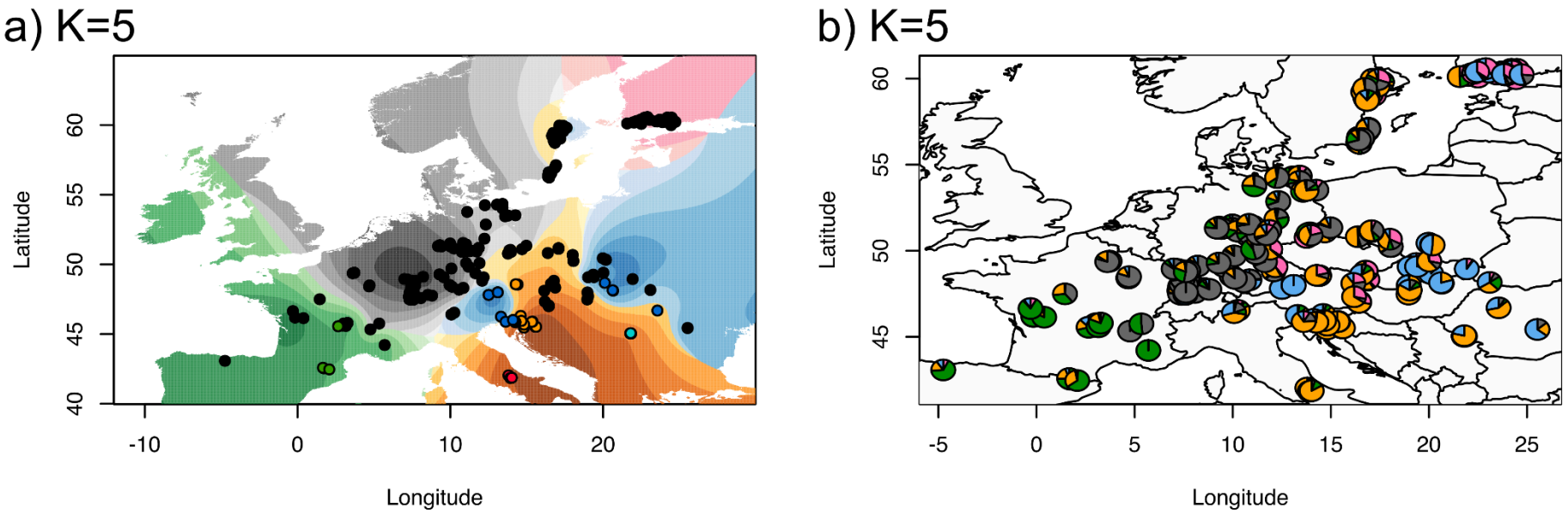


Figure S15. Geographic maps illustrating ancestry coefficients across Europe (K = 5). sNMF results are based on K = 5 genetic clusters, 280 sexual and apomictic *R. auricomus* individuals, and the "min30" unlinked-SNP-RADseq alignment (33,165 loci). (a) Interpolated values of ancestry coefficients were illustrated with method "max", i.e., at each point the cluster for which the ancestry coefficient is maximal and (b) location-wise admixture estimate pie charts were shown. See Figs. S11–16 and Figshare for more detailed sNMF results. In (a), colored circles represent sexual species (coloring according to Fig. 4): blue = *R. cassubicifolius* s.l. (C), turquoise = *R. flabellifolius* (F), red = *R. marsicus* (M), green = *R. envalirensis* s.l. (E), and orange = *R. notabilis* s.l (N). Europe map source: https://maps.ngdc.noaa.gov.

