Variation of candidate genes related to climate change in European beech (*Fagus sylvatica* L.)

Dissertation
zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades
"Doctor rerum naturalium"
an der Georg-August-Universität Göttingen

Vorgelegt von

Sarah Seifert

aus Northeim

Göttingen, 2011

Referent: Prof. Dr. Reiner Finkeldey

Korreferentin: Prof. Dr. Andrea Polle

Tag der mündlichen Prüfung: 10. Januar 2012

Für Andreas

Acknowledgements

First of all, I would like to express my gratitude to Prof. Dr. Reiner Finkeldey for accepting me as his PhD student, for his support, helpful advice and guidance throughout my studies. I am very grateful that he made it possible for me to stay for two additional years as a Postdoc and I am looking forward to beginning with the new research project in this highly motivated and international working group.

Special thanks also to Prof. Dr. Polle who agreed to be my co-referee and examiner and to Prof. Dr. Becker, Prof. Dr. Leuschner, Prof. Dr. Ammer and Prof. Dr. Schütz to be members of my examination committee.

I am especially grateful to Dr. Vornam for her excellent guidance, for numerous invaluable scientific discussions, for her endless optimism and for motivating and believing in me and this project.

My deep gratitude to four persons for being not only colleagues, but for being friends during the last three years: Dr. Kathleen Prinz, Dr. Lesya Kuchma, Alexandra Dolynska and Markus Müller.

I thank Dr. Kathleen Prinz for our valuable scientific discussions and her advice concerning lab and field work, statistical questions, for improving manuscripts and for proof-reading my thesis; Prof. Dr. Ziehe and Dr. Gillet for helping me with questions concerning statistics; Dr. Amaryllis Vidalis for her valuable support and guidance concerning SNP analysis.

Many thanks to the indispensable technical team of the working group: Alexandra Dolynska, Gerold Dinkel, August Capelle and Christine Radler. Special thanks also to Regina Berkeley and Marita Schwahn who always helped me with the complicated paperwork and for their administrative support.

I am grateful to Markus Müller for his collaborative efforts on the KLIFF project, his support in field work, valuable input and for providing microsatellite data for the population in the Harz Mountains.

All current and past members of the working group supported me in some way, either scientifically or helping me with lab or field work: Dr. Ludger Leinemann, Prof. Dr. Hans Heinrich Hattemer, Prof. Dr. Oliver Gailing, Yazar Minn, Randy Villarin, Essy Harnelly,

Table of contents

1   Introduction

The recent global climate change will alter the environmental conditions faster than in the past. In Central Europe, the most important changes will be higher annual mean temperatures and decreasing precipitation during the summer months (EEA 2008, IPCC 2007). These changes require evolutionary adaptation for the conservation of forest tree populations (e.g., Aitken et al. 2008). Trees are main structural and functional elements of forest ecosystems. Therefore, it is important to investigate their ability to adapt to the changing environmental conditions. One of the most important deciduous tree species in Central Europe is *Fagus sylvatica*, European beech. This widely distributed and dominant tree species is of ecological importance but also of great economic interest (e.g., Bolte et al. 2007, Gömöry et al. 2003). Its adaptability to climate change is critically discussed in literature (e.g., Ammer et al. 2004, Rennenberg et al. 2004). It may be concluded that at least in some regions, European beech populations will be out of optimal conditions, but further research is recommended (e.g., Geßler et al. 2007, see chapter 1.2.3-1.2.5). At the moment, mainly studies about the neutral genetic variation of European beech are available (e.g., Buiteveld et al. 2007), but there are almost no studies analysing the adaptive genetic variation (see chapter 1.3). Furthermore, the genetic background of climate change related traits like drought stress tolerance and bud phenology is mainly unknown for European beech.

In this study, the genetic background of these traits is investigated by testing if different populations differ in their genetic constitution and hence in their ability to adapt to a changing climate. European beech populations along a precipitation gradient in the centre of the distribution were selected. Additionally, a controlled drought stress experiment in a climate chamber using beech seedlings was conducted. This study will help to develop strategies for the conservation of forest ecosystems in close collaboration with other workgroups within the research project "KLIFF – climate impact and adaptation research in Lower Saxony" (www.kliff-niedersachen.de).

1.1   Global climate change

The Intergovernmental Panel on Climate Change (IPCC) published their first assessment report in 1990 dealing with enhanced global warming due to human activities and its possible consequences. At this time, global climate change was rarely discussed in public. But over the years, the already changing climate and its impacts became more and more obvious. The Fourth IPCC Assessment Report in 2007 was extensively mentioned and discussed in media and politics. The report revealed that the impacts on ecosystems will have far reaching consequences if the emission of greenhouse gases will not be reduced. The next report composed by more than 800 authors is in progress, and the first part will be published in 2013, part two and three in 2014 (www.ipcc.ch). All predictions in the IPPC reports are based on emission scenarios explained in detail in the Special Report on Emissions Scenarios (SRES) in 2000. Four emission scenario families (A1, A2, B1 and B2) were defined differing in the degree of globalization (scenario families 1 = more global, homogeneous world; scenario families 2 = more regional, heterogeneous world) and degree of environmental friendliness (scenario families A = more economic, scenario families B = more environmental friendly). For example, the A1 scenario describes a globalized world with the focus on economic growth. The different groups within the A1 family differ concerning their technological emphasis (A1Fl = fossil intensive, A1T = focus on non-fossil energy sources). The often applied A1B scenario acts on the assumption on a balance across all energy sources (EEA 2008). The global warming strongly depends on the amount of greenhouse gas emission and therefore on the underlying emission scenario (Fig. 1).

The predicted climate change will influence all ecosystems. There will be a gradual change of average climatic conditions, for example annual mean temperature or precipitation, but also a change of extreme events. Extreme events (such as extended drought periods or storm events) are supposed to be at least as important as the change of annual mean climatic conditions (Fuhrer et al. 2006).
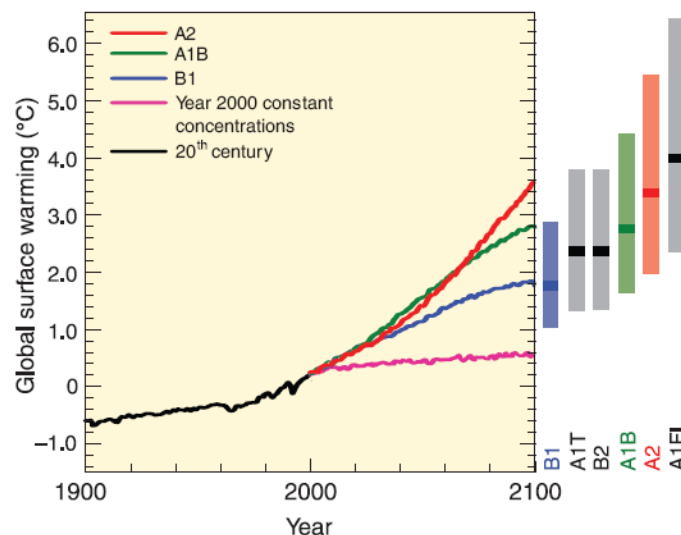
Fig. 1: **Model predictions of surface warming**: solid lines in the graph show the surface warming for the scenarios A2, A1B and B1, the pink line demonstrates the warming where the greenhouse gas concentrations were held constant at year 2000 value. The bars next to the graph show the best estimate for each emission scenario (solid line within each bar) and the range for the emission scenarios at 2090-2099 relative to 1980-1999 (IPCC 2007).

1.1.1 Climate change in Europe

In the last 140 years, temperatures in Europe already increased about 1.2 °C for land alone (EEA 2008) which is above the worldwide average of 1.0 °C for land alone compared to the pre-industrial level (IPCC 2007). The increase was highest for the southwest and the northeast of Europe and mountain areas (EEA 2008). Projections for the end of the century calculated a further temperature increase between 1.0 °C and 5.5 °C for Europe depending on the underlying emission scenario (EEA 2008). The highest increase of temperatures will be in the Mediterranean region (IPCC 2007). This increase is again higher than the projected global warming predicting an increase of 1.8 °C to 4.0 °C (EEA 2008).

Warren et al. (2011) reviewed investigations studying the future impact of higher temperatures on ecosystems and found that the majority are negative, except of enhanced growth of plants at moderate climate change. Some studies already found examples for negative effects concerning an increase of 1.6 °C above pre-industrial level which is the low end of the IPCC emission scenario range. In public and politics, the increase of 2 °C to the end of the century is discussed as a threshold that should not be trespassed. Indeed, Warren et al. (2011) came to the conclusion that negative impacts will rapidly increase behind this threshold and that some ecosystem functions are supposed to start collapsing at

a temperature increase of 2.5 °C. An example for the collapse of an ecosystem is an irreversible damage such as the extinction of key species.

Another important factor influencing ecosystems is the amount of annual precipitation and the distribution over the seasons. For Europe, an increase of precipitation in Northern Europe and a decrease in the Mediterranean region was observed in the time period 1900-2005 (IPCC 2007). Although all projections are based on the increase of greenhouse gases and differ depending on the used emission scenario, the confidence for some projections is higher than for others. Whereas the projections for temperatures are most reliable, it is more difficult to predict the change of precipitation (IPCC 2007). Therefore, often only trends of decrease or increase in precipitation are described. Climate change projections predict a further decrease of precipitation for Southern Europe and a decrease in summer precipitation also for Eastern and Central Europe (e.g., IPCC 2007, Schär et al. 2004). In contrast, the winter precipitation will increase in many regions (Blekinsop and Fowler 2007). In Central Europe, the summer in 2003 was extremely hot and dry with severe drought stress (Löw et al. 2006, Leuzinger et al. 2005), and simulations for Northern Switzerland suggest that about every second summer could be as warm and dry as the summer in 2003 at the end of this century (Schär et al. 2004).

In addition to the change of the climatic averages, the occurrence and frequency of climatic extremes are also important for ecosystems such as extreme drought, precipitation or storm events. Unfortunately, these extreme events are difficult to forecast (IPCC 2007). An increase of these events was already observed compared to pre-industrial levels and the trend is supposed to continue (EEA 2008). Dry and hot extremes will become more frequent in Europe, especially in Southern Europe (Beniston et al. 2007, Kundzewicz et al. 2006, Santos and Corte-Real 2006, Schär et al. 2004). Meehl and Tebaldi (2004) predict that heat waves will not only become more frequent but also more intense and longer lasting in the second half of the 21st century. Projections for the end of the century suggest that Central Europe will experience the same number of hot days that can be found currently in Southern Europe (Beniston et al. 2007).

1.1.2   Climate change in Germany

In the 20th century, the temperature in Germany increased about 0.9 °C, mainly in winter. The increase is about 1.5 °C comparing only the last 30 years to the reference period (Schönwiese et al. 2003). For the end of the century, a further annual mean temperature increase between 2.5 °C (emission scenario B1) and 3.5 °C (emission scenario A2 and A1B) is predicted (Fig. 2; Kreienkamp et al. 2010, Jacob et al. 2008). The increase will be less for the coast and higher for Western and Southern Germany. The temperature increase will also be seasonally different. In winter, the increase will be highest (up to 4.5 °C) and less in spring and autumn (up to 3.0 °C; Kreienkamp et al. 2010, Jacob et al. 2008, Enke et al. 2005).
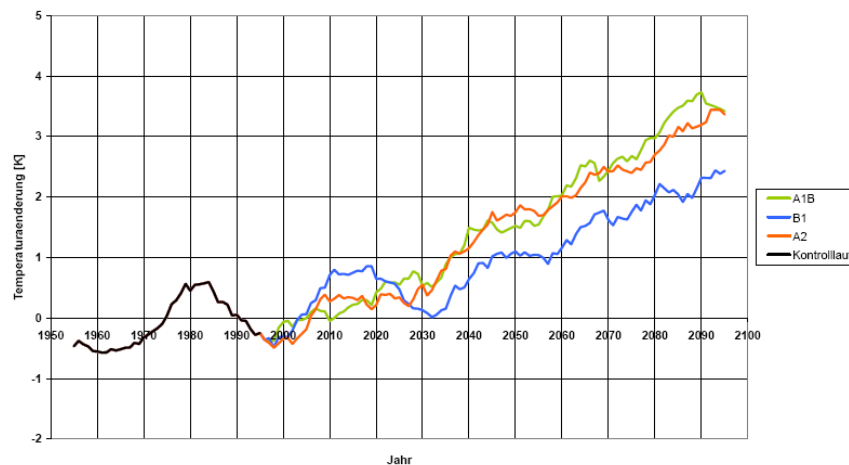
Fig. 2: **Change of the annual mean temperature in Germany**: based on the emission scenarios A1B, B1 and A2 relative to the time period 1961-1990, abscissae: year (1955-2095), ordinate: change of the annual mean temperature [K], black line: control run (Jacob et al. 2008).

In the last 30 years, an increase of precipitation in winter was observed especially in Western and Southern Germany, whereas a decrease in summer precipitation was found, foremost in Eastern Germany (Schönwiese et al. 2003). For the end of the century, this seasonal trend will continue while the annual mean precipitation rate will stay constant (Fig. 3; Kreienkamp et al. 2010, Jacob et al. 2008, Enke et al. 2005). The coast, Southern and Eastern Germany will suffer in particular under summer drought, and higher precipitation in winter is particularly predicted for the low mountain ranges (Kreienkamp et al. 2010).
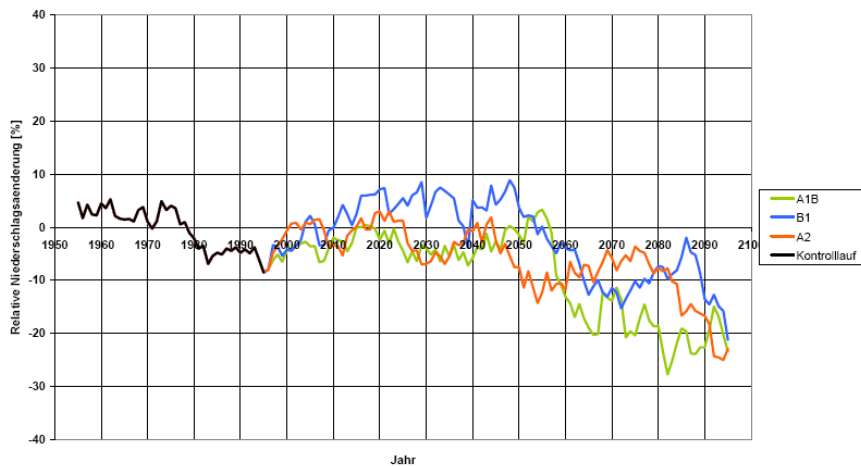
Fig. 3: **Change of the annual mean precipitation in summer in Germany**: based on the emission scenarios A1B, B1 and A2 relative to the time period 1961-1990, abscissae: year (1955-2095), ordinate: change of the annual mean precipitation [%], black line: control run (Jacob et al. 2008).

1.1.3   Adaptation of forest trees to changing environmental conditions

The projected climate change raises the questions which species may adapt to the changing environmental conditions. Many tree species are important for ecosystems playing a foundation or keystone role. Moreover, forests are closely linked with atmospheric carbon budget because they contain a great amount of the total terrestrial biomass (e.g., Aitken et al. 2008). In the past, there have been times when temperatures were higher than the predicted ones for the future. In their evolutionary history, many tree species have survived large-scale environmental changes many times (Hamrick 2004). Therefore, one should assume that forests and tree species are generally able to adapt to the future climate. The major problem is that the climate change will proceed faster this time. Higher temperatures, drought stress and heavy storm events may directly influence the health of forest ecosystems. Furthermore, these abiotic stressors are able to decrease the resistance against secondary damages, for example fungi and insects, which makes it even more difficult to predict the consequences of climate change (Engesser et al. 2008). Species have three different ways to react to the changing climate: migration, adaptation or extirpation (Aitken et al. 2008).

Trees are long-living organisms with long generation times and relatively slow migration rates. Pollen records have suggested post-glacial migration rates of up to 200 m per year, but more recent research names migration rates less than 100 m per year (Aitken et al. 2008). However, different tree species have different mechanisms for seed and pollen dispersal, and therefore, different migration rates. Two mechanisms are important

6

concerning adaptation. The first is the possibility of a species to change its phenotype in response to changing environmental conditions. The second is the ability of genetic adaptation over generations. For the latter one, it is important to conserve a high genetic diversity within species. Compared to herbaceous plants, most of the genetic variation of tree species can be found within populations rather than among them (Hamrick 2004). This is an advantage for trees, because the loss of some populations will not cause a loss of overall genetic diversity in the species. A disadvantage for trees concerning adaptation is the low mortality rate. Modelling the adaptation of birch and pine to climate change, Kuparinen et al. (2010) detected that earlier maturation and long-distance dispersal may improve the adaptability of trees to the recent climate change but this effect is rather small. The increase of mortality would strongly promote the adaptation to changing environmental conditions. However, Hamrick (2004) concludes that the special life-history traits for trees combined with their high genetic diversity allow them to adapt relatively quickly to global climate change.

DeHayes et al. (2000) reviewed the response of forest trees to climate change with the focus on North American forests. He concluded that the responses of tree species are very complex and future predictions for forest composition are difficult and inherently uncertain. Individual species rather than whole communities will disperse, therefore, forest communities are supposed to change in composition. Additionally, better adapted tree species will immigrate and outcompete preexisting species. However, even if tree species are able to adapt to the changing environmental conditions, climate change is not the only factor that will influence forest ecosystems in the future. Another important factor is the anthropogenic influence, for example silvicultural management, fragmentation and the introduction of competitive exotic plant species and new diseases (e.g., Milad et al. 2011). Thus, predicting how forest ecosystems or single tree species will react is very difficult since different factors of global climate change and human interventions need to be considered.

In this study, the tree species European beech (*Fagus sylvatica* L.) was investigated. The species is mainly distributed in temperate oceanic forests in Europe. In general, a positive impact on growth is predicted for the northern and western parts of the temperate oceanic forests in Europe and a negative impact for the southern and eastern parts (Lindner et al. 2010). But this is only a general conclusion without considering single species, such as *F. sylvatica*. Within these regions, different species will react differently to the changing

environmental conditions. The investigated tree species is distributed in parts of the Mediterranean and temperate continental region as well as in the south of the boreal region where the climatic conditions will change differently compared to the centre of the distribution. Therefore, the consequences for beech will be different depending on the considered region.

1.2 Ecology and adaptability of European beech

European beech (*Fagus sylvatica* L.) belongs to the Fagaceae family that comprises seven or eight genera: *Castanea* (chestnuts), *Castanopsis*, *Chrysolepis*, *Fagus* (beeches), *Lithocarpus*, *Quercus* (oaks) and *Trigonobalanus* (Govaerts and Frodin 1998). Govaerts and Frodin (1998) also classify the genus *Nothofagus* to the Fagaceae family. Recent findings classify this genus to an own family (Nothofagaceae; APG III 2009). The genus *Fagus* is divided into the two subgenera *Engleriana* and *Fagus* and includes eleven species (Kremer et al. 2007, Denk 2003). Most of the species are distributed in the northern hemisphere. *Fagus sylvatica* is the only native *Fagus* species in Central Europe. The species is widespread throughout Europe with a distribution range from Southern Sweden and Norway in the north to Sicily (Italy) in the south. The western distribution range is in Spain (Cantabrian Mountains), the eastern in the Carpathians and Balkan Mountains (Ukraine, Romania, Bulgaria; Fig. 4).

Fig. 4: **Distribution range of European beech:** continuous distribution in grey, black dots indicate isolated populations (Bolte et al. 2007).

Around 3000 years ago, the dominance of oak trees was brought to an end in Central Europe and shifted to European beech probably due to anthropogenic influence (Schulze et al. 2010). The potentially natural distribution area of European beech today is more than 910,000 km², but because of recent anthropogenic influence, the preference for conifers such as Norway spruce, European beech is found on a much smaller area. In Germany, as the centre of distribution, about 25% of the European beech tress are located (Knapp et al. 2008). However, the most frequent trees in Germany are Norway spruce (*Picea abies* L., 28.2%) and Scots pine (*Pinus sylvestris* L., 23.2%). European beech is the most frequent deciduous tree (14.8%) followed by oak species (9.6%; Schmitz et al. 2004). The forest inventory in 2004 observed a decrease of spruce and pine compared to the time period 1986 - 1989, whereas the number of European beech trees increased (Schmitz et al. 2004).

Beech trees reach a height of up to 50 m with a trunk diameter of up to 2.6 m (Ellenberg and Leuschner 2010). The lifespan of a beech tree is 150 to 200 years with a maximum of about 300 years (Wilmans 1990). Leaves are 5 to 10 cm long and 3 to 7 cm broad, alternate and simple. European beech starts flowering and seed production at an age of 40 to 50 years (Wagner et al. 2010). The species is highly outcrossing and largely self-incompatible with a selfing rate less than 10% (Merzeau et al. 1994). Although wind-pollinated, pollen dispersal in beech is limited and normally less than 150 m within forests (e.g., Oddou-Muratorio et al. 2011, Oddou-Muratorio et al. 2010, Wang 2001) but long-range pollen transport is possible (e.g., Belmonte et al. 2008). The monoecious plant produces small triangular seeds (beechnuts) in autumn that are primarily dispersed by gravity and secondly by animals (Oddou-Muratorio et al. 2011). Seed dispersal is therefore often limited to about 20 m but up to 80 m is possible (Wagner et al. 2010).

In Central Europe, beech is of high economic and ecological value (e.g., Gömöry et al. 2003). In future, beech will be even more important in Germany because of the transition strategy to transform pure conifer stands into pure beech or mixed deciduous stands comprising beech (e.g., Scharnweber et al. 2011, Fritz 2006, Tarp et al. 2000). One advantage of beech is that the species is considered to be insusceptible to pathogens and insects, apart from the beech bark disease (Ellenberg and Leuschner 2010; beach bark disease in *Fagus sylvatica* e.g., Lunderstädt 2002, Gora et al. 1994). And whereas deer browsing is a problem for most tree species in Central Europe, beech seems to be less attractive for deer. In an experiment, almost 100% of ash, maple and elm seedlings were browsed, but only 15% of the beech seedlings were damaged (Modrý et al. 2004).

Beech has been regarded as "the most successful European plant species" in the distribution area where the tree species is dominant (Leuschner et al. 2006). The dominance of the species is mainly due to shade tolerance (Jarcuska 2009). Seedlings survive long periods at very low light level (Modrý et al. 2004) and adult trees are normally dense that only a low light level reaches the ground. In addition, beech has a high plasticity of morphological traits like biomass allocation, leaf morphology and architectural traits. This allows beech to grow better than oak as a competitor of beech, and even at a high light regime where oak is highly competitive (Kunstler et al. 2005).

European beech colonizes almost all habitats in its area of distribution. Less favourable sites are extremely dry soils, flooding sites and sites with high groundwater level (Ellenberg 1988). A broad range of different soil types are suitable for European beech. Soil pH, for example, may vary between 3.2 and 7.3 (Leuschner et al. 2006). Bolte et al. (2007) summarizes the minimum requirements for beech: at least 500 mm annual mean precipitation or 250 mm in the vegetation period between May and September is necessary, and, additionally, the absence of extreme drought. The mean temperatures in July should be less than 19 °C and the mean January temperature above 3 °C. Altogether, fewer than 141 frost days and more than 217 days with a daily mean above 7 °C are necessary. Furthermore, the absence of extreme heat, extreme frost (less than -35 °C) and severe late frost is required.

1.2.1   Reactions to enhanced CO_2

The global climate change is closely liked to higher CO_2 concentrations in the atmosphere. This may improve the growth of plants if stress factors are absent. An experiment with elevated CO_2 over three years showed that Norway spruce is able to benefit from enhanced CO_2 but European beech not (Kozovits et al. 2005). Körner et al. (2005) found no clear trend in a comparable experiment with enhanced CO_2. The growth of beech in this experiment was improved in some years and in others not. Furthermore, some tree species react differently depending on the soil. There is a positive effect of elevated CO_2 for Norway spruce on acidic and on calcareous soil, but beech trees responded negatively growing on acidic soil while they react positively growing on calcareous soil (Spinnler et al. 2002). Concerning enhanced CO_2, there is no general trend in photosynthetic activity, growth and competitive ability for European beech (Geßler et al. 2007).

1.2.2   Reactions to drought stress

Plants differ in their reaction to drought stress, even closely related species (e.g., *Quercus robur* and Q. *petraea;* Gieger and Thomas 2002). During drought events, predawn water leaf potential, leaf conductance and photosynthesis are reduced in European beech (Fotelli et al. 2001, Tognetti et al. 1995). Even mature leaves became susceptible to oxidative stress during long water stress in summer (Polle et al. 2001). It is suggested that fine root mortality increases in drought stress situations causing a reduction of root biomass. However, compensatory fine root production was observed to be stimulated in particular in dry stands (Meier and Leuschner 2008). Beech saplings respond to drought stress mainly by shortening of the root lifespan, but root system structure and root/shoot carbon partitioning pattern stay constant (Meier and Leuschner 2008). In general, beech seems to be able to vary fine root morphology and soil space sequestration to increase soil resource uptake during drought periods (Bolte and Villanueva 2006).

Drought stress causes a reduction of growth in beech, but the effect is more pronounced in the following year after the drought (Granier et al. 2007, Czajkowski et al. 2005). Furthermore, the observed negative effects on crown condition (crown transparency) may be directly caused by drought stress but also by heat-associated higher ozone concentrations or even biotic agents (Seidling 2007).

1.2.3   Differences between the centre and the edge of the distribution

The reactions of beech to drought stress and predicted consequences concerning climate change are mainly investigated using beech populations at the distribution edge in the south or the east of Europe. The climate there is already warmer and drier in the summer months compared to the centre of the distribution. Modelling studies predict longer vegetation periods for the Mediterranean region enhancing biomass production if water is not limited (Sabaté et al. 2002). However, in general, less precipitation is forecasted for Southern Europe, especially in summer (IPCC 2007, Schär et al. 2004). The observed growth reduction between 1975 and 2003 in Northeastern Spain (Jump et al. 2006b) and the affected health of adults and seedlings on the Iberian Peninsula have been discussed in this context (Aranda et al. 2000). For Eastern Europe, a bioclimatic distribution modelling was used to predict the future distribution of beech in Hungary. Czúcz et al. (2011) concluded that 56-99% of the beech forests might be out of their present bioclimatic niche already by the year 2050. Some authors conclude that beech will probably give way to more drought

tolerant species at the southern and eastern distribution borders (Gärtner et al. 2008, Kölling et al. 2005), although in higher altitudes, an extension of beech up to 70 m was found in the north of Spain (Peñuelas and Boada 2003).

Investigations in the centre of the distribution of beech and the impact of the changing climatic conditions are less prevalent. Peuke et al. (2002) studied seedlings from different German populations. Provenances from the wet habitats all responded like drought-sensitive ecotypes. However, this response was not observed in all provenances from relatively dry habitats (less than 680 mm annual mean precipitation). Beech trees in Northeastern France investigated in the extreme year 2003 showed strong drought stress symptoms with lower radial growth and fall of green leaves already in August, but the crown decline and the death of trees were found to be very limited (Betsch et al. 2011). Charru et al. (2010) found also a general growth decline for beech trees in the time period 1979 to 2007 for Northeastern France. Beside drought stress, another problem increasing with climate change is the biotic stressor *Phytophthora*. In the south of Germany, a stand decline together with typical *Phytophthora* symptoms was observed after the hot and dry year 2003 (Jung 2009). The trees were probably infected during the wet spring and autumn in the year 2002 and were than unable to cope with the biotic stress in combination with the extreme drought in summer in the following year.

1.2.4 The susceptibility of different provenances to drought stress

Several investigations showed that different beech provenances differ in their susceptibility to drought stress (Czajkowski and Bolte 2005, Peuke et al. 2002, Schraml and Rennenberg 2002). In general, provenances from dry regions are better adapted to drought stress (Schraml and Rennenberg 2002, Tognetti et al. 1995). For example, beech at the eastern border of the distribution is less susceptible concerning drought stress in comparison to beech from the centre of the distribution (Czajkowski and Bolte 2005). The ecophysiological reactions in beech populations in Central Europe and the Mediterranean regions were very different in the year 2003, although the stress conditions were comparable. In contrast to the Central European beech populations, the trees in the Mediterranean region are adapted to these conditions (Fotelli et al. 2009). Differences in morphology point out that different provenances are adapted to different ecological conditions. A lower leaf area/fine root ratio was found in drought tolerant populations in Spain allowing seedlings to respond later to drought stress compared to seedlings from other provenances (García-Plazaola and Becerril 2000). The drought stress tolerance of

these seedlings was mainly based on morphological characteristics and less on the photoprotective system. The increase of proline that is assumed to be linked with drought tolerance is also ecotype-specific (Schraml and Rennenberg 2000). There is also a great variability within populations. Drought resistant seedlings can be found in all populations, but provenances from dry regions produce more tolerant seedlings (Czajkowski and Bolte 2005).

1.2.5 The influence of drought stress on competitive interactions

The growth and the survival of beech seedlings depend strongly on the ability to coexist with other highly competitive plant species (Tognetti et al. 1998). One of these competitors is *Rubus fruticosus* (blackberry). Under normal environmental conditions, beech and blackberry coexist without negative effects for beech, but drought stress strongly influences the competitive interactions between these species. Only beech reacts with a biomass reduction and a change in the root/shoot ratio while blackberry shows no water stress reactions. Fotelli et al. (2001) conclude that the growth of beech seedlings may be inhibited by blackberry, especially during summer drought events.

The main competitor tree species to European beech are oaks. In Central Europe, especially *Quercus petraea* (sessile oak) replaces beech in dry environments (Geßler et al. 2007). Because beech is highly competitive were water is not limited, it outcompetes sessile oak in most regions in Western and Central Europe at the moment. Both mechanisms, aboveground and belowground predominance of beech over sessile oak were found (Leuschner et al. 2001b). But sessile oak is often superior in dry environments. It reacts less sensitive in physiological parameters like leaf conductance, photosynthetic activity, stem hydraulic conductivity and fine root-vitality (Leuschner et al. 2001a) which indicates a better adaptation to drought stress. Leuschner et al. (2001a) also observed that all tree organs (leaves, stems, roots) of beech are sensitive to drought in comparison to sessile oak. *Quercus petraea* maintained a better water status during drought periods not only compared to beech but also to *Carpinus betulus*, *Acer campestre* and *Tilia platophyllos* (Leuzinger et al. 2005). Compared to *Q. petraea*, *Q. robur* (pedunculate oak) is more sensitive to drought (Friedrichs et al. 2009) but still more tolerant than beech. Scharnweber et al. (2011) compared the growth of beech and pedunculate oak along a precipitation gradient in Northeastern Germany. The growth of both species strongly depends on the availability of water during early summer (June and July). The study revealed that already small changes

in the amount of precipitation can affect tree growth of both species, but the effect is stronger for beech.

1.2.6 Extension of the vegetation period – advantages and risks

The predicted warming in Europe implicates a longer vegetation period which will, in general, positively influence the growth of plants. Remote sensing data were used to observe an increase in plant productivity in the northern high latitudes. It is assumed that the reason is the extension of the growing season enhancing the net carbon uptake period (Churkina et al. 2005, Zhou et al. 2001, Myneni et al. 1997). In the last 30 years, the mean annual growing season has already lengthened by about eleven days (Menzel 2000, Menzel and Fabian 1999). Many trees species will react with earlier bud burst in spring and/or delay of leaf senescence in autumn (Menzel and Fabian 1999). Earlier bud burst in spring may increase the risk of late frost damage. For beech, the threshold value for frost damage for young leaves is about -3 °C at the beginning of the vegetation period (Dittmar et al. 2006). Beech provenances differ in bud burst date and also in their risk to late frost (Višnjić and Dohrenbusch 2004, Wühlisch et al. 1995). In general, provenances from the east and the south of Europe flush earlier than provenances from the western parts of Europe with very early flushing provenances at the eastern edge of the distribution (Wühlisch et al. 1995). The frost resistance mainly depends on the annual mean temperature of the populations. Furthermore, there is a correlation between flushing date and frost resistance. Višnjić and Dohrenbusch (2004) found a high difference between the flushing dates of populations from Italy and Northern Germany. The provenances from Germany were flushing about seven days later and they were least susceptible to frost (together with provenances from Southeast Europe). In contrast, the early flushing provenances from Italy were most susceptible.

Although the lengthening of the growing season was in general to be found mostly due to earlier bud burst and less to later leaf senescence (Linderholm 2006, Menzel 2000), it is questionable if there will be an earlier flushing in beech. Observations in the time period 1974 to 1996 (in comparison to 1951 – 1973) showed that the growing season of beech was lengthened but the leaf unfolding date stayed unchanged (Menzel and Fabian 2001). Estrella and Menzel (2006) confirmed that warm temperatures in August and September delay leaf senescence in beech and in oak. The growing season of different species was also observed for three years in the Pyrenees Mountains (Southwestern France) and than used for a modelling study. Vitasse et al. (2011) argued that leaf unfolding will be earlier for

14

Faxinus and *Quercus* but not for *Fagus*. In contrast, leaf senescence will be delayed for *Quercus* and especially for *Fagus* at low elevations. The authors concluded that the extension of the growing period will be largest for oak which can possibly modify the competitive interactions between oak and beech. The increase of drought intensity and frequency was not included in this modelling approach (Vitasse et al. 2011), but drought may have a considerable effect on senescence time (Bréda et al. 2006, Vitasse et al. 2010). A longer growing season due to higher temperatures was observed for oak and beech at mid- to high-elevation populations. The low-elevation populations, especially beech, did not delay senescence as a result of drought in contrast to the expectations from the modelling study not including drought stress (Vitasse et al. 2010).

