GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

No sex, No problem?

Mutation accumulation in asexual animals

Dissertation

for the award of the degree

"Doctor rerum naturalium" (Dr. rer. nat.)

of the University of Göttingen

within the Doctoral program Biology

of the Georg August University School of Science (GAUSS)

submitted by

M.Sc.

Alexander Brandt

from Hannover

Göttingen, 30 March 2020

Thesis Committee

Prof. Dr. Stefan Scheu, Animal Ecology, JF Blumenbach Institute

Prof. Dr. Mark Maraun, Animal Ecology, JF Blumenbach Institute

Dr. Marina Schaefer, Animal Ecology, JF Blumenbach Institute

Members of the Examination Board

Reviewer: **Prof. Dr. Stefan Scheu**, Animal Ecology, JF Blumenbach Institute

Second Reviewer: **Prof. Dr. Christoph Bleidorn**, Animal Evolution and Biodiversity, JF Blumenbach Institute

Further Members of the Examination Board

Prof. Dr. Mark Maraun, Animal Ecology, JF Blumenbach Institute

Prof. Dr. Elvira Hörandl, Department of Systematics, Biodiversity and Evolution of Plants, Albrecht von Haller Institute

PD Dr. Sven Bradler, Animal Evolution and Biodiversity, JF Blumenbach Institute

Dr. Jens Bast, Evolution in the Absence of Sex, Institute of Zoology, University of Cologne

Date of the oral examination: 05 June 2020

If there is one thing the history of evolution has taught us it's that life will not be contained. Life breaks free, it expands to new territories and crashes through barriers, painfully, maybe even dangerously, but, uh…well, there it is.

You're implying that a group composed entirely of female animals will… breed?

No I'm, I am simply saying that live uh... finds a way.

Dr. Ian Malcolm and Dr. Henry Wu discussing "ancient asexual scandals", *Jurassic Park*, 1993

Curriculum Vitae

Personal Details

Name	Alexander Brandt
[redacted]	[redacted]
[redacted]	[redacted]

Education

2015-2020	PhD (Biology) *University of Göttingen* Title “No sex, No problem? Mutation accumulation in asexual animals”
[redacted]	[redacted]
[redacted]	[redacted]
[redacted]	[redacted]

Grants/awards

[redacted]	[redacted]
[redacted]	[redacted]

Conference Contributions

2014 – present	GOEvol Network and Symposium on Evolutionary Biology *University of Göttingen* Organisation and hosting of annual conferences on Evolutionary Biology
2020	EVENET 2020 - 10th Annual Symposium on Eco-evolutionary Dynamics *KU Leuven, Royal Belgian Zoological Society* Talk + Poster
2019	Symposium “Frontiers in Molecular Zoology” *University of Göttingen, JF Blumenbach Institute* Talk
2019	SPIRIT Summer school “Molecular Genetics in Zoology” *University of Göttingen, JF Blumenbach Institute* Poster
2018	Invited Seminar *University of Göttingen, Dept. Animal Evolution and Biodiversity* Talk
2018	CNRS Conference Jaques Monod “Sex uncovered” *Roscoff, France* Poster
2018	Sinergia Project retreat “Genome Evolution under Sexual and Asexual Reproduction” *Leysin, Switzerland* Talk
2017	16th Congress of the European Society for Evolutionary Biology (ESEB) *Groningen, The Netherlands* Poster

2016	Graduate meeting on Evolutionary Biology of the German Zoological Society (DZG) *Greifswald, Germany* Talk
2015	15th Congress of the European Society for Evolutionary Biology (ESEB) *Lausanne, Switzerland* Poster
2015	Invited Seminar *University of Lausanne, Dept. of Ecology and Evolution (group Schwander)* Talk
2014	ISeq Symposium *Leipzig, Germany* Poster
2014	Annual congress of the German Zoological Society (DZG) *Göttingen, Germany* Poster

Teaching

2014 – 2020	B.Sc. Seminar “Animal Ecology and Evolution”
2015	Supervision of a B.Sc. thesis (Julien Glanz)
2012	Practical course “Animal anatomy” (tutor)
2010	Practical course “Animal anatomy” (tutor)

Further educational work

[redacted]

Further Activities

[redacted]

[redacted]

Publication list

Peer reviewed articles

Brandt, A., Schaefer, I., Glanz, J., Schwander, T., Maraun, M., Scheu, S. & Bast, J. Effective purifying selection in ancient asexual oribatid mites. *Nat. Comm.* **8**, 873 (2017).

Brandt, A., Bast, J., Scheu, S., Meusemann, K., Donath, A., Schuette, K., Machida, R. & Kraaijeveld, K. No signal of deleterious mutation accumulation in conserved gene sequences of extant asexual hexapods. *Sci. Rep.* **9**, 5338 (2019).

Bast, J., Parker, D. J., Dumas, Z., Jalvingh, K. M., Tran Van, P., Jaron, K. S., Figuet, E., **Brandt, A.**, Galtier, N. & Schwander, T. Consequences of asexuality in natural populations: insights from stick insects. *Mol. Biol. Evol.* **35**, 1668–1677 (2018).

Article/book chapter emerging from this thesis

Brandt, A., Anselmetti, Y., Bluhm, C., Dumas, Z., François, C. M., Galtier, N., Heimburger, B., Jaron, K. S., Labédan, M., Maraun, M., Robinson-Rechavi, M., Schaefer, I., Scheu, S., Simion, P., Tran-Van, P., Schwander, T. & Bast, J. Haplotype divergence supports obligate asexuality in the oribatid mite *Oppiella nova*. *in prep.* (2020).

Hoerandl, E., Bast, J., **Brandt, A.**, Scheu, S., Bleidorn, C., Cordellier, M., Nowrousian, M., Begerow, D., Sturm, A., Verhoeven, K., Boenigk, J., Friedl, T. & Dunthorn, M. Genome evolution of asexual organisms and the paradox of sex in eukaryotes. *in prep.* (2020).

Articles in preparation

Schuppenhauer, M., **Brandt, A.**, Schaefer, I. & Lehmitz, R. Genetic diversity of oribatid mites along running water. *in prep.* (2020).

Heimburger, B., Soto-Maurer, S., Schardt, L., **Brandt, A.** & Hartke, T. The evolution of mount-building in Australian harvester termites. *in prep.* (2020).

Contents

Summary

According to classical predictions, sex is beneficial because it allows for purging of slightly deleterious mutations via effective purifying selection. In asexual organisms mutational load increases over time, eventually leading to their genomic decay. Further, in the absence of sex, haplotypes are expected to accumulate mutations independently and diverge in parallel. This is commonly known as ‘the Meselson effect’ and regarded as strong support for truly asexual evolution. Several single gene studies tested accumulation of slightly deleterious mutations and the Meselson effect. However, genome-based studies of independent evolutionary replicates are essential to account for among-gene variance and lineage-specific effects. Such studies are missing for some important, i.e. well suited and long-term, asexual animals.

In this thesis, I tested if both predictions, mutational load and haplotype divergence, are commonly met in asexual animals. For this, I generated and investigated genome data of different, asexual and closely related sexual arthropod species. While the effectiveness of purifying selection was decreased in asexual *Timema* stick insects, there was no overall difference between asexual and sexual representatives of eight major hexapod groups. Surprisingly, the mutational load in oribatid mites was even lower compared to their sexual relatives. Possibly, effective purging of slightly deleterious mutations is enabled by their very large population sizes. In the asexual oribatid mite species *Oppiella nova*, population genomic patterns including parallel haplotype divergence were consistent with the Meselson effect. This strongly supports its evolution in the absence of sex.

In the discussion, I present a comparative overview of existing studies on mutation accumulation and haplotype divergence in animals. I show that accumulation of slightly deleterious mutations is not commonly found in asexual animals indicating that sex *per se* is not a prerequisite for their

long term persistence. Some “scandalous” asexual animals may well have survived in the absence of sex for millions of years.

Zusammenfassung

Der selektive Vorteil von Sex liegt, gebräuchlichen Annahmen nach, darin, dass Sex es ermöglicht, durch effektive stabilisierende Selektion geringfügig schädliche Mutationen aus dem Genpool zu entfernen. In asexuellen Organismen hingegen, so die Theorie, nimmt die Belastung durch Mutationen nach und nach zu, sodass ihr Genom schließlich “verfällt”. Weiterhin geht man davon aus, dass in Abwesenheit von Sex die Haplotypen (unterschiedliche Varianten desselben Gens) Mutationen unabhängig voneinander ansammeln und sich dadurch parallel verändern. Dieses Phänomen wird als “Meselson Effekt” bezeichnet und gilt als deutlicher Beleg für die Abwesenheit von Sex in der Evolutionsgeschichte eines Organismus. Bisherige Studien zur Ansammlung schädlicher Mutationen und zum Meselson Effekt basieren größtenteils auf der Analyse einzelner Gensequenzen. Ihre Ergebnisse sind jedoch schwer interpretierbar, da sie sich zwischen den Genen eines Genoms stark unterscheiden können. Genomweite Studien unabhängiger evolutionärer Replikate (Organismen, die Sex unabhängig voneinander verloren haben) sind daher besonders wertvoll. Trotz einiger (weniger) Studien fehlen noch Untersuchungen bei einigen wichtigen asexuellen Taxa, insbesondere solchen, die bereits seit langen (evolutionären) Zeiträumen asexuell sind.

In dieser Arbeit habe ich untersucht, ob die beiden genannten Annahmen für asexuelle Tiere weit verbreitete Gültigkeit haben. Dazu habe ich Genomdaten verschiedener asexueller Arthropodenarten und ihrer jeweils nahen sexuellen Verwandten analysiert. Ich konnte zeigen, dass die Effektivität von stabilisierender Selektion bei asexuellen *Timema* Gespenstschrecken reduziert war, während zwischen asexuellen und sexuellen Vertretern acht weiterer Hexapoden-Gruppen kein allgemeiner Unterschied feststellbar war. In “uralt

asexuellen” Hornmilben fand sich sogar eine geringere Belastung mit schädlichen Mutationen als bei ihren sexuellen Verwandten. Das effektive Entfernen geringfügig schädlicher Mutationen könnte durch die sehr großen Populationen der Hornmilben begünstigt worden sein. Ich habe außerdem gezeigt, dass die asexuelle Hornmilbenart *Oppiella nova* verschiedene populationsgenomische Muster aufweist, die auf den Meselson Effekt hindeuten; zum Beispiel, die sich, wie oben beschrieben, parallel verändernden Haplotypen. Dies legt die Abwesenheit von Sex während der Evolutionsgeschichte von *O. nova* nahe.

Unter Berücksichtigung der Gesamt-Studienlage zeige ich, dass die Ansammlung geringfügig schädlicher Mutationen keine weit verbreitete Eigenschaft asexueller Tiere ist. Dies bedeutet, dass Sex *per se* keine zwingende Notwendigkeit für den Fortbestand von Tierarten über lange evolutionäre Zeiträume darstellt. Einige “skandalöse” Arten haben anscheinend über Jahrmillionen ohne Sex überlebt.

Contributions to the chapters of this thesis

Chapter 1

Effective purifying selection in ancient asexual oribatid mites

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

Chapter 2

Consequences of asexuality in natural populations: insights from stick insects

Contribution to data analysis (CUB) and discussion

Chapter 3

No signal of deleterious mutation accumulation in conserved gene sequences of extant asexual hexapods

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

Chapter 4

Haplotype divergence supports obligate asexuality in the oribatid mite *Oppiella nova*

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

I General Introduction

The absence of sex is expected to be accompanied by specific consequences for genome evolution. Some of them are disadvantageous and assumed to drive asexual lineages to extinction in the long term. In this thesis, I tested whether accumulation of slightly deleterious mutations and increased haplotype divergence are common in animals. For this, I compared genome and transcriptome data of sexual and asexual arthropod species of different groups and ages of asexuality.

Diverse reproduction

Nature provides a diverse but sometimes, in public perception, unrecognised variety of ways to reproduce (Bell, 1982; Aanen *et al.*, 2016). In fact, there are so many different and sometimes complicated and curious strategies to produce new individuals that it is even difficult to classify them (Bengtsson, 2009; Aanen *et al.*, 2016). Commonly, a distinction is drawn between processes that change the (genetic) state of individuals through segregation and recombination, called mixis, and those that do not (Ghiselin, 1974; Bell, 1982). However, the purpose of classifying the many ways of reproduction is context-dependent which, in evolutionary biology, is to emphasise the adaptive significance of sexuality (Williams, 1975). In this thesis, I compare two kinds of lineages (continuous temporal series of ancestors and descendants): lineages whose reproduction involves mixis between males and females (biparental sex) and all-female lineages where mixis is restricted to the meiotic production of gametes or entirely absent. Therefore, I will use sex as synonymous with biparental sex from here on.

This definition leaves a plethora of ways to reproduce without sex; one being the development of embryos from unfertilised eggs in animals, called parthenogenesis (Von Siebold, 1856; Schoen *et al.*, 2009). A special case of parthenogenesis is the production of reduced (haploid) eggs which, if fertilised, develop into diploid females and, if not, into haploid males, so called arrhenotoky,

known most prominently from haplodiploid Hymenoptera (Heimpel & de Boer, 2008). Other cases, like hybridogenesis and androgenesis, seem to be even more complex: in hybridogenetic species, a male of a different species contributes his reduced genome to an unreduced diploid egg, but the male genome is excluded again during gamete formation in the offspring (Lamatsch & Stoeck, 2009). Under androgenesis, the maternal genome is excluded from the egg after fertilisation, and females of some species even provide eggs devoid of any nuclear maternal DNA (Schwander & Oldroyd, 2016).

In this thesis, the term parthenogenesis is used synonymously with thelytoky which describes the production of daughters without any contribution from males and which can occur in facultative or cyclical fashion (Jalvingh *et al.*, 2016). However, in the studies of this thesis, reproduction of all analysed lineages never involves mixis with males. Hence, I define asexuality as obligate thelytoky, from here on.

The paradox of sex

About 99% of all animal species reproduce via some kind of sex (Vrijenhoek, 1998; Otto, 2009). This is surprising given that sex is associated with many costs due to the need of searching and finding a mate, exposure to predators through costly courtship behaviour and signalling, and the risk of infection with sexually transmitted diseases (Lehtonen *et al.*, 2012). Most importantly, sex requires allocation of resources to the production of males which usually contribute very little to reproduction themselves (Maynard Smith, 1978; Meirmans *et al.*, 2012). This cost theoretically prevents sexual populations from resisting replacement by asexual offshoots whose reproduction is independent of males.

Hypotheses regarding the selective benefit of sex fall into two categories (Williams, 1975; Kondrashov, 1993). Long-term benefits of sex act at the lineage level and unfold only after many

successive generations. They derive, for example, from relieve of clonal interference, i.e. the ability to effectively combine beneficial alleles arising in different lineages (Fisher-Muller hypothesis; Fisher, 1930; Muller, 1932), from restoration of least-loaded genotypes that are continuously lost by drift (Muller's ratchet; Muller, 1964), from removal of linkage between beneficial and deleterious alleles (Hill & Robertson, 1966), or from enabling repair of deleterious alleles via meiotic gene conversion (Marais, 2003). Short-term benefits of sex unfold under strong fluctuating directional selection at the level of individuals and genes and allow sexual populations to withstand quick replacement by asexual offshoots. The rationale is that sex enables effective selection by generating variation for fitness through break-up of linked loci with opposite fitness effects (Fig. I; Felsenstein, 1974). Such negative linkage disequilibria can be generated via combined effects of drift and selection in finite populations (Hill-Robertson interference; Hill & Robertson, 1966). Thus, for sex to be beneficial in the short-term, the trajectory of selection must change constantly, e.g. through coevolving parasites (Red-Queen; Hamilton, 1980; Hamilton *et al.*, 1990). For details see e.g. Jalvingh *et al.* (2016) and Neiman *et al.* (2018).

Increased genetic variation can also be advantageous under fine-grained spatial heterogeneity (Bell, 1982) or when the availability of resources depends on the speed at which they are exploited and renewed (SRTS; Bell, 1982; Scheu & Drossel, 2007). Notably, hypotheses on the maintenance of sex are not mutually exclusive: several of the above-mentioned mechanisms may act simultaneously in the same lineage and different lineages may benefit from different mechanisms (West *et al.*, 1999). There likely is no "one and only" explanation for the maintenance of sex.

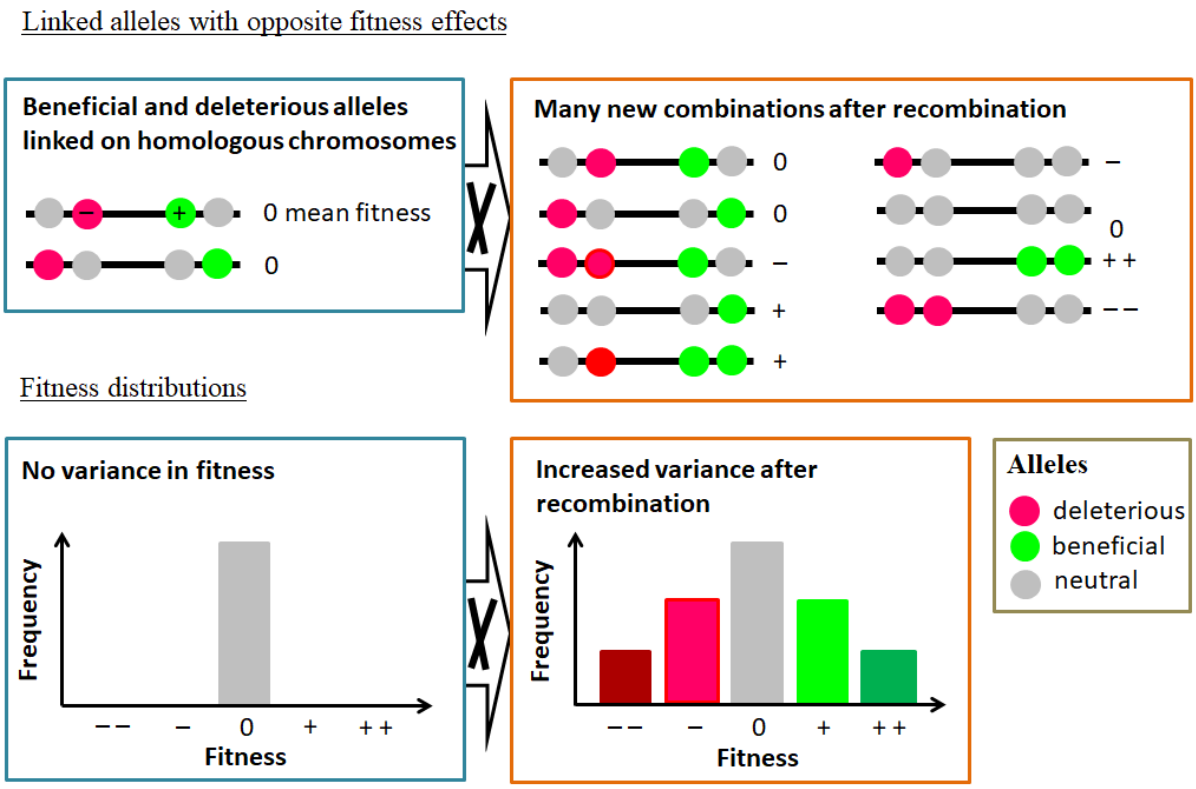


Figure I: The effect of sex on haplotypic variation. Under sexual reproduction recombination enables the break-up of linked loci with opposite fitness effects thereby generating a wide variety of combinations of deleterious and beneficial alleles with different mean fitnesses on the same chromosome. This increases the range of variation selection can act on. X: Recombination; +, -, 0: Fitness effects of alleles. Figure adapted from Jalvingh *et al.* (2016).

Selection-driven genomic consequences of asexuality

Hypotheses on the selective advantage of sex come with an inherent prediction: the absence of sex is accompanied by specific disadvantages offsetting its many costs. One elegant way of testing the various hypothesised benefits of sex is, thus, to investigate the predicted disadvantageous consequences of asexuality. Over evolutionary times, many of these manifest themselves as specific patterns in the genome. The inability of asexuals to break up linked loci with different fitness effects generated by drift (Hill-Robertson effect) and to restore least loaded genotypes in a population (Muller's ratchet), for example, is expected to reduce the effectiveness of purifying selection resulting in accumulation of fixed and segregating nonsynonymous

mutations in asexual species (Muller, 1964; Hill & Robertson, 1966; Felsenstein, 1974; Felsenstein & Yokoyama, 1976; Keightley & Otto, 2006). Further, nonsynonymous substitutions will likely have stronger deleterious effects in asexual as compared to sexual species (Pace *et al.*, 2011; Sharbrough *et al.*, 2018). Effectiveness of selection acting at synonymous sites, which is commonly assumed to be neutral, may also be diminished, thereby inducing a reduced bias in codon usage (Hershberg & Petrov, 2008). Finally, levels of (neutral) polymorphism within populations are expected to be reduced as neutral mutations and mutations with deleterious effects are strongly linked and purged together under asexuality (background selection; Hill & Robertson, 1966; Charlesworth *et al.*, 1993a).

Genomic consequences of asexuality not driven by selection

Evolution under asexuality is not exclusively driven by selection. Instead, patterns of non-selective, purely mechanistic processes can become manifested in the genome as well. An example is the reduction in GC-biased gene conversion (gBGC) through absence of pairing of homologous chromosomes, which is expected to translate into a decreased GC content at synonymous sites and in non-coding regions (Marais, 2003; Galtier *et al.*, 2018). Further, in the absence of sex, homologous chromosomes are expected to accumulate mutations independently of each other which leads to increased heterozygosity levels and a specific pattern of haplotype divergence: in asexual lineages, differentiation of alleles within individuals is expected to go back further in time than differentiation between populations, resulting in resemblance of haplotype subtrees over populations (Fig. II; Birky, 1996; Judson & Normark, 1996). This classically predicted non-selection-driven consequence of obligate asexuality is commonly called the ‘Meselson-effect’.

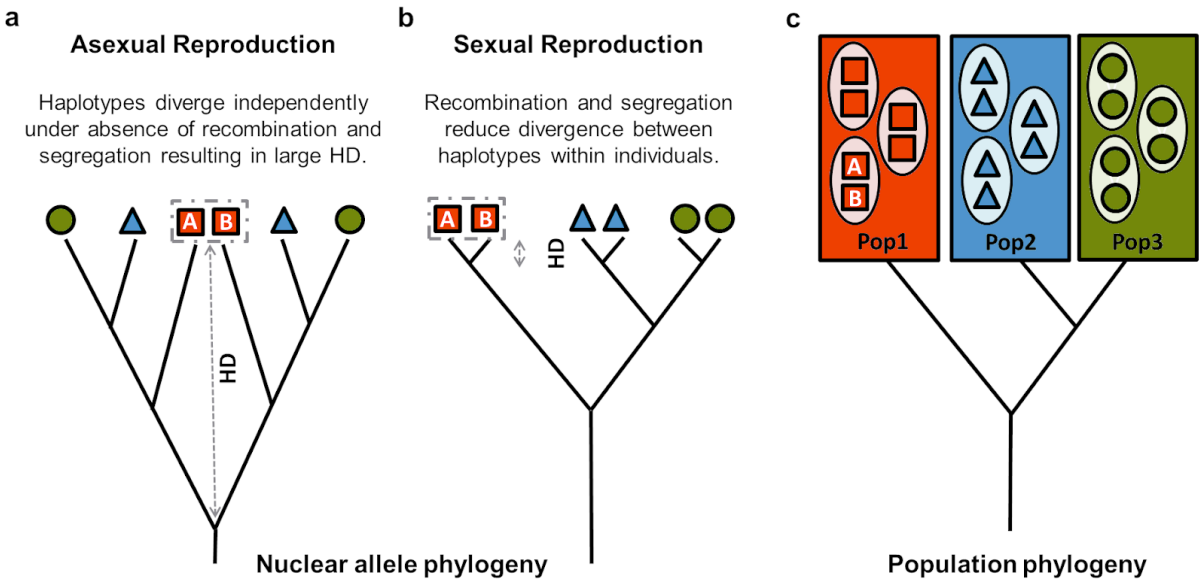


Figure II: Nuclear haplotype phylogenies expected under (long-term) obligate asexual and sexual reproduction. In diploid, effectively asexual organisms, homologous chromosomes accumulate mutations independently of each other and evolve as independent lineages (note that this can be restricted to specific regions sheltered from a loss of heterozygosity). Accordingly, divergence between haplotypes within individuals (HD) is expected to exceed divergence between haplotypes of individuals from different populations, as illustrated in the allele phylogeny of asexuals (a): Here, the splits that separate homologous haplotypes are expected to be older than the splits that separate populations. Also the haplotype subtrees resemble each other because haplotypes are expected to evolve in parallel. In sexual organisms instead, HD is expected to follow population divergence and the haplotype phylogeny of sexuals (b) resembles that of the population (c). Figure adapted from Schwander *et al.* (2011).

Empirical evidence for genomic consequences of asexuality

A plethora of studies regarding genomic consequences of asexuality in various asexual eukaryotes have been published, but only a minority of them are based on whole genome or transcriptome sequencing data (Hartfield, 2016; Glémin *et al.*, 2019). The results of single gene analyses are largely in agreement with the prediction that asexual lineages accumulate slightly deleterious mutations at fixed and segregating sites (Normark & Moran, 2000; Mark Welch & Meselson, 2001; Paland & Lynch, 2006; Barraclough *et al.*, 2007; Johnson & Howard, 2007;

Neiman *et al.*, 2010; Henry *et al.*, 2012). However, genome- and transcriptome-wide analyses have shown extensive among-gene variation, indicating that single gene analyses have to be interpreted with caution (Neiman *et al.*, 2018). Also, the results of the genome- and transcriptome-wide studies are less clear-cut (Hartfield, 2016; Neiman *et al.*, 2018; Glémin *et al.*, 2019; Jaron *et al.*, 2019); e.g. there was no difference in the effectiveness of purifying selection between four sexual and four asexual aphid species (Ollivier *et al.*, 2012) and more effective purifying selection in the asexual parasitoid wasp *Leptopilina clavipes* compared to a sexual relative (Kraaijeveld *et al.*, 2016). Notably, presence and extent of genomic signatures resulting from selection-driven consequences of asexuality have to be interpreted with caution because other factors, such as life history traits and population sizes, can influence genome evolution in similar ways as reproductive mode (Neiman & Schwander, 2011).

Studies that tested for increased levels of heterozygosity have yielded inconsistent results too, with genome-wide heterozygosity estimates ranging from near-complete homozygosity in several hexapods to too-large-to-be-quantifiable heterozygosity in the bdelloid rotifer *Adineta vaga* (Jaron *et al.*, 2019). In addition, genome-based resemblance of haplotype subtrees over populations (Meselson effect) has been shown only for *Trypanosoma* flagellates (Weir *et al.*, 2016), but absence of the Meselson effect can be due to homogenising processes other than sex such as mitotic gene conversion in asexual lineages (Butlin, 2002). The inability to systematically replicate genomic signatures of asexuals across species suggests that the absence of sex has far less extensive consequences for genome evolution than predicted and, hence, calls for further genome- and transcriptome-wide investigations of asexual lineages and their closely related sexual sisters (Jaron *et al.*, 2019). This thesis aims at providing important studies of well-suited and long-term asexual species.

Ancient asexuals

Some consequences of asexuality are widely assumed to be associated with detrimental fitness effects and doom asexual lineages to extinction in the long term through mutational meltdown (Gabriel *et al.*, 1993; Lynch *et al.*, 1993). The general designation of asexuality as evolutionary dead-end is substantiated by the observation that asexual lineages are generally concentrated at the tips of phylogenetic trees ("twiggy" distribution; Bell, 1982; Janko *et al.*, 2008; Schwander & Crespi, 2009). Contrastingly, few asexual lineages have persisted and even diversified in the absence of sex for considerable periods of time (Judson & Normark, 1996; Schoen *et al.*, 2009; Schwander *et al.*, 2011). Examples include the "scandalous" ancient asexual bdelloid rotifers, darwinulid ostracods and several clades of oribatid mites which lost sex tens of millions of years ago (Judson & Normark, 1996; Neiman *et al.*, 2009; Schoen *et al.*, 2009). Among these, bdelloid rotifers have, so far, received most attention (Mark Welch & Meselson, 2000; Flot *et al.*, 2013). Recent studies however, have indicated that cryptic gene exchange might render them quasi sexual (Signorovitch *et al.*, 2015; Debortoli *et al.*, 2016; Schwander, 2016; Vakhrusheva *et al.*, 2018). The "ancient asexual scandals" are particularly valuable because understanding how they managed to cope with the absence of sex will -*vice versa*- yield insight into the selective advantage of sex.

Not necessarily "ancient" but, too, regarded as old are asexual *Timema* stick insects. The stick insect genus features at least seven asexual lineages with lineage ages ranging from < 200,000 years to ~ 2 million years (Schwander *et al.*, 2011). *Timema* stick insects are particularly well suited for analysing genomic consequences of asexuality because they comprise several lineages that lost sex independently, thus providing "evolutionary replicates" of asexuality. For five asexual species sexual sister species are at hand and allow for independent pairwise comparisons of genome evolution under one or the other reproductive mode (Schwander & Crespi, 2009).

Oribatid mites, likewise, feature direct sex-asex sister pairs but additionally include several lineages that lost sex tens of millions of years ago and radiated in the absence of sex forming large, exclusively parthenogenetic clades (Norton & Palmer, 1991; Maraun *et al.*, 2003; Cianciolo & Norton, 2006; Heethoff *et al.*, 2007; Von Saltzwedel *et al.*, 2014).

Outline of this thesis

In this thesis, I tested if predicted consequences of asexuality on genome evolution are commonly present in animals. For this, I generated genome and transcriptome data of a variety of asexual and closely related sexual arthropod lineages and analysed them for commonly expected genomic consequences of asexuality. I utilised genome and transcriptome data to account for expected among-gene variation (Neiman *et al.*, 2018). I mainly focussed on accumulation of slightly deleterious mutations and intra-individual haplotype divergence ('Meselson effect'). My ***main hypothesis*** is that the accumulation of slightly deleterious mutations and haplotype divergence are common features of genome evolution under asexuality.

To test this hypothesis, I analysed if slightly deleterious mutations accumulate in three asexual oribatid mite species that lost sex tens of millions of years ago (*Chapter 1*) and contributed to analyses of various selection- and non-selection-driven consequences (see below) of a loss of sex in five species of asexual *Timema* stick insects (*Chapter 2*). Further, I applied the above-mentioned analyses to pre-existing transcriptome data of several distantly related asexual hexapod species from the 1 K Insect Transcriptome Evolution (1KITE) project (*Chapter 3*). Lastly, I tested if populations of asexual oribatid mites exhibit increased intra-individual haplotype divergence due to the Meselson effect (*Chapter 4*).

More specific summaries of the chapters:

In *Chapter 1*, I investigated if oribatid mites that lost sex tens of millions of years ago accumulated slightly deleterious mutations as a consequence of reduced effectiveness of purifying selection. The results contrast theoretical expectations in that purifying selection was even more effective in asexual as compared to sexual lineages in purging slightly deleterious nonsynonymous and synonymous mutations in the nuclear and mitochondrial genome.

In *Chapter 2*, I tested for presence and extent of various predicted consequences of asexuality, including analyses of within-population polymorphism and gBGC, in natural populations of *Timema* stick insects on a transcriptome-wide scale (see Contributions to the chapters of this thesis). The results strongly contrast those in oribatid mites in that they fit the classically predicted consequences of asexuality: asexual *Timema* lineages showed accumulation of slightly deleterious mutations at fixed and segregating nonsynonymous sites but no difference at synonymous sites, decreased levels of polymorphism, and reduced GC contents consistent with the predicted absence of gBGC in asexuals.

In *Chapter 3*, I expanded on the previous chapters by utilising pre-existing transcriptome data of asexual hexapod species from the 1KITE project. Here, accumulation of slightly deleterious mutations at nonsynonymous sites and their deleteriousness and selection acting on codon usage bias did not differ between asexual and sexual lineages. However, as the absence of a difference may be due to large phylogenetic distances between a majority of the analysed asexual lineages and their sexual counterparts the results have to be interpreted with care. The study implicates the need to account for (large) phylogenetic distances between asexual and sexual sister species by careful taxa selection in future studies.

In *Chapter 4*, I explored haplotype divergence in the asexual oribatid mite species *Oppiella nova* as compared to the closely related sexual species *Oppiella subpectinata*. Analyses based on

transcriptome data showed the presence of several population genomic patterns that indicate haplotype divergence due to obligate asexuality. Additionally, genome-based analyses revealed the presence of multiple, differently heterozygous lineages that coexist in different geographical populations. The results strongly support the presence of the Meselson effect and, hence, obligate asexuality in *O. nova.*

In *III General Discussion*, I present an overview of the studies generated in the framework of this thesis, studies done by others during the period of this thesis (~ 2015 - 2019) and earlier studies. Based on this, I show that neither accumulation of slightly deleterious mutations nor haplotype divergence are common in animals. I discuss potential reasons for this, e.g. the influence of population size and homogenising mechanisms; I illustrate factors that can lead to similar genomic signatures as deleterious mutation accumulation and the Meselson effect; and I highlight the importance of genome-based studies for analysing genomic consequences of asexuality.

II Research chapters

Chapter 1

Effective purifying selection in ancient asexual oribatid mites

Brandt, A., Schaefer, I., Glanz, J., Schwander, T., Maraun, M., Scheu, S. & Bast, J.

Nat. Comm. **8**, 873 (2017)

Sex is beneficial in the long term because it can prevent mutational meltdown through increased effectiveness of selection. This idea is supported by empirical evidence of deleterious mutation accumulation in species with a recent transition to asexuality. Here, we study the effectiveness of purifying selection in oribatid mites which have lost sex millions of years ago and diversified into different families and species while reproducing asexually. We compare the accumulation of deleterious nonsynonymous and synonymous mutations between three asexual and three sexual lineages using transcriptome data. Contrasting studies of young asexual lineages, we find evidence for strong purifying selection that is more effective in asexual as compared to sexual oribatid mite lineages. Our results suggest that large populations likely sustain effective purifying selection and facilitate the escape of mutational meltdown in the absence of sex. Thus, sex per se is not a prerequisite for the long-term persistence of animal lineages.

Introduction

One of the most challenging problems in evolutionary biology is to explain the maintenance of sex (Bell, 1982; Otto, 2009). Sex is coupled with substantial evolutionary costs, such as a twofold demographic cost due to the production of males, costs related to mate searching and mating (e.g. predator exposure and sexually transmitted diseases) and costs coupled with recombination (e.g. reduced likelihood of individual allele transmission and breakup of coadapted gene complexes; Williams, 1975; Maynard Smith, 1978; Lehtonen *et al.*, 2012). Despite these manifold costs, the mode of reproduction for the great majority of animal species is obligate sex. There is little agreement on which mechanisms generate selection for sex in natural populations in the short term (Jalvingh *et al.*, 2016). However, theoretical predictions and empirical evidence for reduced purifying selection in asexual eukaryotes have led to the established consensus that sex and recombination are beneficial for the long-term persistence of lineages. This benefit derives from the ability of sexual reproduction to decouple linked loci with different fitness effects generated by genetic drift, which increases the effectiveness of purifying selection to purge mildly deleterious mutations and reduces Hill–Robertson effects (Hill & Robertson, 1966; Felsenstein, 1974; Felsenstein & Yokoyama, 1976; Keightley & Otto, 2006). Further, sex and recombination enable the restoration of the least loaded genotypes, which would otherwise be lost by chance (Muller's ratchet; Muller, 1964).

Comparing rates of deleterious mutation accumulation in natural populations of sexual and asexual organisms is difficult as other factors than sex and recombination affect the effectiveness of selection. Most importantly, it is also determined by mutation rate and population size. Population size influences the strength of genetic drift in a population, which makes it an important determinant of mutational load and hence mutational decay of a population (Kimura *et al.*, 1963). Nevertheless, empirical estimates showed increased accumulation of deleterious

nonsynonymous mutations in effectively clonal lineages (hereafter referred to as asexual) as compared to closely related sexual lineages, e.g. in natural populations of *Timema* stick insects, *Campeloma* and *Potamopyrgus* freshwater snails and *Oenothera* evening primroses (Johnson & Howard, 2007; Neiman *et al.*, 2010; Henry *et al.*, 2012; Hollister *et al.*, 2015; Hartfield, 2016). Nonsynonymous mutation accumulation occurred at up to 13.4 times the rate in sexual lineages (Henry *et al.*, 2012). Further, there is evidence for accumulation of nonsynonymous and synonymous mutations in non-recombining genomic regions of sexual organisms, such as mitochondria or (neo-) Y chromosomes (Lynch & Blanchard, 1998; Bachtrog *et al.*, 2008).

All asexual lineages that were analysed for mutation accumulation in previous studies lost sex relatively recently between 5000 years and 1.5 million years ago. Whether similar or even more prominent patterns are observed in long-term asexual populations, i.e. after tens of millions of years without sex, remains an open question. Few asexual lineages are known that have persisted and diversified in the absence of sex, most prominently bdelloid rotifers, darwinulid ostracods and various clades of oribatid mites (Judson & Normark, 1996; Schoen *et al.*, 2009). Two studies investigated the accumulation of deleterious nonsynonymous mutations in bdelloid rotifers (Mark Welch & Meselson, 2001; Barraclough *et al.*, 2007), but recent investigations indicate that bdelloid rotifers engage in non-canonical forms of sex and cannot be considered as effectively clonal (Signorovitch *et al.*, 2015; Debortoli *et al.*, 2016; Schwander, 2016). Further, to our knowledge, there are no studies on mutation accumulation for darwinulid ostracods, only evidence of very slow background mutation rates (Schoen *et al.*, 2003). Oribatid mites lost sex multiple times independently, several million years ago, followed by extensive radiation of parthenogenetic clades as indicated by their phylogenetic distribution and high inter- and intraspecific molecular divergence (Norton & Palmer, 1991; Maraun *et al.*, 2003; Cianciolo & Norton, 2006; Heethoff *et al.*, 2007; Von Saltzwedel *et al.*, 2014). Thus this speciose, largely

soil-living animal group (~10,000 species; Schoen *et al.*, 2009) is well suited for comparative investigations of genomic consequences under long-term asexuality.

Here, we test if long-term asexuality in oribatid mites resulted in signatures of reduced effectiveness of selection by comparing the accumulation of deleterious nonsynonymous mutations and changes in codon usage bias (CUB). We analyse nuclear and mitochondrial orthologous genes from newly generated transcriptomic data of the three sexual species *Achipteria coleoptrata* (Linnaeus, 1758), *Hermannia gibba* (Koch, 1839) and *Steganacarus magnus* (Nicolet, 1855) and the three asexual species *Hypochthonius rufulus* (Koch, 1835), *Nothrus palustris* (Koch, 1839) and *Platynothrus peltifer* (Koch, 1839). We use two approaches to compare the effectiveness of purifying selection between asexual and sexual oribatid mite lineages on the transcriptome scale. First, we analyse the accumulation of nonsynonymous point mutations and infer their potential 'deleteriousness'. Second, we infer the effectiveness of selection on CUB since purifying selection also acts on synonymous sites (Hershberg & Petrov, 2008). Analyses are based on 3545 nuclear orthologous genes shared among the six species (see 'Methods'). Additionally, we analyse two nuclear genes (*hsp82* and *ef1α*) of 30 species (nine sexual, 21 asexual species) from different phylogenetic groups (Fig. 1; Domes *et al.*, 2007). We find no evidence for reduced purifying selection in asexual lineages, neither for elevated accumulation of deleterious nonsynonymous point mutations nor for 'relaxed' selection on CUB. Surprisingly, our results indicate even more effective purifying selection in asexual as compared to sexual oribatid mites.

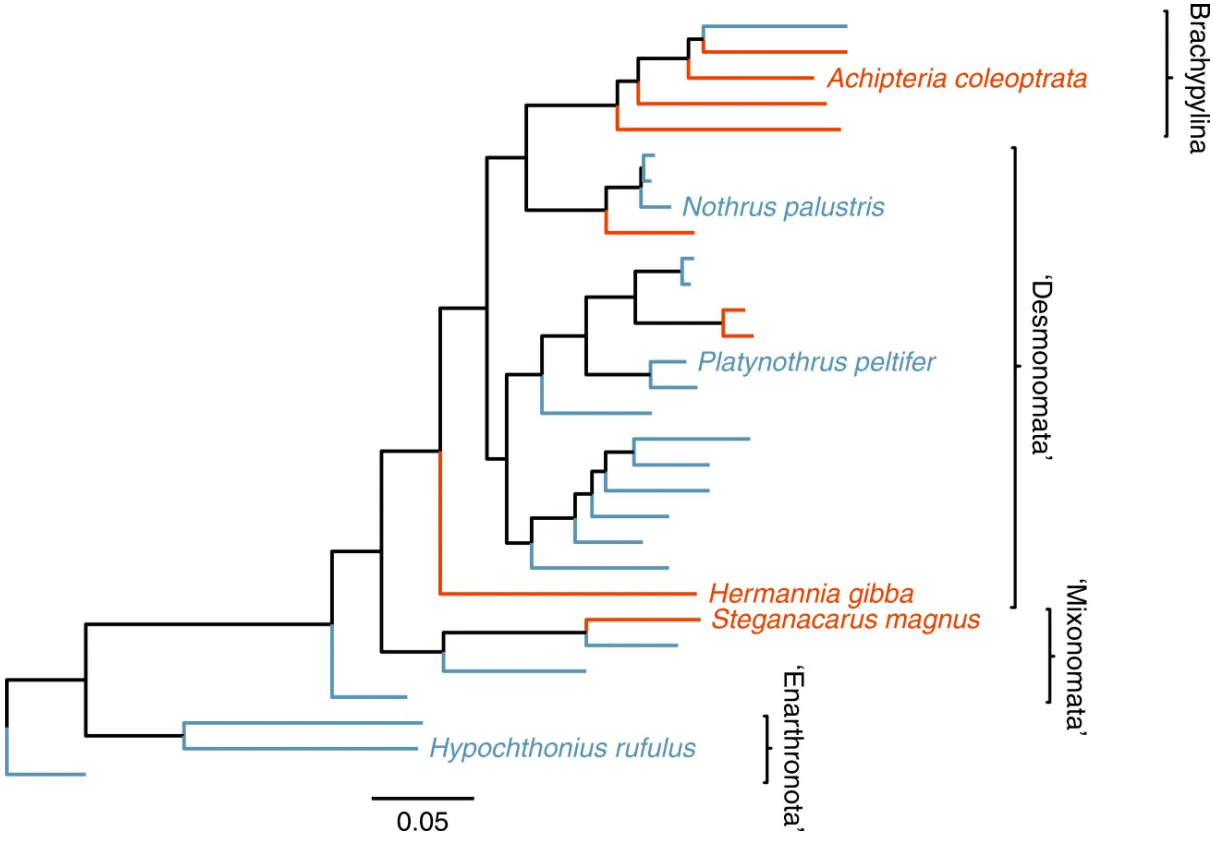


Figure 1: Phylogenetic tree of 30 oribatid mite species analysed in this study. The tree is based on partial sequences of 18S rDNA, *ef1α* and *hsp82* (see 'Methods'). The taxon sampling covers asexual (blue) and sexual (red) species from four out of six major groups of oribatid mites (brackets). Species used for transcriptome-wide analyses are indicated with full name, remaining species names are listed from top to bottom in Supplementary Table 1.

Results

Accumulation of nonsynonymous mutations

To estimate the accumulation rate of nonsynonymous point mutations in orthologs, we computed the ratio of nonsynonymous to synonymous divergence (d*N*/d*S*) as a measure of amino acid changes normalised for background mutation rates (see 'Methods'). Given the long divergence time between the six species used for sequencing of transcriptomes (>200 million years; Schaefer *et al.*, 2010), detectable orthologs shared among the six species are expected to be under strong

purifying selection (d*N*/d*S*<1). In genes under strong purifying selection the majority of nonsynonymous changes are likely mildly deleterious, hence an increased accumulation of deleterious nonsynonymous mutations results in a higher d*N*/d*S* ratio (Li *et al.*, 1985). We analysed nonsynonymous mutation accumulation exclusively at terminal branches of the phylogenetic tree, because character states (i.e. sexual or asexual reproduction) at internal branches are uncertain for methodological reasons (Goldberg & Igić, 2008). We used a model of codon evolution estimating one d*N*/d*S* ratio for internal branches, one for terminal branches of asexual species and one for terminal branches of sexual species (from here referred to as terminal asexual and sexual branches, respectively; three-ratio model; see 'Methods'). As expected, all 3545 orthologs were under purifying selection at terminal branches (mean d*N*/d*S* = 0.084). Surprisingly and contrary to expectations, per-gene d*N*/d*S* ratios were on average lower at asexual branches as compared to sexual branches (mean $\Delta_{asex-sex} = -0.008$; Wilcoxon signed-rank test $V = 4.84*10^6$, $P < 0.001$; Fig. 2). Synonymous substitution rates (d*S*) did not differ between reproductive modes (means of 1.112 and 1.146 for asexual and sexual branches, respectively; gene effect $P = 0.951$, reproductive mode effect $P = 0.218$, interaction $P = 0.999$; permutation ANOVA), indicating that the differences in d*N*/d*S* ratios between reproductive modes were not driven by differences in d*S*.

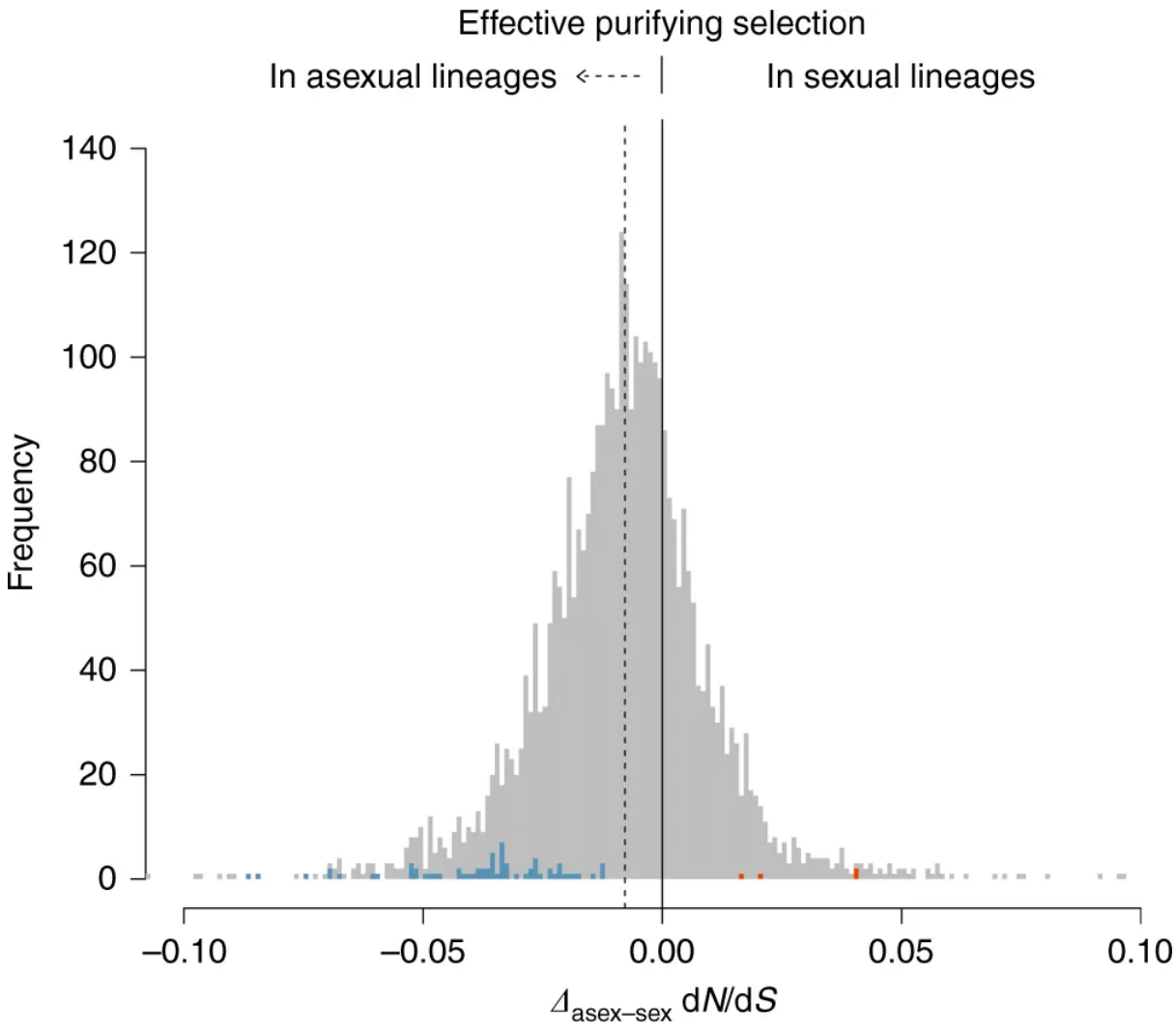


Figure 2: Per-gene differences in d*N*/d*S* ratios between reproductive modes. The histogram shows the distribution of per-gene differences in d*N*/d*S* ratios between asexual and sexual terminal branches in a tree comprising three sexual and three asexual oribatid mite species for 3539 orthologous loci under purifying selection (d*N*/d*S* < 1). To improve data presentation, the histogram range is limited to −0.1 to 0.1, representing 3539 out of 3545 orthologs. The distribution mean (−0.008; dotted line) is shifted to the left indicating overall lower d*N*/d*S* ratios at asexual as compared to sexual branches (Wilcoxon signed-rank test $V = 4.84*10^6$, $P < 0.001$). Frequencies of genes with significantly lower d*N*/d*S* ratios at asexual as compared to sexual branches or sexual as compared to asexual branches are coloured blue and red, respectively.

Although orthologs overall were under stronger purifying selection in asexual as compared to sexual oribatid mites, d*N*/d*S* values as well as their difference between asexuals and sexuals varied widely among orthologs (ranges: d*N*/d*S* 0–0.627; $\Delta_{asex-sex}$ -0.172–0.540; variance: $4.6*10^{-4}$). We therefore also analysed which specific orthologs differed significantly in their effectiveness of purifying selection between reproductive modes. To identify orthologs under significantly stronger purifying selection in asexual and sexual lineages, we tested if the three-ratio model was a better fit to the data than a model allowing for one d*N*/d*S* ratio for terminal and internal branches, without discriminating between reproductive modes (two-ratio model; see 'Methods'). For 67 orthologs, d*N*/d*S* ratios were significantly lower at asexual as compared to sexual branches and for five orthologs at sexual as compared to asexual branches (coloured bars; Fig. 2). The orthologs with strong selection in asexual mites were enriched for Gene Ontology terms related to e.g. immune response, mitotic cell cycle and germline cell division (Supplementary Data 1).

Deleteriousness of nonsynonymous mutations

In addition to elevated rates of nonsynonymous mutation accumulation, reduced purifying selection in sexual oribatid mite lineages is expected to result in fixation of nonsynonymous mutations that have stronger deleterious effects as compared to asexual lineages. We therefore analysed the potential 'deleteriousness' of the amino acid substitutions. For this, we inferred changes in hydrophobicity scores (HS) between ancestral and replacement amino acids as a measurement of the 'deleteriousness' of nonsynonymous mutations (see 'Methods'), because protein folding is influenced by changes in hydrophobic properties at amino acid substitution sites (Pace *et al.*, 2011). For nuclear orthologs, transitions between amino acids with more dissimilar hydrophobicity (HS < 90) involved 46,857 changes (50.74% of overall changes) at asexual and 68,115 (50.98%) at sexual branches. In accordance with the faster nonsynonymous

mutation accumulation in sexual than asexual oribatid mites reported above, amino acids shifted to more dissimilar hydrophobicity at sexual branches (generalised linear mixed model $z = 2.4$, $P = 0.017$; Fig. 3a), indicating stronger 'deleteriousness' of nonsynonymous changes in sexual as compared to asexual oribatid mites.