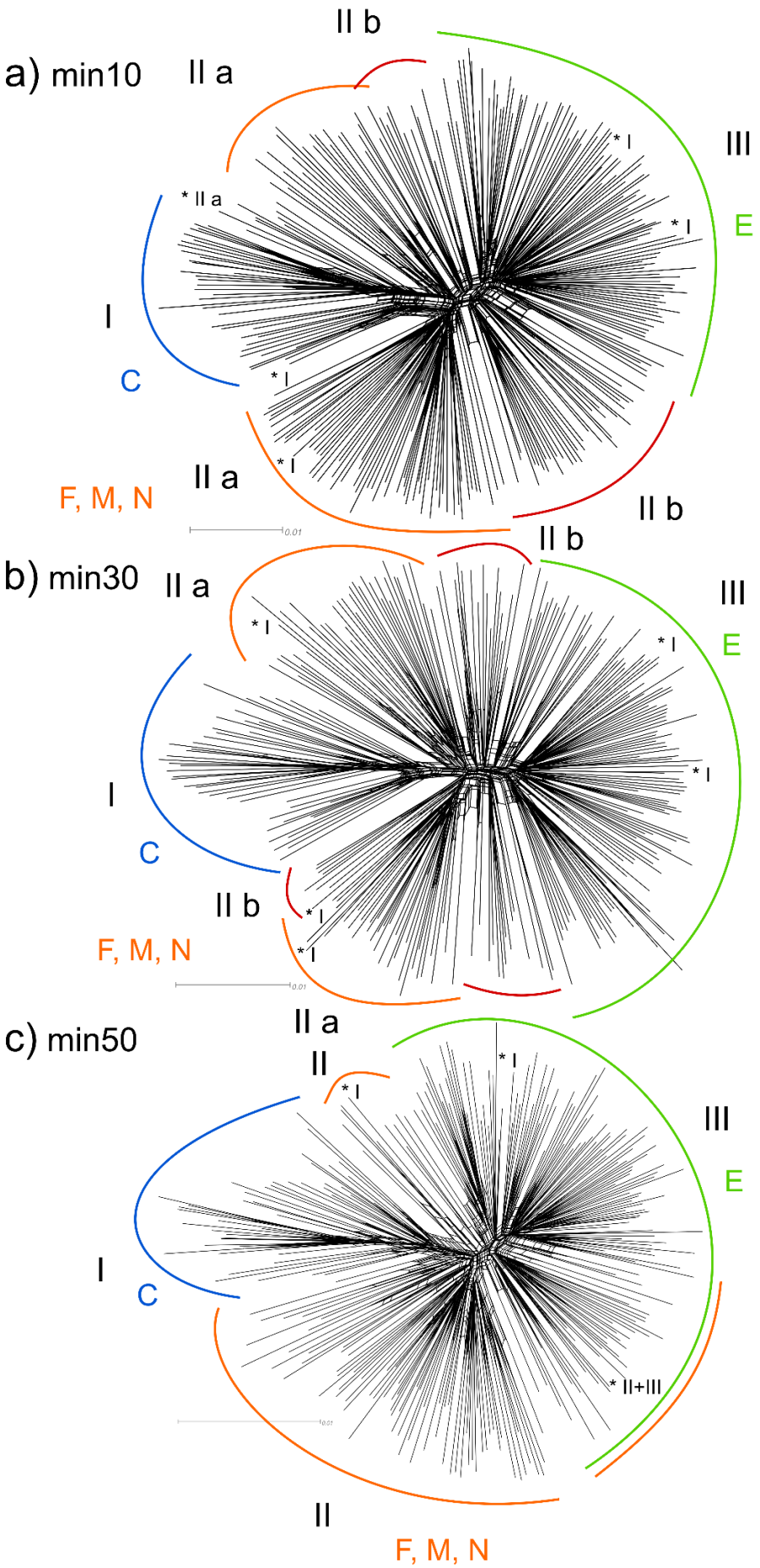

a) min10
II b
II a
III
E
* I
* II a
* I
I
C
* I
* I
II a
F, M, N
II b
0.01
b) min30
II b
II a
III
E
* I
* I
* I
I
C
II b
* I
* I
F, M, N
0.01
c) min50
II a
II
* I
* I
III
E
I
C
* II+III
II
F, M, N