1.2.7   Genetic diversity and differentiation

A high genetic diversity of a species promotes the adaptation to environmental changing conditions. Many studies revealed a positive relationship between genetic diversity and different fitness measures. For example, beech trees in Poland suffering from the fungus *Phytophthora* were less heterozygous compared to healthy beech trees in the same region (Nowakowska and Oszako 2008).

For wind-pollinated species like beech and oak, a high genetic diversity is expected. Indeed, different studies using microsatellite markers revealed in general a high genetic diversity for beech populations (Oddou-Muratorio et al. 2011, K.C. 2011, Gautam 2010, Buiteveld et al. 2007, Vornam et al. 2004). A comparable high genetic diversity was also found for other *Fagus* species like *F. crenata* and *F. japonica* (Asuka et al. 2004, Tanaka et al. 1999) and for *Quercus* species (e.g., Dostálek et al. 2011, Muir et al. 2004, Bruschi et al. 2003). Silvicultural management can have a negative impact on the genetic diversity of species (for example reviewed by Schaberg et al. 2008). Although *F. sylvatica* is intensively managed in Central Europe, there seems to be almost no negative impact on neutral genetic diversity (K.C. 2011, Buiteveld et al. 2007, Hussendörfer and Konnert 2000).

Spatial genetic structures can be found in beech populations, mainly in unmanaged ones, but they are rather weak. Manly due to the limited seed dispersal, spatial genetic structures have been found up to 40 m (K.C. 2011, Jump and Peñuelas 2007, Vornam et al. 2004).

Most of the genetic variation can be found within beech populations. Therefore, the differentiation among stands is generally low but often significant (Buiteveld et al. 2007).

K.C. (2011) was able to clearly separate beech populations from three regions in the north, the middle and the south of Germany.

1.3 Genetic markers to investigate European beech

1.3.1 Genetic markers commonly used for investigations

The three following markers were often applied in the past or are still in use. The first two are based on the analysis of fragment length polymorphisms.

One of the commonly used markers are AFLP markers (Amplified Fragment Length Polymorphism) that can be used to analyse DNA of any origin without sequence information (e.g., Meudt and Clarke 2007, Bussell et al. 2005, Blears et al. 1998, Vos et al. 1995). Restriction enzymes and the PCR technology are used to produce fragments of different sizes from total genomic DNA. The position of the fragments in the genome is unknown. Therefore, this marker is called anonymous. AFLP markers are dominant and only the presence or absence of a fragment can be analysed. It is not possible to distinguish between homozygous and heterozygous individuals which is the major drawback of this method because it complicates the calculation of some population genetic parameters (Bonin et al. 2007, Nybom 2004, Campbell et al. 2003). AFLP markers were for example used to study differences between *Fagus sylvatica* and *Fagus orientalis* (Gailing and von Wühlisch 2004) and to investigate the spatial genetic structure of European beech. Jump and Peñuelas (2007) revealed an extensive spatial genetic structure using AFLP markers that was not found with microsatellite markers. A combination of different marker types including AFLP and microsatellites was used to construct a genetic linkage map for European beech (Scalfi et al 2004). Jump et al. (2006a) used AFLP markers to study the adaptation of European beech to temperature changes. Populations along an altitudinal gradient were analysed and an outlier analysis was conducted that revealed a single outlier locus with unusually high F_{ST}.

Furthermore, microsatellite markers (SSR markers, Simple Sequence Repeat) are often used for different kinds of genetic analysis in plants since 20 years (e.g., Kalia et al. 2011, Ellegren 2004, Weber 1990). Microsatellites are short and simple sequence repeats of one to six bp, but more complex repeat motives are also used. In comparison to AFLP markers, SSRs are codominant markers. Therefore, it is possible to distinguish between homozygous and heterozygous individuals, and to calculate allele frequencies which are the basis for the computation of various population genetic parameters. The position of the microsatellite

markers in the genome is in most cases unknown, and these markers are mostly considered to be selectively neutral. However, microsatellites located in coding regions, EST-SSRs (Expressed Sequence Tag), are also available for many tree species and they have some advantages compared to "normal" SSRs, for example their higher transferability between species (e.g., reviewed by Bouck and Vision 2007). The high level of polymorphisms is a great advantage of microsatellite analysis. Although these two marker types are well established, they have some technical drawbacks (for more detailed information, see the manuscript in appendix 1). Furthermore, their ability to investigate adaptation is very limited as confirmed by Kraj and Sztorc (2009) who investigated European beech populations with early, intermediate and late phenological forms (bud burst) using SSR markers. Microsatellite markers were mainly used to investigate neutral genetic diversity, differentiation and spatial genetic structure of European beech (e.g., Dounavi et al. 2010, Nyári 2010 and literature cited in chapter 1.2.7), but there is recently also an attempt to use them for seed source identification (Hasenkamp et al. 2011).

Isozymes are biochemical markers that analyse structurally different molecular forms of an enzyme with the same catalytic function. Amino acid changes cause changes in the net charge or the confirmation of the enzyme. Therefore, different loci, usually two, can be distinguished (e.g., Kumar et al. 2009). These codominant markers are used to analyse the genetic diversity at gene loci coding for enzymes which serve important functions in the metabolism of plants. Therefore, these markers can be used to study adaptation, but only a very limited number of soluble enzymes can be analysed. Additionally, these markers have a low level of polymorphisms and some of them may be environmentally influenced (Kumar et al. 2009). Furthermore, Müller-Starck and Starke (1993) found that some of the investigated isozyme markers showed substantial tissue-specific expression. In the past, isozymes have been used to study the genetic variation and genetic structure of European beech in different regions (e.g., Konnert and Henkel 1997, Hazler et al. 1997, Larsen et al. 1996). Other studies were conducted to find indications for adaptation, e.g., adaptation to air pollution (Longauer et al. 2001). Müller-Starck (1989) investigated the genetic background of neighbouring pairs of entirely unaffected trees and trees which indicate the onset of dieback. Statistically significant genetic differences between these two types were found and the tolerant type contained a greater amount of genetic variation. Sander et al. (2000) and Hazler-Pilepic et al. (1999) found significant differences investigating isozymes along an altitudinal gradient.

1.3.2 Single Nucleotide Polymorphism (SNP) markers to investigate candidate genes

In comparison to the three previously described marker types, SNP markers (Single Nucleotide Polymorphism) directly analyse the variation within DNA sequences and not only polymorphisms in fragment lengths. A SNP is the substitution of only one nucleobase. Single base indels (insertions/deletions) are formally no SNPs. However, the term SNP is often imprecisely used (Brookes 1999). SNPs are abundant and widespread in the genome and their evolution is based on a simple mutation model (infinite allele model; e.g., Vignal et al. 2002). For humans and model organisms (e.g., *Arabidopsis thaliana)* SNP markers are already well established and frequently used. SNPs are, in comparison to SSR and AFLP markers, valuable markers to study adaptation of plants (e.g., Gailing et al. 2009). An example for the successful identification of only one SNP influencing a phenotypic trait is the loss of seed shattering in rice (Konishi et al. 2006). Ingvarsson et al. (2008) found two SNPs in the *phytochrome B2* gene that were associated with variation in the timing of bud set, but they explained only between 1.5 and 5% of the observed phenotypic variation in European aspen (*Populus tremula*). When SNPs are used to study adaptation by the comparison of populations for example with different environmental conditions, it is important to investigate these populations also with neutral genetic markers, e.g., microsatellites. These neutral markers may be used to identify genetic differentiation not related with selection but caused by demographic or other processes, e.g., human seed transfer (e.g., Eveno et al. 2007). Populations showing strong genetic differentiation already with neutral markers are of limited use because other factors than selection cannot be ruled out as a reason for the significant differentiation of SNP markers in such populations.

The location of a SNP is normally known and based on the objective of the investigation SNPs in non-coding or coding regions of genes can be selected. Furthermore, it is possible to select only non-synonymous SNPs that lead to an amino acid exchange. Non-synonymous SNPs are most interesting for the study of adaptation, but SNPs in non-coding regions can also be of relevance. Whereas non-synonymous SNPs potentially change the protein structure, SNPs in non-coding regions may influence gene splicing and enable a gene to increase its coding capacity leading to several structurally distinct isoforms (Baek et al. 2008).

One way to investigate adaptation is to study candidate genes that are potentially involved in the trait of interest (e.g., Pflieger et al. 2001). At present, the candidate gene approach is a powerful and efficient method to study the genetic architecture of complex traits (Zhu and Zhao 2007). Normally, genes are investigated with known biological functions directly or indirectly influencing the phenotypic trait of interest. Different methods can be used to identify candidate genes of interest (for example reviewed by Kirk and Freeland 2011). A classical approach is the quantitative trait loci (QTL) analysis that uses a large number of individuals with a considerable phenotypic variation at the trait(s) of interest. Normally, the F_2 generation from a known pedigree or a backcrossed family from a known cross is used (Kirk and Freeland 2011). Unfortunately, the resolution of QTL analysis is low. In most cases, it is only possible to identify a region of interest that contains a lot of genes. QTL mapping has been conducted in a number of different tree species and for different complex traits, but it has not been able to reveal the underlying genes for the studied traits as it has been in model systems or a few crop species (Neale and Kremer 2011). However, QTL analysis confirmed that complex traits in trees are controlled by many genes. Another method to study adaptation is the genome-wide association study that is increasingly used for plants (Nordborg and Weigel 2008). The first step in this process is the discovery of a large number of genetic markers, normally SNPs (Myles et al. 2009). Then, hundreds or thousands of unrelated individuals are genotyped for hundreds of genes. This method can be applied for model organisms and is widely used to study human disease. For grapevine (around 475Mb genome), it is estimated that over two million SNPs are necessary to provide a reasonable coverage to apply this technique (Myles et al. 2009). Without sufficient sequence information, this method cannot be applied. Another method used in model species is the screening of mutants to identify gene functions (for example reviewed by Page and Grossniklaus 2002).

For non-model organisms one way of analysing candidate genes is not to identify new candidates, but to validate candidate genes already known in model plants like *Arabidopsis thaliana* or *Populus trichocarpa* for trees. Furthermore, new candidate genes may be identified using expressed sequence tags (ESTs; e.g., Seo and Kim 2009, Alba et al. 2004). Real-time reverse transcription PCR (RT-qPCR) or microarrays may be applied to study the expression profile of potential candidate genes comparing stressed plants with control plants or different developmental stages (RT-qPCR reviewed by e.g., Nolan et al. 2006, Ginzinger 2002, Bustin 2000; microarray technology reviewed by e.g., Alba et al. 2004, Reymond 2001, Wu et al. 2001). Genes that are up-regulated or down-regulated in

comparison to the control plants are supposed to be involved in the reaction to the investigated trait. A microarray approach was for example used to study genes involved in the ripening process of apple fruit (*Malus domestica*; Seo and Kim 2009). Palle et al. (2011) applied RT-qPCR to investigate 111 candidate genes potentially involved in xylem/wood development in the commercially important tree species loblolly pine (*Pinus taeda*).

The relatively new next-generation sequencing technology allows identifying numerous new candidate genes even in non-model organisms. These techniques enable to analyse the whole transcriptome of a plant without previous sequence information (e.g., Martínez-Gómez et al. 2011, Wang et al. 2009). Because this method is still expensive, already identified candidate genes for the traits drought stress and bud phenology were selected from the literature for this investigation. Until now, only a few studies investigate candidate genes in European beech involved in these traits (Lalagüe et al. 2010). The genome of European beech is not sequenced yet. Therefore, an increasing but still limited number of beech sequences are available in databases (e.g., Schlink 2011, Olbrich et al. 2010, Jimenez et al. 2008, Olbrich et al. 2005). For this investigation, both published *Fagus sylvatica* sequences and orthologous sequences identified in the related genera oak were used (Gailing et al. 2009). Because oak and beech belong to the same family, it is possible to transfer some genetic markers from one species to the other.

1.4   Aims and objectives

The present work aims to investigate the genetic basis of adaptation of beech to global climate change using a candidate gene approach. On the one hand, natural beech populations along a precipitation gradient in Northern Germany were selected for the analysis. On the other hand, a controlled drought stress experiment was conducted in a climate chamber with seedlings from two different regions.

The main objectives of the present study are:

- to select and sequence candidate genes potentially involved in the adaptive response to drought stress and in bud phenology,
- to assess genetic diversity patterns at these genes,
- to provide evidence that selection shaped genetic structures at these genes in adult European beech populations.
- The following hypotheses were tested in order to reach these specific objectives:
- Genetic variation in European beech populations is high at putatively neutral genetic markers and at genes potentially involved in adaptive responses to climate change.
- Genetic differentiation in European beech populations growing along a gradient of decreasing precipitation is higher at genes putatively involved in adaptation to drought stress (candidate genes) than at putatively neutral genetic markers.
- Some of the SNPs identified in the candidate genes are significantly different between the populations along the precipitation gradient.
- Some of the SNPs identified in the candidate genes are significantly different between damaged and not damaged/slightly damaged European beech seedlings of a controlled drought stress experiment confirming the results obtained analysing the populations along the precipitation gradient.

2 Material and methods

2.1 Study areas and sampling design

The sampling areas are located in the north of Germany in or near the Lüneburger Heide (Lower Saxony and Saxony-Anhalt) along a precipitation gradient. They are jointly investigated by several research groups within the collaborative project 'Climate Impact and Adaptation Research in Lower Saxony' (KLIFF; www.kliff-niedersachsen.de).

The annual mean precipitation ranges from 543.7 mm to 765.8 mm and the precipitation in the vegetation period (May-September) ranges from 293.8 mm to 374.4 mm (1971-2000; Table 1). The annual amount of precipitation at the driest area near Calvörde is close to the minimum value necessary for beech (Bolte et al. 2007). The annual mean temperature and the elevation are comparable between the stands (Table 1). More detailed information can be found in climate diagrams (Appendix 2).

At each location, two neighbouring stands were selected with less than 4 km distance between them, one on a sandy soil and one on a rather loamy soil. The available water storage capacity of the soil (AWSC) is almost the same for the different sandy sampling areas and Göhrde loam (78-81 mm/120 cm). It is considerably higher for Unterlüß loam and Calvörde loam (Table 1). The AWSC was calculated separately for each soil horizon from water retention curves established for the matrix potential range of -300 hPa to -1.5 MPa and the data was summed up to a profile depth of 120 cm (Müller-Haubold pers. comm.).

The aim of this study area design was to vary mostly the parameter annual mean precipitation but to keep other environmental factors as constant as possible to study the influence of drought stress on beech forests. Therefore, the largest distance between the different areas is about 90 km linear distance (Calvörde sand – Göhrde sand; Table 2). One exception is the reference area with a high amount of precipitation and a lower annual mean temperature that is located in the Harz Mountains near the village Bad Grund (Tables 1 and 2).

Table 1: Characteristics of the study areas in Germany in or near the Lüneburger Heide (Lower-Saxony and Saxony-Anhalt) and the reference area in the Harz Mountains (Bad Grund; Lower-Saxony).

Closest village	Position	Annual amount of precipitation (1971-2000)*	Precipitation in the vegetation period (May-September; 1971-2000)*	Annual mean temperature (1971-2000)*	Soil type	Available water storage capacity (AWSC; mm/120 cm)	Elevation	Age in years	Average height (dbh >23 cm)°, (beech trees)	Average dbh°, (beech trees)	Percentage of beech trees°	Stem density N/ha° (all trees)	Stand basal area m²/ha° (all trees)
Calvörde	N52 22.819 E11 17.406	543.7 mm	293.8 mm	9.19 °C	sand	81	65m	95	23.8 m	23.4 cm	100	711.1	33.2
	N52 24.238 E11 15.661	543.2 mm	294.0 mm	9.12 °C	loam	140	75m	129	28.3 m	36.6 cm	96.6	300.0	33.3
Göhrde	N53 08.660 E10 52.003	664.5 mm	346.6 mm	8.74 °C	sand	80	90m	131	24.6 m	30.7 cm	94.3	288.9	24.4
	N53 07.379 E10 49.224	674.8 mm	349.2 mm	8.65 °C	loam	78	85m	131	30.2 m	51.0 cm	100	122.2	26.6
Unterlüß	N52 49.831 E10 18.985	765.8 mm	374.4 mm	8.47 °C	sand	79	130m	90	25.3 m	18.6 cm	81.3	611.1	24.3
	N52 49.894 E10 19.183	765.8 mm	374.4 mm	8.47 °C	loam	95	130m	90	28.4 m	26.1 cm	100	411.1	28.5
Bad Grund	N51 49.180 E10 15.213	1169.9 mm	559.8 mm	7.21 °C	n.a.	n.a.	458m	136	n.a.	n.a.	n.a.	n.a.	n.a.

n.a.: data not available; *provided by National Climate Monitoring of Deutscher Wetterdienst (DWD); °Measurements are only available for the jointly investigated areas (30x30m; Müller-Haubold pers. comm.)

Table 2: Linear distances between the populations.

Population pair	Linear distance [km]	Population pair	Linear distance [km]
Calvörde sand - Calvörde loam	3.4	Calvörde - Bad Grund	95*
Göhrde sand - Göhrde loam	3.9	Göhrde - Bad Grund	150*
Unterlüß sand - Unterlüß loam	0.3	Unterlüß - Bad Grund	113*
Calvörde - Göhrde	85*		
Calvörde - Unterlüß	80*		
Göhrde - Unterlüß	47*		

*using the populations (sand or loam) that are closest

All investigated stands are between 90 and 136 years old. The elevation is low, between 65 and 130 m. One exception is the reference area near Bad Grund that is located in the Mountains (458 m). Most of the study areas are pure beach stands or mixed with a few oak trees. More details about the characteristics of the study areas can be found in Table 1. The origin of the stands is unknown. Natural regeneration is used at the moment for all stands and was probably also used to establish the investigated stands, but planting activity cannot be excluded. Detailed information for the reference area is not available yet because this population is not jointly investigated by the KLIFF research groups.

For this investigation, at least 100 adult trees per stand were sampled in early summer 2009. Only the stand in the Harz Mountains was sampled in summer 2011. Altogether, leaves from 707 different trees were used for this study (Table 3). In autumn 2009, 100 beechnuts were sampled under every tree to raise seedlings mainly for the establishment of a translocation experiment that will be investigated in the next phase of the project, but also for a climate chamber experiment

2.2 Climate chamber experiment

Beechnuts were sampled between October and November 2009. Without cleaning, the beechnuts were dried at room temperature to a moisture content of 10% (+/- 2%). The seeds were stored in normal plastic bags at -10 °C until the end of January 2010. Starting from February, the seeds were stored in darkness at +5 °C for stratification and were watered from time to time not to run dry for a total time period of seven weeks. The firstly germinated beechnuts were transferred into plastic bags and stored at 0 °C to delay their growth. After stratification, all seeds, germinated or not, were planted into plastic containers (100 cm³, HerkuPlast-Kubern GmbH, type QP D 84 T/11,5) using normal

potting soil (Fruhstorfer Erde, type P25). Plants were regularly watered and fertilised (Wuxal®Super), and were grown in the greenhouse under natural conditions. In July 2010, the height of the plants was measured and one leaf per plant was sampled. Leaves were stored at -20 °C. 400 healthy plants from the sandy soil stand Göhrde and 400 plants from the reference area in the Harz Mountains were selected for the experiment and replanted into slightly larger plastic pots. In August 2010, all plants were transferred to the climate chamber. In this experiment, a warm summer day was imitated under the following conditions: 14 hours light (100 µE) at 20 °C and 45% air humidity, 10 hours darkness at 12 °C and 55% air humidity. All plants were grouped according to their height (between three and 29 cm): small (up to 13 cm), middle (up to 20 cm) and tall (up to 29 cm). After one week for acclimatization, the drought stress experiment started. No watering at all was done for two weeks, and then, all plants were watered again. Control plants were normally watered twice per week in the greenhouse. Within the groups, the containers were interchanged regularly and randomly.

After two weeks of drought treatment, the status of the leaves was documented as strongly damaged, slightly damaged or not damaged in comparison to the control plants (Fig. 5). The status of the seedlings was defined as follows: not damaged seedlings = up to one leaf is slightly damaged; slightly damaged seedling = at least two leaves are slightly damaged, up to one leaf is strongly damaged; strongly damaged seedling = at least two leaves are strongly damaged. For the genetic analysis, 100 slightly or not damaged plants and 100 strongly damaged plants were randomly selected equally distributed over the two provenances from the group middle (height; Table 3). The plants were transferred back to the greenhouse for hibernation.

Fig. 5: from left to right: not damaged leaf, slightly damaged leaf, strongly damaged leaf

2.3 Summary of the investigated material

In total, 707 trees from seven different sampling areas were available for the investigations. All trees were analysed with microsatellite markers. For the SNP analysis, only the adult

trees from the sandy soil stands (Table 3) were used because the soil conditions are more comparable at these sites. Soil scientists, one of the KLIFF research groups, found small scale soil differences for the loamy sampling areas outside the jointly investigated 30x30 m area. In contrast, the sandy soil sampling areas were more homogeneous. Furthermore, the available water storage capacity of the soil is almost the same for the three sandy soil populations (Table 1). Thus, the different drought stress risks of the populations may be directly explained by the different amounts of precipitation.

For the investigation of the phenotypic trait "bud phenology", it was intended to rate the bud burst in spring for all adult trees. It was planned to group the trees into early, late and intermediate flushing groups and to investigate SNPs in the early and the late flushing trees. Unfortunately, it was not possible to conduct the rating of the flushing behaviour due to financial and time reasons. Therefore, this study concentrated on the phenotypic trait "drought stress tolerance".

Table 3: A summary of the investigated stands and the sampling size used for the different experiments.

Experiment/ marker type	Calvörde sand	Göhrde sand	Unterlüß sand	Calvörde loam	Göhrde loam	Unterlüß loam	Bad Grund	Total sampling size
Annual mean precipitation	543.7 mm	664.5 mm	765.8 mm	543.2 mm	674.8 mm	765.8 mm	1169.9 mm	
adult trees								
Microsatellite analysis	104	103	100	101	100	100	100	707
Comparative sequencing (candidate genes)	3	3	3	3	3	3	--	18
SNP analysis	104	103	100	--	--	--	--	307
seedlings								
SNP analysis (climate chamber experiment)	--	100	--	--	--	--	100	200

2.4 Methods

2.4.1 DNA isolation

Total DNA was extracted from leaves using the DNeasy™ 96 Plant Kit (Qiagen, Hilden, Germany). The amount and the quality of the DNA were analysed by 0.8% agarose gel electrophoresis with 1x TAE as running buffer (Sambrook et al. 1989). DNA was stained with ethidium bromide, visualized by UV illumination and compared to a Lambda DNA size marker (Roche).

2.4.2 Microsatellites

Nine highly polymorphic microsatellite markers were used to screen all sampled trees. Four of them were originally developed for *Fagus crenata* (sfc markers, Asuka et al. 2004) and two of them were directly developed for *F. sylvatica* (FS 3-04, Pastorelli et al. 2003; mfs 11, Vornam et al. 2004). Additionally, three EST microsatellite markers were applied originally developed and transferred from *Quercus robur* (GOT006, FIR065, FIR004; Durand et al. 2010). Multiplexing of two to four primers, labelled with different fluorescent dyes, was performed to save time and costs for PCR amplification and subsequent capillary electrophoresis (set 1: all sfc loci, set 2: FS 3-04 and mfs 11, set 3: GOT006, FIR065, FIR004).

PCR amplifications were conducted in a 15 µl volume containing 2 µl of genomic DNA (about 10 ng), 10x reaction buffer (0.8 M Tris-HCl pH 9.0, 0.2 M $(NH_4)_2SO_4$, 0.2% w/v Tween-20; Solis BioDyne, Estonia), 2.5 mM $MgCl_2$, 0.2 mM of each dNTP, 1 unit of *Taq* DNA polymerase (HOT FIREPol® DNA Polymerase, Solis BioDyne, Estonia), 0.3 µM of each forward and reverse primer. The PCR protocol consisted of an initial denaturation step of 95 °C for 15 min followed by 30 cycles of 94 °C for 1 min (denaturation), 47 °C (for the EST primer set 3) or 55 °C (for primer set 1 and 2) for 30 sec (annealing), 72 °C for 1 min (denaturation) and a final extension step of 72 °C for 20 min. Microsatellite fragments were separated on an ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems). Data were collected and aligned with the help of the internal size standard GS 500 ROX^Tm using GeneScan 3.7® (Applied Biosystems), and fragments were scored with the software Genotyper 3.7® (Applied Biosystems).

2.4.3 Selection of candidate genes

All candidate genes have been chosen based on literature surveys suggesting an impact of the genes on either drought stress or bud phenology (Table 4). The Evoltree EST database (www.evoltree.org) and the EMBL Nucleotide Sequence Database (http://www.ebi.ac.uk/embl/) were mainly used to find corresponding *F. sylvatica* sequences. Alternatively, sequences of *Quercus petraea* were transferred to *F. sylvatica* (Vidalis 2011, Vornam et al. 2007). The selected sequences were verified by a TBLASTX search (Washington University Basic Local Alignment Search Tool Version 2.0) and used for primer design in order to amplify the corresponding genomic regions in beech. Additional information on the selection and the analysis of the candidate genes may be found in the manuscript "DNA Sequence Variation and Development of SNP Markers in Beech (*Fagus sylvatica* L.)" (Appendix 1).

2.4.4 Amplification, cloning and sequencing of the candidate genes

In total, 18 trees from six different populations were selected for the comparative sequencing (Table 3). Primers for amplification and direct sequencing of the amplification product were designed by using the program Primer3 (v.0.4.0; Rozen and Skaletsky 2000; http://frodo.wi.mit.edu/primer3/; Table 5). Primers were checked for self-annealing, dimer and hairpin formations using the program Oligo calc: Oligonucleotide Properties Calculator (http://www.basic.northwestern.edu/biotools/oligocalc.html). PCR amplifications were conducted in a 15 µl volume containing 2 µl of genomic DNA (about 10 ng), 7.5 µl HotStar*Taq* Master Mix Kit (Qiagen, Hilden, Germany) and 0.3 µM of each forward and reverse primer. The PCR protocol consisted of an initial denaturation step of 95 °C for 15 min followed by 35 cycles of 94 °C for 60 sec (denaturation), different temperatures according to the primers for 45 sec (annealing; Table 5), 72 °C for 90 sec (extension) and a final extension step of 72 °C for 20 min.

PCR products were analysed by 1% agarose gel electrophoresis with 1x TAE as running buffer (Sambrook et al. 1989). DNA was stained with ethidium bromide and visualized by UV illumination. PCR products were excised from gel and purified using the Geneclean®kit (MP Biomedicals, Illkirch, France). The purified products were cloned into a pCR2.1 vector using the TOPO TA Cloning® kit (Invitrogen, Carlsbad, CA) with slight modifications. The inserts were amplified by colony PCR using M13 forward (-20)(5'-GTAAAACGACGGCCAG-3') and M13 reverse (5'-CAGGAAACAGCTATGAC-3')

primers, visualized by agarose gel electrophoresis, excised from the gel and purified (see above). Three to four different clones of the fragments were sequenced using both M13 forward and M13 reverse primers in order to identify the presence of different haplotypes within individuals (heterozygotes) and to control for sequencing errors. The sequencing reaction was carried out with the Big Dye® Terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems) based on the dideoxy-mediated chain termination method (Sanger et al. 1977). Sequencing reactions were run on an ABI PRISM® 3100*xl* Genetic Analyzer (Applied Biosystems). The sequenced fragments were verified by a TBLASTX search. Putative introns and exons were determined following the GT-AG rule (Breathnach et al. 1978).

Table 4: Selected candidate genes related to drought stress response or bud phenology.

Name (abbreviation)	Gene	Drought stress/ bud phenology	Reference with investigated species
ALDH	Aldehyde dehydrogenase	Drought stress	Gao and Han 2009 (*Oryza sativa*) Guo et al. 2009 (*Hordeum vulgare)* Sathyan et al. 2005 (*Pinus halepensis*)
Cry	Cryptochrome	Bud phenology	Muleo et al. 2001 (*Prunus cerasifera*)
Dhn	Dehydrin	Drought stress and bud phenology	Vornam et al. 2011 (*Quercus petraea*), Luo et al. 2010 (*Zea mays*), Wachowiak et al. 2009 (*Pinus sylvestris*), Jimenez et al. 2008 (*Fagus sylvatica*), Beck et al. 2007 (*Cicer pinnatifidum)*, Gonzáles-Martínez et al. 2006a (*Pinus taeda)*, Ramanjalu and Bartels 2002 (*Picea glauca*)
ERD	early response to dehydration	Drought stress	Eveno et al. 2008 (*Pinus pinaster*), Gonzáles-Martínez et al. 2006a (*Pinus taeda)*, Street et al. 2006 (*Populus* spp.)
IDH	Isocitrate dehydrogenase	Drought stress	Liu et al. 2010 (*Zea mays)*
*APX*1*, APX*4*, GPX*	Peroxidases (ascorbate and glutathione)	Drought stress	Lu et al. 2010 (*Zea mays)*, Street et al. 2006 (*Populus* spp.)
PhyB	Phytochrome B	Drought stress and bud phenology	Boggs et al. 2010 (*Arabidopsis thaliana*), Ingvarsson et al. 2006 (*Populus tremula*), Frewen et al. 2000 (*Populus* spp.)
CHZFP	cys-his-zinc finger protein	Drought stress	Lu et al. 2010 (*Zea mays)*, Street et al. 2006 (*Populus* spp.)

Table 5: Primer sequences and corresponding annealing temperatures for the selected candidate genes (Accession No: EMBL Nucleotide Sequence Database (http://www.ebi.ac.uk/embl/)).

Name	Gene	Primer sequence (5'-3')	Annealing temperature	Accession No
ALDH	Aldehyde dehydrogenase	F: AAG ATC TGG TGT TGA AAA TGG AG	53 °C	FR774766
		R: TGC ATT CTT CAA AGG AGT GAC		
APX1[2]	Ascorbate peroxidase	F Part 1: AGG CGA AGA GAA AGC TCA GG	55 °C	FR774767
		R Part 1: AAG AAA GCA ACT ATC AGC CTC A		
		F Part 2: AAG CAG ATT TGT TGA CAT TAA TAT TTC		
		R Part 2: GCA AAG AAG GCA TCC TCA TC		
APX4 (Part 1)[1]	Ascorbate peroxidase	F Part 1: ATC AAG GGA ACG CTT TCT ACG	55 °C	FR775801
		R Part 1: TCC ACA TCA CAT CTC AAC AGC		
APX4 (Part 2)[1]	Ascorbate peroxidase	F Part 2: GGC CTC TTA AGT GCC AAT TC	55 °C	FR775801
		R Part 2: CTC CCC TCT GGA TCT GGT TC		
Cry[2]	Cryptochrome	F Part 1: CTT GAG ATG ATG CTC TTG GTT G	53 °C	FR775802
		R Part 1: ATG GGC TCA ATT GGA GAA TC		
		F Part 2: TTT TCT CCA CAG GGA TCA CG		
		R Part 2: AAG TCA TGC TTG GGA CCA TC		
Dhn	Dehydrin	F: TGC ACC CCA AAA TGA TGA AT	54 °C	FR772355
		R: TGA TCC CCT TCT TCT CAT GG		
ERD	early response to dehydration	F: GGC AAT GGA CGT AAT TTC TCA	51 °C	FR775803
		R: CTG GGC TGC TGA ATC GTC		
GPX	Glutathione peroxidase	F: GGC TGC CAT GCC TTT CTC	55 °C	FR796394
		R: GAA ATC ATA GAT AGT CTT CTC CGT AGC		
IDH (Part 1)[1]	Isocitrate dehydrogenase	F: GTG ATC AGT ACA GGG CAA CTG	50 °C	FR796392
		R: AAG GTA CAA GAG GGG CTT TG		
IDH (Part 2)[1]	Isocitrate dehydrogenase	F: GAT GAT ATG GTT GCT TAT GCC ATG	50 °C	FR796392
		R: GGT TTC ACC ACC TTT CTG ATG GAC		
PhyB	Phytochrome B	F: CAG GCA TCG AGG TTT TTG TT	50 °C	FR774765
		R: GAA GGG AAT GCA CCT AGC AG		
CHZFP	cys-his-zinc finger protein	F: CTT TGC AAG GAT GAG ACT GG	50 °C	FR796395
		R: ACG CAT CTG ATG AGC ATT TG		

[1]Both parts belong to the same gene but the sequenced parts do not overlap. [2]Both parts belong to the same gene and the two parts overlap.

2.4.5   Analysis of SNPs

After the comparative sequencing, SNPs appearing at least in two different individuals were selected as correct SNPs to exclude polymerase mistakes. Seventeen SNPs in coding and non-coding regions were chosen from eight different candidate genes for further analysis (Table 6). Primers were designed according to the SNaPshot® Multiplex Kit (Applied Biosystems) by addition of nonhomologous polynucleotides (poly (dT)) of different lengths. Therefore, it was possible to analyse all SNPs in two different multiplexes (Table 6). Primers were checked for self-annealing, dimer and hairpin formations using the program Oligo calc: Oligonucleotide Properties Calculator.