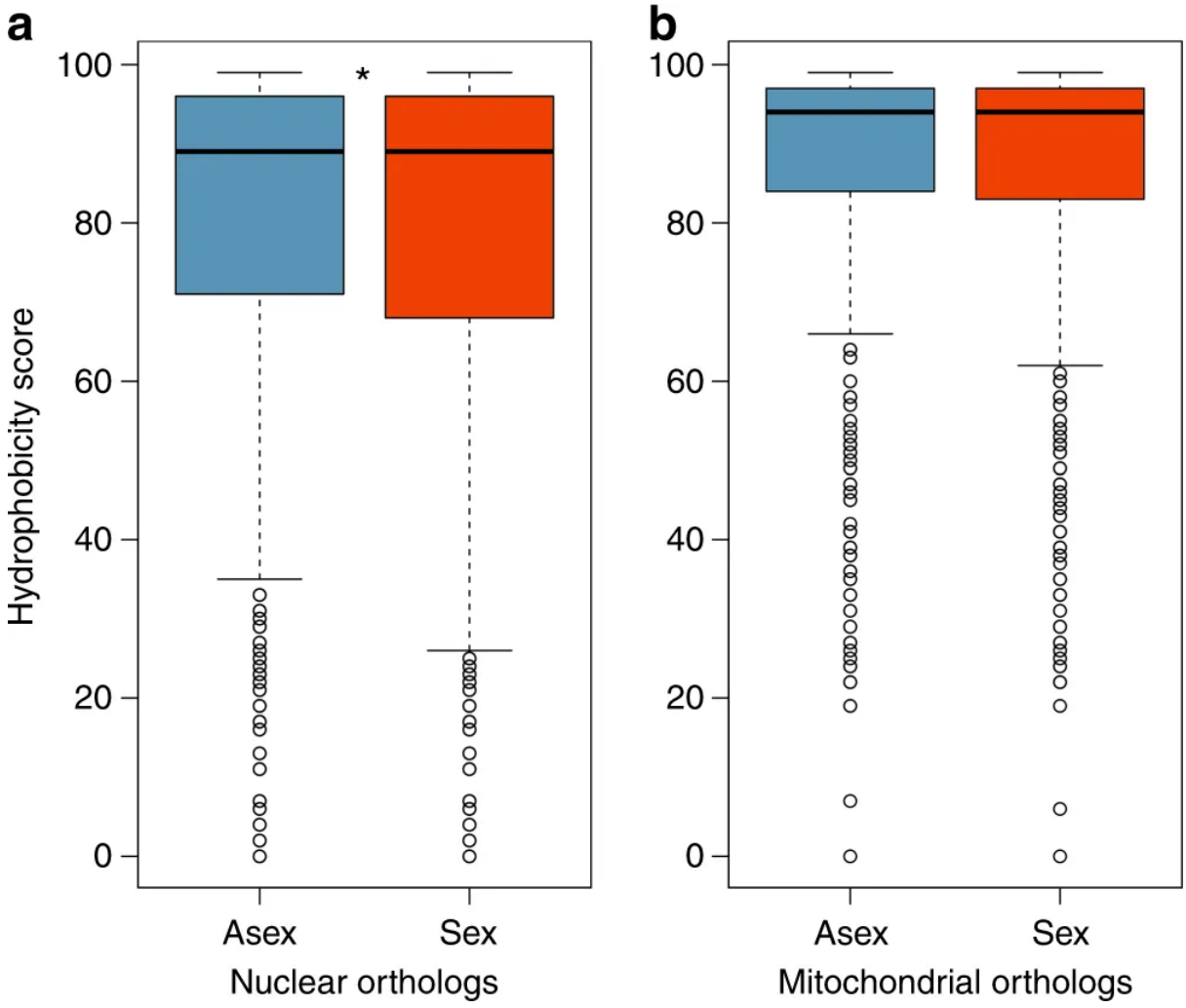


Figure 3: Hydrophobicity scores at asexual and sexual terminal branches. The boxplots show hydrophobicity scores (HS) at three asexual (blue) and three sexual (red) terminal branches for (a) 3545 nuclear orthologous genes (involving 92,351 and 133,606 HS at asexual and sexual terminal branches, respectively) and (b) ten mitochondrial orthologous genes (involving 1494 and 2468 HS at asexual and sexual terminal branches, respectively). HS measures the strength in changes of hydrophobicity from ancestral to replacement amino acids and indicates the 'deleteriousness' of a nonsynonymous mutation. The lower the HS the stronger is the change in hydrophobicity and 'deleteriousness' of a nonsynonymous mutation. For nuclear orthologs, amino acids shifted to more dissimilar hydrophobicity at sexual branches (generalised linear mixed model $z = 2.4$, $P = 0.017$), indicating stronger 'deleteriousness' of nonsynonymous changes in sexual as compared to asexual oribatid mites. For mitochondrial orthologs, there was no difference (generalised linear mixed model $z = 0.81$, $P = 0.421$). Significant differences in HS are marked with an asterisk ($^*P < 0.05$). Whiskers correspond to 1.5 times the interquartile range.

Analyses of the larger taxon sampling

Overall, based on the analyses of transcriptomic orthologs shared among six species, asexual oribatid mite lineages accumulated less deleterious nonsynonymous mutations as compared to sexual lineages. However, the small number of taxa can generate branch length uncertainties which could affect inferences of d*N*/d*S* and ancestral amino acid sequences. To reduce this problem, we extended the taxon sampling to a total of 30 species. Using the same approaches, we compared nonsynonymous mutation accumulation among 21 asexual and nine sexual oribatid mite species (Fig. 1, Supplementary Table 1) in two nuclear genes (*ef1α* and *hsp82*) generated in a previous study (Domes *et al.*, 2007). Both genes were under strong purifying selection at terminal branches as indicated by d*N*/d*S* ratios (0.045 and 0.034 for *ef1α* and *hsp82*, respectively). There were no significant differences between asexual and sexual branches as the two-ratio model provided a better fit to the data than the three-ratio model (Table 1). Synonymous substitution rates (d*S*) did not differ between reproductive modes, ($t = 0.627$, $P = 0.540$ and $t = 1.246$, $P = 0.237$ for *ef1α* and *hsp82*, respectively; *t*-test). This lack of a significant difference between d*N*/d*S* ratios is unlikely the result of low variation in *ef1α* and *hsp82*, as mean pairwise divergence of amino acid sequences was high in both genes (0.069 and 0.109 for *ef1α* and *hsp82*, respectively). Consistent with d*N*/d*S* ratios, there were no differences in hydrophobicity changes at asexual and sexual branches for both genes ($W = 1637$, $P = 0.542$ and $W = 2235.5$, $P = 0.624$ for *ef1α* and *hsp82*, respectively; Wilcoxon rank-sum test).

As indicated by the transcriptome analyses described above, the d*N*/d*S* ratio differences between sexual and asexual taxa varied considerably among orthologs and only analyses of large gene sets are likely to reveal putative differences between reproductive modes. Indeed the gene *ef1α*, although present in the transcriptome ortholog set, did not feature significantly different d*N*/d*S* ratios between reproductive modes (0.015 and 0.026 at asexual and sexual branches, respectively;

$\chi^2 = 3.157$, $P = 0.076$; likelihood ratio test). *Hsp82* was absent from the ortholog data set, preventing estimation of d*N*/d*S* ratios from the transcriptomes for this gene.

Table 1: Results of the branch-specific dN/dS ratio analyses of *ef1α* and *hsp82* for 21 asexual and 9 sexual oribatid mite species. $asex_T$: terminal asexual branches; I: internal branches; lnL: ln(Likelihood); sex_T: terminal sexual branches; T: terminal branches. Non-significant *P*-values indicate that the two-ratio model provides the better fit to the data than the three-ratio model for both genes. Therefore, asexual and sexual branches did not differ in the effectiveness of selection on nonsynonymous sites in these two genes.

Model	**Free parameters**	**lnL**	**Likelihood ratio test**	**dN/dS**
ef1α				
2 ratios (T/I)	2	-6115.434		T: 0.045 I: 0.036
3 ratios (sex_T/$asex_T$/I)	3	-6114.837	vs. 2 ratios (T/I): $X^2 = 1.193$ $P = 0.275$	sex_T: 0.051 $asex_T$: 0.041 I: 0.036
hsp82				
2 ratios (T/I)	2	-4973.451		T: 0.034 I: 0.026
3 ratios (sex_T/$asex_T$/I)	3	-4972.167	vs. 2 ratios (T/I): $X^2 = 2.567$ $P = 0.109$	sex_T: 0.04 $asex_T$: 0.03 I: 0.026

Accumulation of synonymous mutations

Although mutations at synonymous sites are generally considered to be neutral, they can be under selection for example because different codons influence the speed and accuracy of protein translation (Hershberg & Petrov, 2008). Hence, we also assessed the effectiveness of selection acting at synonymous sites by estimating CUB. This method is particularly robust because measurements of CUB do not depend on likelihood estimates, tree topologies or branch length estimates. As a metric, we used the codon deviation coefficient (CDC; Zhang *et al.*, 2012), which calculates the deviation from expected CUB and accounts for background nucleotide composition, thus allowing for cross-species comparisons. A lower CDC value indicates more 'relaxed' selection on CUB (see 'Methods'). We compared gene-specific CDC values for each of

the 3545 orthologous nuclear genes between reproductive modes. Consistent with the results for nuclear d*N*/d*S* ratios, per-gene CDC was slightly but significantly higher in asexual species (Fig. 4a; means of 0.130 and 0.128 for asexual and sexual species, respectively; gene effect $P < 0.001$, reproductive mode effect $P = 0.008$, interaction $P = 0.616$; permutation ANOVA), indicating more effective purifying selection at synonymous sites in asexuals.

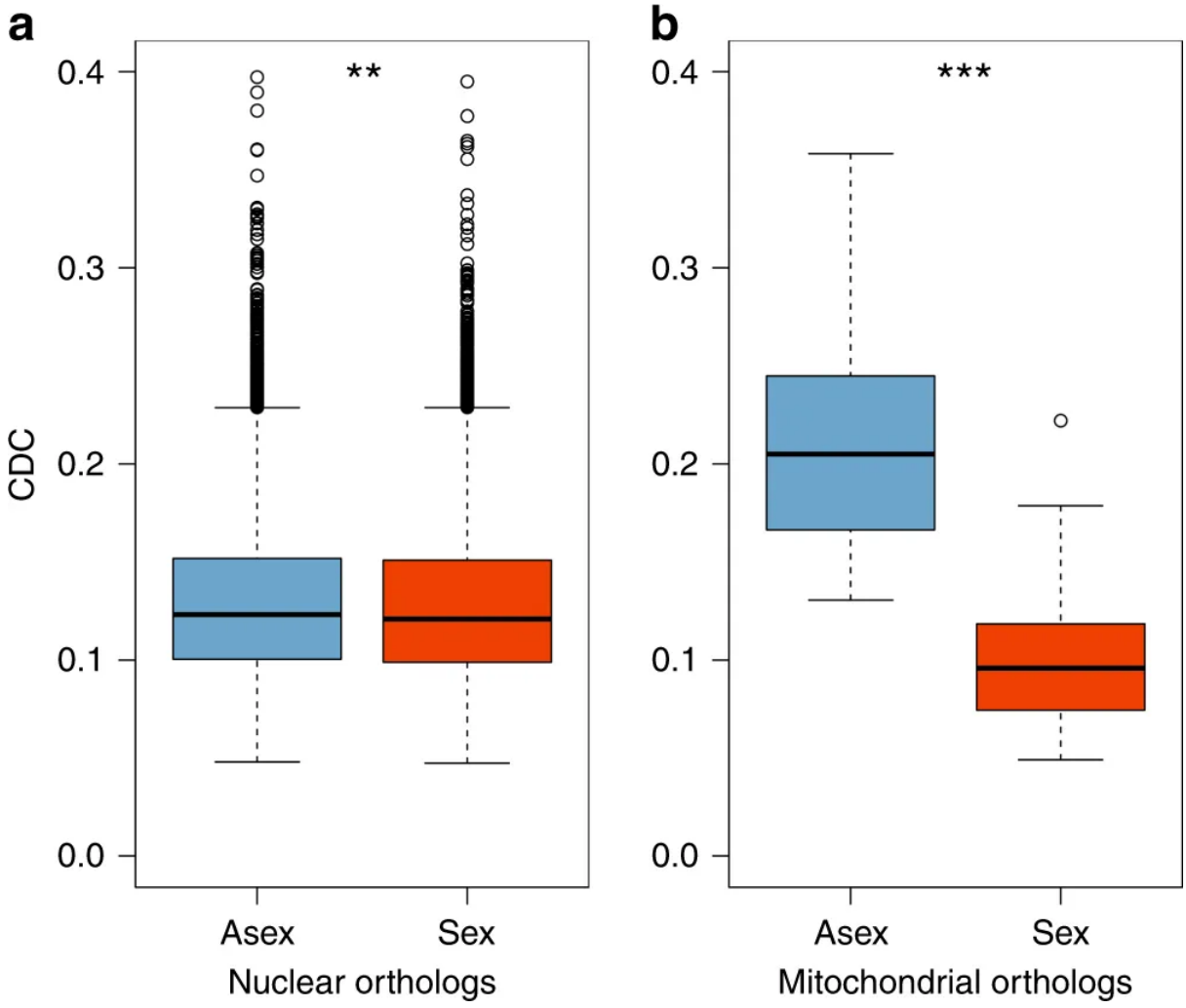


Figure 4: Comparison of CDC values between asexual and sexual oribatid mite species. Boxplots show CDC values of three asexual (blue) and three sexual (red) oribatid mite species for (a) 3545 nuclear orthologous genes (involving 10,635 CDC values for both asexual and sexual species) and (b) ten mitochondrial orthologous genes (involving 30 CDC values for both asexual and sexual species). CDC measures the deviation of observed from expected codon usage and allows for inference of the effectiveness of purifying selection acting on synonymous sites. A lower CDC value corresponds to more 'relaxed' selection on codon usage bias (see 'Methods'). For nuclear orthologs, per-gene CDC was slightly but significantly higher in asexual species (gene effect $P < 0.001$, reproductive mode effect $P = 0.008$, interaction $P = 0.616$; permutation ANOVA). For mitochondrial orthologs, per-gene CDC differed more strongly between reproductive modes (gene effect $P = 0.002$, reproductive mode effect $P < 0.001$, interaction $P = 0.960$; permutation ANOVA). Significant differences in CDC are marked with asterisks (***$P < 0.001$ and **$P < 0.01$). Whiskers correspond to 1.5 times the interquartile range.

Accumulation of mutations in the mitochondria

Next, we analysed mitochondrial genes, which allows for comparing the effectiveness of purifying selection among non-recombining genomes with different nuclear genomic backgrounds. We analysed the effectiveness of purifying selection in ten orthologous mitochondrial genes (*atp6, cob, cox1, cox2, cox3, nd1, nd2, nd3, nd4* and *nd5*) using the same approaches as for nuclear genes. Surprisingly, and contrasting to the results of nuclear orthologs, d*N*/d*S* ratios were lower at terminal sexual branches as compared to terminal asexual branches (mean $\Delta_{\text{asex–sex}} = 0.027$; Wilcoxon signed-rank test $V = 2$, $P = 0.006$). However, d*S* rates were highly elevated for sexual as compared to asexual branches (means of 2.814 and 5.577 for asexual and sexual branches, respectively; gene effect $P = 0.044$, reproductive mode effect $P = 0.002$, interaction $P = 0.876$; permutation ANOVA). This raises the question if the observed differences in mitochondrial d*S* between reproductive modes could be influenced by selection at synonymous sites (Ticher & Graur, 1989). Indeed, synonymous mitochondrial substitution rates (d*S*) and selection at synonymous sites as measured by CDC were negatively correlated in sexual lineages but not in asexual lineages ($r = -0.37$, $P = 0.042$ and $r = 0.16$, $P = 0.396$, respectively; Spearman's rank correlation). Because d*S* rates strongly differed and were adversely influenced by selection at synonymous sites between reproductive modes, mitochondrial d*N*/d*S* ratios are not comparable. Based on d*N* rates only, there was no difference in asexual and sexual nonsynonymous changes (means of 0.173 and 0.231 for asexual and sexual branches, respectively; gene effect $P = 0.037$, reproductive mode effect $P = 0.138$, interaction $P = 0.994$; permutation ANOVA). Consistent with d*N* rates, the analysis of hydrophobicity changes ('deleteriousness') of amino acids for mitochondrial orthologs showed no overall difference between reproductive modes (generalised linear mixed model $z = 0.81$, $P = 0.421$; Fig. 3b). At mitochondrial synonymous sites, effectiveness of selection was strongly reduced for mitochondria in sexual lineages as compared to those in asexual lineages (Fig. 4b; mean CDC of

0.211 and 0.104 for asexual and sexual species, respectively; gene effect $P = 0.002$, reproductive mode effect $P < 0.001$, interaction $P = 0.960$; permutation ANOVA).

Discussion

In metazoans it is believed that sex is important for the long-term persistence of a lineage because it facilitates effective purifying selection and rapid adaptation despite the effect of drift in finite populations. In contrast to this established consensus, our study showed no evidence for accumulation of deleterious nonsynonymous and synonymous point mutations in oribatid mite lineages that survived without sex for several million years. Moreover, our results suggest that purifying selection acts even more effectively in asexual as compared to sexual oribatid mite lineages. It is thus unlikely that asexual oribatid mite lineages drift towards extinction because of mutation accumulation. While our taxon sampling for d*N*/d*S* ratio estimates of nuclear orthologs in the transcriptome scale analyses was small, the overall pattern was substantiated by results from the more extended taxon set (Figs. 1 and 2, Table 1). Moreover, synonymous substitution analyses (CDC, Fig. 4) are independent of assumptions about phylogenetic relationships among taxa and these analyses likewise supported the result of more effective purifying selection in asexuals.

Our finding of more effective purifying selection despite the lack of sex is unique among investigated asexual eukaryotes. Previous studies that compared sexual and asexual eukaryotes typically found increased rates of nonsynonymous mutation accumulation in asexual as compared to sexual lineages, as predicted by theory (reviewed by Hartfield, 2016). Where no differences were found, sex was lost probably too recently to detect accumulation of mutations (Ollivier *et al.*, 2012; Ament-Velásquez *et al.*, 2016). Increased rates of nonsynonymous mutation

accumulation and ‘relaxed’ CUB further extends to self-fertilising eukaryotes, although selfers often maintain low rates of outcrossing (Jarne & Auld, 2006; Hartfield, 2016).

What mechanisms could account for the escape of asexual oribatid mites from the ‘mutational meltdown’? Comparing the peculiarities of oribatid mites with those of animal groups including other old asexual lineages may help resolving this question. Oribatid mites are small animals (150–1400 µm) with a wide geographical distribution (Maraun *et al.*, 2012). Both features also hold for other old asexual groups, such as darwinulid ostracods and meloidogyne root-knot nematodes, and contrast most asexual groups that accumulate deleterious mutations (van Doninck *et al.*, 2003; Sahu *et al.*, 2015). Generally, small body size and large geographic range are indicative of high abundances of organisms (Gaston *et al.*, 1997; White *et al.*, 2007), which suggests that large population sizes might alleviate the negative effects of loss of sex.

Traditional models that predict mutation accumulation in asexuals are based on population sizes that are typical for many macroorganisms, but that might not be applicable to asexuals with large populations (Muller, 1964; Hill & Robertson, 1966; Felsenstein, 1974). Indeed, large population sizes were proposed to maintain effective purifying selection in the absence of sex (Normark & Johnson, 2011; Ross *et al.*, 2013), because the speed of mutation accumulation in asexual populations is predicted to drop substantially with increasing population sizes and eventually to become biologically irrelevant (Lynch *et al.*, 1993; Gordo & Charlesworth, 2000; Schoen *et al.*, 2009).

Consistent with the idea that large population size can protect asexual lineages from mutation accumulation, total oribatid mite densities are often very large and can reach up to ~350,000 ind. m^{-2} in temperate and boreal forests. At high densities (>100,000 ind. m^{-2}) asexual species typically contribute >55% to total species numbers but >80% to total density (Maraun *et al.*, 2012), indicating substantially larger population sizes of asexual as compared to sexual

oribatid mites. This observation is reflected by our findings of purifying selection being more effective in asexual oribatid mite lineages as compared to sexual lineages in both the nucleus and mitochondria. The observation of a strong difference for selection at synonymous sites for non-recombining mitochondria in sexual and asexual nuclear backgrounds supports the conclusion that asexual oribatid mite populations are larger than those of sexual oribatid mites (Fig. 4b). Remarkably, both animal population sizes and the frequency of parthenogenetic reproduction are generally high in soil animals (including protozoans, nematodes, enchytraeids, collembolans, isopods and oribatid mites ;Maraun *et al.*, 2012; Veresoglou *et al.*, 2015), suggesting that more ancient asexual scandals may be hidden among these poorly studied groups.

Large population sizes can theoretically be sufficient to explain effective purifying selection in asexual oribatid mites, but additional non-mutually exclusive mechanisms could contribute to the reduction of the mutational load. Substitution rates can be altered via molecular mechanisms. For example, gene conversion has prevented mutation accumulation in the human Y chromosome and higher plant chloroplasts (Khakhlova & Bock, 2006; Hughes *et al.*, 2012) and DNA-repair contributes to the maintenance of DNA integrity in animal mitochondria (Kang & Hamasaki, 2002). However, as the CDC mainly reflects the impact of purifying selection rather than differences in mutation rates, large population size is likely the most important factor in preventing the accumulation of deleterious mutations in asexual oribatid mites. Finally, some non-canonical form of sex (e.g. horizontal gene transfer between bdelloid rotifers; Debortoli *et al.*, 2016), in oribatid mites cannot formally be excluded, although it seems unlikely that this would result in more effective selection than common bisexual reproduction.

In conclusion, we conducted several detailed analyses of mutation accumulation in transcriptomes of three asexual and sexual oribatid mite lineages and found purifying selection to be more effective in asexual lineages. This contrasts empirical evidence from younger asexual

lineages and non-recombining genomic regions. Additional studies in other ancient asexual animal groups would allow testing of whether escaping deleterious mutation accumulation is a general feature of ancient asexuals. To what extent population sizes and/or molecular mechanisms contribute to the escape of mutational meltdown remains to be investigated. The results here indicate that deleterious mutation accumulation is unlikely to drive extant asexual oribatid mite lineages to extinction and challenge the established consensus that loss of sex necessarily leads to reduced effectiveness of purifying selection and mutational meltdown in the long term.

Methods

Animal sampling and sequencing of transcriptomes

For analyses of transcriptomic data, we selected three sexual species (*A. coleoptrata:* Brachypylina, *H. gibba:* Desmonomata, *S. magnus:* Mixonomata) and three asexual species (*H. rufulus:* Enarthronota, *N. palustris:* Desmonomata, *P. peltifer:* Desmonomata) representing four of the six major oribatid mite groups (Domes *et al.*, 2007; Fig. 1). Animals were collected from litter in different forests in Germany (Goettinger Wald, Solling), extracted from litter alive using heat gradient extraction (Kempson *et al.*, 1963), identified following Weigmann (2006) and starved for at least one week to avoid contamination by gut content RNA. For RNA extraction, between 1 and 50 individuals were pooled, depending on body size of the species, to obtain sufficient RNA for library construction while minimising the numbers of individuals pooled. These numbers were not biased towards sexual and asexual species. RNA extraction was done using Qiagen and ZymoResearch RNA extraction kits, following manufacturer's instructions and RNA concentrations were measured using the Qubit 2.0 fluorometer (Invitrogen). Library preparation and sequencing of *A. coleoptrata*, *H. rufulus*, *N. palustris* and *H. gibba* was performed on Illumina HiSeq 2500 platforms, yielding 125 bp reads, by GATC Biotech

(Constance, Germany) and of *P. peltifer* and *S. magnus* on Illumina MiSeq platforms, yielding 250 bp reads, by the Transcriptome and Genome Analysis Laboratory (TAL, Georg-August-University Goettingen, Germany). For detailed information, see Supplementary Table 2.

Transcriptome assembly and data processing

Quality of raw reads was controlled, checked for adaptors and the first 12 bases of each read were clipped and bases with quality scores below 30 were trimmed from the end of the reads using Trimmomatic version 0.33 (Bolger *et al.*, 2014). Trimmed read pairs were assembled with Trinity version 2.0.3 (Grabherr *et al.*, 2011) on the scientific computing clusters of the Georg-August-University Goettingen. The Trinity assembly process includes the reconstruction of genetic polymorphism resulting in different isoforms and is effective in avoiding the assembly of chimeric isoforms. Remains of adaptors and sequences matching to non-oribatid mite species in the NCBI database were removed. The cleaned assemblies were deposited under NCBI BioProject PRJNA339058. To avoid spurious contigs from sequencing errors, only contigs with RPKM (Reads Per Kilobase per Million mapped reads) > 2 were selected and the most abundant isoform per gene extracted following Harrison *et al.* (2015). The most abundant isoform represents the allele that is shared by most individuals in a pooled sample. As we analysed (long-term) fixed effects of mutation accumulation and not (short-term) segregating polymorphisms, the most abundant allele in the population is most suitable. To assess the quality of the generated transcriptomes and the subsequent RPKM filtered transcript set, we used BUSCO version 2.0 (Simão *et al.*, 2015) to score completeness and fragmentation of the assemblies (Supplementary Table 3). BUSCO uses a core set of genes that are present in single copy in 90% of arthropods for detection in the transcriptome set. All filtered assemblies contained a high fraction of full-length arthropod core genes (‘complete’; >90%) and low levels

of fragmentation, with the exception of *P. peltifer* (resulting from adaptor contamination in the reads that were trimmed). However, all mite assemblies were missing only very few core genes (<5%). Using the filtered assemblies, open reading frames were detected using TransDecoder version 2.0.1 (Haas *et al.*, 2013). For subsequent analyses of branch-specific d*N*/d*S* ratios, CUB and hydrophobicity changes, orthologous ORFs were detected using reciprocal best alignment heuristics implemented in Proteinortho version 5.11 (Lechner *et al.*, 2011) with amino acid sequences of detected ORFs as input. Amino acid sequences that were shared by all species were aligned with MUSCLE version 3.8.31 (Edgar, 2004) and back-translated using RevTrans version 1.4 (Wernersson & Pedersen, 2003). Alignments were curated (including deletion of triplet gaps) using Gblocks version 0.91b with parameters *-t = c*, *-b4* = 4 (Castresana, 2000), alignments < 200 bp were excluded from further analyses, leaving 3545 alignments (mean alignment length 943 bp) for d*N*/d*S* ratio and codon usage analysis. As all analyses were based on alignments of orthologous ORFs from all six species, and curated to yield a uniform alignment size, the described differences in the assemblies did not affect selection estimates. For detailed information, see Supplementary Table 3. Mitochondrial genes were identified with blastn and tblastx and standard parameters using mitochondrial genes derived from genome data (Accession: PRJNA280488) and mitochondrial genes of *S. magnus* (Accession: EU935607) as query sequences (Domes *et al.*, 2008; Bast *et al.*, 2015). Alignment and alignment processing was done as described above for nuclear orthologous genes. For the general phylogeny and the analyses on the increased taxon sampling, we similarly aligned partial coding sequences of *ef1α*, *hsp82* and 18S rDNA of 30 oribatid mite species, generated in a previous study (Domes *et al.*, 2007). Sequences of these genes were available from NCBI, except for *N. palustris* for which sequences were extracted from the transcriptome data generated in this study using blast (Fig. 1, Supplementary Table 1).

Phylogenetic tree estimations for d*N*/d*S* ratio analysis

To reconstruct the phylogeny of the large taxon sampling of all 30 species (that includes the six species for which transcriptomes were generated), the alignments of *ef1α*, *hsp82* and 18S were concatenated and passed to RAxML version 3.1 (Stamatakis, 2014) with the GTRGAMMAI model set as model of sequence evolution (Fig. 1). The non-coding 18S rDNA was included in a combined alignment of *ef1α* and *hsp82* for construction of the general phylogeny, as 18S rDNA is highly conserved among oribatid mites and improved the reliability of the resulting topology. This topology was used in subsequent analyses. For d*N*/d*S* ratio analyses, the fixed unrooted species tree, gene-specific alignments and gene-specific branch lengths are required. For nuclear and mitochondrial orthologs derived from the transcriptome data, the 30-taxa topology was compacted to contain the six species only and per-gene branch lengths were calculated using this fixed tree together with the processed alignments with RAxML and the GTRGAMMAI model of sequence evolution. Similarly, for the large taxon set, branch lengths of *ef1α* and *hsp82* were calculated using the 30-species tree. For all gene-specific trees, branch lengths were multiplied with three to match the requirements of CodeML for testing models of codon evolution.

Branch-specific d*N*/d*S* ratio analyses

To infer nonsynonymous mutation accumulation in the three sexual and asexual lineages on transcriptome level, the alignments of 3545 orthologous genes were analysed using CodeML, implemented in the PAML package version 4.8 (Yang, 2007). To get gene-specific d*N*, d*S* and d*N*/d*S* ratios, a custom script was used to pass each gene alignment together with the fixed species tree, appropriate branch labels (according to the model) and the respective individual branch lengths generated with RAxML to CodeML (available at https://github.com/ptranvan/mites2codeml). CodeML utilises a Maximum Likelihood framework to estimate the goodness of fit of a codon substitution model to the tree and the sequence

alignment and calculates branch-specific d*N*/d*S* ratios. First, a model allowing for one d*N*/d*S* ratio for sexual, asexual and internal branches each was used to calculate d*N*/d*S* ratios (three-ratio model). The d*N*/d*S* ratios of terminal asexual and sexual branches were compared using a Wilcoxon signed-rank test in R version 3.0.2 (R Core Team, 2015). To investigate if the results were influenced by high synonymous substitution rates rather than by low nonsynonymous substitution rates, values of d*S* and d*N* were compared between terminal sexual and asexual branches using a non-parametric permutation ANOVA with 1000 bootstrap replicates, that does not require specific assumptions for distributions of data (as utilised in previous studies; Manly, 1997; Henry *et al.*, 2012). To identify orthologous genes that were under significantly stronger purifying selection either at asexual or at sexual branches, the goodness of fit of the three-ratio model was compared to that of a model allowing for only one d*N*/d*S* ratio at terminal and internal branches, respectively (two-ratio model), using a likelihood ratio test (coloured bars, Fig. 2). Resulting *P* values were corrected for multiple testing using the R package qvalue (Storey, 2015). To corroborate these analyses with a larger taxon set, branch-specific d*N*/d*S* ratios were predicted using the alignments of *ef1α* and *hsp82* with the respective phylogenetic trees. Differences between terminal sexual and asexual branches were tested for significance as described above. To test if high synonymous mutation rates rather than low nonsynonymous mutation rates affected the results, values of d*S* and d*N* were compared between terminal sexual and asexual branches using a *t*-test. Additionally, mean pairwise amino acid divergences were calculated for *ef1α* and *hsp82* to test if divergence was sufficiently high to detect possible differences between d*N*/d*S* ratios at terminal sexual and asexual branches.

Analysis of mutation 'deleteriousness'

To infer the 'deleteriousness' of nonsynonymous mutations, changes in hydrophobicity at amino acid replacement sites were analysed. To infer the ancestral states of amino acids for the nuclear

orthologous genes, first Proteinortho was used to predict orthologous loci among translated ORFs of the six oribatid mite species and the transcriptome of *Tetranychus urticae*, that was used as outgroup (Accession: PRJNA315122; Grbić *et al.*, 2011). Orthologous ORFs were aligned using MUSCLE, back-translated using RevTrans and curated using Gblocks with similar parameters as above. Alignments < 200 bp were excluded from further analysis. Branch lengths were calculated for each gene, individually, using the processed alignments with fixed topologies including *T. urticae* as outgroup and the GTRGAMMAI model of sequence evolution. For subsequent prediction of ancestral amino acids, branch lengths of individual trees were multiplied by three and the processed alignments were translated to amino acid using EMBOSS version 6.6.0 (Rice, 2000). A custom script was used to pass each translated alignment together with the fixed species tree and individual calculated branch lengths to CodeML with RateAncestor parameter set to 1 to predict ancestral amino acid sequences for each internal node (https://github.com/ptranvan/mites2codeml). Transitions from ancestral to replacement amino acids were scored for changes in hydrophobicity according to a hydrophobicity scoring matrix (Riek *et al.*, 1995). To infer the ancestral states of amino acids for mitochondrial genes, mitochondrial genes of *T. urticae* were downloaded from NCBI and added to the alignments as outgroup (Accession: NC_010526; Van Leeuwen *et al.*, 2008). Alignment and alignment processing was done as described above for nuclear orthologous genes. To infer the ancestral states of amino acids for the large taxon set, sequences of *ef1α* and *hsp82* of the basal oribatid mite *Palaeacarus hystricinus* (Palaeosomata) were downloaded from NCBI (Accession: EF203793 and DQ090809, respectively) and added to the alignments as outgroup. Alignment and alignment processing was done as described above for nuclear orthologous genes. Hydrophobicity transitions were compared between modes of reproduction using generalised linear mixed models implemented in the R package lme4 (Bates *et al.*, 2016) with random effect of gene nested in species while correcting for overdispersion and fitted to Poisson distribution.

Analysis of CUB

Bias in codon usage was calculated with the CDC metric in Composition Analysis Toolkit version 1.3 (Zhang *et al.*, 2012). CDC estimates expected codon usage from observed positional GC and purine contents and calculates the deviation from observed codon usage using a cosine distance matrix, ranging from 0 (no deviation; no detectable, i.e. 'relaxed' selection on codon usage) to 1 (maximum deviation; effective selection on codon usage). To analyse CUB of nuclear and mitochondrial orthologous genes, we calculated CDC from processed alignments and compared per gene between reproductive modes using a permutation ANOVA similar to d*S* and d*N* comparisons (see above).

Annotation and enrichment analysis

To analyse enrichment of GO terms in genes that had significantly lower d*N*/d*S* ratios either at asexual as compared to sexual branches or at sexual as compared to asexual branches, the 3545 orthologous genes under purifying selection were annotated for the species *H. rufulus* using Blast2GO (Conesa *et al.*, 2005). Automatic annotation was successful for 2587 genes, of which 67 genes yielded significantly lower d*N*/d*S* ratios at asexual as compared to sexual branches and five genes yielded significantly lower d*N*/d*S* ratios at sexual as compared to asexual branches. Based on the 2587 orthologous loci as reference set and the 67 and five genes as test sets, enrichment of GO terms was tested for the category Biological Process with Fisher's exact test in Blast2GO.

Code availability

The script for passing the input data to codeml and collecting the d*N*/d*S* estimate and hydrophobicity output is available under https://github.com/ptranvan/mites2codeml.

Data availability

The transcriptomic data generated during this study are deposited in NCBI with the BioProject PRJNA339058. Further processed data that support the study are available upon request.

Acknowledgements

We thank Christian Bluhm for help with species identification, Darren J. Parker for help with statistics, Patrick Tran Van for help with bioinformatics and Casper van der Kooi, Ken Kraaijeveld, Nicolas Galtier and Roy A. Norton for valuable comments on the manuscript. This study was supported by core funding of S.S., by DFG research fellowship BA 5800/1-1 to J.B. and Swiss SNF grant PP00P3_170627 to T.S.

Contributions

A.B. and J.B. conceived and designed the study. A.B., J.G. and J.B. collected samples. A.B. and J.G. performed wet lab work. A.B., J.G. and J.B. performed data analysis. T.S., I.S., M.M., S.S. contributed to data interpretation and analyses and A.B., J.B., I.S. wrote the paper with input from all authors.

Chapter 2

Consequences of asexuality in natural populations: insights from stick insects

Bast, J., Parker, D. J., Dumas, Z., Jalvingh, M. J., Tran Van, P., Jaron, K. S., Figuet, E., Brandt, A., Galtier, N. & Schwander, T.

Mol. Biol. Evol. **35**, 1668–1677 (2018)

Recombination is a fundamental process with significant impacts on genome evolution. Predicted consequences of the loss of recombination include a reduced effectiveness of selection, changes in the amount of neutral polymorphisms segregating in populations, and an arrest of GC-biased gene conversion. Although these consequences are empirically well documented for non-recombining genome portions, it remains largely unknown if they extend to the whole genome scale in asexual organisms. We identify the consequences of asexuality using *de novo* transcriptomes of five independently derived, obligately asexual lineages of stick insects and their sexual sister-species. We find strong evidence for higher rates of deleterious mutation accumulation, lower levels of segregating polymorphisms and arrested GC-biased gene conversion in asexuals as compared to sexuals. Taken together, our study conclusively shows that predicted consequences of genome evolution under asexuality can indeed be found in natural populations.

Introduction

The absence of recombination is a common feature in eukaryotic genomes and it can span over a few hundred bases, entire chromosomes and, as in the case of asexual organisms, over complete genomes. The reasons for an absence of recombination can vary. For example, centromeric regions do not recombine because of their mechanistic role as the attachment points for kinetochores during meiosis (Talbert & Henikoff, 2010). Recombination can be reduced in genome portions with nearby transposable elements as a consequence of heterochromatin formation through local epigenetic silencing (Castel & Martienssen, 2013), or because of extensive structural differences between homologous chromosomes, a trait notably characterizing many supergenes and sex chromosomes (Bachtrog, 2013; Schwander *et al.*, 2014). Finally, meiotic recombination is completely lacking in asexual organisms because of clonal reproduction.

Theory predicts two major consequences for selection on genome portions and complete genomes that lack recombination. First, selection is less effective, because physical linkage among loci hinders selection's ability to act upon loci independently (Muller, 1964; Hill & Robertson, 1966; Felsenstein, 1974; Keightley & Otto, 2006). This should translate into decreased rates of adaptation and increased accumulation of mildly deleterious mutations, including point mutations and repetitive elements. A related second consequence is that population levels of polymorphisms change (Hill & Robertson, 1966; Charlesworth *et al.*, 1993). Whether the lack of recombination increases or decreases polymorphism depends on the relative importance of multiple counteracting factors, and especially the level of heterozygosity in the non-recombining region. A classical prediction is that heterozygosity (and thus polymorphism) increases over time in the absence of recombination, as allelic sequences diverge independently of each other (Birky, 1996). Polymorphism levels can therefore be higher in non-recombining than recombining genomes and

genome portions (Balloux *et al.*, 2003). By contrast, mitotic gene conversion and selection at linked sites (background selection) lead to a loss of heterozygosity and diversity in the absence of recombination, which can translate to lower polymorphism in recombining than non-recombining genome portions (Hill & Robertson, 1966; Charlesworth *et al.*, 1993).

In addition to selection-driven consequences of linkage, mechanistic effects that are independent of selection or drift can differ between recombining and non-recombining genome portions. For example, recombination *per se* is thought to be mutagenic, which can affect polymorphism levels and the rate of divergence (Webster & Hurst, 2012). Finally, base compositions are expected to differ between recombining and non-recombining genomes and genome portions because recombination-associated gene conversion will tend to increase GC content. Gene conversion is a molecular mechanism associated with DNA double-strand breaks and mismatch repair during recombination, where one allele is replaced by its homolog (Van den Bosch *et al.*, 2002). In many species, gene conversion is non-random with respect to the base composition of an allele; the bases A and T are more likely to be converted into G and C than the reverse (Marais, 2003; Mugal *et al.*, 2015). In the absence of meiotic recombination, such GC-biased gene conversion (gBGC) should occur at greatly reduced rates, and the base composition of neutral sequences should shift towards the mutational equilibrium (Galtier *et al.*, 2018). There is accumulating empirical evidence for several of the above mentioned consequences in non-recombining genome portions of sexual organisms. For instance, studies of Y-chromosomes revealed lower levels of neutral polymorphism, increased rates of deleterious mutation accumulation and lower rates of adaptive evolution compared to recombining parts of genomes (Bachtrog & Charlesworth, 2002; Skaletsky *et al.*, 2003; Bachtrog, 2013). Furthermore, as a result of reduced gBGC, non-recombining genome portions typically feature low GC content (Spencer *et al.*, 2006; Duret & Galtier, 2009; Pessia *et al.*, 2012; Glémin *et al.*, 2015).

In contrast to these non-recombining, ‘asexual’ genome portions in sexual genomes, the consequences of a genome-wide lack of recombination in natural populations of asexual species remain less clear. In particular, the effects of asexuality on segregating polymorphisms and gBGC remain largely untested. The reduced effectiveness of selection in asexuals was investigated in several studies via analyses of deleterious mutation accumulation, but with mixed results. Most studies based on few individual genes indeed found increased rates of mutation accumulation in asexual as compared to sexual lineages, but others reported no differences (reviewed in Hartfield, 2016). Similarly, the three available genome-wide studies all reported different patterns, including one case where selection was more effective in asexual than sexual species (Hollister *et al.*, 2015; Ament-Velásquez *et al.*, 2016; Brandt *et al.*, 2017). Given these conflicting results, additional whole-genome studies are warranted.

Timema stick insects are an ideal group for identifying signatures of asexual genome evolution. Because asexuality is a lineage level-trait, independently derived asexuals are required to disentangle consequences of asexuality from species-specific factors. Seven such independently derived asexuals are known in *Timema*, with an ecologically equivalent sexual sister species at hand for comparison (Fig. 1). Sexual and asexual sister species are morphologically alike, live and feed on similar host plants and have a generation time of one year (Law & Crespi, 2002). The asexual lineages are obligate apomicts, diploid, not of hybrid origin and differ in age, ranging from 0.2 to 2 my (Schwander & Crespi, 2009; Schwander *et al.*, 2011).

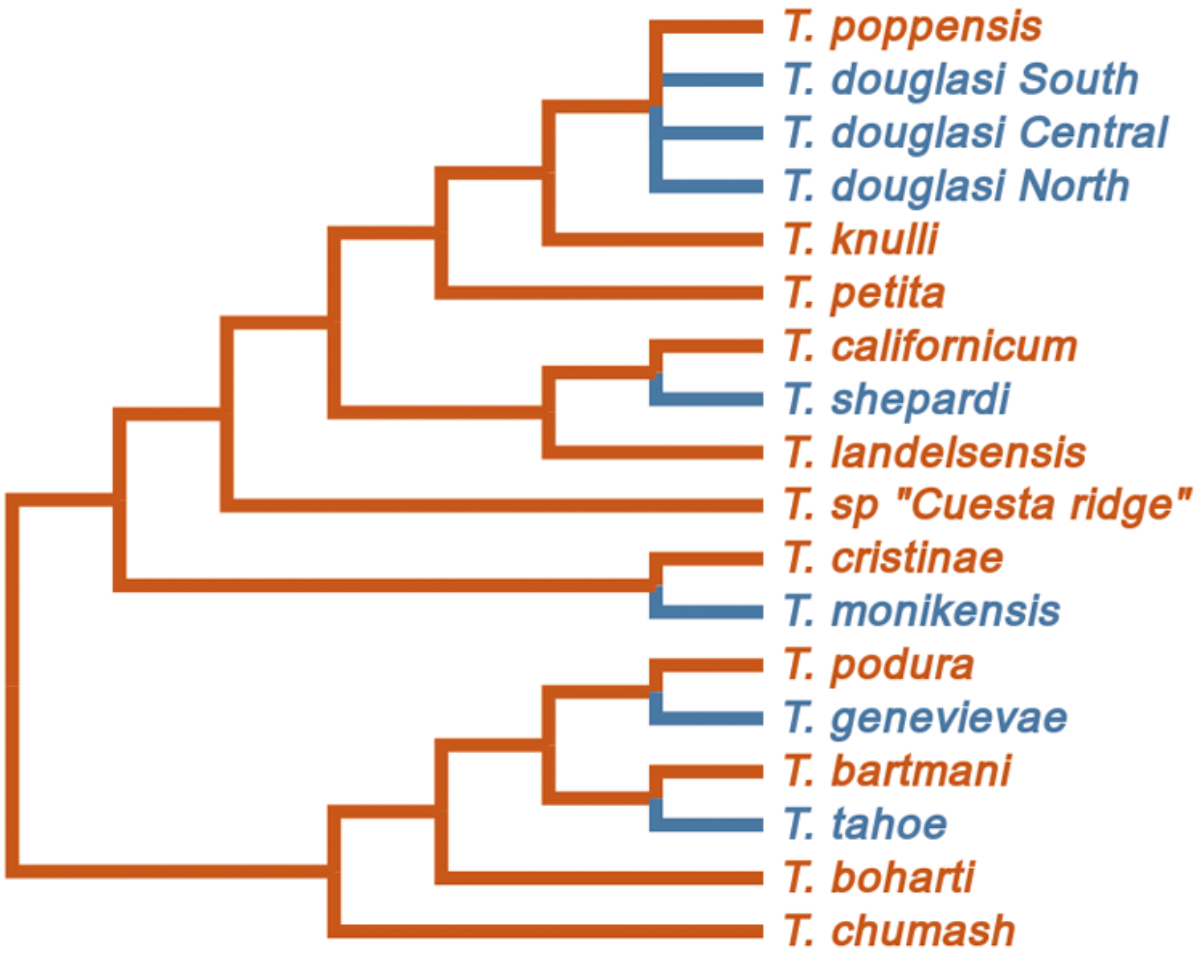


Figure 1: Phylogeny of *Timema* species (sexual species phylogeny redrawn from Riesch *et al.*, 2017); asexual lineages added from Schwander *et al.* (2011). Sexual species are colored red, asexual species blue. For this study, five sex-asex species pairs were used (for *T. douglasi* only the southern lineage).

Here, we tested whether the lack of recombination in *Timema* asexuals is associated with I) changes in the strength of gBGC; II) changes in the amount of genetic variation within populations; and III) increased rates of deleterious mutation accumulation and more segregating deleterious variants. We generated *de novo* transcriptomes for ten *Timema* species (see Materials and Methods for details), including transcriptomes of five asexual lineages (*T. tahoe, T. monikensis, T. douglasi 'South', T. shepardi* and *T. genevievae*), and their sexual sister-species (*T. bartmani, T. cristinae, T. poppensis, T. californicum and T. podura*).

Results

De novo transcriptome assembly and ortholog prediction

Depending on the species, the quality-filtered assemblies contained between 31,747 and 45,655 transcripts with a N_{50} between 1,301 and 1,836 bp. Open reading frames (ORFs) were identified in 7,329 - 10,436 of the transcripts, and these ORFs were used to determine orthologs shared between the two species of a pair (between 5329 and 5908 per pair) as well as orthologs shared across all 10 species (3010 orthologs; see Materials and Methods for details, Supplementary Data 1 and Supplementary Table 1 and 2).

The ages of the asexual lineages were previously estimated based on intra- and interspecific divergence at a single mtDNA locus (Schwander *et al.*, 2011). We used orthologs within each species-pair to update the ranking of asexuals from youngest to oldest, with interspecific divergence as a proxy for lineage age (Fig. 2).

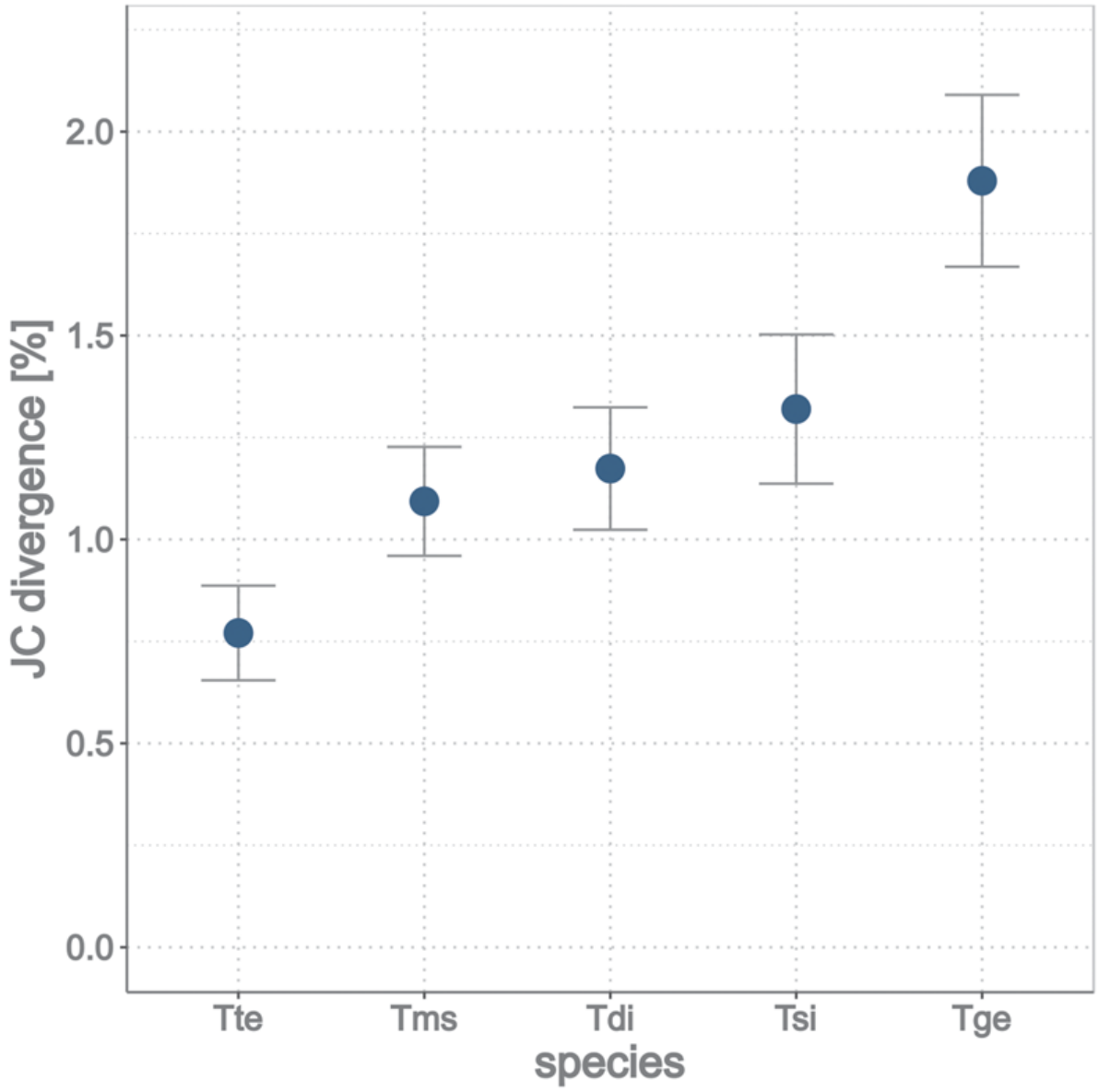


Figure 2: Asexual lineages ranked from youngest to oldest as estimated from Jukes-Cantor corrected divergence between sexual-asexual sister-species (depending on the species pair, using 5329 to 5908 pairwise orthologs). Tte: *Timema tahoe*, Tms: *T. monikensis*, Tdi: *T. douglasi*, Tsi: *T. shepardi*, Tge: *T. genevievae*. The oldest asexual lineage *T. genevievae*, was previously estimated to be approximately 1.5 my old (Schwander *et al.*, 2011).

To evaluate the robustness of the 3010 ten-species orthologs, we screened each transcriptome for variants of these orthologs. Variants are transcripts with the same ORF, but that contain sufficiently many SNPs, deletions or insertions of variable length in any part of the ORF to result

in separately assembled transcripts in the *de novo* assembly (our ortholog detection pipeline retained the longest ortholog if multiple options are available, see Materials and Methods for details). Surprisingly, the proportion of orthologs with variants was significantly higher for the asexual than sexual *Timema* species (mean asex: 0.076, mean sex: 0.013; paired t-test, t = 7.29, df = 4, $P < 0.001$; Supplementary Data 2). Given this systematic difference between sexual and asexual species, we evaluated whether ortholog variants represent transcripts from multiple gene copies by mapping genomic reads from single individuals generated in a different study (SRR5248877 - SRR5248936) to our 3010 ten-species orthologs (see Supplementary Material). If ortholog variants were generated by gene duplications, the coverage for orthologs with variants would be higher than the coverage for orthologs without variants (i.e., an ortholog with two copies in the genome would have twice the coverage relative to an ortholog with single copy). Our analyses revealed that ortholog variants do not stem from gene duplicates, as the coverage for orthologs was the same, independently of whether an ortholog had a variant or not in a given transcriptome (Supplementary Fig. 1). By consequence, ortholog variants are most likely generated by transcriptional noise and error, or perhaps by different alleles (see discussion for details). Indeed half of the variants (725 out of 1353) differed in length from each other (up to 7-fold, mean length difference: 13%) and 27% of the variants further featured SNP divergences (mean pairwise distance 1.55%). While the enrichment of orthologs with variants in asexuals is interesting *per se*, it is important to note that these variants did not affect our subsequent analyses and results. This is because all polymorphism, pN/pS and dN/dS analyses were based on one 'best' ortholog variant (i.e. longest) and estimates of segregating polymorphism were conservative, as the presence of variants could potentially inflate polymorphisms in asexuals compared to sexuals but sexuals had more polymorphisms than asexuals (see below).

Reduced gBGC in asexuals

Because meiotic recombination has stopped in asexuals, GC-biased gene conversion should be strongly reduced if the process was at work in *Timema*. We tested for evidence of this mechanistic consequence of asexuality via various approaches. First, we analyzed GC content at third codon positions (GC3 content), which are most likely neutral. In agreement with the hypothesis that gBGC favors the increase of GC content and generates variation among genes experiencing different recombination rates, sexual species within each pair had a higher GC3 content (Wilcoxon signed-rank test, $P < 0.01$ in all pairs) and higher variance among genes than asexual species (Fig. 3). We also compared per gene GC3 in each of the sexual-asexual species pairs. We calculated the proportion of genes in which GC3 was higher in the sexual than in the asexual species, and found that in all five pairs this proportion was above 0.5. Interestingly, the proportion was maximal (0.62) in the oldest pair (*T. podura* vs. *T. genevieve*), minimal (0.51) in the youngest pair (*T. bartmani* vs. *T. tahoe*), and intermediate in the other three pairs (0.52, 0.54, 0.55).

Then we investigated the substitution processes that shaped base composition in *Timema* using maximum likelihood approaches. We estimated branch-specific equilibrium GC3 using the non-homogeneous maximum likelihood method developed by Galtier & Gouy (1998). Similar to the results from the previous approach, the asexual species were characterized by a lower equilibrium GC3 than their sexual counterparts (likelihood ratio test, P < 2.2e-16 between stationary and nonstationary per-branch model of sequence evolution), meaning that current substitution processes drive GC3 towards lower values in asexuals than in sexuals. We mapped synonymous substitutions along the *Timema* phylogeny and estimated the per-lineage number of weak-to-strong (A/T->G/C), strong-to-weak (G/C->A/T) and GC-conservative (A<->T, G<->C) substitutions. As expected under gBGC arrest, asexuals always harbored a higher proportion of strong-to-weak substitutions and a lower proportion of weak-to-strong substitutions than to

sexuals (Supplementary Data 3). Finally, we calculated GC-content in UTR regions and correlated it to GC3. We obtained a strong and significantly positive correlation between GC_UTR and GC3 in all 10 species (r between 0.37 and 0.44, $P < 0.001$ in all cases), demonstrating that GC3 is governed by evolutionary forces independent of the process of translation - most likely gBGC. In combination, the results from these different approaches strongly support the hypothesis of a reduced impact of gBGC in *Timema* asexuals.