Figure S16. Neighbor-net analysis (SplitsTree) of RADseq data. We included 280 sexual and apomictic *R. auricomus* individuals. Results are based on genetic distances (GTR model with estimated site frequencies and ML) and the unlinked-SNP-RADseq datasets, a) min10 (97,312 loci), b) min30 (33,165 loci), and c) min50 (11,196 loci). Major clusters were highlighted with colors according to Fig. 4 and designated with numbers, according to clusters found in the respective RADpainter analyses (min10, min30, and min50). Superscript asterisks (*) show exemplary exceptions to RADpainter clusters. "min10": Cluster I contains *R. flabellifolius* samples that are situated within RADpainter cluster IIa; cluster III contains *R.* × *pseudocassubicus* and *R.* × *megacarpus* samples that are situated within RADpainter cluster I; cluster IIa contains *R.* × *platycolpoides* that is situated within RADpainter cluster I. "min30": Cluster III contains *R.* × *pseudocassubicus* and *R.* × *megacarpus* samples that are situated within RADpainter cluster I; cluster IIa contains *R.* × *platycolpoides*, *R.* × *czywczynensis*, and *R.* × *dispar* that are situated within RADpainter cluster I. "min50": Cluster III contains *R.* × *pseudocassubicus* and *R.* × *megacarpus* samples that are situated within RADpainter cluster I, a subcluster that contains a mixture of RADpainter clusters II + III; cluster IIa contains *R.* × *platycolpoides* that is situated within RADpainter cluster I. Colored letters indicate sexual species and also dominant subgenomes of each group, i.e., C = (subgenome of) *R. cassubicifolius* s.l., E = (subgenome of) *R. envalirensis* s.l., and F, M, N = (subgenomes of) *R. flabellifolius*, *R. marsicus*, and *R. notabilis* s.l. See Fig. S10 (min30) in Karbstein et al. (2021) and data on Figshare for files in better resolution. Scale bar = no. of changes.

Table S7. Genetic Structure and Phylogenetic Network results of tested tetraploid *R. auricomus* accessions (H_1-H_{10}). Each row (H_1-H_{10}) represents a separately analyzed individual. Results are based on RADseq (RADpainter+fineRADstructure, PhyloNetworks) and TE (STACEY, PhyloNetworks, Phylonet) datasets. P_1 is always the parent with the largest/likeliest genomic contribution (either indicated by coancestry/posterior probability values of structure analyses or by inheritance probabilities of network analysis) followed by the other parental contributions. Genetic structure analyses (RADpainter+fineRADstructure, STACEY): Parents in brackets indicate genomic contributions below specific thresholds (i.e., only some sexual accessions of a species show significant subgenome contributions, and not the mean of all accessions, see Material and methods for details). For phylogenetic networks, percentages of subgenome contribution are given in brackets. Criteria for building consensus results of parental subgenome contribution(s) are: (i) take the most abundant parent within column; (ii) if there are two equal abundant parents (e.g., two-times 'C' and 'F') within a column, both parental subgenome contributions were taken for the consensus result ('C/F'); (iii) if there are two parental subgenome contributions for one parent, we included them with a value of '0.5' (instead of '1.0') in consensus calculations; (iv) parental subgenome contributions in brackets were ignored for consensus calculations. The row 'final result' indicate the final subgenome contribution(s), i.e., consensus results corrected by the full likelihood approach followed by AIC calculations in Phylonet (likel+AIC, AIC = Akaike Information Criterion; different results if AIC network difference was less than 10 units) and plastome analysis results (CP type; C/F= plastid type shared by the diploid sexual species *R. cassubicifolius* and *R. flabellifolius*, * = not the same sample between CP analyses and phylogenetic network analyses, U# = haplotype from an unknown/extinct sexual progenitor species of Central Europe). Parental subgenomes not inferred by both analyses were removed from final results. Concerning the final results of H_9, we classified P_1 as "E(U)" because of the *R. envalirensis*-like U plastid type (see above). According to final results, we classified the genome evolution of investigated and the number of involved subgenomes in polyploid formation. See also Figs. 5a,b, 8a-h, and data on Figshare for sample IDs, genetic structure, and phylogenetic network results. Letters indicate sexual progenitor subgenomes, i.e., C = *R. cassubicifolius* s.l., F = *R. flabellifolius*, N = *R. notabilis* s.l., E = *R. envalirensis* s.l., and U = probably extinct sexual progenitor.