After amplification of the candidate genes, the PCR products were cleaned using 1 unit Exonuclease I (Affymetrix, Santa Clara, USA) and 2.5 units SAP (Shrimp Alkaline Phosphatase; Affymetrix, Santa Clara, USA), 37 °C for one hour, 75 °C for 15 min. SNaPshot® Multiplex Kit (Applied Biosystems) PCR amplifications were conducted in a 10 µl volume containing 5 µl of cleaned PCR product from the different genes, 5 µl Reaction Mix (SNaPshot® Multiplex Kit (Applied Biosystems)) and 0.2 µM of each primer. The PCR protocol consisted of 25 cycles of 96 °C for 10 sec (denaturation), 50 °C for 5 sec (annealing), 60 °C for 30 sec (extension). After the SNaPshot® Multiplex Kit (Applied Biosystems) PCR amplifications, the product was again cleaned using 1 unit SAP (Affymetrix, Santa Clara, USA). Preparations for the SNP analysis were done according to the protocol. SNP analysis reactions were run on an ABI PRISM® 3100*xl* Genetic Analyzer (Applied Biosystems) and scoring was done according to the protocol. No automatic scoring was used.

Table 6: Characterization of the 17 SNP markers with information about the related genes; H_o: observed heterozygosity, H_e: expected heterozygosity tested with 50 samples.

Multiplex/ position	SNP name	Gene name	EMBL Accession No	5' tail length - Primer sequence	Forward/ reverse	Variation	Characteristic	H_o	H_e
1/5	ALDH1			(T:38) – GGC AAC TAA CAC TAA CTG TAA CT	F	C/T	non-coding	0.380	0.358
1/6	ALDH2	Aldehyde Dehydrogenase	FR774766	(T:43) – GAT ACG AAG AGC AAA TTC CTC AC	F	C/A	non-synonymous	0.400	0.407
1/7	ALDH3			(T:51) – ACC CCA GCA GCA AGC CCA TA	R	C/G	non-synonymous	0.080	0.078
1/8	ALDH4			(T:56) – CGT GCA TTA CGA GTT GGG AC	F	T/A	synonymous	0.080	0.078
1/1	IDH1		FR796392 (part 1)	(T:21) – GGA CGT GAT GAG AAG AAA GAG TT	F	C/A	synonymous	0.420	0.379
1/2	IDH2	Isocitrate Dehydrogenase		(T:23) – CCA TGA AAA GTG AAG GAG GTT AT	F	G/A	non-synonymous	0.200	0.182
1/3	IDH3		FR796392 (part 2)	(T:27) – GAT TTT TTA GCC CAA GGT TTG TAT	F	C/G	non-coding	0.380	0.398
1/4	IDH4			(T:35) – C AGC AAG ATT ACC AGC ACT GA	R	G/A	synonymous	0.420	0.379
2/5	APX1.1	Ascorbate Peroxidase	FR774767 (part 1)	(T:39) – CCA CCA TCT CTT ACG CCG ATT T	F	C/T	synonymous	0.440	0.453
2/9	APX1.2			(T:63) – CAT GTC ATC GTG TCT TTA TCC TT	F	G/T	non-coding	0.400	0.465
2/3	APX4.1	Ascorbate Peroxidase	FR775801	(T:30) – CAG GTT AGC TCA TGC TAG ATA	R	G/T	non-coding	0.260	0.228
2/6	APX4.2			(T:53) – G ATG TAG CAC AAG GTG CC	F	C/G	non-synonymous	0.280	0.272
2/7	ERD	early responsive to dehydration	FR775803	(T:50) – GCA AAG CGA ACG AGT TGA AAC AGA	R	G/A	non-coding	0.480	0.505
2/1	dhn1	Dehydrin	FR772355	(T:21) – ACG GCA ACC GAG TTC GCA CT	F	C/G	non-synonymous	0.520	0.498
2/2	dhn2			(T:24) – AGA AGA TAC CAG GTG TTG GGC A	F	C/G	non-synonymous	0.280	0.272
2/8	GPX	Glutathione Peroxidase	FR796394	(T:56) – CGA TAC CAT CCA TCA AAA CCT CA	F	C/T	non-synonymous	0.060	0.059
2/4	PhyB	Phytochrome B	FR774765	(T:34) – CAA TCA TCC TAA CCC TAT TCT G	R	G/A	synonymous	0.460	0.505

2.4.6   Data analysis

Microsatellite analysis

GenAlEx 6.4.1 (Peakall and Smouse 2006) was used to calculate the population diversity parameters number of effective alleles (N_e), observed heterozygosity (H_o), expected heterozygosity (H_e) and fixation index (F). Furthermore, Nei's genetic distance (Nei 1972, 1978), the number of private alleles and allele frequencies were calculated with this program. Furthermore, one AMOVA (Analysis of MOlecular VAriance) was done with this program by using 999 permutations. The dendrogram based on Nei's genetic distance ((Nei 1972, 1978) was created with the program NTSYSpc ver. 2.02 (Rohlf 1998) applying the Unweighted Pair Group Method with Arithmetic mean (UPGMA) logarithm (Sneath and Sokal 1973). To test if the populations are significantly different from each other, a pairwise F_{ST} analysis was done with Arlequin 3.1 (number of permutation: 10,000; Excoffier et al. 2005).

Analysis of the candidate genes

For visual examination, editing and alignments of the sequences, Codon Code Aligner (CodonCode cooperation, www.codoncode.com) and BioEdit version 7.0.9.0 (Hall 1999) using ClustalW multiple alignment (Thompson et al. 1994) were applied. Haplotype diversity and nucleotide diversity (π) of the candidate gene sequences were calculated excluding indels using DnaSP v.5.0 (Librado and Rozas 2009).

Analysis of the single nucleotide polymorphisms

GenAlEx 6.4.1 (Peakall and Smouse 2006) was used to calculate the population diversity parameters observed heterozygosity (H_o), expected heterozygosity (H_e) and fixation index (F). Analysis of allele frequencies and an AMOVA (number of permutations 999) was done with this program, too. Genotype frequencies were calculated using Microsoft Excel. Linkage disequilibrium was tested using Tassel, 1,000 permutations (Bradbury et al. 2007). To test if the populations are significantly different from each other, a pairwise F_{ST} analysis was done with Arlequin 3.1 (number of permutation: 10,000; Excoffier et al. 2005). The same comparison was done for every SNP pair. An interactive chi-square test of independence for k groups (Preacher 2001; http://www.quantpsy.org/chisq/chisq.htm) was used to detect population differences using frequency (count) data (allele frequencies for the SNP data). The dendrograms were created with the program NTSYSpc ver. 2.02 (Rohlf

1998) applying the UPGMA logarithm (Sneath and Sokal 1973). For the association of the genotyped seedlings from the controlled drought stress experiment with the phenotypic trait strongly damaged/undamaged and slightly damaged, the program Tassel 3.0 was used to conduct a GLM (general linear model) analysis (Bradbury et al. 2007).

3 Results

3.1 Microsatellites

3.1.1 Molecular diversity indices

The genetic diversity was found to be high for all locations (mean H_e: 0.617; Table 7). Neither the four different locations nor the different soils (sand and loam) differ significantly concerning genetic diversity. Lower genetic diversity was found for two of the three EST markers in comparison with the six normal microsatellite markers (Table 8). The fixation index is close to zero for all stands (Table 7). Most private alleles were found in population Bad Grund (four alleles at three different loci) and in Göhrde sand (three alleles at two different loci). For the other populations, only one or two (Calvörde loam) private alleles were detected. Allele frequencies for all loci and populations may be found in appendix 3.

Table 7: Molecular diversity indices for the six sampling areas (N_e: Number of effective alleles, H_o: observed heterozygosity, H_e: expected heterozygosity, F: fixation index).

Sampling area	N_e	H_o	H_e	F
Calvörde loam	3.219	0.580	0.599	0.020
Calvörde sand	3.387	0.599	0.619	0.015
Göhrde loam	3.437	0.643	0.635	-0.020
Göhrde sand	3.552	0.602	0.638	0.077
Unterlüß loam	3.299	0.607	0.607	-0.004
Unterlüß sand	3.177	0.594	0.595	0.009
Bad Grund	3.510	0.612	0.630	0.042
Mean	3.369	0.612	0.617	0.020

Table 8: Locuswise molecular diversity indices based on all seven populations, the last three loci are EST-microsatellites.

	sfc0018	sfc0161	sfc1063	sfc1143	FS 3-04	mfs11	GOT066	FIR065	FIR004
N_e	3.887	4.841	5.259	4.234	1.468	3.744	1.209	3.256	2.421
H_o	0.698	0.781	0.800	0.779	0.309	0.739	0.172	0.673	0.498
H_e	0.736	0.787	0.807	0.761	0.313	0.729	0.167	0.686	0.572
F	0.052	0.009	0.009	-0.025	0.018	-0.013	-0.015	0.018	0.126

N_e: No of effective alleles, H_o: observed heterozygosity, H_e: expected heterozygosity, F: Fixation index

3.1.2 Genetic differentiation

Most of the genetic variation was found within populations (96%) rather than between them. The variation among populations differed depending on the locus between 1 and 7% (Table 9). Nevertheless, using pairwise F_{ST} analysis, all populations were significantly different from each other (significance level 0.05; 10,000 permutations). The dendrogram based on Nei's genetic distance revealed that the two plots located next to each other in Unterlüß are genetically very close (0.016; Fig. 6). Even more similar are the sand populations from Calvörde and Göhrde (0.015). Furthermore, the populations in Unterlüß are clearly separated from the other populations (0.032 - 0.066). Calvörde loam is an outlier population clearly separated from all the other populations (0.043 - 0.068). Bad Grund in the Harz Mountains is not an outlier population but grouped together with Calvörde sand (0.024) and Göhrde sand (0.028).

Table 9: Locuswise AMOVA based in all seven populations analysed with microsatellite markers.

Locus	Variation among populations	Locus	Variation among populations
sfc0018	2%	mfs11	1%
sfc0161	4%	GOT066	4%
sfc1063	4%	FIR065	5%
sfc1143	4%	FIR004	7%
FS 3-04	3%		
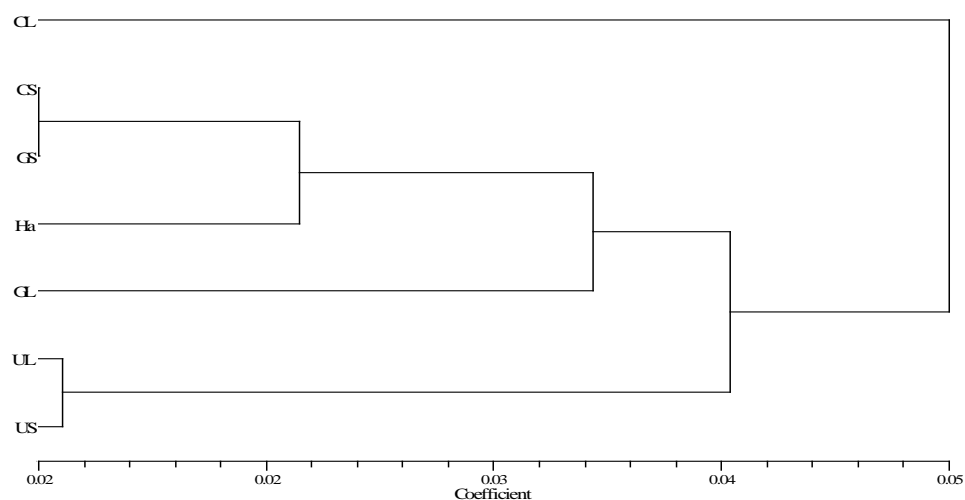
Fig. 6: UPGMA dendrogram using Nei's genetic distance (Nei 1972, 1978) based on microsatellite data (CL: Calvörde loam, CS: Calvörde sand; GL: Göhrde loam; GS: Göhrde sand; Ha: Bad Grund, Harz Mountains; UL: Unterlüß loam; US: Unterlüß sand).

3.2   Analysis of the candidate genes

Initially, fragments from eleven different candidate genes were successfully amplified, identified and analysed. One of the gene fragments (*Glucan endo-1,3-beta-glucosidase*) was found to be extremely variable compared to the other gene fragments. Later on, four SNPs were analysed in this gene. Two of the SNPs were found to be heterozygous in all samples. Most probably, it was not possible to amplify only one gene of this gene family with the designed primers, but more than one gene was amplified. Therefore, the gene fragment *Glucan endo-1,3-beta-glucosidase* and also the analysed SNPs were excluded from further analyses.

Considering only the ten remaining gene fragments, 9468 bp were analysed with 4418 bp in exon regions and 5050 bp in intron regions (Table 10). All exons and introns could be determined following the GT-AG rule. No alternative splicing was found. In seven different genes, 11 indels (insertions/deletions) were identified, mainly in intron regions (Table 10). Some of them showed a microsatellite repeat motif (Appendix 1, supplementary material of the manuscript). Only two indels also represented by microsatellite motives were found within coding regions (gene *ERD* and *CHZFP*). The lengths of these indels were multiples of 3 bps, thus, the reading frame is not shifted.

In total, 63 SNPs were found differently distributed over the analysed gene fragments. More SNPs were found in non-coding regions (1 SNP every 112 bp) than in coding regions (1 SNP every 245 bp). Eighteen SNPs were found in coding regions, and seven of them were non-synonymous. All non-synonymous SNPs led to an amino acid exchange, no one caused an early stop codon. The number of haplotypes ranged from one to eleven. The nucleotide diversity (π) was higher at non-coding sites than at coding sites for most genes. Exceptions are the genes *GPX* and *PhyB* for which the investigated non-coding regions were very short (Table 10). Furthermore, the nucleotide diversity at synonymous sites was in most cases higher than at non-synonymous sites (Table 10).

Detailed information on all 63 SNPs including the exact position of the SNPs, the character (non-coding, synonymous or non-synonymous), the substitution, indel sequences, microsatellite motives and, where appropriate, amino acid changes may be found in the manuscript "DNA Sequence Variation and Development of SNP Markers in Beech (*Fagus sylvatica* L.)" (Appendix 1, supplementary material).

Table 10: Length, exons, introns, indels, SNPs, haplotype and nucleotide diversity for the genes (syn= synonymous; *π x10⁻³, -: gene fragment is completely coding)

Gene name	Total length (bp)	No and length (bp) of exons	No and lengths (bp) of introns	No of indels	No of non-coding SNPs	No of syn. SNPs	No of non-syn. SNPs	No of haplotypes	Haplotype diversity	Total nucleotide diversity*	Nucleotide diversity Non-coding sites	Coding sites	Syn. sites	Non-syn. sites
ALDH	519	3/350	2/169	1	4	2	2	4	0.613	5.91	10.59	3.82	8.60	2.36
APX1	1566	7/577	7/989	4	12	3	0	11	0.584	2.61	3.42	1.3	5.42	0
APX4 (Part 1)	608	2/107	3/501	1	6	0	0	7	0.756	3.91	4.81	0	0	0
APX4 (Part 2)	814	2/238	3/576	1	9	1	1	5	0.756	3.94	4.58	2.41	4.82	1.61
Cry	1439	2/379	1/1060	1	1	0	0	2	0.157	0.11	0.15	0	0	0
dhn	546	2/455	1/91	0	1	0	2	3	0.627	1.83	2.7	1.66	0	2.12
ERD (Part 1)	317	1/242	1/75	0	2	0	0	3	0.629	2.52	10.67	0	0	0
ERD (Part 2)	152	1/145	1/7	1	0	0	0	1	0	0	0	0	0	0
GPX	224	1/224	0/0	0	0	0	1	2	0.157	0.7	0	0.71	0	0.95
IDH (Part 1)	646	3/259	3/387	0	8	2	0	11	0.883	6.62	8.41	4.0	17.7	0
IDH (Part 2)	474	3/222	2/252	1	2	1	1	5	0.744	3.56	3.7	3.4	9.65	1.45
PhyB	301	1/301	0/0	0	0	2	0	3	0.656	2.65	-	2.65	11.83	0
CHZFP	539	1/539	0/0	1	0	0	0	1	0	0	0	0	0	0
Total	8145	29/4038	25/4107	11	45	11	7							

3.3　Single Nucleotide Polymorphism analysis

3.3.1　Molecular diversity indices and linkage disequilibrium

Seventeen SNPs from eight different candidate genes comprising five non-coding, five synonymous and seven non-synonymous SNPs were selected for this investigation (Table 6). The mean expected heterozygosity (H_e) analysed with the 17 SNPs varied between 0.302 and 0.349 (Table 11) for the different populations and depended strongly on the locus (Table 12). The fixation index (F) was negative for one population and close to zero for the other populations (Table 12). The adult trees and the investigated seedling were not different from each other.

Table 11: Molecular diversity indices for all investigated populations and for both marker sets. H_o: observed heterozygosity, H_e: expected heterozygosity, F: fixation index; --: not analysed-

Population	Precipitation (annual mean)	Microsatellite markers			SNP markers		
		H_o	H_e	F	H_o	H_e	F
Calvörde	543.7mm	0.599	0.619	0.015	0.331	0.339	0.016
Göhrde	664.5mm	0.602	0.638	0.077	0.326	0.327	0.006
Unterlüß	765.8mm	0.594	0.595	0.009	0.337	0.316	-0.066
Mean		0.598	0.617	0.034	0.331	0.327	-0.015
Göhrde seedlings	664.5mm	--	--	--	0.332	0.349	0.031
Bad Grund seedlings	1169.9mm	--	--	--	0.297	0.302	0.014
Total mean					0.325	0.327	0.000

The analysis of grouping the SNPs according to non-coding, synonymous or non-synonymous characteristics (Table 13) showed that the observed and expected heterozygosities are lowest for non-synonymous SNPs, except in the population Bad Grund. The highest values for H_o and H_e were found for non-coding SNPs in the populations Göhrde, Unterlüß and Göhrde seedlings and highest for synonymous SNPs in the populations Calvörde and Bad Grund seedlings (Table 13). The fixation index (F) was negative or close to zero for all three groups.

Table 12: Locuswise genetic diversity indices based on the SNP markers

	IDH1	IDH2	IDH3	IDH4	ALDH1	ALDH2	ALDH3	ALDH4	dhn1
N_e	1.901	1.337	1.401	1.903	1.656	1.806	1.171	1.091	1.886
H_o	0.462	0.251	0.269	0.462	0.388	0.460	0.142	0.083	0.503
H_e	0.472	0.246	0.280	0.472	0.392	0.444	0.145	0.082	0.469
F	0.013	-0.042	0.031	0.013	0.006	-0.037	0.017	-0.015	-0.071

	dhn2	APX1.1	APX1.2	APX4.1	APX4.2	PhyB	ERD	GPX
N_e	1.246	1.5	1.567	1.297	1.565	1.899	1.924	1.176
H_o	0.193	0.325	0.362	0.219	0.338	0.454	0.457	0.154
H_e	0.193	0.325	0.354	0.225	0.354	0.473	0.479	0.147
F	0.004	-0.003	-0.024	0.036	0.043	0.038	0.049	-0.051

N_e: No of effective alleles, H_o: observed heterozygosity, H_e: expected heterozygosity, F: Fixation index

Table 13: Molecular diversity indices for non-coding, synonymous and non-synonymous SNPs.

	Non-coding			Synonymous			Non-synonymous		
Population	H_o	H_e	F	H_o	H_e	F	H_o	H_e	F
Calvörde	0.337	0.328	-0.036	0.343	0.369	0.077	0.320	0.326	0.011
Göhrde	0.339	0.362	0.053	0.361	0.361	-0.004	0.291	0.278	-0.022
Unterlüß	0.398	0.370	-0.069	0.372	0.359	-0.047	0.268	0.245	-0.076
Göhrde Seedlings	0.364	0.397	0.084	0.358	0.394	0.064	0.292	0.284	-0.029
Bad Grund seedlings	0.256	0.274	0.067	0.352	0.341	-0.044	0.288	0.295	0.018
Mean	0.339	0.346	0.02	0.357	0.365	0.009	0.292	0.285	-0.02

The linkage disequilibrium test was made separately per population for the 136 SNP pairs. Between 13 (Bad Grund seedlings) and 30 SNP pairs (Calvörde sand) were found to be significantly linked. Most of these SNP pairs were found within genes. Nine of the SNP pairs were significantly linked in all five populations, all within genes (*IDH*, *ALDH* and *APX*1). Two SNP pairs were significantly linked in three populations, both within genes (*ALDH* and *dhn*). Two other SNP pairs between genes were significantly linked in two populations (in Calvörde sand and Göhrde sand seedlings; *ALDH*4 and *IDH*1, *ALDH*4 and

*IDH*4). All the other SNP pairs were significantly linked only in one population. Detailed information about the linkage disequilibrium test is available in Appendix 4.

3.3.2 Genetic differentiation

The AMOVA revealed that most of the genetic variation was found within populations (95%). The variation among populations strongly depended on the analysed SNP and varied between 0 and 11% (Table 14). However, all population pairs were significantly different from each other (Table 15).

Table 14: AMOVA based in all four populations analysed with SNP markers

Locus	Variation among populations	Locus	Variation among populations
*ALDH*1	4%	*APX*1.1	5%
*ALDH*2	5%	*APX*1.2	5%
*ALDH*3	0%	*APX*4.1	2%
*ALDH*4	0%	*APX*4.2	4%
*IDH*1	11%	*dhn*1	2%
*IDH*2	3%	*dhn*2	3%
*IDH*3	3%	*ERD*	2%
*IDH*4	11%	*GPX*	1%
		PhyB	4%

Table 15: Genetic differentiation of the populations investigated with SNP markers based on pairwise F_{ST} (Arlequin 3.1, Excoffier et al. 2005), exact F_{ST} values in the upper part, significance values in the lower part (*= $p<0.05$, **= $p<0.01$, ***= $p<0.001$).

	Calvörde sand	Göhrde sand	Unterlüß sand	Göhrde seedlings	Bad Grund seedlings
Calvörde sand	--	0.00538	0.05392	0.02054	0.00589
Göhrde sand	*	--	0.02654	0.00847	0.01045
Unterlüß sand	***	***	--	0.02884	0.06492
Göhrde seedlings	***	**	***	--	0.03170
Bad Grund seedlings	*	**	***	***	--

A dendrogram based on Nei's genetic distance was calculated for the SNP data, but an unambiguous construction was not possible due to the same distance of two different populations to the population Calvörde sand. Therefore, pairwise F_{ST} values were used for the dendrogram (Fig. 7). Unterlüß is the outlier population (0.0265 - 0.0649) while Calvörde sand, Göhrde sand and the seedlings from Bad Grund are very close (0.0054 - 0.0105). Furthermore, the adult trees from Göhrde sand and the seedlings from this population are also very similar (0.0085; Table 15).
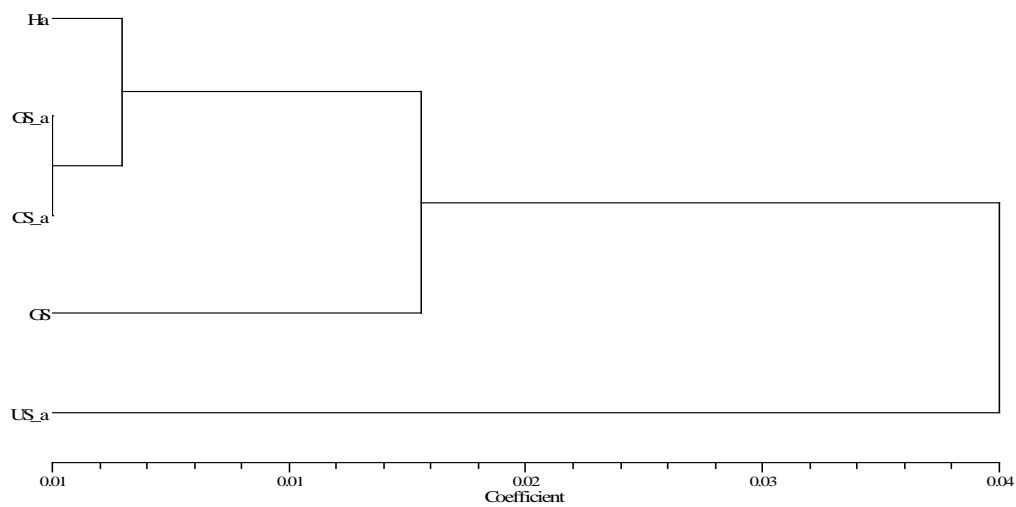
Fig. 7a: UPGMA dendrogram using pairwise F_{ST} based on SNP data (CS_a: Calvörde sand adult trees; GS: Göhrde loam seedlings; GS_a: Göhrde sand adult trees; Ha: Bad Grund, Harz Mountains, seedlings; US_a: Unterlüß sand adult trees).

The three adult populations were investigated with microsatellite and SNP markers. All populations were significantly different from each other, independent of the marker type (Table 16). The F_{ST} values for the populations Göhrde and Unterlüß were almost the same. For the other two population pairs, higher F_{ST} values were found using SNP markers (Table 16).

Table 16: Population comparison based on pairwise F_{ST} values (Arlequin 3.1, Excoffier et al. 2005) for microsatellite and SNP markers for the populations along the precipitation gradient. P-value = * p<0.05, ** p<0.01, *** p<0.001.

	Calvörde - Unterlüß 543.7 mm - 765.8 mm*		Göhrde - Unterlüß 664.5 mm - 765.8 mm*		Göhrde - Calvörde 664.5 mm - 543.7 mm*	
	F_{ST} value	significance level	F_{ST} value	significance level	F_{ST} value	significance level
Microsatellites	0.02206	***	0.02694	***	0.00397	*
SNPs	0.05392	***	0.02654	***	0.00538	*

*annual mean precipitation

3.3.3 Population comparison for the populations along the environmental gradient

A comparison of the two populations with the highest and the lowest amount of precipitation (Calvörde and Unterlüß) revealed that ten out of the 17 analysed SNPs were significantly different from each other (Table 17). These SNPs were located in six out of the eight investigated genes. Four out of the ten SNPs were non-synonymous. Four SNPs from three different genes were found to be significantly different between the population with the medium amount of precipitation (Göhrde) and the highest amount of precipitation (Unterlüß). All four SNPs were synonymous or non-coding SNPs. Only two SNPs were significantly different for the population pair with the medium amount of precipitation (Göhrde) and the lowest amount of precipitation (Calvörde) comprising two genes and both SNPs were non-synonymous. A third SNP (*IDH* 1, non-coding) was close to significant differences between the two populations (Table 17). Allele and genotype frequencies for all SNPs may be found in Table 18a and 18b. The chi-square test of independence mainly confirmed the results of the pairwise F_{ST} analysis (Appendix 5a and 5b). For the population pair Calvörde-Unterlüß, the SNP *ALDH*2 was not significantly different applying the chi-square test. However, the SNP *GPX* was slightly significantly different using the chi-square

test but not using the F_{ST} analysis. For the population pair Göhrde-Unterlüß, the SNP *ALDH*1 was found to be significantly different only using the F_{ST} analysis.

Table 17: Population comparison based on pairwise F_{ST} values (Arlequin 3.1, Excoffier et al. 2005) of SNPs for the pairs of the populations along the environmental gradient. P-value = * p<0.05, ** p<0.01, *** p<0.001, ns: not significant.

Location		Calvörde - Unterlüß 543.7 mm - 765.8 mm°		Göhrde - Unterlüß 664.5 mm - 765.8 mm°		Göhrde - Calvörde 664.5 mm - 543.7 mm°	
SNP	Characteristic	F_{ST} value	Significance level	F_{ST} value	Significance level	F_{ST} value	Significance level
ALDH 1	non-coding	0.00247	ns	-0.00390	ns	-0.00202	ns
ALDH 2	non-syn.	0.01566	*	-0.00017	ns	0.00065	ns
ALDH 3	non-syn.	0.01315	ns	0.00822	ns	-0.00444	ns
ALDH 4	synonymous	0.00239	ns	-0.00329	ns	0.01062	ns
IDH 1	synonymous	0.16693	***	0.07771	***	0.01585	ns (0.0515)
IDH 2	non-syn.	0.04703	**	0.00219	ns	0.01740	*
IDH 3	non-coding	0.02847	**	0.00011	ns	0.00803	ns
IDH 4	synonymous	0.15271	***	0.07771	***	0.01037	ns
*APX*1.1	synonymous	0.04786	**	0.03682	**	-0.00415	ns
*APX*1.2	non-coding	0.04245	**	0.00970	ns	0.00506	ns
*APX*4.1	non-coding	-0.0494	ns	0.00364	ns	0.00313	ns
*APX*4.2	non-syn.	0.07388	***	0.00928	ns	0.02355	*
dhn 1	non-syn.	0.03708	**	0.00865	ns	0.00333	ns
dhn 2	non-syn.	0.00084	ns	-0.00109	ns	0.01387	ns
ERD	non-coding	-0.00109	ns	0.00900	ns	-0.00171	ns
GPX	non-syn.	0.01315	ns	0.00268	ns	-0.00243	ns
PhyB	synonymous	0.07245	***	0.08115	***	-0.00458	ns

non-syn.: non-synonymous,°annual mean precipitation

Table 18a: Allele and genotype frequencies of all SNPs for the populations along the environmental gradient (N= sample size by populations; CS= Calvörde sand, US= Unterlüß sand, GS= Göhrde sand).

SNP	Allele/ genotype	CS	US	GS	SNP	Allele/ genotype	CS	US	GS
ALDH 1	N	104	100	102	*IDH* 2	N	103	100	103
	T	0.288	0.235	0.255		A	0.199	0.085	0.121
	C	0.712	0.765	0.745		G	0.801	0.915	0.879
	TT	0.087	0.020	0.098		AA	0.068	0.000	0.000
	CT	0.404	0.430	0.314		AG	0.262	0.170	0.243
	CC	0.510	0.550	0.588		GG	0.670	0.830	0.757
ALDH 2	N	104	100	100	*IDH* 3	N	103	100	103
	A	0.370	0.275	0.320		C	0.874	0.775	0.816
	C	0.630	0.725	0.680		G	0.126	0.225	0.184
	AA	0.125	0.050	0.110		CC	0.748	0.610	0.660
	AC	0.490	0.450	0.420		CG	0.252	0.330	0.311
	CC	0.385	0.500	0.470		GG	0.000	0.060	0.029
ALDH 3	N	104	100	102	*IDH* 4	N	103	100	103
	C	0.913	0.960	0.922		A	0.437	0.725	0.524
	G	0.087	0.040	0.078		G	0.563	0.275	0.476
	CC	0.837	0.920	0.863		AA	0.214	0.490	0.272
	CG	0.154	0.080	0.118		AG	0.447	0.470	0.505
	GG	0.010	0.000	0.020		GG	0.340	0.040	0.223
ALDH 4	N	104	100	102	*APX1.1*	N	103	100	103
	A	0.933	0.960	0.971		T	0.830	0.690	0.816
	T	0.067	0.040	0.029		C	0.170	0.310	0.184
	AA	0.875	0.920	0.941		TT	0.709	0.470	0.670
	AT	0.115	0.080	0.059		TC	0.243	0.440	0.291
	TT	0.010	0.000	0.000		CC	0.049	0.090	0.039
IDH 1	N	103	100	103	*APX1.2*	N	103	100	103
	T	0.422	0.725	0.524		T	0.189	0.325	0.248
	C	0.578	0.275	0.476		G	0.811	0.675	0.752
	TT	0.204	0.490	0.272		TT	0.039	0.090	0.058
	TC	0.437	0.470	0.505		TG	0.301	0.470	0.379
	CC	0.359	0.040	0.223		GG	0.660	0.440	0.563

Table 18b: Allele and genotype frequencies of all SNPs for the populations along the environmental gradient (N= sample size by populations; CS= Calvörde sand, US= Unterlüß sand, GS= Göhrde sand).

SNP	Allele/ genotype	CS	US	GS	SNP	Allele/ genotype	CS	US	GS
*APX*4.1	N	103	100	103	*ERD*	N	103	100	102
	T	0.884	0.885	0.840		A	0.539	0.495	0.578
	G	0.117	0.115	0.160		G	0.461	0.505	0.422
	TT	0.777	0.780	0.709		AA	0.282	0.220	0.363
	TG	0.214	0.210	0.262		AG	0.515	0.550	0.431
	GG	0.010	0.010	0.029		GG	0.204	0.230	0.206
*APX*4.2	N	103	99	102	*GPX*	N	104	100	103
	C	0.650	0.828	0.760		T	0.913	0.960	0.932
	G	0.350	0.172	0.240		C	0.087	0.040	0.068
	CC	0.447	0.677	0.559		TT	0.827	0.920	0.864
	CG	0.408	0.303	0.402		TC	0.173	0.080	0.136
	GG	0.146	0.020	0.039		CC	0.000	0.000	0.000
dhn 1	N	104	100	103	*PhyB*	N	104	100	103
	C	0.332	0.475	0.393		A	0.351	0.550	0.340
	G	0.668	0.525	0.607		G	0.649	0.450	0.660
	CC	0.106	0.200	0.117		AA	0.115	0.350	0.117
	CG	0.452	0.550	0.553		AG	0.471	0.400	0.447
	GG	0.442	0.250	0.330		GG	0.413	0.250	0.437
dhn 2	N	104	100	103					
	C	0.832	0.870	0.898					
	G	0.168	0.130	0.102					
	CC	0.683	0.750	0.816					
	CG	0.298	0.240	0.165					
	GG	0.019	0.010	0.019					

3.3.4 Population comparison for the populations in the climate chamber experiment

The seedlings from Göhrde and Bad Grund were significantly different from each other using SNP markers (Table 15). Both populations were not analysed with microsatellite markers. Eleven out of the 17 SNPs were significantly different between the seedlings from the two populations (Table 19). A comparison of the strongly damaged with the undamaged

(and slightly damaged) seedlings after the drought stress experiment showed that only one SNP (*APX*4.1) was slightly significantly different between the two groups (Table 19). The association analysis using Tassel confirmed that the SNP *APX*4.1 was associated with the phenotypic trait. Another SNP was slightly significantly different comparing only the strongly with the undamaged seedlings from Bad Grund (*ALDH2*). Allele and genotype frequencies for all SNPs for the damaged and the not damaged/slightly damaged seedlings are shown in Appendix 6. The chi-square test of independence confirmed the results of the F_{ST} analysis (Appendix 5b).