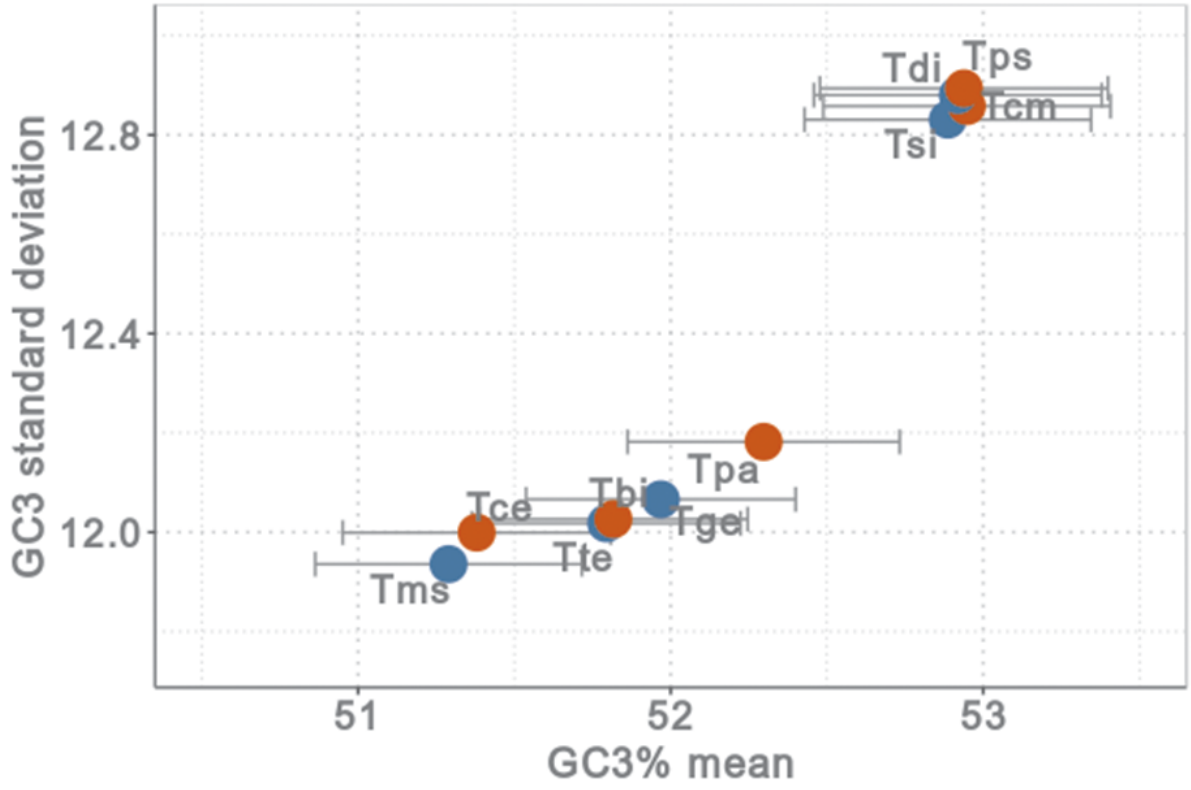


Figure 3: Mean GC3 content (with 95% CI) and GC3 standard deviation of the 3010 ten-species orthologs of sexual (in red) and asexual (in blue) *Timema* species. Tbi: *Timema bartmani*, Tte: *T. tahoe*, Tce: *T. cristinae*, Tms: *T. monikensis*, Tps: *T. poppensis*, Tdi: *T. douglasi*, Tcm: *T. californicum*, Tsi: *T. shepardi*, Tpa: *T. podura*, Tge: *T. genevievae*.

Reduced levels of polymorphisms in asexuals

Polymorphism levels at individual loci can be higher or lower in asexual than sexual populations, depending on the amount of heterozygosity in asexuals and the number of different clones

(genotype diversity). To quantify the amount of segregating polymorphisms in sexuals and asexuals, we identified SNPs in transcripts by mapping the read data derived from three pooled individuals (all collected from the same population) back to the 3010 ten-species orthologs that include only one 'best' variant per ortholog (see Materials and Methods for details). Since alternative variants of each ortholog were not included, reads corresponding to these variants mapped to the 'best' ortholog. Specifically, for 63% of the orthologs, the sum of reads corresponding to the different variants was exactly identical to the number of reads mapping to the best ortholog; for 93%, the read numbers differed by 5% or less (Supplementary Fig. 2). As expected given the larger number of variants in asexuals than sexuals, mapping quality was somewhat lower in asexuals than sexuals, however not significantly so (paired t-test, $t = 1.3$, $df = 7.9$, $P = 0.215$). We found that a higher proportion of transcripts contained SNPs in sexuals than asexuals (paired t-test, $t = 13.46$, $df = 4$, $P < 0.001$; Fig. 4a), in spite of the enrichment for ortholog variants in asexual transcriptomes. Similarly, among these transcripts that contained SNPs, the proportion of variable sites per transcript was higher for sexuals than asexuals (paired t-test, $t = 3.850$, $df = 4$, $P = 0.018$; Fig. 4b).

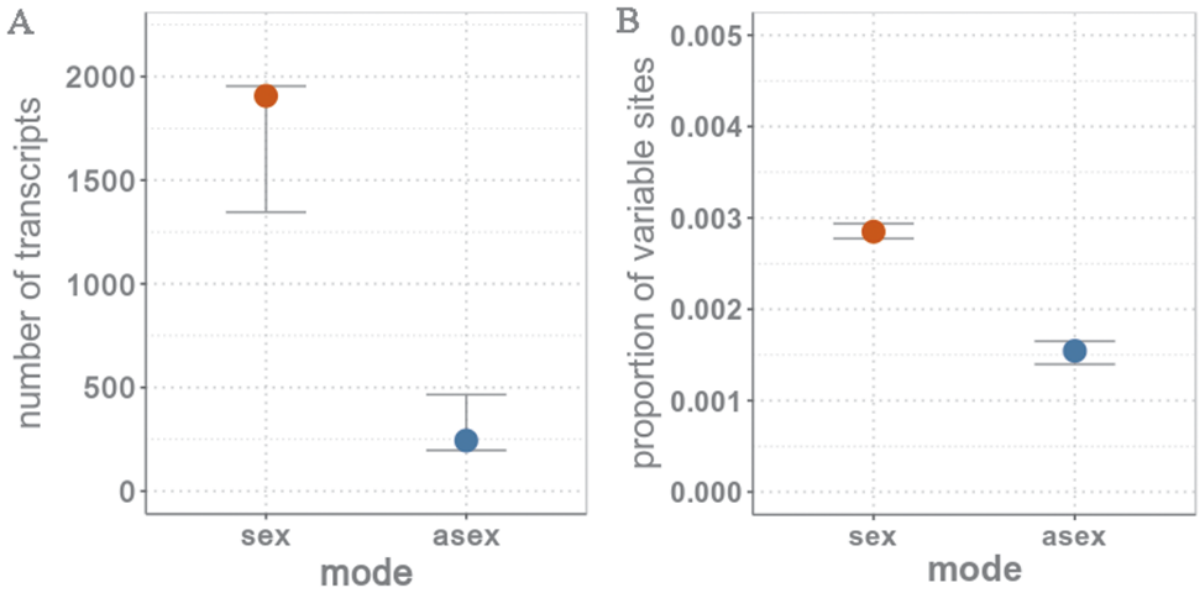


Figure 4: medians (and 95% CI with 5000 bootstrap replicates) of a) number of transcripts containing SNPs and b) proportion of variable sites per transcript for sexual (red) and asexual (blue) *Timema* species.

Purifying selection is less effective in asexuals.

To test the prediction that selection acts less effectively on asexuals than sexuals, we focused on mutations that are expected to affect protein function and are likely deleterious. This should be the case for non-synonymous mutations in the 3010 orthologs shared among the ten species. Indeed, since the *Timema* species diverged approximately 30 my ago, the vast majority of these orthologs should be under purifying selection (or they would be too diverged to be identified as orthologs). Thus, we quantified whether more non-synonymous mutations segregate in asexual than sexual populations and whether these mutations accumulate faster in asexual than sexual species.

We identified non-synonymous and synonymous SNPs per transcript by mapping the read data (derived from three pooled individuals) to the 3010 ten-species orthologs (see Materials and Methods). As expected, the ratio of non-synonymous to synonymous polymorphisms (pN/pS) was larger in asexual than sexual populations (paired t-test, t = -3.02, df = 4, P = 0.039; Fig. 5).

This indicates that deleterious variants are rapidly removed by selection in sexual populations but persist longer in asexual populations.

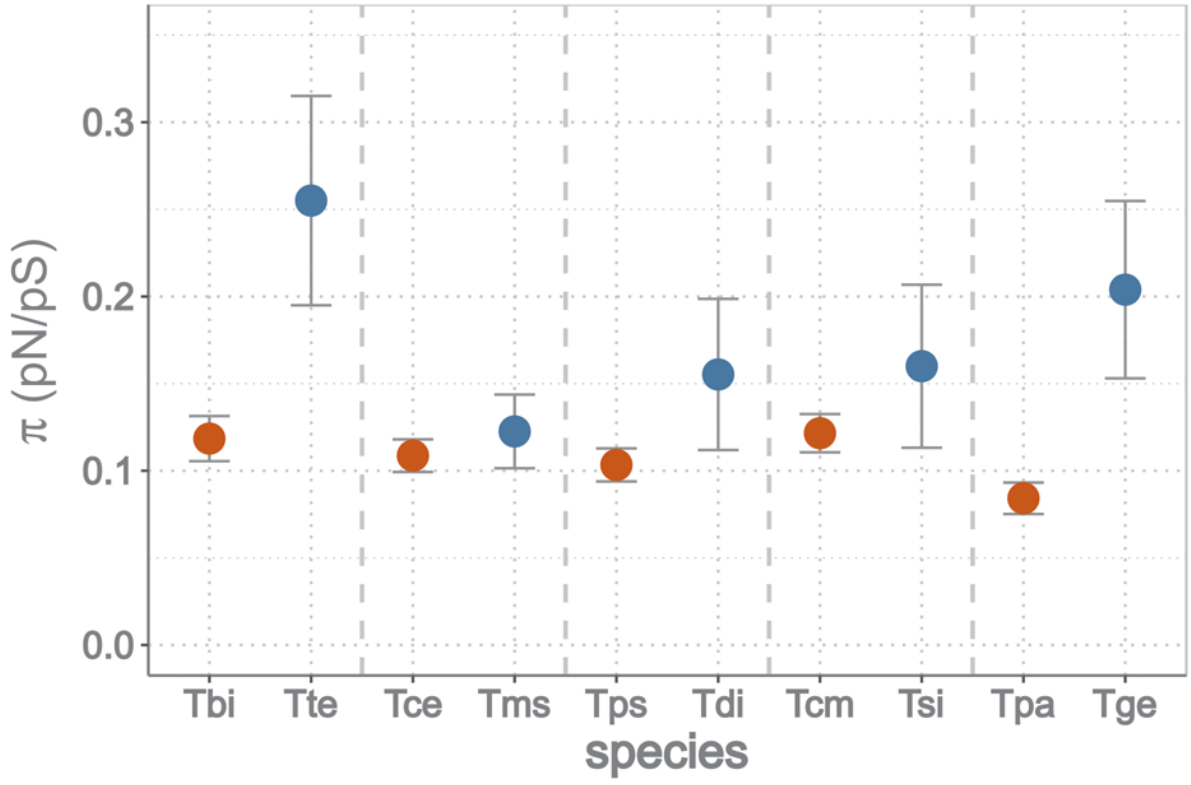


Figure 5: Means (and 95% CI) of non-synonymous segregating polymorphisms (pN/pS) of sexual (in red) and asexual (in blue) *Timema* species. Tbi: *Timema bartmani*, Tte: *T. tahoe*, Tce: *T. cristinae*, Tms: *T. monikensis*, Tps: *T. poppensis*, Tdi: *T. douglasi*, Tcm: *T. californicum*, Tsi: *T. shepardi*, Tpa: *T. podura*, Tge: *T. genevievae*. Species pairs are ranked by age of the asexual lineage (youngest to oldest).

To elucidate whether deleterious variants also become fixed faster in asexual than sexual *Timema* species, we estimated the ratio of non-synonymous to synonymous divergence (ω=dN/dS) along the branches of the *Timema* phylogeny, using the maximum likelihood methods implemented in PAML (Yang, 2007; see Materials and Methods for details). Again, for transcripts under purifying selection (i.e., the majority of the ten-species orthologs), most non-synonymous changes are likely mildly deleterious (Li *et al.*, 1985) and a higher ω ratio therefore indicates a

higher rate of non-synonymous mutation accumulation and less effective purifying selection. As the *Timema* system provides evolutionary replicates, we allowed ω to vary independently for each gene and branch on the phylogeny (i.e., applying a 'free model'), to take these replicates into account. Out of the 3010 ten-species orthologs, we removed 200 without variation and 323 with $\omega \geq 1$ or without synonymous substitutions, which left 2487 transcripts for comparisons. Consistent with the predictions of less effective purifying selection in asexual species, we found that asexuals accumulated non-synonymous mutations at a higher rate than sexuals (permutation ANOVA: gene effect, species pair effect, reproductive mode effect, and interaction between pair and mode all $P < 0.001$; Fig 6). We additionally ran a simpler three-ratio model that estimates one dN/dS ratio for internal branches and one for terminal branches of each reproductive mode, which corroborated these findings (see Supplementary Material).

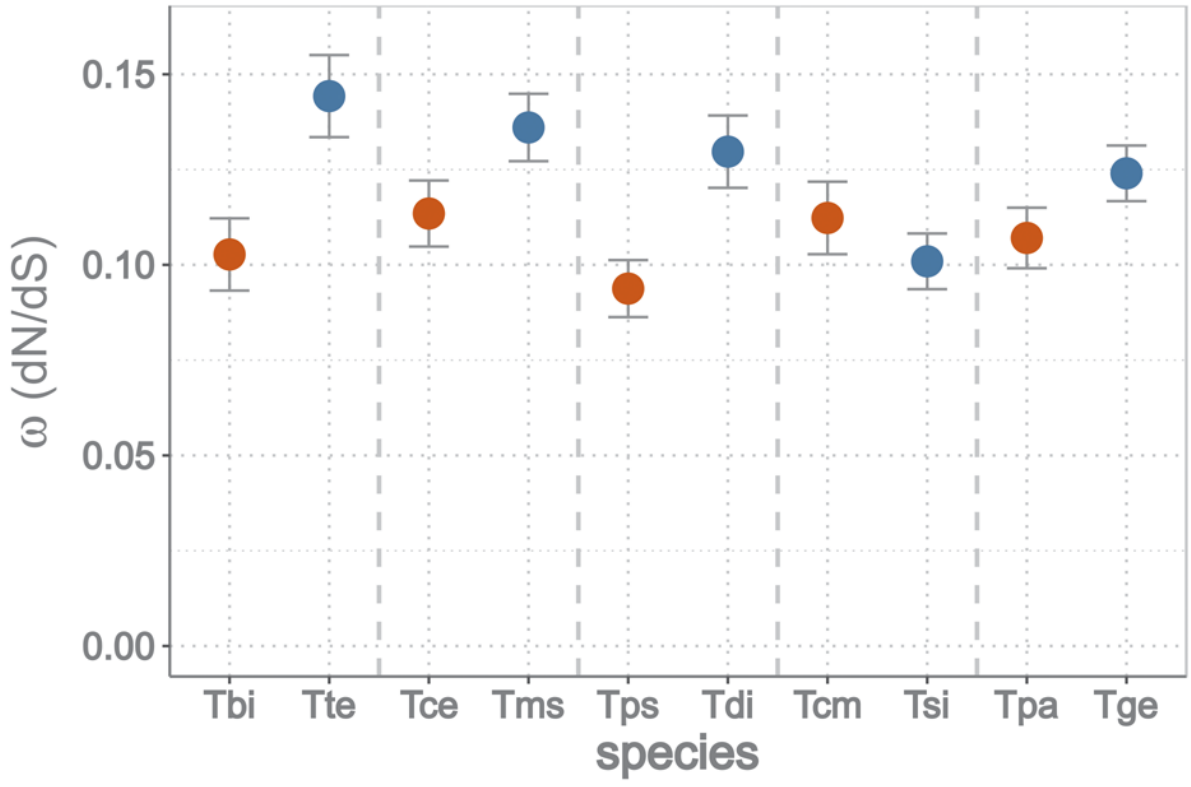


Figure 6: Means (and 95% CI) of ω=dN/dS ratios from 2487 orthologs in sexual (in red) and asexual (in blue) *Timema* species. Tbi: *Timema bartmani*, Tte: *T. tahoe*, Tce: *T. cristinae*, Tms: *T. monikensis*, Tps: *T. poppensis*, Tdi: *T. douglasi*, Tcm: *T. californicum*, Tsi: *T. shepardi*, Tpa: *T. podura*, Tge: *T. genevievae*. Species pairs are ranked by age of the asexual lineage (youngest to oldest).

Increased substitution rates at synonymous sites in asexuals

Synonymous substitutions (dS) are commonly assumed to be neutral, and are used to scale non-synonymous substitutions in the above-mentioned analyses. However, purifying selection can also act on synonymous sites, because different synonymous codons influence for example the speed and accuracy of translation (Hershberg & Petrov, 2008). As a consequence, the frequencies by which different codons are used can differ (codon usage bias). We compared the effective number of codons (Enc), a quantitative estimate for the level of codon usage bias, as well as synonymous substitution rates (dS) between sexual and asexual *Timema* pairs. *Timema* featured Enc values that indicate weaker codon usage bias than in animals with large population

sizes (e.g. nematodes), but the bias level was similar for species with different reproductive modes (gene effect $P < 0.001$, species pair effect $P < 0.001$, reproductive mode $P = 0.241$, and interaction between pair and mode all $P = 0.928$; permutation ANOVA; Supplementary Fig. 3). Estimates of dS were systematically elevated in asexuals (gene effect, species pair effect, reproductive mode, and interaction between pair and mode all $P < 0.001$; permutation ANOVA; Supplementary Fig. 6).

We tested if the difference in dS between sexuals and asexuals stems from different effectiveness of selection on codon usage bias using the Codon Deviation Coefficient (CDC) as a metric (Zhang *et al.*, 2012). This metric calculates the deviation from expected codon usage bias and accounts for per-gene background nucleotide composition, thus allowing for cross-species comparisons. A lower CDC value would indicate less effective selection on codon usage bias (see Materials and Methods), but we did not detect any differences in CDC value between reproductive modes (gene effect $P < 0.001$, species pair effect $P < 0.001$, reproductive mode effect $P = 0.595$, interaction between pair and mode $P=0.991$; permutation ANOVA; Fig. 7). Thus, elevated substitution rates at synonymous sites in asexuals are unlikely to stem from less effective selection on codon usage bias in asexuals than sexuals.

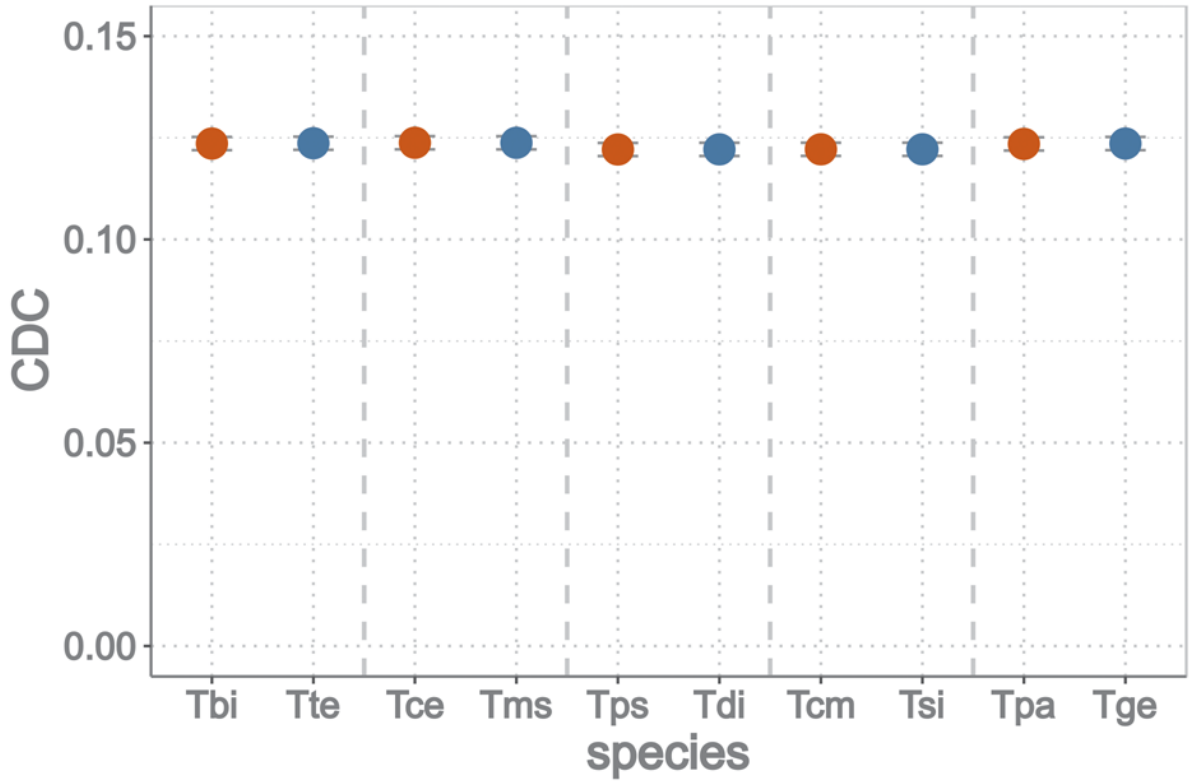


Figure 7: Means (and 95% CI) of the codon deviation coefficient (CDC) of sexual (in red) and asexual (in blue) *Timema* species. Tbi: *Timema bartmani*, Tte: *T. tahoe*, Tce: *T. cristinae*, Tms: *T. monikensis*, Tps: *T. poppensis*, Tdi: *T. douglasi*, Tcm: *T. californicum*, Tsi: *T. shepardi*, Tpa: *T. podura*, Tge: *T. genevievae*. Species pairs are ranked by age of the asexual lineage (youngest to oldest).

Evaluation of the consequences of asexuality over time

Given the age variation of the asexual *Timema* species (Fig. 2), possible consequences of asexuality can be studied over a range of recent to old asexual organisms. For all of the measured effects of asexuality, lineage age should *a priori* only impact the consequences of arrested gBGC, because the return to equilibrium base composition should happen gradually over time. Indeed, the effect of arrested gBGC on GC composition was stronger for older than younger asexual *Timema* species (Pearson's product-moment correlation; t = 3.29, cor 0.88, *P* = 0.046, Supplementary Fig. 4). There was no age effect for dN/dS (Pearson's product-moment

correlation; t = -1.03, cor -0.51, *P* = 0.686), pN/pS (Pearson's product-moment correlation; t = -1.23, cor -0.13, *P* = 0.829) and dN estimates (Supplementary Fig. 5), but surprisingly, dS increased with age of asexuality (Pearson's product-moment correlation; t = 5.83, cor 0.96, *P* = 0.010, Supplementary Fig. 6). However, any correlation (or absence thereof) should be interpreted carefully, as there are only five data points and correlates might be driven by only one old 'outlier' asexual (*T. genevievae*).

Discussion

The absence of recombination is predicted to be associated with reduced GC content (because GC-biased gene conversion does not occur), reduced effectiveness of selection on short and long evolutionary timescales, and affect levels of neutral genetic variation segregating in populations (Muller, 1964; Hill & Robertson, 1966; Felsenstein, 1974; Marais, 2003; Keightley & Otto, 2006). Using multiple, independently evolved asexual lineages, we provide for the first time empirical evidence for all of these predicted consequences, in full agreement with theory.

GC-biased gene conversion (gBGC) influences the base composition in sexual *Timema* stick insects, but is unlikely to occur (or at least is strongly reduced) in asexuals. Consistent with theory, the effect of arrested gBGC on base composition is stronger in older asexual *Timema* species compared to their sister species (Supplementary Fig. 4). These findings represent the first documented case of arrested gBGC in asexual eukaryotes. Although there is extensive variation in GC content and in the strength of gBGC among sexual *Timema* species, we consistently detect lower levels of gBGC in the asexual sister lineages. Given this lineage-specific variation, sister-pair comparisons are essential to detect arrested gBGC in asexual species. This might be why all evidence for arrested gBGC in the absence of recombination thus far stems from

non-recombining parts of genomes and selfers, where appropriate comparisons are available (Glémin *et al.*, 2006; Spencer *et al.*, 2006; Duret & Galtier, 2009; Glémin & Galtier, 2012).

Asexual stick insects exhibit lower levels of segregating genetic variation within populations as compared to sexuals. This finding suggests that asexual *Timema* species are more homozygous than their sexual counterparts, as intra-individual heterozygosity should otherwise generate high genetic variation at individual loci, at least in the old asexual species such as *T. genevievae*. Low genetic diversity in asexuals is also expected because of background selection. Although asexual *Timema* populations are far from being genetically uniform, there are eight-fold more transcripts containing polymorphisms in sexual populations and within these transcripts, variants are twice as frequent (Fig. 4). Thus far, comparative studies investigating segregating polymorphisms in sexual and asexual populations are complicated by confounding effects like hybridization events and/or found no differences of the amount of genetic variation between reproductive modes, likely because of young asexual lineage ages (Hollister *et al.*, 2015; Ament-Velásquez *et al.*, 2016). Our findings of reduced levels of standing genetic variation in asexual compared to sexual *Timema* sister species is direct evidence that reproductive mode can strongly influence genetic diversity in natural populations.

Purifying selection is less effective in asexual than sexual *Timema*, resulting in more deleterious variants segregating in populations and a higher rate of deleterious mutation accumulation in the long term. The ratio of non-synonymous to synonymous segregating variants is approximately two times higher in asexual than sexual populations (range 1.1 to 2.4, depending on the species pair). In spite of these differences, there is currently no evidence that the average fitness of asexual *Timema* females is lower than the fitness of sexual females, but fitness differences may require more time to build up and become measurable. Indeed some of the deleterious variants become fixed over time as a consequence of drift, resulting in a faster rate of deleterious mutation

accumulation in asexual than sexual *Timema* species, with ω (dN/dS) on average being 1.2 times higher in asexuals than sexuals (Fig. 6).

It is important to note however that ω ratios are highly variable among genes (variance = 0.043, range 0.001 - 0.998), with sometimes opposite effects ($\omega_{sex} > \omega_{asex}$) to the general pattern. Thus, given this variance, inferring overall genomic patterns from studies that are based on small gene numbers might result in arbitrary and wrong conclusions. Indeed most previous studies on mutation accumulation in asexual metazoans, including a previous study in *Timema*, are based on few genes (Johnson & Howard, 2007; Neiman *et al.*, 2010; Henry *et al.*, 2012; Ollivier *et al.*, 2012; Hartfield, 2016). To date, only three studies exist that use genome-scale analyses to address mutation accumulation in asexuals, with mixed results (Hollister *et al.*, 2015; Ament-Velásquez *et al.*, 2016; Brandt *et al.*, 2017). Investigated asexual lineages were either very young, and no difference in the effectiveness of purifying selection between reproductive modes was found (Ament-Velásquez *et al.*, 2016). Or very old (tens of millions of years) and accumulating deleterious mutations at lower rates than sexual species, opposite to theoretical predictions (Brandt *et al.*, 2017). In addition to the present study, there is thus far only one study in selfing (but functionally asexual) plants that converged with theoretical predictions of more effective purifying selection in sexual taxa compared to asexuals (Hollister *et al.*, 2015).

Our results also suggest that the reduced effectiveness of selection in asexuals may result in an increase of transcriptional errors (or transcriptional noise). Indeed, an interesting and unexpected finding is that asexual species harbor more variants of orthologous transcripts than sexuals. Analyses using genomic data suggest that these variants do not stem from duplicated genes in the genomes. Moreover, the majority of variants are unlikely to represent different haplotypes, given that they vary considerably in the length of the ORF and the extremely low levels of segregating polymorphism in populations of *Timema* asexuals. Additional studies are required to corroborate

our interpretation of higher transcriptional error rates, reduced stability or more frequent post-transcriptional alteration of transcripts in asexuals.

There is evidence that selection can also act on synonymous sites (Lawrie *et al.*, 2013), for example through altering the frequencies by which different codons are used (Duret, 2002; Hershberg & Petrov, 2008). Interestingly, rates of synonymous substitutions are increased in asexual *Timema*, compared to sexuals. The difference between sexuals and asexuals further increases with the age of asexual lineages. This pattern could be indicative of less effective selection on codon usage bias in asexual than sexual *Timema*. However, while there is weak codon usage bias overall, we could not detect differences in the strength of the bias between sexual and asexual species. This is possibly because evolution of codon bias might be too slow such that differences are too subtle to be detectable in the relatively 'young' *Timema* species. Another mechanism that could potentially influence rates of synonymous substitutions is gBGC. gBGC arrest in asexuals results in a sudden change in equilibrium GC3, which under certain conditions is expected to increase substitution rates (Bolívar *et al.*, 2016). Moreover, increased synonymous substitution rates in asexuals could also be caused by increased mutation rates.

In conclusion, this is the first genome scale study on asexual evolution that documents consequences for gBGC, polymorphism levels and effectiveness of purifying selection. This study is particularly robust, because the *Timema* system allows us to utilize evolutionary replicates of independent sexual-asexual sister pairs with different asexual lineage ages. It remains unknown and a challenge for future studies to determine if the accumulation of deleterious mutations may eventually result in the extinction of asexual *Timema* lineages and thereby contribute to the long-term maintenance of sex in this system.

Materials and Methods

Taxon sampling and sequencing

Timema individuals from ten different species (five sexual and five asexual) were collected in California, USA in spring 2014. For details on the sampling locations see Supplementary Table 3. Prior to RNA extraction, animals were fed with artificial medium for two days to avoid contamination with gut content and then frozen at -80°C. Total RNA was extracted from whole bodies of a pool of three individuals from the same species and location. This was done by first freezing the individuals in liquid nitrogen, followed by addition of Trizol (Life Technologies) and mechanical bead crushing (Sigmund Lindner). The homogenized tissue was then treated with Chloroform and Ethanol and the aqueous layer transferred to RNeasy MinElute Columns (Qiagen). Following the RNeasy Mini Kit protocol, potential DNA in the sample was digested, RNA eluted into water and stored at -80°C. RNA quantity and quality was measured using NanoDrop (Thermo Scientific) and Bioanalyzer (Agilent). RNA extracts were pooled and fragmented to 120 nt for strand-specific library preparation. Single-end sequencing with a read length of 100 bp was performed on a HiSeq2000 platform at the CIG (Centre of Integrative Genomics, Lausanne, Switzerland). CutAdapt was used to remove adapter sequences from the raw reads (Martin, 2011). Reads longer than 80 bp were then quality trimmed using trimmomatic v 0.36 (Bolger *et al.*, 2014; first clipping leading or trailing bases with a phred score of <10 from the read, before using a sliding window from the 5’ end to clip the read if 4 consecutive bases had an average phred score of <20). Following quality trimming, any reads <80 bp in length were discarded. For assembly, reads found to contain adapter sequence were not used, but for mapping all trimmed reads were used.

Assembly and assembly cleaning

All available trimmed reads per species were pooled and used as input for assembly with Trinity v2.2.0 with the addition of the following option: --min_kmer_cov 2 (Grabherr *et al.*, 2011). Assembly often results in an unrealistically high number of transcripts. To remove likely erroneous contigs, we applied a minimum expression filter following Moghadam *et al.* (2013) and Harrison *et al.* (2015). We mapped the pooled reads for each sample separately back to the assembly with RSEM v1.2.20 using bowtie2 v2.2.4 (Li & Dewey, 2011; Langmead & Salzberg, 2012). Ambiguously-mapping reads were assigned to the most likely transcript by RSEM. We discarded contigs shorter than 300 bp and with low coverage (reads per kilobase per million mapped, RPKM ≤ 2). After filtering, all assemblies had a comparable number of isotigs and N_{50} (Supplementary Table 1). For subsequent analyses in this study, the longest isoform per gene/graph was selected. Open reading frames (ORFs) were identified from the filtered transcriptome set using transdecoder v2.1 running TransDecoder.LongOrfs and TransDecoder.Predict (with strand-specific option). In cases where multiple ORFs were predicted for a single transcript the longest ORF per transcript was selected (Supplementary Table 1).

Annotation and contamination filtering

NCBI's blast client (v. 2.2.30+) was used to blast local versions of the nt (using blastN, default options except -task blastn, -max_target_seqs 10) and nr (using blastX, default options except, -max_target_seqs 10) databases (downloaded: 07/08/2016). Blast hits were first filtered so that any hit with an e-value >0.0000001 was discarded. Contigs were assigned to a Domain if >=50% of blast hits came from a domain. In the event of a tie, the taxa with the highest e-value was used as a tiebreaker. Contigs annotated as 'non-eukaryote' were discarded prior to submission to GenBank (the number of discarded transcripts ranged from 64-133 among transcriptomes).

Orthologs

To identify orthologous sequences shared by any combination of the ten species, OMA v1.0.6 (Altenhoff *et al.*, 2013) was run on the identified ORF protein sequences with default parameters and fixed species tree (Fig. 1). Further, orthologs were selected that were shared between all *Timema* species, yielding 3010 orthologous groups (with one sequence per *Timema* species in each group). To identify the variants of these orthologs from each of the ten transcriptome assemblies, a custom script, provided by Clément Train, a member of the OMA development team, was used (see Code availability).

Alignments

Alignments of the orthologs shared between sister species were generated to calculate divergence estimates. Alignments were also generated for each of the 3010 ten-species orthologs for analyses of dN/dS and CDC. For all types of alignments, orthologous sequences were aligned on the protein level using M-coffee for a consensus alignment from multiple software (clustalw2_msa muscle_msa kalign_msa mafftgins_msa t_coffee_msa) implemented in t-coffee v11.00.8 (Wallace *et al.*, 2006). In almost all cases, all methods converged to a similar alignment, as indicated by a quality score > 75. Resulting protein alignments were back-translated into nucleotide alignments using t-coffee and curated with Gblocks v0.91b (type=codons; minimum block length set to 4; Talavera & Castresana, 2007) to remove gap-rich and poorly aligned regions.

Distance estimates

Pairwise Jukes-Cantor distances of curated alignments from orthologous sequences between sister species were calculated using the distmat program from the EMBOSS suite v6.6.0.0 (Rice *et al.*, 2000).

gBGC

Ancestral GC3 was estimated for all nodes of the *Timema* tree for the 3,010 orthologs using the NHML program (Galtier & Gouy, 1998). This method makes use of a nonhomogeneous and non-stationary Markov model of DNA evolution to estimate branch-specific equilibrium GC-content in a maximum-likelihood framework (Romiguier *et al.*, 2010; Figuet *et al.*, 2014). The substitution mapping procedure was used to calculate the number of synonymous substitutions in the terminal branch of each *Timema* species, and categorized them into weak-to-strong (A/T->G/C), strong-to-weak (G/C->A/T) and GC-conservative (others; Dutheil *et al.*, 2012; Romiguier *et al.*, 2012). This consisted in first fitting the YN98 codon model, as implemented in the bppML program (Dutheil & Boussau, 2008), to each alignment of protein-coding sequences. The estimated model parameters were then used to map substitutions on each branch of the tree in the empirical Bayesian framework. The total number of substitutions in a branch for a given category was obtained by summing substitution counts across genes.

Polymorphism (SNPs)

To identify polymorphism from SNP data, we mapped reads from a pool of three females from the same species and population back to the 3010 10-species orthologs (that contained only one ‘best’ variant for each ortholog per orthologous group) with RSEM using bowtie2 with default parameters and --fragment-length-mean 200 --fragment-length-sd 100 (Li & Dewey, 2011; Langmead & Salzberg, 2012). Samtools v1.2 was used to sort the mapped reads and samtools mpileup to identify SNPs. Resulting polymorphic sites were filtered using VarScan v2.3.2 mpileup2snp with a 20-fold minimum coverage cutoff, 10% minor allele freq and a minimum average phred quality of 20 following Hollister *et al.* (2015). To check for mapping bias generated by the presence of variants for certain orthologs, mapping quality and read count was

extracted from the RSEM output table and bam files for both the ten-species ortholog set that included the best variant only and a set that included all variants.

Divergence ω (dN/dS) analyses

Curated back-translated nucleotide alignments of the 3010 ten-species orthologs were used as input for dN/dS ratio estimates. To estimate gene-specific ω (dN/dS) ratios, we first calculated per-gene branch-lengths using RAxML v8.2.8 (Stamatakis, 2014) given a fixed, unrooted species tree and the GTRCATI model. Gene-specific branch-lengths, respective alignments and the fixed, unrooted species tree were then parsed as input for codeml, which is included in the paml v4.8 package (Yang, 2007), using a custom script that also subsequently collected the output for each gene (see Code availability). We ran both a 'free' model, allowing one ω value for every branch, and a 'three-rate' model, allowing one rate for all asexual and one for all sexual terminal branches as well as one rate for internal (sexual) branches to separate mutations occurring at terminal branches from internal ones. Gene and branch-specific ω values were further analyses with R (R Core Team, 2013). We restricted analyses to transcripts with overall ω values $\omega < 1$ and $dS \neq 0$, to capture only transcripts that are under purifying selection. Within this gene set, we removed 200 orthologs that had no ω variation between species, leaving 2487 orthologs to compare. Here, we did not analyse positive selection on a $\omega > 1$ gene set, because this would require an entirely different approach that would be based on branch-site-specific models. Given that in the current study, we only used orthologs shared across the ten species, it seems unlikely that we could detect consistently adaptive signals.

Codon Usage Bias

To infer codon usage bias we calculated the effective number of codons (Enc) per gene using codonW v1.4 (Wright, 1990; Peden, 2005). Enc quantifies how strongly codon usage deviates from equal usage of all synonymous codons, taking values from 20 (each amino acid is encoded

by only one codon exclusively) to 61 (equal use of all possible alternative codons). Following, Enc values below 61 indicate codon usage bias.

We estimated signatures of relaxed selection on synonymous sites using the codon deviation coefficient (CDC; Zhang *et al.*, 2012). CDC estimates the expected codon usage from observed GC and purine contents and calculates the deviation from observed codon usage. CDC ranges from 0 (no detectable, i.e. 'relaxed' selection on codon usage) to 1 (effective selection on codon usage). We calculated CDC using the same alignments as used for ω analyses.

Polymorphism (pN/pS) analyses

We used the SNP data (details above under 'polymorphism') to identify non-synonymous and synonymous segregating polymorphism in the ten *Timema* populations. For this, we identified the fold-degenerate position of resulting filtered nucleotide variants using a custom script following Li *et al.* (1985) from which pN, pS and pN/pS was calculated per gene (see Code availability).

Statistics

For statistical tests on dN/dS, dS, Enc and CDC estimates we used a permutation ANOVA (Manly, 2006) with 5000 bootstrap replicates (script available in Henry *et al.*, 2012). For other analyses, statistical methods are given in the text.

Data availability

Raw reads are deposited in SRA under accession codes SRR5748941-SRR5749000. The transcriptome assemblies are deposited at DDBJ/EMBL/GenBank under the the BioProject PRJNA380865 with the following accession codes: GFPP00000000, GFPR00000000, GFPS00000000, GFPT00000000, GFPU00000000, GFPV00000000, GFPW00000000, GFPX00000000, GFPY00000000, and GFPZ00000000. Data will be publically released upon acceptance.

Code availability

The custom scripts used in this study are deposited at https://github.com/jensbast/TimemaConsOfAsex/.

Acknowledgements

We thank David Dylus and Clément Train for helping with orthology detection and Chloé Larose and Bart Zijlstra for helping with sample collection. This study was supported by a DFG research fellowship (grant number BA 5800/11 to J.B.); by funding from the University of Lausanne and Swiss SNF (grant number PP00P3_170627 to T.S).

Author contributions

T.S., J.B. conceived and designed the study. T.S., K.M.J. collected samples, Z.D. performed wet lab work, J.B., D.P., K.S.J., P.T.V., E.F., A.B. performed data analyses, J.B., T.S., N.G. interpreted data, J.B., T.S. wrote the paper with input from all authors.

Chapter 3

No signal of deleterious mutation accumulation in conserved sequences of extant asexual hexapods

Brandt, A., Bast, J., Scheu, S., Meusemann, K., Donath, A., Schuette, K., Machida, R., Kraaijeveld, K.

Sci. Rep. **9**, 5338 (2019)

Loss of sex and recombination is generally assumed to impede the effectiveness of purifying selection and to result in the accumulation of slightly deleterious mutations. Empirical evidence for this has come from several studies investigating mutational load in a small number of individual genes. However, recent whole transcriptome based studies have yielded inconsistent results, hence questioning the validity of the assumption of mutational meltdown in asexual populations. Here, we study the effectiveness of purifying selection in eight asexual hexapod lineages and their sexual relatives, as present in the 1 K Insect Transcriptome Evolution (1KITE) project, covering eight hexapod groups. We analyse the accumulation of slightly deleterious nonsynonymous and synonymous point mutations in 99 single copy orthologue protein-coding loci shared among the investigated taxa. While accumulation rates of nonsynonymous mutations differed between genes and hexapod groups, we found no effect of reproductive mode on the effectiveness of purifying selection acting at nonsynonymous and synonymous sites. Although the setup of this study does not fully rule out nondetection of subtle effects, our data does not support the established consensus of asexual lineages undergoing 'mutational meltdown'.

Introduction

The ubiquitous prevalence of sex among eukaryotes is surprising given that sexual reproduction involves manifold evolutionary costs as compared to obligate asexuality (Bell, 1982; Otto, 2009; Lehtonen *et al.*, 2012). One prediction for the benefit of sex in the long-term is the increased effectiveness of purifying selection (Kondrashov, 1988). The rationale is that segregation, recombination and outcrossing enable the uncoupling of linked loci with different selection coefficients, such that selection can act on different loci independently (Felsenstein, 1974). This accelerates adaptation and the purging of slightly deleterious mutations and facilitates the restoration of least loaded genotypes that are continuously lost by drift (Muller, 1964; Hill & Robertson, 1966; Felsenstein, 1974; Keightley & Otto, 2006). Asexual lineages lack these benefits and are therefore predicted to succumb to 'mutational meltdown' (Lynch *et al.*, 1993).

A multitude of studies have tested the prediction of impeded effectiveness of purifying selection (i.e. selective removal of deleterious mutations) in non-recombining genomic regions, such as mitochondria or (neo-) Y chromosomes as well as different lineages of asexual eukaryotes (Lynch & Blanchard, 1998; Bachtrog *et al.*, 2008; Hartfield, 2016; McDonald *et al.*, 2016). Their results have led to the established consensus that slightly deleterious mutations accumulate in the absence of sex. However, many of the studies that have investigated purifying selection in asexual species were based on only few individual genes and recent studies based on whole transcriptome comparisons between asexual and related sexual lineages did not find consistent support: while accumulation of slightly deleterious mutations was found in asexual *Timema* stick insects, *Oenothera* evening primroses and *Boechera* rockcress, it was absent in *Lineus* ribbon worms as well as four aphid species and, opposite to predictions, reduced in asexual as compared to sexual oribatid mites (Ollivier *et al.*, 2012; Hollister *et al.*, 2014; Ament-Velásquez *et al.*, 2016; Brandt *et al.*, 2017; Lovell *et al.*, 2017; Bast *et al.*, 2018). Moreover, all whole

transcriptome based studies found excessive variation among genes, and hence doubts have been raised about the robustness of inferences drawn from single gene analyses (Neiman *et al.*, 2018). These conflicting results highlight the need for more studies of larger gene sets along with broader taxonomic sampling to infer whether or not accumulation of deleterious mutations is indeed a consequence of asexual reproduction.

Here, we study the effectiveness of purifying selection in obligately asexual and related sexual hexapod lineages covering eight hexapod groups using transcriptome data generated by the 1KITE project (1 K Insect Transcriptome Evolution, www.1kite.org). As parthenogenesis is a lineage-level trait, we use the term 'lineage', whenever to distinguish between the two reproductive modes (sexual and asexual) within a hexapod group, from here on. With the term 'hexapod group' we refer to one of the following analysed taxa: Collembola, Zygentoma, Phasmatodea, Mantodea, Thysanoptera, Sternorrhyncha, Hymenoptera and Psocodea, respectively.

We compared the accumulation of slightly deleterious mutations in nuclear protein-coding genes between eight asexual lineages and their sexual relatives as available from 1KITE (see Fig. 1). To this end, we first inferred divergence at nonsynonymous sites normalised for background mutation rates (dN/dS), and the potential 'deleteriousness' of nonsynonymous substitutions. Second, we investigated selection on Codon Usage Bias (CDC), because selection also acts at synonymous sites (Hershberg & Petrov, 2008). We based all analyses on 99 single copy orthologues which we found to be under purifying selection in the analysed lineages. We found extensive variation in dN/dS and CDC among genes and between hexapod groups, but no overall difference between reproductive modes.

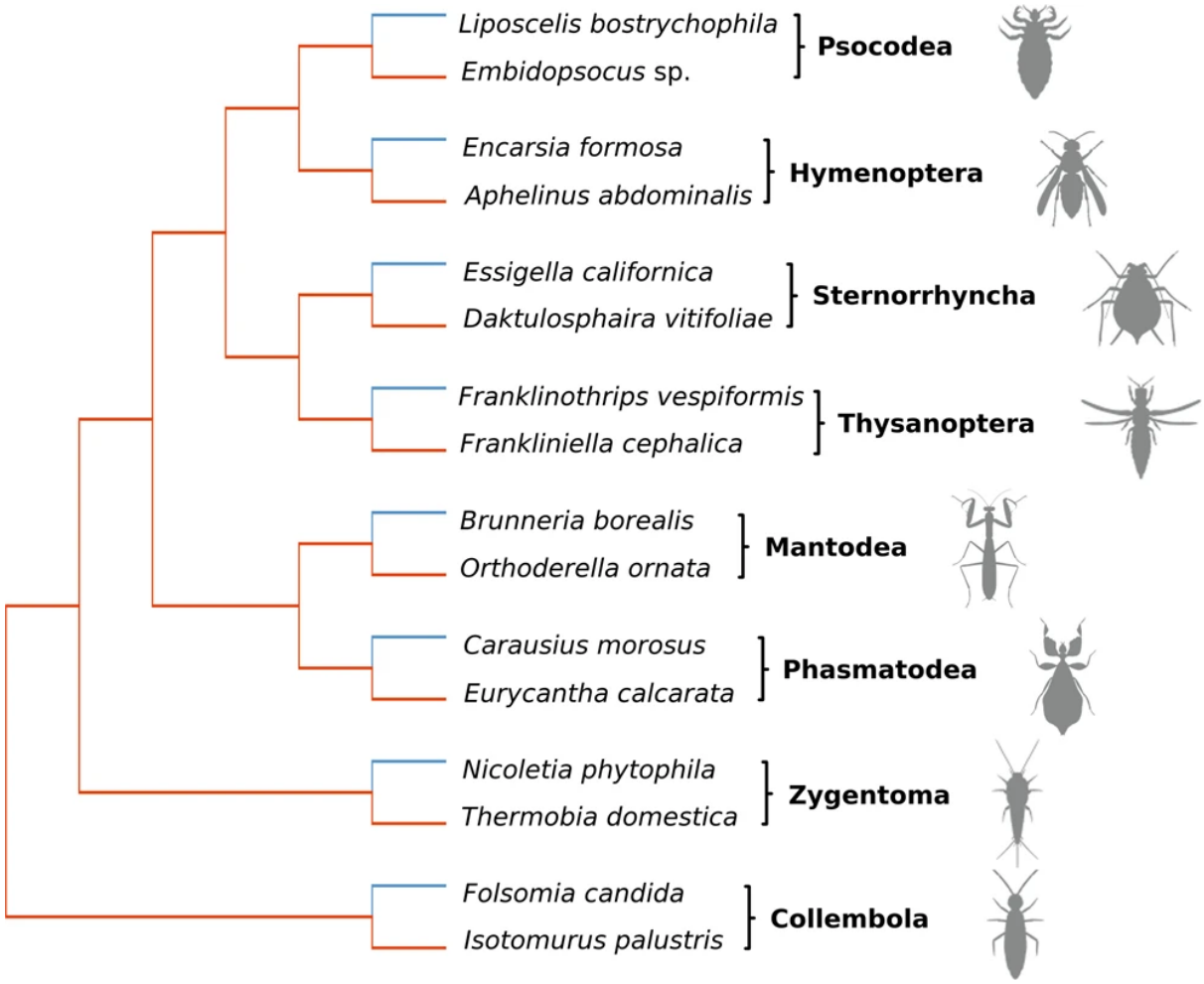


Figure 1: Cladogram of 16 hexapod species analysed in this study. The cladogram was manually built based on the phylogeny published by Misof et al. (2014; see Methods). The taxon sampling includes eight asexual lineages along with their closest sexual relatives as present in 1KITE, covering eight hexapod groups. Sexual and asexual lineages are depicted in red and blue, respectively. Silhouettes courtesy of Hans Pohl.

Methods

Species selection

We searched for obligately asexual lineages within the 1KITE species list based on primarily van der Kooi & Schwander (2014), Vershinina & Kuznetsova (2016) and the Tree of Sex database (Tree of Sex Consortium, 2014). We found eight hexapod groups with obligately asexual lineages represented by the following species: *Folsomia candida* (Collembola), *Nicoletia phytophila*

(Zygentoma), *Carausius morosus* (Phasmatodea), *Brunneria borealis* (Mantodea), *Franklinothrips vespiformis* (Thysanoptera), *Essigella californica* (Sternorrhyncha), *Encarsia formosa* (Hymenoptera) and *Liposcelis bostrychophila* (Psocodea). In all analysed species, parthenogenesis is thelytokous. In *C. morosus* and *E. californica* it is obligately apomictic whereas in *F. candida*, *F. vespiformis* and *E. formosa* parthenogenesis is induced by bacteria of the genus *Wolbachia* and in *L. bostrychophila* by the genus *Ricketsia* (White, 1948; Pijnacker, 1966; Simon *et al.*, 2002; Stouthamer & Mak, 2002; Frati *et al.*, 2004; Nguyen *et al.*, 2015; Yang *et al.*, 2015). We found no information concerning the mode of parthenogenesis in *B. borealis* and *N. phytophila* in the literature. Further, we selected per hexapod group the closest sexual relatives to the asexual lineages present in 1KITE, namely *Isotomurus palustris* (Collembola), *Thermobia domestica* (Zygentoma), *Eurycantha calcarata* (Phasmatodea), *Orthoderella ornata* (Mantodea), *Frankliniella cephalica* (Thysanoptera), *Daktulosphaira vitifoliae* (Sternorrhyncha), *Aphelinus abdominalis* (Hymenoptera) and *Embidopsocus* sp. (Psocodea). For information on first describers of species used, see Supplementary Table S1.

Data acquisition

We downloaded published and most current transcriptome assemblies from Genbank and TSA of the following species: *A. abdominalis*, *E. formosa*, *E. californica*, *F. candida*, *F. cephalica*, *L. bostrychophila*, *T. domestica* and *Xibalbanus* cf. *tulumensis* (von Reumont *et al.*, 2012; Misof *et al.*, 2014; Peters *et al.*, 2018). For remaining species, we obtained unpublished transcriptome assemblies from the 1KITE consortium (RNA extraction, cDNA library generation, sequencing, assembly and contaminant removal was done as described in detail in Misof *et al.*, 2014 and Peters *et al.*, 2017. Assemblies were used as input for identification of orthologous sequences among all studied species. For information on References, BioSample IDs and Bioproject IDs of assembled transcriptomes, see Supplementary Table S1.

Orthologue detection and alignment processing

To infer orthologue protein-coding genes among the 16 hexapod species, we used the Orthologous MAtrix (OMA) version 2.1.1 installed on the Vital-IT computing cluster (Train *et al.*, 2017). To generate the input for OMA, we extracted long Open Reading Frames (ORFs) and predicted likely coding regions with Transdecoder version 2.0.1 using default options for each transcript and species (Haas *et al.*, 2013). Amino acid sequences of likely coding regions of each of the 16 species were passed to OMA together with an unrooted cladogram built by reducing a published phylogeny of hexapods to the eight hexapod groups analysed in this study (Misof *et al.*, 2014), which resulted in 286 orthologues shared among all analysed species. We aligned the amino acid sequences (ORFs) of each of the 286 orthologues with M-Coffee using a combination of different alignment methods creating a consensus alignment (option clustalw2_msa muscle_msa kalign_msa mafftgins_msa t_coffee_msa; Wallace *et al.*, 2006). Based on these, we generated corresponding codon alignments from the original nucleotide sequences with T-Coffee version 11.00.8 (Wallace *et al.*, 2006). As orthologue detection based on naturally incomplete datasets such as transcriptome data can lead to the detection of false positives, e.g. due to loci with paralogous sequences, we tested all protein sequences of the 286 orthologues of all 16 species for their presence in a precompiled set of orthologues of insects (insecta_hmmer3.1; www.deep-phylogeny.org/hamstr/download/datasets/hmmer3) using HaMStR version 13.2.6 (Ebersberger *et al.*, 2009). We only kept multiple sequence alignments of orthologues for which all species had an equivalent in the precompiled orthologue set leaving a set of 153 loci for further analyses. We curated these multiple sequence alignments using Gblocks version 0.91b with sequence type set to codons (*t* = *c*) and minimum block length set to 4 (*b*4 = *4*; Castresana, 2000).