Analysis	H_1 ("*R.* × *platycolpoides*")			H_2 ("*R.* × *elatior*")				H_3 ("*R.* × *pseudocassubicus*")				H_4 ("*R.* × *hungaricus*")				H_5 ("*R.* × *fissifolius*")			H_6 ("*R.* × *glechomoides*")			H_7 ("*R.* × *pilisiensis*")				H_8 ("*R.* × *indecorus*")				H_9 ("*R.* × *subglechomoides*")				H_{10} ("*R.* × *leptomeris*")		
	P_1	P_2	P_3	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4	P_1	P_2	P_3	P_4	P_1	P_2	P_3
R+F (RADseq)	C	N	F	C	F	(N)	(E)	C	(N)			C	F	N		E	N	(F)	N	F	(E)	F	N	C	(E)	N	F	(C)		E				E	(N)	(F)
STACEY (TE)	C	F	N	C	F	(N)	(E)		E	F	N	F	C	N	(E)	E	F	N	E	F	N	F	C	N	E	N	C	F	E	E	C	F	N	E	F	N
Phylo-Networks (RADseq)	C (63)	N (37)		C (97)	E (3)			C (51)	E (49)			C (88)	F (12)			N (54)	E (46)		C (71)	F (29)		N (67)	F (33)			N (92(	F (8)			E (82)	C (18)			N (60)	E	E (40)
Phylo-Networks (TE)	C (99)	E (1)		C (68)	F (32)			C (63(	F (37)			C (51)	F (49)			F (61)	E (39)		F (93)	E (7)		F (91)	E (9) N			E (70)	C (30)			F (53)	E (47)			E (90)		F (10)
PhyloNet (TE)	E (64)	C (36)		C (78)	C (22)			C (60)	C (40)			C (96)	C (4)			E (88)	F (12)		E (72)	F (28)		E (96)	N F (4)			E (80)	C (20)			F (53)	E (47)			F (67)		E (33)
consensus	C	N	F	C	F			C	E	F	N	C	F	N		E	E F	N	E	F	N	F	F	C N	E	N	C	F	E	E	C F	F	N	E	E F	N
liked+AIC (PhyloNet)	reticulate (N, C)			reticulate (C, F)				reticulate C=E)				reticulate (F, C > C, C)				reticulate (3x E, F)			reticulate (E, F)			tree-like (F)				reticulate (E, C > N, C)				reticulate (E, F) > tree like (E)				reticulate (E, F)		
CP type	C	F		C	F			C	F			C	F			N*			E			C*	F*			N				U#				E		
final results	C	N		C	F			C	E			C	F			E	F	N	E	F		F				N	C	E		E	(F)			E	F	
genome evolution	allopolyploid			allopolyploid				allopolyploid				allopolyploid				allopolyploid			allopolyploid			autopolyploid				allopolyploid				allo- vs. autopolyploid				allopolyploid		
no. sub-genome/s	2			2				2				2				3			2			1				3				1-2				2		

Table S8. SNP discovery based on RADseq-SNiPLOID results. SNPs of different categories are shown in percents, and numbers of SNP positions are given (in total, without the category "others"). SNP percents of Cat1–5 were calculated based on all SNP positions without the category "others". In brackets, we gave SNPs percent concerning all SNPs (see Material and Methods for evaluation of SNP category 'others'). The category gives the percent of non-defined SNPs (i.e., heterozygous SNP calls for DIPLOID2) in relation to all SNP positions. Cat 1 = SNPs identical to DIPLOID2, cat 2 = SNPs identical to DIPLOID1/reference, cat ¾ = derived SNPs, cat 5 = homeo-SNPs (hybrid heterozygous for homeologous alleles of both parental genomes; see also Materials and methods, and Wagner et al., 2020). Concerning H_6 and H_8, we calculated two SNiPloid analyses because three parents have contributed to its origin. Coloring is according to Fig. 3. Subgenomes of C = *R. cassubicifolius* s.l., F = *R. flabellifolius*, E = *R. envalirensis* s.l., and N = *R. notabilis* s.l.