Table 19: Population comparison (pairwise F_{ST}; Arlequin 3.1, Excoffier et al. 2005),) for all 17 SNPs of the seedlings from Göhrde and Bad Grund (controlled drought stress experiment; *= p<0.05, **= p<0.01, ***= p<0.001, ns= not significant)

	Bad Grund - Göhrde		Strongly damaged - not damaged*		Bad Grund Strongly damaged - not damaged*		Göhrde Strongly damaged - not damaged*	
SNP	F_{ST} value	Signifi-cance value	F_{ST} value	Signifi-cance value	F_{ST} value	Signifi-cance value	F_{ST} value	Signifi-cance value
ALDH 1	0.07331	***	-0.00354	ns	0.00920	ns	-0.00820	ns
ALDH 2	0.06116	***	-0.00418	ns	0.02395	*	0.00277	ns
ALDH 3	-0.00444	ns	-0.00444	ns	-0.00028	ns	-0.00786	ns
ALDH 4	0.00668	ns	-0.00372	ns	-0.00989	ns	-0.00928	ns
IDH 1	0.04523	**	-0.00422	ns	-0.00926	ns	-0.00928	ns
IDH 2	0.01805	*	-0.00357	ns	-0.00238	ns	-0.00964	ns
IDH 3	0.04265	**	-0.00419	ns	-0.00528	ns	0.00044	ns
IDH 4	0.04835	**	-0.00457	ns	-0.00989	ns	-0.00928	ns
*APX*1.1	0.05950	***	-0.00115	ns	-0.00964	ns	0.00617	ns
*APX*1.2	0.05981	***	-0.00027	ns	0.00489	ns	-0.00912	ns
*APX*4.1	0.04205	**	0.01927	*	0.03288	ns	0.00618	ns
*APX*4.2	0.02200	*	-0.00136	ns	-0.01013	ns	0.00956	ns
dhn 1	-0.00416	ns	-0.00503	ns	-0.00816	ns	-0.00812	ns
dhn 2	0.02072	*	-0.00213	ns	-0.00893	ns	0.00325	ns
ERD	-0.00365	ns	-0.00453	ns	-0.00814	ns	-0.00203	ns
GPX	-0.00022	ns	-0.00449	ns	-0.00248	ns	0.01186	ns
PhyB	-0.00373	ns	0.00125	ns	0.00020	ns	-0.00679	ns

*and slightly damaged

4 Discussion

The present study is the first analysing additionally to the neutral genetic diversity also the adaptive genetic diversity using a candidate gene approach of the ecologically and economically important tree species *Fagus sylvatica*. Global climate change and the increased demand for wood require genetic investigations not only for a few model tree species (Neale and Kremer 2011). At the moment, the most intensively studied tree species belong to four families: Pinaceae (different species), Salicaceae (*Populus*), Myrtaceae (*Eucalyptus*) and Fagaceae (*Quercus* and *Castanea*). The two already intensively investigated Fagaceae species *Quercus* and *Castanea* have a low genetic divergence, thus, molecular markers can be easily transferred from one species to another (Kremer et al. 2007). A project supported by the Forest Health Initiative (FHI) has already started to sequence the genome of the Chinese chestnut (*Castanea mollissima*; http://foresthealthinitiative.org/genomics.html). Unfortunately, Fagus diverged earlier from the common ancestor and the rate of successfully transferred molecular markers is not that high (Neale and Kremer 2011).

The rapid development of molecular techniques, in particular next-generation sequencing, will enable to study the genetic background of important phenotypic traits in detail also for non-model tree species in the near future (Neale and Kremer 2011). Until now, the scarcity of sequence information available for non-model organisms with no sequenced reference genome limits the development of suitable markers. Therefore, a small set of candidate genes and SNP markers were applied in this study.

4.1 Molecular diversity indices, linkage disequilibrium and differentiation

4.1.1 Genetic diversity

The neutral genetic diversity analysed with microsatellite markers is high for all populations (H_e mean: 0.617). The mean value is lower compared to other studies using microsatellite loci to analyse the genetic diversity of beech in Central Europe (Oddou-Muratorio et al. 2011: H_e = 0.72; Buiteveld et al. 2007: H_e = 0.829; Vornam et al. 2004: H_e = 0.765). However, these studies used different loci, and the genetic diversity measures strongly depend on the selected loci. The EST-SSRs used in this investigation are known to be less polymorphic. K.C. (2011) applied the same marker set and found comparable results for beech stands in three different and widely separated regions in Germany (H_e mean: 0.622, range: 0.596 - 0.656).

The analysis of the SNP markers confirmed the high genetic diversity of all populations (H_e mean: 0.327). The absolute values are lower because microsatellite markers are highly polymorphic markers in comparison to SNP markers. Therefore, it is not possible to compare the values directly. There are no comparable studies until now investigating the important adaptive genetic diversity of beech based on SNP analysis. Neutral and adaptive genetic diversity differ in their impact on the survival of organisms. Only the adaptive genetic diversity enables populations to adapt to the changing environmental conditions in the future (e.g., Gebremedhin et al. 2009). This study was able to confirm the hypothesis that the genetic variation is high at putatively neutral genetic markers and at genes potentially involved in adaptive responses to climate change.

For a common, wind-pollinated and highly outcrossing species like beech, the detected high values of genetic diversity are expected. The retention of a high genetic diversity is an important basis for the adaptation of beech to changing environmental conditions. It may be assumed that the genetic diversity of beech in Germany will not decrease in the near future because of the ongoing extension of beech forests. Additionally, there is no main impact of management activities at least on the neutral genetic diversity of the species (K.C. 2011, Buiteveld et al. 2007). Nevertheless, Oddou-Muratorio et al. (2010) suggest some strategies to further increase the genetic diversity of European beech at a local scale, e.g., to spread the recruitment process across several years, and to promote regeneration plots close to and far from the parent trees.

4.1.2   Nucleotide and haplotype diversity of the candidate genes

Parts of ten different candidate genes related to drought stress and/or bud phenology were successfully analysed in this study. As expected, most of the 63 SNPs and 11 indels were found in intron regions and most of the SNPs found in coding regions were synonymous. Two of the detected indels were located in coding regions. They are of special interest because they may alter protein structures and function which may also influence phenotypic traits (e.g., reviewed by Li et al. 2002). However, only short indels were found in this study, and they are supposed to have only minor impact on protein structures (Kim and Guo 2010).

A broad range of nucleotide diversity ($\pi \, x10^{-3}$) was found among genes in this study (0 to 6.62) which is comparable to the range found for other deciduous tree species and conifers (Table 20). No comparable results concerning the nucleotide diversity of *Fagus sylvatica*

have been published. The mean nucleotide diversity of 2.64 found in this study is low in comparison to most other already investigated tree species (Table 20). Because of the life history traits of beech (e.g., highly outcrossing and wind-pollination), a lower mean nucleotide diversity is unexpected. However, the mean values for nucleotide diversity strongly depends on the investigated genes (or gene fragments), the number of the investigated genes and the length of the sequences. Because all investigations examined different candidate genes and most of them analysed only a few genes, the results for the mean values are not directly comparable (Table 20). Furthermore, the number of SNPs in the presented data is most likely underestimated. To avoid false positive SNPs due to sequencing errors, all SNPs appearing only once were excluded from the analysis. In addition, only a limited number of trees and clones per tree (*Escherichia coli* transformants, see chapter 2.3.4) were used for comparative sequencing.

Table 20: Nucleotide diversity of different species analysed based on different candidate genes.

Study species	Reference	Mean nucleotide diversity (π x10^{-3})	Nucleotide diversity range (π x10^{-3})	Number of investigated genes	Length of investigated sequences [bp]
Fagus sylvatica	this study	2.64	0.0 – 6.62	10	8,145
Populus tremula	Ingvarsson 2005	11.1	2.7 - 18.8	5	6,188
Quercus crispula	Quang et al. 2008	6.93	6.67 - 7.21	11	6,477
Quercus petraea	Derory et al. 2010	6.15	1.09 - 14.7	9	4,638
Quercus petraea	Gailing et al. 2009	5.42	3.02 - 11.96	12	4,981
Picea abies	Heuertz et al. 2006	2.08	0.24 - 6.81	22	15,836
Pinus sylvestris	Wachowiak et al. 2009	6.0	0.6 - 20.9	14	9,635
Pinus taeda	González-Martínez et al. 2006a	7.5	0.1 - 11.79	18	10,116
Pinus taeda	Brown et al. 2004	3.98	0.27- 17.28	19	17,580
Pseudotsuga menziesii	Krutovsky and Neale 2005	6.55	2.37 - 13.78	18	15,183

4.1.3 Linkage disequilibrium

Linkage disequilibrium (also known as gametic phase disequilibrium and allelic association) is the "nonrandom association of alleles at different loci" (Flint-Garcia et al. 2003). It is important to differentiate between linkage disequilibrium between genes and even chromosomes and the linkage disequilibrium within genes. Linkage disequilibrium within genes is determined by the correlated inheritance of two or more variable sites due to their physical connection on a chromosome and is therefore more frequent than linkage disequilibrium between genes and chromosomes (e.g., Flint-Garcia et al. 2003).

For the investigation of the linkage disequilibrium, only the 17 SNPs were analysed that were screened in the three adult populations and the drought stress experiment. An analysis of all SNPs found by the comparative sequencing was not possible, because the total sample size (18 individuals) was too low. In this study, between 13 and 30 SNP pairs were found to be significantly linked. As expected, most of these pairs were found within genes separated by only a few hundred basepairs. Linkage disequilibrium tends to decay with distance and depends among other things on demographic history of the population and the degree of inbreeding (Nordborg et al. 2002). For forest trees, a rapid decay of linkage disequilibrium (LD) within genes has been frequently reported in forest tree populations (González-Martínez et al. 2006b). Neale and Savolainen (2004) reported that most SNPs separated by more than 1,500 bp are most likely not in linkage disequilibrium in *Pinus taeda*. In this study, some linked SNP pairs were also found between genes. However, the pairs that are linked in at least three out of the five populations were within genes. Significantly linked SNPs between genes were found only in one or two populations. Therefore, this is probably only an artefact produced by the low sample size. The SNPs found to be in linkage disequilibrium are not completely linked and it is still reasonable to investigate them.

4.1.4 Genetic differentiation analysed with microsatellites and SNPs

The average level of genetic differentiation was low to moderate, but significantly different from zero at the 95% confidence interval for all populations. This is not unusual for trees. Namroud et al. (2008) found even lower genetic differentiation for *Picea abies* but still significantly different. For *Pinus sylvestris*, comparable low values were found analysing candidate genes, although populations from all over Europe were investigated (Pyhäjärvi et al. 2007).

There was almost no difference concerning differentiation levels applying microsatellite or SNP markers, even though, the differentiation was slightly higher for two of the population pairs using SNP markers. Derory et al. (2010) investigated oaks using SNPs and microsatellites and also detected higher F_{ST} values for SNPs in comparison to microsatellite markers. For beech, a lot of studies (e.g., K.C. 2011, Oddou-Muratorio et al. 2011, Papageorgiou et al. 2008, Sander et al. 2000, Konnert 1995) confirm that most of the genetic variation is found within populations and therefore, the genetic differentiation between populations is low. However, the investigation of 30 beech stands from three different regions in Germany revealed that almost all population pairs are significantly different from each other (K.C. 2011).

A comparison of genetic differentiation patterns revealed similarities between the analysis of microsatellites and SNPs (described in chapter 3.1.2 and 3.3.1). Unterlüß appeared as an outlier population with both investigated marker types. Therefore, the highly significant differentiation revealed by SNPs between Unterlüß, the population with the highest amount of precipitation, and the two other populations probably adapted to drought stress at least in the summer months, is not necessarily an indication for adaptation. The differences might be explained by different population histories in the past. The populations Calvörde, Göhrde and Bad Grund are less differentiated at both marker types. Although the trees in Bad Grund are apparently adapted to other environmental conditions (Table 1), some of the investigated SNPs are not significantly differentiated and hence most likely not affected by selection. Therefore, it is not unexpected that the population with the highest amount of precipitation is genetically close to the population with the lowest amount of precipitation, even at SNP markers.

The genetic differentiation patterns at microsatellite and SNP markers (Fig. 6 and 7) do not reflect the geographic distances of the populations. K.C. (2011) were able to show with a dendrogram based on Nei's genetic distance (Nei 1972, 1978) that almost all investigated beech populations grouped together according to their geographical clearly distinguishable regions (Schorfheide-Chorin, Hainich-Dün, Schwäbische Alb). In contrast, in this study, the population near Bad Grund is geographically most separated from the other populations but the genetic distance to the sand populations in Göhrde and Calvörde is low. Furthermore, the population Calvörde loam was found to be an outlier population based on microsatellites, although this population is geographically very close to Calvörde sand (Table 2; this population was not analysed with SNP markers). Although natural

regeneration is and was mainly used in Germany for the regeneration of beech stands, it is possible that some of the populations were planted in the past. The origin of the investigated populations is unknown. However, K.C. (2011) found also, that within regions on a fine scale, the populations do not group according to their geographic distance. The reason is mainly that the variation within populations is higher than the variation among populations. Additionally, there was no linear recolonization of beech from the south to the north after the last ice age. It may be concluded that the investigated beech populations in this study do not group according to their geographical distance, because they do not belong to clearly distinguishable regions. The largest distance between the studied populations is around 150 km (Göhrde - Bad Grund; Table 2), whereas the regions investigated by K.C. (2011) are separated by around 300 km.

The analysis of genetic differentiation was largely based on the "classical" measure F_{ST} (Wright 1951, 1965). The measurement of genetic differentiation at SNPs with the help of F_{ST} is beyond controversy because F-statistics were developed for biallelic markers (Meirmans and Hedrick 2011). F-statistics have also been used for microsatellites as a standard tool since years, but several authors recently questioned if F_{ST} (or related measurements like G_{ST}) may be applied for microsatellite markers (e.g., Leng and Zhang 2011, Whitlock 2011, Gerlach et al. 2010, Jost 2009, Ryman and Leimar 2009). Two alternative measurements were developed exclusively for microsatellite markers: G'_{ST} and Jost's D (Hedrick 2005, Jost 2008). Ryman and Leimar (2009) came to the conclusion that Jost's D is not suitable to compare SNP markers with microsatellite markers. Furthermore, White et al. (2010) analysed microsatellite markers comparing F_{ST} with Jost's D. They found very similar results. G'_{ST} may cause problems analysing SNP markers because of the characteristic of having normally only two allelic stages (Meirmans and Hedrick 2011). Therefore, F_{ST} is used in this study for both marker types. Although the direct comparison is under dispute, it is frequently used in comparable experimental studies (e.g., De Carvalho et al. 2010, Hall et al. 2007, Ryynänen et al. 2007).

It may be concluded that the hypothesis that the genetic differentiation in European beech populations is higher at genes putatively involved in adaptation than at putatively neutral genetic markers was not confirmed. The differentiation was only slightly higher for two of the three population pairs using SNP markers. However, the direct comparison of F_{ST} values between microsatellite and SNP markers is questionable. That is why the hypothesis may not be answered thoroughly.

4.2   Population comparisons

4.2.1   Population comparisons along the environmental gradient

The investigated populations are located relatively close together in the centre of the distribution of *Fagus sylvatica*. Nevertheless, the annual mean precipitation differs significantly between the three regions. In contrast, other factors of environmental importance are comparable. These are good pre-conditions for the investigation of the genetic bases of the phenotypic trait "resistance to drought stress".

A comparison of the populations with the highest (Unterlüß) and the lowest (Calvörde) amount of precipitation revealed that ten of the seventeen SNPs comprising non-coding, synonymous and non-synonymous SNPs from six genes were significantly different. However, the analysis with putatively neutral microsatellite markers showed that these two populations are genetically rather different even at these markers. Thus, different population histories including different postglacial recolonization routes or human seed transfer cannot be ruled out as a reason for the significant differentiation at numerous SNP markers if only Unterlüß and Calvörde are compared. Rather strong differentiation at microsatellite markers was also observed for the population pair Unterlüß and Göhrde. Accordingly, the observation of significant differentiation between these populations at four SNP markers might also be related to different population histories. The observation of significant differentiation at two SNPs and strong, though not significant differentiation at a third SNP marker for the population pair Calvörde and Göhrde is particularly remarkable since very low differentiation was observed at presumably neutral microsatellite markers for these two populations. These SNPs were found in two different genes, *isocitrate dehydrogenase* and *ascorbate peroxidase* 4, which are possibly involved in the adaptation to drought stress. Since the low differentiation at microsatellite markers suggests a common population history for these stands, selection might be involved in shaping the genetic structures at these genes. Genetic differentiation patterns at both genes suggest their involvement in genetic responses to drought stress. Accordingly, these two genes are of special interest for further investigations. However, SNPs found to be significantly different in the comparison of the other population pairs should not be ignored, mainly the SNPs found in *ascorbate peroxidase* 1 and *phytochrome B*. SNPs in these two genes were found to be highly significantly different in the two other population pairs. These results confirm the hypothesis that some of the SNPs identified in the candidate genes are significantly different between the populations along the precipitation gradient.

Namroud et al. (2008) investigated natural populations of *Picea glauca* exhibiting variation at different quantitative traits (e.g., wood density, phenology and growth) in Canada. F_{ST} analyses revealed 73 SNPs that were significantly different between the populations. Outlier analysis validated between two and 49 of these SNPs as outlier SNPs, depending on the applied method and parameters. Natural populations of *Picea mariana* in eastern Canada were investigated by Prunier et al. (2011) at traits related to temperature and precipitation. They used two different outlier analysis methods and found between one and 16 SNPs that were significantly different. Drought stress candidate genes were analysed for natural populations of *Pinus pinaster* in France, Spain, Morocco and Tunisia (Eveno et al. 2008). Between one and 15 SNPs were found to be outliers depending on the applied method. Derory et al. (2010) were looking for significantly different SNPs for the trait "bud burst" in a common garden experiment with *Quercus petraea* trees from nine different European populations. None of the investigated SNPs was significantly different. All these studies investigated a larger number of SNPs compared to this study and showed that the identification of SNPs involved in phenotypic traits is a challenging task for forest trees. The situation is even more complicated because different statistical methods are available for outlier analyses providing different SNPs as outliers and therefore, potentially involved in adaptation.

In contrast to this study, all of the above mentioned candidate gene investigations on tree populations along an environmental gradient were conducted in populations from a wide range of their distribution suggesting a complex impact of multiple environmental factors including different temperature and precipitation regimes on the observed phenotypic traits. This approach enhances the possibility to find significantly different SNPs but makes it more difficult to functionally link these SNPs to a specific phenotypic trait or a single environmental factor. The experimental design of this study allows to focus on data interpretation of genetic adaptation mechanisms to drought stress.

4.2.2 Population comparison for the climate chamber experiment

The climate chamber experiment was designed in order to create an association population where the phenotypic trait of interest, resistance to drought stress, could be easily scored and associated with variation in candidate genes. Seedlings are more sensitive to drought stress than adult trees, and it is possible to keep them under controlled conditions and to phenotype them easily. Thus, the experiment may be used for the validation of interesting SNPs found in populations along the environmental gradient, or to identify further

potentially selective genes. Only one SNP was found to be significantly different comparing the strongly damaged with the healthy (and slightly damaged) seedlings. The number of 200 seedlings from two different populations (Bad Grund and Göhrde sand) might be too low to find significant differences at more loci. This is also stated by Neale (2007) who mentioned that in trees the association genetic approach to complex trait dissection requires genotyping and phenotyping of a large number of individuals. Thus, the hypothesis that some of the SNPs identified in the candidate genes are significantly different between damaged and not damaged/slightly damaged seedlings may not be completely answered.

The significantly different SNP found in the climate chamber experiment in the *ascorbate peroxidase* gene was not significantly different comparing the populations along the precipitation gradient. Drought stress tolerance is a complex, quantitative trait, like the vast majority of traits in forest trees (Neale and Kremer 2011). These traits are under polygenic control and different genes might be involved in controlling trait expressions during different life stages, for example seedlings and adult trees (see below). Variation in wood quality also seems to be based on many gene loci, each with only small effects (Brown et al. 2003). The *APX*4.1 SNP may be involved in the adaptation to drought stress while the potentially detrimental allele is still present in drought stress adapted populations because this SNP could be compensated by other important genes which are already adapted. In this climate chamber experiment, young seedlings were investigated because of practical constraints to establish experiments with old trees and because they are most sensible to drought stress. Although the trees in Calvörde and Göhrde are most probably adapted to drought stress, nothing is known about the stress conditions in their seedling stage. Furthermore, the detected SNP is located in a non-coding region and thus not necessarily a functionally important variant, but it might be linked to another functionally important SNP. The rapid decay of linkage disequilibrium within genes in forest trees is advantageous in this case. The detected SNP of interest is most likely physical close to the functional variant (e.g., Neale and Kremer 2011). Although the hypothesis that the climate chamber experiment may support the results obtained by the analysis of the populations along the precipitation gradient was not confirmed, the gene *ascorbate peroxidase* and the observed SNP in this gene are of interest for further investigation.

4.2.3 Methodological aspects

For the detection of loci that are under selection, outlier analyses may be used (e.g., Excoffier et al. 2009, reviewed by Nosil et al. 2009). The aim of these analyses is to identify loci that do not behave "neutral". The first method that is used is based on linkage disequilibria between markers. Regions under selection are expected to have a high degree of linkage disequilibrium around the locus under selection (the "hitch-hiking" effect, Smith and Haigh 1974). This method is only suitable to detect recent selection because the linkage breaks up rapidly over evolutionary times (Kirk and Freeland 2011). Furthermore, the position of the loci has to be known. This is not always the case for SNPs, for example for a small number of SNPs extracted from EST databases. The second method based on comparison of genetic variation in random sets of markers. This analysis is used for markers with a large number of loci, for example for AFLP data. New statistical programs offer outlier analyses based on simulations that are also suitable for analyses with less markers, especially for SNPs (see Helyar et al. 2011 for an overview). In this study, a set of 17 SNPs was genotyped and some of the SNPs were in linkage disequilibrium which may bias results using outlier analyses (Helyar et al. 2011). Therefore, this kind of analysis was not applied. In this investigation, a comparison based on pairwise F_{ST} and the significance of this differentiation measure was used, for example also applied by Moen et al. (2008) and Namroud et al. (2008) and discussed above.

One of the severest problems analysing SNPs is the ascertainment bias (e.g., Anderson 2010, Rosenblum and Novembre 2007, Clark et al. 2005, Nielsen 2004). The ascertainment set is the set of individuals used for the detection of SNPs. An ascertainment bias occurs most probably if only a small set of individuals from only a part of the species' range is used for the SNP detection and later on, a large set of individuals is genotyped (e.g., Helyar et al. 2011). The small ascertainment set does not enable to identify SNPs of low frequency which will later on bias all statistical measurements based on allele frequencies (e.g., nucleotide diversity, population size and linkage disequilibrium). This is a widespread problem, because of the costs, most studies are using only a small ascertainment set, including studies based on next-generation sequencing studies (see Seeb et al. 2011 for an overview). Special statistical methods are available that help to correct this bias (e.g., Marth et al. 2004, Nielsen et al. 2004, Polanski and Kimmel 2003). However, to avoid ascertainment bias, it is recommended to select a relatively large sample of individuals from all populations that will be genotyped later (Rosenblum and Novembre 2077, Morin et

al. 2004). In this study, trees from six different populations were selected for the ascertainment set and only three of these populations were genotyped with the detected SNPs. Therefore, an ascertainment bias is unlikely in this study.

Another recently recognized problem is the generation of false positives in outlier detection approaches because of population structures (Excoffier et al. 2009). Although most of the genetic variation was found within the investigated populations based on microsatellite markers (Table 9), all populations were significantly different from each other, even though the F_{ST} values are low (see chapter 3.1.2). However, the use of F-statistics to measure the differentiation of populations by microsatellite markers is questionable (see chapter 4.1.4). Because population structures cannot be excluded in this study, it is possible that some of the significantly different SNPs are false positives, especially for the population pairs Calvörde-Unterlüß and Göhrde-Unterlüß. Unterlüß is highly significantly different from the other populations whereas Calvörde and Göhrde are genetically close (see discussion in chapter 4.2.1).

4.3   The future of European beech under global climate change

European beech is a dominant, highly competitive and widespread tree species in Europe that is well adapted to the present climatic conditions in its range. The adaptability of European beech to the changing climatic conditions is discussed in literature with contrasting opinions. On the one hand, it is suggested that beech is a drought sensitive species, and further increase of drought periods will negatively influence the growth and also the competitive ability of beech (Betsch et al. 2011, Geßler et al. 2004, Rennenberg et al. 2004). Beech seedlings are most sensitive to drought stress. For example, 65% of beech seedlings planted in an afforestation experiment in spring 2003 died after the severe drought in summer that year. The mortality rate in the following year (without severe drought during summer) was reduced to 33% (Don et al. 2007). The experiment also showed that the mortality rate for beech was high in comparison to other studied tree species. Another study showed that in the following year after the severe drought in 2003, seedlings reacted with reduced growth (Czajkowski et al. 2005). In addition, dry air alone is stressful for seedlings reducing biomass production and leaf growth (Lendzion and Leuschner 2008). It can be concluded that the competitiveness of beech will possibly decrease if the length and also the frequency of drought periods in summer will increase. The main competitor tree species, oaks, seems to be better adapted to drier conditions in the summer months (see chapter 1.2.5).

Migration would be another option to react to the changing environmental conditions. Oddou-Muratorio et al. (2010) investigated the seed dispersal of different *Fagus* species and found only a few tens of meters per generation. They conclude that hundreds of meters are necessary to keep up with the 21st century climate change.

However, beech has a high genetic diversity and also a high adaptive potential. Thus, in Germany, most probably only populations in the eastern parts will be threatened by drought stress (Bolte 2005). The observed crown transparency after the extreme dry year 2003 in Germany was found out to be only temporarily (Kölling et al. 2005). The investigated beech trees rapidly recovered. Furthermore, Kölling et al. (2005) even argue that a further distribution to the north and to higher altitudes is possible. Higher temperatures and less cloudy climatic conditions are assumed to improve growth and competitiveness at many locations (Ammer et al. 2005). Simulations confirmed that a warmer climate will increase growth in beech in contrast to Norway spruce where the productivity will seriously decrease (Pretzsch and Ďurský 2002). Specific management strategies may also help to increase the adaptive potential by short recruitment intervals and many mother trees contributing to the next generation (Kramer et al. 2008). Even beech seedlings believed to be most susceptible were able to survive the drought and hot years in 1934 and 1976 (Ammer et al. 2005). Climate envelopes developed by Kölling (2007) demonstrate that almost all regions in Germany will maintain optimal climatic conditions for European beech, but it has to be mentioned that the B1 scenario was used for this model. This scenario is the one with the smallest temperature rise (Fig. 2). Additionally, the use of only annual mean temperature and precipitation for climate envelopes is critically discussed in literature (e.g., Bolte et al. 2008).

Concerning the lengthening of the growing period, later leaf senescence is expected to positively influence the growth of beech (Menzel and Fabian 2001). Although higher temperature was found to induce later leaf senescence, there are still uncertainties regarding the effect of drought potentially preventing a longer growing season (Vitasse et al. 2011 and 2010, Bréda et al. 2006). More severe drought stress in summer may even lead to earlier leaf senescence. Concerning the flushing date in spring, it is supposed that earlier bud burst as well as constant bud burst date may be disadvantageous for beech due to intensified frost risk and stronger competitive interaction with oak (e.g., Vitasse et al. 2011 and see chapter 1.2.5 and 1.2.6).

In general, it has to be taken into account that there are still a lot of uncertainties concerning the climate change predictions, and it is therefore difficult to predict the consequences for European beech (Bolte 2005). Most studies apply a moderate scenario (A1B) with only a moderate change of the climatic conditions. These moderate scenarios suppose that the production of greenhouse gas emissions will decrease latest in the year 2050. More investigations on global climate change and on European beech are necessary to finally estimate its potential to adapt to the changing conditions. At the moment, it may be concluded that probably not all beech forests will be negatively influenced by the changing environmental conditions, but some regions will be definitely out of optimal conditions. In these regions, oak species may be better adapted to dry conditions and may replace beech (e.g., Kölling et al. 2005, Aranda et al. 2000; cf. chapter 1.2.3 and 1.2.5). Even in regions where beech will not be displaced by oak, drought periods in summer can damage trees or slow down their growth. This study shows that evolutionary changes of genetic structures will contribute to the adaptation of these beech populations to changing environmental conditions. Therefore, there is no ecological problem, but there may be an economic one. Most of the European beech forests in Germany are used for timber production and fast growing trees with a good wood quality are required because the sustainable yield of beech timber compared to Norway spruce timber (*Picea abies*) has been low in the last 80 years (e.g., Möhring et al. 2008). Although still in an early phase of basic research, the new sequencing techniques will accelerate the investigations on non-model tree species and results will be available in the near future that may be used for example for tree breeding. Once SNP markers are found to be involved in the trait of interest, marker-assisted breeding may be applied. Compared to traditional breeding, there are some advantages like fewer costs because of the reduction of expensive field experiments and reduced breeding cycle time (e.g., Neale and Kremer 2011).

The results of this study are a first important step to further investigate the adaptability of European beech to climate change. First of all, it is important to identify genes significantly involved in the adaptation to drought stress. Then, drought resistant genotypes have to be identified, for example sets of SNP markers. However, drought stress resistance is a complex trait and many genes and loci will be involved in the adaptation. The selected SNP markers have to be tested in different natural populations and the results should be confirmed additionally with controlled drought stress experiments. In the end, sets of SNP markers significantly involved in the adaptation to drought stress may be used to screen

European beech populations to test their adaptability. Furthermore, these sets may be used for tree breeding.

4.4 Conclusions and outlook

The genetic background of drought stress tolerance in the non-model species *Fagus sylvatica* was investigated in this study. Due to the limited sequences that are available for beech, only a few candidate genes and only a few SNPs within these genes were analysed. Two different approaches were used to investigate the genetic differentiation of the SNPs: an investigation of natural populations along an environmental gradient and a controlled climate chamber experiment. In the last years, more and more researchers investigated the variation in candidate genes in natural populations, especially for trees (e.g., Campbell and Narum 2011, Keller et al. 2011, Williams and Olesiak 2011, Eckert et al. 2010, Derory et al. 2010, Hall et al. 2007). However, caution is recommended investigating only natural populations. Random events or processes like isolation by distance or founder effects may create an illusion of an adaptive effect (Gilchrist and Meats 2010). Kirk and Freeland (2011) therefore recommend validating results by common garden experiments or replicates in other natural populations. However, the controlled climate chamber experiment was not able to validate the SNPs found to be of interest investigating the populations along the precipitation gradient, probably due to the low number of investigated seedlings. Therefore, these genes should not be excluded from the list of candidate genes for drought stress tolerance in European beech, although their significance for phenotypic responses to drought stress is not yet proven (Brookes 1999). Since only fragments of the candidate genes were investigated in this study, it might be possible that SNPs in other regions of these genes are under selection or SNPs that were found by the comparative sequencing but which were not selected for further analysis.

For model species, candidate gene analysis is no longer the method of choice. The selection of candidate genes is risky since stronger adaptive responses might be missed in genomic regions which are not observed. This holds in particular if the information about the genetic control of a phenotypic trait is very limited (Myles et al. 2009). Hudson already suggested in the year 2008, that ecological geneticists will soon be able to perform population genetics at a genome level using next-generation sequencing techniques (the different techniques are for example reviewed by Glenn 2011, Niedringhaus et al.2011 and Pareek et al. 2011). In a review three years later, Helyar et al. (2011) concluded that this prediction will not become reality in a foreseeable future. The sequencing of whole genomes is still

too expensive, especially for population genetics, where a large number of samples have to be investigated. Neale and Kremer (2011) are more optimistic. They formulated five future directions for the forest tree genomic research. The most important one is the sequencing of more reference genomes. They think that this will be accomplished in the near future. Furthermore, research should concentrate more on ecologically important species and not only on species with a high economic value.