Accumulation of nonsynonymous mutations

As a measure of purifying selection, we analysed the divergence at nonsynonymous sites normalised for background substitution rates (dN/dS) using CodeML as implemented in PAML version 4.9 (Yang, 2007). For this, we first manually constructed an unrooted cladogram of the 16 analysed species based on a published hexapod phylogeny (Misof *et al.*, 2014). To exclude orthologues that were under positive selection, we ran BUSTED as implemented in HYPHY version 2.3.10 (Pond *et al.*, 2005) providing as input the unrooted cladogram described above and the multiple sequence alignments of the 153 orthologues. We found 54 loci showing signatures of episodes of positive selection, which left 99 loci for subsequent analyses (for GenBank Accession numbers, see Supplementary Table S2). CodeML relies on a Maximum Likelihood framework to estimate the goodness of fit of a codon substitution model to a sequence alignment and an unrooted species tree with gene-specific branch lengths for inference of branch-specific dN/dS ratios. We calculated per-gene branch lengths for the unrooted cladogram based on each of the 99 codon alignments, accordingly, using RAxML version 8.2.8 (Stamatakis, 2014) with GTRGAMMAI (with four GAMMA rate categories) set as model of sequence evolution. We modified a custom script used by Brandt et al. (2017; see Supplementary information) to pass the loci-specific branch lengths and fixed species tree together with each codon alignment to CodeML for divergence rate estimations. Due to the 1KITE taxon sampling, the asexual and sexual lineages of each hexapod group used in our analysis are likely not natural sister lineages (i.e. not the closest extant relatives) and therefore the time of transition to asexuality (or the split from the closest sexual relative) was unknown. As this could potentially lead to an overestimation of dN/dS under a model restricted to only two rates (one for asexual and one for sexual branches), we chose a free model allowing for different dN/dS ratios, one for each branch in the tree. For between-species comparisons of dN/dS ratios, we excluded all dN/dS ratios of internal branches and four terminal branch dN/dS ratios that were > 1, indicating positive selection acting at one

branch, from statistical analyses. We then tested whether branch-specific dN/dS ratios differed according to (I) gene (II) reproductive mode, or (III) hexapod group using a permutation ANOVA with 5,000 bootstrap replicates (available at https://gist.github.com/KamilSJaron/358c997698b67486be47d4e8eef2921d; Manly, 1997; Henry *et al.*, 2012). Differences in dN/dS can be driven by differing synonymous substitution rates and, in the long-term, different levels of saturation at synonymous sites. Given the old age of the splits between sexual and asexual lineages (~40 myo for Phasmatodea – ~160 myo for Zygentoma; Misof *et al.*, 2014; Bradler *et al.*, 2015) we tested for differences in branch-specific dS as described above. To infer whether or not sexual and asexual lineages within individual hexapod groups differed in dN/dS ratios, we compared between reproductive modes using Wilcoxon signed-rank tests. All statistical analyses were done in R version 3.4.4 (R Core Team, 2013).

'Deleteriousness' of nonsynonymous mutations

To infer the 'deleteriousness' of nonsynonymous substitutions we analysed hydrophobicity changes from ancestral to replacement amino acids along the terminal branches of the phylogenetic tree. Hydrophobic interactions are the main determinants of the 3D conformation of proteins and thus an indicator of protein stability (Pace *et al.*, 2011). Inference of ancestral amino acids relies on the presence of an outgroup sequence included in the input amino acid alignments and the phylogenetic tree (cladogram with loci-specific branch lengths) used in analyses with CodeML. Therefore, we first searched for orthologues shared among all 16 hexapod species plus the crustacean *Xibalbanus* cf. *tulumensis* (previously *Speleonectes* cf. *tulumensis*), as a representative of Remipedia, the sister-group of hexapods (von Reumont *et al.*, 2012; Misof *et al.*, 2014). For this, we predicted ORFs from the assembled transcriptome of *X.* cf. *tulumensis* and checked the ORFs of *X.* cf. *tulumensis* for presence of orthologues in the precompiled orthologue set of insects as described above. We found 73 ORFs of *X.* cf. *tulumensis*, each of them

orthologous to one of the previously inferred 99 clusters of orthologues of the 16 hexapod species (for GenBank Accession numbers, see Supplementary Table S2). We again aligned the amino acid sequences (ORFs, including now the sequences of *X*. cf. *tulumensis*) of the 73 clusters of orthologues and subsequently generated corresponding codon alignments from the original nucleotide sequences as described above. Further, we added *X*. cf. *tulumensis* to the unrooted manually constructed cladogram used for the calculation of branch lengths for dN/dS ratio analyses as outgroup to all hexapods. Using this tree as a fixed topology, we calculated branch lengths for each orthologue locus individually as described above for analyses of nonsynonymous mutation accumulation, and then translated the curated alignments into amino acids using EMBOSS version 6.6.0 (Rice, 2000). To predict ancestral amino acid sequences for each internal node in each inferred tree from each gene, we passed the 73 amino acid codon alignments individually with its respective species trees (and estimated loci-specific branch lengths) to CodeML using the modified custom script mentioned above (see Supplementary information). We determined the strength of hydrophobicity changes (Hydrophobicity Scores; HS) for each of the amino acid transitions along the terminal branches of the species tree using a hydrophobicity scoring (HS) matrix (Henry *et al.*, 2012; Riek *et al.*, 1995). HS indicates the 'deleteriousness' of a nonsynonymous mutation by measuring the strength in hydrophobicity change from ancestral to replacement amino acid. The lower the HS, the stronger is the change in hydrophobicity and, hence, the deleteriousness of the underlying nonsynonymous mutation. We compared values of HS between the two reproductive modes using Generalised Linear Mixed Models (GLMM) implemented in the R package lme4 with gene nested in species set as random effect, correction for overdispersion and Poisson distribution fitting (Bates *et al.*, 2014).

Accumulation of synonymous mutations

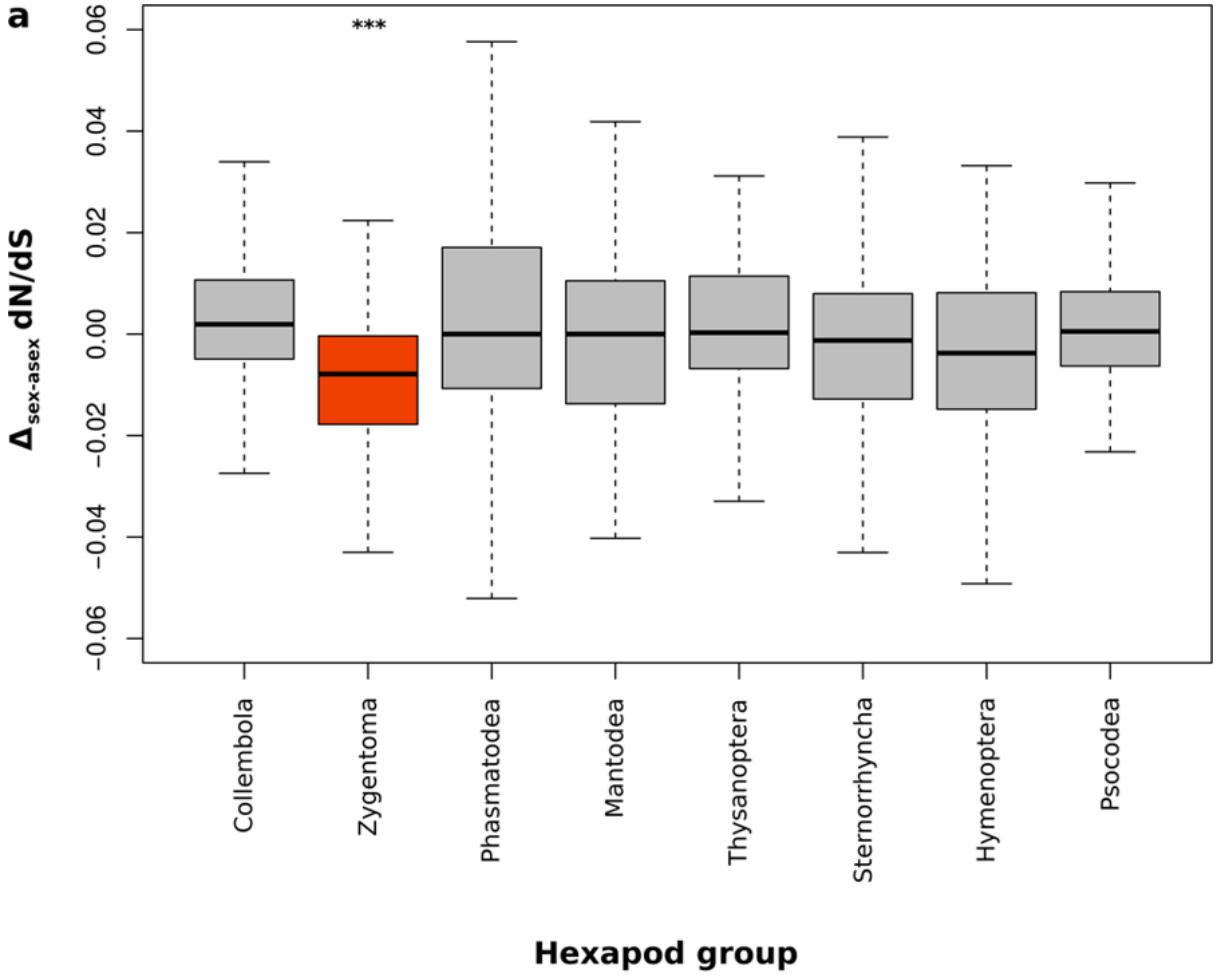
Synonymous mutations are generally assumed to be neutral but can be subject to purifying selection because different codons can influence the speed and accuracy of translation (Hershberg & Petrov, 2008). Hence, we also analysed the effectiveness of selection acting on Codon Usage Bias (CUB). First, we inferred the existence of CUB for each of the 99 orthologues of the 16 species using the Effective Number of Codons (Nc) as measure with the software codonW version 1.4 (Wright, 1990; Peden, 2005). The Nc specifies the deviation of observed codon usage from equal usage of all codons ranging from 20 (each amino acid is encoded by one codon only; strong CUB) to 61 (equal use of all possible codons; no CUB). Next, we inferred selection on CUB using the Codon Deviation Coefficient (CDC; Zhang *et al.*, 2012). Measurement of CDC allows for cross-species comparisons by correcting for background nucleotide composition and is particularly robust because, unlike dN/dS ratio analyses, it does not rely on likelihood and branch length estimates. The CDC represents the deviation of expected CUB based on observed positional GC and purine contents from observed CUB, ranging from 0 (no deviation; no detectable selection on CUB) to 1 (maximum deviation; strong selection on CUB). We estimated CDC for the processed alignments using Composition Analysis Toolkit version 1.3 (Zhang *et al.*, 2012) and analysed it for an effect of I) gene, II) reproductive mode, and III) hexapod group and inferred within-group differences as described above for statistical analyses of dN/dS ratios.

Results

Accumulation of nonsynonymous mutations

We estimated nonsynonymous to synonymous divergence (dN/dS) along individual branches of a phylogenetic tree comprising eight asexual and eight sexual hexapod species for 99 single-copy orthologous protein-coding genes under purifying selection (see Fig. 1; Methods). In genes under

purifying selection, nonsynonymous mutations have likely deleterious effects, hence, a higher dN/dS ratio indicates less effective purifying selection (Li *et al.*, 1985). Consistent with the expectation for loci under purifying selection, dN/dS ratios at terminal branches were generally low (mean dN/dS = 0.032). Contrasting the established consensus on deleterious mutation accumulation in asexual lineages, there was no difference in dN/dS when compared between sexual and asexual branches over all hexapod groups (gene effect $P < 0.001$, reproductive mode effect $P = 0.488$, hexapod group effect $P = 0.048$, interaction reproductive mode * hexapod group $P = 0.145$; Permutation ANOVA). The absence of a difference between reproductive modes was not driven by a difference in dS (gene effect $P < 0.001$, reproductive mode effect $P = 0.278$, hexapod group effect $P < 0.001$, interaction reproductive mode * hexapod group $P = 0.004$; Permutation ANOVA). There was significant among-gene variation in dN/dS (range 0–0.454; variance $1.04 * 10^{-3}$) and a significant difference in dN/dS among hexapod groups. To detect whether or not the effectiveness of purifying selection differed between reproductive modes within each hexapod group, we compared dN/dS between reproductives modes in each hexapod group on a per-gene basis. For Zygentoma, we found significantly lower per-gene dN/dS in the sexual as compared to the asexual terminal branch indicating more effective purifying selection for the sexual branch (see Table 1; Fig. 2a red box).


a
***
$\Delta_{sex\text{-}asex}$ dN/dS
0.06
0.04
0.02
0.00
−0.02
−0.04
−0.06
Collembola
Zygentoma
Phasmatodea
Mantodea
Thysanoptera
Sternorrhyncha
Hymenoptera
Psocodea
Hexapod group

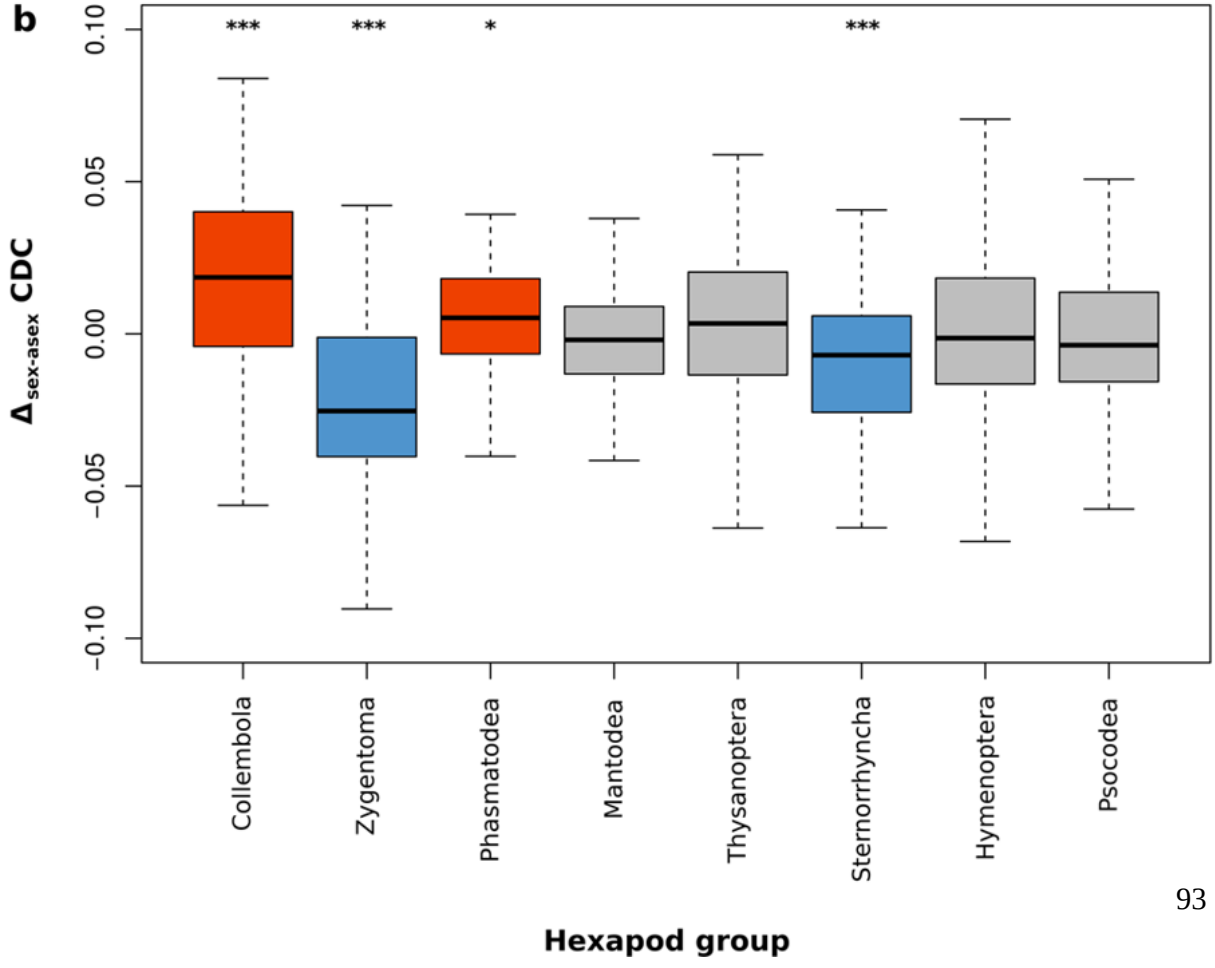

b
***
***
*
***
$\Delta_{sex\text{-}asex}$ CDC
0.10
0.05
0.00
−0.05
−0.10
Collembola
Zygentoma
Phasmatodea
Mantodea
Thysanoptera
Sternorrhyncha
Hymenoptera
Psocodea
Hexapod group

Figure 2: Per-gene differences in (a) dN/dS and (b) CDC between reproductive modes within each hexapod group. The boxplots show the distribution of per-gene differences in dN/dS between the sexual and asexual terminal branch ($\Delta_{sex\text{-}asex}$ dN/dS) and of per-gene differences in CDC between the sexual and asexual species ($\Delta_{sex\text{-}asex}$ CDC) of each of eight hexapod groups for 99 orthologues. For better representation, the ordinate is restricted to between −0.06 and 0.06 and −0.1 and 0.1, respectively, and outliers outside of 1.5 times the interquartile range (whiskers) are excluded. For dN/dS seven of eight within-hexapod-group comparisons between sexual and asexual branches yielded non-significant results (see Table 1). The red plot denotes a significantly lower per-gene dN/dS in sexual as compared to asexual terminal branches of Zygentoma ($V = 3786$, $P < 0.001$***; Wilcoxon signed-rank test). For CDC four of eight within-hexapod-group comparisons between sexual and asexual branches yielded significant results (see Table 1). There was significantly higher per-gene CDC in sexual as compared to asexual Collembola and Phasmatodea (red plots; $V = 1191$, $P < 0.001$*** and $V = 1832$, $P = 0.025$*, respectively; Wilcoxon signed-rank test) contrasting a significantly lower per-gene CDC in sexual as compared to asexual Zygentoma and Sternorrhyncha species (blue plots; $V = 4198$, $P < 0.001$ and $V = 3562$, $P < 0.001$, respectively; Wilcoxon signed-rank test).

Table 1: *V* and *P* values of within-hexapod group comparisons of dN/dS and CDC between reproductive modes. Values were inferred by comparing dN/dS and CDC per gene between the sexual and asexual lineages in each of eight hexapod groups using Wilcoxon signed-rank tests. Underlined and bold measures indicate more effective purifying selection in sexual species (underlined) and asexual species (bold), respectively (see Fig. 2a, b). Coll.: Collembola; Zyg.: Zygentoma; Phas.: Phasmatodea; Man.: Mantodea; Thys.: Thysanoptera; Stern.: Sternorrhyncha; Hym.: Hymenoptera; Psoc.: Psocodea.

Measure	**Hexapod group**							
	Coll.	**Zyg.**	**Phas.**	**Man.**	**Thys.**	**Stern.**	**Hym.**	**Psoc.**
dN/dS	*V* = 1886	*V* = 3786	*V* = 1747	*V* = 1263	*V* = 2008	*V* = 2382	*V* = 2845	*V* = 2056
	P = 0.107	*P* < 0.001	*P* = 0.381	*P* = 0.585	*P* = 0.314	*P* = 0.179	*P* = 0.092	*P* = 0.407
CDC	*V* = 1191	***V* = 4198**	*V* = 1832	*V* = 2593	*V* = 2040	***V* = 3562**	*V* = 2471	*V* = 2737
	P < 0.001	***P* < 0.001**	*P* = 0.025	*P* = 0.682	*P* = 0.129	***P* < 0.001**	*P* = 0.99	*P* = 0.361

‘Deleteriousness’ of nonsynonymous mutations

In addition to nonsynonymous mutation accumulation in asexual hexapod lineages, purifying selection is expected to lead to more deleterious amino acid transitions in asexual lineages as reflected by stronger hydrophobicity changes from ancestral to replacement amino acids along asexual and sexual terminal branches (Sharbrough *et al.*, 2018). Contrasting this expectation, but consistent with our dN/dS estimates, HS were similar between asexual and sexual branches ($z = -0.152$; $P = 0.879$; GLMM; see Supplementary Fig. S1). Percentages of ancestral to replacement amino acid transitions with more dissimilar hydrophobicity (HS < 90) were similar between asexual and sexual branches (44.936% and 44.893% of all asexual and sexual transitions, respectively).

Accumulation of synonymous mutations

We assessed whether or not purifying selection on synonymous sites was less effective in asexual as compared to sexual hexapod lineages by inferring selection on CUB. To infer whether or not the species investigated use some codons preferentially over others, we first analysed the Effective Number of Codons (Nc). Nc ranges from 20 (each amino acid is encoded by one codon only; strong CUB) to 61 (equal use of all possible codons; no CUB). Overall, all species showed CUB, with Nc means of species ranging from 40.6 in *D. vitifoliae* (Sternorrhyncha) to 54.1 in *F. candida* (Collembola; see Supplementary Fig. S2) which is comparable to values of Nc found in other invertebrate groups, e.g. bivalves and nematodes (Mitreva *et al.*, 2006; Gerdol *et al.*, 2015). Afterwards, we directly inferred selection on CUB using CDC as a measure. CDC calculates the deviation of observed from predicted CUB by correcting for background nucleotide composition ranging from 0 (no selection on CUB) to 1 (effective selection on CUB). Consistent with the results from analyses of nonsynonymous mutation accumulation and ‘deleteriousness’ of nonsynonymous mutations, there was no difference in per-gene CDC between reproductive

modes (gene effect $P < 0.001$, reproductive mode effect $P = 0.283$, hexapod group effect $P < 0.001$, interaction reproductive mode * hexapod group $P < 0.001$; Permutation ANOVA). Further, there was significant among gene variation for CDC (range 0.064–0.329; variance 1.653 * 10^{-3}) and a significant difference between hexapod groups but, unlike for dN/dS estimates, there was also a significant interaction between hexapod group and reproductive mode. As for dN/dS, we compared per-gene CDC between reproductive modes in each hexapod group, individually. Four of eight within-hexapod-group comparisons between sexual and asexual species yielded significant results (see Table 1; Fig. 2b red and blue boxes, respectively). For Collembola and Phasmatodea, there was a significantly higher per-gene CDC in sexual as compared to asexual species indicating more effective selection on CUB in sexual species, whereas for Zygentoma and Sternorrhyncha there was a significantly lower per-gene CDC in sexual as compared to asexual species indicating more effective selection on CUB in asexual species.

Discussion

It has become established consensus among evolutionary biologists that sex and recombination increase the effectiveness of purifying selection, based on theoretical considerations and empirical evidence derived from a multiplicity of studies (Hartfield, 2016). Our results do not match these studies: we find no evidence for accumulation of deleterious mutations in asexual hexapod species. Overall, asexual and sexual lineages neither differed in nonsynonymous site divergences corrected for background substitution rates (dN/dS), in synonymous substitution rates (dS) potentially influencing dN/dS ratios, in the ‘deleteriousness’ of nonsynonymous mutations nor in selection acting on CUB (CDC). Comparisons between sexual and asexual lineages within each group differed for several hexapod groups which may hint at group-specific

differences in the consequences of asexuality on effectiveness of purifying selection. However, here, we refrain from drawing conclusions based on the within-group comparisons because the taxon sampling of 1KITE did not allow for analysing multiple within-group replicates but restrict the discussion to the overall result of more effective selection being absent in our data of sexual hexapods: which (non-mutually exclusive) mechanisms might be responsible for the discrepancy between this finding and the established consensus?

First, analyses of purifying selection acting at nonsynonymous sites and on their 'deleteriousness' may have been affected due to data limitations. dN/dS ratio analyses and ancestral state reconstructions rely on branch length estimates (see Methods). Due to the limited nature of the taxon sampling of the 1KITE data set for this study, the asexual and sexual species analysed are most likely not sister species, but rather more distantly related (Munro *et al.*, 2011; Bradler *et al.*, 2015). Therefore, the loss of sex did not occur with the split of the sex-asex species pair as present in the given phylogenetic tree here, such that evolution over some fraction of the asexual branch was likely sexual. If mutations occuring in asexual lineages did not accumulate at greatly increased rates compared with sexuals, a change in deleterious mutation accumulation along with the transition to asexuality might be masked by the rates that occured in the sexual fraction of the branch. Thus, the power to detect an effect of reproductive mode on the effectiveness of purifying selection might be low. Further, the orthologue search among phylogenetically distantly related hexapod groups and the stringent control for false positive orthologues and loci under positive selection resulted in a rather small orthologue set for analysis (99 orthologue loci). This biases the analyses towards strongly conserved loci and excludes recently evolved orthologues which might differ in accumulation of deleterious mutations between reproductive modes. Also, the within hexapod group comparisons between reproductive modes for CDC did not resemble those for dN/dS; in the case of Zygentoma they even opposed them (see Fig. 2a, b; Table 1). This is

surprising, given that translational selection acting at synonymous sites is assumed to be weak and effective purging of synonymous mutations likely occurs at lower rates as compared to that of nonsynonymous mutations. Hence, within hexapod group comparisons of purifying selection acting at nonsynonymous sites should not oppose those at synonymous sites.

Second, the analysed asexual lineages may have lost sex too recently to have fixed enough deleterious mutations to be detected. This was for example also assumed to be responsible for absence of deleterious mutation accumulation observed in whole transcriptome data of fissiparous *Lineus* ribbon worms and four aphid species (Ollivier *et al.*, 2012; Ament-Velásquez *et al.*, 2016). In fact, the occurrence of abundant males in some locally restricted populations of *N. phytophila*, *E. californica*, and *L. bostrychophila* analysed in this study is in line with a rather recent loss of sex (Ollivier *et al.*, 2012). Additionally, interference of rare, furtive, or cryptic sex with asexual genome evolution may explain the observed absence of less effective selection in asexual hexapods because rare events of sex are assumed to be sufficient to compensate for predicted consequences of asexuality (Hurst *et al.*, 1992).

A third reason for the absence of deleterious mutation accumulation in the analysed data of asexual hexapod lineages may be effective homogenising mechanisms. Gene conversion and DNA repair have been shown to maintain DNA integrity, e.g. within the human Y chromosome, higher plant chloroplasts and animal mitochondria (Kang & Hamasaki, 2002; Khakhlova & Bock, 2006; Hughes *et al.*, 2012). If homogenising mechanisms play a role in the analysed species remains to be investigated.

Fourth, besides reproductive mode, population size acts as a major determinant of effectiveness of selection with the speed of mutation accumulation being inversely related to population size, as shown by modelling approaches (Lynch *et al.*, 1993; Gordo & Charlesworth, 2000; Rice & Friberg, 2009). Large population sizes have been suggested to maintain effective purifying

selection in asexual organisms, e.g. in oribatid mites and polyphagous pest insects, such as scale insects (Normark & Johnson, 2011; Ross *et al.*, 2013; Brandt *et al.*, 2017). Further, large population sizes have been shown to increase the effectiveness of selection acting on CUB in a variety of animal species and on nonsynonymous mutations with large deleterious effects in *Caenorhabditis elegans* (Katju *et al.*, 2015; Galtier *et al.*, 2018). As information on population sizes of the analysed species is absent, any correlation between population sizes and the observed absence of more effective purifying selection acting at nonsynonymous and synonymous sites has to remain speculative.

In conclusion, our results do not support the established consensus of reduced effectiveness of purifying selection in asexual species, contrasting earlier analyses of few individual genes in a variety of animal and plant species. However, whether or not this is due to the limited nature of our data or due to non-mutually exclusive biological mechanisms has to remain elusive at this point. Future studies, hence, need not only to include large gene sets but also carefully selected closely related sexual and asexual focal lineages to study the effectiveness of purifying selection in asexual organisms in more detail.

Data Availability

Sequence data analysed in this study is available at NCBI GenBank under Accession numbers MH551269-MH551284, MH602437-MH602956, MH637812-MH638065 and MH799322-MH800185. Supplementary data is available for download from the digital repository DRYAD under https://doi.org/10.5061/dryad.5501rv4 (see Supplementary Archives S1–S4; Supplementary information).

Acknowledgements

We thank Kamil S. Jaron for help with statistics, Patrick Tran Van for help with bioinformatics, Hans Pohl for permission of using silhouettes of hexapod groups and Ina Schaefer, Mark Maraun and Tamara Hartke for helpful comments on the manuscript. We also thank the 1KITE community, in particular Xin Zhou and Bernhard Misof, the 1KITE Basal Hexapod group, the 1KITE Polyneoptera group, the 1KITE Hymenoptera group and the 1KITE Paraneoptera group for collecting material, sequencing and assembly, cleaning of transcriptomes and granting us access to unpublished transcriptome data. Further, we thank Beijing Genomics Institute for funding sequencing of the 1KITE transcriptomes. This study was supported by core funding of S.S. and by DFG research fellowship BA 5800/1-1 to J.B.

Contributions

A.B., J.B. and K.K. conceived and designed the study. K.M., K.S., A.D. and R.M. provided transcriptome data and respective taxa specific information. A.B. and J.B. performed data analysis with the help of K.M. for identifying orthologs. A.B., J.B., K.M., K.K. and S.S. contributed to data interpretation and analyses and A.B. wrote the paper with input from all other authors. All co-authors approved the manuscript.

Chapter 4

Haplotype divergence supports obligate asexuality in the oribatid mite *Oppiella nova*

Brandt, A.*, Anselmetti, Y., Bluhm, C., Dumas, Z., François, C. M., Galtier, N., Heimburger, B., Jaron, K. S., Labédan, M., Maraun, M., Robinson-Rechavi, M., Schaefer, I., Scheu, S., Simion, P.,Tran-Van, P., Schwander, T.† & Bast, J.†

in prep. (2020)

* Corresponding author
† Shared senior authorship
Co-authors in alphabetical order

Sex strongly impacts genome evolution via recombination and chromosomal segregation. In the absence of these processes haplotypes within lineages of diploid organisms are predicted to accumulate mutations independently of each other and diverge over time. This so-called 'Meselson effect' is regarded as strong support for the evolution under obligate asexuality. Here, we test for its presence in a putatively ancient asexual oribatid mite species by comparatively analysing haplotype divergences in different populations of the asexual *Oppiella nova* and its close sexual relative *Oppiella subpectinata*. In the asexual species, we find a high proportion of shared heterozygous variants between individuals of different populations which indicates that haplotypes diverged prior to the separation of populations. Further, we find close topological resemblance of haplotype subtrees, indicating parallel divergence of haplotypes. Our findings of the Meselson effect in a putatively ancient asexual oribatid mite species are consistent with the strict absence of sex over evolutionary times and highlight the potential of ancient asexual oribatid mites as model organisms to investigate the singularities that allow for long-term persistence under asexuality.

Introduction

Sexual reproduction is considered as a prerequisite for the long-term persistence of the vast majority of species because it facilitates adaptation and purifying selection (Lehtonen *et al.*, 2012; Neiman *et al.*, 2017). Contrary to this scientific consensus, some exceptional taxa likely have persisted and diversified in the absence of sex for millions of years, the so-called 'ancient asexuals scandals' (sensu Judson & Normark, 1996; Schoen *et al.*, 2009a; Schurko *et al.*, 2009). The most notorious examples are the bdelloid rotifers, the darwinulid ostracods and several taxa of oribatid mites (Heethoff *et al.*, 2009; Mark Welch *et al.*, 2009; Schoen *et al.*, 2009b). Parthenogenetic lineages of these groups have persisted for tens of millions of years and diversified into ecologically different species (Birky & Barraclough, 2009; Schurko *et al.*, 2009). These evolutionary scandals are invaluable, because understanding how they have persisted as asexuals can help to identify the adaptive value of sex (Butlin, 2002). However, testing for patterns of evolution under ancient asexuality in a lineage requires the exclusion of rare events of sex, which is challenging.

One of the strongest predictions for the evolution of truly obligate asexual organisms is that the two haplotypes (each stemming from one homologous chromosome copy) within a diploid clonal lineage accumulate mutations independently of each other and evolve as independent lineages. Thus, after the loss of sex, the haplotype sequences diverge over time and levels of intra-individual heterozygosity increase (Fig. 1). This intra-individual **haplotype divergence (HD)** is more commonly known as the 'Meselson effect' (Birky, 1996; Normark *et al.*, 2003).

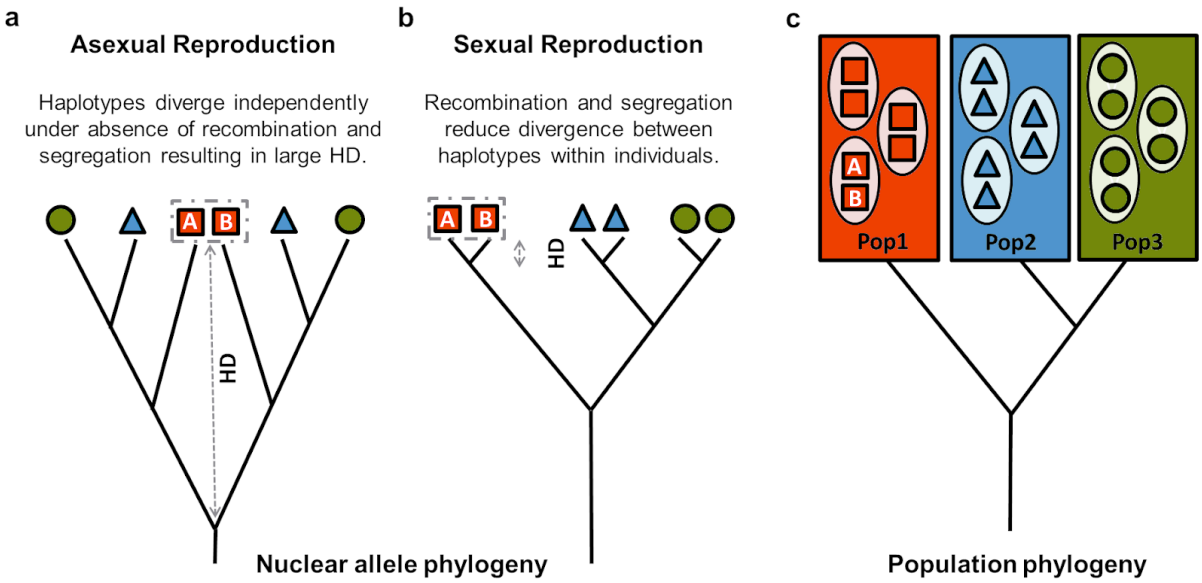


Figure 1: Nuclear haplotype phylogenies expected under (long-term) obligate asexual and sexual reproduction. In diploid, effectively asexual organisms, homologous chromosomes accumulate mutations independently of each other and evolve as independent lineages (note that this can be restricted to specific regions sheltered from a loss of heterozygosity). Accordingly, divergence between haplotypes within individuals (HD) is expected to exceed divergence between haplotypes of individuals from different populations, as illustrated in the allele phylogeny of asexuals (a): Here, the splits that separate homologous haplotypes are expected to be older than the splits that separate populations. Also the haplotype subtrees resemble each other because haplotypes are expected to evolve in parallel. In sexual organisms instead, HD is expected to follow population divergence and the haplotype phylogeny of sexuals (b) resembles that of the population (c). Figure adapted from Schwander *et al.* (2011).

Surprisingly, this strong theoretical prediction has thus far only met equivocal empirical validation. In several organisms the Meselson effect was not found (e.g., darwinulid ostracods; Schoen & Martens, 2003) or could be explained by mechanisms other than haplotype divergence after the transition to asexuality, such as a hybrid origin (e.g., *Meloidogyne* nematodes; Lunt, 2008) or divergence between paralogs (ohnologs) rather than between haplotypes (e.g. bdelloid rotifers; Mark Welch *et al.*, 2008; Flot *et al.*, 2013). Moreover, gene conversion can contribute to the loss of heterozygosity (LOH) in genomic regions and reduce HD. Support for the Meselson

effect was found in the fissiparous *Dugesia* flatworms but based only on one nuclear and one mitochondrial gene (Leria *et al.*, 2019). The, as yet, strongest support comes from a whole-genome study of obligately asexual trypanosomes, unicellular parasitic flagellates, in which some genomic regions which are not affected by LOH show the expected HD and parallel divergence (Weir *et al.*, 2016). For parthenogenetic organisms, however, the current literature still lacks any support of the Meselson effect.

One of the most promising eukaryotic systems for understanding the evolutionary persistence in the absence of sex are oribatid mites (Schwander, 2016). Many lineages lost sex several million years ago independently, providing the possibility for comparative analyses (Norton & Palmer, 1991; Cianciolo & Norton, 2006). Some of them diversified into clades comprising up to 150 morphologically distinct species, illustrating their evolutionary success in the absence of sex (Norton & Palmer, 1991; Norton *et al.*, 1993; Maraun *et al.*, 2003; Birky & Barraclough, 2009; Heethoff *et al.*, 2009; Von Saltzwedel *et al.*, 2014).

In this study, we test if HD in an asexual oribatid mite species is consistent with the ‘Meselson effect’. For this, we first generated *de novo* genomes based on single individuals of the asexual oribatid mite species *Oppiella nova* and its close sexual relative *Oppiella subpectinata* for comparison. To study haplotype divergence patterns on a broad geographical scale, we sequenced transcriptomes of three individuals per species from three separate geographical locations (‘geographical populations’; nine individuals in total per species, see Supplementary Figure 1). Additionally, we generated genome data to support part of the transcriptome-based analyses on a genome-wide scale and enable comparison of genome-wide heterozygosity levels between reproductive modes (one “geographical population” per species; four individuals *O. nova*, five individuals *O. subpectinata*).

Using polymorphism data, we tested for four population genomic signatures exclusively expected under obligate asexuality ('Meselson effect'; see Fig. 1). These signatures should be present in asexual *O. nova* but absent in sexual *O. subpectinata*: (I) variation between haplotypes within individuals exceeds variation between haplotypes from different populations; (II) an excess of shared heterozygous variants among individuals of different populations (indicating that haplotypes diverged prior to separation of populations); (III) the deepest split in haplotype phylogenies separates haplotypes, not populations (as opposed to sexual organisms where the deepest split should typically correspond to populations unless there is no recombination between haplotypes); (IV) the topologies of haplotype subtrees resemble each other due to parallel divergence of haplotypes during population separation. Also we tested if genome-wide heterozygosity levels of *O. nova* exceed those of the closely related sexual species *O. subpectinata* as a consequence of HD.

Results

De novo genomes

The quality- and contaminant-filtered *de novo* genome assemblies (v03, see Data availability) spanned a total size of 213 Mb for *O. subpectinata* and 197 Mb for *O. nova* and contained 60,250 and 63,118 scaffolds, with an N_{50} of 7,017 and 6,753 bp, with 23,555 and 23,761 genes annotated that spanned 24% of assemblies (see Supplementary Tables S1, S2, S3, Supplementary Figure 2 and Methods for details). Despite the relatively low assembly contiguity (derived from whole genome amplification), individual genes were assembled and annotated with confidence. This is reflected by the high completeness scores of core genes (C), with very few fragmented (F) or missing (M) genes (C:86.2%, F:6.4%, M:7.4% for *O. subpectinata* and C:87.5%, F:6.6%,

M:5.9% for *O. nova*; see Supplementary Table S1), and high quality annotation metrics (see Supplementary Figure 2).

(I) Variation within individuals exceeds variation between populations

Under HD, within-individual divergence is expected to exceed divergence between populations. Therefore, we analysed within-individual and between-population differentiation of three geographically distinct populations per species, consisting of three sampled individuals, each. The analyses were based on polymorphism data generated from RNAseq reads mapped to the reference genomes (for location of sampling sites, see Supplementary Figure 1 and Methods). For *O. nova*, multi-dimensional scaling (MDS) analysis showed two distinct clusters of four and three individuals, each comprising individuals from different geographical populations, respectively (Fig. 2a). For *O. subpectinata*, individuals separated into three distinct clusters, each corresponding to one geographical population (Fig. 2b). Accordingly, between-population variation contributed a minority to overall species-wide genetic variation in *O. nova* (12.79%) but was considerably larger in *O. subpectinata* (57.08%; for detailed results of the AMOVA see Supplementary Table S4). In *O. nova* most variation (51.26%) was explained by variation within individuals (due to heterozygous variants), as predicted under HD. Additionally, variation among individuals within geographical populations was substantial (35.94%). The clustering of individuals from different populations in the MDS plot in two groups indicates the presence of at least two coexisting divergent lineages in *O. nova*. Therefore, we also analysed variation between lineages of the seven individuals that formed two distinct clusters, only. Variation between lineages was small (26.05%), while variation within individuals was considerably larger (55.32%; see Supplementary Table S4).

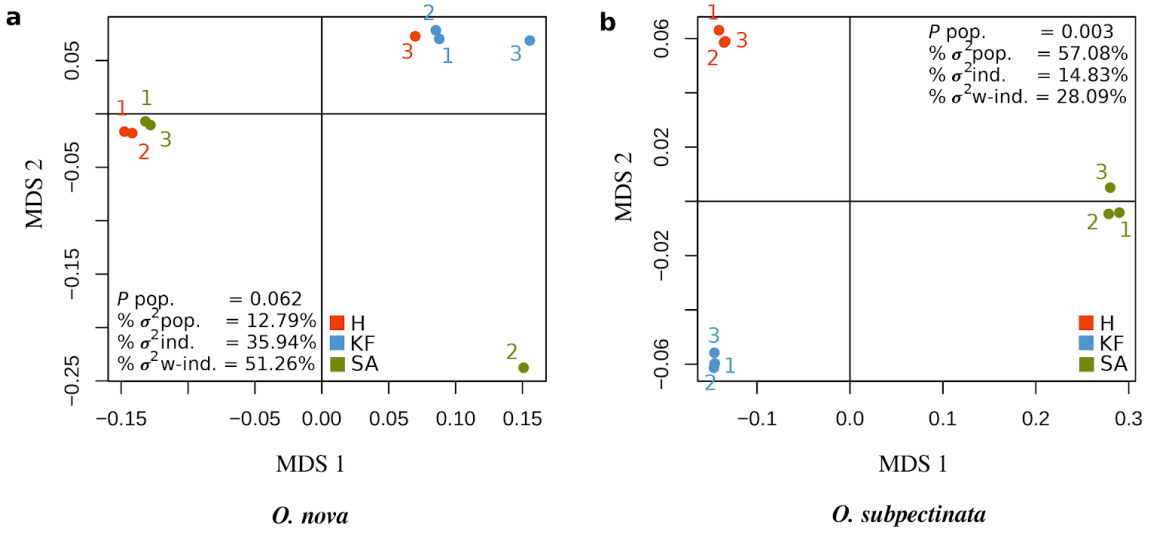


Figure 2: MDS plots and AMOVA results of the asexual *O. nova* (a) and sexual *O. subpectinata* (b). MDS plots display clustering of individuals by population for *O. subpectinata* (rand-test *P* pop. = 0.003) but not for *O. nova* (rand-test *P* pop. = 0.062) where two individuals from Schwäbische Alb (SA) and Hainich (H), and one individual from H and two individuals from Kranichstein forest (KF) clustered together. This is consistent with the large variation among individuals within populations (% σ^2 ind.) to total genetic variation. Notably, the majority of total genetic variation was explained by differences between populations (% σ^2 pop.) in *O. subpectinata* but by within-individual differences in *O. nova* (% σ^2 w-ind.).

(II) Excess of shared heterozygous SNPs between individuals of different populations

Loss of sex prior to population separation is expected to result in the presence of shared heterozygous variants between individuals of different populations. To test this, we analysed the distribution of genotype profiles generated from RNAseq data combined over all three populations using Genotype Profile Site Frequency Spectra (GPSFS). Site Frequency Spectra depict the number of sites with a given number of alternative (non-reference) variants over all individuals (e.g. nine diploid individuals can display a maximum of 18 alternative variants). Genotype profiles, i.e. distinct combinations of alternative alleles over individuals for each number of alternative variants, are indicated with different colours in the Site Frequency Spectra (see Fig. 3). *Oppiella nova* showed an overrepresentation of sites with heterozygous SNPs (more

than twice than expected under Hardy-Weinberg equilibrium) exclusively shared among seven individuals from all populations (grey bar; two individuals that lacked shared heterozygous SNPs with the other individuals were excluded from the analyses: Fig. 3a). Further, we found an overrepresentation of sites with heterozygous SNPs exclusively shared among a group of four individuals (two from H and SA, each) and of three individuals (two from KF and one individual from H: see Fig. 3a, green and blue bars, respectively). These groups corresponded to individuals that formed the two distinct clusters in the MDS plot of *O. nova* and likely represent separate genetic lineages (Fig. 2a). The results are consistent with an accumulation of shared heterozygous variants after the loss of sex, followed by lineage separation and independent accumulation of shared heterozygous variants among individuals of each lineage (Fig. 3b). One individual from SA and one individual from KF lacked an excess of heterozygous SNPs shared with the other individuals but homozygous sites were overrepresented for the individual from SA (see Supplementary Figure 4). The GPSFS of *O. subpectinata* showed an excess of homozygous alternative genotypes separating the population SA from the populations H and KF indicating that the large contribution of between-population variation to total variation as inferred from AMOVA is largely due to differences in the population from SA (see Figs 2b, 3c; for the full GPSFS of *O. nova* and GPSFSs of both species with detailed genotype assignment, see Supplementary Figures 3-5).

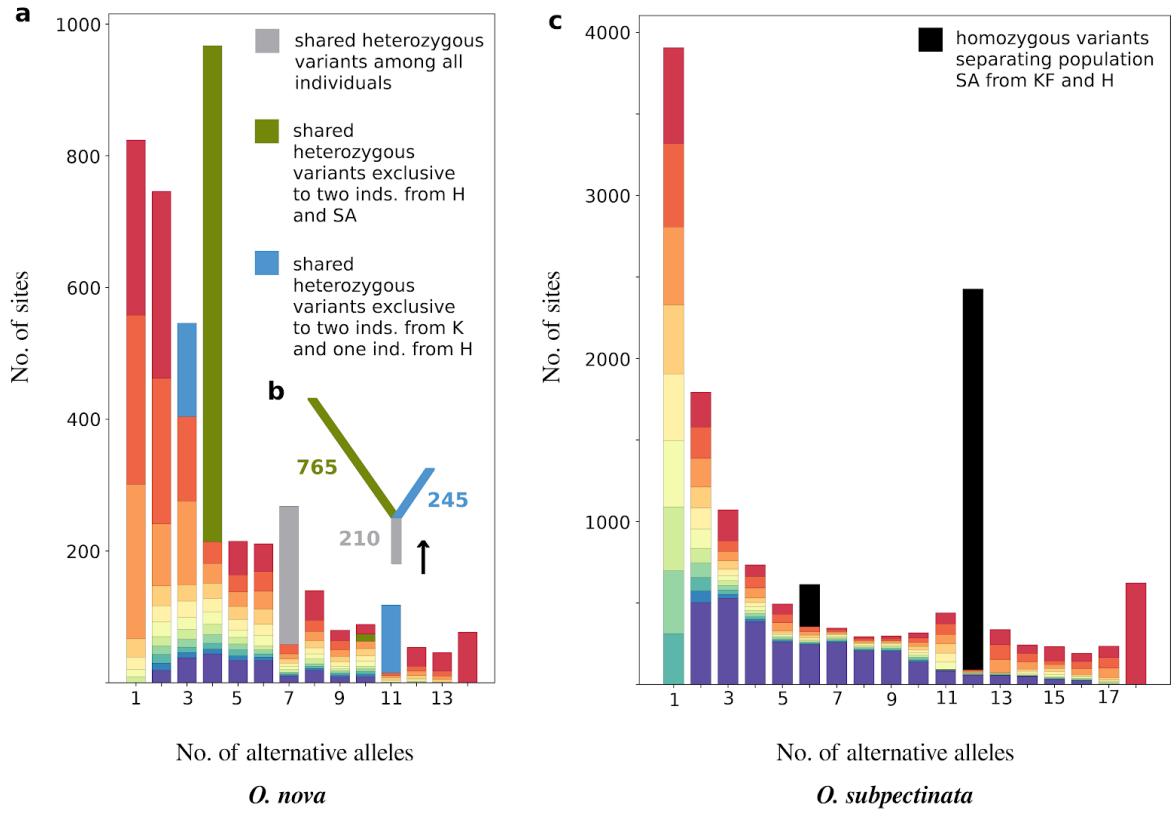


Figure 3: Genotype profile site frequency spectra (GPSFSs) of seven and nine diploid individuals, stemming from three geographical populations, for the asex-sex species pair *O. nova* (a, b) and *O. subpectinata* (c). For better representation of shared heterozygous SNPs in *O. nova* one individual from SA and KF were excluded, as both lacked an excess of heterozygous SNPs shared with the other individuals (for the full GPSFS, see Supplementary Figure 4). The grey coloured bar (a) indicates shared heterozygous SNPs among all seven individuals, the green coloured bars indicate shared heterozygous SNPs private to one divergent lineage comprising individuals one and two from H and individuals one and three from SA, and the blue coloured bars represent shared heterozygous SNPs private to a different lineage comprising individuals one and two from KF and three from H. The pattern in *O. nova* is consistent with the accumulation of shared heterozygous variants after the loss of sex, followed by lineage separation and independent accumulation of shared heterozygous variants among individuals of each lineage (b; schematic phylogeny displaying lineage separation and numbers of shared heterozygous SNPs accumulated prior (grey) and after lineage separation (green and blue). Black bars in (c) indicate an excess of homozygous alternative genotypes separating the population from SA from the populations from H and KF.

The number of polymorphic sites with genotypes present in all individuals and with a genotype coverage ≥ 10 in all individuals derived from RNASeq data was limited. Therefore, we tested for overrepresentation of heterozygous variants shared among individuals within one population at the genome-wide scale using resequencing data of four individuals of *O. nova* and five individuals of *O. subpectinata*. In *O. nova* there was an excess of homozygous SNPs that separated individual one from the three remaining individuals (black bar; Fig. 4a) indicating that this individual represents a different genetic lineage that coexists within the same geographical population. Additionally, there was an excess of heterozygous SNPs shared between individuals two, three and four (grey bars; Figs 4a, b). In *O. subpectinata* there was an excess of homozygous SNPs separating individuals two and five from the three remaining individuals (black bars; Fig. 4c) also indicating the presence of strongly divergent lineages within the geographical population. For the GPSFSs of *O. nova* comprising four and three individuals and the GPSFS of *O. subpectinata* with detailed genotype assignment, see Supplementary Figures 6-8).

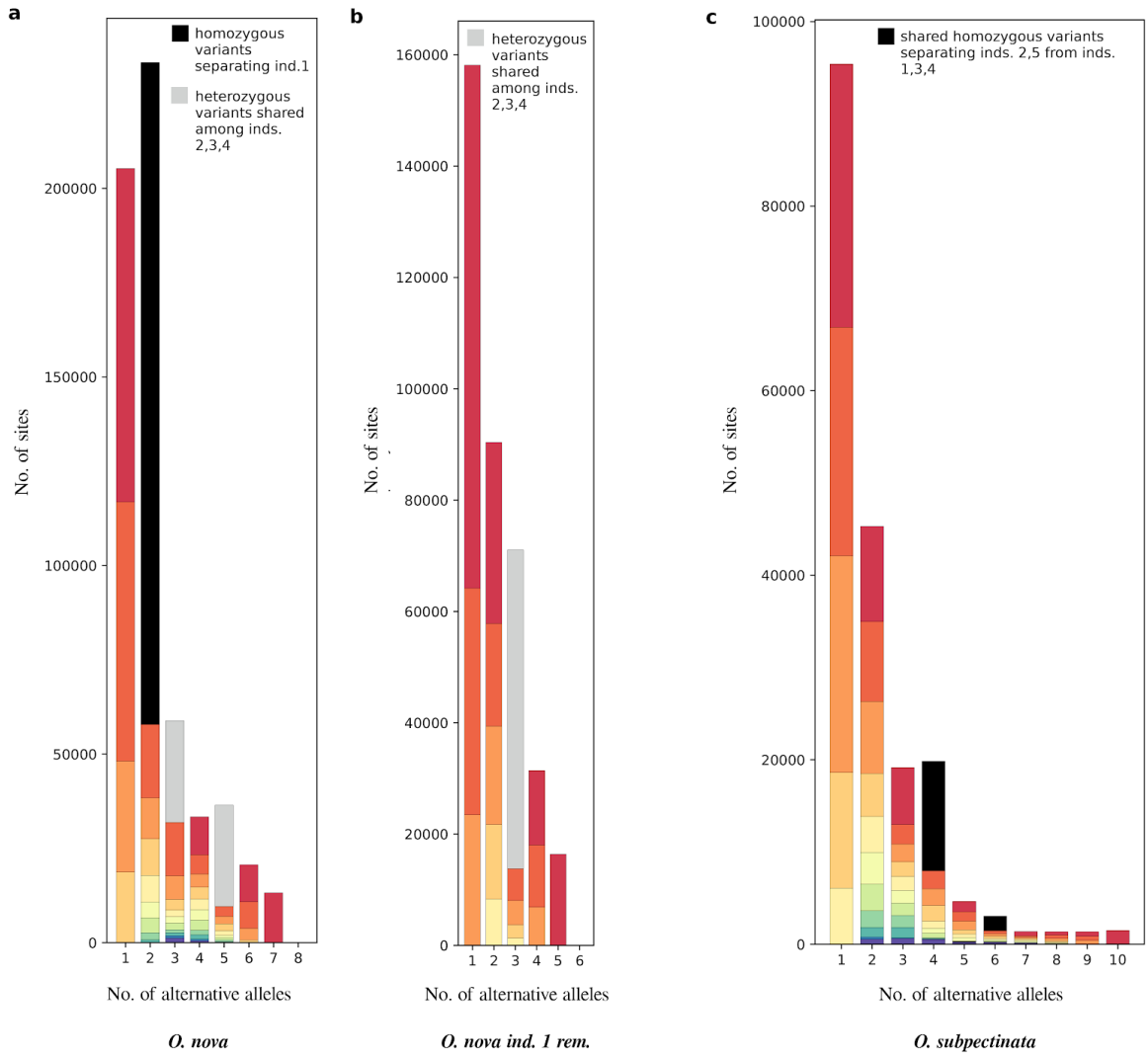


Figure 4: Genotype profile site frequency spectra (GPSFSs) based on resequencing data of four and three individuals from one population (KF) of *O. nova* and five individuals from one population (GF) of *O. subpectinata*. The black coloured bar in *O. nova* (a) shows excess of homozygous SNPs separating individual one from the remaining three individuals. The grey coloured bars indicate excess of heterozygous SNPs shared among the remaining individuals. For better representation of the overrepresentation of shared heterozygous SNPs, individual one was removed prior to construction of the GPSFS (b). The black coloured bars in the GPSFS of *O. subpectinata* indicate excess of homozygous SNPs separating individuals two and five from the remaining three individuals (c).