Analysis	**H_1** ("*R. × platycolpoides*")		**H_2** ("*R. × elatior*")		**H_3** ("*R. × pseudocassubicus*")		**H_4** ("*R. × hungaricus*")		**H_5** ("*R. × fissifolius*")				**H_6** ("*R. × glechomoides*")		**H_8** ("*R. × indecorus*")				**H_{10}** ("*R. × leptomeris*")	
	D1(REF) = C	D2 = N	D1(REF) = C	D2 = F	D1(REF) = C	D2 = E	D1(REF) = C	D2 = F	D1(REF) = E	D2 = F	D1(REF) = C	D2 = N	D1(REF) = E	D2 = F	D1(REF) = N	D2 = C	D1(REF) = N	D2 = E	D1(REF) = E	D2 = F
Cat 1 & 2 (interspecific SNPs)	73.65 (71.50)		84.51 (82.57)		64.35 (62.48)		65.34 (62.99)		91.46 (89.07) (88.43)		92.53		91.6 (89.27)		86.94 (84.22) (82.39)		85.08		87.47 (84.09)	
Cat 3/4 (derived SNPs)	3.08 (2.14)		3.86 (2.58)		2.81 (1.77)		2.63 (1.82)		2.78 (1.99)		4.67 (1.72)		2.72 (1.88)		4.84 (2.14)		4.41 (2.64)		3.08 (2.16)	
Cat 5 (homeo-SNPs)	23.26 (26.36)		11.64 (14.84)		32.84 (35.76)		32.04 (35.18)		5.76 (8.94)		2.80 (9.84)		5.69 (8.85)		8.22 (13.64) (14.68)		10.03		9.45 (13.75)	
others to all SNP positions (not considered in calculations)	31.94% (2.31)		34.73 (2.62)		39.14 (2.72)		32.00 (2.05)		29.92 (1.97) (3.75)		64.47		32.11 (1.96)		57.50 (4.02)		41.67 (2.60)		31.49 (2.18)	
SNP positions (without others)	5133 (7368)		1633 (2438)		8650 (13826)		6208 (8947)		3700 (5178) (11431)		4214		2831 (4090)		2665 (6019)		2493 (4163)		5872 (8384)	
SNP positions (all)	7542 (7542)		2502 (2502)		14212 (14212)		9130 (9130)		5280 (5280) (11860)		11860		4170 (4170)		6271 (6271)		4274 (4274)		8571 (8571)	

References

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Borowiec, M. L. (2016). AMAS: A fast tool for alignment manipulation and computing of summary statistics. *PeerJ*, *4*, e1660. https://doi.org/10.7717/peerj.1660

Dann, M., Bellot, S., Schepella, S., Schaefer, H., & Tellier, A. (2017). Mutation rates in seeds and seed-banking influence substitution rates across the angiosperm phylogeny. *BioRxiv*. https://doi.org/10.1101/156398

Frichot, E., Mathieu, F., Trouillon, T., Bouchard, G., & François, O. (2014). Fast and efficient estimation of individual ancestry coefficients. *Genetics*, *196*, 973–983. https://doi.org/10.1534/genetics.113.160572

Karbstein, K., Tomasello, S., Hodač, L., Dunkel, F. G., Daubert, M., & Hörandl, E. (2020). Phylogenomics supported by geometric morphometrics reveals delimitation of sexual species within the polyploid apomictic *Ranunculus auricomus* complex (Ranunculaceae). *Taxon*, *69*, 1191–1220. https://doi.org/10.1002/tax.12365