Nevertheless, the new techniques may be very helpful to study adaptation. Even without a reference genome, it is possible to study the transcriptomes of species, for example comparing stressed with control plants. These kinds of studies are not only helpful to identify new candidate genes, but are also necessary to confirm candidate genes for a trait of interest identified in a model species like *Arabidopsis*. The RNA-Seq (RNA-Sequencing) method is the latest and the most powerful tool for transcriptome analysis. It has some advantages over the previously developed methods such as microarrays (e.g., Wang et al. 2009). In comparison to microarrays, RNA-Seq does not require prior knowledge of the sequences to be profiled (Martínez-Gómez et al. 2011). Furthermore, transcripts that are expressed at very low levels may be investigated (Flintoft 2008). The technique has already been applied to tree species, e.g. the model tree species *Populus trichocarpa* (Geraldes et al. 2011) and the non-model species *Acacia auriculiformis* and *Acacia mangium* (Wong et al. 2011). Next-generation sequencing may also be used to detect SNPs in candidate genes more efficiently. After the successful identification of genes (or gene fragments), the genes of interest can be sequenced for a high number of individuals in a shorter time comparing to Sanger sequencing. This is possible by multiplexing using different barcodes. Furthermore, the investigation of a larger number of individuals to detect SNPs may help to avoid an ascertainment bias (see chapter 4.2.3)

It can be concluded that next-generation sequencing is a very promising technique that is more and more used also for the investigation of non-model species. Studies like the one presented here will be soon replaced by investigations using larger sample sizes, more candidate genes and more SNPs. Independent of the use of the new techniques, another factor is challenging, too. Most geneticists concentrate on the molecular part of the investigation, but the analysis of phenotypes must not be disregarded. Forest trees are difficult to phenotype. Therefore, the development and improvement of phenotyping technologies should be in the centre of attention (Neale and Kremer 2011).

5 Summary

European beech (*Fagus sylvatica*) is an ecologically and economically important tree species in Central Europe. It covers a large geographical range in Central Europe, with Germany as the centre of the current distribution. Although beech is a highly competitive species, it does not colonize all habitats in its area of distribution, e.g., very dry habitats. Climate change scenarios for Europe and Germany predict higher annual mean temperatures. The overall precipitation in Germany is supposed to change less, but the amount of precipitation will increase in winter, whereas there will be a decrease in the summer months. This will increase the drought stress risk for European beech. Additionally, earlier flushing dates in spring due to higher temperatures may be disadvantageous for beech.

In future, beech will become more important in Germany because of the transition strategy to transform pure conifer stands into pure beech or mixed deciduous stands comprising beech. The adaptability of European beech to the changing environmental conditions is critically discussed in literature. At least in some regions, beech, in particular seedlings, will suffer under increasing drought stress in summer. Therefore, it is important to investigate the genetic background of climate change related traits.

A candidate gene approach and the analysis of variation within these candidate genes were used in this study to investigate the genetic background of drought stress and bud phenology. European beech populations along a precipitation gradient in Northern Germany were studied. Additionally, a controlled drought stress experiment in a climate chamber was conducted comparing strongly damaged seedlings with healthy/slightly damaged seedlings.

First of all, a set of nine neutral microsatellite markers were used to study the neutral genetic diversity of the populations. Then, partial sequences of ten candidate genes probably involved in drought stress and/or bud phenology were identified at the genomic level, and SNPs (Single Nucleotide Polymorphisms) and indels (insertions/deletions) in coding and non-coding regions were analysed. The neutral genetic diversity was found to be high for all locations (mean H_e: 0.617) and is comparable to other beech populations in Germany. In total, 8145 bp were sequenced and analysed, 4038 bp were located in exon and 4107 bp in intron regions. 63 SNPs and 11 indels were detected, which are differently distributed over the studied gene regions. The nucleotide diversity ranged from 0 to 6.62

(π x10^{-3}) and is comparable to other tree species, whereas the mean nucleotide diversity (2.64) for *F. sylvatica* is comparatively low. A set of 17 SNPs selected from eight of the identified candidate genes was established for further investigation. All populations were found to be significantly different with respect to genetic differentiation. Ten of the seventeen SNPs comprising non-coding, synonymous and non-synonymous SNPs from six genes were found to be significantly different comparing the populations with the highest (Unterlüß) and the lowest amount of precipitation (Calvörde). The analysis with neutral microsatellite markers showed that these two populations are genetically very different already at neutral markers. Four SNPs were found to be significantly different between Unterlüß and Göhrde (medium amount of precipitation), a subset of the SNPs significantly different between Unterlüß and Calvörde. Unterlüß and Göhrde are also genetically very different at neutral markers. Thus, different population histories including different postglacial recolonization routes or human seed transfer cannot be ruled out as a reason for the significant differentiation at numerous SNP markers analysing these population pairs. In contrast, Calvörde and Göhrde are genetically very close, analysed with neutral markers. As expected, only two SNPs were significantly different, a third one was almost significant. These SNPs are from two different genes, *isocitrate dehydrogenase* and *ascorbate peroxidase* 4 and are possibly involved in the adaptation to drought stress. Although these two genes are of special interest for further investigations, the SNPs found to be different in the comparison of the other population pairs should not be disregarded. Interesting are mainly *ascorbate peroxidase* 1 and *phytochrome B* that were highly significantly different between the other two populations. The comparison of the damaged and undamaged seedlings from the controlled drought stress experiment revealed that only one SNP in the *ascorbate peroxidase* 4 gene is significantly different. Although this SNP was not significantly different comparing the populations along the precipitation gradient, it is still interesting for further research. Because of the small set of analysed SNPs, only a comparison based on pairwise F_{ST} was conducted. The application of outlier tests was not possible.

The microsatellite analysis confirmed the high neutral genetic diversity of European beech found in former studies. Although this is a good basis for adaptation, it is necessary to study the genetic background of climate change related traits to assess the genetic potential to adapt to the changing environmental conditions. The number of analysed candidate genes and SNPs was low in this study, mainly due to the limited available sequences for *F. sylvatica*. However, it was possible to detect SNPs that are significantly different

between the investigated populations and are therefore possibly involved in the adaptation to drought stress. This study is only a first step for the investigation of the genetic background of drought tolerance in the non-model species European beech. In future, the now available next-generation sequencing techniques may be used for further investigation. RNA-Sequencing is a useful tool to study the transcriptome, for example comparing drought stressed plants with control plants to identify new candidate genes or to validate candidates already known in other plant species. Furthermore, next-generation sequencing may be used to identify and to genotype SNPs. In contrast, the sequencing of large numbers of whole genomes for population genetic studies seems to be out of reach for the near future.

6 Zusammenfassung

Die Rotbuche (*Fagus sylvatica*) ist eine ökonomisch und ökologisch bedeutende Baumart in Mitteleuropa. Das Verbreitungsgebiet erstreckt sich über weite Teile Mitteleuropas mit Deutschland im Zentrum der Verbreitung. Obwohl die Rotbuche eine sehr konkurrenzstarke Baumart ist, besiedelt sie nicht alle Habitate in ihrem Verbreitungsgebiet, ausgenommen sind z.B. sehr trockene Habitate. Klimawandelszenarien prognostizieren für Europa und Deutschland steigende Jahresdurchschnittstemperaturen. Der durchschnittliche Niederschlag soll sich zwar kaum verändern, allerdings wird er in den Wintermonaten zunehmen, wohingegen er in den Sommermonaten abnehmen wird, was das Trockenstressrisiko für die Rotbuche erhöhen wird. Zusätzlich könnten die steigenden Temperaturen einen früheren Austrieb im Frühling verursachen, was sich möglicherweise zusätzlich negativ auf die Buche auswirken wird.

In Deutschland wird die Rotbuche in der Zukunft aufgrund der Waldumbaustrategien eine noch wichtigere Rolle übernehmen. Reine Nadelwälder sollen in Buchenwälder oder Mischwälder mit Buchenanteil umgewandelt werden. Die Anpassungsfähigkeit der Rotbuche an die sich ändernden klimatischen Bedingungen wird in der Literatur kritisch diskutiert. Zumindest in einigen Regionen muss davon ausgegangen werden, dass die Rotbuche, vor allem Jungpflanzen, unter erhöhtem Trockenstress in den Sommermonaten leiden wird. Daher ist es äußerst wichtig die genetischen Grundlagen von klimawandelrelevanten Merkmalen zu erforschen. Ein Kandidatengenansatz und die Analyse der Variation in Kandidatengenen wurden in dieser Studie verwendet, um den genetischen Hintergrund von Trockenstresstoleranz und Austriebsverhalten zu untersuchen. Rotbuchenpopulationen entlang eines Niederschlagsgradienten in Norddeutschland wurden dafür ausgewählt. Zusätzlich wurde ein kontrolliertes Trockenstressexperiment in einer Klimakammer durchgeführt bei dem stark geschädigte Jungpflanzen mit nicht geschädigten/nur schwach geschädigten Sämlingen verglichen wurden.

Zunächst wurden neun neutrale Mikrosatelliten verwendet um die neutrale genetische Diversität der ausgewählten Populationen zu ermitteln. Anschließend wurden Fragmente von zehn Kandidatengenen auf genomischer Ebene identifiziert, die höchstwahrscheinlich einen Einfluss auf Trockenstresstoleranz und/oder Austriebsverhalten haben. SNPs (Single Nucleotide Polymorphisms) und Indels (Insertion/Deletion) in kodierenden und nicht-kodierenden Bereichen in diesen Kandidatengenen wurden analysiert. Die neutrale genetische Diversität war in allen Populationen hoch (durchschnittliche erwartete

Heterozygotie: 0.617) und ist vergleichbar mit anderen Rotbuchenpopulationen. Insgesamt wurden 8145 bp sequenziert und analysiert, 4038 bp davon in Exon- und 4107 bp in Intronregionen. 63 SNPs und elf Indels wurden stark unterschiedlich verteilt über die verschiedenen Genregionen gefunden. Die Nukleotiddiversität reichte von 0 bis 6,62 ($\pi \times 10^{-3}$) und ist damit vergleichbar mit anderen Baumarten, wohingegen die durchschnittliche Nukleotiddiversität (2,64) vergleichsweise niedrig ist. Für weiterführende Untersuchungen wurden insgesamt 17 SNPs von acht verschiedenen Kandidatengenen ausgewählt. Die Analyse der genetischen Differenzierung anhand von Mikrosatelliten und SNPs zeigte, dass sich alle Populationen signifikant voneinander unterscheiden. Zehn der siebzehn SNPs aus sechs verschiedenen Genen, darunter SNPs in nicht-kodierenden Bereichen, synonyme und nicht-synonyme SNPs, waren signifikant unterschiedlich zwischen der Population mit dem höchsten Jahresniederschlag (Unterlüß) und der Population mit dem niedrigsten Jahresniederschlag (Calvörde). Die Analyse der neutralen genetischen Diversität ergab allerdings, dass die beiden Populationen sich genetisch stark voneinander unterscheiden. Der Vergleich von Unterlüß und Göhrde (mittlerer Jahresniederschlag) zeigte, dass vier der siebzehn SNPs signifikante Unterschiede aufwiesen. Es handelt sich dabei um eine Teilmenge der SNPs, die signifikant unterschiedlich waren bei dem Vergleich der Populationen Unterlüß und Calvörde. Unterlüß und Göhrde weisen jedoch auch starke genetische Unterschiede auf bei der Analyse der neutralen genetischen Diversität. Für diese beiden Populationspaare kann daher nicht ausgeschlossen werden, dass die signifikanten Unterschiede an den SNP-Markern auf unterschiedliche Populationsgeschichten, einschließlich nacheiszeitlicher Rückwanderungsrouten, oder menschlichem Samentransfer zurückzuführen sind. Demgegenüber konnten die neutralen Marker zeigen, dass sich die beiden Populationen Calvörde und Göhrde genetisch sehr ähnlich sind. Wie zu erwarten, waren nur zwei SNPs im Vergleich dieser beiden Populationen signifikant unterschiedlich, ein dritter SNP erreichte fast signifikante Werte. Die beiden SNPs befinden sich in den Genen *Isocitrat-Dehydrogenase* und *Ascorbat Peroxidase* 4 und haben möglicherweise einen Einfluss auf die Trockenstresstoleranz der Rotbuche. Diese beiden SNPs sind besonders interessant für weitere Analysen, aber auch die SNPs, die signifikante Unterschiede im Vergleich der anderen Populationen zeigten, sollten nicht außer Acht gelassen werden. Interessant sind hier vor allem die Gene *Ascorbat Peroxidase* 1 und *Phytochrom B* zu nennen, die hochsignifikante Unterschiede im Vergleich der Populationen aufwiesen. Der Vergleich von geschädigten und nicht geschädigten Jungpflanzen im Rahmen des kontrollierten

Trockenstressexperimentes ergab, dass nur einer der untersuchten SNPs leicht signifikant unterschiedlich zwischen den beiden Gruppen war. Es handelt sich dabei um einen SNP im Gen *Ascorbat Peroxidase* 4. Obwohl dieser SNP bei dem Vergleich der natürlichen Populationen keine signifikanten Unterscheide zeigte, ist er trotzdem interessant für nachfolgende Studien. Aufgrund der geringen Anzahl der untersuchten SNPs, wurde in dieser Studie nur ein Vergleich basierend auf paarweisem F_{ST} durchgeführt. Die Anwendung von Outlier-Analysen war nicht möglich.

Die Mikrosatellitenanalyse in dieser Studie hat die hohe neutrale genetische Diversität der Rotbuche aus früheren Studien bestätigt. Obwohl dies eine gute Grundlage für Anpassung bildet, ist es wichtig den genetischen Hintergrund von klimawandelrelevanten Merkmalen zu untersuchen, um das genetische Anpassungspotential an das sich verändernde Klima einschätzen zu können. Die Anzahl der untersuchten Kandidatengene und SNPs in dieser Studie ist niedrig, vor allem, weil nur wenige Sequenzinformationen für die Rotbuche zur Verfügung standen. Trotzdem war es möglich SNPs zu finden, die statistisch signifikante Unterschiede zwischen den Populationen aufweisen und die daher möglicherweise an der Anpassung an Trockenstress beteiligt sind. Diese Untersuchung ist nur ein erster Schritt für die Erforschung der genetischen Grundlagen von Trockenstresstoleranz bei der nicht-Modellbaumart Rotbuche. In der Zukunft können neue nun verfügbare Sequenziertechniken für weiterführende Analysen verwendet werden. Vor allem RNA-Sequenzierung ist eine hilfreiche Methode um Transkriptome zu studieren, z.B. um trockengestresste Pflanzen mit Kontrollpflanzen zu vergleichen und dadurch neue Kandidatengene zu identifizieren oder in anderen Pflanzen bekannte Kandidatengene zu verifizieren. Des Weiteren können die neuen Sequenziertechniken auf zur Identifizierung und Genotypisierung von SNPs genutzt werden. Die Sequenzierung einer großen Anzahl von ganzen Genomen, wie sie für populationsgenetische Untersuchungen nötig wären, wird jedoch in naher Zukunft nicht möglich sein.

7 References

Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. Evolutionary Applications 1: 95-111

Alba R, Fei ZJ, Payton P, Liu Y, Moore SL, Debbie P, Cohn J, D'Ascenzo M, Gordon JS, Rose JKC, Martin G, Tanksley SD, Bouzayen M, Jahn MM, Giovannoni J (2004) ESTs, cDNA microarrays, and gene expression profiling: tools for dissecting plant physiology and development. Plant Journal 39: 697-714

Ammer C, Albrecht L, Borchert H, Brosinger F, Dittmar C, Elling W, Ewald J, Felbermeier B, von Gilsa H, Huss J, Kenk G, Kölling C, Kohnle U, Meyer P, Mosandl R, Moosmayer HU, Palmer S, Reif A, Rehfuess KE, Stimm B (2005) Future suitability of beech (*Fagus sylvatica* L.) in Central Europe: critical remarks concerning a paper of Rennenberg et al. (2004). Allgemeine Forst- und Jagdzeitung 176: 60-67

Anderson E (2010) Assessing the power of informative subsets of loci for population assignment: standard methods are upwardly biased. Molecular Ecology 10: 701-710

APG III - The Angiosperm Phylogeny Group: Bremer B, Bremer K, Chase MW, Fay MF, Reveal JL, Soltis DE, Soltis PS, Stevens PF, Anderberg AA, Moore MJ, Olmstead RG, Rudall PJ, Sytsma KJ, Tank DC, Wurdack K, Xiang JQY, Zmarzty S (2009) An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. Botanical Journal of the Linnean Society 161: 105-121

Aranda I, Gil L, Pardos JA (2000) Water relations and gas exchange in *Fagus sylvatica* L. and *Quercus petraea* (Mattuschka) Liebl. in a mixed stand at their southern limit of distribution in Europe. Trees - Structure and Function 14: 344-352

Asuka Y, Tomaru N, Nisimura N, Tsumura Y, Yamamoto S (2004) Heterogeneous genetic structure in a *Fagus crenata* population in an old-growth beech forest revealed by microsatellite markers. Molecular Ecology 13: 1241-1250

Baek JM, Han P, Iandolino A, Cook DR (2008) Characterization and comparison of intron structure and alternative splicing between *Medicago truncatula*, *Populus trichocarpa*, *Arabidopsis* and rice. Plant Molecular Biology 67: 499-510

Beck EH, Fettig S, Knake C, Hartig K, Bhattarai T (2007) Specific and unspecific responses of plants to cold and drought stress. Journal of Biosciences 32: 501-510

Belmonte J, Alarcon M, Avila A, Scialabba E, Pino D (2008) Long-range transport of beech (*Fagus sylvatica* L.) pollen to Catalonia (north-eastern Spain). International Journal of Biometeorology 52: 675-687

Beniston M, Stephenson DB, Christensen OB, Ferro CAT, Frei C, Goyette S, Halsnaes K, Holt T, Jylha K, Koffi B, Palutikof J, Schoell R, Semmler T, Woth K (2007) Future extreme events in European climate: an exploration of regional climate model projections. Climatic Change 81: 71-95

Betsch P, Bonal D, Breda N, Montpied P, Peiffer M, Tuzet A, Granier A (2011) Drought effects on water relations in beech: the contribution of exchangeable water reservoirs. Agricultural and Forest Meteorology 151: 531-543

Blears MJ, De Grandis SA, Lee H, Trevors JT (1998) Amplified fragment length polymorphism (AFLP): a review of the procedure and its applications. Journal of Industrial Microbiology & Biotechnology 21: 99-114

Blenkinsop S and Fowler HJ (2007) Changes in European drought characteristics projected by the PRUDENCE regional climate models. International Journal of Climatology 27: 1595-1610

Boggs JZ, Loewy K, Bibee K, Heschel MS (2010) Phytochromes influence stomatal conductance plasticity in *Arabidopsis thaliana*. Plant Growth Regulation 60: 77-81

Bolte A (2005) Zur Zukunft der Buche in Mitteleuropa. AFZ - Der Wald 60: 1077-1078

Bolte A and Villanueva I (2006) Interspecific competition impacts on the morphology and distribution of fine roots in European beech (*Fagus sylvatica* L.) and Norway spruce (*Picea abies* (L.) Karst.). European Journal of Forest Research 125: 15-26

Bolte A, Czajkowski T, Kompa T (2007) The north-eastern distribution range of European beech - a review. Forestry 80: 413-429

Bolte A, Ibisch PL, Menzel A, Rothe A (2008): Anpassung der Wälder an den Klimawandel. Was Klimahüllen uns verschweigen. AFZ - Der Wald 15: 800-803

References

Bonin A, Ehrich D, Manel S (2007) Statistical analysis of amplified fragment length polymorphism data: a toolbox for molecular ecologists and evolutionists. Molecular Ecology 16: 3737-3758

Bouk A and Vision T (2007) The molecular ecologist's guide to expressed sequence tags. Molecular Ecology 16: 907-924

Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES (2007) TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23: 2633-2635

Breathnach R, Benoist C, O'Hare K, Gannon F, Chambon P (1978) Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. Proceedings of the National Academy of Sciences of the United States of America 75: 4853-4857

Bréda N, Huc R, Granier A, Dreyer E (2006) Temperate forest trees and stands under severe drought: a review of ecophysiological responses, adaptation processes and long-term consequences. Annals of Forest Science 63: 625-644

Brookes A (1999) The essence of SNPs. Gene 234: 177-186

Brown GR, Bassoni DL, Gill GP, Fontana JR, Wheeler NC, Megraw RA, Davis MF, Sewell MM, Tuskan GA, Neale DB (2003) Identification of quantitative trait loci influencing wood property traits in loblolly pine (*Pinus taeda* L.). III. QTL verification and candidate gene mapping. Genetics 164: 1537-1546

Brown GR, Gill GP, Kuntz RJ, Langley CH, Neale DB (2004) Nucleotide diversity and linkage disequilibrium in loblolly pine. Proceedings of the National Academy of Sciences of the United States of America 101: 15255-15260

Bruschi P, Vendramin GG, Bussotti F, Grossoni P (2003) Morphological and molecular diversity among Italian populations of *Quercus petraea* (Fagaceae). Annals of Botany 91: 707-716

Buiteveld J, Vendramin GG, Leonardi S, Kamer K, Geburek T (2007) Genetic diversity and differentiation in European beech (*Fagus sylvatica* L.) stands varying in management history. Forest Ecology and Management 247: 98-106

Bussell JD, Waycott M, Chappill JA (2005) Arbitrarily amplified DNA markers as characters for phylogenetic inference. Perspectives in Plant Ecology, Evolution and Systematics 7: 3-26

Bustin SA (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. Journal of Molecular Endocrinology 25: 169-193

Campbell D, Duchesne P, Bernatchez L (2003) AFLP utility for population assignment studies: analytical investigation and empirical comparison with microsatellites. Molecular Ecology 12: 1979-1991

Campbell NR and Narum S (2011) Development of 54 novel single-nucleotide polymorphism (SNP) assays for sockeye and coho salmon and assessment of available SNPs to differentiate stocks within the Columbia River. Molecular Ecology Resources 11: 20-30

Charru M, Seynave I, Morneau F, Bontemps JD (2010) Recent changes in forest productivity: an analysis of national forest inventory data for common beech (*Fagus sylvatica* L.) in north-eastern France. Forest Ecology and Management 260: 864-874

Churkina G, Schimel D, Braswell BH, Xiao XM (2005) Spatial analysis of growing season length control over net ecosystem exchange. Global Change Biology 11: 1777-1787

Clark AG, Hubisz MJ, Bustamante CD, Williamson SH, Nielsen R (2005) Ascertainment bias in studies of human genome-wide polymorphism. Genome Research 15: 1496-1502

Czajkowski T, Kühling M, Bolte A (2005) Impact of the 2003 summer drought on growth of beech sapling natural regeneration (*Fagus sylvatica* L.) in north-eastern Central Europe. Allgemeine Forst- und Jagdzeitung 176: 133-143

Czajkowski T and Bolte A (2005) Unterschiedliche Reaktion deutscher und polnischer Herkünfte der Buche (*Fagus sylvatica* L.) auf Trockenheit. Allgemeine Forst- und Jagdzeitung 177: 30-40

Czúcz B, Gálhidy L, Mátyás C (2011) Present and forecasted xeric climatic limits of beech and sessile oak distribution at low altitudes in Central Europe. Annals of Forest Science 68: 99-108

De Carvalho D, Ingvarsson PK, Joseph J, Suter L, Sedivy C, Macaya-Sanz D, Cottrell J, Heinze B, Schanzer I, Lexer C (2010) Admixture facilitates adaptation from standing variation in the

European aspen (*Populus tremula* L.), a widespread forest tree. Molecular Ecology 19: 1638-1650

DeHayes DH, Jacobson GL Jr, Schaberg PG, Bongarten B, Iverson L, Dieffenbacher-Krall AC (2000) Forest responses to changing climate: lessons from the past and uncertainty for the future. In: Mickler RA, Birdsey RA, Hom J (eds.) Responses of northern U.S. forests to environmental change. Ecological Studies 139. Springer-Verlag, New York, pp 495-540

Denk T (2003) Phylogeny of *Fagus* L. (Fagaceae) based on morphological data. Plant Systematics and Evolution 240: 55-81

Derory J, Scotti-Saintagne C, Bertocchi E, Le Dantec L, Graignic N, Jauffres A, Casasoli M, Chancerel E, Bodenes C, Alberto F, Kremer A (2010) Contrasting relationships between the diversity of candidate genes and variation of bud burst in natural and segregating populations of European oaks. Heredity 104: 438-448

Dittmar C, Fricke W, Elling W (2006) Impact of late frost events on radial growth of common beech (*Fagus sylvatica* L.) in Southern Germany. European Journal of Forest Research 125: 249-259

Don A, Arenhoevel W, Jacob R, Scherer-Lorenzen M, Schulze ED (2007) Establishment success of 19 different tree species on afforestations - results of a biodiversity experiment. Allgemeine Forst- und Jagdzeitung 178: 164-172

Dostálek J, Frantík T, Lukášová M (2001) Genetic differences within natural and planted stands of *Quercus petraea*. Central European Journal of Biology 6: 597-605

Dounavi A, Koutsias N, Ziehe M, Hattemer HH (2010) Spatial patterns and genetic structures within beech populations (*Fagus sylvatica* L.) of forked and non-forked individuals. European Journal of Forest Research 129: 1191-1202

Durand J, Bodenes C, Chancerel E, Frigerio JM, Vendramin G, Sebastiani F, Buonamici A, Gailing O, Koelewijn HP, Villani F, Mattioni C, Cherubini M, Goicoechea PG, Herran A, Ikaran Z, Cabane C, Ueno S, Alberto F, Dumoulin PY, Guichoux E, de Daruvar A, Kremer A, Plomion C (2010) A fast and cost-effective approach to develop and map EST-SSR markers: oak as a case study. BMC Genomics 11: No 570

Eckert AJ, Bower AD, González-Martínez SC, Wegrzyn JL, Coop G, Neale DB (2010) Back to nature: ecological genomics of loblolly pine (*Pinus taeda*, Pinaceae). Molecular Ecology 19: 3789-3805

EEA - European Environmental Agency (2008) Impacts of Europe's changing climate - 2008 indicator-based assessment. Joint EEA-JRC-WHO report. http://www.eea.europa.eu

Ellegren H (2004) Microsatellites: simple sequences with complex evolution. Nature Reviews Genetics 5: 435-445

Ellenberg H (1988) Vegetation ecology of Central Europe. 4th edn. Cambridge University Press, Cambridge

Ellenberg H and Leuschner C (2010) Vegetation Mitteleuropas mit den Alpen. Ulmer Verlag, Stuttgart

Engesser R, Forster B, Meier F, Wermelinger B (2008) Forstliche Schadorganismen im Zeichen des Klimawandels. Schweizerische Zeitschrift für Forstwesen 159: 344-351

Enke W, Deutschländer T, Schneider F, Küchler W (2005) Results of five regional climate studies applying a weather pattern based downscaling method to ECHAM4 climate simulations. Meteorologische Zeitschrift 14: 247-257

Estrella N and Menzel A (2006) Responses of leaf colouring in four deciduous tree species to climate and weather in Germany. Climate Research 32: 253-267

Eveno E, Collada C, Guevara MA, Leger V, Soto A, Diaz L, Leger P, González-Martínez SC, Cervera MT, Plomion C, Garnier-Gere PH (2008) Contrasting patterns of selection at *Pinus pinaster* Ait. drought stress candidate genes as revealed by genetic differentiation analyses. Molecular Biology and Evolution 25: 17-437

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. Evolutionary Bioinformatics 1: 47-50

Excoffier L, Hofer T, Foll M (2009) Detecting loci under selection in a hierarchically structured population. Heredity 103: 285-298

References

Flint-Garcia SA, Thornsberry JM, Buckler ES (2003) Structure of linkage disequilibrium in plants. Annual Review of Plant Biology 54: 357-374

Flintoft L (2008) Transcriptomics: digging deep with RNA-Seq. Nature Review Genetics 9: 413

Fotelli MN, Geßler A, Peuke AD, Rennenberg H (2001) Drought affects the competitive interactions between *Fagus sylvatica* seedlings and an early successional species, *Rubus fruticosus*: responses of growth, water status and delta C-13 composition. New Phytologist 151: 427-435

Fotelli MN, Nahm M, Radoglou K, Rennenberg H, Halyvopoulos G, Matzarakis A (2009) Seasonal and interannual ecophysiological responses of beech (*Fagus sylvatica*) at its south-eastern distribution limit in Europe. Forest Ecology and Management 257: 1157-1164

Frewen BE, Chen THH, Howe GT, Davis J, Rohde A, Boerjan W, Bradshaw HD (2000) Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. Genetics 154: 837-845

Friedrichs DA, Büntgen U, Frank DC, Esper J, Neuwirth B, Löffler J (2009) Complex climate controls on 20th century oak growth in Central-West Germany. Tree Physiology 29: 39-51

Fritz P (2006) Ökologischer Waldumbau in Deutschland: Fragen, Antworten, Perspektiven. Oekom Verlag, München

Fuhrer J, Beniston M, Fischlin A, Frei C, Goyette S, Jasper K, Pfister C (2006) Climate risks and their impact on agriculture and forests in Switzerland. Climatic Change 79: 79-1002

Gärtner S, Reif A, Xystrakis F, Sayer U, Bendagha N, Matzarakis A (2008) The drought tolerance limit of *Fagus sylvatica* forest on limestone in southwestern Germany. Journal of Vegetation Science 19: 757-768

Gailing O, Vornam B, Leinemann L, Finkeldey R (2009) Genetic and genomic approaches to assess adaptive genetic variation in plants: forest trees as a model. Physiologia Plantarum 137: 509-519

Gailing O and von Wühlisch G (2004) Nuclear markers (AFLPs) and chloroplast microsatellites differ between *Fagus sylvatica* and *F. orientalis*. Silvae Genetica 53: 105-110

Gao CX and Han B (2009) Evolutionary and expression study of the aldehyde dehydrogenase (ALDH) gene superfamily in rice (*Oryza sativa*). Gene 431: 86-94

García-Plazaola JI and Becerril JM (2000) Effects of drought on photoprotective mechanisms in European beech (*Fagus sylvatica* L.) seedlings from different provenances. Trees - Structure and Function 14: 485-490

Garvin MR, Saitoh K, Gharrett AJ (2010) Application of single nucleotide polymorphisms to non-model species: a technical review. Molecular Ecology Resources 10: 915-934

Gautam DR (2010) Effects of competition during development of European beech (*Fagus sylvatica* L.) regeneration on genetic structures. M.Sc. thesis. Georg-August-University Göttingen, Germany

Gebremedhin B, Ficetola GF, Naderi S, Rezaei HR, Maudet C, Rioux D, Luikart G, Flagstad O, Thuiller W, Taberlet P (2009) Frontiers in identifying conservation units: from neutral markers to adaptive genetic variation. Animal Conservation 12: 107-109

Geraldes A, Pang J, Thiessen N, Cezard T, Moore R, Zhao Y, Tam A, Wang S, Friedmann M, Birol I, Jones SJM, Cronk QCB, Douglas CJ (2011) SNP discovery in black cottonwood (*Populus trichocarpa*) by population transcriptome resequencing. Molecular Ecology Resources 11: 81-92

Gerlach G, Jueterbock A, Kraemer P, Deppermann J, Harmand P (2010) Calculations of population differentiation based on G_{ST} and D: forget G_{ST} but not all of statistics! Molecular ecology 19: 3845-3852

Geßler A, Keitel C, Nahm M, Rennenberg H (2004) Water shortage affects the water and nitrogen balance in central European beech forests. Plant Biology 6: 289-298

Geßler A, Keitel C, Kreuzwieser J, Matyssek R, Seiler W, Rennenberg H (2007) Potential risks for European beech (*Fagus sylvatica* L.) in a changing climate. Trees - Structure and Function 21: 1-11

Gieger T and Thomas FM (2002) Effects of defoliation and drought stress on biomass partitioning and water relations of *Quercus robur* and *Quercus petraea*. Basic and Applied Ecology 3: 171-181

References

Gilchrist AS and Meats AW (2010) The genetic structure of populations of an invading pest fruit fly, *Bactrocera tryoni*, at the species climatic range limit. Heredity 105: 165-172

Ginzinger DG (2002) Gene quantification using real-time quantitative PCR: An emerging technology hits the mainstream. Experimental Hematology 30: 503-512

Glenn TC (2011) Field guide to next-generation DNA sequencers. Molecular Ecology Resources 11: 759-769

Gömöry D, Paule L, Shvadchak IM, Popescu F, Sulkowska M, Hynek V, Longauer R (2003) Spatial patterns of the genetic differentiation in European beech (*Fagus sylvatica* L.) at allozyme loci in the Carpathians and the adjacent regions. Silvae Genetica 52: 78-83

González-Martínez SC, Ersoz E, Brown GR, Wheeler NC, Neale DB (2006a) DNA sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in *Pinus taeda* L.. Genetics 172: 1915-1926

González-Martínez SC, Krutovsky KV, Neale DB (2006b) Forest-tree population genomics and adaptive evolution. New Phytologist 170: 227-238

Gora V, Starke R, Ziehe M, König J, Müller-Starck G, Lunderstädt J (1994) Influence on genetic structures and silvicultural treatments in a beech stand (*Fagus sylvatica*) on the population dynamics of beech scale (*Cryptococcus fagisuga*). Forest Genetics 1: 157-164

Govaerts R and Frodin DG (1998) World checklist and bibliography of Fagales. Royal Botanical Gardens, Kew

Granier A, Reichstein M, Bréda N, Janssens IA, Falge E, Ciais P, Grünwald T, Aubinet M, Berbigier P, Bernhofer C, Buchmann N, Facini O, Grassi G, Heinesch B, Ilvesniemi H, Keronen P, Knohl A, Köstner B, Lagergren F, Lindroth A, Longdoz B, Loustau D, Mateus J, Montagnani L, Nys C, Moors E, Papale D, Peiffer M, Pilegaard K, Pita G, Pumpanen J, Rambal S, Rebmann C, Rodrigues A, Seufert G, Tenhunen J, Vesala I, Wang Q (2007) Evidence for soil water control on carbon and water dynamics in European forests during the extremely dry year: 2003. Agricultural and Forest Meteorology 143: 123-145