(III) Separation of haplotypes in haplotype phylogenies

Another consequence of HD predating population separation is the separation of haplotypes A and B at the base of a haplotype phylogeny as compared to the separation of populations expected

for panmictic sexual organisms. To test this, we first phased genomic regions using RNAseq polymorphism data from the two genetic lineages of *O. nova* (seven individuals) and the different geographic populations of *O. subpectinata* (Fig. 2). A total of 386 regions could be phased for *O. nova* and 470 regions for *O. subpectinata*. To exclude the possibility of erroneously comparing paralogs instead of haplotypes in our analyses, we tested for genomic coverage differences (see Methods). While identified single-copy BUSCO genes, phased regions and scaffolds (from which phased regions derive) had all very comparable coverage, we found two-fold coverage of identified duplicated BUSCO genes (Supplementary Figure 9). This indicates that the phased regions we base all following analyses on are not resulting from paralogs, but are indeed haplotypes.

We next aligned the phased haplotype sequences and calculated best fitting ML trees and scored the trees distances (combined distance score including topology and branch length differences; Kuhner & Felsenstein, 1994) of these ML trees to two constrained trees: (1) separating haplotypes and, hence, consistent with HD (HD-tree; Fig. 1a), and (2) separating populations as consistent with sex (sex-tree; Fig. 1b; for detailed information on phaseable regions, see Supplementary Table S5 and Methods). The delta of the two scores is indicative of a phaseable region being more consistent with HD ($\mathbf{\Delta}_{\text{dist. HD-tree - dist. sex-tree}} < 0$) or sex ($\mathbf{\Delta}_{\text{dist. HD-tree - dist. sex-tree}} > 0$). The MDS plot of *O. nova* indicated genetic separation of two coexisting lineages of three and four individuals sharing heterozygous variants rather than separation by geographical population (see Figs 2a, 3a). Therefore, we restricted the analysis to these seven individuals and accounted for lineage separation in the constrained trees (see Methods for additional information). The comparison of tree distances is based on 280 phaseable regions for which reconstruction of constrained trees was possible for *O. nova* (combined length 210,375 bp, median length 473 bp) and 393 regions for *O. subpectinata* (combined length 410,610 bp, median length 628 bp; for

detailed information, see Supplementary Table S5). For *O. nova*, 231 of 280 best fitting ML trees (82.5%) were more consistent with their respective HD-trees as compared to only 34 of 393 trees (8.7%) for the sexual *O. subpectinata* (Fig. 5). This is supported by tree topology tests showing 166 of 280 phaseable regions being significantly more consistent with a haplotype- and lineage-separating tree in *O. nova* but only seven of 393 phaseable regions in *O. subpectinata* (for details, see Supplementary Table S6, Methods).

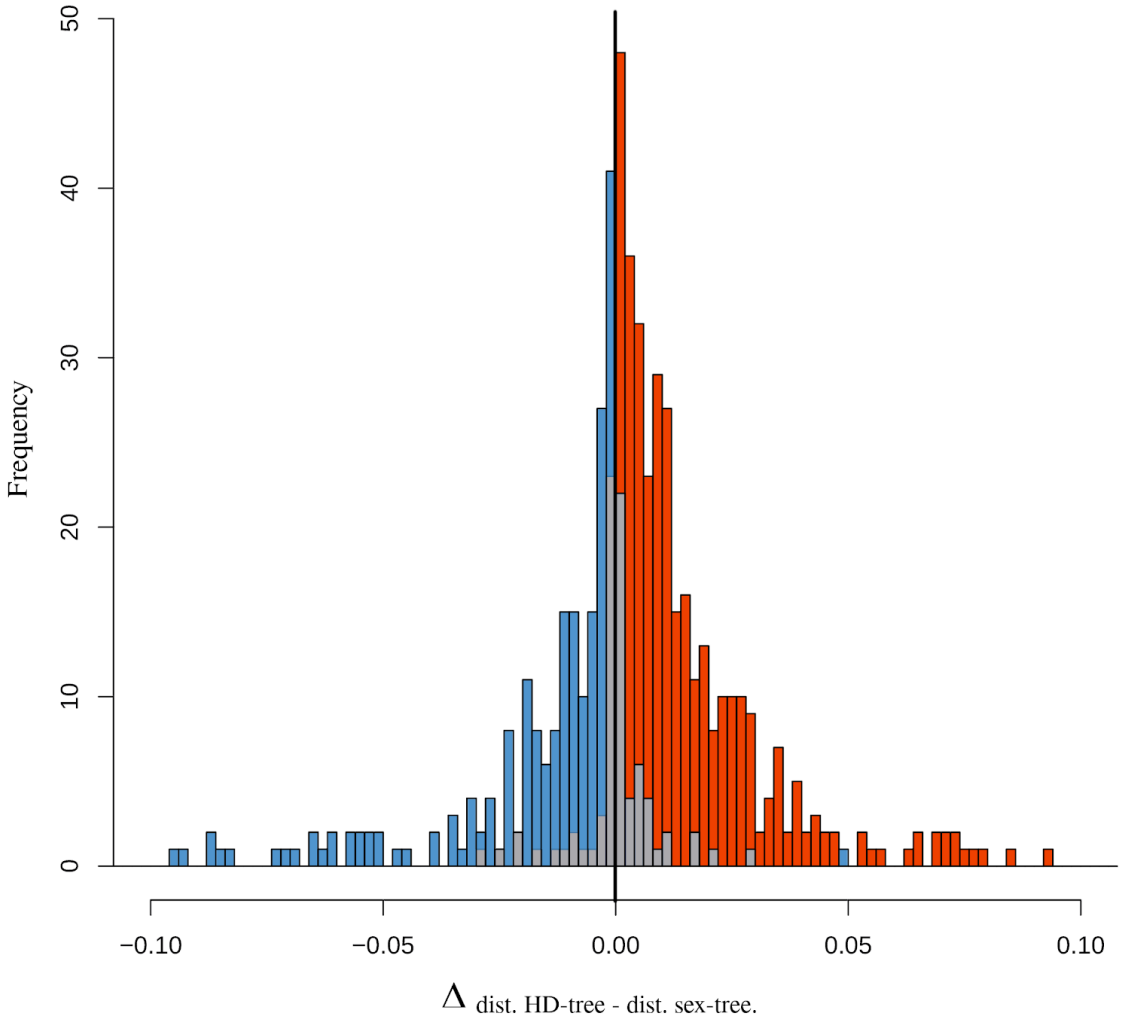


Figure 5: Frequency distribution of per-region tree-distance score comparisons. The score measures the combined distance in topology and branch lengths between two trees. For each phaseable region we calculated two scores measuring the distance of a best fitting ML tree to a constrained tree separating haplotypes and, hence, consistent with HD (HD-tree) and to a constrained tree separating populations as consistent with sex (sex-tree). A negative delta of the two scores ($\Delta_{\text{dist. HD-tree - dist. sex-tree}} < 0$; zero indicated by black line) indicates that a phaseable region's best fitting ML tree is more similar to the HD- than to the sex-tree. This is true for the majority (231) of 280 phaseable regions (82.5%) in *O. nova* (blue) whereas in *O. subpectinata* (red) only 34 of 393 phaseable regions (8.7%) are more similar to the haplotype-separating tree (overlapping frequencies in grey). To improve viewer-friendliness, the histogram range is limited to −0.1 to 0.1, thereby excluding 25 regions below and four regions above this range for *O. nova* and twelve regions above the range for *O. subpectinata*.

(IV) Parallel divergence of haplotypes

Under obligate asexuality haplotypes are expected to evolve in parallel which leads to topological resemblance of haplotype subtrees (see Fig. 1a). We reconstructed haplotype-specific (A & B) phylogenetic trees for each of the 166 phaseable regions that had been found to be significantly more consistent with a haplotype- and lineage-separating tree (see above). We based the trees exclusively on heterozygous SNPs and then tested whether haplotype tree-topologies A and B were more similar than expected by chance for each phaseable region using co-phylogeny analyses. We restricted the analyses to phaseable regions with a mean number of pairwise differences > 4 in both haplotype alignments A and B (each comprising haplotype sequences of seven individuals) because co-phylogeny analyses were rarely informative below this threshold due to an increased number of polytomies in the resulting trees. Accordingly, nine out of 166 phaseable regions were analysed, of which seven regions were statistically more similar than expected by chance and five regions showed parallel divergence over all branches (perfect parallel divergence; see Supplementary Table S7, Methods). The most polymorphic phaseable region (mean number of pairwise differences haplotype A 18.57, haplotype B 31.05, length 3,919 bp) showed perfect parallel divergence (see Fig. 6a). The other four phaseable regions with perfect parallel divergence of haplotype trees included polytomies and low bootstrap support at some nodes, likely as a result of the small numbers of informative sites (see Supplementary Figure 10). Also, co-phylogenies of the five regions with perfect parallel divergence differed in topology (but note that each topology separated the two diverged lineages as described above). To support the findings of the co-phylogeny analysis, and to use the information of all available phaseable regions, we tested parallel divergence of two “species” trees that combine the phylogenies of the 166 haplotype trees (“gene” trees; mean number of pairwise differences summed over all 166 regions haplotype A 248.57; haplotype B 876.57). Combining haplotype trees of different phaseable regions is potentially problematic because it is uncertain whether

haplotype assignment reflects “true biological” haplotypes, i.e. it is uncertain whether haplotype A from one region stems from the same homologous chromosome as haplotype A from a different region. If the “true biological” haplotypes A and B differed in the topology of their haplotype trees, combining them into “species” trees will result in haplotype tree conflict around most branches. To test for haplotype tree conflict, we estimated the percentage of quartet subtrees (all possible subtrees of four taxa) in the set of 166 haplotype trees that support an internal branch of the “species” tree (quartet support values). Quartet support values were high (see Fig. 6b), indicating an absence of gene tree topology conflict around most branches and, hence, that uncertain haplotype assignment had little influence on the topology of “species” trees of haplotypes A and B. Haplotype “species” trees were more similar than expected by chance (random tip assignment $P = 0.007$; random parasite tree $P = 0.002$). However, parallel divergence was restricted to five of six nodes (but note very low quartet support values at the conflicting branches; Fig. 6b). With the exception of the conflicting branch, divergence of the “species” trees corresponded to the divergence of the most informative region (see Figs 6a, b). Additionally in both trees, haplotypes formed clades corresponding to the two genetic lineages with an overrepresentation of shared heterozygous SNPs (see Figs 2a, 3a, b).

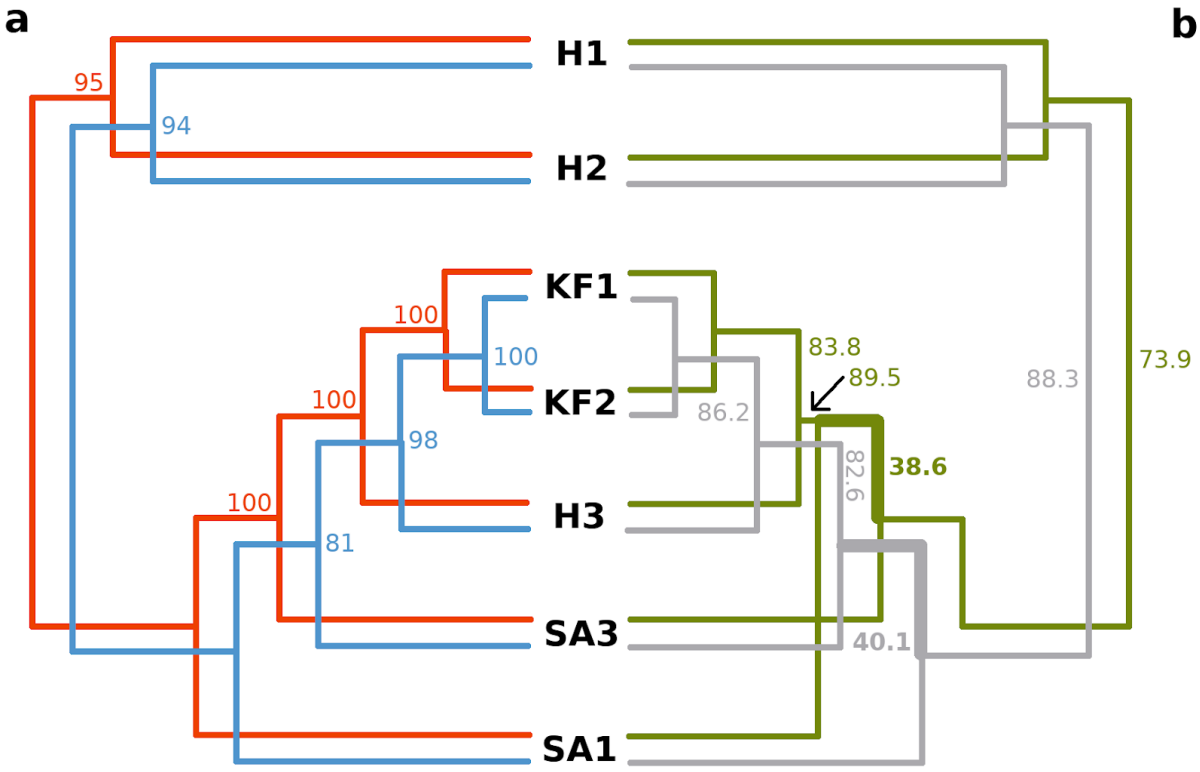


Figure 6: Results of the co-phylogeny analysis of haplotype trees for the most polymorphic phaseable region (a) and haplotype "species" trees (b). For the most informative phaseable region haplotypes A (blue) and B (red) diverged in parallel. Haplotype "species" trees (haplotype A grey; haplotype B green) diverged in parallel with the exception of one branch (highlighted in bold, however, note the low quartet support values). The parallel divergence pattern of the most informative phaseable region fits the "species" trees with the exception of the conflicting branch. For (a), bootstrap support is indicated at nodes, for (b) quartet support values are indicated at branches (ranging from 0 to 100). H: Hainich; KF: Kranichstein forest; SA: Schwäbische Alb.

Genome-wide heterozygosity

Species that persisted in the absence of sex for a very long time should display increased levels of genome-wide heterozygosity. We estimated genome-wide heterozygosity (GWH) per individual, using polymorphism data (SNPs and indels) generated from genomic resequencing reads of four individuals of *O. nova* and five of *O. subpectinata* that were sampled from the same population as the individual used to generate the reference genome. There was no difference in mean GWH between reproductive modes (*O. nova* 0.387%, *O. subpectinata* 0.303%; Mann-Whitney *U*-test *W*

= 11, *P* = 0.9; see Fig. 7). The GWHs estimated for the asexual individuals are in the typical range of asexual animals with a non-hybrid origin (Jaron *et al.*, 2019). However, GWH differed strongly among individuals in both the sexual and the asexual species, consistent with previous analyses in oribatid mites based on few allozyme markers (Palmer & Norton, 1992; Fig. 7). The differences in GWH were consistent with genetic differences among individuals as displayed in an MDS plot (see Fig. 7), and the GPSFSs of resequencing data (see Fig. 4), indicating coexistence of differently heterozygous, strongly divergent lineages within the same population for both species.

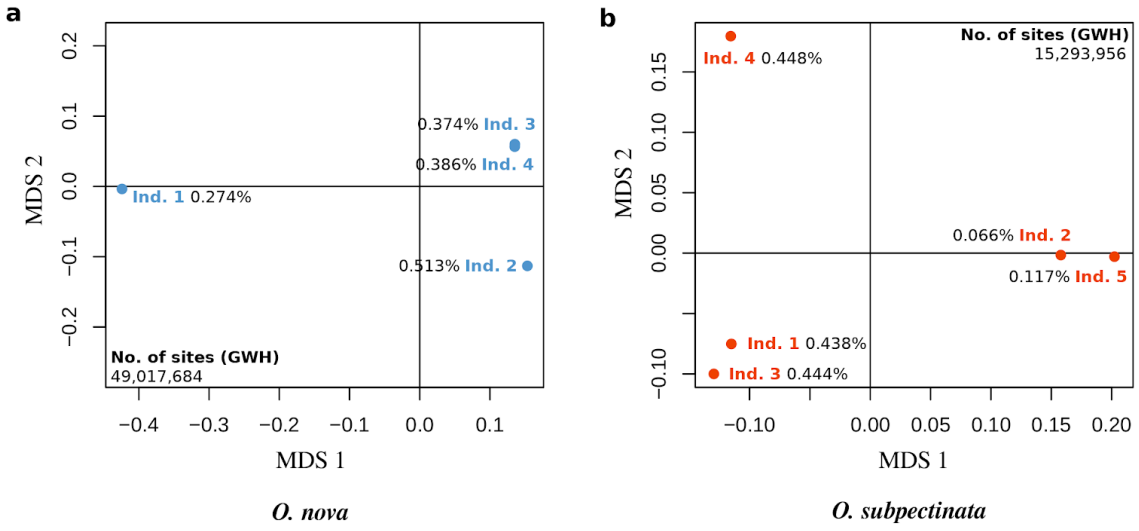


Figure 7: MDS plots and genome-wide heterozygosity (GWH) levels for the asex-sex species pair *O. nova* (a) - *O. subpectinata* (b). The MDS plot displays large genetic differences between groups of individuals in both species (largest differences in *O. nova* between individuals one and individuals two, three, four and in *O. subpectinata* between individuals two, five and individuals one, three, four). These groups strongly differ in levels of GWH (percentages per individual given in the plots).

Discussion

Haplotype divergence due to independent mutation accumulation within lineages of diploid asexual organisms is considered as strong support for evolution under obligate asexuality. Surprisingly, empirical evidence for this ‘Meselson effect’ in parthenogenetic organisms is, as

yet, either lacking or equivocal. Here, we report population genomic signatures consistent with the presence of the Meselson effect in the ancient asexual oribatid mite species *Oppiella nova*, namely: (I) excess of between-haplotype variation over between-population variation; (II) overrepresentation of heterozygous variants shared among individuals of different lineages and geographic populations and; (III) separation of haplotypes rather than lineages in allele phylogenies. In the sexual species *O. subpectinata*, these signatures were absent. Although results of co-phylogeny analyses were partly ambiguous, this was likely due to the low number of informative sites per phaseable region and overall they were supportive of (IV) parallel divergence of haplotypes. To our knowledge this study is the first to provide strong support for evolution in the complete absence of sex over evolutionary time scales in a parthenogenetic organism.

A number of reports of haplotype divergence after transition to asexuality in different eukaryotes have been attributed to processes mimicking the Meselson effect in its outcome. Large within-individual divergences, e.g. in asexual *Meloidogyne* root-knot nematodes, *Lineus* ribbon worms, *Lasaea* marine clams, *Campeloma* freshwater snails and the aphid *Rhopalosiphum padi* are presumed to be the result of interspecific hybridisation (Taylor & Foighil, 2000; Delmotte *et al.*, 2003; Johnson, 2006; Lunt, 2008; Ament-Velásquez *et al.*, 2016). In the ancient asexual bdelloid rotifers a previous report of the Meselson effect has been attributed to divergence between paralogs following an ancient genome duplication event by later studies (Mark Welch & Meselson, 2000; Mark Welch *et al.*, 2008; Flot *et al.*, 2013). Although we cannot formally exclude that *O. nova* is an interspecific hybrid, explaining the observation of shared heterozygous SNPs among seven individuals and shared heterozygous SNPs private to each of two diverged lineages of four and three individuals with hybridisation is non-parsimonious as it cannot be explained by one hybridisation event preceding loss of sex, alone (see Figs 3a, b). Also, to our

knowledge, there are no described instances of hybridisation in oribatid mites. Moreover, previous studies erroneously compared paralogs (ohnologs) stemming e.g. from genome duplication events instead of diverged haplotypes. Our findings of haplotype divergence in *O. nova* are unlikely to be the result of such paralog comparisons, as suggested by coverage analyses (Supplementary Figure 9). In addition, the observation of similar percentages of duplicate BUSCO genes in both *O. nova* and *O. subpectinata* (see Supplementary Table S1) in the absence of haplotype divergence in *O. subpectinata* renders it unlikely that the observed overrepresentation of shared heterozygous variants in *O. nova* is due to false identification of strongly divergent paralogs as haplotypes.

Our findings of haplotype divergence are in line with a number of observations on how thelytoky is realised in oribatid mites: cytological and allozyme-based studies have suggested presence of holokinetic chromosomes (2n=18) that undergo terminal fusion automixis with an inverted sequence of meiotic divisions (Taberly, 1987; Palmer & Norton, 1992; Norton *et al.*, 1993; Wrensch *et al.*, 1994; Heethoff *et al.*, 2006; Laumann *et al.*, 2008; Bergmann *et al.*, 2018). Further, recombination has not been demonstrated in asexual oribatid mite species (Palmer & Norton, 1992; Schaefer *et al.*, 2006). Taken together, this suggests that oribatid mites are functionally apomictic and can retain heterozygosity despite terminal fusion (in regions sheltered from homogenising mechanisms). However, heterozygosity can be partly maintained, even in the absence of inverted meiosis if recombination is present (Engelstaedter, 2017).

A previous study of HD based on within-individual cloning data in oribatid mites failed to identify the Meselson effect in three asexual oribatid mite species, likely due to limitations of data and methods (Schaefer *et al.*, 2006). First, *O. nova* is only very distantly related to the species analysed by the previous study (Schaefer *et al.*, 2010). Also, the previous study was based on partial sequences of only two nuclear genes (*ef1α, hsp82*), and absence of the Meselson effect

in parts of the genome is in line with our findings of varying levels of haplotype divergence among phaseable regions in *O. nova*. Finally, per species, the authors analysed individuals from only a single population. Generating conclusive evidence of the Meselson effect, however, requires individuals from multiple divergent populations or lineages, to avoid a sampling bias towards heterozygous individuals that share the same haplotypes by chance. Notably, *Nothrus palustris*, the only parthenogenetic species with individuals from two geographic populations did show phylogenetic separation of haplotypes over populations as consistent with the Meselson effect in Schaefer *et al.* (2006).

Besides haplotype divergence in *O. nova*, our study indicates the presence of coexisting, strongly diverged, and differently heterozygous lineages in both *O. nova* and *O. subpectinata*. Coexistence of strongly diverged lineages in *O. nova* has been shown previously based on the mitochondrial gene COI (separation of lineages was estimated to have occurred 6-16 mya) and was attributed to frozen niche variation between forest and grassland genotypes (Von Saltzwedel *et al.*, 2014). The occurrence of *O. nova* over a wide variety of habitat types and its cosmopolitan distribution (Subias, 2004) is also consistent with the existence of strongly divergent lineages. Generally, coexistence of strongly divergent lineages is a common phenomenon in oribatid mites. The sexual species *Steganacarus magnus* exhibits COI nucleotide divergences exceeding 25% within geographical populations (Rosenberger *et al.*, 2013). Similarly, the morphospecies *Cymbaeremaeus cymba* has been reported to represent a species complex of eight genetically distinguishable “cryptic” species (Schaeffer *et al.*, 2019).

Differences in heterozygosity between lineages are indicative of varying levels of homogenising forces acting among lineages. While the observed loss of heterozygosity between two lineages of the sexual *O. subpectinata* might reflect reproductive isolation and inbreeding, different heterozygosity levels in the asexual *O. nova* require an explanation at the within-individual level,

such as gene conversion and mitotic recombination (Birky, 1996; Marais, 2003). Gene conversion, e.g., has been assumed to account for loss of heterozygosity in the darwinulid ostracod *Darwinula stevensoni* and aphids of the tribe *Tramini* (Normark, 1999; Schoen *et al.*, 2003).

Taken together, we interpret our results as strongly supportive of genome evolution in the complete absence of sex over evolutionary times in the asexual oribatid mite species *O. nova*. This is in line with previous studies which have shown that oribatid mites are able to overcome some major predicted selective disadvantages of asexuality. Unlike other asexual animal taxa their genomes showed reduced accumulation of slightly deleterious mutations compared to their sexual relatives, possibly facilitated by large population sizes (Maraun *et al.*, 2012; Hartfield, 2016; Brandt *et al.*, 2017; Bast *et al.*, 2018). Also, similar to other asexual organisms, oribatid mites showed no increased abundance and activity of transposable elements compared to sexuals (Bast *et al.*, 2016; Bast *et al.*, 2019). The findings suggest that asexual oribatid mites indeed escaped the dead-end fate usually associated with asexual lineages. In conclusion, *O. nova* truly appears to be an “ancient asexual scandal”, hence placing oribatid mites into the spotlight as invaluable model organisms to identify the adaptive value of sex.

Methods

Animal sampling and DNA/RNA extraction

Animals were sampled in fall of 2015 and 2017 from leaf litter and soil of four different forest sites in Germany (Göttingen forest (GF), Hainich (H), Kranichstein forest (KF), Schwäbische Alb (SA); for details see Supplementary Table S8). Living animals were separated from leaf litter with heat gradient extraction (Kempson *et al.*, 1963) and identified (Weigmann, 2006), followed by at least one week of starving to reduce potential contaminants derived from gut-contents.

Afterwards, animals were cleaned by removing surface particles in sterile water, several minutes of washing in a solution of hexane:bleach:detergent:water (25:25:1:49) and rinsing with sterile water before extraction. Note that animals were alive after cleaning.

For generating reference genomes and genomic resequencing data, DNA was extracted from six single individuals of *O. subpectinata* collected in 2015 from GF leaf litter and *O. nova* collected in 2015 from KF leaf litter (for an overview see Supplementary Table S9) using the QIAamp DNA Micro kit according to manufacturer's instructions. To generate transcriptomes for annotation of reference genomes, RNA was extracted from five pooled individuals per species from the same collection batch. For this, individuals were frozen in liquid nitrogen and, after addition of Trizol (Life Technologies), mechanically crushed with beads (Sigmund Lindner). Next, Chloroform and Ethanol-treatment was applied to the homogenised tissue and the aqueous layer transferred to RNeasy MinElute Columns (Qiagen). Subsequent steps of RNA extraction were done following the RNeasy Mini Kit protocol, including DNAse digestion. Finally, RNA was eluted into water and stored at -80°C. To infer haplotype divergence (HD), RNA was extracted from single individuals of *O. nova* and *O. subpectinata* from H, SA and KF (three individuals from each forest site for each species; see Supplementary Table S8) from the 2017 collection batch. RNA extraction was done as described above for transcriptome generation. DNA and RNA quantity and quality was measured using NanoDrop (Thermo Scientific) and Bioanalyzer (Agilent).

Reference genome assemblies, contaminant removal and annotation

For genome sequencing, extracted DNA from single individuals was amplified using the REPLI-g kit (Qiagen) and sequenced on an Illumina HiSeq2500 platform with insert sizes of 180 bp, 350 bp and 550 bp for paired reads and 3000 bp for mate pairs with read sizes of 125 bp at Fasteris facilities. This resulted in a total number of $387*10^6$ reads for *O. subpectinata* and

$451*10^6$ reads for *O. nova* with a total read coverage of 420- and 490-fold per sample (for details see Supplementary Table S2). Read quality trimming and adapter clipping of paired reads was done using Trimmomatic v0.36 (Bolger *et al.*, 2014) with the following options: ILLUMINACLIP:/all-PE.fa:2:30:10 LEADING:20 TRAILING:20 SLIDINGWINDOW:3:20 MINLEN:100. This resulted in 46% and 56% surviving read pairs (for details see Supplementary Table S2). For mate pair quality trimming, Nxtrim v0.4.1 (O'Connell *et al.*, 2015) with options --separate --preserve-mp --minlength 40, followed by Trimmomatic v0.36 with options ILLUMINACLIP:/all-PE.fa:2:30:10 LEADING:20 TRAILING:20 SLIDINGWINDOW:4:20 MINLEN:60 were used to identify properly paired reads and to remove low quality bases and adapters. This resulted in 48% to 54% surviving read pairs (for details see Supplementary Table S2).

With the available read data, we tested a range of assembly strategies (see GitHub repository for details; https://github.com/AsexGenomeEvol/ASD_mites/). The best assemblies were generated using normalised, overlapped reads, because whole genome amplification introduces overrepresented genomic regions, which leads to coverage bias that is problematic for assembly. Overlapped read libraries were generated by merging the paired forward and reverse reads of the 180 bp read libraries and additionally merging unpaired reads, followed by normalisation using BBnorm v37.82 (Bushnell, 2014). These normalized, overlap read libraries were assembled into contigs using SPAdes v3.10.1, a multi k-mer, haplotype-aware assembler (Bankevich *et al.*, 2012), with options -m 400 --careful -k 21, 33, 55, 77, 99, 111, 127. The resulting haplotype-resolved contigs were ordered into scaffolds using the 350 bp, 500 bp and 3000 bp read libraries with SSPACE v3.0 (Boetzer *et al.*, 2011) and option -z 200. To close gaps emerging during scaffolding, GapCloser v1.12 (Luo *et al.*, 2012) with option -l 125 was run.

Scaffolds that are likely a source from contaminants (e.g. bacteria, fungi) were removed by first annotating and visualising contaminations using BlobTools v1.0 (Laetsch & Blaxter, 2017), followed by custom filtering. For this, coverage of each scaffold was estimated by mapping reads back to the scaffolds using BWA mem v0.7.15 (Li & Durbin, 2010) and coverage calculated with BBTools v73.82 (Bushnell, 2014). Additionally, for annotation, scaffolds were blasted using ncbi-blast v2.7.1+ blastn with options -outfmt '6 qseqid staxids bitscore evalue std sscinames sskingdoms stitle' -max_target_seqs 10 -max_hsps 1 -evalue 1e-25, after which BlobTools-plots and -tables that list all hits were generated. Contaminants were identified by filtering the lists containing scaffold information and ‘best hit’ annotation and removing entries that have one or more eukaryote phylum hits using a custom script (see GitHub for details: https://github.com/AsexGenomeEvol/ASD_mites/). Scaffolds were removed that only annotated to identified contaminants. Next, scaffolds were sorted by decreasing length, scaffold headers renamed and scaffolds shorter than 500 bp removed, resulting in the final assemblies (v03). The assemblies were checked for quality and completeness calculating standard genome statistics and by checking presence, fragmentation and duplication of arthropod core genes using CEGMA v2.5 and BUSCO v3.0.2 (Parra *et al.*, 2007; Simão *et al.*, 2015). For details see Supplementary Table S1.

Genome annotation

The de-contaminated genome assemblies (v03) were annotated using MAKER v2.31.8 (Holt & Yandell, 2011), a hybrid *de novo* evolution-based and transcript-based method (for details see GitHub: https://github.com/AsexGenomeEvol/ASD_mites/). For this, repetitive genomic regions were first masked using RepeatMasker v4.0.7 (Smit *et al.*, 2013-2015) as implemented in MAKER. Protein coding genes were then predicted in a 2-iterative way described in Campbell *et al.* (2014) with minor modifications following author recommendations. For the first iteration,

genes were predicted using Augustus v3.2.3 (Stanke *et al.*, 2006) trained with the BUSCO v3.0.2 results (arthropoda_odb9 lineage with the -long option). A combination of UniProtKB/Swiss-Prot (release 2018_01) and the BUSCO arthropoda_odb9 proteome were used as protein evidence. The Trinity assembled mRNA-seq sequences (described below) were used as transcript evidence. The resulting gene models from iteration 1 were then used to retrain Augustus as well as SNAP v2013.11.29 (Korf, 2004) and a second iteration was performed. Following, predicted protein coding genes were functionally annotated using Blast2GO v5.5.1 (Conesa *et al.*, 2005) with default parameters against the NCBI non-redundant arthropod protein database. The MAKER configuration files will be made available at https://github.com/AsexGenomeEvol/ASD_mites/.

For the MAKER annotation, RNAseq reads were quality trimmed with Trimmomatic v0.36 with options adapters.fa:2:30:12:1:true LEADING:3 TRAILING:3 MAXINFO:40:0.4 MINLEN:80. For generating genome-guided transcriptome assemblies, trimmed reads were first mapped against the genomes using STAR v2.5.3a (Dobin *et al.*, 2013) under the '2-pass mapping' mode and default parameters. Following, the outputs were used with Trinity v2.5.1 (Haas *et al.*, 2013) set to 'genome guided' mode (Parameters: --genome_guided_max_intron 100000 --SS_lib_type RF --jaccard_clip). For quality filtering of the resulting transcriptomes, the trimmed RNAseq reads were mapped back against the transcriptomes using Kallisto v0.43.1 (Bray *et al.*, 2016) with options --bias and --rf-stranded then transcripts with at least 1 TPM in any samples were retained.

Reseq: sequencing, variant calling and processing

Similarly to genome sequencing, extracted DNA from single individuals was amplified using the REPLI-g kit (Qiagen) and sequenced on an Illumina HiSeq2500 platform with read sizes of 125 bp at Fasteris facilities. For each species, resequenced reads of every individual were trimmed using Trimmomatic v0.36 and mapped onto the newly assembled genome of reference using

BWA mem version 0.7.12. Mapping files corresponding to paired-end were then sorted and filtered (i.e. minimal mapping quality = 20) using samtools v1.9 (Li *et al.*, 2009) and duplicates were removed using Picard v2.10.5 MarkDuplicates (Broad Institute, 2020). Individual haplotypes were called using GATK HaplotypeCaller version 3.8 (McKenna *et al.*, 2010), with parameters as follows: --emitRefConfidence GVCF --genotyping_mode DISCOVERY -hets 0.001) and were subsequently transformed into multi-individual SNP calls for each species using GATK GenotypeGVCFs (default parameters). Following GATK best practice guides, we visualized the distribution of SNP characteristics in order to set the following list of hard thresholds that were used within GATK VariantFiltration to set a list of good quality calls (thresholds as follows: QD < 3; FS > 20; SOR > 4; MQ < 57; MQRankSum < -0.3; ReadPosRankSum < -2). These putative known SNPs were used for quality base recalibration using GATK BaseRecalibrator. Quality base recalibrated mapping files then underwent the same initial step as above (GATK HaplotypeCaller, GenotypeGVCFs and VariantFiltration, with an additional filter: ExcessHet > 10).

Resequencing: genome-wide heterozygosity, GPSFS, and MDS

To reduce the amount of missing sites for calculation of genome-wide heterozygosity, Genotype Profile Site Frequency Spectra (GPSFS), and for multi-dimensional scaling analysis (MDS) one individual of *O. nova* with > 50% of missing sites after VariantFiltration was excluded *a priori*. Calculation of the individual`s genome-wide heterozygosity was based on variants that had passed VariantFiltration and for which genotype information was present and genotype coverage was > 10 in all individuals using vcftools v0.1.15 (Danecek *et al.*, 2011). Construction of GPSFS's was done using Pop-Con with standard parameters and based on SNPs only (Anselmetti, 2019). For multi-dimensional scaling (MDS), sites monomorphic for the reference variant and polyallelic sites were excluded, additionally, using vcftools. Multi-dimensional

scaling (two scales for two-dimensional representation) was done using plink v1.90b6.10 (Purcell *et al.*, 2007) and resulting coordinates plotted in R (R Core Team, 2013). All commands used will be made available under https://github.com/AsexGenomeEvol/ASD_mites/.

HD: RNAseq, quality control and mapping

RNA extracts were fragmented to 175 nt for strand-specific library preparation using the NEBNext® Ultra™ II Directional RNA Library Prep Kit. Paired-end sequencing with a read length of 100 bp was performed on a HiSeq2000 platform at the CIG (Centre of Integrative Genomics, Lausanne, Switzerland). Data processing was done using the high performance computing cluster of the Gesellschaft für Wissenschaftliche Datenverarbeitung Göttingen (GWDG). RNAseq reads were adapter- and quality-trimmed (Phred quality threshold 20; sequences < 35 bp deleted) using TrimGalore v0.6.5 (Martin, 2011; Krueger, 2012). For each individual, reads flagged as contaminating sequences were removed using kraken2 (--db minikraken2_v2; Wood & Salzberg, 2014). Surviving paired end reads per individual were mapped simultaneously against the respective species reference genome, the scaffolds flagged as contaminating sequences (identified as described above), the human reference genome GRCh38.p12 (GenBank assembly accession: GCA_000001405.27) and the human microbiome (downloaded from https://www.hmpdacc.org/hmp/HMREFG/all/index.php) using bbmap v37.66 (bbsplit; maxindel=100k; Bushnell, 2014). The portions of reads that were found to be derived from contaminating RNA of human and microbial origin ranged from 40.36% to 90.31% in *O. subpectinata* and 53.33% to 93.04% in *O. nova* (see Supplementary Table S10). Only oribatid-mite-exclusive reads, i.e. read pairs that mapped best and unambiguously against the respective mite reference genome were kept for further processing. All analyses using RNAseq data are based on these cleaned reads that were mapped to the mite reference genomes using

STAR v2.7.3a with standard parameters. All commands used will be made available under https://github.com/AsexGenomeEvol/ASD_mites/.

HD: Variant calling and processing

PCR and optical duplicates were removed from mapped reads using Picard v2.10.5 and sequences spanning intronic regions removed using GATK v4.0.3.0 SplitNCigarReads. Next, these nine modified alignments per species were merged with samtools v1.9 for subsequent SNP calling. For this, GATK HaplotypeCaller was run with standard parameters and called variants were subsequently filtered by coverage < 10 using VariantFiltration (GATK v4.0.3.0) and removed. Next, only variants were kept that were biallelic, using vcftools v0.1.15. Variants that had failed VariantFiltration were excluded and indels deleted to avoid mistaken assignment of alternatively spliced sites as indels. All commands used will be made available under https://github.com/AsexGenomeEvol/ASD_mites/.

Population differentiation and GPSFS

Population differentiation was tested based on the filtered set of SNPs with an Analysis of Molecular Variance (AMOVA) and a randomness test with the packages vcfR and poppr in R (Kamvar *et al.*, 2014; Mateus & Caeiro, 2014; Knaus & Gruenwald, 2017). To visualise genotype composition of populations, multi-dimensional scaling (MDS; two scales for two dimensional representation) was done as described above for resequencing data after removing individual genotypes with coverage < 10. Allele frequency spectra computation and visualisation of genotype profiles was done with Pop-Con utilising only genotypes with coverage >= 10.

Haplotype phasing

Phasing of alleles was done per individual using phASER v1.1.1 with minimum mapping quality of reads set to 30, minimum base quality set to 20 and bottom cut-off to quantile for alignment score set to 0 in paired-end mode for each individual, separately, utilising heterozygous variants

with minimum coverage of 10 for each individual (Castel *et al.*, 2016). Haplotype sequences of phaseable regions were reconstructed based on the reference genome using a custom script (will be made available under https://github.com/AsexGenomeEvol/ASD_mites/). First, at positions of homozygous alternative SNPs the reference base was changed to the alternative base. Additionally, as some SNPs might have been missed by HaplotypeCaller due to insufficient coverage or absence of coverage in non-coding regions, all bases of the reference genome with coverage < 10 (the coverage threshold for VariantFiltration) were excluded from further analysis (changed to N; coverage estimated with bedtools v2.26.0; Quinlan & Hall, 2010). This was done to exclude non-called homozygous reference SNPs and thereby avoid underrepresentation of homozygous alternative SNPs. Second, in regions where phasing had been successful, bases were changed to the base that had been assigned to one and the other haplotype, respectively, during prior phasing, resulting in two modified reference genomes per individual. Finally, regions where at least one individual had been phased successfully were extracted from the resulting modified reference genomes and aligned (detailed information will be made available under https://github.com/AsexGenomeEvol/ASD_mites/).

To identify the presence of paralogs in the phased regions (stemming from e.g. duplication events), these regions were tested for double coverage compared to the genomic baseline (detailed information will be made available under https://github.com/AsexGenomeEvol/ASD_mites). Reads used to assemble genomes were mapped back to single copy genes and duplicated genes identified by BUSCO (see above), and additionally to the phased regions and to scaffolds from which the phased regions were derived (but that were masked in the phased regions), using bowtie2 v2.3.4.1 with standard parameters (Langmead & Salzberg, 2012). The mapped alignments were quality filtered (MAPQ score > 10)

using samtools v1.9 and optical and PCR duplicates removed using Picard Tools Mark Duplicates v2.22.0. Following, coverage was calculated using bedtools genomecov v2.26.0.

Topology testing

To enable testing whether alignments of phaseable regions are better explained by a topology separating the haplotypes (as expected with HD) as compared to a topology separating populations (expected under sexual reproduction) or vice versa, a constrained tree search was done. Because assignment of haplotype names during phasing with phASER does not necessarily reflect the distance of the haplotype to the reference genome, and (assuming complete separation of haplotypes) to ensure that those haplotypes of different individuals that cluster together are labeled alike, the haplotypes were renamed according to their distance to the reference genome. For this, positions including degenerate bases were deleted, the pairwise distance of each haplotype to the reference genome was calculated using snp-dist (Seeman, 2020), and haplotypes were renamed accordingly (the haplotype with the smaller distance to the reference genome was assigned haplotype A). Afterwards, two constrained Maximum Likelihood (ML) trees, one complying with a fixed haplotype-separating topology (HD-tree), the other with a fixed population-separating topology (sex-tree) were calculated for each phased region using iqtree v1.6.10 with 1000 bootstrap replicates and model-testing included (Nguyen *et al.*, 2015). As the MDS plot of *O. nova* indicated genetic separation of two coexisting lineages of three and four individuals sharing heterozygous variants rather than population separation (see Figs 2a, 3a), we restricted the analysis to these seven individuals. Further, we constrained the HD-trees of *O. nova* to separate haplotypes at the base and the two lineages visible in the MDS plots and GPSFSs within each haplotype subtree. We subsequently constrained the sex-tree of *O. nova* to separate lineages at the base and geographical populations within each lineage subtree. Next, the combined distance in topology and branch lengths between the two resulting trees to an

unrestricted best fitting ML tree (calculated alike but without enforcing a fixed topology) was estimated according to Kuhner & Felsenstein (1994) with the dist.topo function implemented in the R package ape (Paradis & Schliep, 2019). To enable the comparison between distances of different phaseable regions, the topological distances of the best fitting ML tree to the haplotype-separating tree and to the population-separating tree were combined as difference (Delta value) for each phaseable region. To test if the haplotype-separating tree (including consideration of branch length estimates) is a significantly better fit to the alignment than the population-separating tree, we compared them using RELL approximation with 10,000 RELL replicates and an approximately unbiased (AU) test with iqtree (Kishino *et al.*, 1990; Shimodaira, 2002). For detailed information, see https://github.com/AsexGenomeEvol/ASD_mites/.

Parallel divergence testing

To test if haplotypes of *O. nova* evolved in parallel, haplotype sequences were reconstructed as described above using exclusively heterozygous SNPs, and named according to their distance to the reference genome. For each phaseable region, best fitting ML trees were reconstructed for haplotypes A and B using iqtree including testing for nucleotide substitution models and allowing for branch lengths of zero. Afterwards, branch lengths of zero were resolved to allow for polytomies using the function di2multi implemented in the R package phytools (Revell, 2012). Following, both haplotype trees per phaseable region were analysed for co-phylogeny using Jane (Conow *et al.*, 2010) with default cost settings following Weir *et al.* (2016). Co-phylogeny testing was restricted to phaseable regions where the mean number of pairwise distances among the seven sequences of both haplotype A or B were larger than four. Obtained results were tested for robustness, i.e. whether haplotype phylogenies were more similar than expected by chance, using a sample size of 1000 and two randomisation methods: Random tip mapping and random parasite tree building (beta = -1.0). Species trees were reconstructed based on all 166 polytomy-resolved

trees of haplotypes A and B, separately with quartet support scoring (-t 1) using Astral 3 (Zhang *et al.*, 2018).

Code availability

All scripts and commands as referred to in the article will be made available at https://github.com/AsexGenomeEvol/ASD_mites upon submission.

Data availability

Reference genomes of the four species generated during this study will be deposited in NCBI upon submission. Resequencing reads of five individuals per species will be deposited in NCBI upon submission. Transcriptomes and RNAseq reads of the four species used for annotation will be deposited in NCBI upon submission. RNAseq reads of nine individuals per species used for inferring allelic sequence divergence will be deposited in NCBI upon submission. Further processed data supporting the study will be made available at DRYAD upon submission.

Acknowledgements

We like to thank Juergen Brandt for adjusting figures to the needs of colour-blind people. This study was supported by core funding of S.S., by DFG research fellowship BA 5800/1-1 and DFG Emmy Nöther fellowship BA 5800/3-1 to J.B. and Swiss SNF grant PP00P3_170627 to T.S.

Contributions

A.B., J.B. and T.S. conceived and designed the study. A.B. and C.B. collected samples and determined species. J.B., Z.D. and M.L. performed wet lab work. A.B., J.B., P.T-V., Y.A., P.S. performed data analysis. T.S., N.G., M.R.R., K.S.J., C.M.F., M.M., B.H., I.S. and S.S. contributed to data interpretation and analyses and A.B. and J.B. wrote the paper with input from all authors.

III General Discussion

Asexual species are indispensable for generating insight into the selective advantage of sex: similar to a geneticist knocking out a gene to study its function, the evolutionary biologist studies asexual species to tell why sex is maintained despite its manifold costs (Maderspacher, 2011). And just like the geneticist who studies several different knock-out organisms to test if the gene's function is similar in many organisms, the evolutionary biologist analyses many different "natural sex knock-out replicates" to test if one selective advantage can commonly account for the maintenance of sex.

This thesis put to the test whether the selective advantage of sex lies in enabling the purging of slightly deleterious mutations via effective purifying selection (Muller, 1964; Hill & Robertson, 1966; Felsenstein, 1974; Kondrashov, 1988). The "knock-out replicates" were asexual *Timema* stick insects, oribatid mites and eight hexapod species. Asexual *Timema* stick insects, indeed, accumulated slightly deleterious mutations as compared to their sexual sister species (Chapter 2; Bast *et al.*, 2018) and, hence, support the above mentioned prediction. Contrastingly, purifying selection in asexual oribatid mites was even more effective than in their sexual relatives (Chapter 1; Brandt *et al.*, 2017). The eight asexual hexapod species did not provide a clear picture: although there was no difference when taken together, some did and some did not accumulate slightly deleterious mutations (Chapter 3; Brandt *et al.*, 2019).

Mutational decay in asexuals

The observed effective purifying selection in asexual oribatid mites indicates that accumulation of slightly deleterious mutations is no universal feature of asexual animals. However, deleterious mutation accumulation may be an important factor for many other asexual animals. Table I lists significant earlier studies, studies generated by others during the period of this thesis (~2015 -

2019) and studies generated within the framework of this thesis. It serves as basis for the following discussion.

Table I: Single gene-based and genome-based studies that include inference of accumulation of slightly deleterious mutations in asexual animals. The table is based on Hartfield (2016), Neiman *et al.* (2018), Glémin *et al.* (2019) and Jaron *et al.* (2019) and contains earlier studies, studies generated within the framework of this thesis (marked with *) and studies generated by others during the period of this thesis (~2015 - 2019). In case of the presence of multiple studies on the same species and of the same category (single genes / genome-based) with similar results, one representative study was chosen. Studies are sorted by year of publication. **pur. sel.:** purifying selection; **green:** support for accumulation of slightly deleterious mutations in asexuals; **red:** no support; **light red:** support but mutations inherited from parental species of hybrid asexuals; **orange:** ambiguous/differing results.

study	system	group	mode of parthenogenesis	pur. sel. single genes	pur. sel. genome-based
Normark & Moran, 2000	Aphididae spp.	aphids	apomixis	one lineage yes/ one no	
Mark Welch & Meselson, 2001	Bdelloidea spp.	rotifers	ameiotic	no difference	
Paland & Lynch, 2006	*Daphnia pulex*	water fleas	apomixis	less effective - hybridisation	
Barraclough *et al.*, 2007	Bdelloidea spp.	rotifers	ameiotic	less effective	
Johnson & Howard, 2007	*Campeloma* spp.	snails	apomixis	less effective	
Neiman *et al.*, 2010	*Potamopyrgus antipodarum*	snails	apomixis	less effective	
Henry *et al.*, 2012	*Timema* spp.	stick insects	apomixis	less effective	
Ollivier *et al.*, 2012	Aphididae spp.	aphids	apomixis		no difference
Tucker *et al.*, 2013	*Daphnia pulex*	water fleas	apomixis		no difference - hybridisation
Ament-Velásquez *et al.*, 2016	*Lineus pseudolacteus*	ribbon worms	fissiparous		less effective - hybridisation
Kraaijeveld *et al.*, 2016	*Leptopilina clavipes*	wasps	Wolbachia, gamete duplication		more effective
Brandt *et al.*, 2017 *	oribatid mites	oribatid mites	yes - effectively ameiotic?		more effective
Bast *et al.*, 2018 *	*Timema* spp.	stick insects	apomictic		less effective

Lindsey *et al.*, 2018	*Trichogramma pretiosum*	wasps	Wolbachia, gamete duplication		no difference
Warren *et al.*, 2018	*Poecilia formosa*	fish	gynogenetic		no difference
Leria *et al.*, 2019	*Dugesia spp.*	flat worms	fissiparous	no difference	
Brandt *et al.*, 2019 *	hexapods	hexapods	differs		no difference - differs between hexapod groups

At first sight, support for accumulation of slightly deleterious mutations is far from common in asexual animals (see Table I). While five studies reported less effective purifying selection in asexuals compared to their sexual relatives, ten studies did not (after accounting for the hybrid origin of asexuality) and two studies yielded ambiguous or differing results among analysed species. Notably, of nine genome-based studies only one study supported the theoretical prediction that the selective advantage of sex lies in effective purging of slightly deleterious mutations (Bast *et al.*, 2018). In single gene studies, the pattern is ambiguous: four out of eight showed accumulation of slightly deleterious mutations in their focal asexual species. This, however, is not surprising as a number of genome-based studies, including the studies generated within the framework of this thesis, found excessive among-gene variation in the effectiveness of purifying selection (see also Neiman *et al.*, 2018). This thesis, among others, clearly showed that interpreting the effectiveness of purifying selection acting on single genes as representative for the genome is problematic. Conclusive insight into selective advantages of sex at the genome-level needs to be based on whole genome or transcriptome studies; and those largely do not support effective purging of slightly deleterious mutations as a selective advantage of sex.

Another benefit of genome-based analyses of closely related asexual and sexual species is that they enable to distinguish if slightly deleterious mutations accumulated as a consequence of asexuality or if they were already present before sex was lost. The latter was demonstrated for obligately asexual lineages of *Daphnia pulex* (Paland & Lynch, 2006; Tucker *et al.*, 2013). Here,

slightly deleterious mutations had accumulated already in cyclically asexual lineages that later became “infected” with obligate asexuality by different obligately asexual lineages. Also, slightly deleterious mutations in the hybrid ribbon worm *Lineus pseudolacteus* were present in its sexual parent species and “transmitted” during hybridisation (Ament-Velásquez *et al.*, 2016).

Factors alleviating mutational decay

Notably, two genome-based studies in *Wolbachia*-infected asexual parasitoid wasps were consistent with other genome-based studies that found no mutational decay in effectively ameiotic asexuals (Kraaijeveld *et al.*, 2016; Lindsey *et al.*, 2018). This is surprising because in both wasp species, meiosis is retained and diploidy restored via gamete duplication which leads to complete homozygotisation of the genomes (Stouthamer & Kazmer, 1994; Pannebakker *et al.*, 2004). Homozygotisation can reduce the accumulation rate of slightly deleterious mutations by exposing recessive mutations to selection, thereby allowing for their effective removal from the gene pool (Charlesworth *et al.*, 1993b; c). The observed effective purifying selection in most of the analysed, effectively ameiotic species, hence, signals that other, meiosis-independent homogenising mechanisms, like mitotic recombination or gene conversion, might play an important role (Marais, 2003).