Karbstein, K., Tomasello, S., Hodač, L., Lorberg, E., Daubert, M., & Hörandl, E. (2021). Moving beyond assumptions: Polyploidy and environmental effects explain a geographical parthenogenesis scenario in European plants. *Molecular Ecology*, *00*, 1–17. https://doi.org/10.1111/mec.15919

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, *30*, 772–780. https://doi.org/10.1093/molbev/mst010

Kent, W. J. (2002). BLAT-The BLAST-like alignment tool. *Genome Research*, *12*, 656–664. https://doi.org/10.1101/gr.229202

Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics*, *26*, 589–595. https://doi.org/10.1093/bioinformatics/btp698

Paradis, E., Blomberg, S., Bolker, B., Brown, J., Claude, J., Cuong, H. S., … de Vienne, D. (2020). *APE: Analyses of phylogenetics and evolution*. Retrieved from http://ape-package.ird.fr/

Paradis, Emmanuel, & Schliep, K. (2019). ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, *35*, 526–528. https://doi.org/10.1093/bioinformatics/bty633

Pease, J. B., Brown, J. W., Walker, J. F., Hinchliff, C. E., & Smith, S. A. (2018). Quartet Sampling distinguishes lack of support from conflicting support in the green plant tree of life. *American Journal of Botany*, *105*, 385–403. https://doi.org/10.1002/ajb2.1016

Wagner, N. D., He, L., & Hörandl, E. (2020). Phylogenomic relationships and evolution of polyploid *Salix* species revealed by RAD sequencing data. *Frontiers in Plant Science*, *11*, 36–41. https://doi.org/10.3389/fpls.2020.01077

Wang, L.-G., Lam, T. T.-Y., Xu, S., Dai, Z., Zhou, L., Feng, T., … Yu, G. (2020). TREEIO: An R package for phylogenetic tree input and output with richly annotated and associated data. *Molecular Biology and Evolution*, *37*, 599–603. https://doi.org/10.1093/molbev/msz240

Weitemier, K., Straub, S. C. K., Cronn, R. C., Fishbein, M., Schmickl, R., McDonnell, A., & Liston, A. (2014). Hybseq: Combining target enrichment and genome skimming for plant phylogenomics. *Applications in Plant Sciences*, *2*, 1400042. https://doi.org/10.3732/apps.1400042

Winston, C., Lionel, H., Thomas Lin, P., Kohske, T., Claus, W., Kara, W., & Hiroaki, Y. (2020). *GGPLOT2: Create elegant data visualisations using the grammar of graphics*. Retrieved from https://cran.r-project.org

Xu, H., Luo, X., Qian, J., Pang, X., Song, J., Qian, G., … Chen, S. (2012). FastUniq: A fast de novo duplicates removal tool for paired short reads. *PLoS ONE*, *7*, e52249. https://doi.org/10.1371/journal.pone.0052249

Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T. (2017). *GGTREE*: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, *8*, 28–36. https://doi.org/10.1111/2041-210X.12628

Acknowledgments

First of all, I like to thank Elvira Hörandl for raising my interest in the polyploid apomictic *R. auricomus* complex and the most complicated, puzzling but evolutionarily tremendously important plant topics polyploidy and apomixis. This PhD position gave me so much for my further road of life, from learning new scientific techniques, writing internationally accepted scientific papers, improving communicative skills, but also enabling networking with scientists across Germany and other countries. It is not an exaggeration to say that this position opened up my horizons and that this work would not have been possible at the same scientific level in many other working groups. Therefore, I like to further thank Elvira Hörandl and all members of the Systematic Botany working group in Göttingen for valuable (and sometimes hard-fought species-concept) discussions, for giving me the opportunity to develop and improve my teaching abilities, for trusting me to lead teaching courses, for great (alpine) excursions and working group parties (and volleyball events!), for letting me participate on international conferences and Germany-wide SPP workshops, and for essential future scientific career tips.