Guo P, Baum M, Grando S, Ceccarelli S, Bai G, Li R, von Korff M, Varshney RK, Graner A, Valkoun J (2009) Differentially expressed genes between drought-tolerant and drought-

78

sensitive barley genotypes in response to drought stress during the reproductive stage. Journal of Experimental Botany 60: 3531-3544

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symposium Series 41: 95-98

Hall D, Luquez V, Garcia VM, St Onge KR, Jansson S, Ingvarsson PK (2007) Adaptive population differentiation in phenology across a latitudinal gradient in European Aspen (*Populus tremula*, L.): a comparison of neutral markers, candidate genes and phenotypic traits. Evolution 61: 2849-2860

Hamrick JL (2004) Response of forest trees to global environmental changes. Forest Ecology and Management 197: 323-335

Hasenkamp N, Ziegenhagen B, Mengel C, Schulze L, Schmitt HP, Liepelt S (2011) Towards a DNA marker assisted seed source identification: a pilot study in European beech (*Fagus sylvatica* L.). European Journal of Forest Research 130: 513-519

Hazler K, Comps B, Šugar I, Melovski L, Tashev A, Gračan J (1997) Genetic structure of *Fagus sylvatica* L. populations in southeastern Europe. Silvae Genetica 46: 229-236

Hazler-Pilepic K, Comps B, Sugar I, Solic E (1999) Isozyme polymorphism of Croatian beech populations (*Fagus sylvatica* L.). Periodicum Biologorum 101: 165-170

Hedrick PW (2005) A standardized genetic differentiation measure. Evolution 59: 1633-1638

Helyar SJ, Hemmer-Hansen J, Bekkevold D, Taylor MI, Ogden R, Limborg MT, Cariani A, Maes GE, Diopere E, Carvalho GR, Nielsen EE (2011) Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Molecular Ecology Resources 11: 123-136

Heuertz M, De Paoli E, Kallman T, Larsson H, Jurman I, Morgante M, Lascoux M, Gyllenstrand N (2006) Multilocus patterns of nucleotide diversity, linkage disequilibrium and demographic history of Norway spruce [*Picea abies* (L.) Karst]. Genetics 174: 2095-2105

Hudson M (2008) Sequencing breakthroughs for genomic ecology and evolutionary biology. Molecular Ecology Resources 8: 3-17

References

Hussendörfer E and Konnert M (2000) Untersuchungen zur Bewirtschaftung von Weisstannen- und Buchenbeständen unter dem Aspekt der Erhaltung genetischer Variation. Forest, Snow and Landscape Research 75: 187-204

Ingvarsson PK (2005) Nucleotide polymorphism and linkage disequilbrium within and among natural populations of European Aspen (*Populus tremula* L., Salicaceae). Genetics 169: 945-953

Ingvarsson PK, Garcia MV, Hall D, Luquez V, Jansson S (2006) Clinal variation in *phyB*2, a candidate gene for day-length-induced growth cessation and bud set, across a latitudinal gradient in European aspen (*Populus tremula*). Genetics 172: 1845-1853

Ingvarsson PK, Garcia MV, Luquez V, Hall D, Jansson S (2008) Nucleotide polymorphism and phenotypic associations within and around the *phytochrome B2* locus in European aspen (*Populus tremula*, Salicaceae). Genetics 178: 2217-2226

IPCC - Intergovernmental Panel on Climate Change (2000) Special reports: emission scenarios, edited by Nakicenovic N and Swart R, Cambridge University Press, UK

IPCC - Intergovernmental Panel on Climate Change (2007) Climate Change 2007: the physical science basis, contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change, edited by Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, Miller HL, Cambridge University Press, Cambridge

Jacob D, Göttel H, Kotlarski S, Lorenz P, Sieck K (2008) Forschungsbericht: Klimaauswirkungen und Anpassung in Deutschland - Phase 1: Erstellung regionaler Klimaszenarien für Deutschland. In: Umweltbundesamt (ed.), Reihe Climate Change 11/08, Dessau, http://www.umweltbundesamt.de/

Jarcuska B (2009) Growth, survival, density, biomass partitioning and morphological adaptations of natural regeneration in *Fagus sylvatica*. Dendrobiology 61: 3-11

Jimenez JA, Alonso-Ramerez A, Nicolas C (2008) Two cDNA clones (FsDhn1 and FsClo1) up-regulated by ABA are involved in drought responses in *Fagus sylvatica* L. seeds. Journal of Plant Physiology 165: 1798-1807

Jost L (2008) G_{ST} and its relatives do not measure differentiation. Molecular Ecology 17: 4015-4026

Jost L (2009) D vs. G_{ST}: Response to Heller and Siegismund (2009) and Ryman and Leimar (2009). Molecular Ecology 18: 2088-2091

Jump AS, Hunt JM, Martínez-Izquierdo JA, Peñuelas J (2006a) Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in *Fagus sylvatica*. Molecular Ecology 15: 3469-3480

Jump AS, Hunt JM, Peñuelas J (2006b) Rapid climate change-related growth decline at the southern range edge of *Fagus sylvatica*. Global Change Biology 12: 2163-2174

Jump AS and Peñuelas J (2007) Extensive spatial genetic structure revealed by AFLP but not SSR molecular markers in the wind-pollinated tree, *Fagus sylvatica*. Molecular Ecology 16: 925-936

Jung T (2009) Beech decline in Central Europe driven by the interaction between *Phytophthora* infections and climatic extremes. Forest Pathology 39: 73-94

Kalia RK, Rai MK, Kalia S, Singh R, Dhawan AK (2011) Microsatellite markers: an overview of the recent progress in plants. Euphytica 177: 309-334

K.C. R (2011) Spatial dynamics of intraspecific genetic variation in European beech (*Fagus sylvatica* L.). Ph.D. Dissertation. Georg-August-University Göttingen

Keller SR, Levsen N, Ingvarsson PK, Olson MS, Tiffin P (2011) Local selection across a latitudinal gradient shapes nucleotide diversity in balsam poplar, *Populus balsamifera* L.. Genetics 188: 941-U318

Kim R and Guo JT (2010) Systematic analysis of short internal indels and their impact on protein folding. BMC Structural Biology 10: 24

Kirk H and Freeland JR (2011) Applications and implications of neutral versus non-neutral markers in molecular ecology. International Journal of Molecular Sciences 12: 3966-3988

References

Knapp HD, Emde FA, Engels B, Lehrke S, Hendrischke O, Klein M, Kluttig H, Krug A, Schäfer HJ, Scherfose V, Schröder E, Schweppe-Kraft B (2008) Naturerbe Buchenwälder - Situationsanalyse und Handlungserfordernisse. Bundesamt für Naturschutz, Bonn, Vilm

Kölling C, Walentowski H, Borchert H (2005) Die Buche in Mitteleuropa - Eine Waldbaumart mit grandioser Vergangenheit und sicherer Zukunft. AFZ - Der Wald 13: 696-701

Kölling C (2007) Klimahüllen für 27 Waldbaumarten. AFZ - Der Wald 23: 1242-1245

Körner C, Asshoff R, Bignucolo O, Hättenschwiler S, Keel SG, Peláez-Riedl S, Pepin, S, Siegwolf RTW, Zotz G (2005) Carbon flux and growth in mature deciduous forest trees exposed to elevated CO_2. Science 309: 1360-1362

Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M (2006) An SNP caused loss of seed shattering during rice domestication. Science 312: 1392–1396

Konnert M (1995) Investigations on the genetic variation of beech (*Fagus sylvatica* L.) in Bavaria. Silvae Genetica 44: 346-351

Konnert M and Henkel W (1997) Investigations on the genetic variation of beech (*Fagus sylvatica* L.) in Thuringia. Allgemeine Forst- und Jagdzeitung 168: 182-190

Kozovits AR, Matyssek R, Winkler JB, Göttlein A, Blaschke H, Grams TEE (2005) Above-ground space sequestration determines competitive success in juvenile beech and spruce trees. New Phytologist 167: 181-196

Kumar P, Gupta V K, Misra AK, Modi DR, Pandey BK (2009) Potential of molecular markers in plant biotechnology. Plant Omics 2: 141-162

Kundzewicz ZW, Radziejewski M, Pinskwar I (2006) Precipitation extremes in the changing climate of Europe. Climate Research 31: 51-58

Kunstler G, Curt T, Bouchaud M, Lepart J (2005) Growth, mortality, and morphological response of European beech and downy oak along a light gradient in sub-Mediterranean forest. Canadian Journal of Forest Research 35: 1657-1668

Kuparinen A, Savolainen O, Schurr FM (2010) Increased mortality can promote evolutionary adaptation of forest trees to climate change. Forest Ecology and Management 259: 1003-1008

Kraj W and Sztorc A (2009) Genetic structure and variability of phenological forms in the European beech (*Fagus sylvatica* L.). Annals of Forest Science 66: No 203

Kramer K, Buiteveld J, Forstreuter M, Geburek T, Leonardi S, Menozzi P, Povillon F, Schelhaas Mj, du Cros ET, Vendramin GG, van der Werf DC (2008) Bridging the gap between ecophysiological and genetic knowledge to assess the adaptive potential of European beech. Ecological Modelling 216: 333-353

Kreienkamp F, Spekat A, Enke W (2010) Ergebnisse eines regionalen Szenarienlaufs für Deutschland mit dem statistischen Modell WETTREG2010. Technical Report, Climate and Environment Consulting Potsdam GmbH on a contract of the Federal Environment Agency (UBA), Potsdam, Dessau

Kremer A, Casasoli M, Barreneche T, Bodénès C, Sisco P, Kubisiak T, Scalfi M, Leonardi S, Bakker EG, Buiteveld J, Romero-SeversoJ, Arumugnianathan K, Derory J, Scotti-Saintagne C, Roussel G, Bertocchi ME, Lexer C, Porth I, Hebard F, Clark C, Carlson J, Plomion C, Koelewijn HP, Villani F (2007) Comparative genetic mapping in Fagaceae. In: Kole C (ed.) Genome mapping & molecular breeding in plants Vol. 7 - Forest trees. Springer, Heidelberg, Berlin, New York, Tokyo, pp 161-187

Krutovsky KV and Neale DB (2005) Nucleotide diversity and linkage disequilibrium in cold-hardiness- and wood quality-related candidate genes in Douglas fir. Genetics 171: 2029-2041

Lalagüe H, Fady B, Garnier-Géré P, González-Martínez SC, Lin YC, Oddou-Muratorio S, Sebastiani F, Vendramin GG (2010) Candidate gene variation in common beech (*Fagus sylvatica* L.) along an altitudinal gradient. In: Vinceti B and Neate P (comps.) Conference on "Forest Ecosystem Genomics and Adaptation". San Lorenzo de El Escorial (Madrid), Spain, 9 - 11 June 2010. Book of abstracts. Bioversity International (Rome, Italy) and INIA (Madrid, Spain), pp 242

Larsen AB (1996) Genetic structure of populations of beech (*Fagus sylvatica* L) in Denmark. Scandinavian Journal of Forest Research 11: 220-232

References

Lendzion J and Leuschner C (2008) Growth of European beech (*Fagus sylvatica* L.) saplings is limited by elevated atmospheric vapour pressure deficits. Forest Ecology and Management 256: 648-655

Leng L and Zhang DX (2011) Measuring population differentiation using G_{ST} or D? A simulation study with microsatellite DNA markers under a finite island model and nonequilibrium conditions. Molecular Ecology 20: 2494-2509

Leuschner C, Backes K, Hertel D, Schipka F, Schmitt U, Terborg O, Runge M (2001a) Drought responses at leaf, stem and fine root levels of competitive *Fagus sylvatica* L. and *Quercus petraea* (Matt.) Liebl. trees in dry and wet years. Forest Ecology and Management 149: 33-46

Leuschner C, Hertel D, Coners H, Buttner V (2001b) Root competition between beech and oak: a hypothesis. Oecologica 126: 276-284

Leuschner C, Meier IC, Hertel D (2006) On the niche breadth of *Fagus sylvatica*: soil nutrient status in 50 Central European beech stands on a broad range of bedrock types. Annals of Forest Science 63: 355-368

Leuzinger S, Zotz G, Asshoff R, Körner C (2005) Responses of deciduous forest trees to severe drought in Central Europe. Tree Physiology 25: 641-650

Li YC, Korol AB, Fahima T, Beiles A, Nevo E (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Molecular Ecology 11: 2453-2465

Librado P and Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452

Linderholm HW (2006) Growing season changes in the last century. Agricultural and Forest Meteorology 137: 1-14

Lindner M, Maroschek M, Netherer S, Kremer A, Barbati A, Garcia-Gonzalo J, Seidl R, Delzon S, Corona P, Kolström M, Lexer MJ, Marchetti M (2010) Climate change impacts, adaptive capacity, and vulnerability of European forest ecosystems. Forest Ecology and Management 259: 698-709

Liu YH, Shi YS, Song YC, Wang TY, Li Y (2010) Characterization of a stress-induced NADP-isocitrate dehydrogenase gene in maize confers salt tolerance in *Arabidopsis*. Journal of Plant Biology 53: 107-112

Löw M, Herbinger K, Nunn AJ, Haeberle KH, Leuchner M, Heerdt C, Werner H, Wipfler P, Pretzsch H, Tausz M, Matyssek R (2006) Extraordinary drought of 2003 overrules ozone impact on adult beech trees (*Fagus sylvatica*). Trees - Structure and Function 20: 539-548

Longauer R, Gömöry D, Paule L, Karnosky DF, Maňkovská B, Müller-Starck G, Percy K, Szaro R (2001) Selection effects of air pollution on gene pools of Norway spruce, European silver fir and European beech. Environmental Pollution 115: 405-411

Lu Y, Zhang S, Shah T, Xie C, Hao Z, Li X, Farkhari M, Ribaut JM, Cao M, Rong T, Xu Y (2010) Joint linkage–linkage disequilibrium mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize. Proceedings of the National Academy of Sciences of the United States of America 107: 19585-19590

Lunderstädt J (2002) Long term research on the infestation dynamics of beech scale (*Cryptococcus fagisuga* LIND.) and on the formation of necroses in a mixed stand of beech and valuable broad leaved trees. Allgemeine Forst- und Jagdzeitung 173: 193-200

Luo M, Liu J, Lee D, Scully BT, Guo B (2010) Monitoring the expression of maize genes in developing kernels under drought stress using oligo-microarray. Journal of Integrative Plant Biology 52: 1059-1074

Marth GT, Czabarka E, Murvai J, Sherry ST (2004) The allele frequency spectrum in genome-wide human variation data reveals signals of differential demographic history in three large world populations. Genetics 166: 351-372

Martínez-Gómez P, Crisosto CH, Bonghi C, Rubio M (2011) New approaches to *Prunus* transcriptome analysis. Genetica 139: 755-769

Meehl GA and Tebaldi C (2004) More intense, more frequent, and longer lasting heat waves in the 21st century. Science 305: 994-997

References

Meier IC and Leuschner C (2008) Belowground drought response of European beech: fine root biomass and carbon partitioning in 14 mature stands across a precipitation gradient. Global Change Biology 14: 2081-2095

Meirmans PG and Hedrick PW (2011) Assessing population structure: F_{ST} and related measures. Molecular Ecology Resources 11: 5-18

Menzel A and Fabian P (1999) Growing season extended in Europe. Nature 397: 659-659

Menzel A (2000) Trends in phenological phases in Europe between 1951 and 1996. International Journal of Biometeorology 44: 76-81

Merzeau D, Comps B, Thiébaut B, Letouzey J (1994) Estimation of *Fagus sylvatica* L. mating system parameters in natural populations. Annals of Forest Science 51: 163-173

Meudt HM and Clarke AC (2007) Almost forgotten or latest practice? AFLP applications, analyses and advances. Trends in Plant Science 12: 106-117

Milad M, Schaich H, Bürgi M, Konold W (2011) Climate change and nature conservation in Central European forests: a review of consequences, concepts and challenges. Forest Ecology and Management 261: 829-843

Modrý M, Hubený D, Rejšek K (2004) Differential response of naturally regenerated European shade tolerant tree species to soil type and light availability. Forest Ecology and Management 188: 185-195

Moen T, Hayes B, Nilsen F, Delghandi M, Fjalestad KT, Fevolden SE, Berg PR, Lien S (2008) Identification and characterisation of novel SNP markers in Atlantic cod: evidence for directional selection. BMC Genetics 9: No 18

Möhring B, Leefken G, Gutsche C (2008) Economic valuation of beech forests. In: Nordwestdeutsche Forstliche Versuchsanstalt (ed.) Beiträge aus der Nordwestdeutschen Forstlichen Versuchsanstalt Band 3 - Ergebnisse angewandter Forschung zur Buche. Universitätsdrucke Göttingen, Göttingen

Morin PA, Luikart G, Wayne RK (2004) SNPs in ecology, evolution and conservation. Trends in Ecology & Evolution 19: 208-216

86

Müller-Starck G and Ziehe M (1991) Genetic variation in populations of *Fagus sylvatica* L., *Quercus robur* L., and *Q. petraea* Liebl. in Germany. In: Müller-Starck G and Ziehe M (eds.) Genetic variation in European populations of forest trees. J. D. Sauerländer's Verlag, Frankfurt am Main

Müller-Starck G and Starke R (1993) Inheritance of isoenzymes in European beech (*Fagus sylvatica* L.). Journal of Heredity 84: 291-296

Muleo R, Morini S, Casano S (2001) Photoregulation of growth and branching of plum shoots: Physiological action of two photosystems. In Vitro Cellular & Developmental Biology - Plant 37: 609-617

Muir G, Lowe AJ, Fleming CC, Vogl C (2004) High nuclear genetic diversity, high levels of outcrossing and low differentiation among remnant populations of *Quercus petraea* at the margin of its range in Ireland. Annals of Botany 93: 691-697

Myles S, Peiffer J, Brown PJ, Ersoz ES, Zhang Z, Costich DE, Buckler ES (2009) Association mapping: critical considerations shift from genotyping to experimental design. Plant Cell 21: 2194-2202

Myneni RB, Keeling CD, Tucker CJ, Asrar G, Nemani RR (1997) Increased plant growth in the northern high latitudes from 1981 to 1991. Nature 386: 698-702

Namroud MC, Beaulieu J, Junge N, Laroche J, Bousquet J (2008) Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. Molecular Ecology 17: 3599-3613

Neale DB and Kremer A (2011) Forest tree genomics: growing resources and applications. Nature Review Genetics 12: 111-122

Neale DB and Savolainen O (2004) Association genetics of complex traits in conifers. Trends in Plant Science 9: 325–330

Neale DB (2007) Genomics to tree breeding and forest health. Current Opinion in Genetics & Development 17: 539-544

Nei M (1972) Genetic distance between populations. American Naturalist 106: 283-392

References

Nei M (1978) Estimation of average heterozygosity and genetic distance from a number of individuals. Genetics 89: 538-590

Niedringhaus TP, Milanova D, Kerby MB, Snyder MP, Barron AE (2011) Landscape of next-generation sequencing technologies. Analytical Chemistry 83: 4327-4341

Nielsen R (2004) Population genetic analysis of ascertained SNP data. Human Genomics 1: 218-224

Nielsen R, Hubisz MJ, Clark AG (2004) Reconstituting the frequency spectrum of ascertained single-nucleotide polymorphism data. Genetics 168: 2373-2382

Nolan T, Hands RE, Bustin SA (2006) Quantification of mRNA using real-time RT-PCR. Nature Protocols 1: 1559-1582

Nordborg M, Borevitz JO, Bergelson J, Berry CC, Chory J, Hagenblad J, Kreitman M, Maloof JN, Noyes T, Oefner PJ, Stahl EA, Weigel D (2002) The extent of linkage disequilibrium in *Arabidopsis thaliana*. Nature Genetics 30: 190-193

Nordborg M and Weigel D (2008) Next-generation genetics in plants. Nature 456: 720-723

Nosil P, Funk D J, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. Molecular Ecology 18: 375-402

Nowakowska JA and Oszako T (2008) Health condition and genetic differentiation level of beech in the Siewierz Forest District assessed with cpDNA markers. Sylwan 152: 11-20

Nyári L (2010) Genetic diversity, differentiation and spatial genetic structures in differently managed adult European beech (*Fagus sylvatica* L.) stands and their regeneration. Forstarchiv 81: 156-164

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Molecular Ecology 13: 1143-1155

Oddou-Muratorio S, Bontemps A, Klein EK, Chybicki I, Vendramin GG, Suyama Y (2010) Comparison of direct and indirect genetic methods for estimating seed and pollen dispersal in *Fagus sylvatica* and *Fagus crenata*. Forest Ecology and Management 259: 2151-2159

Oddou-Muratorio S, Klein EK, Vendramin GG, Fady B (2011) Spatial vs. temporal effects on demographic and genetic structures: the roles of dispersal, masting and differential mortality on patterns of recruitment in *Fagus sylvatica*. Molecular Ecology 20: 1997-2010

Olbrich M, Betz G, Gerstner E, Langebartels C, Sandermann H, Ernst D (2005) Transcriptome analysis of ozone-responsive genes in leaves of European beech (*Fagus sylvatica* L.). Plant Biology 7: 670-676

Olbrich M, Gerstner E, Bahnweg G, Haberle KH, Matyssek R, Welzl G, Heller W, Ernst D (2010) Transcriptional signatures in leaves of adult European beech trees (*Fagus sylvatica* L.) in an experimentally enhanced free air ozone setting. Environmental Pollution 158: 977-982

Page DR and Grossniklaus U (2002) The art and design of genetic screens: *Arabidopsis thaliana*. Nature Review Genetics 3: 124-136

Palle SR, Seeve C, Eckert AJ, Cumbie WP, Goldfarb B, Loopstra CA (2001) Natural variation in expression of genes involved in xylem development in loblolly pine (Pinus taeda L.). Tree Genetics & Genomes 7: 193-206

Papageorgiou AC, Vidalis A, Gailing O, Tsiripidis I, Hatziskakis S, Boutsios S, Galatsidas S, Finkeldey R (2008) Genetic variation of beech (*Fagus sylvatica* L.) in Rodopi (NE Greece). European Journal of Forest Research 127: 81-88

Pareek CS, Smoczynski R, Tretyn A (2011) Sequencing technologies and genome sequencing. Journal of Applied Genetics 52: 413-435

Pastorelli R, Smulders MJM, Van't Westende WPC, Vosman B, Giannini R, Vettori C, Vendramin GG (2003) Characterization of microsatellite markers in *Fagus sylvatica* L. and *Fagus orientalis* Lipsky. Molecular Ecology Notes 3: 76-78

Peakall R and Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6: 288-295

Peñuelas J and Boada M (2003) A global change-induced biome shift in the Montseny mountains (NE Spain). Global Change Biology 9: 131-140

References

Peuke AD, Schraml C, Hartung W, Rennenberg H (2002) Identification of drought-sensitive beech ecotypes by physiological parameters. New Phytologist 154: 373-387

Pflieger S, Lefebvre V, Causse M (2001) The candidate gene approach in plant genetics: a review. Molecular Breeding 7: 275-291

Polanski A and Kimmel M (2003) New explicit expressions for relative frequencies of single-nucleotide polymorphisms with application to statistical inference on population growth. Genetics 165: 427-436

Polle A, Schwanz P, Rudolf C (2001) Developmental and seasonal changes of stress responsiveness in beech leaves (*Fagus sylvatica* L.). Plant Cell and Environment 24: 821-829

Preacher K J (2001) Calculation for the chi-square test: An interactive calculation tool for chi-square tests of goodness of fit and independence [Computer software]. Available from http://quantpsy.org

Pretzsch H and Ďurský J (2002) Growth reaction of Norway spruce (*Picea abies* (L.) Karst.) and European beech (*Fagus silvatica* L.) to possible climatic changes in Germany. A sensitivity study. Forstwissenschaftliches Centralblatt 121: 145-154

Prunier J, Laroche J, Beaulieu J, Bousquet J (2011) Scanning the genome for gene SNPs related to climate adaptation and estimating selection at the molecular level in boreal black spruce. Molecular Ecology 20: 1702-1716

Pyhäjärvi T, García-Gil MR, Knürr T, Mikkonen M, Wachowiak W, Savolainen O (2007) Demographic history has influenced nucleotide diversity in European *Pinus sylvestris* populations. Genetics 177: 1713-1724

Quang ND, Ikeda S, Harada K (2008) Nucleotide variation in *Quercus crispula* Blume. Heredity 101: 166-174

Ramanjalu S and Bartels D (2002) Drought- and dessication-induced modulation of gene expression in plants. Plant Cell and Environment 25: 141-151

Reymond P (2001) DNA microarrays and plant defence. Plant Physiology and Biochemistry 39: 313-321

Rennenberg H, Seiler W, Matyssek R, Gessler A, Kreuzwieser J (2004) European beech (*Fagus sylvatica* L.) - a forest tree without future in the south of Central Europe?. Allgemeine Forst- und Jagdzeitung 175: 210-224

Rohlf FJ (1998) NTSYSpc: Numerical taxonomy and multivariate analysis system version 2.02. Exeter Software, Setauket, NY

Rosenblum EB and Novembre J (2007) Ascertainment bias in spatially structured populations: a case study in the eastern fence lizard. Journal of Heredity 98: 331-336

Rozen S and Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S and Misener S (eds.) Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, NJ, pp 365-386

Ryman N and Leimar O (2009) G_{ST} is still a useful measure of genetic differentiation - a comment on Jost's D. Molecular Ecology 18: 2084-2087

Ryynänen H, Tonteri A, Vasemägi A, Primmer CR (2007) A comparison of biallelic markers and microsatellites for the estimation of population and conservation genetic parameters in Atlantic salmon (*Salmo salar*). Journal of Heredity 98: 692-704

Sabaté S, Gracia CA, Sánchez A (2002) Likely effects of climate change on growth of *Quercus ilex*, *Pinus halepensis*, *Pinus pinaster*, *Pinus sylvestris* and *Fagus sylvatica* forests in the Mediterranean region. Forest Ecology and Management 162: 23-37

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor N.Y., Cold Spring Harbor Laboratory

Sander T, König S, Rothe GM, Janßen A, Weisgerber H (2000) Genetic variation of European beech (*Fagus sylvatica* L.) along an altitudinal transect at mount Vogelsberg in Hesse, Germany. Molecular Ecology 9: 1349-1361

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America 74: 5463-5467

Santos J and Corte-Real J (2006) Temperature extremes in Europe and wintertime large-scale atmospheric circulation: HadCM3 future scenarios. Climate Research 31: 3-18

Sathyan P, Newton RJ, Loopstra CA (2005) Genes induced by WDS are differentially expressed in two populations of aleppo pine (*Pinus halepensis*). Tree Genetics & Genomes 1: 166-173

Scalfi M, Troggio M, Piovani P, Leonardi S, Magnaschi G, Vendramin GG, Menozzi P (2004) A RAPD, AFLP and SSR linkage map, and QTL analysis in European beech (*Fagus sylvatica* L.). Theoretical and Applied Genetics 108: 433-441

Schaberg PG, DeHayes DH, Hawley GJ, Nijensohn SE (2008) Anthropogenic alterations of genetic diversity within tree populations: Implications for forest ecosystem resilience. Forest Ecology and Management 256: 855-862

Schär C, Vidale PL, Lüthi D, Frei C, Häberli C, Liniger MA, Appenzeller C (2004) The role of increasing temperature variability in European summer heatwaves. Nature 427: 332-336

Scharnweber T, Manthey M, Criegee C, Bauwe A, Schröder C, Wilmking M (2011) Drought matters - Declining precipitation influences growth of *Fagus sylvatica* L. and *Quercus robur* L. in north-eastern Germany. Forest Ecology and Management 262: 947-961

Schlink K (2011) Gene expression profiling in wounded and systemic leaves of *Fagus sylvatica* reveals up-regulation of ethylene and jasmonic acid signalling. Plant Biology 13: 445-452

Schmitz F, Polley H, Hennig P, Schwitzgebel F, Kriebitzsch WU (2004) Die zweite Bundeswaldinventur - BWI2. Bundesministerium für Verbraucherschutz, Ernährung und Landwirtschaft (Hrsg.), Bonn, http://www.bundeswaldinventur.de

Schönwiese CD, Bader S, Böhm R, Claussen M, Cubasch U, Fischer H, Gärtner U, Graßl H, Rahmstorf S, Sündermann S, Kromp-Kolb H, Richner H (2003) Klimastatement der Deutschen Meteorologischen Gesellschaft (DMG), der Österreichischen Gesellschaft für Meteorologie (ÖGM) und der Schweizerischen Gesellschaft für Meteorologie (SGM)

Schraml C and Rennenberg H (2000) Sensitivity of different ecotypes of beech trees (*Fagus sylvatica* L.) to drought stress. Forstwissenschaftliches Centralblatt 119: 51-61

Schraml C and Rennenberg H (2002) The different reactions of beech tree (*Fagus sylvatica* L.) ecotypes to drought stress. Forstwissenschaftliches Centralblatt 121: 59-72

Schulze ED, Hessenmöller D, Seele C, Wäldchen J, von Lüpke N (2010) Die Buche - Eine Kultur- und Wirtschaftsgeschichte. Biologie in unserer Zeit 40: 171-18

Seeb JE, Carvalho G, Hauser L, Naish K, Roberts S, Seeb LW (2011) Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organisms. Molecular Ecology Resources 11: 1-8

Seidling W (2007) Signals of summer drought in crown condition data from the German Level I network. European Journal of Forest Research 126: 529-544

Seo YS and Kim WT (2009) A genomics approach using expressed sequence tags and microarrays in ripening apple fruit (*Malus domestica* Borkh.). Journal of Plant Biology 52: 35-40

Sneath PHA and Sokal RR (1973) Numeric taxonomy: the principles and practice of numerical classification. W.H. Freeman, San Francisco

Smith JM and Haigh J (1974) The hitch-hiking effect of a favourable gene. Genetic Resources 23: 23-35

Spinnler D, Egh P, Körner C (2002) Four-year growth dynamics of beech-spruce model ecosystems under CO_2 enrichment on two different forest soils. Trees - Structure and Function 16: 423-436

Street NR, Skogstrom O, Sjodin A, Tucker J, Rodriguez-Acosta M, Nilsson P, Jansson S, Taylor G (2006) The genetics and genomics of the drought response in *Populus*. Plant Journal 48: 321-341

Tanaka K, Tsumura Y, Nakamura T (1999) Development and polymorphism of microsatellite markers for *Fagus crenata* and the closely related species, *F. japonica*. Theoretical and Applied Genetics 99: 11-15

Tarp P, Helles F, Holten-Andersen P, Larsen JB, Strange N (2000) Modelling near-natural silvicultural regimes for beech - an economic sensitivity analysis. Forest Ecology and Management 130: 187-198

References

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Research 22: 4673-4680

Tognetti R, Johnson JD, Michelozzi M (1995) The response of European beech (*Fagus sylvatica* L.) seedlings from two Italian populations to drought and recovery. Trees - Structure and Function 9: 348-354

Tognetti R, Minotta G, Pinzauti S, Michelozzi M, Borghetti M (1998) Acclimation to changing light conditions of long-term shade-grown beech (*Fagus sylvatica* L.) seedlings of different geographic origins. Trees - Structure and Function 12: 326-333

Vidalis A (2011) Patterns of nucleotide variation and gene-associated SNP analysis in a *Quercus* spp. forest at isocitrate dehydrogenase genes. Ph.D. Dissertation. Georg-August-University Göttingen, Germany

Vignal A, Milan D, SanCristobal M, Eggen A (2002) A review on SNP and other types of molecular markers and their use in animal genetics. Genetics Selection Evolution 34: 275-305

Višnjić C and Dohrenbusch A (2004) Frost resistance and phenology of European beech provenances (*Fagus sylvatica* L.). Allgemeine Forst- und Jagdzeitung 175: 101-108

Vitasse Y, Bresson CC, Kremer A, Michalet R, Delzon S (2010) Quantifying phenological plasticity to temperature in two temperate tree species. Functional Ecology 24: 1211-1218

Vitasse Y, François C, Delpierre N, Dufrêne E, Kremer A, Chuine I, Delzon S (2011) Assessing the effects of climate change on the phenology of European temperate trees. Agricultural and Forest Meteorology 151: 969-980

Vornam B, Decarli N, Gailing O (2004) Spatial distribution of genetic variation in a natural beech stand (*Fagus sylvatica* L.) based on microsatellite markers. Conservation Genetics 5: 561-570

Vornam B, Gailing O, Finkeldey R, Collada C, Guevera M, Soto Á. de María N, González-Martínez S, Díaz L, Alia R, Aranda I, Climent J, Cervera MT, Goicoechea P, Léger V, Eveno E, Derory J, Garnier-Géré P, Kremer A, Plomion C (2007) Naturally occurring nucleotide diversity in candidate genes for forest tree adaptation: magnitude, distribution and association

with quantitative trait variation. GABI - The German Plant Genome Research Program Progress Report 2004 - 2007, pp 116-120

Vornam B, Gailing O, Derory J, Plomion C, Kremer A, Finkeldey R (2011) Characterisation and natural variation of a dehydrin gene in *Quercus petraea* (Matt.) Liebl.. Plant Biology 13: 881-887

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: A new technique for DNA fingerprinting. Nucleic Acids Research 23: 4407-4414

Wachowiak W, Balk PA, Savolainen O (2009) Search for nucleotide diversity patterns of local adaptation in dehydrins and other cold-related candidate genes in Scots pine (*Pinus sylvestris* L.). Tree Genetics & Genomes 5: 117-132

Wagner S, Collet C, Madsen P, Nakashizuka T, Nyland RD, Sagheb-Talebi K (2010) Beech regeneration research: from ecological to silvicultural aspects. Forest Ecology and Management 259: 2172-2182

Wang K (2001) Gene flow and mating system in European beech. Ph.D. Dissertation. Georg-August-University Göttingen, Germany

Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nature Review Genetics 10: 57-63

Warren R, Price J, Fischlin A, de la Nava Santos, Midgley G (2011) Increasing impacts of climate change upon ecosystems with increasing global mean temperature rise. Climatic Change 106: 141-177

Weber JL (1990) Human DNA polymorphisms and methods of analysis. Current Opinion in Biotechnology 1: 166-71

White C, Selkoe KA, Watson J, Siegel DA, Zacherl DC, Toonen RJ (2010) Ocean currents help explain population genetic structure. Proceedings of the Royal Society B-Biological Sciences 277: 1685-1694

Whitlock MC (2011) G'_{ST} and D do not replace F_{ST}. Molecular Ecology 20: 1083-1091

References

Williams LM and Oleksiak MF (2011) Ecologically and evolutionarily important SNPs identified in natural populations. Molecular Biology and Evolution 28: 1817-1826

Wilmans O (1990) Pflanzen prägen Lebensräume – Die Rotbuche, *Fagus sylvatica* L.. Biologie in unserer Zeit 20: 60-62

Wong MML, Cannon CH, Wickneswari R (2011) Identification of lignin genes and regulatory sequences involved in secondary cell wall formation in *Acacia auriculiformis* and *Acacia mangium* via *de novo* transcriptome sequencing. BMC Genomics 12: 342

Wright S (1951) The genetical structure of populations. Annals of Eugenics 15: 323-354

Wright S (1965) The interpretation of population-structure by F-statistics with special regard to systems of mating. Evolution 19: 395-420

Wu SH, Ramonell K, Gollub J, Somerville S (2001) Plant gene expression profiling with DNA microarrays. Plant Physiology and Biochemistry 39: 917-926

Wühlisch G, Krusche D, Muhs HJ (1995) Variation in temperature sum requirement for flushing of beech provenances. Silvae Genetica 44: 343-346

Zhou LM, Tucker CJ, Kaufmann RK, Slayback D, Shabanov NV, Myneni RB (2001) Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999. Journal of Geophysical Research - Atmospheres 106: 20069-20083

Zhu MJ and Zhao SH (2007) Candidate gene identification approach: progress and challenges. International Journal of Biological Sciences 3: 420-427

8 Appendices

Appendix 1: Manuscript

DNA Sequence Variation and Development of SNP Markers in Beech (*Fagus sylvatica* L.)