What other factors can allow asexual animals to evade mutational decay? Large population sizes have been discussed as an important factor maintaining effective purifying selection under asexuality (Gordo & Charlesworth, 2000; Rice & Friberg, 2009; Normark & Johnson, 2011; Ross *et al.*, 2013). Many small and widely distributed animals have potentially very large population sizes (Gaston *et al.*, 1997; White *et al.*, 2007). Indeed, census population densities of very old asexual taxa (oribatid mites and darwinulid ostracods) can exceed 10^5 individuals per square metre and often feature larger population sizes than their sexual relatives (Van Doninck *et al.*,

2003; Maraun *et al.*, 2012). In fact, *Timema* stick insects, the as yet only example of genome-wide deleterious mutation accumulation (Bast *et al.*, 2018) have restricted geographic ranges, large body sizes and low population densities (Sandoval, 1994; Law & Crespi, 2002). Future genome-wide estimates of slightly deleterious mutation accumulation in asexual animals are required to take abundances and effective population sizes into account. This will provide empirical information on the importance of population size for the long-term persistence of asexual species.

Mutational decay in ancient asexuals

A different factor predicted to influence the extent of slightly deleterious mutation accumulation is the time that has passed since sex was lost in an asexual organism. As mutation accumulation occurs gradually, asexuals that lost sex recently may show no signs of decreased effectiveness of purifying selection at all, as has been suggested for asexual aphids and the water flea *Daphnia pulex* (Ollivier *et al.*, 2012; Tucker *et al.*, 2013). However, in some animals that have persisted and diversified in the absence of sex for tens of millions of years, time seems to have limited impact on mutational decay - otherwise they should have long gone extinct by now (Judson & Normark, 1996); but note that predicting mutational decay rates requires extensive modelling approaches (Loewe & Lamatsch, 2008). These “ancient asexual scandals” classically comprised the bdelloid rotifers, darwinulid ostracods, apomictic *Meloidogyne* root-knot nematodes and parthenogenetic clades of oribatid mites (Schoen *et al.*, 2009). However, studies have shown that apomictic *Meloidogyne* root-knot nematodes were of hybrid origin and lost sex more recently than previously assumed (Lunt, 2008; Szitenberg *et al.*, 2017). Among the remainder, bdelloid rotifers have, so far, received most attention. Recent studies, however, have indicated that cryptic gene exchange might render them quasi sexual (Signorovitch *et al.*, 2015; Debortoli *et al.*, 2016;

Schwander, 2016; Vakhrusheva *et al.*, 2018). The amount of gene exchange and its mechanism remain controversial (Flot *et al.*, 2018; Wilson *et al.*, 2018).

Data on genome evolution in asexual oribatid mites and darwinulid ostracods are scarce. For asexual oribatid mites only two studies showed effective purifying selection and decreased load of transposable elements (Bast *et al.*, 2016; Brandt *et al.*, 2017), and no genome-based studies in darwinulid ostracods exist, so far. More studies on these two animal groups, especially regarding mutation accumulation and haplotype divergence in darwinulid ostracods are strongly required as these truly ancient asexual lineages are invaluable for generating insight into the long-term selective advantage of sex.

The Meselson effect

Another genomic consequence of asexuality analysed in this thesis is the Meselson effect. Homologous chromosomes in asexual organisms are expected to accumulate mutations independently of each other in regions sheltered from loss of heterozygosity and diverge in parallel which leads to high levels of heterozygosity and resemblance of haplotype subtrees over populations (Birky, 1996; Judson & Normark, 1996; Mark Welch & Meselson, 2000). Testing for the Meselson effect is important because its presence is regarded as strong support for the complete absence of sex (Normark *et al.*, 2003). Also, it opens the possibility for dating the transition to asexuality. This thesis showed the presence of the Meselson effect in the putatively ancient asexual oribatid mite species *Oppiella nova*, supporting that sex was truly absent for a long period of time (Chapter 4). To discuss whether the Meselson effect can be regarded as a common feature of asexual animals, consultation of further studies is required. Table II lists studies of major importance and their findings with regard to haplotype divergence and serves as basis for the following discussion.

Table II: Single gene-based and genome-based studies that include inference of within-individual genetic divergence in asexual animals. The table is based on Hartfield (2016), Neiman *et al.* (2018), Glémin *et al.* (2019) and Jaron *et al.* (2019) and contains earlier studies, studies generated within the framework of this thesis (marked with *) and studies generated by others during the period of this thesis (~2015 - 2019). In case of the presence of multiple studies on the same species and of the same category (single genes / genome-based) with similar results, one representative study was chosen. Studies are sorted by year of publication. **WID**: within-individual divergence; **green:** support for increased WID in asexuals; **red:** no support; **light red:** support but observed WID likely due to divergence between paralogs.

study	system	group	mode of parthenogenesis	WID single genes	WID genome-based
Normark, 1999	Tramini spp.	aphids	apomixis	no	
Mark Welch & Meselson, 2000	Bdelloidea spp.	rotifers	ameiotic	yes - ohnologs	
Taylor & Foighil, 2000	*Lasaea* spp.	clams	apomixis/obligate selfing?	yes - hybridisation	
Delmotte *et al.*, 2003	*Rhopalosiphum padi*	aphids	apomixis	yes - hybridisation	
Schoen *et al.*, 2003	*Darwinula stevensoni*	ostracods	likely apomixis	strongly homozygous	
Johnson, 2006	*Campeloma* spp.	snails	apomixis	yes - hybridisation	
Schaefer *et al.*, 2006	Desmonomata spp.	oribatid mites	yes - effectively ameiotic?	no	
Abad *et al.*, 2008	*Meloidogyne incognita*	nematodes	apomixis		yes - hybridisation
Lunt, 2008	*Meloidogyne* spp.	nematodes	apomixis	yes - hybridisation	
Schwander *et al.*, 2011	*Timema* spp.	stick insects	apomixis	yes	
Flot *et al.*, 2013	*Adineta vaga*	rotifers	ameiotic		yes - ohnologs
Ament-Velásquez *et al.*, 2016	*Lineus pseudolacteus*	ribbon worms	fissiparous		yes - hybridisation
Fradin *et al.*, 2017	*Diploscapter pachys*	nematodes	central fusion auto-mixis equivalent		yes - hybridisation
Hiraki *et al.*, 2017	*Diploscapter coronatus*	nematodes	central fusion automixis		yes - hybridisation
Szitenberg *et al.*, 2017	*Meloidogyne* spp.	nematodes	apomixis/one automixis?		yes - hybridisation
Nowell *et al.*, 2018	Bdelloidea spp.	rotifers	ameiotic		yes - ohnologs

Leria *et al.*, 2019	*Dugesia* spp.	flat worms	fissiparous	yes	
Brandt *in prep.* *	*Oppiella nova*	oribatid mites	yes - effectively ameiotic?		yes

The trojan horse that is the Meselson effect

Support for the Meselson effect in animals is even less common than for the accumulation of slightly deleterious mutations: as yet, only three out of 18 studies have reported the expected haplotype divergence pattern, including two based on single genes (*Timema* stick insects and fissiparous *Dugesia* flatworms; out of a total of 10 single gene-based studies; Schwander *et al.*, 2011; Leria *et al.*, 2019). Although large within-individual genetic divergence levels were found in a number of different invertebrates, e.g. the apomictic *Meloidogyne* root-knot nematodes and the ribbon worm *Lineus pseudolacteus*, they were (later) attributed to divergence between homeologs derived from hybridisation (Lunt, 2008; Ament-Velásquez *et al.*, 2016; Jaron *et al.*, 2019). Similarly, large within-individual divergence in bdelloid rotifer species has been shown to result from an ancient genome duplication event and reflect divergence between ohnologs instead of haplotypes (Mark Welch *et al.*, 2008; Flot *et al.*, 2013; Nowell *et al.*, 2018).

In other animal species which show no sign of the Meselson effect, e.g. darwinulid ostracods or tramini aphids, haplotype divergence has putatively been reduced by homogenising processes like mitotic gene conversion (Normark, 1999; Schoen & Martens, 2003). Surprisingly, extensive re-analyses of genome data from a wide array of asexual animals has revealed that elevated heterozygosity levels are predominantly driven by a hybrid origin of asexuality (Jaron *et al.*, 2019). In contrast, the mode of parthenogenesis seems to have only limited impact on heterozygosity. Apomictic species of non-hybrid origin showed heterozygosity levels similar to species with automixis or gamete duplication (Jaron *et al.*, 2019). This, yet again, hints at frequent action of homogenising mechanisms like mitotic gene conversion in apomictic species.

An integrative view on the influence of population size on mutational decay and the Meselson effect

This thesis entails the first genome-based support of the Meselson effect in a parthenogenetic organism (Chapter 4). The ancient asexual oribatid mite species *Oppiella nova* shows several population genomic patterns consistent with independent evolution of haplotypes and their parallel divergence. Unpublished results of shared heterozygous variants in the ancient asexual *Darwinula stevensoni* similarly indicate strong haplotype divergence (Simion, P. - personal communication). However, evolution without sex for millions of years is not necessarily a prerequisite for detectable levels of haplotype divergence: the only published genome-based example of the Meselson effect comes from the unicellular parasitic flagellate *Trypanosoma brucei gambiense*, which lost sex only ~ 10,000 years ago (Weir *et al.*, 2016). Notably, the Meselson effect was restricted to singular regions whereas genomes of different strains of *Trypanosoma brucei gambiense* were largely homozygous. Population sizes of *Trypanosoma brucei* are large but they go through several severe bottlenecks when transmitted between hosts and during migration within hosts, e.g., population sizes can drop below only a couple (> five) individuals when migrating from gut to salivary gland tissue in the host fly (Oberle *et al.*, 2010). I hypothesise that frequent bottlenecks after the loss of sex have lead to an increase in mutational load which in turn selected for higher rates of gene conversion in *Trypanosoma* flagellates and resulted in the observed high levels of homozygosity. I suggest that via this mechanism genomes will eventually become completely homozygous erasing any detectable Meselson effect in *Trypanosoma* flagellates. Contrastingly, in ancient asexuals (asexual oribatid mites, darwinulid ostracods) effective purging of deleterious mutations is largely driven by their large and stable population sizes and homogenising mechanisms likely play a secondary role. Accordingly, only low rates of gene conversion allow to detect the Meselson effect even after millions of years.

A general rule to be drawn from these conclusions is that large population sizes will generate effective purifying selection and to some extent allow for haplotype divergence in some asexual animals. Contrastingly, animals with smaller population sizes require homogenising mechanisms to counter deleterious mutation accumulation which leave their genomes largely homozygous. Direct comparisons of population sizes, the effectiveness of purifying selection, genome-wide loss of heterozygosity, and positive selection acting on e.g. the gene conversion machinery among small and large asexual animals may be suited to yield conclusive insight.

Synthesis and outlook

The classical prediction that sex is imperative for effective purging of deleterious mutations is not commonly met in animals. It remains an open question whether large population sizes and/or effective repair mechanisms facilitate effective selection in asexuals. Haplotype divergence is no common feature of asexual animals, either. Presence of the Meselson effect has to be interpreted with care regarding the potential confounding effects of paralogs.

Generally, most of the classical predictions on genome evolution under asexuality need (re-)evaluation on a whole genome basis. For this, it is important to compare replicates of independently derived asexual lineages to closely related sexual species on both the population- and species-level to disentangle true consequences of asexuality from confounding lineage-specific patterns. In doing so, the understudied “truly ancient asexuals” - oribatid mites and darwinulid ostracods - deserve special attention.

Clearly, asexual animals feature complex genomic mosaics shaped by different factors, some of which may well facilitate their long-term evolutionary success, not their extinction.

IV References

General References

Aanen, D., Beekman, M. & Kokko, H. Weird sex: the underappreciated diversity of sexual reproduction. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371**, 20160262 (2016).

Abad, P. et al. Genome sequence of the metazoan plant-parasitic nematode *Meloidogyne incognita*. *Nat. Biotechnol.* **26**, 909–915 (2008).

Ament-Velásquez, S. L. et al. Population genomics of sexual and asexual lineages in fissiparous ribbon worms (*Lineus*, Nemertea): hybridization, polyploidy and the Meselson effect. *Mol. Ecol.* **25**, 3356–3369 (2016).

Barraclough, T. G., Fontaneto, D., Ricci, C. & Herniou, E. A. Evidence for inefficient selection against deleterious mutations in cytochrome oxidase I of asexual bdelloid rotifers. *Mol. Biol. Evol.* **24**, 1952–1962 (2007).

Bast, J. et al. Consequences of asexuality in natural populations: insights from stick insects. *Mol. Biol. Evol.* **35**, 1668–1677 (2018).

Bast, J. et al. No accumulation of transposable elements in asexual arthropods. *Mol. Biol. Evol.* **33**, 697–706 (2015).

Bell, G. *The Masterpiece of Nature: The Evolution and Genetics of Sexuality* 1–635 (Croom Helm Ltd., 1982).

Bengtsson, B. O. Asex and evolution: a very large-scale overview In *Lost sex* (eds Schoen, I., Martens, K. & Van Dijk, P.) 1–19 (Springer Netherlands, 2009).

Birky, C. W. Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. *Genetics* **144**, 427–437 (1996).

Brandt, A. et al. No signal of deleterious mutation accumulation in conserved gene sequences of extant asexual hexapods. *Sci. Rep.* **9**, 5338 (2019).

Brandt, A. et al. Effective purifying selection in ancient asexual oribatid mites. *Nat. Commun.* **8**, 873 (2017).

Butlin, R. The costs and benefits of sex: new insights from old asexual lineages. *Nat. Rev. Genet.* **3**, 311–317 (2002).

Charlesworth, B., Morgan, M. T. & Charlesworth, D. The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289–1303 (1993a).

Charlesworth, D., Morgan, M. T. & Charlesworth, B. Mutation accumulation in finite outbreeding and inbreeding populations. *Genet. Res.* **61**, 39–56 (1993b)

Charlesworth, D., Morgan, M. T. & Charlesworth, B. Mutation accumulation in finite populations. *J. Hered.* **84**, 321–325 (1993c).

Cianciolo, J. M. & Norton, R. A. The ecological distribution of reproductive mode in oribatid mites, as related to biological complexity. *Exp. Appl. Acarol.* **40**, 1–25 (2006).

Debortoli, N. et al. Genetic exchange among bdelloid rotifers is more likely due to horizontal gene transfer than to meiotic sex. *Curr. Biol.* **26**, 723–732 (2016).

Delmotte, F. et al. Phylogenetic evidence for hybrid origins of asexual lineages in an aphid species. *Evolution* **57**, 1291–1303 (2003).

Felsenstein, J. The evolutionary advantage of recombination. *Genetics* **78**, 737–756 (1974).

Felsenstein, J. & Yokoyama, S. The evolutionary advantage of recombination. II. Individual selection for recombination. *Genetics* **83**, 845–859 (1976).

Fisher, R. A. *The genetical theory of natural selection* 1–356 (Oxford University Press, 1930).

Flot, J. F., Debortoli, N., Hallet, B., Narayan, J. & Van Doninck, K. Reply to cross-contamination explains“ inter and intraspecific horizontal genetic transfers” between asexual bdelloid rotifers (Wilson, Nowell & Barraclough 2018). *bioRxiv*, https://doi.org/10.1101/368209 (2018).

Flot, J.-F. et al. Genomic evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. *Nature* **500**, 453–457 (2013).

Fradin, H. et al. Genome architecture and evolution of a unichromosomal asexual nematode. *Curr. Biol.* **27**, 2928–2939.E6 (2017).

Gabriel, W., Lynch, M. & Buerger, R. Muller’s ratchet and mutational meltdowns. *Evolution* **47**, 1744–1757 (1993).

Galtier, N. et al. Codon usage bias in animals: disentangling the effects of natural selection, effective population size and GC-biased gene conversion. *Mol. Biol. Evol.* **35**, 1092–1103 (2018).

Gaston, K. J., Blackburn, T. M. & Lawton, J. H. Interspecific abundance-range size relationships: an appraisal of mechanisms. *J. Anim. Ecol.* **66**, 579–601 (1997).

Ghiselin, M. T. *The economy of nature and the evolution of sex* 1–346 (University of California Press, 1974).

Glémin, S., François, C. M. & Galtier, N. Genome evolution in outcrossing vs. selfing vs. asexual species In *Evolutionary genomics: statistical and computational methods, Volume 1 (*ed Anisimova, M.) 331–369 (Humana Press, 2019).

Gordo, I. & Charlesworth, B. The degeneration of asexual haploid populations and the speed of Muller's ratchet. *Genetics* **154**, 1379–1387 (2000).

Hamilton, W. D. Sex versus non-sex versus parasite. *Oikos* **35**, 282–290 (1980).

Hamilton, W. D., Axelrod, R. & Tanese, R. Sexual reproduction as an adaptation to resist parasites (a review). *Proc. Natl. Acad. Sci. USA* **87**, 3566–3573 (1990).

Hartfield, M. Evolutionary genetic consequences of facultative sex and outcrossing. *J. Evol. Biol.* **29**, 5–22 (2016).

Heethoff, M. et al. High genetic divergences indicate ancient separation of parthenogenetic lineages of the oribatid mite *Platynothrus peltifer* (Acari, Oribatida). *J. Evol. Biol.* **20**, 392–402 (2007).

Heimpel, G. E. & de Boer, J. G. Sex determination in the Hymenoptera. *Annu. Rev. Entomol.* **53**, 209–230 (2008).

Henry, L., Schwander, T. & Crespi, B. J. Deleterious mutation accumulation in asexual *Timema* stick insects. *Mol. Biol. Evol.* **29**, 401–408 (2012).

Hershberg, R. & Petrov, D. A. Selection on codon bias. *Annu. Rev. Genet.* **42**, 287–299 (2008).

Hill, W. G. & Robertson, A. The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294 (1966).

Hiraki, H. *et al.* Genome analysis of *Diploscapter coronatus*: insights into molecular peculiarities of a nematode with parthenogenetic reproduction. *BMC Genomics* **18**, 478 (2017).

Jalvingh, K., Bast, J. & Schwander, T. Sex, evolution and maintenance of In *Encyclopedia of Evolutionary Biology* (ed. Kliman, R. M.) 89–97 (Academic Press, 2016).

Janko, K., Drozd, P., Flegr, J. & Pannell, J. R. Clonal turnover versus clonal decay: a null model for observed patterns of asexual longevity, diversity and distribution. *Evolution* **62**, 1264–1270 (2008).

Jaron, K. S. et al. Genomic features of asexual animals. b*ioRxiv*, https://doi.org/10.1101/497495 (2019).

Johnson, S. G. Geographic ranges, population structure, and ages of sexual and parthenogenetic snail lineages. *Evolution* **60**, 1417–1426 (2006).

Johnson, S. G. & Howard, R. S. Contrasting patterns of synonymous and nonsynonymous sequence evolution in asexual and sexual freshwater snail lineages. *Evolution* **61**, 2728–2735 (2007).

Judson, O. P. & Normark, B. B. Ancient asexual scandals. *Trends Ecol. Evol.* **11**, 41–46 (1996).

Keightley, P. D. & Otto, S. P. Interference among deleterious mutations favours sex and recombination in finite populations. *Nature* **443**, 89–92 (2006).

Kondrashov, A. S. Classification of hypotheses on the advantage of amphimixis. *J. Hered.* **84**, 372–387 (1993).

Kondrashov, A. S. Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**, 435–440 (1988).

Kraaijeveld, K. et al. Decay of sexual trait genes in an asexual parasitoid wasp. *Genome Biol. Evol.* **8**, 3685–3695 (2016).

Lamatsch, D. K. & Stoeck, M. Sperm-dependent parthenogenesis and hybridogenesis in teleost fishes In *Lost sex* (eds Schoen, I., Martens, K. & Van Dijk, P.) 399–432 (Springer Netherlands, 2009).

Law, J. H. & Crespi, B. J. The evolution of geographic parthenogenesis in *Timema* walking-sticks. *Mol. Ecol.* **11**, 1471–1489 (2002).

Lehtonen, J., Jennions, M. D. & Kokko, H. The many costs of sex. *Trends Ecol. Evol.* **27**, 172–178 (2012).

Leria, L., Vila-Farré, M., Solà, E. & Riutort, M. Outstanding intraindividual genetic diversity in fissiparous planarians (*Dugesia*, Platyhelminthes) with facultative sex. *BMC Evol. Biol.* **19**, 130 (2019).

Lindsey, A. R. I. et al. Comparative genomics of the miniature wasp and pest control agent *Trichogramma pretiosum*. *BMC Biol.* **16**, 54 (2018).

Loewe, L. & Lamatsch, D. K. Quantifying the threat of extinction from Muller's ratchet in the diploid Amazon molly (*Poecilia formosa*). *BMC Evol. Biol.* **8**, 88 (2008).

Lunt, D. H. Genetic tests of ancient asexuality in root knot nematodes reveal recent hybrid origins. *BMC Evol. Biol.* **8**, 194 (2008).

Lynch, M., Buerger, R., Butcher, D. & Gabriel, W. The mutational meltdown in asexual populations. *J. Hered.* **84**, 339–344 (1993).

Maderspacher, F. Asexuality: the insects that stick with it. *Curr. Biol.* **21**, R495–R497 (2011).

Marais, G. Biased gene conversion: implications for genome and sex evolution. *Trends Genet.* **19**, 330–338 (2003).

Maraun, M. et al. Radiation in sexual and parthenogenetic oribatid mites (Oribatida, Acari) as indicated by genetic divergence of closely related species. *Exp. Appl. Acarol.* **29**, 265–277 (2003).

Maraun, M., Norton, R. A., Ehnes, R. B., Scheu, S. & Erdmann, G. Positive correlation between density and parthenogenetic reproduction in oribatid mites (Acari) supports the structured resource theory of sexual reproduction. *Evol. Ecol. Res.* **14**, 311–323 (2012).

Mark Welch, D. B., Mark Welch, J. L. & Meselson, M. Evidence for degenerate tetraploidy in bdelloid rotifers. *Proc. Natl Acad. Sci. USA* **105**, 5145–5149 (2008).

Mark Welch, D. B. & Meselson, M. S. Rates of nucleotide substitution in sexual and anciently asexual rotifers. *Proc. Natl Acad. Sci. USA* **98**, 6720–6724 (2001).

Mark Welch, D. & Meselson, M. Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. *Science* **288**, 1211–1215 (2000).

Maynard Smith, J. *The Evolution of Sex* 1–222 (Cambridge University Press, 1978).

Meirmans, S., Meirmans, P. G. & Kirkendall, L. R. The costs of sex: facing real-world complexities. *Q. Rev. Biol.* **87**, 19–40 (2012).

Muller, H. J. Some genetic aspects of sex. *Am. Nat.* **66**, 118–138 (1932).

Muller, H. J. The relation of recombination to mutational advance. *Mutat. Res.* **1**, 2–9 (1964).

Neiman, M., Hehman, G., Miller, J. T., Logsdon, J. M. & Taylor, D. R. Accelerated mutation accumulation in asexual lineages of a freshwater snail. *Mol. Biol. Evol.* **27**, 954–963 (2010).

Neiman, M., Meirmans, P. G., Schwander, T. & Meirmans, S. Sex in the wild: how and why field-based studies contribute to solving the problem of sex. *Evolution* **72**, 1194–1203 (2018).

Neiman, M., Meirmans, S. & Meirmans, P. G. What can asexual lineage age tell us about the maintenance of sex? *Ann. N. Y. Acad. Sci.* **1168**, 185–200 (2009).

Neiman, M. & Schwander, T. Using parthenogenetic lineages to identify advantages of sex. *Evol. Biol.* **38**, 115–123 (2011).

Normark, B. B. Evolution in a putatively ancient asexual aphid lineage: recombination and rapid karyotype change. *Evolution* **53**, 1458-1469 (1999).

Normark, B. B. & Johnson, N. A. Niche explosion. *Genetica* **139**, 551–564 (2011).

Normark, B. B., Judson, O. P. & Moran, N. A. Genomic signatures of ancient asexual lineages. *Biol. J. Linn. Soc. Lond.* **79**, 69–84 (2003).

Normark, B. B. & Moran, N. A. Testing for the accumulation of deleterious mutations in asexual eukaryote genomes using molecular sequences. *J. Nat. Hist.* **34**, 1719–1729 (2000).

Norton, R. A. & Palmer, S. C. The distribution, mechanisms and evolutionary significance of parthenogenesis in oribatid mites In *The Acari: Reproduction, Development, and Life-History Strategies* (eds Schuster, R. & Murphy, P. W.) 107–136 (Chapman & Hall, 1991).

Nowell, R. W. et al. Comparative genomics of bdelloid rotifers: insights from desiccating and nondesiccating species. *PLoS Biol.*, https://doi.org/10.1371/journal.pbio.2004830 (2018).

Oberle, M., Balmer, O., Brun, R. & Roditi, I. Bottlenecks and the maintenance of minor genotypes during the life cycle of *Trypanosoma brucei*. *PLoS Pathog.* **6**, e1001023 (2010).

Ollivier, M. et al. Comparison of gene repertoires and patterns of evolutionary rates in eight aphid species that differ by reproductive mode. *Genome Biol. Evol.* **4**, 155–167 (2012).

Otto, S. P. The evolutionary enigma of sex. *Am. Nat.* **174**, S1–S14 (2009).

Pace, C. N. et al. Contribution of hydrophobic interactions to protein stability. *J. Mol. Biol.* **408**, 514–528 (2011).

Paland, S. & Lynch, M. Transitions to asexuality result in excess amino acid substitutions. *Science* **311**, 990–992 (2006).

Pannebakker, B. A., Pijnacker, L. P., Zwaan, B. J. & Beukeboom, L. W. Cytology of *Wolbachia*-induced parthenogenesis in *Leptopilina clavipes* (Hymenoptera: Figitidae). *Genome* **47**, 299–303 (2004).

Rice, W. R. & Friberg, U. A graphical approach to lineage selection between clonals and sexuals In *Lost sex* (eds Schoen, I., Martens, K. & Van Dijk, P.) 75–97 (Springer Netherlands, 2009).

Ross, L., Hardy, N. B., Okusu, A. & Normark, B. B. Large population size predicts the distribution of asexuality in scale insects. *Evolution* **67**, 196–206 (2013).

Sandoval, C. P. The effects of the relative geographic scales of gene flow and selection on morph frequencies in the walking-stick *Timema cristinae*. *Evolution* **48**, 1866–1879 (1994).

Schaefer, I. et al. No evidence for the “Meselson effect” in parthenogenetic oribatid mites (Oribatida, Acari). *J. Evol. Biol.* **19**, 184–193 (2006).

Scheu, S. & Drossel, B. Sexual reproduction prevails in a world of structured resources in short supply. *Proc. Biol. Sci.* **274**, 1225–1231 (2007).

Schoen, I. & Martens, K. No slave to sex. *Proc. Biol. Sci.* **270**, 827–833 (2003).

Schoen, I., Martens, K. & Van Dijk, P. *Lost Sex-The Evolutionary Biology of Parthenogenesis* 1–617 (Springer, 2009).

Schoen, I., Martens, K., Van Doninck, K. & Butlin, R. K. Evolution in the slow lane: molecular rates of evolution in sexual and asexual ostracods (Crustacea: Ostracoda). *Biol. J. Linn. Soc. Lond.* **79**, 93–100 (2003).

Schwander, T. Evolution: the end of an ancient asexual scandal. *Curr. Biol.* **26**, R233–R235 (2016).

Schwander, T. & Crespi, B. J. Multiple direct transitions from sexual reproduction to apomictic parthenogenesis in *Timema* stick insects. *Evolution* **63**, 84–103 (2009).

Schwander, T. & Crespi, B. J. Twigs on the tree of life? Neutral and selective models for integrating macroevolutionary patterns with microevolutionary processes in the analysis of asexuality. *Mol. Ecol.* **18**, 28–42 (2009).

Schwander, T., Henry, L. & Crespi, B. J. Molecular evidence for ancient asexuality in *Timema* stick insects. *Curr. Biol.* **21**, 1129–1134 (2011).

Schwander, T. & Oldroyd, B. P. Androgenesis: where males hijack eggs to clone themselves. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371**, 20150534 (2016).

Sharbrough, J., Luse, M., Boore, J. L., Logsdon, J. M. Jr. & Neiman, M. Radical amino acid mutations persist longer in the absence of sex. *Evolution* **72**, 808–824 (2018).

Signorovitch, A., Hur, J., Gladyshev, E. & Meselson, M. Allele sharing and evidence for sexuality in a mitochondrial clade of bdelloid rotifers. *Genetics* **200**, 581–590 (2015).

Stouthamer, R. & Kazmer, D. J. Cytogenetics of microbe-associated parthenogenesis and its consequences for gene flow in *Trichogramma* wasps. *Heredity* **73**, 317–327 (1994).

Szitenberg, A. et al. Comparative genomics of apomictic root-knot nematodes: hybridisation, ploidy, and dynamic genome change. *Genome Biol. Evol.* **9**, 2844–2861 (2017).

Taylor, D. J. & Foighil, D. O. Transglobal comparisons of nuclear and mitochondrial genetic structure in a marine polyploid clam (*Lasaea*, Lasaeidae). *Heredity* **84**, 321–330 (2000).

Tucker, A. E., Ackerman, M. S., Eads, B. D., Xu, S. & Lynch, M. Population-genomic insights into the evolutionary origin and fate of obligately asexual *Daphnia pulex*. *Proc. Natl. Acad. Sci. USA* **110**, 15740–15745 (2013).

Vakhrusheva, O. A. et al. Recombination in a natural population of the bdelloid rotifer *Adineta vaga*. *bioRxiv*, https://doi.org/10.1101/489393 (2018).

Van Doninck, K., Schoen, I., Martens, K. & Goddeeris, B. The life-cycle of the asexual ostracod *Darwinula stevensoni* (Brady & Robertson 1870) (Crustacea, Ostracoda) in a temperate pond. *Hydrobiologia* **500**, 331–340 (2003).

Von Saltzwedel, H., Maraun, M., Scheu, S. & Schaefer, I. Evidence for frozen-niche variation in a cosmopolitan parthenogenetic soil mite species (Acari, Oribatida). *PLoS ONE* **9**, e113268 (2014).

Von Siebold, C. T. E. *Wahre Parthenogenesis bei Schmetterlingen und Bienen: ein Beitrag zur Fortpflanzungsgeschichte der Thiere* 1–156 (W. Engelmann, 1856).

Vrijenhoek, R. C. Animal clones and diversity. *Bioscience* **48**, 617–628 (1998).

Warren, W. C. et al. Clonal polymorphism and high heterozygosity in the celibate genome of the Amazon molly. *Nat. Ecol. Evol.* **2**, 669–679 (2018).

Weir, W. et al. Population genomics reveals the origin and asexual evolution of human infective trypanosomes. *elife* **5**, e11473 (2016).

West, S. A., Lively, C. M. & Read, A. F. A pluralist approach to sex and recombination. *J. Evol. Biol.* **12**, 1003–1012 (1999).

White, E. P., Ernest, S. K. M., Kerkhoff, A. J. & Enquist, B. J. Relationships between body size and abundance in ecology. *Trends Ecol. Evol.* **22**, 323–330 (2007).

Williams, G. C. *Sex and Evolution* 1–200 (Princeton University Press, 1975).

Wilson, C. G., Nowell, R. W. & Barraclough, T. G. Cross-contamination explains “inter and intraspecific horizontal genetic transfers” between asexual bdelloid rotifers. *Curr. Biol.* **28**, 2436–2444 (2018).

Research Chapter 1

Ament-Velásquez, S. L. et al. Population genomics of sexual and asexual lineages in fissiparous ribbon worms (*Lineus*, Nemertea): hybridization, polyploidy and the Meselson effect. *Mol. Ecol.* **25**, 3356–3369 (2016).

Bachtrog, D., Hom, E., Wong, K. M., Maside, X. & de Jong, P. Genomic degradation of a young Y chromosome in *Drosophila miranda*. *Genome Biol.* **9**, R30 (2008).

Barraclough, T. G., Fontaneto, D., Ricci, C. & Herniou, E. A. Evidence for inefficient selection against deleterious mutations in cytochrome oxidase I of asexual bdelloid rotifers. *Mol. Biol. Evol.* **24**, 1952–1962 (2007).

Bast, J. et al. No accumulation of transposable elements in asexual arthropods. *Mol. Biol. Evol.* **33**, 697–706 (2015).

Bates, D., Maechler, M., Bolker, B. & Walker, S. Package 'lme4': Linear Mixed-Effects Models using 'Eigen' and S4. R Package Version 1.1-12 https://cran.r-project.org/web/packages/lme4/ (2016).

Bell, G. *The Masterpiece of Nature: The Evolution and Genetics of Sexuality* 1–635 (Croom Helm Ltd., 1982).

Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552 (2000).

Cianciolo, J. M. & Norton, R. A. The ecological distribution of reproductive mode in oribatid mites, as related to biological complexity. *Exp. Appl. Acarol.* **40**, 1–25 (2006).

Conesa, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005).

Debortoli, N. et al. Genetic exchange among bdelloid rotifers is more likely due to horizontal gene transfer than to meiotic sex. *Curr. Biol.* **26**, 723–732 (2016).

Domes, K., Maraun, M., Scheu, S. & Cameron, S. L. The complete mitochondrial genome of the sexual oribatid mite *Steganacarus magnus*: genome rearrangements and loss of tRNAs. *BMC Genom.* **9**, 532 (2008).

Domes, K., Norton, R. A., Maraun, M. & Scheu, S. Reevolution of sexuality breaks Dollo's law. *Proc. Natl. Acad. Sci. USA* **104**, 7139–7144 (2007).

Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

Felsenstein, J. The evolutionary advantage of recombination. *Genetics* **78**, 737–756 (1974).

Felsenstein, J. & Yokoyama, S. The evolutionary advantage of recombination. II. Individual selection for recombination. *Genetics* **83**, 845–859 (1976).

Gaston, K. J., Blackburn, T. M. & Lawton, J. H. Interspecific abundance-range size relationships: an appraisal of mechanisms. *J. Anim. Ecol.* **66**, 579–601 (1997).

Goldberg, E. E. & Igić, B. On phylogenetic tests of irreversible evolution. *Evolution* **62**, 2727–2741 (2008).

Gordo, I. & Charlesworth, B. The degeneration of asexual haploid populations and the speed of Muller's ratchet. *Genetics* **154**, 1379–1387 (2000).

Grabherr, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).

Grbić, M. et al. The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. *Nature* **479**, 487–492 (2011).

Haas, B. J. et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).

Harrison, P. W. et al. Sexual selection drives evolution and rapid turnover of male gene expression. *Proc. Natl. Acad. Sci. USA* **112**, 4393–4398 (2015).

Hartfield, M. Evolutionary genetic consequences of facultative sex and outcrossing. *J. Evol. Biol.* **29**, 5–22 (2016).

Heethoff, M. et al. High genetic divergences indicate ancient separation of parthenogenetic lineages of the oribatid mite *Platynothrus peltifer* (Acari, Oribatida). *J. Evol. Biol.* **20**, 392–402 (2007).

Henry, L., Schwander, T. & Crespi, B. J. Deleterious mutation accumulation in asexual *Timema* stick insects. *Mol. Biol. Evol.* **29**, 401–408 (2012).

Hershberg, R. & Petrov, D. A. Selection on codon bias. *Annu. Rev. Genet.* **42**, 287–299 (2008).

Hill, W. G. & Robertson, A. The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294 (1966).

Hollister, J. D. et al. Recurrent loss of sex is associated with accumulation of deleterious mutations in *Oenothera*. *Mol. Biol. Evol.* **32**, 896–905 (2014).

Hughes, J. F. et al. Strict evolutionary conservation followed rapid gene loss on human and rhesus Y chromosomes. *Nature* **483**, 82–87 (2012).

Jalvingh, K., Bast, J. & Schwander, T. Sex, evolution and maintenance of In *Encyclopedia of Evolutionary Biology* (ed. Kliman, R. M.) 89–97 (Academic Press, 2016).

Jarne, P. & Auld, J. R. Animals mix it up too: the distribution of self-fertilization among hermaphroditic animals. *Evolution* **60**, 1816–1824 (2006).

Johnson, S. G. & Howard, R. S. Contrasting patterns of synonymous and nonsynonymous sequence evolution in asexual and sexual freshwater snail lineages. *Evolution* **61**, 2728–2735 (2007).

Judson, O. P. & Normark, B. B. Ancient asexual scandals. *Trends Ecol. Evol.* **11**, 41–46 (1996).

Kang, D. & Hamasaki, N. Maintenance of mitochondrial DNA integrity: repair and degradation. *Curr. Genet.* **41**, 311–322 (2002).

Keightley, P. D. & Otto, S. P. Interference among deleterious mutations favours sex and recombination in finite populations. *Nature* **443**, 89–92 (2006).

Kempson, D., Lloyd, M. & Ghelardi, R. A new extractor for woodland litter. *Pedobiologia* **3**, 1–21 (1963).

Khakhlova, O. & Bock, R. Elimination of deleterious mutations in plastid genomes by gene conversion. *Plant J.* **46**, 85–94 (2006).

Kimura, M., Maruyama, T. & Crow, J. F. The mutation load in small populations. *Genetics* **48**, 1303–1312 (1963).

Lechner, M. et al. Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinform.* **12**, 124 (2011).

Lehtonen, J., Jennions, M. D. & Kokko, H. The many costs of sex. *Trends Ecol. Evol.* **27**, 172–178 (2012).

Li, W. H., Wu, C. I. & Luo, C. C. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* **2**, 150–174 (1985).

Lynch, M. & Blanchard, J. L. Deleterious mutation accumulation in organelle genomes. *Genetica* **102/103**, 29–39 (1998).

Lynch, M., Buerger, R., Butcher, D. & Gabriel, W. The mutational meltdown in asexual populations. *J. Hered.* **84**, 339–344 (1993).

Manly, B. F. J. *Randomization, Bootstrap and Monte Carlo Methods in Biology* 1–480 (Chapman & Hall/CRC, 1997).

Maraun, M. et al. Radiation in sexual and parthenogenetic oribatid mites (Oribatida, Acari) as indicated by genetic divergence of closely related species. *Exp. Appl. Acarol.* **29**, 265–277 (2003).

Maraun, M., Norton, R. A., Ehnes, R. B., Scheu, S. & Erdmann, G. Positive correlation between density and parthenogenetic reproduction in oribatid mites (Acari) supports the structured resource theory of sexual reproduction. *Evol. Ecol. Res.* **14**, 311–323 (2012).

Mark Welch, D. B. & Meselson, M. S. Rates of nucleotide substitution in sexual and anciently asexual rotifers. *Proc. Natl. Acad. Sci. USA* **98**, 6720–6724 (2001).

Maynard Smith, J. *The Evolution of Sex* 1–222 (Cambridge University Press, 1978).

Muller, H. J. The relation of recombination to mutational advance. *Mutat. Res.* **1**, 2–9 (1964).

Neiman, M., Hehman, G., Miller, J. T., Logsdon, J. M. & Taylor, D. R. Accelerated mutation accumulation in asexual lineages of a freshwater snail. *Mol. Biol. Evol.* **27**, 954–963 (2010).

Normark, B. B. & Johnson, N. A. Niche explosion. *Genetica* **139**, 551–564 (2011).

Norton, R. A. & Palmer, S. C. The distribution, mechanisms and evolutionary significance of parthenogenesis in oribatid mites In *The Acari: Reproduction, Development, and Life-History Strategies* (eds Schuster, R. & Murphy, P. W.) 107–136 (Chapman & Hall, 1991).

Ollivier, M. et al. Comparison of gene repertoires and patterns of evolutionary rates in eight aphid species that differ by reproductive mode. *Genome Biol. Evol.* **4**, 155–167 (2012).

Otto, S. P. The evolutionary enigma of sex. *Am. Nat.* **174**, S1–S14 (2009).

Pace, C. N. et al. Contribution of hydrophobic interactions to protein stability. *J. Mol. Biol.* **408**, 514–528 (2011).

R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*, http://www.R-project.org/ (2015).

Rice, P. EMBOSS: the European molecular biology open software suite. *Trends Genet.* **16**, 276–277 (2000).

Riek, R. P. et al. Evolutionary conservation of both the hydrophilic and hydrophobic nature of transmembrane residues. *J. Theor. Biol.* **172**, 245–258 (1995).

Ross, L., Hardy, N. B., Okusu, A. & Normark, B. B. Large population size predicts the distribution of asexuality in scale insects. *Evolution* **67**, 196–206 (2013).

Sahu, G., Gautam, S. K. & Poddar, A. N. Suitable hosts of root knot nematode attack: an assessment on the basis of morphological size variations and population density under field conditions. *Int. J. Phytopathol.* **4**, 87–92 (2015).

Schaefer, I., Norton, R. A., Scheu, S. & Maraun, M. Arthropod colonization of land - linking molecules and fossils in oribatid mites (Acari, Oribatida). *Mol. Phylogenet. Evol.* **57**, 113–121 (2010).

Schoen, I., Martens, K. & Van Dijk, P. *Lost Sex-The Evolutionary Biology of Parthenogenesis* 1–617 (Springer, 2009).

Schoen, I., Martens, K., Van Doninck, K. & Butlin, R. K. Evolution in the slow lane: molecular rates of evolution in sexual and asexual ostracods (Crustacea: Ostracoda). *Biol. J. Linn. Soc. Lond.* **79**, 93–100 (2003).

Schwander, T. Evolution: the end of an ancient asexual scandal. *Curr. Biol.* **26**, R233–R235 (2016).

Signorovitch, A., Hur, J., Gladyshev, E. & Meselson, M. Allele sharing and evidence for sexuality in a mitochondrial clade of bdelloid rotifers. *Genetics* **200**, 581–590 (2015).

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).

Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

Storey, J. D. qvalue: q-value estimation for false discovery rate control. R Package Version 2.0.0 https://github.com/StoreyLab/qvalue (2015).

Ticher, A., Aharon, T. & Dan, G. Nucleic acid composition, codon usage, and the rate of synonymous substitution in protein-coding genes. *J. Mol. Evol.* **28**, 286–298 (1989).

Van Doninck, K., Schoen, I., Martens, K. & Goddeeris, B. The life-cycle of the asexual ostracod *Darwinula stevensoni* (Brady & Robertson 1870) (Crustacea, Ostracoda) in a temperate pond. *Hydrobiologia* **500**, 331–340 (2003).

Van Leeuwen, T. et al. Mitochondrial heteroplasmy and the evolution of insecticide resistance: non-Mendelian inheritance in action. *Proc. Natl. Acad. Sci. USA* **105**, 5980–5985 (2008).

Veresoglou, S. D., Halley, J. M. & Rillig, M. C. Extinction risk of soil biota. *Nat. Commun.* **6**, 8862 (2015).

Von Saltzwedel, H., Maraun, M., Scheu, S. & Schaefer, I. Evidence for frozen-niche variation in a cosmopolitan parthenogenetic soil mite species (Acari, Oribatida). *PLoS ONE* **9**, e113268 (2014).

Weigmann, G., Miko L. *Die Tierwelt Deutschlands und der angrenzenden Meeresteile nach ihren Merkmalen und nach ihrer Lebensweise 76. Teil* 1–520 (Goecke & Evers, 2006).

Wernersson, R. & Pedersen, A. G. RevTrans: multiple alignment of coding DNA from aligned amino acid sequences. *Nucleic Acids Res.* **31**, 3537–3539 (2003).

White, E. P., Ernest, S. K. M., Kerkhoff, A. J. & Enquist, B. J. Relationships between body size and abundance in ecology. *Trends Ecol. Evol.* **22**, 323–330 (2007).

Williams, G. C. *Sex and Evolution* 1–200 (Princeton University Press, 1975).

Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).

Zhang, Z. et al. Codon deviation coefficient: a novel measure for estimating codon usage bias and its statistical significance. *BMC Bioinform.* **13**, 43 (2012).

Research Chapter 2

Altenhoff, A. M., Gil, M., Gonnet, G. H. & Dessimoz, C. Inferring hierarchical orthologous groups from orthologous gene pairs. *PLoS One* **8**, e53786 (2013).

Ament-Velásquez, S. L. et al. Population genomics of sexual and asexual lineages in fissiparous ribbon worms (*Lineus*, Nemertea): hybridization, polyploidy and the Meselson effect. *Mol. Ecol.* **25**, 3356–3369 (2016).

Bachtrog, D. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat. Rev. Genet.* **14**, 113–124 (2013).

Bachtrog, D. & Charlesworth, B. Reduced adaptation of a nonrecombining neo-Y chromosome. *Nature* **416**, 323–326 (2002).

Balloux, F., Lehmann, L. & de Meeus, T. The population genetics of clonal and partially clonal diploids. *Genetics* **164**, 1635–1644 (2003).

Birky, C. W. Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. *Genetics* **144**, 427–437 (1996).

Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

Bolivar, P., Mugal, C. F., Nater, A. & Ellegren, H. Recombination rate variation modulates gene sequence evolution mainly via GC-biased gene conversion, not Hill–Robertson interference, in an avian system. *Mol. Biol. Evol.* **33**, 216–227 (2016).

Brandt, A. et al. Effective purifying selection in ancient asexual oribatid mites. *Nat. Commun.* **8**, 873 (2017).

Castel, S. E. & Martienssen, R. A. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. *Nat. Rev. Genet.* **14**, 100–112 (2013).

Charlesworth, B., Morgan, M. T. & Charlesworth, D. The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289–1303 (1993).

Duret, L. Evolution of synonymous codon usage in metazoans. *Curr. Opin. Genet. Dev*. **12**, 640–649 (2002).

Duret, L. & Galtier, N. Biased gene conversion and the evolution of mammalian genomic landscapes. *Annu. Rev. Genomics Hum. Genet.* **10**, 285–311 (2009).

Dutheil, J. & Boussau, B. Non-homogeneous models of sequence evolution in the Bio++ suite of libraries and programs. *BMC Evol. Biol.* **8**, 255 (2008).

Dutheil, J. Y. et al. Efficient selection of branch-specific models of sequence evolution. *Mol. Biol. Evol.* **29**, 1861–1874 (2012).

Felsenstein, J. The evolutionary advantage of recombination. *Genetics* **78**, 737–756 (1974).

Figuet, E., Ballenghien, M., Romiguier, J. & Galtier, N. Biased gene conversion and GC-content evolution in the coding sequences of reptiles and vertebrates. *Genome Biol. Evol.* **7**, 240–250 (2014).

Galtier, N. & Gouy, M. Inferring pattern and process: maximum likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis. *Mol. Biol. Evol.* **15**, 871–879 (1998).

Galtier, N. et al. Codon usage bias in animals: disentangling the effects of natural selection, effective population size and GC-biased gene conversion. *Mol. Biol. Evol.* **35**, 1092–1103 (2018).

Glemin, S. et al. Quantification of GC-biased gene conversion in the human genome. *Genome Res.* **25**, 1215–1228 (2015).

Glemin, S., Bazin, E. & Charlesworth, D. Impact of mating systems on patterns of sequence polymorphism in flowering plants. *Proc. Biol. Sci.* **273**, 3011–3019 (2006).

Glemin, S. & Galtier, N. Genome evolution in outcrossing versus selfing versus asexual species In *Evolutionary genomics: statistical and computational methods, Volume 1 (*ed Anisimova, M.) 311–335 (Springer Science + Business Media, 2012).

Grabherr, M. G. et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).

Harrison, P. W. et al. Sexual selection drives evolution and rapid turnover of male gene expression. *Proc. Natl. Acad. Sci. USA* **112**, 4393–4398 (2015).

Hartfield, M. Evolutionary genetic consequences of facultative sex and outcrossing. *J. Evol. Biol.* **29**, 5–22 (2016).

Henry, L., Schwander, T. & Crespi, B. J. Deleterious mutation accumulation in asexual *Timema* stick insects. *Mol. Biol. Evol.* **29**, 401–408 (2012).

Hershberg, R. & Petrov, D. A. Selection on codon bias. *Annu. Rev. Genet.* **42**, 287–299 (2008).

Hill, W. G. & Robertson, A. The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294 (1966).

Hollister, J. D. et al. Recurrent loss of sex is associated with accumulation of deleterious mutations in *Oenothera*. *Mol. Biol. Evol.* **32**, 896–905 (2015).

Johnson, S. G. & Howard, R. S. Contrasting patterns of synonymous and nonsynonymous sequence evolution in asexual and sexual freshwater snail lineages. *Evolution* **61**, 2728–2735 (2007).

Keightley, P. D. & Otto, S. P. Interference among deleterious mutations favours sex and recombination in finite populations. *Nature* **443**, 89–92 (2006).

Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–360 (2012).

Law, J. H. & Crespi, B. J. The evolution of geographic parthenogenesis in *Timema* walking-sticks. *Mol. Ecol.* **11**, 1471–1489 (2002).

Lawrie, D. S., Messer, P. W., Hershberg, R. & Petrov, D. A. Strong purifying selection at synonymous sites in *D. melanogaster*. *PLoS Genet.* **9**, e1003527 (2013).

Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).

Li, W. H., Wu, C. I. & Luo, C. C. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* **2**, 150–174 (1985).

Manly, B. F. J. *Randomization, Bootstrap and Monte Carlo Methods in Biology, Third Edition* 1–480 (Chapman & Hall/CRC, 2006).

Marais, G. Biased gene conversion: implications for genome and sex evolution. *Trends Genet.* **19**, 330–338 (2003).

Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).

Moghadam, H. K., Harrison, P. W., Zachar, G., Szekely, T. & Mank, J. E. The plover neurotranscriptome assembly: transcriptomic analysis in an ecological model species without a reference genome. *Mol. Ecol. Resour.* **13**, 696–705 (2013).

Mugal, C. F., Weber, C. C. & Ellegren, H. GC-biased gene conversion links the recombination landscape and demography to genomic base composition. *Bioessays* **37**, 1317–1326 (2015).

Muller, H. J. The relation of recombination to mutational advance. *Mutat. Res.* **1**, 2–9 (1964).

Neiman, M., Hehman, G., Miller, J. T., Logsdon, J. M. & Taylor, D. R. Accelerated mutation accumulation in asexual lineages of a freshwater snail. *Mol. Biol. Evol.* **27**, 954–963 (2010).

Ollivier, M. et al. Comparison of gene repertoires and patterns of evolutionary rates in eight aphid species that differ by reproductive mode. *Genome Biol. Evol.* **4**, 155–167 (2012).

Peden, J. CodonW version 1.4.2, http://codonw.sourceforge.net/culong.html.

Pessia, E. et al. Evidence for widespread GC-biased gene conversion in eukaryotes. *Genome Biol. Evol.* **4**, 675–682 (2012).

R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*, http://www.R-project.org/ (2013).

Rice, P. EMBOSS: the European molecular biology open software suite. *Trends Genet.* **16**, 276–277 (2000).

Riesch, R. et al. Transitions between phases of genomic differentiation during stick-insect speciation. *Nat. Ecol. Evol.* **1**, 0082 (2017).

Romiguier, J. et al. Fast and robust characterization of time-heterogeneous sequence evolutionary processes using substitution mapping. *PLoS ONE* **7**, e33852 (2012).

Romiguier, J., Ranwez, V., Douzery, E. J. P. & Galtier, N. Contrasting GC-content dynamics across 33 mammalian genomes: relationship with life-history traits and chromosome sizes. *Genome Res.* **20**, 1001–1009 (2010).

Schwander, T. & Crespi, B. J. Multiple direct transitions from sexual reproduction to apomictic parthenogenesis in *Timema* stick insects. *Evolution* **63,** 84–103 (2009).

Schwander, T., Henry, L. & Crespi, B. J. Molecular evidence for ancient asexuality in *Timema* stick insects. *Curr. Biol.* **21**, 1129–1134 (2011).

Schwander, T., Libbrecht, R. & Keller, L. Supergenes and complex phenotypes. *Curr. Biol.* **24**, R288–R294 (2014).

Skaletsky, H. et al. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **423**, 825–837 (2003).

Spencer, C. C. A. et al. The influence of recombination on human genetic diversity. *PLoS Genet.* **2**, e148 (2006).

Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

Talavera, G. & Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564–577 (2007).

Talbert, P. B. & Henikoff, S. Centromeres convert but don't cross. *PLoS Biol.* **8**, e1000326 (2010).

Van den Bosch, M., Lohman, P. H. M. & Pastink, A. DNA double-strand break repair by homologous recombination. *Biol. Chem.* **383**, 873–892 (2002).

Wallace, I. M., O'Sullivan, O., Higgins, D. G. & Notredame, C. M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res.* **34**, 1692–1699 (2006).

Webster, M. T. & Hurst, L. D. Direct and indirect consequences of meiotic recombination: implications for genome evolution. *Trends Genet.* **28**, 101–109 (2012).

Wright, F. The effective number of codons used in a gene. *Gene* **87**, 23–29 (1990).

Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).

Zhang, Z. et al. Codon Deviation Coefficient: a novel measure for estimating codon usage bias and its statistical significance. *BMC Bioinformatics* **13**, 43 (2012).

Research Chapter 3

Ament-Velásquez, S. L. et al. Population genomics of sexual and asexual lineages in fissiparous ribbon worms (*Lineus*, Nemertea): hybridization, polyploidy and the Meselson effect. *Mol. Ecol.* **25**, 3356–3369 (2016).

Bachtrog, D., Hom, E., Wong, K. M., Maside, X. & de Jong, P. Genomic degradation of a young Y chromosome in *Drosophila miranda*. *Genome Biol.* **9**, R30 (2008).

Bast, J. et al. Consequences of asexuality in natural populations: insights from stick insects. *Mol. Biol. Evol.* **35**, 1668–1677 (2018).

Bates, D., Maechler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects models using lme4. *J. Stat. Soft.* **67**, 1–48 (2014).

Bell, G. *The masterpiece of nature: the evolution and genetics of sexuality* 1–635 (University of California Press, 1982).