I particularly thank Salvatore Tomasello and Ladislav Hodač for the relentless support of my PhD and the *R. auricomus* project, for valuable discussions about phylogenetics and geometric morphometrics, tips on how I should proceed with scientific and organizational issues, but also to talk about anything and everything (I hope that I haven't spammed you too much about facts and sights of Thuringia!), and of course about the best beer. Without your support, I would never have been able to complete my PhD so successfully. Thanks also to Natascha Wagner for the overwhelming support in every sense, Marc Appelhans for herbarium voucher ordering and great tips for student excursions around Göttingen, and John Paul Bradican for support in English scientific manuscript writing.

The support of Silvia Friedrichs should also be mentioned here. She tirelessly took care of the watering, weed removal, plant disease control, and care of our goldilock buttercups in the Botanical Garden and climate chambers. This enabled the success of our project - thank you for that and for hours and hours talking and discussing horticulture.

I also want to thank Jennifer Krüger for scanning fresh leaf material, cutting images, and preparing these images for geometric morphometrics for weeks and months, which took a lot of work off my feet. I enjoyed the chats about hobbies, from marathon-topics to everything that turned on dogs.

I also want to mention and thank Gabi Ließmann (and the volunteers) for taking care of mounting thousands of *R. auricomus* individuals (not without a joke about these "wonderful" plants ☺). Sorry for this torture!

Moreover, I like to thank Claudia Pätzold for the bioinformatical support. Thanks also to Birthe Barke, who taught me a lot about flow cytometry and cytology.

Many students and voluntary helpers were involved in the project. They made valuable contributions, but it would take too long to list them all. I would particularly like to mention and thank the bachelor and master students: Mareike Daubert, Elisabeth Rahmsdorf, Ellen Lorberg, and Lea Hansen. I also acknowledge Franz Dunkel, since decades a passionate *R. auricomus* researcher. He provided tons of herbarium specimens, valuable comments, and DNA plant material to the project. I further want to express my gratitude to Sven Bradler for great discussions about phylogenetics (and his nice stick insects) and the support of my PhD.

Considering mental support, I have to mention our "Mensa group" - thanks to Natascha Wagner, Marc Appelhans, Salvatore Tomasello, Eleni Syngelaki, and the numerous scientific guest visitors for great, sometimes hot discussions, and fun during lunch, and a good coffee afterward outside in front of the Cafétaria. The corona pandemic forced us to break with this tradition. Nevertheless, I really enjoyed the lunch outside in the botanical garden with you and the new PhD students (John Paul Bradican, Lee Ping Ang, and Loïc Pittet) later on. When talking about mental and any support, I have also to thank Kathleen Prinz, one of my former bachelor and master thesis supervisors. She nurtured my scientific path from the beginning, drew my attention to this PhD position, and motivated me during this time. Thanks for everything!

And finally, my greatest gratitude goes to my family and friends. Isabell Schellin, thanks for the overwhelming support in every sense for 10 years, for the funny, beautiful moments, and for overcoming difficult situations, and for enduring the many moves that a scientific career entails, demanding fieldwork (e.g., empty car batteries, thunderstorms, and driving several hundred kilometers for a single plant population ☺), talk preparations, and my never-ending wanting urge to talk about plants during excursions and trips; my mother Manuela and my father Ronald, for fostering any of my interests from the cradle, starting with an own herbal garden (yes, I was really proud of it!), hiking, and finally the desire to study biology, and the innumerable excursions and fieldwork you joined. Without your contributions, this work would have not been possible. Thanks for having supported me through the good and occasional rough times, and for always being there for me.

Statutory Declaration

I, hereby, declare that this dissertation has not been presented to any other examining body either in its present or a similar form. Furthermore, I also affirm that I have not applied for a doctoral program at any other higher school of education. I declare that this dissertation was undertaken independently.

Göttingen, 31st August 2021

Kevin Karbstein