Seifert S[1], Vornam B[1], Finkeldey R[1]

[1]Georg-August-University Göttingen, Faculty of Forest Sciences and Forest Ecology, Büsgen Institute, Forest Genetics and Forest Tree Breeding

Büsgenweg 2, 37077 Göttingen, Germany

Abstract

European beech (*Fagus sylvatica* L.) is one of the most important deciduous tree species in Central Europe. The potential of beech to adapt to climate change, higher temperatures and less precipitation in the summer months is still unknown. Most studies in beech used microsatellite, AFLP (Amplified Fragment Length Polymorphism) or isozyme markers which have only a restricted potential to analyse adaptation. Only few studies investigated genes probably involved in the adaptation to drought stress and bud phenology in beech. In this study, SNP (single nucleotide polymorphisms) markers were developed in order to analyse adaptation and their technical advantages compared to microsatellites and AFLPs were discussed. Partial sequences of ten candidate genes probably involved in drought stress and/ or bud phenology were identified at the genomic level, and SNPs and indels (insertions/deletions) in coding and non-coding regions were analysed. Plant material was sampled along a precipitation gradient in Germany. In total, 8145 bp were sequenced and analysed, 4038 bp were located in exon and 4107 bp in intron regions. 63 SNPs and 11 indels were detected, which are differently distributed over the studied gene regions. The nucleotide diversity ranged from 0 to 6.62 (π x10^{-3}) and is comparable to other tree species, whereas the mean nucleotide diversity (2.64) for *F. sylvatica* is comparatively low. These results will help to investigate the genetic basis of drought stress and bud burst, and to conduct association mapping in natural populations. Furthermore, the detected SNPs can also be used for population genetic studies.

Key words: Climate change, adaptation, candidate gene, *Fagus sylvatica*, SNPs

Author contributions: S.S., B.V. and R.F. designed research; S.S. performed research; S.S. analysed data; S.S., B.V. and R.F. wrote the paper.

Introduction

Forest tree populations consist of sessile, long-lived organisms which must survive temporally varying environmental conditions that are presently also affected by accelerated global climate change. Hence, the presence and maintenance of genetic variation at genes controlling adaptive traits is important for the long-term persistence and stability of forest tree populations in order to survive heterogeneous conditions.

Genetic markers are ultimately based on the variation of DNA sequences. However, the sequences of the currently most commonly used genetic markers in beech (*Fagus sylvatica* L.) are not directly observed and are usually unknown. Only particular aspects of the variation are investigated within PCR amplified DNA fragments such as the number of tandem repeats in microsatellite motives (e.g., Pastorelli et al. 2003) or the absence or presence of restriction sites in Amplified Fragment Length Polymorphisms (AFLPs; e.g., Gailing and v. Wühlisch 2004). The amplified genomic regions are usually either unknown such as in anonymous AFLPs or are located in non-coding regions of the DNA (most microsatellites). Accordingly, most of the variation at molecular DNA-based markers is assumed to be selectively neutral. In addition, the accurate scoring of microsatellite markers (SSRs) and AFLPs can be difficult due to PCR and electrophoresis artifacts, and the comparability between different laboratories is problematic. Although different SSR loci can be multiplexed for a higher throughput, the multiplex process is complex, not always successful, and limited to a low number of loci.

Furthermore, isozymes are also important markers. These biochemical markers can be used to assess genetic diversity at gene loci coding for enzymes which serve important functions in the metabolism of plants. For beech, the genetic diversity has been studied extensively at selected isozyme gene loci (e.g., Müller-Starck and Ziehe 1991; Müller-Starck and Starke 1993). However, the analysis of isozymes allows to explore only a fraction of the underlying sequence variation, and only few gene loci coding for selected soluble enzymes can be investigated by means of enzyme electrophoresis. Furthermore, it is questionable whether isozymes are suitable to detect adaptive variation or if most of the markers are neutral (e.g., Eriksson 1998 and references therein).

Comparative sequencing is the ultimate method to detect variation within any DNA fragment. Today, it is already possible to analyse and compare whole genomes of organisms using high-throughput sequencing technologies, also called next-generation sequencing. The most frequently used techniques at the moment are 454 (Roche), SOLiD (Life Technologies) and Illumina (e.g., reviewed by Glenn 2011, Deschamps and Campbell 2010). However, this method is still too expensive to analyse a sufficient number of individuals for population genetic studies or the study of adaptation in natural populations. For non-model organisms like trees, where most of the genomes are not sequenced yet, next generation sequencing is not an established technique. Considering these limitations, the most promising markers for the study of adaptation at the moment are SNPs (Single Nucleotide Polymorphism), i.e., the substitutions of only one nucleobase. SNPs are the most frequent variations found in DNA (Brookes 1999) and the analysis is not restricted to special enzymes. SNP markers can, unlike isozymes, also be used to analyse regions controlling the transcription of genes, e.g., transcription factor binding sites. In comparison to SSRs and AFLPs, they are valuable markers to study adaptation of plants, for example to changing environmental conditions (Gailing et al. 2009). For human and plant model organisms this type of marker is already establishment and often used (e.g., *Populus tremula*, Ingvarsson 2004). However, SNP markers are nowadays more and more applied for non-model organisms like most of the forest trees (e.g., Seeb et al. 2011, Helyar et al. 2011).

Unfortunately, SNP analyses in human populations revealed that only few SNPs can be associated with phenotypic traits (Yoshiura et al. 2006). Some of these SNPs with a direct impact on phenotypes are likely to be under selection, while the vast majority of SNPs are likely to behave selectively neutral. However, besides the study of adaptation, SNPs in non-coding regions can also be used instead or additional to other neutral markers. The analysis of an unprecedented number of mostly selectively neutral SNP loci allows new insights in the population genetic structure of species that cannot be found with other genetic markers. For example, the observation of more than 500,000 SNPs in over 3,000 Europeans revealed overall genetic differentiation patterns among humans on the continent closely resembling their spatial distribution on the continent (Novembre et al. 2008). Furthermore, comparing AFLP and microsatellite markers with SNP markers, the latter markers have some important advantages. The scoring is unambiguously and comparable between laboratories, even if different platforms are used for the analysis. Jones et al. (2007) compared SSR and SNP markers in maize and concluded that SNP markers have a lower level of missing data

and are more reliable. For the analysis of SNPs, multiplexing can be conducted easily, and thus, the throughput is very high. Estimations show that SNP costs are lower in comparison to SSR markers (Jones et al. 2007). However, efficiency and costs strongly depend on the platform used for SNP scoring, but it is predictable that the efficiency costs will decrease in the future.

The aim of this study is to detect SNPs within candidate genes, related to phenotypic traits in beech. European beech (*Fagus sylvatica* L.) is one of the predominant and most important tree species in European forests and covers a large geographic range in Central Europe. The species is wind-pollinated, predominantly outcrossing, a monoecious tree with heavy fruits and therefore with limited seed dispersal.

So far, in beech most studies on the genetic diversity and differentiation were focused on the spatial genetic structure or on the impact of different silvicultural treatments using AFLP and microsatellite markers (Vornam et al. 2004; Buiteveld et al 2007; Nyári 2010; Oddou-Muratorio et al. 2011). In the context of global climatic changes predicting less precipitation in summer and higher precipitation in winter contradictory opinions exist whether beech will be adaptable to the enhanced drought stress conditions in the summer months (Gessler et al. 2004; Rennenberg et al. 2004; Ammer et al. 2005). Another effect of the predicted global change is the extending growing season influencing the growth of beech in the future. Earlier bud burst is supposable which will lead to an increasing risk of late frost damage. The analysis of the variation within 'candidate' genes potentially involved in adaptation to a phenotypic trait is one possibility to investigate the genetic background of adaptation. Until now, only few studies aim to identify genes which are involved in drought stress response and bud phenology in beech (Lalagüe et al. 2010), and only a limited number of beech sequences are available (Jimenez et al. 2008; Olbrich et al. 2005; Olbrich et al. 2010; Schlink 2011). Therefore, the candidate gene approach described here is based on both published *Fagus sylvatica* sequences and orthologous sequences identified in other plant species such as oaks (Gailing et al. 2009; Vornam et al. 2011).

SNPs were analysed in both coding (exons) and non-coding regions (introns) of the identified genes. For the purpose of using SNP markers additionally or in place of microsatellite markers, it is necessary to analyse both regions. For the study of adaptation, SNPs in coding regions changing the amino acid composition of the gene products (non-synonymous SNPs) are most interesting, but non-coding regions can also be of relevance. Whereas non-synonymous SNPs potentially lead to changes of protein structures, SNPs in

intron regions potentially influence gene splicing and enable a single gene to increase its coding capacity producing several structurally distinct isoforms (Baek et al. 2008).

The results described here are a prerequisite for association mapping in natural populations in order to identify SNPs correlated to phenotypic traits like drought stress response and bud phenology. Other applications of the analysis of SNPs are, for example, population genetic studies concerning the history, structure and demography of populations or molecular systematic studies and parentage analyses (Garvin et al. 2010; Morin et al. 2004). The SNPs identified in this study are suitable for population genetic investigations complementing other frequently used markers such as microsatellites and AFLPs. Furthermore, this study provides the first estimates of nucleotide and haplotype diversity in *Fagus sylvatica*.

Material and Methods

Plant material: Fresh leaves were sampled in early summer 2009 in three different regions of northern Germany along a rainfall gradient (Table 1). All stands are jointly investigated by several research groups within the collaborative project 'Climate Impact and Adaptation Research in Lower Saxony' (KLIFF; http://www.kliff-niedersachsen.de.vweb5-test.gwdg.de/?page_id=26). Each region is represented by two populations differing in their soil type. Three trees per population were used for SNP identification. Thus, the total sample size was 3 (regions) x 2 (populations/region) x 3 (trees/population) = 18 trees. The investigated trees were separated by a distance of at least 50m to minimize the risk of sampling related plants (Vornam et al. 2004).

Table 1: Sampling sites in Germany, Lower-Saxony and Saxony-Anhalt

Closest village	Position	Elevation	Annual amount of precipitation*	Amount of precipitation from April-September*	Soil type
Calvörde	N52 22.819 E11 17.406	65m	543mm	294mm	sand
	N52 24.238 E11 15.661	75m	543mm	294mm	loam
Göhrde	N53 08.660 E10 52.003	90m	665mm	347mm	sand
	N53 07.379 E10 49.224	85m	675mm	349mm	loam
Unterlüß	N52 49.831 E10 18.985	130m	766mm	374mm	sand
	N52 49.894 E10 19.183	130m	766mm	374mm	loam

*provided by National Climate Monitoring of Deutscher Wetterdienst (DWD)

Selection of candidate genes: All candidate genes have been chosen based on literature surveys suggesting an impact of the genes on either drought stress or bud phenology (Table 2). The Evoltree EST database (www.evoltree.org) and the EMBL Nucleotide Sequence Database (http://www.ebi.ac.uk/embl/) were mainly used to find corresponding *Fagus sylvatica* sequences. Alternatively, sequences of *Quercus petraea* were transferred to *F. sylvatica* (Vornam et al. 2007; Vidalis 2011). The selected sequences were verified by a TBLASTX search (Washington University Basic Local Alignment Search Tool Version 2.0) and used for primer design in order to amplify the corresponding genomic regions in beech.

Table 2: Selected candidate genes related to drought stress response or bud phenology

Name (abbreviation)	Gene	Drought stress/bud phenology	Reference with investigated species
ALDH	*Aldehyde dehydrogenase*	Drought stress	Gao and Han 2009 (*Oryza sativa*), Guo et al. 2009 (*Hordeum vulgare)*, Sathyan et al. 2005 (*Pinus halepensis*)
Cry	*Cryptochrome*	Bud phenology	Muleo et al. 2001 (*Prunus cerasifera*)
Dhn	*Dehydrin*	Drought stress and bud phenology	Beck et al. 2007 (*Cicer pinnatifidum*), Gonzáles-Martínez et al. 2006 (*Pinus taeda)*, Jimenez et al. 2008 (*Fagus sylvatica*), Ramanjalu and Bartels 2002 (*Picea glauca*), Wachowiak et al. 2009 (*Pinus sylvestris*), Vornam et al. 2011 (*Quercus petraea*)
ERD	*early response to dehydration*	Drought stress	Eveno et al. 2008 (*Pinus pinaster*), Street et al. 2006 (*Populus* spp.)
IDH	*Isocitrate dehydrogenase*	Drought stress	Liu et al. 2010 (*Zea mays)*
APX1, APX4, GPX	*Peroxidases (ascorbate and glutathione)*	Drought stress	Lu et al. 2010 (*Zea mays)*, Street et al. 2006 (*Populus* spp.)
PhyB	*Phytochrome B*	Drought stress and bud phenology	Boggs et al. 2010 (*Arabidopsis thaliana*), Frewen et al. 2000 (*Populus* spp.), Ingvarsson et al. 2006 (*Populus tremula*)
CHZFP	*cys-his-zinc finger protein*	Drought stress	Lu et al. 2010 (*Zea mays)*, Street et al. 2006 (*Populus* spp.)

DNA isolation, amplification, cloning and sequencing: Total DNA was extracted from leaves using the DNeasy™ 96 Plant Kit (Qiagen, Hilden, Germany). The amount and the quality of the DNA were analysed by 0.8% agarose gel electrophoresis with 1 x TAE as

running buffer (Sambrook et al. 1989). DNA was stained with ethidium bromide, visualized by UV illumination and compared to a Lambda DNA size marker (Roche).

Primers for amplification and direct sequencing of the amplification product (Table 3) were designed by using the program Primer3 (v.0.4.0; Rozen and Skaletsky 2000; http://www-genome.wi.mit.edu/cgi-bin/primer/ primer3_www.cgi). Primers were checked for self-annealing, dimer and hairpin formations using the program Oligo calc: Oligonucleotide Properties Calculator (http://www.basic.northwestern.edu/biotools/oligocalc.html). PCR amplifications were conducted in a 15µl volume containing 2µl of genomic DNA (about 10ng), 7.5µl HotStar*Taq* Master Mix Kit (Qiagen, Hilden, Germany) and 0.3µM of each forward and reverse primer. The PCR protocol consisted of an initial denaturation step of 95°C for 15min, followed by 35 cycles of 94°C for 60sec (denaturation), different temperatures according to the primers (Table 3) for 45sec (annealing), 72°C for 90sec (extension) and a final extension step of 72°C for 20min.

PCR products were analysed by 1% agarose gel electrophoresis with 1 x TAE as running buffer (Sambrook et al. 1989). DNA was stained with ethidium bromide and visualized by UV illumination. PCR products were excised from gel and purified using the Geneclean®kit (MP Biomedicals, Illkirch, France). The purified products were cloned into a pCR2.1 vector using the TOPO TA Cloning® kit (Invitrogen, Carlsbad, CA) with slight modifications. The inserts were amplified by colony PCR using M13 forward (-20)(5'-GTAAAACGACGGCCAG-3') and M13 reverse (5'-CAGGAAACAGCTATGAC-3') primers, visualized by agarose gel electrophoresis, excised from the gel and purified (see above). Three to four different clones of the fragments were sequenced using both M13 forward and M13 reverse primers in order to identify the presence of different haplotypes within individuals (heterozygotes) and to control for sequencing errors. The sequencing reaction was carried out with the Big Dye® Terminator v.3.1. Cycle Sequencing Kit (Applied Biosystems) based on the dideoxy-mediated chain termination method (Sanger et al. 1977). Sequencing reactions were run on an ABI 3100*xl* Genetic Analyser (Applied Biosystems). The sequenced fragments were verified by a TBLASTX search. Putative introns and exons were determined following the GT-AG rules (Breathnach et al. 1978).

Data analysis: For editing and visual examination of the sequences as well as for the analysis of SNPs and indels (insertions/deletions) within the genes, the sequences were aligned using Codon Code Aligner (CodonCode cooperation, www.codoncode.com) and BioEdit version 7.0.9.0 (Hall 1999) using ClustalW multiple alignment (Thompson et al.

1994). Only polymorphisms with Phred scores above 25 in the chromatograms were considered (Ewing et al. 1998). Only SNPs appearing at least twice were analysed in order to avoid sequencing errors. Haplotype diversity, nucleotide diversity (π) and F_{ST} values were calculated excluding indels using DnaSP v.5.0 (Librado and Rozas 2009).

Table 3: Primer sequences and corresponding annealing temperatures for the selected candidate genes

Name	Gene	Primer sequence (5'-3')	Annealing temperature	Accession No[3]
ALDH	*Aldehyde dehydrogenase*	F: AAG ATC TGG TGT TGA AAA TGG AG	53°C	FR774766
		R: TGC ATT CTT CAA AGG AGT GAC		
APX1[2]	*Ascorbate peroxidase*	F Part 1: AGG CGA AGA GAA AGC TCA GG	55°C	FR774767
		R Part 1: AAG AAA GCA ACT ATC AGC CTC A		
		F Part 2: AAG CAG ATT TGT TGA CAT TAA TAT TTC		
		R Part 2: GCA AAG AAG GCA TCC TCA TC		
APX4 (Part 1)[1]	*Ascorbate peroxidase*	F Part 1: ATC AAG GGA ACG CTT TCT ACG	55°C	FR775801
		R Part 1: TCC ACA TCA CAT CTC AAC AGC		
APX4 (Part 2)[1]	*Ascorbate peroxidase*	F Part 2: GGC CTC TTA AGT GCC AAT TC	55°C	FR775801
		R Part 2: CTC CCC TCT GGA TCT GGT TC		
Cry[2]	*Cryptochrome*	F Part 1: CTT GAG ATG ATG CTC TTG GTT G	53°C	FR775802
		R Part 1: ATG GGC TCA ATT GGA GAA TC		
		F Part 2: TTT TCT CCA CAG GGA TCA CG		
		R Part2: AAG TCA TGC TTG GGA CCA TC		
Dhn	*Dehydrin*	F: TGC ACC CCA AAA TGA TGA AT	54°C	FR772355
		R: TGA TCC CCT TCT TCT CAT GG		
ERD	*early response to dehydration*	F: GGC AAT GGA CGT AAT TTC TCA	51°C	FR775803
		R: CTG GGC TGC TGA ATC GTC		
GPX	*Glutathione peroxidase*	F: GGC TGC CAT GCC TTT CTC	55°C	FR796394
		R: GAA ATC ATA GAT AGT CTT CTC CGT AGC		
IDH (Part 1)[1]	*Isocitrate dehydrogenase*	F: GTG ATC AGT ACA GGG CAA CTG	50°C	FR796392
		R: AAG GTA CAA GAG GGG CTT TG		
IDH (Part 2)[1]	*Isocitrate dehydrogenase*	F: GAT GAT ATG GTT GCT TAT GCC ATG	50°C	FR796392
		R: GGT TTC ACC ACC TTT CTG ATG GAC		
PhyB	*Phytochrome B*	F: CAG GCA TCG AGG TTT TTG TT	50°C	FR774765
		R: GAA GGG AAT GCA CCT AGC AG		
CHZFP	*cys-his-zinc finger protein*	F: CTT TGC AAG GAT GAG ACT GG	50°C	FR796395
		R: ACG CAT CTG ATG AGC ATT TG		

[1]Both parts belong to the same gene but the sequenced parts do not overlap. [2]Both parts belong to the same gene and the two parts overlap. [3] EMBL Nucleotide Sequence Database (http://www.ebi.ac.uk/embl/)

Results

Fragments from ten different genes were successfully amplified, identified and analysed. After sequencing, all fragments were verified using TBLASTX search. Any similarity with an E-Value of less than 10^{-3} was considered to be a hit. In total 9468 bp were analysed with 4418 bp in exon regions and 5050 bp in intron regions (Table 4). All exons and introns could be determined following the GT-AG rule. No alternative splicing was found. The reading frame was assessed according to the TBLASTX results (see above).

Insertions/Deletions

In seven different genes, 11 indels (insertions/deletions) were identified, mainly in intron regions (Table 4). Some of them showed a microsatellite repeat motif (see supplementary material). Only two indels also represented by microsatellite motives were found within coding regions (gene *ERD* and *CHZFP*). The lengths of these indels were multiples of 3 bps, thus the reading frame is not shifted.

Single Nucleotide Polymorphisms

SNPs only appearing once were excluded from the analyses in order to avoid the selection of false positives caused by sequencing errors, although they could be true SNPs. Therefore, only common SNPs are presented here that may be also present in *Fagus sylvatica* trees in other regions in Europe than investigated in this study. Considering these limitations, in total 63 SNPs were found differently distributed over the analysed gene fragments. The results indicate that numerous of these SNPs are linked (see supplementary material). Excluding the potentially linked SNPs from the analysis, 45 SNPs remain. However, because of the low number of investigated trees, the linkage of these SNPs is not unambiguous and it is not possible to clearly define a set of tag SNPs.

More SNPs were found in non-coding regions (1 SNP every 112 bp) than in coding regions (1 SNP every 245 bp). Eighteen SNPs were found in coding regions, and seven of them were non-synonymous. All non-synonymous SNPs led to an amino acid exchange, no one caused an early stop codon. The number of haplotypes ranged from one to eleven. The nucleotide diversity (π) was higher at non-coding sites than at coding sites for most genes. Exceptions are the genes *GPX* and *PhyB* for which the investigated non-coding regions were very short (Table 5). Furthermore, the nucleotide diversity at synonymous sites was in most cases higher than at non-synonymous sites (Table 5).

Table 4: Length, exons, introns, indels and SNPs of the amplified candidate genes

Gene name	Total length (bp)	No and length (bp) of exons	No and lengths (bp) of introns	No indels	No SNPs	No of non-coding SNPs	No of synonymous SNPs	No of non-synonymous SNPs
ALDH	519	3/ 350	2/ 169	1	8	4	2	2
APX1	1566	7/ 577	7/ 989	4	15	12	3	
APX4 (Part 1)	608	2/ 107	3/ 501	1	6	6	0	0
APX4 (Part 2)	814	2/ 238	3/ 576	1	11	9	1	1
Cry	1439	2/ 379	1/ 1060	1	1	1	0	0
Dhn	546	2/ 455	1/ 91	0	3	1	0	2
ERD (Part 1)	317	1/ 242	1/ 75	0	2	2	0	0
ERD (Part 2)	152	1/ 145	1/ 7	1	0	0	0	0
GPX	224	1/ 224	0/ 0	0	1	0	0	1
IDH (Part 1)	646	3/ 259	3/ 387	0	10	8	2	0
IDH (Part 2)	474	3/ 222	2/ 252	1	4	2	1	1
PhyB	301	1 /301	0/ 0	0	2	0	2	0
CHZFP	539	1/ 539	0/ 0	1	0	0	0	0
total	8145	29/ 4038	25/ 4107	11	63	45	11	7

F_{ST} was analysed grouping the studied trees according to their region (Calvörde, Göhrde or Unterlüß), each region includes trees from two different populations. The detected values were rather low, between 0 and 0.157 with a mean value of 0.012 (Table 5). This mean value is comparable to the results of a study analysing the same populations with nine microsatellite markers (F_{ST}: 0.022; Seifert, unpublished). However, the strongest differentiation with microsatellite loci was 0.032, whereas some candidate genes showed a considerable higher differentiation. The highest differentiations were found investigating the genes *ALDH*, *ERD* (Part1) and *IDH* (Part 2) with values above 0.05 (Table 5). Derory et al. (2010) found comparable results for SNPs analysed in candidate genes and microsatellites for *Quercus petraea*.

The partial sequence encoding aldehyde dehydrogenase was found to be of special interest. All but one of the detected SNPs (non-coding, synonymous and non-synonymous) were represented in three different haplotypes. The indel found in the non-coding region is also linked to two other non-coding SNPs. Within this gene fragment, two non-synonymous SNPs were identified in the same codon which were not linked. Therefore, three different amino acids are encoded depending on the combination of the SNPs indicating different lineages of the alleles. The *dehydrin* sequence is also interesting because the larger part of the sequence represents an exon region in which two SNPs were detected. Both of them are non-synonymous, and one is linked to the third non-coding SNP.

The position of the SNPs in the gene regions and additional information about the composition of the indels can be found in the supplementary material.

Table 5: Haplotype diversity and nucleotide diversity for the different genes

Gene name	No of haplotypes	Haplotype diversity	Total Nucleotide diversity (π x10^{-3})	Nucleotide diversity (π x10^{-3})				F_{ST}
				non-coding sites	coding sites	syn. sites	non-syn. sites	
ALDH	4	0.613	5.91	10.59	3.82	8.60	2.36	0.157
*APX*1	11	0.584	2.61	3.42	1.30	5.42	0.00	0.017
*APX*4 (Part 1)	7	0.756	3.91	4.81	0.00	0.00	0.00	-0.039
*APX*4 (Part 2)	5	0.756	3.94	4.58	2.41	4.82	1.61	0.016
Cry	2	0.157	0.11	0.15	0.00	0.00	0.00	-0.091
Dhn	3	0.627	1.83	2.70	1.66	0.00	2.16	0.044
ERD (Part 1)	3	0.629	2.52	10.67	0.00	0.00	0.00	0.058
ERD (Part 2)	1	0.000	0.00	0.00	0.00	0.00	0.00	0.000
GPX	2	0.157	0.70	0.00	0.71	0.00	0.95	0.005
IDH (Part 1)	11	0.883	6.62	8.41	4.00	17.7	0.00	-0.006
IDH (Part 2)	5	0.744	3.56	3.70	3.40	9.65	1.45	0.062
PhyB	3	0.656	2.65	-	2.65	11.83	0.00	-0.027
CHZFP	1	0.000	0.00	0.00	0.00	0.00	0.00	-0.043

Discussion

In this study, parts of ten different candidate genes have been investigated. Because of the limited sequence information for *Fagus sylvatica* it was not possible to sequence whole genes. However, this study was able to detect numerous SNPs and indels in non-coding and, probably more important, in coding regions of genes potentially involved in drought stress response and bud phenology. Most of the indels were found in intron regions. Only two were located in exon regions. Indels in exon regions are important due to their influence on protein structures and thus, on phenotypic trait changes (for example reviewed by Li et al. 2002). However, short indels, like the ones that were found in this study, seem to have only minor impact on protein structures (Kim and Guo 2010). Because SNPs appearing only once were excluded from the analysis, the presented data most likely underestimate the number of SNPs. Other reasons for underestimating the number of SNPs are the limited number of investigated samples and sequencing of only three to four clones which does not allow to correctly identify all heterozygotes. Nevertheless, 63 reliable SNPs were found. As expected, more SNPs were found in non-coding regions than in coding regions and the nucleotide diversity was higher non-coding sites than in coding sites.

Some of the non-synonymous SNPs detected in this study are of special interest because they might have an influence on the protein structure and protein function. For example, one non-synonymous SNP found in the partial sequence encoding aldehyde dehydrogenase is coding for proline, which leads to confirmation changes of the protein (Chou and Fasman 1974). The first non-synonymous SNP found in the partial dehydrin gene sequence leads to an amino acid substitution from aspartic acid to histidine implicating also a changed charge profile of the different genotypes from negatively charged to positively charged.

The nucleotide diversity (π x10^{-3}) found in this study ranged from 0 to 6.62 and is comparable to the nucleotide diversities analysed in other tree species, for example *Quercus petraea* (1.09 - 14.7, Derory et al. 2010; 3.02 - 11.96, Gailing et al. 2009), *Quercus crispula* (6.67 - 7.21, Quang et al. 2008), *Populus tremula* (2.7 - 18.8, Ingvarsson 2004), *Pinus taeda* (0.1 - 11.79, González-Martínez et al. 2006) and *Pseudotsuga menziesii* (2.37 - 13.78, Krutovsky and Neale 2005). The mean nucleotide diversity of 2.64 for *Fagus sylvatica* is comparatively low (*Q. petraea*: 6.15 or 5.42; *Q. crispula*: 6.93; *P. tremula*: 11.1; *P. taeda*: 7.5; *P. menziesii*: 6.55). One reason for lower nucleotide diversity values may be the exclusion of all SNPs appearing only once from the analysis (see above). However, the significance of mean values for nucleotide diversity depends on the analysed candidate

gene. Olson et al. (2010) also found in *Populus balsamifera* that the diversity is affected by the functional classification of the analysed candidate genes. They found higher diversity in gene fragments with insertion∕deletion length variation (indels) than in fragments that did not contain indels. Studies that do not include regions with length variation may slightly underestimate the overall level of nucleotide diversity.

The variation found in this study can be used to develop SNP markers and to apply them additionally or instead of neutral SSR or AFLP markers. SNP markers are more optimal markers for many applications because they are suitable for high-throughput analysis, inexpensive, highly reproducible, easy to score, comparable between different laboratories and some SNPs clearly show higher differentiation values. Although SNP markers have some advantages, there are also some drawbacks that have to be discussed. A disadvantage of SNPs is their normally biallelic character. Thus, they are less polymorphic than SSRs. To replace ten to twenty highly polymorphic SSR markers, around 100 neutral SNP markers are necessary (Kalinowski 2002). However, the virtually unlimited number of SNP markers in the different parts of the genome of higher organisms creates opportunities for the investigation of genetic variation within species with numerous applications in population genomics. Furthermore, an ascertainment bias can occur, i.e., the deviation from the expected allele frequency distribution for the case that the SNPs are identified based on only a few individuals and later used for the genotyping of a large sample set. This problem can be overcome, for example, by a direct correction of the statistical estimators (Helyar et al. 2011).

Another important application of SNPs is the study of adaptation. The SNPs found in this study can be useful, for example, to extend the investigation of Kraj and Sztorc (2009) who analysed the variability of phenological forms (bud burst) in beech using a set of five microsatellite markers. They pointed out that the neutral microsatellite loci are not directly linked with adaptive genetic variation and the genetic differences between the phenological forms of beech (early-, intermediate- and late-flushing individuals) have therefore no direct effect on the fitness of these forms. But genetic diversity and fitness are the basis for the ability of forest tree populations to adapt to changes of the environment (Krutovsky and Neale 2005). Because forest trees are continuously challenged by changing environmental conditions during their life time, adaptive genetic variation in relevant genes and phenotypic plasticity are essential for the long-term adaptation to stressful conditions. Thus, the knowledge of adaptive genetic variation is a basis for future management and

conservation strategies of forests (Krutovsky and Neale 2005), and can assist in breeding in combination with traditional phenotypic selection (Neale 2007). Furthermore, the results presented here are a prerequisite for association mapping studies in order to identify genomic regions and even individual nucleotides underlying phenotypic variation. The success of such an approach largely depends on the reasonable selection of candidate genes. This study revealed huge differences in diversity among the investigated candidate genes. Whereas the genes with regulatory function such as the cys-his-zinc finger protein (*CHZFP*) representing a transcription factor show low or moderate SNP variation, genes with a structural function as the ascorbate peroxidase show comparatively high SNP variation.