Bradler, S., Cliquennois, N. & Buckley, T. R. Single origin of the Mascarene stick insects: ancient radiation on sunken islands. *BMC Evol. Biol.* **15**, 196 (2015).

Brandt, A. et al. Effective purifying selection in ancient asexual oribatid mites. *Nat. Commun.* **8**, 873 (2017).

Castresana, J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol. Biol. Evol.* **17**, 540–552 (2000).

Ebersberger, I., Strauss, S. & von Haeseler, A. HaMStR: profile hidden markov model based search for orthologs in ESTs. *BMC Evol. Biol.* **9**, 157 (2009).

Felsenstein, J. The evolutionary advantage of recombination. *Genetics* **78**, 737–756 (1974).

Frati, F. et al. High levels of genetic differentiation between *Wolbachia*-infected and non-infected populations of *Folsomia candida* (Collembola, Isotomidae). *Pedobiologia* **48**, 461–468 (2004).

Galtier, N. et al. Codon usage bias in animals: disentangling the effects of natural selection, effective population size and GC-biased gene conversion. *Mol. Biol. Evol.* **35**, 1092–1103 (2018).

Gerdol, M., De Moro, G., Venier, P. & Pallavicini, A. Analysis of synonymous codon usage patterns in sixty-four different bivalve species. *PeerJ* **3**, e1520 (2015).

Gordo, I. & Charlesworth, B. The degeneration of asexual haploid populations and the speed of Muller's ratchet. *Genetics* **154**, 1379–1387 (2000).

Haas, B. J. et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).

Hartfield, M. Evolutionary genetic consequences of facultative sex and outcrossing. *J. Evol. Biol.* **29**, 5–22 (2016).

Henry, L., Schwander, T. & Crespi, B. J. Deleterious mutation accumulation in asexual *Timema* stick insects. *Mol. Biol. Evol.* **29**, 401–408 (2012).

Hershberg, R. & Petrov, D. A. Selection on codon bias. *Annu. Rev. Genet.* **42**, 287–299 (2008).

Hill, W. G. & Robertson, A. The effect of linkage on limits to artificial selection. *Genet. Res.* **8**, 269–294 (1966).

Hollister, J. D. et al. Recurrent loss of sex is associated with accumulation of deleterious mutations in *Oenothera*. *Mol. Biol. Evol.* **32**, 896–905 (2014).

Hughes, J. F. et al. Strict evolutionary conservation followed rapid gene loss on human and rhesus Y chromosomes. *Nature* **483**, 82–86 (2012).

Hurst, L. D., Hamilton, W. D. & Ladle, R. J. Covert sex. *Trends Ecol. Evol.* **7**, 144–145 (1992).

Kang, D. & Hamasaki, N. Maintenance of mitochondrial DNA integrity: repair and degradation. *Curr. Genet.* **41**, 311–322 (2002).

Katju, V., Packard, L. B., Bu, L., Keightley, P. D. & Bergthorsson, U. Fitness decline in spontaneous mutation accumulation lines of *Caenorhabditis elegans* with varying effective population sizes. *Evolution* **69**, 104–116 (2015).

Keightley, P. D. & Otto, S. P. Interference among deleterious mutations favours sex and recombination in finite populations. *Nature* **443**, 89–92 (2006).

Khakhlova, O. & Bock, R. Elimination of deleterious mutations in plastid genomes by gene conversion. *Plant J.* **46**, 85–94 (2006).

Kondrashov, A. S. Deleterious mutations and the evolution of sexual reproduction. *Nature* **336**, 435–440 (1988).

Lehtonen, J., Jennions, M. D. & Kokko, H. The many costs of sex. *Trends Ecol. Evol.* **27**, 172–178 (2012).

Li, W. H., Wu, C. I. & Luo, C. C. A new method for estimating synonymous and nonsynonymous rates of nucleotide substitution considering the relative likelihood of nucleotide and codon changes. *Mol. Biol. Evol.* **2**, 150–174 (1985).

Lovell, J. T., Williamson, R. J., Wright, S. I., McKay, J. K. & Sharbel, T. F. Mutation accumulation in an asexual relative of *Arabidopsis*. *PLoS Genet.* **13**, e1006550 (2017).

Lynch, M. & Blanchard, J. L. Deleterious mutation accumulation in organelle genomes. *Genetica* **102-103**, 29–39 (1998).

Lynch, M., Buerger, R., Butcher, D. & Gabriel, W. The mutational meltdown in asexual populations. *J. Hered.* **84**, 339–344 (1993).

Manly, B. F. J. *Randomization, Bootstrap and Monte Carlo Methods in Biology* 1–480 (Chapman & Hall/CRC, 1997).

McDonald, M. J., Rice, D. P. & Desai, M. M. Sex speeds adaptation by altering the dynamics of molecular evolution. *Nature* **531**, 233–236 (2016).

Misof, B. et al. Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**, 763–767 (2014).

Mitreva, M. et al. Codon usage patterns in Nematoda: analysis based on over 25 million codons in thirty-two species. *Genome Biol.* **7**, R75 (2006).

Muller, H. J. The relation of recombination to mutational advance. *Mutat. Res.* **1**, 2–9 (1964).

Munro, J. B. et al. A molecular phylogeny of the Chalcidoidea (Hymenoptera). *PLoS One* **6**, e27023 (2011).

Neiman, M., Meirmans, P. G., Schwander, T. & Meirmans, S. Sex in the wild: how and why field-based studies contribute to solving the problem of sex. *Evolution* **72**, 1194–1203 (2018).

Nguyen, D. T., Spooner-Hart, R. N. & Riegler, M. Polyploidy versus endosymbionts in obligately thelytokous thrips. *BMC Evol. Biol.* **15**, 23 (2015).

Normark, B. B. & Johnson, N. A. Niche explosion. *Genetica* **139**, 551–564 (2011).

Ollivier, M. et al. Comparison of gene repertoires and patterns of evolutionary rates in eight aphid species that differ by reproductive mode. *Genome Biol. Evol.* **4**, 155–167 (2012).

Otto, S. P. The evolutionary enigma of sex. *Am. Nat.* **174**, S1–S14 (2009).

Pace, C. N. et al. Contribution of hydrophobic interactions to protein stability. *J. Mol. Biol.* **408**, 514–528 (2011).

Peden, J. CodonW version 1.4.2, http://codonw.sourceforge.net/culong.html.

Peters, R. S. et al. Evolutionary history of the Hymenoptera. *Curr. Biol.* **27**, 1013–1018 (2017).

Peters, R. S. et al. Transcriptome sequence-based phylogeny of chalcidoid wasps (Hymenoptera: Chalcidoidea) reveals a history of rapid radiations, convergence, and evolutionary success. *Mol. Phylogenet. Evol.* **120**, 286–296 (2018).

Pijnacker, L. P. The maturation divisions of the parthenogenetic stick insect *Carausius morosus* Br. (Orthoptera, Phasmidae). *Chromosoma* **19**, 99–112 (1966).

Pond, S. L. K., Frost, S. D. W. & Muse, S. V. HyPhy: hypothesis testing using phylogenies. *Bioinformatics* **21**, 676–679 (2005).

R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*, http://www.R-project.org/ (2013).

Rice, P. EMBOSS: the european molecular biology open software suite. *Trends Genet.* **16**, 2–3 (2000).

Rice, W. R. & Friberg, U. A graphical approach to lineage selection between clonals and sexuals In *Lost sex* (eds Schoen, I., Martens, K. & Van Dijk, P.) 75–97 (Springer Netherlands, 2009).

Riek, R. P. et al. Evolutionary conservation of both the hydrophilic and hydrophobic nature of transmembrane residues. *J. Theor. Biol.* **172**, 245–258 (1995).

Ross, L., Hardy, N. B., Okusu, A. & Normark, B. B. Large population size predicts the distribution of asexuality in scale insects. *Evolution* **67**, 196–206 (2013).

Sharbrough, J., Luse, M., Boore, J. L., Logsdon, J. M. Jr. & Neiman, M. Radical amino acid mutations persist longer in the absence of sex. *Evolution* **72**, 808–824 (2018).

Simon, J.-C., Rispe, C. & Sunnucks, P. Ecology and evolution of sex in aphids. *Trends Ecol. Evol.* **17**, 34–39 (2002).

Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

Stouthamer, R. & Mak, F. Influence of antibiotics on the offspring production of the *Wolbachia*-infected parthenogenetic parasitoid *Encarsia formosa*. *J. Invertebr. Pathol.* **80**, 41–45 (2002).

Train, C.-M., Glover, N. M., Gonnet, G. H., Altenhoff, A. M. & Dessimoz, C. Orthologous Matrix (OMA) algorithm 2.0: more robust to asymmetric evolutionary rates and more scalable hierarchical orthologous group inference. *Bioinformatics* **33**, i75–i82 (2017).

Tree of Sex Consortium. Tree of Sex: a database of sexual systems. *Sci Data* **1**, 140015, https://doi.org/10.1038/sdata.2014.15 (2014).

Van der Kooi, C. J. & Schwander, T. On the fate of sexual traits under asexuality. *Biol. Rev. Camb. Philos. Soc.* **89**, 805–819 (2014).

Vershinina, A. O. & Kuznetsova, V. G. Parthenogenesis in Hexapoda: Entognatha and non-holometabolous insects. *J. Zoolog. Syst. Evol. Res.* **54**, 257–268 (2016).

Von Reumont, B. M. et al. Pancrustacean phylogeny in the light of new phylogenomic data: support for Remipedia as the possible sister group of Hexapoda. *Mol. Biol. Evol.* **29**, 1031–1045 (2012).

Wallace, I. M., O'Sullivan, O., Higgins, D. G. & Notredame, C. M-Coffee: combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res.* **34**, 1692–1699 (2006).

White, M. J. D. The chromosomes of the parthenogenetic mantid *Brunneria borealis*. *Evolution* **2**, 90–93 (1948).

Wright, F. The effective number of codons used in a gene. *Gene* **87**, 23–29 (1990).

Yang, Q. et al. Morphological and molecular characterization of a sexually reproducing colony of the booklouse *Liposcelis bostrychophila* (Psocodea: Liposcelididae) found in Arizona. *Sci. Rep.* **5**, 10429 (2015).

Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).

Zhang, Z. et al. Codon Deviation Coefficient: a novel measure for estimating codon usage bias and its statistical significance. *BMC Bioinformatics* **13**, 43 (2012).

Research Chapter 4

Ament-Velásquez, S. L. et al. Population genomics of sexual and asexual lineages in fissiparous ribbon worms (*Lineus*, Nemertea): hybridization, polyploidy and the Meselson effect. *Mol. Ecol.* **25**, 3356–3369 (2016).

Anselmetti, Y. Pop-Con: a tool to visualize genotype profiles on site frequency spectrum from populational genomics data, https://github.com/YoannAnselmetti/Pop-Con.

Bankevich, A. et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **19**, 455–477 (2012).

Bast, J., Jaron, K. S., Schuseil, D., Roze, D. & Schwander, T. Asexual reproduction reduces transposable element load in experimental yeast populations. *elife* **8**, e48548 (2019).

Bast, J. et al. Consequences of asexuality in natural populations: insights from stick insects. *Mol. Biol. Evol.* **35**, 1668–1677 (2018).

Bast, J. et al. No accumulation of transposable elements in asexual arthropods. *Mol. Biol. Evol.* **33**, 697–706 (2015).

Bergmann, P., Laumann, M., Norton, R. A. & Heethoff, M. Cytological evidence for automictic thelytoky in parthenogenetic oribatid mites (Acari, Oribatida): synaptonemal complexes confirm meiosis in *Archegozetes longisetosus*. *Acarologia* **58**, 342–356 (2018).

Birky, C. W. Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. *Genetics* **144**, 427–437 (1996).

Birky, C. W. & Barraclough, T. G. Asexual Speciation In *Lost sex* (eds Schoen, I., Martens, K. & Van Dijk, P.) 201–216 (Springer Netherlands, 2009).

Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D. & Pirovano, W. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **27**, 578–579 (2011).

Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

Brandt, A. et al. Effective purifying selection in ancient asexual oribatid mites. *Nat. Commun.* **8**, 873 (2017).

Bray, N., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-Seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).

Broad Institute. Picard Toolkit, http://broadinstitute.github.io/picard/.

Bushnell, B. BBTools software package, https://sourceforge.net/projects/bbmap.

Butlin, R. The costs and benefits of sex: new insights from old asexual lineages. *Nat. Rev. Genet.* **3**, 311–317 (2002).

Campbell, M. S., Holt, C., Moore, B. & Yandell, M. Genome annotation and curation using MAKER and MAKER-P. *Curr. Protoc. Bioinformatics* **48**, 4.11.1–39 (2014).

Castel, S. E., Mohammadi, P., Chung, W. K., Shen, Y. & Lappalainen, T. Rare variant phasing and haplotypic expression from RNA sequencing with phASER. *Nat. Commun.* **7**, 12817 (2016).

Cianciolo, J. M. & Norton, R. A. The ecological distribution of reproductive mode in oribatid mites, as related to biological complexity. *Exp. Appl. Acarol.* **40**, 1–25 (2006).

Conesa, A. et al. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**, 3674–3676 (2005).

Conow, C., Fielder, D., Ovadia, Y. & Libeskind-Hadas, R. Jane: a new tool for the cophylogeny reconstruction problem. *Algorithms Mol. Biol.* **5**, 16 (2010).

Danecek, P. et al. The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).

Delmotte, F. et al. Phylogenetic evidence for hybrid origins of asexual lineages in an aphid species. *Evolution* **57**, 1291–1303 (2003).

Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

Engelstaedter, J. Asexual but not clonal: evolutionary processes in automictic populations. *Genetics* **206**, 993–1009 (2017).

Flot, J.-F. et al. Genomic evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. *Nature* **500**, 453–457 (2013).

Haas, B. J. et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).

Hartfield, M. Evolutionary genetic consequences of facultative sex and outcrossing. *J. Evol. Biol.* **29**, 5–22 (2016).

Heethoff, M., Bergmann, P. & Norton, R. A. Karyology and sex determination of oribatid mites. *Acarologia* **46**, 127–131 (2006).

Heethoff, M., Norton, R. A., Scheu, S. & Maraun, M. Parthenogenesis in oribatid mites (Acari, Oribatida): evolution without sex In *Lost sex* (eds Schoen, I., Martens, K. & Van Dijk, P.) 241–257 (Springer Netherlands, 2009).

Holt, C. & Yandell, M. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* **12**, 491 (2011).

Jaron, K. S. et al. Genomic features of asexual animals. *bioRxiv*, https://doi.org/10.1101/497495 (2019).

Johnson, S. G. Geographic ranges, population structure, and ages of sexual and parthenogenetic snail lineages. *Evolution* **60**, 1417–1426 (2006).

Judson, O. P. & Normark, B. B. Ancient asexual scandals. *Trends Ecol. Evol.* **11**, 41–46 (1996).

Kamvar, Z. N., Tabima, J. F. & Gruenwald, N. J. Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ* **2**, e281 (2014).

Kempson, D., Lloyd, M. & Ghelardi, R. A new extractor for woodland litter. *Pedobiologia* **3**, 1–21 (1963).

Kishino, H., Miyata, T. & Hasegawa, M. Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *J. Mol. Evol.* **31**, 151–160 (1990).

Knaus, B. J. & Gruenwald, N. J. vcfr: a package to manipulate and visualize variant call format data in R. *Mol. Ecol. Resour.* **17**, 44–53 (2017).

Korf, I. Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004).

Krueger, F. Trim Galore: a wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files, with some extra functionality for MspI-digested RRBS-type (Reduced Representation Bisulfite-Seq) libraries, https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

Kuhner, M. K. & Felsenstein, J. A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates. *Mol. Biol. Evol.* **11**, 459–468 (1994).

Laetsch, D. R. & Blaxter, M. L. BlobTools: Interrogation of genome assemblies. *F1000Res.* **6**, 1287 (2017).

Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–360 (2012).

Laumann, M., Bergmann, P. & Heethoff, M. Some remarks on the cytogenetics of oribatid mites. *Soil Org.* **80**, 223–232 (2008).

Lehtonen, J., Jennions, M. D. & Kokko, H. The many costs of sex. *Trends Ecol. Evol.* **27**, 172–178 (2012).

Leria, L., Vila-Farré, M., Solà, E. & Riutort, M. Outstanding intraindividual genetic diversity in fissiparous planarians (*Dugesia*, Platyhelminthes) with facultative sex. *BMC Evol. Biol.* **19**, 130 (2019).

Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).

Li, H. et al. The sequence alignment/map format and samtools. *Bioinformatics* **25**, 2078–2079 (2009).

Lunt, D. H. Genetic tests of ancient asexuality in root knot nematodes reveal recent hybrid origins. *BMC Evol. Biol.* **8**, 194 (2008).

Luo, R. et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* **1**, 18 (2012).

Marais, G. Biased gene conversion: implications for genome and sex evolution. *Trends Genet.* **19**, 330–338 (2003).

Maraun, M. et al. Radiation in sexual and parthenogenetic oribatid mites (Oribatida, Acari) as indicated by genetic divergence of closely related species. *Exp. Appl. Acarol.* **29**, 265–277 (2003).

Maraun, M., Norton, R. A., Ehnes, R. B., Scheu, S. & Erdmann, G. Positive correlation between density and parthenogenetic reproduction in oribatid mites (Acari) supports the structured resource theory of sexual reproduction. *Evol. Ecol. Res.* **14**, 311–323 (2012).

Mark Welch, D. B., Mark Welch, J. L. & Meselson, M. Evidence for degenerate tetraploidy in bdelloid rotifers. *Proc. Natl Acad. Sci. USA* **105**, 5145–5149 (2008).

Mark Welch, D. B., Ricci, C. & Meselson, M. Bdelloid rotifers: progress in understanding the success of an evolutionary scandal In *Lost sex* (eds Schoen, I., Martens, K. & Van Dijk, P.) 259–279 (Springer Netherlands, 2009).

Mark Welch, D. M. & Meselson, M. Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. *Science* **288**, 1211–1215 (2000).

Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).

Mateus, A. & Caeiro, F. An R implementation of several randomness tests. *AIP Conf. Proc.* **1618**, 531–534 (2014).

McKenna, A. et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).

Neiman, M., Lively, C. M. & Meirmans, S. Why sex? A pluralist approach revisited. *Trends Ecol. Evol.* **32**, 589–600 (2017).

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).

Normark, B. B. Evolution in a putatively ancient asexual aphid lineage: recombination and rapid karyotype change. *Evolution* **53**, 1458-1469 (1999).

Normark, B. B., Judson, O. P. & Moran, N. A. Genomic signatures of ancient asexual lineages. *Biol. J. Linn. Soc. Lond.* **79**, 69–84 (2003).

Norton, R. A., Kethley, J. B., Johnston, D. E. & O'Connor, B. M. Phylogenetic perspectives on genetic systems and reproductive modes of mites In *Evolution and diversity of sex ratio in insects and mites* (eds Wrensch, D. L. & Ebbert, M. A.) 8–99 (Chapman & Hall, 1993).

Norton, R. A. & Palmer, S. C. The distribution, mechanisms and evolutionary significance of parthenogenesis in oribatid mites In *The Acari: Reproduction, Development, and Life-History Strategies* (eds Schuster, R. & Murphy, P. W.) 107–136 (Chapman & Hall, 1991).

O'Connell, J. et al. NxTrim: optimized trimming of Illumina mate pair reads. *Bioinformatics* **31**, 2035–2037 (2015).

Palmer, S. C. & Norton, R. Genetic diversity in thelytokous oribatid mites (Acari; Acariformes: Desmonomata). *Biochem. Syst. Ecol.* **20**, 219–231 (1992).

Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).

Parra, G., Bradnam, K. & Korf, I. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* **23**, 1061–1067 (2007).

Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

R Core Team. R: A language and environment for statistical computing. *R Foundation for Statistical Computing, Vienna, Austria*, http://www.R-project.org/ (2013).

Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).

Rosenberger, M., Maraun, M., Scheu, S. & Schaefer, I. Pre- and post-glacial diversifications shape genetic complexity of soil-living microarthropod species. *Pedobiologia* **56**, 79–87 (2013).

Schaefer, I. et al. No evidence for the “Meselson effect” in parthenogenetic oribatid mites (Oribatida, Acari). *J. Evol. Biol.* **19**, 184–193 (2006).

Schaefer, I., Norton, R. A., Scheu, S. & Maraun, M. Arthropod colonization of land–linking molecules and fossils in oribatid mites (Acari, Oribatida). *Mol. Phylogenet. Evol.* **57**, 113–121 (2010).

Schaeffer, S., Kerschbaumer, M. & Koblmueller, S. Multiple new species: cryptic diversity in the widespread mite species *Cymbaeremaeus cymba* (Oribatida, Cymbaeremaeidae). *Mol. Phylogenet. Evol.* **135**, 185–192 (2019).

Schoen, I. & Martens, K. No slave to sex. *Proc. Biol. Sci.* **270**, 827–833 (2003).

Schoen, I., Martens, K. & Van Dijk, P. *Lost Sex-The Evolutionary Biology of Parthenogenesis* 1–617 (Springer, 2009a).

Schoen, I., Martens, K., Van Doninck, K. & Butlin, R. K. Evolution in the slow lane: molecular rates of evolution in sexual and asexual ostracods (Crustacea: Ostracoda). *Biol. J. Linn. Soc. Lond.* **79**, 93–100 (2003).

Schoen, I., Rossetti, G. & Martens, K. Darwinulid ostracods: ancient asexual scandals or scandalous gossip? In *Lost sex* (eds Schoen, I., Martens, K. & Van Dijk, P.) 217–240 (Springer Netherlands, 2009b).

Schurko, A. M., Neiman, M. & Logsdon, J. M. Signs of sex: what we know and how we know it. *Trends Ecol. Evol.* **24**, 208–217 (2009).

Schwander, T. Evolution: the end of an ancient asexual scandal. *Curr. Biol.* **26**, R233–R235 (2016).

Schwander, T., Henry, L. & Crespi, B. J. Molecular evidence for ancient asexuality in *Timema* stick insects. *Curr. Biol.* **21**, 1129–1134 (2011).

Seemann, T. snp-dists, https://github.com/tseemann/snp-dists.

Shimodaira, H. An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* **51**, 492–508 (2002).

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V. & Zdobnov, E. M. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).

Smit, A., Hubley, R. & Green, P. RepeatMasker Open-4.0, http://www.repeatmasker.org/.

Stanke, M. et al. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res.* **34**, W435–W439 (2006).

Subias, L. S. Listado sistimatico, sininimico y biogeografico de los Acaros Oribatidos (Acariformes, Oribatida) del mundo (1748–2002). *Graellsia* **60**, 3–305 (2004).

Taberly, G. Recherches sur la parthénogenèse thélytoque de deux espèces d'acariens Oribates: *Trhypochthonius tectorum* (Berlese) et *Platynothrus peltifer* (Koch). II: Etude anatomique, histologique et cytologique des femelles parthénogénétiques. *Acarologia* **28**, 285–293 (1987).

Taylor, D. J. & Foighil, D. O. Transglobal comparisons of nuclear and mitochondrial genetic structure in a marine polyploid clam (*Lasaea*, Lasaeidae). *Heredity* **84**, 321–330 (2000).

Von Saltzwedel, H., Maraun, M., Scheu, S. & Schaefer, I. Evidence for frozen-niche variation in a cosmopolitan parthenogenetic soil mite species (Acari, Oribatida). *PLoS ONE* **9**, e113268 (2014).

Weigmann, G., Miko L. *Die Tierwelt Deutschlands und der angrenzenden Meeresteile nach ihren Merkmalen und nach ihrer Lebensweise 76. Teil* 1–520 (Goecke & Evers, 2006).

Weir, W. et al. Population genomics reveals the origin and asexual evolution of human infective trypanosomes. *elife* **5**, e11473 (2016).

Wood, D. E. & Salzberg, S. L. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* **15**, R46 (2014).

Wrensch, D. L., Kethley, J. B. & Norton, R. A. Cytogenetics of holokinetic chromosomes and inverted meiosis: keys to the evolutionary success of mites, with generalizations on eukaryotes In *Mites: ecological and evolutionary analyses of life-history pattern*s (ed. Houck, M. A.) 282–343 (Chapman & Hall, 1994).

Yandell, M. & Ence, D. A beginner's guide to eukaryotic genome annotation. *Nat. Rev. Genet.* **13**, 329–342 (2012).

Zhang, C., Rabiee, M., Sayyari, E. & Mirarab, S. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**, 153 (2018).

V Supplementaries

Research Chapter 1

Effective purifying selection in ancient asexual oribatid mites

For a complete collection of Supplementary Data, see CD, attached.

Supplementary Table 1: NCBI accession numbers of partial coding sequences of *ef1α* and *hsp82* and partial sequences of 18S rDNA of 30 oribatid mite species generated by Domes *et al.* (2007). Sequences were used for construction of a phylogenetic tree (Fig. 1) and branch-specific dN/dS ratio analyses. Because sequences of *Nothrus palustris* were not part of the study of Domes *et al.* (2007) and not available from NCBI, they were extracted from the transcriptome data. Order of species reflects the arrangement of species in the phylogenetic tree from top to bottom.

Species	**Reproductive mode**	**Accession number**		
		ef1α	***hsp82***	**18S rDNA**
Tectocepheus velatus	asexual	EF093763	EF093770	EF093781
Eupelops plicatus	sexual	AY632797	DQ090783	EF091419
Achipteria coleoptrata	sexual	AY632776	EF081335	EF091418
Eutegaeus curviseta	sexual	EF081326	DQ090789	EF081297
*Carabodes femoralis**	sexual	EF081325	DQ090786	EF091429
Nothrus silvestris bistilus	asexual	EF081323	EF081333	EF081305
Nothrus truncatus	asexual	EF081322	EF081334	EF081306
Nothrus palustris	asexual	-	-	-
Novonothrus flagellatus	sexual	EF081324	DQ090801	EF081307
Camisia biurus	asexual	EF081312	EF081331	EF081302
Camisia spinifer	asexual	EF081313	EF081332	EF091420
Crotonia brachyrostrum	sexual	EF081314	DQ090796	EF081303
Crotonia cf. *caudata*	sexual	EF081315	DQ090795	EF081304
Platynothrus peltifer	asexual	AY632851	DQ090793	EF091422
Heminothrus paolianus	asexual	EF081316	DQ090794	EF091423

Nanhermannia coronata	asexual	AY632825	DQ090799	EF091421
Archegozetes longisetosus	asexual	EF081321	DQ090798	AF022027
Trypochthonius americanus	asexual	EF081317	EF081337	EF081298
Trypochthoniellus crassus	asexual	EF081320	EF081336	EF081300
Mainothrus badius	asexual	EF081318	EF081338	EF081301
Mucronothrus nasalis	asexual	EF081319	DQ090797	EF081299
Malaconothrus gracilis	asexual	EF081311	EF081339	EF091424
Hermannia gibba	sexual	EF081327	DQ090800	EF091426
Steganacarus magnus	sexual	AY632837	DQ090781	AF022040
Atropacarus striculus	asexual	EF081309	DQ090782	EF091416
Rhysotritia duplicata	asexual	EF081310	DQ090780	EF091417
Nehypochthonius porosus	asexual	EF081328	DQ090779	EF081308
Lohmannia banksi	asexual	EF081330	DQ090777	AF022036
Hypochthonius rufulus	asexual	AY632861	DQ090776	EF091427
Eniochthonius minutissimus	asexual	EF081329	DQ090773	EF091428

* *Carabodes subarcticus* for 18S rDNA

Supplementary Table 2: Details on oribatid mite sampling and transcriptome sequencing.

Species	Reproductive mode	Sampling locality	No. of pooled individuals	Sequencing provider	Sequencing platform
Achipteria coleoptrata	sexual	Solling	20	GATC Biotech	HiSeq 2500 125 bp PE
Hermannia gibba	sexual	Göttinger Wald	10	GATC Biotech	HiSeq 2500 125 bp PE
Hypochthonius rufulus	asexual	Göttinger Wald	50	GATC Biotech	HiSeq 2500 125 bp PE
Nothrus palustris	asexual	Göttinger Wald	3	GATC Biotech	HiSeq 2500 125 bp PE
Platynothrus peltifer	asexual	Solling	1	TAL	MiSeq 250 bp PE
Steganacarus magnus	sexual	Göttinger Wald	1	TAL	MiSeq 250 bp PE

Supplementary Table 3: Overview of the assembly and filtering processes. After assembly, transcripts were cleaned and then filtered for the most abundant isoform (i.e. most abundant allele in the population) using an RPKM based approach. The filtered transcript set was assessed for quality using BUSCO ('complete' are core arthropod genes assembled to full length). For subsequent analyses of purifying selection, ORFs were predicted.

Species	No of raw reads [M]	Contig N50 [bp]	No. of tran-scripts	No. of filtered tran-scripts	No. of ORFs	BUSCO complete [%]	BUSCO fragments [%]	BUSCO missing [%]
Achipteria coleoptrata	f: 67.03 r: 70.30	1,318	123,336	34,262	16,070	93.20	1.80	5.00
Hermannia gibba	f: 65.45 r: 68.47	1,481	142,723	39,974	19,108	95.70	0.50	3.80
Hypochthonius rufulus	f: 54.42 r: 60.23	1,204	59,116	35,595	15,507	90.10	5.40	4.50
Nothrus palustris	f: 68.44 r: 71.76	1,703	118,984	30,608	16,397	95.00	0.80	4.20
Platynothrus peltifer	f: 8.97 r: 11.35	702	85,238	68,779	24,108	83.40	11.70	4.90
Steganacarus magnus	f: 13.42 r: 16.83	1,761	73,855	35,286	21,131	92.60	2.30	5.10

Research Chapter 2

Consequences of asexuality in natural populations: insights from stick insects

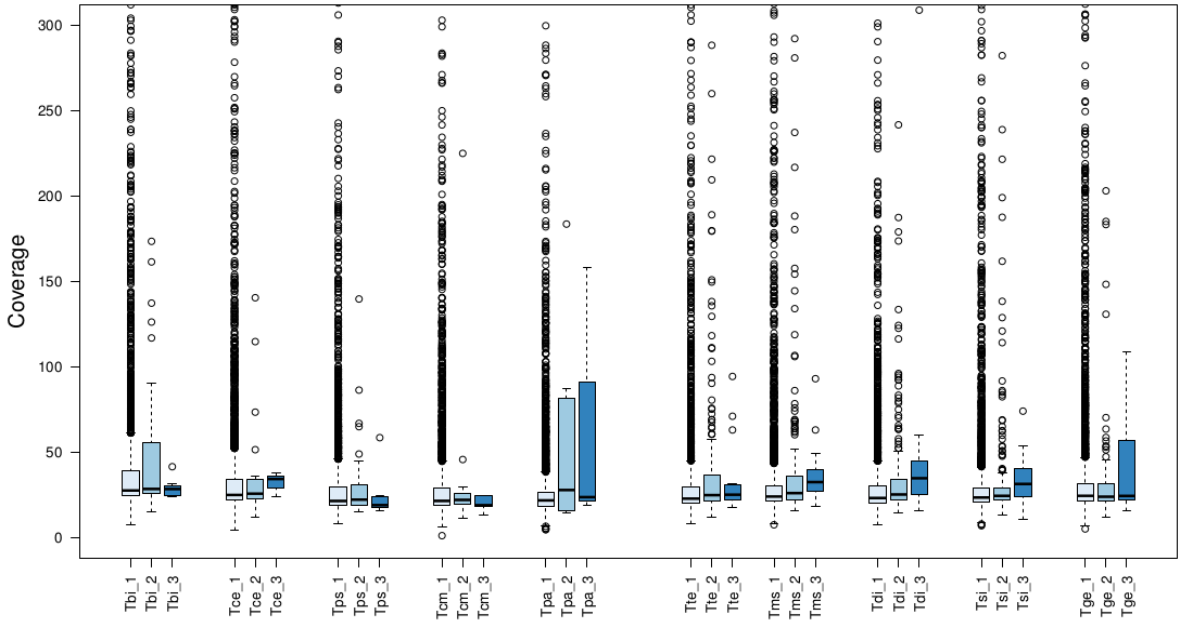
For a complete collection of Supplementary Data, see CD, attached.

Three-rate model

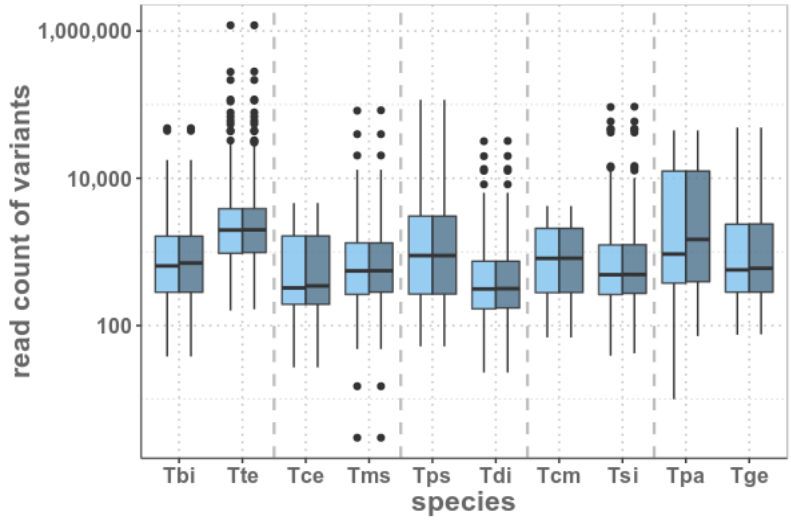
We used a three-ratio model, which calculates one dN/dS ratio for internal branches and one for terminal branches leading to asexual species and one for terminal branches leading to sexual species. In this way, only mutations occurring after the split of sexuals and asexuals are taken into account. After filtering data for which ω did not converge (0.0001 or 999 ω values for both terminal and internal branches) or where ω<1, 1647 genes remained for this model. We detected a small but significant difference for ω, with a faster rate of mutation accumulation for asexual *Timema* species (paired Wilcox test $P < 0.01$; median $\Delta_{sex\text{-}asex} = -0.008$). This indicates that asexual *Timema* species overall accumulate non-synonymous mutations at a higher rate.

Variant coverage estimation

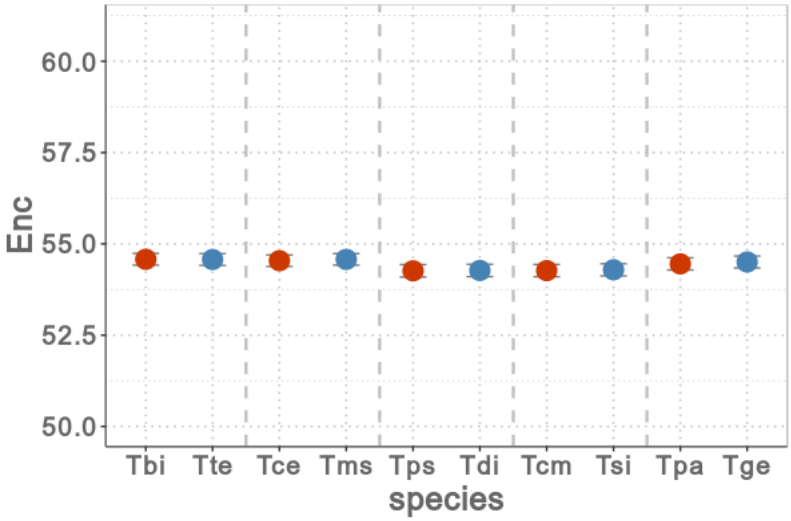
Genomic reads from SRA (SRR5248877 - SRR5248936) were mapped to the transcriptome assemblies using Bowtie2 v2.2.4. We used the local alignment mode to allow reads to be soft-clipped when mapping, since genomic reads contain intronic sequences not found in the transcriptomes. Coverage for each ortholog and ortholog variant was obtained using Bedtools v2.26. Coverage for each ortholog and its variants were summed together.



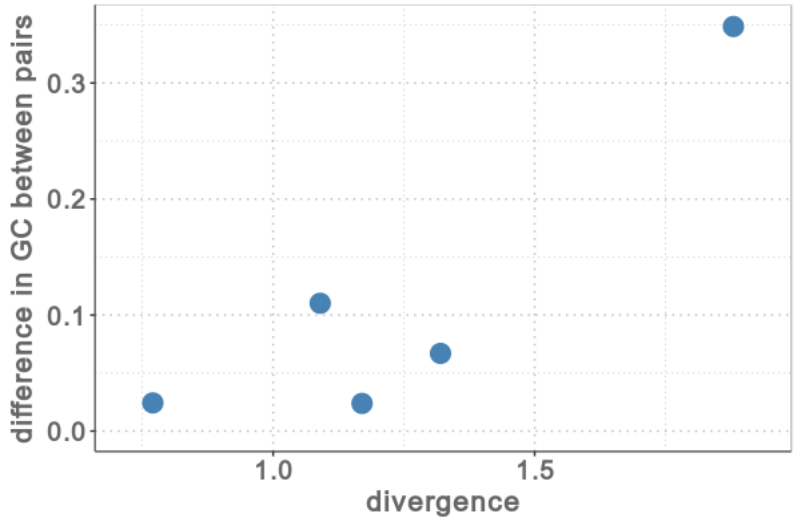
Supplementary Figure 1: Orthologous variants do not stem from gene duplicates. Genomic per-base read coverage for orthologs with one, two or three variants per species. Only orthologs or ortholog variants with coverage between 1-2000x were used for plotting (discarding a total of 83 ortholog or ortholog variants). Orthologs with ≥ 4 variants were a maximum of 14 per species and were also not plotted. Tbi: *Timema bartmani*, Tte: *T. tahoe*, Tce: *T. cristinae*, Tms: *T. monikensis*, Tps: *T. poppensis*, Tdi: *T. douglasi*, Tcm: *T. californicum*, Tsi: *T. shepardi*, Tpa: *T. podura*, Tge: *T. genevievae*.



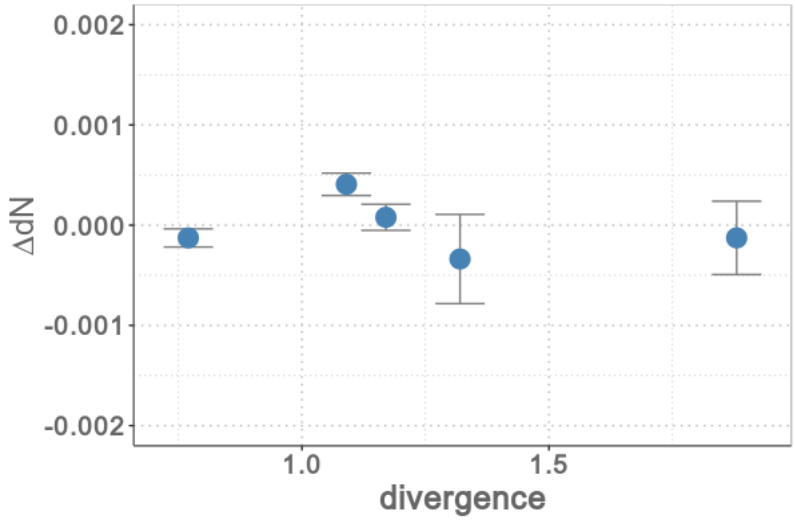
Supplementary Figure 2: Reads corresponding to variants of specific orthologs mapped to the best ortholog if the variants were excluded. Read counts for the best ortholog only (in light blue) and summed counts for reads mapping to all variants, including the best ortholog (in dark blue) were similar (paired t-test, t = -0.023, df = 2704, *P* = 0.982). As expected given the larger number of variants in asexuals than sexuals, mapping quality was somewhat lower in asexuals than sexuals, however not significantly so (asexuals: range 99.84 - 99.89, sexuals: range 99.86 - 99.91; paired t-test, t = 1.3, df = 7.9, *P* = 0.215). Tbi: *Timema bartmani*, Tte: *T. tahoe*, Tce: *T. cristinae*, Tms: *T. monikensis*, Tps: *T. poppensis*, Tdi: *T. douglasi*, Tcm: *T. californicum*, Tsi: *T. shepardi*, Tpa: *T. podura*, Tge: *T. genevievae*.



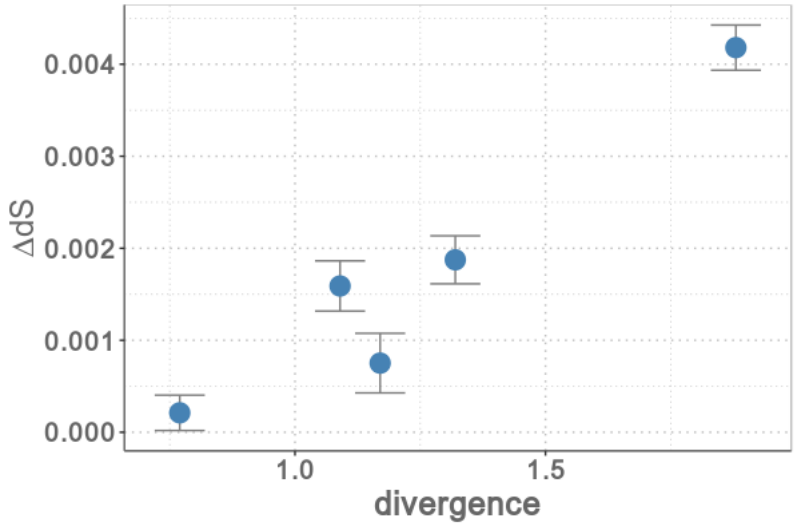
Supplementary Figure 3: *Timema* species exhibit codon usage bias and the effective number of codons used does not differ between sexuals asexuals (means and 95% CI). Usage of fewer than 61 codons indicates that some codons are used more frequently than others. Note that the axis does not start at zero. Tbi: *Timema bartmani*, Tte: *T. tahoe*, Tce: *T. cristinae*, Tms: *T. monikensis*, Tps: *T. poppensis*, Tdi: *T. douglasi*, Tcm: *T. californicum*, Tsi: *T. shepardi*, Tpa: *T. podura*, Tge: *T. genevievae*.



Supplementary Figure 4: Base composition difference increases with asexual lineage age. Euclidian distance between sister-species for GC3% means and GC3 std correlated with divergence between the sexual and asexual species within a pair.



Supplementary Figure 5: The rate of dN does not increase with asexual lineage age (Pearson's product-moment correlation; t = -0.450, cor -0.249, *P* = 0.686). Means (and 95% CI) of $\Delta dN_{asex\text{-}sex}$ correlated with divergence between the sexual and asexual species within a pair.



Supplementary Figure 6: The rate of dS increases with asexual lineage age. Means (and 95% CI) of $\Delta dS_{asex\text{-}sex}$ correlated with divergence between the sexual and asexual species within a pair.

Supplementary Table 1: Statistics for the quality filtered assemblies.

Species	**Pair**	**Reproductive mode**	**No. of transcripts**	**Longest transcript length (bp)**	**N50**	**No. of ORFs**
T. bartmani	**1**	sex	31,747	21,998	1,835	7,741
T. cristinae	**2**	sex	34,177	2,9940	1,729	7,703
T. californicum	**4**	sex	32,416	28,616	1,737	8,026
T. douglasi	**3**	asex	37,652	24,333	1,530	8,910
T. genevievae	**5**	asex	44,740	28,013	1,499	9,409
T. monikensis	**2**	asex	43,326	27,229	1,512	10,061
T. podura	**5**	sex	45,655	30,228	1,301	9,392
T. poppensis	**3**	sex	32,167	30,132	1,756	7,136
T. shepardi	**4**	asex	41,642	29,999	1,493	10,196
T. tahoe	**1**	asex	42,372	24,302	1,640	9,634
mean sexual species			35,232	28,183	1,672	8,000
mean asexual species			41,946	26,775	1,535	9,642

Supplementary Table 2: Number of orthologs per species pair.

Species pair	**No. of pairwise orthologs**
T. bartmani - T. tahoe	5,908
T. cristinae - T. monikensis	5,806
T. californicum - T. shepardi	5,754
T. poppensis - T. douglasi	5,476
T. podura - T. genevievae	5,329

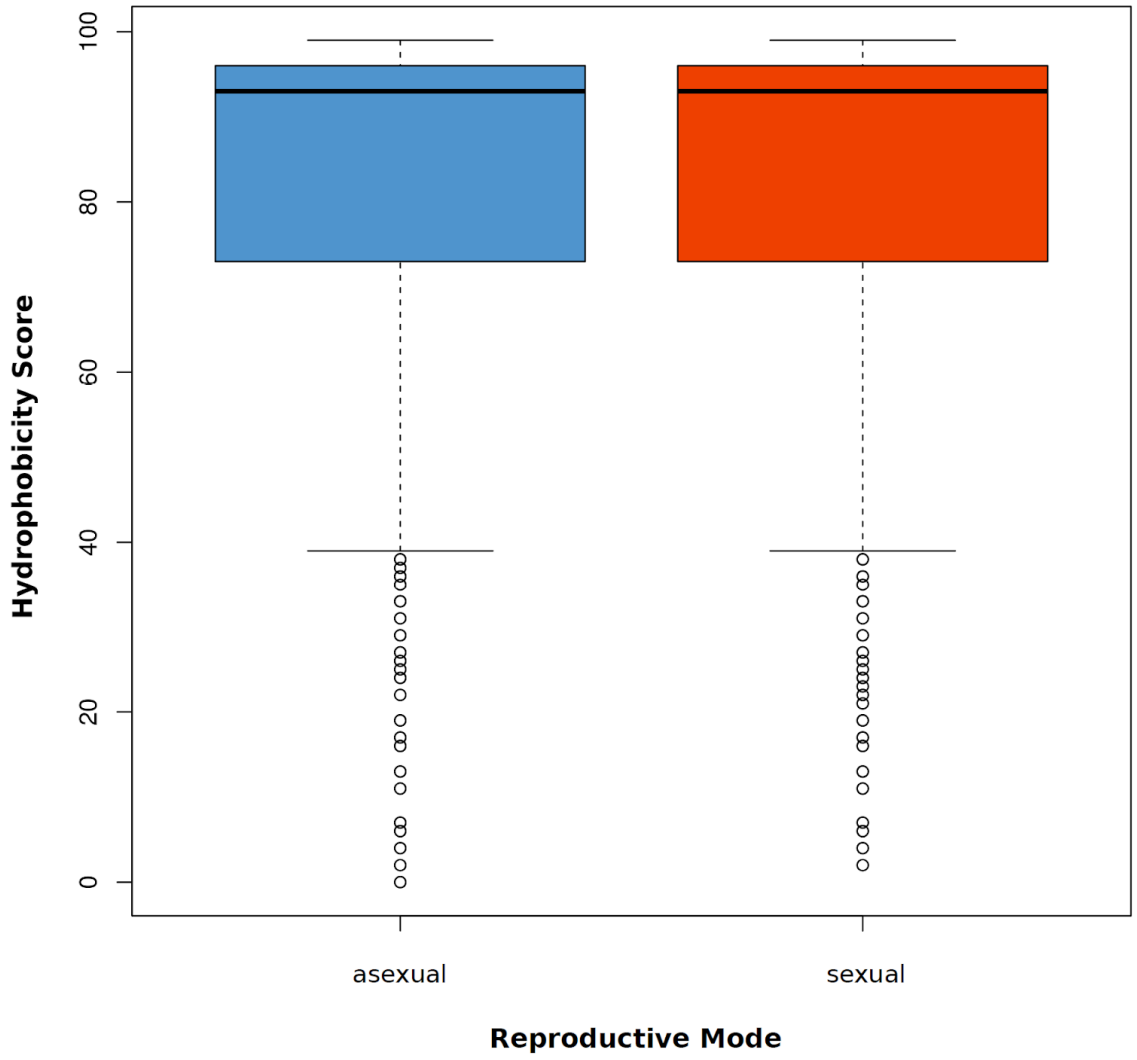
Supplementary Table 3: Coordinates of the sampling locations.

Species	**Latitudes**	**Longitudes**
T. bartmani	34.170000	-117.002017
T. californicum	37.222860	-122.087231
T. cristinae	34.522300	-119.831283
T. douglasi	39.200783	-123.293933
T. genevievae	38.995783	-122.925767
T. monikensis	34.102933	-118.861783
T. podura	33.797603	-116.776878
T. poppensis	36.996433	-121.717783
T. shepardi	38.982550	-123.469750
T. tahoe	38.761011	-120.160053

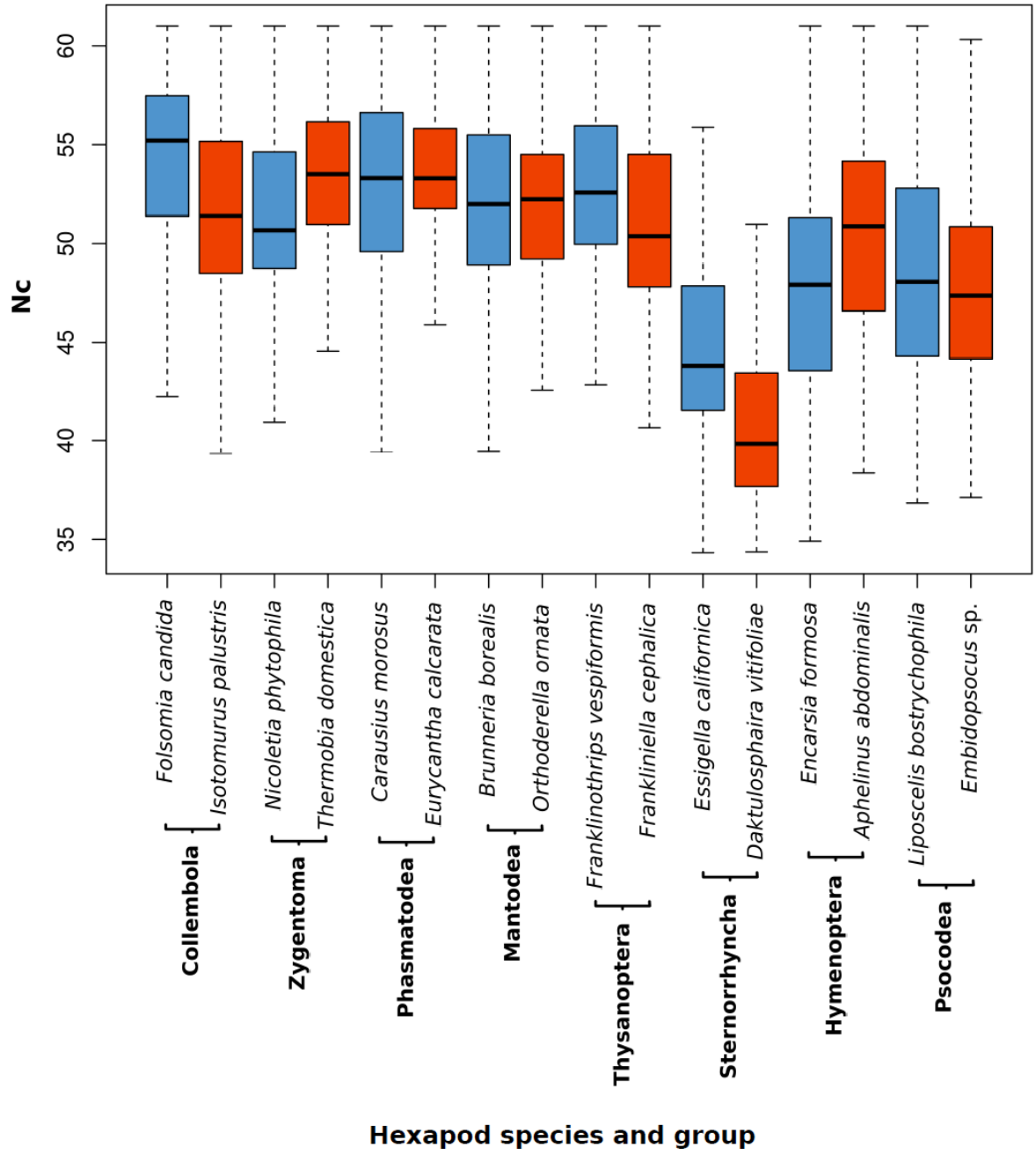
Research Chapter 3

No signal of deleterious mutation accumulation in conserved sequences of extant asexual hexapods

For a complete collection of Supplementary Data, see CD, attached.



Supplementary Figure S1: Hydrophobicity scores at asexual and sexual terminal branches. The boxplots show Hydrophobicity Scores (HS) of eight asexual (blue) and eight sexual (red) terminal branches for 73 nuclear orthologous genes shared among the 16 hexapod species and the outgroup *Xibalbanus* cf. *tulumensis* (involving 3,841 and 3,172 HS of asexual and sexual terminal branches, respectively). HS indicates the 'deleteriousness' of a nonsynonymous mutation by measuring the strength in hydrophobicity change from ancestral to replacement amino acid. The lower the HS the stronger is the change in hydrophobicity from ancestral to replacement amino acid and, hence, the deleteriousness of the underlying nonsynonymous mutation. There was no difference in hydrophobicity changes between reproductive modes (z = -0.152; P = 0.879; GLMM). Whiskers correspond to 1.5 times the interquartile range.



Supplementary Figure S2: Effective Numbers of Codons (Nc) of 16 hexapod species. The boxplots show the distributions of Nc of 99 orthologue loci for eight asexual (obligate asexual species present in the 1kite transcriptome data set; blue boxes) and eight sexual (red boxes) hexapod species, covering eight hexapod groups. For better representation, the ordinate is restricted to between 34 and 60; outliers outside of 1.5 times the interquartile range (whiskers) are excluded. Nc specifies the deviation of observed codon usage from equal usage of all codons ranging from 20 (each amino acid is encoded by one codon only; strong CUB) to 61 (equal use of all possible codons; no CUB).

Supplementary Table S1: Additional information on sequence data analysed in this study.

Species	**First describer**	**Transcriptome assembly previously published in:**	**BioProject Accession of already published transcriptomes**	**BioSample Accession of species samples used in this study**
Aphelinus abdominalis	Dalman, 1820	Peters *et al.*, 2018	PRJNA252351	SAMN02870213
Brunneria borealis	Scudder, 1896	not applicable		SAMN03339328
Carausius morosus	Sinety, 1901	not applicable		SAMN04005132
Daktulosphaira vitifoliae	Fitch, 1855	not applicable		SAMN03331961
Embidopsocus sp.	Genus by Hagen, 1866	not applicable		SAMN03331967
Encarsia formosa	Gahan, 1924	Peters *et al.*, 2018	PRJNA252167	SAMN02870290
Essigella californica	Essig, 1909	Misof *et al.*, 2014	PRJNA219554	SAMN02047099
Eurycantha calcarata	Lucas, 1869	not applicable		SAMN04005157
Folsomia candida	Willem, 1902	Misof *et al.*, 2014	PRJNA219557	SAMN02047120
Frankliniella cephalica	Crawford, 1910	Misof *et al.*, 2014	PRJNA219559	SAMN02047110
Franklinothrips vespiformis	Crawford, 1909	not applicable		SAMN03331968
Isotomurus palustris	Mueller, 1776	not applicable		SAMN03142432
Liposcelis bostrychophila	Badonnel, 1931	Misof *et al.*, 2014	PRJNA219573	SAMN02047187
Nicoletia phytophila	Gervais, 1844	not applicable		SAMN03142452
Orthoderella ornata	Giglio-Tos, 1897	not applicable		SAMN04005206

Thermobia domestica	Packard, 1873	Misof *et al.*, 2014	PRJNA219608	SAMN02047119
Xibalbanus cf. *tulumensis*	Yager, 1987	Von Reumont *et al.*, 2012	PRJNA66987	not available

Supplementary Table S2: GenBank Accession numbers of sequences (Open Reading Frames; ORFs) analysed in this study sorted by OMA orthologue group IDs. 70 orthologue groups include sequences of *X*. cf. *tulumensis* (used for analyses of 'deleteriousness' of nonsynonymous mutations; 17 sequences per orthologue group). For three orthologue groups, sequences of *X*. cf. *tulumensis* remained unchanged after ORF extraction and represent duplicates of sequences of the published transcriptome (Von Reumont *et al.*, 2012).

OMA orthologue group ID	**GenBank Accession**	**No. of sequences per orthologoue group**	**Remarks**
9945	MH638049 - MH638065	17	
9834	MH638032 - MH638048	17	
9690	MH638015 - MH638031	17	
9567	MH637998 - MH638014	17	
9559	MH637981 - MH637997	17	
9512	MH637964 - MH637980	17	
9320	MH637947 - MH637963	17	
9266	MH637930 - MH637946	17	
9067	MH637913 - MH637929	17	
9051	MH637896 - MH637912	17	
8832	MH637879 - MH637895	17	
8688	MH637862 - MH637878	17	
8442	MH637845 - MH637861	17	

8063	MH637828 - MH637844	17	
7631	MH637812 - MH637827	16	
7586	MH800169 - MH800185	17	
7514	MH800152 - MH800168	17	
7342	MH800135 - MH800151	17	
7199	MH800118 - MH800134	17	
6996	MH800102 - MH800117	16	
6954	MH800085 - MH800101	17	
6755	MH800068 - MH800084	17	
6720	MH800051 - MH800067	17	
6343	MH800035 - MH800050	16	OG6343_X_tulumensis_s is identical to TSA Accession JL198982 (Von Reumont *et al.*, 2012).
6321	MH800018 - MH800034	17	
6087	MH800001 - MH800017	17	
6076	MH799984 - MH800000	17	
5435	MH799968 - MH799983	16	
5260	MH799951 - MH799967	17	
5220	MH799934 - MH799950	17	
4377	MH799918 - MH799933	16	

4250	MH799902 - MH799917	16	OG4250_X_tulumensis_s is identical to TSA Accession JL138502 (Von Reumont *et al.*, 2012).
4143	MH799885 - MH799901	17	
4101	MH799869 - MH799884	16	
3827	MH799853 - MH799868	16	
3814	MH799836 - MH799852	17	
3745	MH799820 - MH799835	16	
3632	MH799803 - MH799819	17	
3464	MH799786 - MH799802	17	
3341	MH799770 - MH799785	16	
3278	MH799753 - MH799769	17	
2885	MH799737 - MH799752	16	
2874	MH799721 - MH799736	16	
2845	MH799704 - MH799720	17	
2780	MH799688 - MH799703	16	
2453	MH799671 - MH799687	17	
2434	MH799655 - MH799670	16	
2366	MH799638 - MH799654	17	
2249	MH799622 - MH799637	16	

1942	MH799605 - MH799621	17	
1925	MH799589 - MH799604	16	
1874	MH799573 - MH799588	16	
1858	MH799556 - MH799572	17	
1815	MH799539 - MH799555	17	
1641	MH799523 - MH799538	16	
1518	MH799507 - MH799522	16	OG1518_X_tulumensis_s is identical to TSA Accession JL174507 (Von Reumont *et al.*, 2012).
1499	MH799490 - MH799506	17	
1031	MH799474 - MH799489	16	
375	MH551269 - MH551284, MH799473	17	
17763	MH799456 - MH799472	17	
17620	MH799439 - MH799455	17	
17527	MH799423 - MH799438	16	
16696	MH799406 - MH799422	17	
16029	MH799390 - MH799405	16	
15828	MH799373 - MH799389	17	
15479	MH799356 - MH799372	17	
15365	MH799339 - MH799355	17	

14989	MH799322 - MH799338	17	
14697	MH602941 - MH602956	16	
14338	MH602924 - MH602940	17	
14182	MH602907 - MH602923	17	
14102	MH602890 - MH602906	17	
13754	MH602874 - MH602889	16	
13719	MH602857 - MH602873	17	
13695	MH602840 - MH602856	17	
13624	MH602823 - MH602839	17	
13609	MH602806 - MH602822	17	
12963	MH602790 - MH602805	16	
12890	MH602774 - MH602789	16	
12725	MH602758 - MH602773	16	
12223	MH602742 - MH602757	16	
11984	MH602725 - MH602741	17	
11941	MH602708 - MH602724	17	
11819	MH602691 - MH602707	17	
11738	MH602674 - MH602690	17	
11485	MH602657 - MH602673	17	

11151	MH602641 - MH602656	16	
10970	MH602624 - MH602640	17	
10837	MH602607 - MH602623	17	
10733	MH602590 - MH602606	17	
10625	MH602573 - MH602589	17	
10598	MH602556 - MH602572	17	
10571	MH602539 - MH602555	17	
10467	MH602522 - MH602538	17	
10418	MH602505 - MH602521	17	
10379	MH602488 - MH602504	17	
10272	MH602471 - MH602487	17	
10257	MH602454 - MH602470	17	
10011	MH602437 - MH602453	17	

List of archives available for download from DRYAD under doi:10.5061/dryad.5501rv4 (Supplementary Archives S1-S4):

Supplementary Archive S1 includes:

- curated codon alignments of 99 orthologue groups (OGs; FASTA-format) used for analyses of nonsynonymous and synonymous mutation accumulation (see Methods). Each alignment filename corresponds to the respective OG (see Supplementary Table S2).
- CodeML control file (PAML version 4.9) used for analyses of nonsynonymous mutation accumulation
- Python script for passing fixed species trees with loci-specific branch lengths and curated codon alignments to CodeML for analyses of nonsynonymous mutation accumulation (detailed information is given in the script)

Supplementary Archive S2 includes:

- curated codon alignments of 73 orthologue groups (OGs; FASTA-format) used for analyses of 'deleteriousness' of nonsynonymous mutations (see Methods). Each alignment filename corresponds to the respective OG (see Supplementary Table S2).
- CodeML control file (PAML version 4.9) used for prediction of ancestral amino acid states
- Python script for passing fixed species trees with loci-specific branch lengths and curated codon alignments to CodeML for prediction of ancestral amino acid states (detailed information is given in the script)

Supplementary Archive S3 includes:

- R-script for statistical analyses of Hydrophobicity Scores ('deleteriousness' of nonsynonymous mutations; detailed information is given in the R-script)

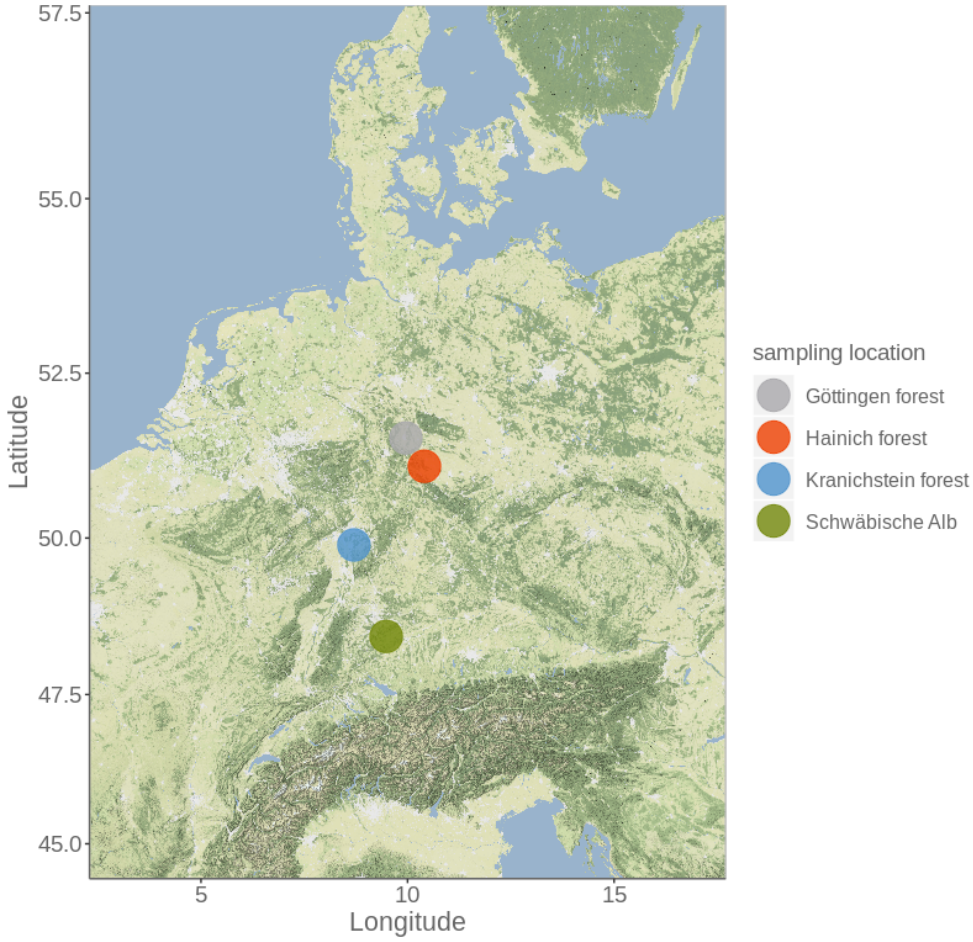
Supplementary Archive S4 includes:

- codon alignments of 286 orthologue groups resulting from initial orthology inference with OMA
- codon alignments of 153 orthologue groups resulting from searching the 286 initial orthologue groups against a precompiled set of orthologues with HaMStR (see Methods)

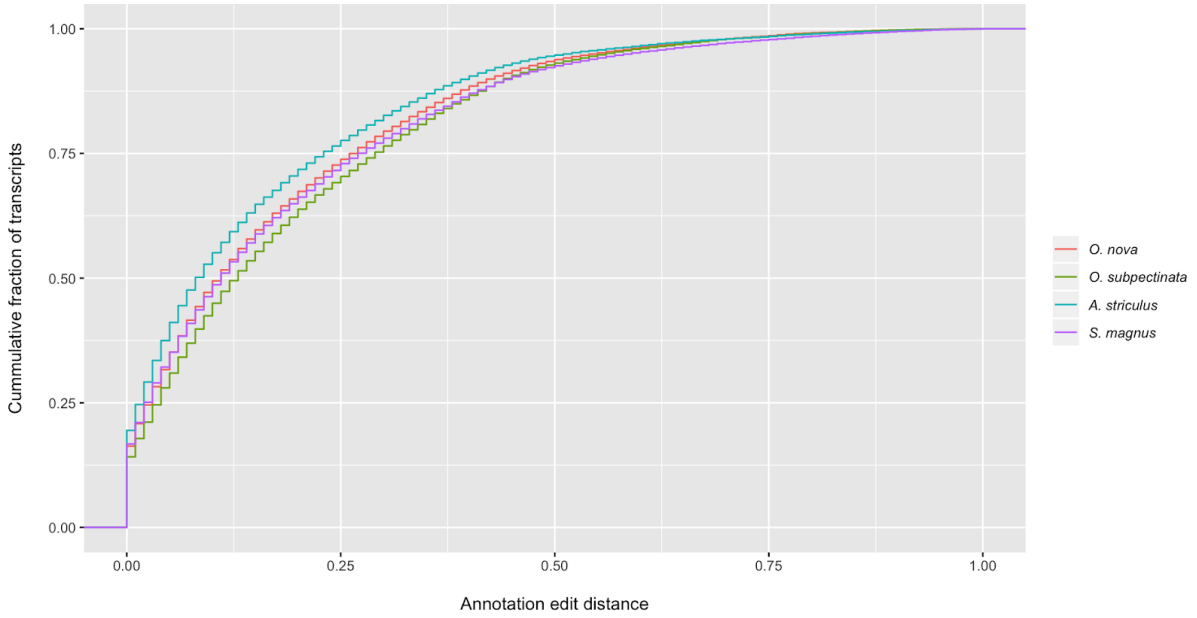
Research Chapter 4

Haplotype divergence supports obligate asexuality in the oribatid mite *Oppiella nova*

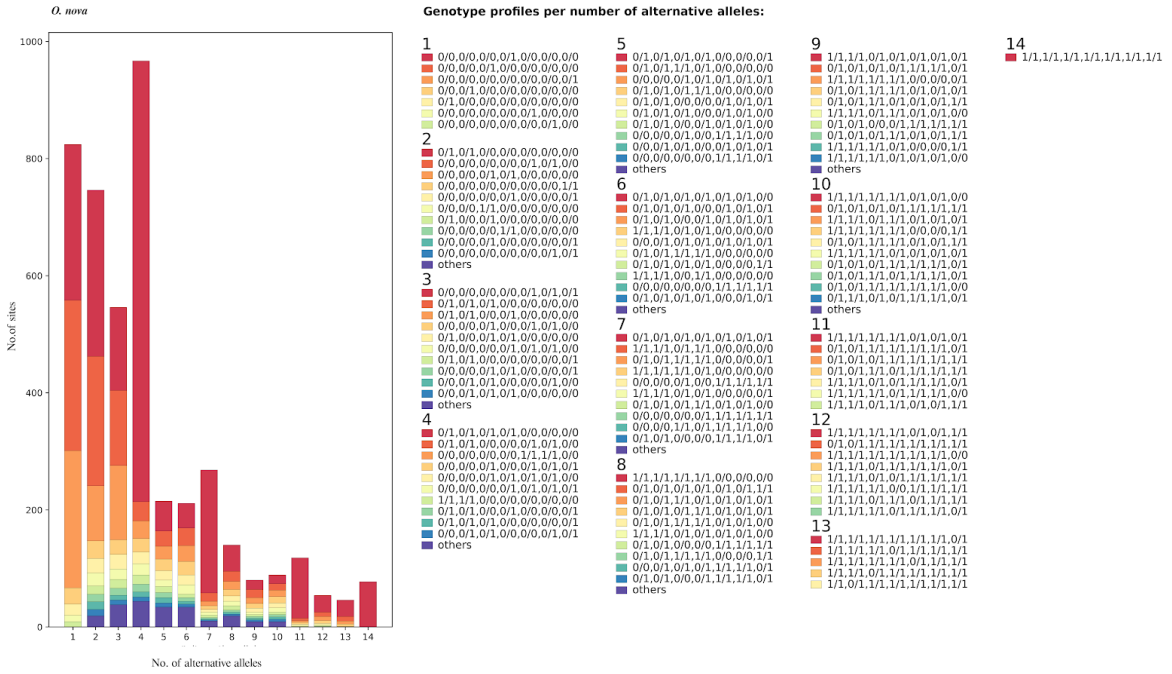
For a complete collection of Supplementary Data, see CD, attached.



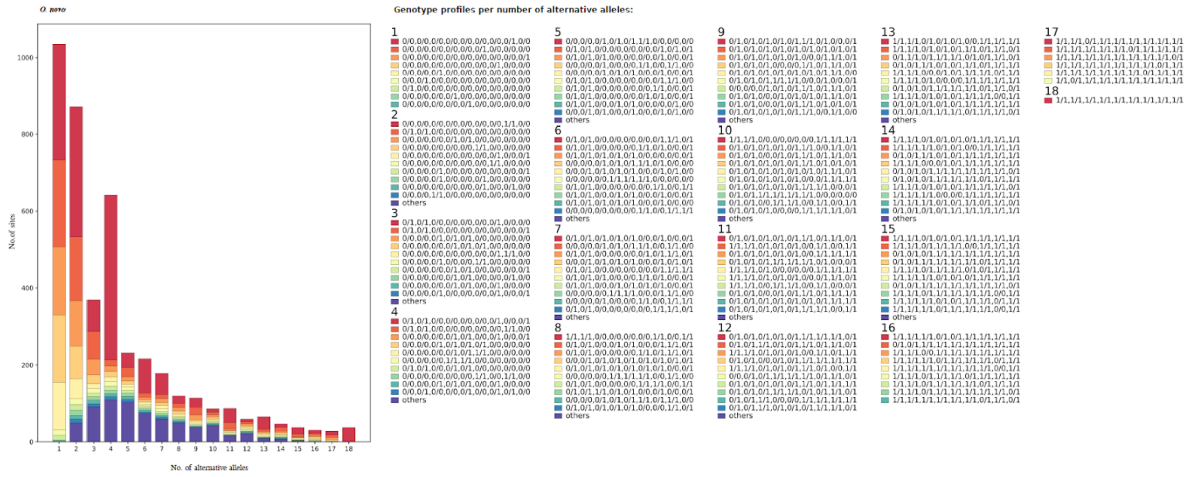
Supplementary Figure 1: Sampling sites. For detailed information on sampling, see Supplementary Table S8.



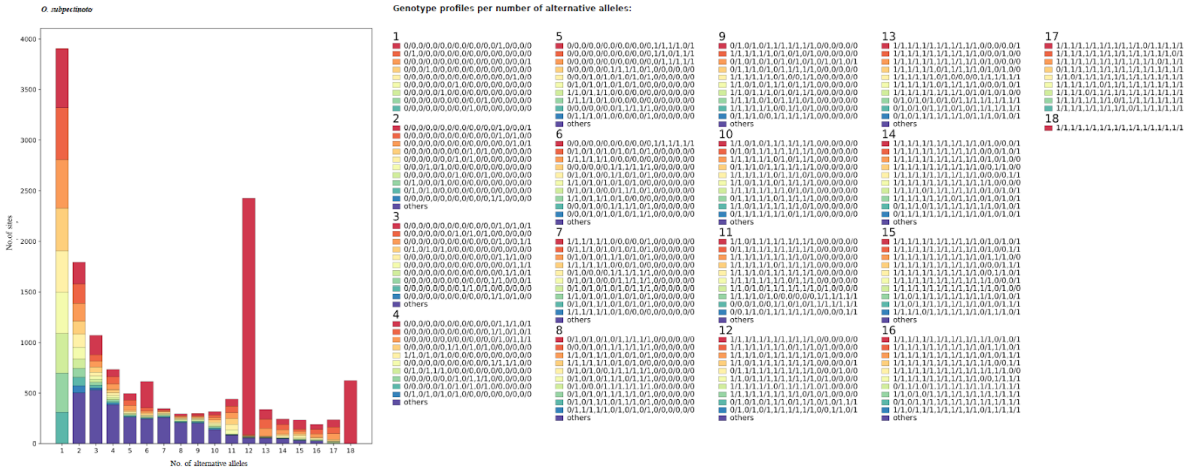
Supplementary Figure 2: Annotation edit distance (AED) distribution. The AED ranges from 0 to 1 and a low AED reflects good concordance between the models predicted and the supported evidence (transcript and protein evidence; Yandell & Ence, 2012). The annotation process resulted in between 20,498 and 23,761 protein coding genes depending on the species and between 93% and 95% of total transcripts have an AED below 0.5, which is consistent with a well annotated genome (Holt and Yandell, 2011).



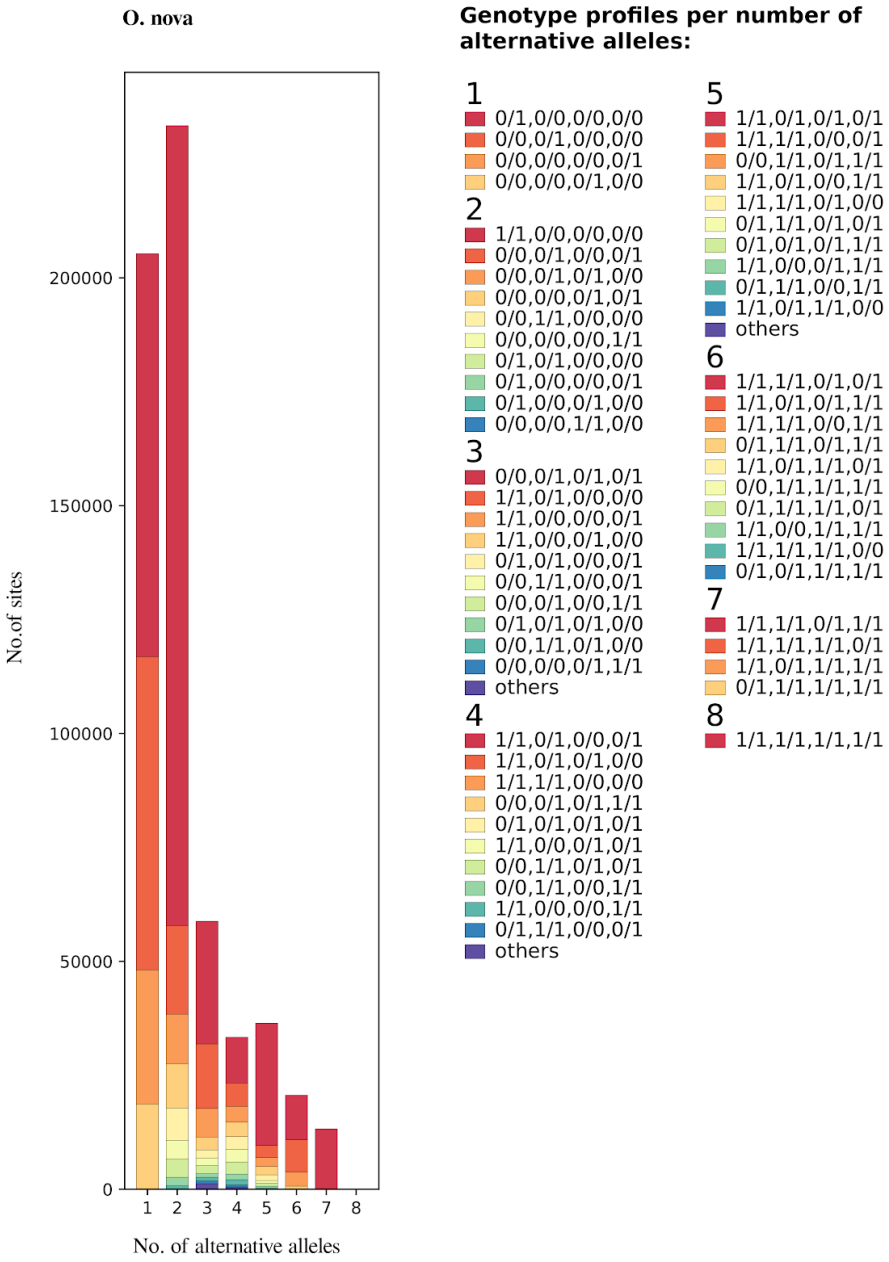
Supplementary Figure 3: Genotype profile site frequency spectrum (GPSFS) of seven individuals, stemming from three populations, for the asexual species *O. nova*. For better representation of shared heterozygous SNPs the GPSFS was constructed excluding one ind. from SA and KF, that both lacked an excess of heterozygous SNPs shared with the other individuals (for the full GPSFS, see Supplementary Figure 4). The GPSFS shows an overrepresentation of shared heterozygous SNPs among all seven individuals (genotypes of individuals from the three populations H, KF and SA in the following order H1,H2,SA1,SA3,KF1,KF2,H3; 0/1,0/1,0/1,0/1,0/1,0/1,0/1). Further, there is an overrepresentation of shared heterozygous SNPs private to one diverged lineage comprising individuals one and two from H and individuals one and three from SA (0/1,0/1,0/1,0/1,0/0,0/0,0/0 and 0/1,0/1,0/1,0/1,1/1,1/1,1/1). Also, there is an overrepresentation of shared heterozygous SNPs private to a different lineage comprising individuals one and two from KF and three from H (0/0,0/0,0/0,0/0,0/1,0/1,0/1 and 1/1,1/1,1/1,1/1,1/0,1/0,1/0,1/1). The reference genome of *O. nova* was generated from one individual from KF resulting in an asymmetric overrepresentation of genotypes, i.e. stronger overrepresentation of 0/1,0/1,0/1,0/1,0/0,0/0,0/0 and 1/1,1/1,1/1,1/1,0/1,0/1,0/1, respectively. For detailed interpretation of the results, see ‘Results’ section of the main article.



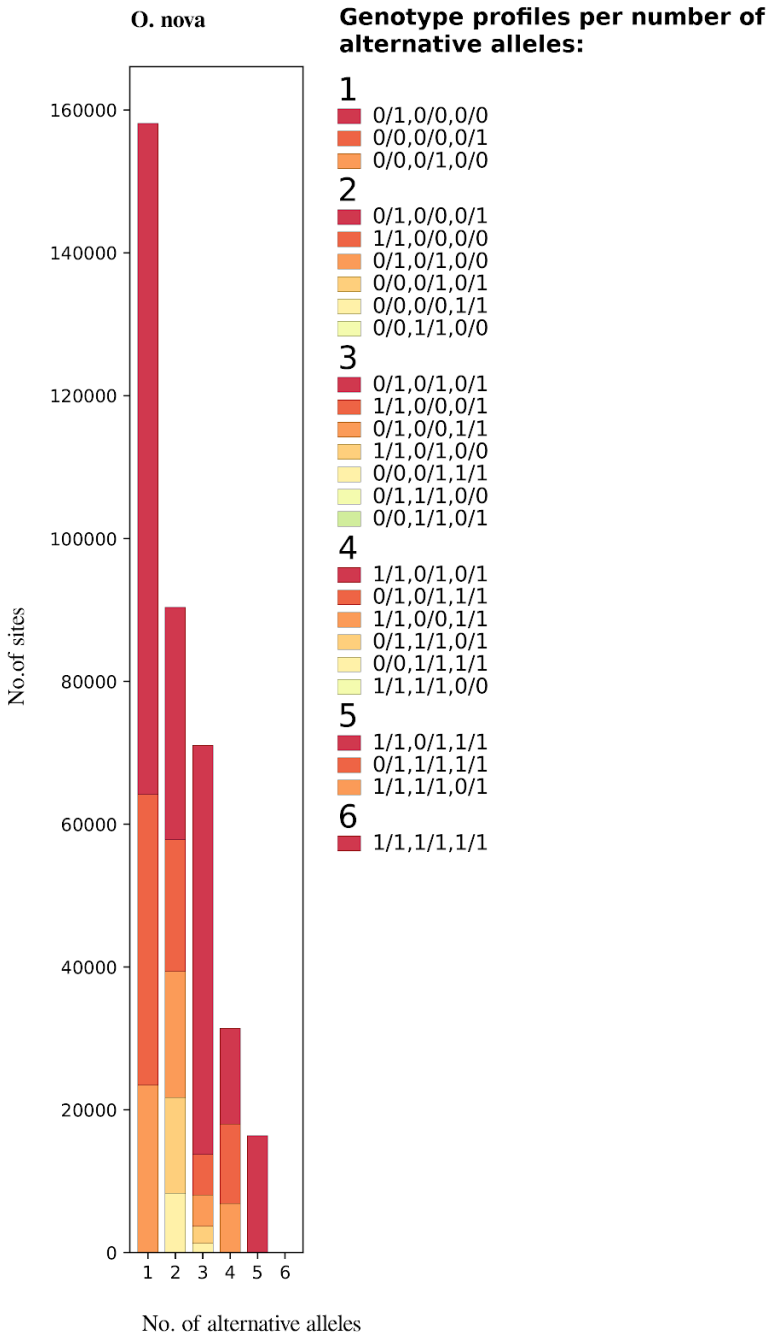
Supplementary Figure 4: Genotype profile site frequency spectrum (GPSFS) based on SNPs derived from RNAseq data of three individuals from each three populations of *O. nova*. The histogram depicts the number of sites with a given number of alternative variants over all individuals (nine diploid individuals can display a maximum of 18 alternative variants). Colours indicate distinct combinations of alternative alleles over individuals, i.e. genotype profiles for each number of alternative alleles (see colour legend; individuals ordered by populations Hainich, Kranichstein forest, Schwäbische Alb). Note the overrepresentation of homozygous SNPs private to individual two from SA (genotypes of individuals from the three populations H, KF and SA in the following order H1,H2,H3,KF1,KF2,KF3,SA1,SA2,SA3; 0/0,0/0,0/0,0/0,0/0,0/0,0/0,0/0,0/0,1/1,0/0). For detailed interpretation of the results, see 'Results' section of the main article.



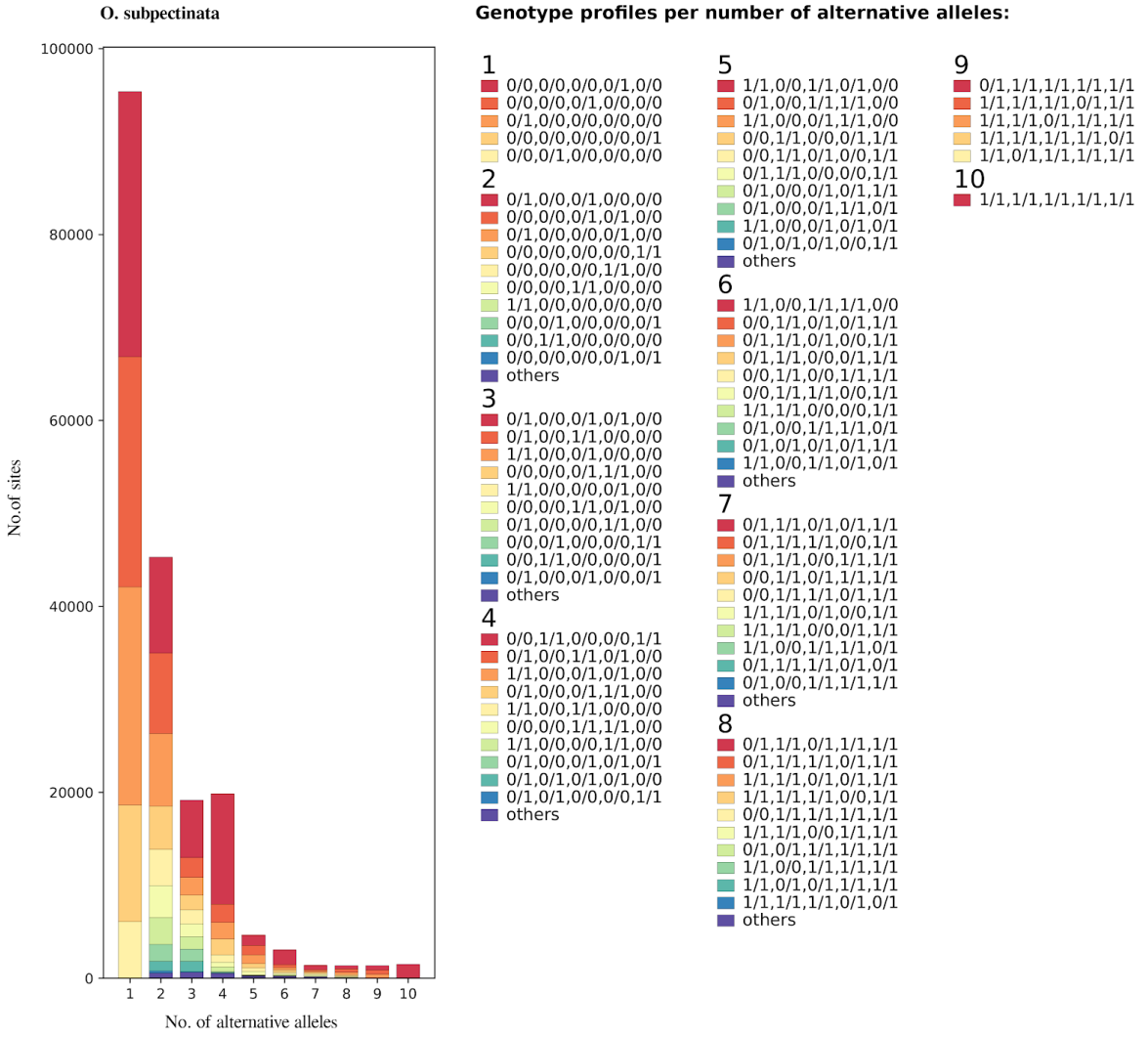
Supplementary Figure 5: Genotype profile site frequency spectrum (GPSFS) based on SNPs derived from RNAseq data of three individuals from each three populations of *O. subpectinata*. The histogram depicts the number of sites with a given number of alternative variants over all individuals (nine diploid individuals can display a maximum of 18 alternative variants). Colours indicate distinct combinations of alternative alleles over individuals, i.e. genotype profiles for each number of alternative alleles (see colour legend; individuals ordered by populations Hainich, Kranichstein forest, Schwäbische Alb). There is an excess of homozygous alternative genotypes separating the population from SA from the populations from H and KF (H1,H2,H3,KF1,KF2,KF3,SA1,SA2,SA3; 0/0,0/0,0/0,0/0,0/0,0/0,0/0,1/1,1/1,1/1 and 1/1,1/1,1/1,1/1,1/1,1/1,1/1,0/0,0/0,0/0). The reference genome of *O. subpectinata* was generated from one individual from GF resulting in overrepresentation of homozygous alternative genotypes shared among all individuals (from H, KF, SA; 18 alternative alleles). For detailed interpretation of the results, see 'Results' section of the main article.



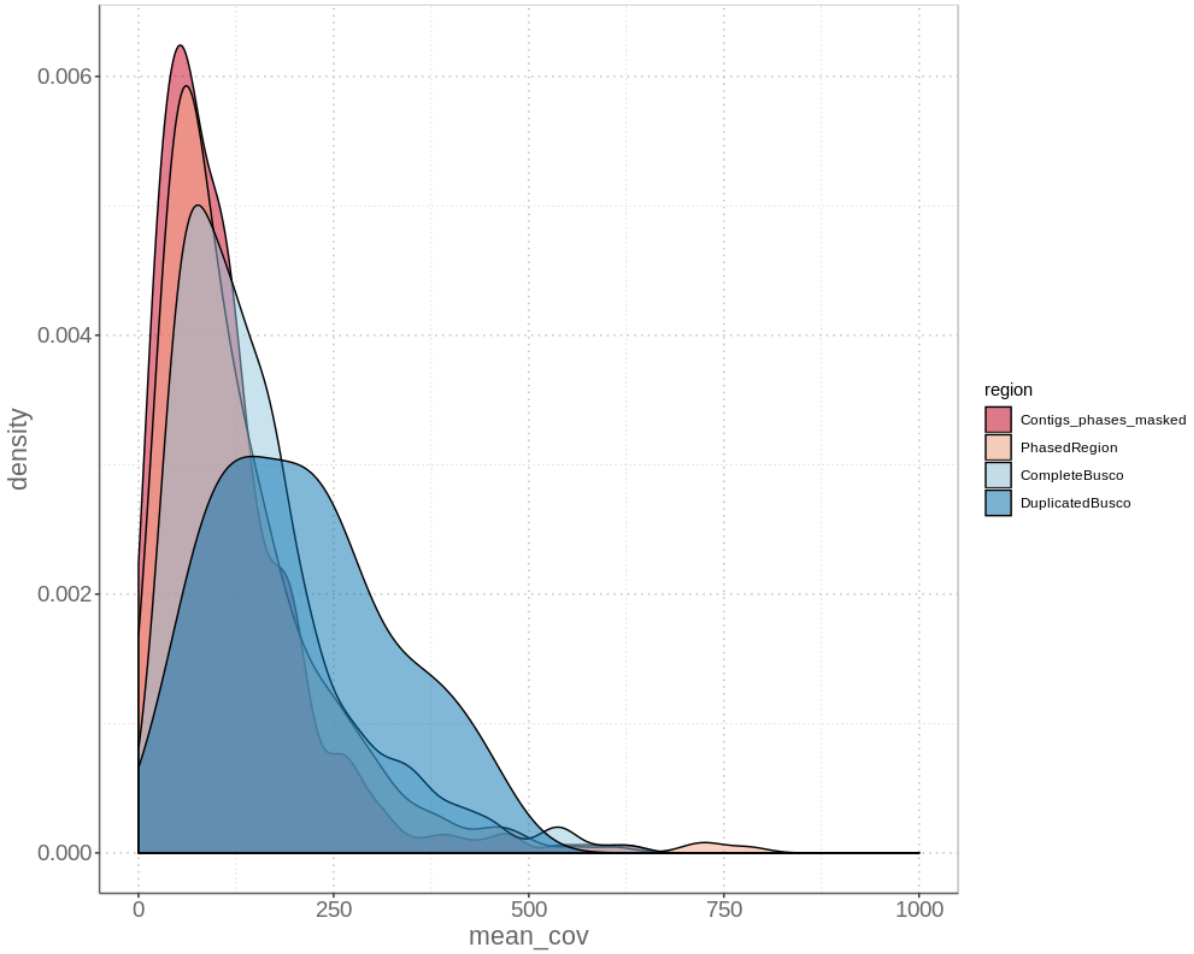
Supplementary Figure 6: Genotype profile site frequency spectrum (GPSFS) based on SNPs derived from reseq data of four individuals of *O. nova*. The histogram depicts the number of sites with a given number of alternative variants over all individuals (four diploid individuals can display a maximum of eight alternative variants). Colours indicate distinct combinations of alternative alleles over individuals, i.e., genotype profiles for each number of alternative alleles (see colour legend). There is an excess of homozygous SNPs separating individual one from the remaining three individuals (genotypes of individuals from KF in the following order KF1,KF2,KF3,KF4; 1/1,0/0,0/0,0/0). Further, there is an excess of heterozygous SNPs shared among the remaining individuals (individuals two, three and four; 0/0,0/1,0/1,0/1 and 1/1,0/1,0/1,0/1). For detailed interpretation of the results, see 'Results' section of the main article.



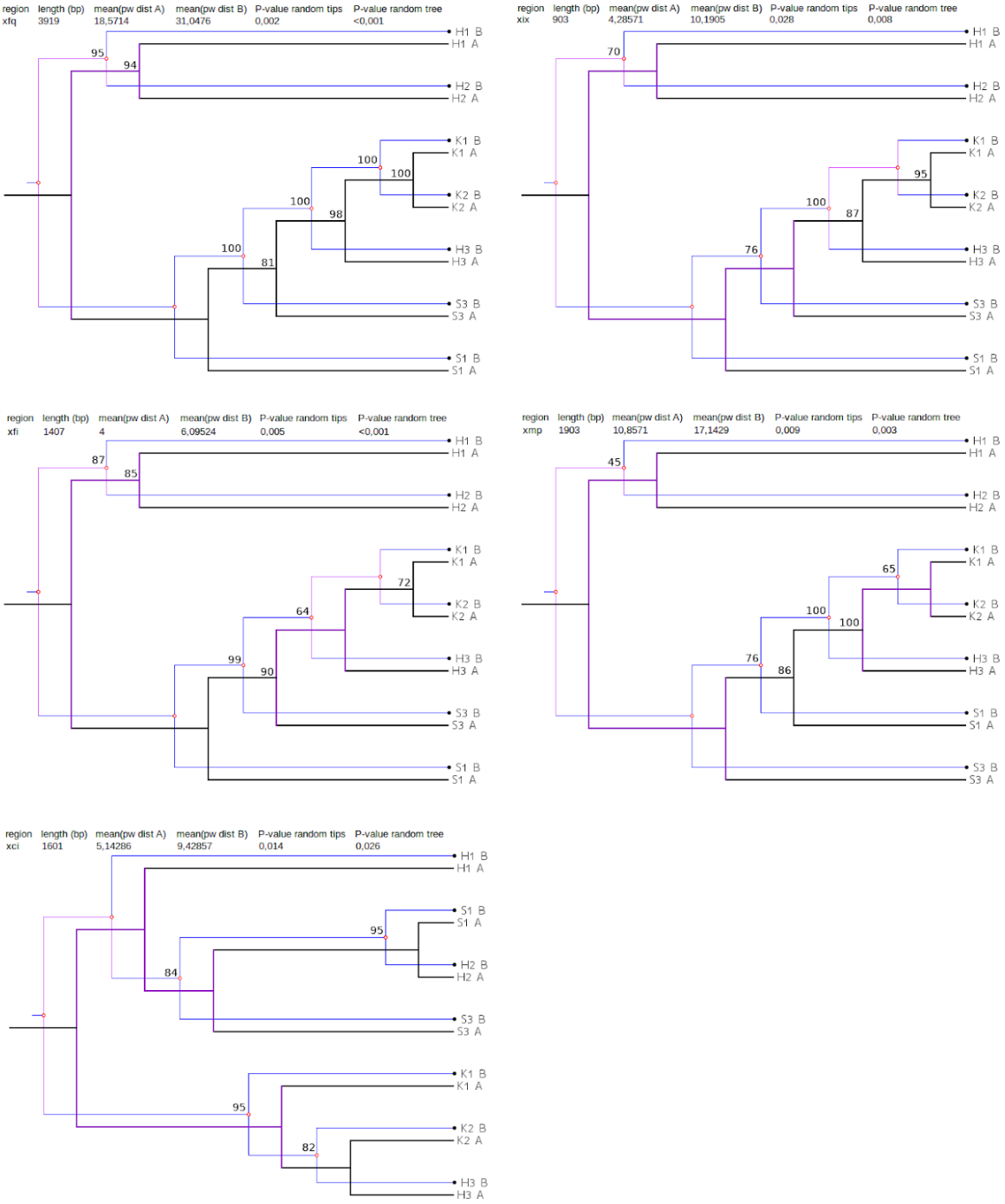
Supplementary Figure 7: Genotype profile site frequency spectrum (GPSFS) based on SNPs derived from reseq data of three individuals of *O. nova* (individual one removed). The histogram depicts the number of sites with a given number of alternative variants over all individuals (three diploid individuals can display a maximum of six alternative variants). Colours indicate distinct combinations of alternative alleles over individuals, i.e., genotype profiles for each number of alternative alleles (see colour legend). The histogram shows an excess of heterozygous sites shared among all individuals (genotypes from KF in the following order KF2,KF3,KF4; 0/1,0/1,0/1). For detailed interpretation of the results, see 'Results' section of the main article.



Supplementary Figure 8: Genotype profile site frequency spectrum (GPSFS) based on SNPs derived from reseq data of five individuals of *O. subpectinata*. The histogram depicts the number of sites with a given number of alternative variants over all individuals (five diploid individuals can display a maximum of ten alternative variants). Colours indicate distinct combinations of alternative alleles over individuals, i.e., genotype profiles for each number of alternative alleles (see colour legend). There is an excess of homozygous SNPs separating individuals two and five from the remaining three individuals (genotypes of individuals from GF in the following order GF1,GF2,GF3,GF4,GF5; 0/0,1/1,0/0,0/0,1/1 and 1/1,0/0,1/1,1/1,0/0). For detailed interpretation of the results, see 'Results' section of the main article.



Supplementary Figure 9: Coverage comparisons indicate that phased regions are not derived from paralogs. Kernel densities of coverages of the different regions: contigs that contained the phased regions but were masked for these, the phased regions, complete single-copy genes and duplicated genes as identified by BUSCO.



Supplementary Figure 10: Results of the co-phylogeny analysis and additional information for five phaseable regions that showed perfect parallel divergence (see Results). Violet branches indicate polytomies that were automatically resolved in favour of cospeciation. Bootstrap support at nodes indicated. H: Hainich; K: Kranichstein forest; S: Schwäbische Alb.

Supplementary Table S1: Genome statistics and completeness scores. v02: decontaminated assembly, v03: decontaminated assembly with > 500bp scaffold length size filter. C: complete; P: partial; S: single-copy; D: duplicate; F: fragmented; M: missing; n: No. BUSCO genes.

Species	CEGMA	BUSCO	Genome size (Mb)	Scaffold N50 (bp)	Number of scaffolds	Largest scaffold length (bp)	GC content (%)	N (%)
1 Os v01			217.82	6,873	73,173			1.44
Blobtools v02	C: 88.31 P: 97.58	C:86.6%[S:78.6%,D:8.0%], F:6.3%,M:7.1%,n:1066	217.02	6,852	73,101	165,769	30.84	1.44
Filtered v03	C: 88.31 P: 97.58	C:86.2%[S:78.3%,D:7.9%], F:6.4%,M:7.4%,n:1066	213.17	7,017	60,250	165,769	30.83	1.47
1 On v01			207.08	6,604	83,865			1.64
Blobtools v02	C: 88.31 P: 95.56	C:87.1%[S:78.8%,D:8.3%], F:6.6%,M:6.3%,n:1066	202.44	6,493	82,784	406,770	30.41	1.68
Filtered v03	C: 88.31 P: 95.56	C:87.5%[S:78.9%,D:8.6%], F:6.6%,M:5.9%,n:1066	196.72	6,753	63,118	406,770	30.37	1.73

Supplementary Table S2: Information on the read library insert sizes, number of reads generated and their coverage and number of surviving read pairs after filtering used for genome assembly.

Species	Individual	Library	No. of raw read pairs	Coverage	No. of read pairs surviving
O. nova	On6	180	125,889,255	137	95,460,017 (75.83%)
	On6	350	73,798,941	80	44,494,871 (60.29%)
	On6	550	76,545,499	83	34,976,565 (45.69%)
	On6	3,000	174,946,393	190	33,642,006 (53.99%)
O. subpectinata	Os5	180	112,721,137	122	89,023,834 (78.98%)
	Os5	350	87,673,896	95	65,281,720 (74.46%)
	Os5	550	71,743,398	78	40,465,571 (56.40%)
	Os5	3,000	115,207,485	125	21,898,456 (53.96%)

Supplementary Table S3: Annotation statistics for the four mite genomes.

Species	Gene number	mRNA number	Gene min length [bp]	Gene mean length [bp]	Gene max length [bp]	Genome covered by genes [%]	Exon mean length [bp]	Intron mean length [bp]
1 Os	23,555	23,806	49	2,148	53,142	23.7	258	246
1 On	23,761	24,140	30	1,956	42,876	23.6	258	270

Supplementary Table S4: AMOVA statistics of RNAseq data per species.

Species	**Stratification levels**	**% variance**	**φ-statistic**	***P*-value (randtest)**
O. nova (9 individuals)	between populations	12.791	0.128	0.06
	between individuals within populations	35.945	0.412	0.005
	within individuals	51.265	0.487	< 0.001
O. nova (7 clustering individuals)	between lineages	26.053	0.261	0.337
	between populations within lineages	14.396	0.195	0.116
	between individuals within populations	4.236	0.071	0.41
	within individuals	55.315	0.447	< 0.001
O. subpectinata	between populations	57.08	0.571	0.003
	between individuals within populations	14.83	0.346	< 0.001
	within individuals	28.093	0.719	< 0.001

Supplementary Table S5: Detailed information on successfully phased regions.

Species	**No. phased regions**	**Total length**	**No. phased regions > 100bp after removal of sites with coverage < 10**	**Total length**	**No. phased regions after removal of regions with sequence similarity too large for constrained tree calculation**	**Total length**	**Median length**
O. nova	386	1,453,429	298	214,720	280	210,375	473
O. subpec-tinata	470	3,217,849	401	415,031	393	410,610	628

Supplementary Table S6: Results of ML based tree topology (AU) tests. The table lists numbers of haplotype sequence alignments of phased regions that were fit significantly better by a constrained tree separating haplotypes than by a constrained tree separating populations (HD-tree + sex-tree -) and vice versa (HD-tree - sex-tree +) for the asexual species *O. nova* and its sexual sister *O. subpectinata*. The majority of phased regions of *O. nova* is fit better by a constrained tree separating haplotypes than by a constrained tree separating populations whereas in *O. subpectinata* the majority of phased regions is fit better by a population-separating tree. A number of phaseable regions in both species shows no significant difference in the goodness of fit between the constrained trees (HD-tree + sex-tree +). For some phased regions AU tests failed because of too low numbers of unique sequences (HD-tree - sex-tree -).

	O. nova	***O. subpectinata***
HD-tree + sex-tree -	166	7
HD-tree - sex-tree +	40	315
HD-tree + sex-tree +	73	68
HD-tree - sex-tree -	1	3

Supplementary Table S7: Additional information on haplotype-separating regions used for co-phylogeny analysis (with mean number of pairwise differences for both haplotypes > 4) and results of co-phylogeny analysis.

Region	**Length (bp)**	**Mean (pw diff. A)**	**Mean (pw diff B)**	**P-value random tips**	**P-value random tree**	**Cospeciation events**
xfq	3,919	18.5714	31.0476	0.002	<0.001	6
xmp	1,903	10.8571	17.1429	0.009	0.003	6
xdm	5,392	8.57143	22.381	0.212	0.07	5
xkh	1,843	8.57143	12.0952	0.132	0.065	5
xka	843	6.7619	14.4762	0.011	0.006	5
xci	1,601	5.14286	9.42857	0.014	0.026	6
xep	1,663	4.95238	18.1905	0.01	<0.001	5
xix	903	4.28571	10.1905	0.028	0.008	6
xfi	1,407	4	6.09524	0.005	<0.001	6

Supplementary Table S8: Additional information on sampling sites and species.

Sampling site	**GPS coordinates**	**Species for WGS and transcriptomes (sampled fall 2015)**	**Species for RNAseq (sampled fall 2017)**
Göttingen forest	51.533778,9.959861	*Oppiella subpectinata*	
Hainich	51.104556,10.408500		*Oppiella subpectinata*
			Oppiella nova
Kranichstein forest	49.892194,8.701889	*Oppiella nova*	*Oppiella subpectinata*
			Oppiella nova
Schwäbische Alb	48.442389,9.482250		*Oppiella subpectinata*
			Oppiella nova

Supplementary Table S9: Numbers of individuals used for generating different sets of sequencing data.

Species	**Reproductive mode**	**Number of individuals**			
		Reference genome	Reseq	Transcriptomes	HD
O. subpectinata	sexual	1	5	5 pooled	3x3 from diff. populations
O. nova	asexual	1	5	5 pooled	3x3 from diff. populations

Supplementary Table S10: numbers of read pairs/reads after different steps of quality trimming, contamination removal and mapping.

Species	**Population**	**Indivi-dual**	**No. raw read pairs**	**No. read pairs after trimming**	**No. read pairs after contaminant removal**	**% contam-inating read pairs**	**No. mapped reads after duplicate removal**
O. nova	Hainich	ONH1	12,305,768	10,922,239	5,073,080	53.33	1,538,502
		ONH3	11,653,023	10,150,566	1,546,946	84.76	434,144
		ONH7	17,293,884	14,456,993	3,959,499	72.61	866,110
	Kranichstein forest	ONK2	14,818,283	12,790,718	1,426,174	88.85	444,750
		ONK5	14,514,107	12,380,574	4,874,135	60.63	1,780,426
		ONK6	22,424,703	20,094,950	7,214,778	64.1	2,138,804
	Schwäbische Alb	ONS1	10,460,975	9,570,865	1,339,805	86	567,076
		ONS2	21,489,919	19,714,647	1,371,202	93.04	500,764
		ONS4	23,061,245	21,030,385	7,930,217	62.29	786,268
O. subpectinata	Hainich	OSH3	21,756,064	20,776,464	5,399,072	74.01	2,828,074
		OSH4	18,105,719	15,780,873	5,727,395	63.71	1,320,928
		OSH5	12,584,357	11,011,539	2,464,755	77.62	997,412
	Kranichstein forest	OSK1	26,438,129	23,518,688	4,829,639	79.46	3,472,828
		OSK2	36,588,456	34,945,277	3,386,946	90.31	4,238,092
		OSK6	28,975,922	23,341,626	11,877,834	49.11	3,279,550
	Schwäbische Alb	OSS3	21,616,073	20,271,910	12,090,364	40.36	4,230,422
		OSS4	12,186,958	10,229,333	2,695,017	73.65	739,384
		OSS5	15,924,247	13,972,229	3,810,147	72.73	1,178,238

Acknowledgements

[redacted]

Statutory Declaration

I, hereby, declare that this Ph.D. 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 Ph.D. at any other higher school of education.

Göttingen, 30 March 2020

Alexander Brandt

I, hereby, solemnly declare that this dissertation was undertaken independently and without any unauthorised aid.

Göttingen, 30 March 2020

Alexander Brandt