In the view of the above considerations, we propose to apply the genomic resources developed for beech by the identification and characterization of SNPs in coding and non-coding regions of candidate genes to investigate both the genetic basis of adaptive variation and the population structure of beech at the ultimate level of genetic resolution. In future, there will be the possibility to use whole genome sequencing for these applications. But considering the costs and the possibilities at the moment for non-model organisms, the comparable sequencing of (partial) genes and the identification of SNPs presented in this study is the best available method.

Conflict of interest - The authors declare that they have no conflict of interest.

References

Ammer C, Albrecht L, Borchert H, Brosinger F, Dittmar C, Elling W, Ewald J, Felbermeier B, von Gilsa H, Huss J, Kenk G, Kölling C, Kohnle U, Meyer P, Mosandl R, Moosmayer HU, Palmer S, Reif A, Rehfuess KE, Stimm, B (2005) Future suitability of beech (*Fagus sylvatica* L.) in Central Europe: Critical remarks concerning a paper of Rennenberg et al. (2004). Allg Forst- u J-Ztg 176:60–67

Baek JM, Han P, Iandolino A, Cook DR (2008) Characterization and comparison of intron structure and alternative splicing between *Medicago truncatula, Populus trichocarpa*, *Arabidopsis* and rice. Plant Mol Biol 67:499–510

Beck EH, Fettig S, Knake C, Hartig K, Bhattarai T (2007) Specific and unspecific responses of plants to cold and drought stress. J Biosci 32:501-510

Boggs JZ, Loewy K, Bibee K, Heschel MS (2010) Phytochromes influence stomatal conductance plasticity in *Arabidopsis thaliana*. Plant Growth Regul 60:77-81

Breathnach R, Benoist C, O'Hare K, Gannon F, Chambon P (1978) Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. Proc Natl Acad Sci USA 75:4853-4857

Brookes A (1999) The essence of SNPs. Gene 234:177-186

Buiteveld J, Vendramin GG, Leonardi S, Kamer K, Geburek T (2007) Genetic diversity and differentiation in European beech (*Fagus sylvatica* L.) stands varying in management history. Forest Ecol Manag 247:98-106

Chou PY and Fasman GD (1974) Conformational parameters for amino acids in helical, β-sheet, and random coil regions calculated from proteins. Biochemistry 13:211-222

Derory J, Scotti-Saintagne C, Bertocchi E, Le Dantec L, Graignic N, Jauffres A, Casasoli M, Chancerel E, Bodenes C, Alberto F, Kremer A (2010) Contrasting relations between diversity of candidate genes and variation of bud burst in natural and segregating populations of European oaks. Heredity 105:401-411

Deschamps S and Campbell MA (20120) Utilization of next-generation sequencing platforms in plant genomics and genetic variant discovery. Mol Breeding 25:553-570

Eriksson G (1998) Evolutionary forces influencing variation among populations of *Pinus sylvestris*. Silva Fenn 32:173-184

Eveno E, Collada C, Guevara MA, Leger V, Soto A, Diaz L, Leger P, González-Martínez SC, Cervera MT, Plomion C, Garnier-Gere PH (2008) Contrasting patterns of selection at Pinus pinaster

Ait. drought stress candidate genes as revealed by genetic differentiation analyses. Mol Biol Evol 25:417-437

Ewing B, Hillier LD, Wendl MC, Green P (1998) Base-calling of automated sequencer traces using phred. I. accuracy assessment. Genome Res 8:175-185

Frewen BE, Chen THH, Howe GT, Davis J, Rohde A, Boerjan W, Bradshaw HD (2000) Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. Genetics 154:837-845

Gailing O and von Wühlisch G (2004) Nuclear markers (AFLPs) and chloroplast microsatellites differ between *Fagus sylvatica* and *F. orientalis*. Silvae Genet 53:105-110

Gailing O, Vornam B, Leinemann L, Finkeldey R (2009) Genetic and genomic approaches to assess adaptive genetic variation in plants: forest trees as a model. Physiol Plantarum 137:509-519

Gao CX and Han B (2009) Evolutionary and expression study of the aldehyde dehydrogenase (ALDH) gene superfamily in rice (Oryza sativa). Gene 431:86-94

Garvin MR, Saitoh K, Gharrett AJ (2010) Application of single nucleotide polymorphisms to non-model species: a technical review. Mol Ecol Resour 10:915-934

Gessler A, Keitel, C, Kreuzwieser J, Matyssek R, Seiler W, Rennenberg H (2007) Potential risks for European beech (*Fagus sylvatica* L.) in a changing climate. Trees-Struct Funct 21:1-11

Glenn TC (2011) Field guide to next-generation DNA sequencers. Mol Ecol Resour 11:759-769

González-Martínez SC, Ersoz E, Brown GR, Wheeler NC, Neale DB (2006a) DNA sequence variation and selection of tag single-nucleotide polymorphisms at candidate genes for drought-stress response in *Pinus taeda* L.. Genetics 172:1915-1926

Guo P, Baum M, Grando S, Ceccarelli S, Bai G, Li R, von Korff M, Varshney RK, Graner A, Valkoun J (2009) Differentially expressed genes between drought-tolerant and drought-sensitive barley genotypes in response to drought stress during the reproductive stage. J Exp Bot 60:3531-3544

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acid S 41:95-98

Helyar SJ, Hemmer-Hansen J, Bekkevold D, Taylor MI, Ogden R, Limborg MT, Cariani A, Maes GE, Diopere E, Carvalho GR, Nielsen EE (2011) Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Mol Ecol Resour 11:123-136

Ingvarsson PK (2004) Nucleotide polymorphism and linkage disequilbrium within and among natural populations of European Aspen (*Populus tremula* L., Salicaceae). Genetics 169:945-953

Ingvarsson PK, Garcia MV, Hall D, Luquez V, Jansson S (2006) Clinal variation in phyB2, a candidate gene for day-length-induced growth cessation and bud set, across a latitudinal gradient in European aspen (*Populus tremula*). Genetics 172:1845-1853

Jimenez JA, Alonso-Ramerez A, Nicolas C (2008) Two cDNA clones (FsDhn1 and FsClo1) up-regulated by ABA are involved in drought responses in *Fagus sylvatica* L. seeds. J Plant Physiol 165:1798-1807

Jones ES, Sullivan H, Bhattramakki D, Smith JSC (2007) A comparison of simple sequence repeat and single nucleotide polymorphism marker technologies for the genotypic analysis of maize (*Zea mays* L.). Theor Appl Genet 115:361-371

Kalinowski ST (2002) How many alleles per locus should be used to estimate genetic distances? Heredity 88:62-65

Kim R and Guo JT (2010) Systematic analysis of short internal indels and their impact on protein folding. BMC Struct Biol 10:24

Kraj W and Sztorc A (2009) Genetic structure and variability of phenological forms in the European beech (*Fagus sylvatica* L.). Ann For Sci 66:203

Krutovsky KV and Neale DB (2005) Forest genomics and new molecular genetic approaches to measuring and conserving adaptive genetic diversity in forest trees. In: Geburek T, Turok J (eds.) Conservation and Management of Forest Genetic Resources in Europe. Arbora Publishers, Zvolen, pp 369-390

Lalagüe H, Fady B, Garnier-Géré P, González-Martínez SC, Lin YC, Oddou-Muratorio S, Sebastiani F, Vendramin GG (2010) Candidate gene variation in common beech (*Fagus sylvatica* L.) along an altitudinal gradient. In: Vinceti B, Neate P (comps.) Conference on "Forest Ecosystem Genomics and Adaptation". San Lorenzo de El Escorial (Madrid), Spain, 9 - 11 June 2010. Book of abstracts. Bioversity International (Rome, Italy) and INIA (Madrid, Spain), pp 242

Li YC, Korol AB, Fahima T, Beiles A, Nevo E (2002) Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. Mol Ecol 11:2453-2465

Librado P and Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25**:**1451-1452

Liu YH, Shi YS, Song YC, Wang TY, Li Y (2010) Characterization of a stress-induced NADP-isocitrate dehydrogenase gene in maize confers salt tolerance in *Arabidopsis*. J Plant Biol 53:107-112

Lu Y, Zhang S, Shah T, Xie C, Hao Z, Li X, Farkhari M, Ribaut JM, Cao M, Rong T, Xu Y (2010) Joint linkage–linkage disequilibrium mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize. Proc Natl Acad Sci USA 107:19585-19590

Morin PA, Luikart G, Wayne RK (2004) SNPs in ecology, evolution and conservation. Trends Ecol Evol 19:208-216

Muleo R, Morini S, Casano S (2001) Photoregulation of growth and branching of plum shoots: Physiological action of two photosystems. In Vitro Cell Dev-Pl 37:609-617

Müller-Starck G and Starke R (1993) Inheritance of isoenzymes in European beech (*Fagus sylvatica* L.). J Hered 84:291-296

Müller-Starck G and Ziehe M (1991) Genetic variation in populations of *Fagus sylvatica* L., *Quercus robur* L., and *Q. petraea* Liebl. in Germany. In: Müller-Starck G and Ziehe M (eds.) Genetic Variation in European Populations of Forest Trees. J. D. Sauerländer's Verlag. Frankfurt am Main

Neale DB (2007) Genomics to tree breeding and forest health. Curr Opin Genet Dev 17:539-544

Novembre J, Johnson T, Bryc K, Kutalik Z, Boyko AR, Auton A, Indap A, King KS, Bergmann S, Nelson MR, Stephens M, Bustamante CD (2008) Genes mirror geography within Europe. Nature 456:98-101

Nyári L (2010) Genetic diversity, differentiation and spatial genetic structure in differently managed adult European beech (*Fagus sylvatica* L.) stands and their regeneration. Forstarchiv 81:156-164

Oddou-Muratorio S, Klein EK, Vendramin GG, Fady B (2011) Spatial vs. temporal effects on demographic and genetic structures: the roles of dispersal, masting and differential mortality on patterns of recruitment in *Fagus sylvatica*. Mol Ecol 20:1997-2010

Olbrich M, Betz G, Gerstner E, Langebartels C, Sandermann H, Ernst D (2005) Transcriptome analysis of ozone-responsive genes in leaves of European beech (*Fagus sylvatica* L.). Plant Biol 7:670-676

Olbrich M, Gerstner E, Bahnweg G, Haberle KH, Matyssek R, Welzl G, Heller W, Ernst D (2010) Transcriptional signatures in leaves of adult European beech trees (*Fagus sylvatica* L.) in an experimentally enhanced free air ozone setting. Environ Pollut 158:977-982

Olson MS, Robertson AL, Takebayashi N, Silim S, Schroeder WR, Tiffin P (2010) Nucleotide diversity and linkage disequilibrium in balsam poplar (*Populus balsamifera*). New Phytol 186:526-536

Quang ND, Ikeda S, Harada K (2008) Nucleotide variation in *Quercus crispula* Blume. Heredity 101:166-174

Pastorelli R, Smulders MJM, Van't Westende WPC, Vosman B, Giannini R, Vettori C, Vendramin GG (2003) Characterization of microsatellite markers in *Fagus sylvatica* L. and *Fagus orientalis* Lipsky. Mol Ecol Notes 3:76-78

Ramanjalu S and Bartels D (2002) Drought- and dessication-induced modulation of gene expression in plants. Plant Cell Environ 25:141-151

Rennenberg H, Seiler W, Matyssek R, Gessler A, Kreuzwieser J (2004) European beech (*Fagus sylvatica* L.) - a forest tree without future in the south of Central Europe? Allg Forst- u J-Ztg 175:210-224

Rozen S and Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds.) Bioinformatics: Methods and Protocols (Methods in Molecular Biology). Humana Press, Totowa, NJ, pp 365-386

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. 2nd ed. Cold Spring Harbor N.Y., Cold Spring Harbor Laboratory

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463-5467

Sathyan P, Newton RJ, Loopstra CA (2005) Genes induced by WDS are differentially expressed in two populations of aleppo pine (*Pinus halepensis*). Tree Genet Genomes 1:166-173

Schlink K (2011) Gene expression profiling in wounded and systemic leaves of *Fagus sylvatica* reveals up-regulation of ethylene and jasmonic acid signalling. Plant Biol 13:445-452

Seeb JE, Carvalho G, Hauser L, Naish K, Roberts S, Seeb LW (2011) Single-nucleotide polymorphism (SNP) discovery and applications of SNP genotyping in nonmodel organism. Mol Ecol Resour 11:1-8

Street NR, Skogstrom O, Sjodin A, Tucker J, Rodriguez-Acosta M, Nilsson P, Jansson S, Taylor G (2006) The genetics and genomics of the drought response in *Populus*. Plant J 48:321-341

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673-4680

Vidalis A (2011) Patterns of nucleotide variation and gene-associated SNP analysis in a *Quercus* spp. forest at isocitrate dehydrogenase genes. Ph.D. Dissertation. Georg-August-University Göttingen

Vornam B, Decarli N, Gailing O (2004) Spatial distribution of genetic variation in a natural beech stand (*Fagus sylvatica* L.) based on microsatellite markers. Conserv Genet 5:561-570

Vornam B, Gailing O, Finkeldey R, Collada C, Guevera M, Soto Á. de María N, González-Martínez S, Díaz L, Alia R, Aranda I, Climent J, Cervera MT, Goicoechea P, Léger V, Eveno E, Derory J, Garnier-Géré P, Kremer A, Plomion C (2007) Naturally occurring nucleotide diversity in

candidate genes for forest tree adaptation: magnitude, distribution and association with quantitative trait variation. GABI – The German Plant Genome Research Program Progress Report 2004 - 2007, pp 116-120

Vornam B, Gailing O, Derory J, Plomion C, Kremer A, Finkeldey R (2011) Characterisation and natural variation of a dehydrin gene in *Quercus petraea* (Matt.) Liebl.. Plant Biol 13:881-887

Wachowiak W, Balk PA, Savolainen O (2009) Search for nucleotide diversity patterns of local adaptation in dehydrins and other cold-related candidate genes in Scots pine (*Pinus sylvestris* L.). Tree Genet Genomes 5:117-132

Yoshiura K, Kinoshita A, Ishida T, Ninokata A, Ishikawa T, Kaname T, Bannai M, Tokunaga K, Sonoda S, Komaki R, Ihara M, Saenko VA, Alipov GK, Sekine I, Komatsu K, Takahashi H, Nakashima M, Sosonkina N, Mapendano CK, Ghadami M, Nomura M, Liang DS, Miwa N, Kim DK, Garidkhuu A, Natsume N, Ohta T, Tomita H, Kaneko A, Kikuchi M, Russomando G, Hirayama K, Ishibashi M, Takahashi A, Saitou N, Murray JC, Saito S, Nakamura Y, Niikawa N (2006) A SNP in the ABCC11 gene is the determinant of human earwax type. Nat Genet 38:324-330

Supplementary material: Candidate genes with detailed information about SNPs and indels

Gene; SNP/ indel No	Position (bp)	Haplotype	Character	Substitution/ indel sequence/ microsatellite motif	Amino acid change
ALDH					
1	66-76	1	non-coding indel	TAATTGTACCA	
2	93	1	non-coding	A/T	
3	127	1	non-coding	C/T	
4	251	2	non-coding	A/G	
5	257	3	non-coding	A/G	
6	316		non-synonymous	A/C	histidine/glutamine/proline
7	317	2	non-synonymous	C/G	histidine/glutamine/proline
8	335	3	synonymous	G/T	
9	395	2	synonymous	A/T	
APX1					
1	182		synonymous	C/G	
2	272		synonymous	C/T	
3	294		non-coding	C/T	
4	302	1	non-coding	A/T	
5	302-344		non-coding high variable indel/ SNP site		
6	344	1	non-coding	A/G	
7	347-348	1	non coding indel	AA	
8	352	1	non-coding	A/T	
9	525	2	non-coding	C/G	
10	545		non-coding	A/T	
11	554		non-coding	A/G	
12	560	2	non-coding	C/T	
13	614-615		non-coding indel/ microsatellite motif	T	
14	706	3	non-coding	A/G	
15	809	3	non-coding	G/T	
16	822	3	non-coding	A/T	
17	872	4	non-coding	C/T	
18	1000-1006		non-coding indel	AATTAAG	
19	1297	4	synonymous	C/T	
APX4, **Part1**					
1	63		non-coding	G/T	
2	235	1	non-coding	A/G	
3	258-288	1	non-coding indel		
4	307		non-coding	A/G	
5	426		non-coding	A/G	
6	517	+	non-coding	A/G	

7	574		non-coding	A/G	

APX4, Part2

1	38	1	non-coding	C/T	
2	73		non-coding	G/T	
3	78	1	non-coding	A/T	
4	177	2	non-coding	C/T	
5	178-180	2	non-coding indel	TCC	
6	208	1	synonymous	C/T	
7	386	3	non-coding	C/T	
8	450	3	non-coding	G/T	
9	508	1	non-synonymous	C/T	leucine/valine
10	687		non-coding	C/G	
11	692	+	non-coding	A/T	
12	693	+	non-coding	A/T	

Cry

1	95-98		non-coding indel/ microsatellite motif	CA	
2	228		non-coding	A/G	

Dhn

1	68		non-synonymous	C/G	aspartic acid/histidine
2	325	1	non-coding	A/T	
3	455	1	non-synonymous	C/G	histidine/glutamine

ERD, Part 1

1	274		non-coding	A/G	
2	308		non-coding	A/C	

ERD, Part 2

1	21-26		coding indel/ microsatellite motif	AGA	

GPX

1	117		non-synonymous	C/T	phenylalanine/leucine

IDH , Part 1

1	83		non-coding	C/T	
2	94		non-coding	A/C	
3	120		non-coding	C/T	
4	156		non-coding	A/T	
5	270		synonymous	C/T	
6	309		synonymous	A/G	
7	400	1	non-coding	A/C	
8	430	1	non-coding	A/G	
9	440		non-coding	G/T	
10	604		non-coding	A/G	

IDH,

Part 2

1	43	non-synonymous	A/G	valine/isoleucine
2	106	non-coding	C/G	
3	202	non-coding	C/T	
4	278	synonymous	A/G	
5	335-336	non-coding indel/ microsatellite	T	

PhyB

1	25	synonymous	A/G	
2	196	synonymous	A/G	

CHZFP

1	390-395	coding indel/ microsatellite	CAG	

Appendix 2: Climate diagrams for the sampling areas.

Unterlüß (113 m)
Zeitraum: 1971-2000   8.5 °C   765.9 mm

Harz (458 m)
Zeitraum: 1971-2000   7.2 °C   1170 mm
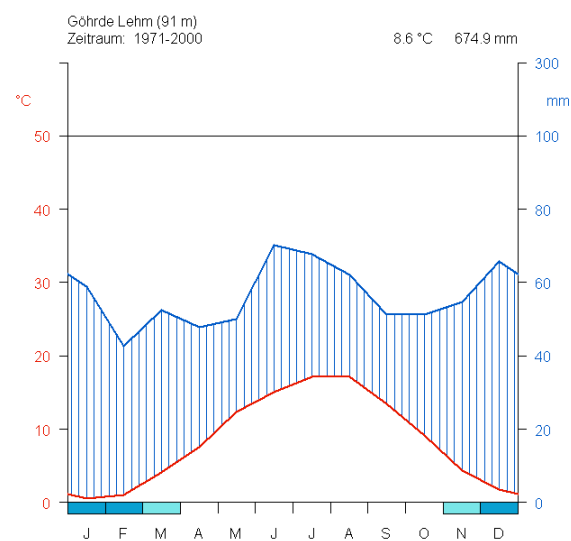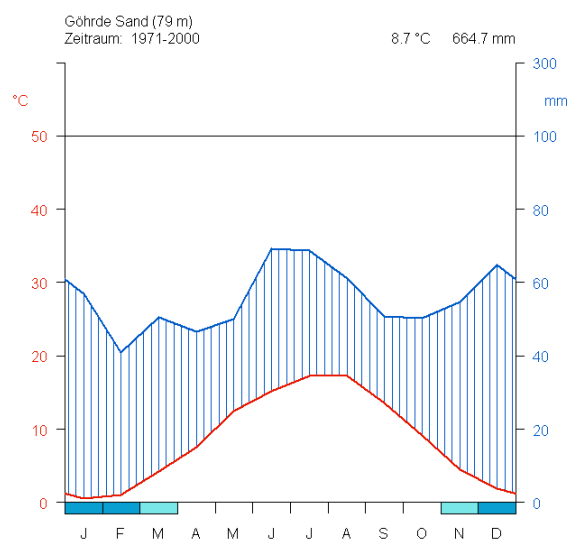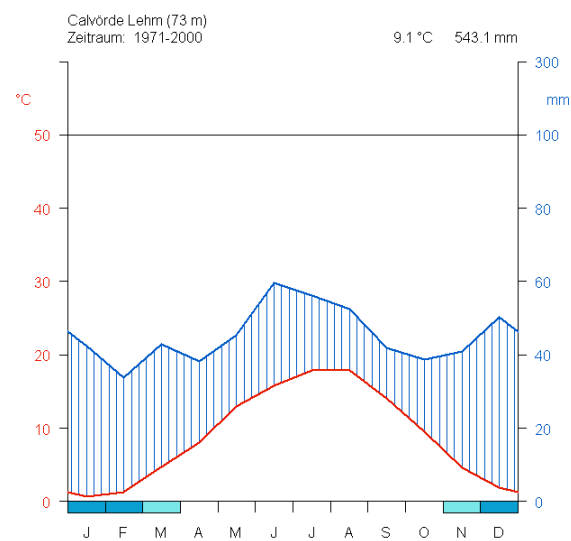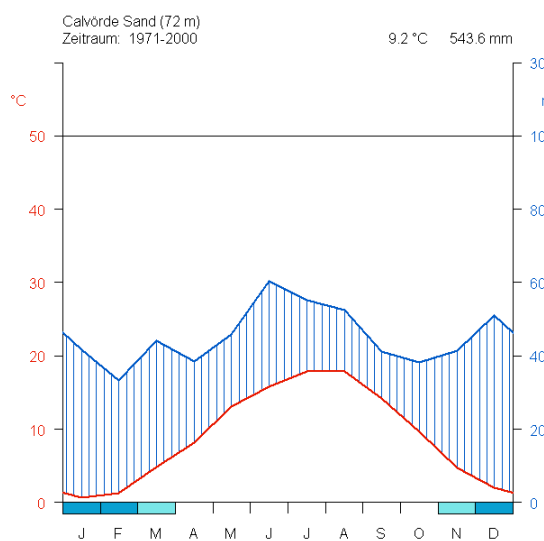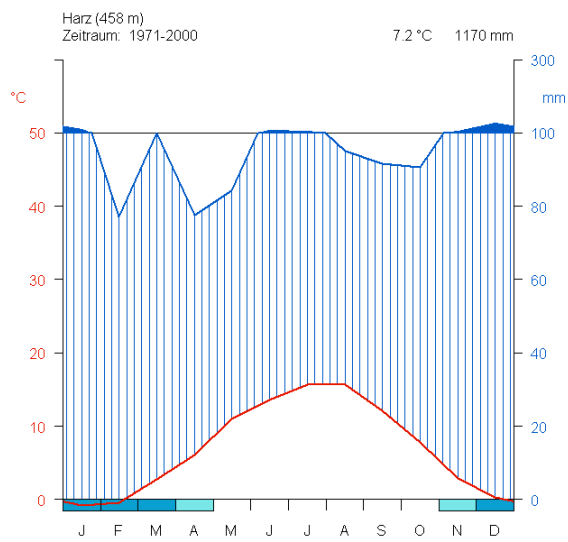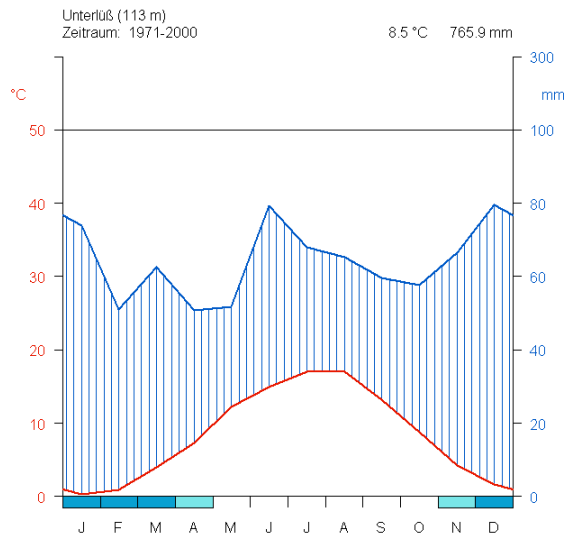
Appendix 3: Locuswise allele frequencies for all microsatellite alleles (N= sample size by populations, CL= Calvörde loam, CS= Calvörde sand, GL= Göhrde loam, GS= Göhrde sand, UL= Unterlüß loam, US= Unterlüß sand, Ha= Bad Grund).

Locus	Allele	CL	CS	GL	GS	UL	US	Ha
sfc0018	**N**	101	104	100	103	100	99	99
	151	0.114	0.034	0.010	0.044	0.025	0.000	0.020
	158	0.020	0.024	0.015	0.029	0.035	0.000	0.015
	162	0.005	0.005	0.050	0.010	0.005	0.030	0.010
	164	0.059	0.019	0.035	0.044	0.020	0.096	0.035
	166	0.079	0.053	0.060	0.058	0.055	0.061	0.030
	168	0.000	0.000	0.000	0.000	0.010	0.000	0.000
	170	0.554	0.466	0.365	0.393	0.480	0.465	0.475
	172	0.035	0.111	0.070	0.097	0.145	0.141	0.086
	174	0.020	0.101	0.195	0.097	0.050	0.040	0.086
	176	0.059	0.144	0.150	0.160	0.140	0.101	0.146
	178	0.054	0.043	0.045	0.068	0.035	0.066	0.091
	180	0.000	0.000	0.000	0.000	0.000	0.000	0.005
	188	0.000	0.000	0.005	0.000	0.000	0.000	0.000
		CL	CS	GL	GS	UL	US	Ha
sfc0161	**N**	101	104	100	103	100	99	99
	106	0.054	0.144	0.095	0.087	0.045	0.056	0.061
	108	0.000	0.005	0.000	0.010	0.005	0.000	0.000
	110	0.010	0.000	0.000	0.000	0.000	0.010	0.000
	112	0.000	0.000	0.010	0.015	0.000	0.000	0.000
	114	0.050	0.216	0.140	0.136	0.180	0.207	0.162
	116	0.000	0.014	0.070	0.005	0.005	0.015	0.010
	118	0.015	0.000	0.005	0.010	0.000	0.005	0.005
	122	0.000	0.010	0.020	0.005	0.005	0.005	0.020
	124	0.465	0.245	0.220	0.223	0.440	0.359	0.303
	126	0.243	0.202	0.310	0.330	0.175	0.177	0.222
	128	0.015	0.024	0.065	0.049	0.030	0.005	0.045
	130	0.040	0.019	0.010	0.019	0.035	0.030	0.030
	132	0.010	0.010	0.000	0.000	0.000	0.005	0.000
	134	0.069	0.063	0.045	0.063	0.020	0.051	0.045
	136	0.015	0.010	0.000	0.010	0.000	0.005	0.051

		CL	CS	GL	GS	UL	US	Ha
	138	0.005	0.000	0.000	0.010	0.005	0.020	0.005
	140	0.000	0.005	0.000	0.000	0.015	0.015	0.010
	142	0.010	0.024	0.010	0.019	0.035	0.025	0.030
	144	0.000	0.005	0.000	0.005	0.005	0.000	0.000
	146	0.000	0.000	0.000	0.005	0.000	0.000	0.000
	150	0.000	0.005	0.000	0.000	0.000	0.010	0.000
		CL	**CS**	**GL**	**GS**	**UL**	**US**	**Ha**
sfc1063	**N**	101	104	100	103	100	98	98
	187	0.040	0.024	0.060	0.039	0.020	0.015	0.046
	193	0.109	0.192	0.165	0.252	0.270	0.240	0.204
	195	0.045	0.091	0.025	0.073	0.105	0.163	0.122
	197	0.347	0.332	0.370	0.282	0.185	0.255	0.276
	199	0.153	0.063	0.270	0.102	0.090	0.097	0.092
	201	0.050	0.163	0.080	0.175	0.200	0.148	0.097
	203	0.015	0.000	0.000	0.000	0.000	0.000	0.000
	205	0.050	0.096	0.020	0.058	0.085	0.010	0.107
	207	0.025	0.029	0.010	0.010	0.020	0.041	0.036
	209	0.149	0.005	0.000	0.010	0.000	0.026	0.020
	211	0.020	0.005	0.000	0.000	0.025	0.005	0.000
		CL	**CS**	**GL**	**GS**	**UL**	**US**	**Ha**
sfc1143	**N**	101	104	100	103	100	99	99
	106	0.000	0.005	0.000	0.015	0.000	0.000	0.005
	110	0.010	0.000	0.000	0.000	0.005	0.005	0.000
	112	0.000	0.000	0.000	0.005	0.000	0.000	0.005
	114	0.025	0.063	0.045	0.053	0.100	0.081	0.056
	116	0.114	0.063	0.085	0.087	0.025	0.025	0.035
	118	0.005	0.014	0.055	0.049	0.025	0.035	0.035
	120	0.030	0.014	0.020	0.015	0.020	0.005	0.015
	122	0.050	0.014	0.000	0.005	0.005	0.000	0.020
	124	0.084	0.111	0.010	0.092	0.150	0.177	0.116
	126	0.252	0.361	0.410	0.354	0.390	0.495	0.389
	128	0.327	0.308	0.200	0.252	0.195	0.136	0.247
	130	0.099	0.029	0.150	0.053	0.040	0.020	0.020
	132	0.005	0.019	0.015	0.019	0.040	0.020	0.056
	136	0.000	0.000	0.010	0.000	0.005	0.000	0.000

		CL	CS	GL	GS	UL	US	Ha
GOT066	**N**	101	104	99	103	100	100	99
	235	0.000	0.000	0.000	0.000	0.000	0.000	0.010
	238	0.045	0.053	0.121	0.049	0.040	0.045	0.071
	241	0.951	0.923	0.823	0.932	0.950	0.885	0.899
	243	0.000	0.000	0.000	0.000	0.000	0.000	0.020
	244	0.005	0.019	0.020	0.019	0.005	0.000	0.000
	246	0.000	0.005	0.035	0.000	0.005	0.070	0.000
		CL	**CS**	**GL**	**GS**	**UL**	**US**	**Ha**
FIR065	**N**	101	104	100	103	100	100	98
	168	0.356	0.361	0.305	0.432	0.250	0.250	0.332
	171	0.356	0.288	0.340	0.262	0.485	0.600	0.321
	177	0.163	0.207	0.110	0.136	0.095	0.050	0.173
	180	0.124	0.144	0.245	0.170	0.170	0.100	0.173
		CL	**CS**	**GL**	**GS**	**UL**	**US**	**Ha**
FIR004	**N**	101	104	100	103	100	100	99
	143	0.104	0.135	0.115	0.102	0.150	0.075	0.096
	145	0.475	0.688	0.730	0.544	0.565	0.675	0.535
	147	0.010	0.010	0.005	0.053	0.055	0.090	0.056
	148	0.000	0.000	0.000	0.010	0.000	0.000	0.000
	149	0.050	0.019	0.005	0.019	0.040	0.005	0.263
	150	0.307	0.130	0.105	0.228	0.150	0.145	0.040
	152	0.000	0.000	0.000	0.000	0.005	0.005	0.000
	158	0.000	0.000	0.000	0.005	0.000	0.000	0.000
	162	0.040	0.000	0.035	0.010	0.030	0.005	0.000
	164	0.015	0.019	0.005	0.029	0.005	0.000	0.010
		CL	**CS**	**GL**	**GS**	**UL**	**US**	**Ha**
FS 3-04	**N**	101	104	100	103	100	100	99
	195	0.059	0.010	0.010	0.019	0.015	0.020	0.010
	201	0.876	0.760	0.775	0.752	0.850	0.855	0.818
	204	0.040	0.202	0.160	0.214	0.120	0.100	0.157
	207	0.025	0.029	0.055	0.015	0.015	0.025	0.015
		CL	**CS**	**GL**	**GS**	**UL**	**US**	**Ha**
mfs11	**N**	101	104	99	103	100	100	99
	128	0.015	0.000	0.000	0.000	0.000	0.000	0.000
	132	0.163	0.188	0.141	0.184	0.185	0.185	0.126

134	0.465	0.457	0.399	0.466	0.410	0.385	0.510
136	0.035	0.010	0.071	0.019	0.125	0.085	0.005
138	0.005	0.005	0.000	0.005	0.005	0.000	0.000
140	0.129	0.125	0.182	0.160	0.115	0.110	0.192
142	0.129	0.154	0.131	0.126	0.120	0.145	0.091
144	0.005	0.038	0.035	0.010	0.015	0.005	0.000
146	0.054	0.024	0.040	0.029	0.020	0.065	0.071
148	0.000	0.000	0.000	0.000	0.005	0.015	0.000
150	0.000	0.000	0.000	0.000	0.000	0.005	0.000
152	0.000	0.000	0.000	0.000	0.000	0.000	0.005

Appendix 4: List of SNPs with significant linkage disequilibrium values for every population.

Calvörde sand			Göhrde sand			Göhrde sand seedlings		
SNP1	SNP2	Significance value	SNP1	SNP2	Significance value	SNP1	SNP2	Significance value
IDH2	IDH1	***	IDH2	IDH1	***	IDH2	IDH1	***
IDH3	IDH1	***	IDH3	IDH1	***	IDH3	IDH1	***
IDH4	IDH1	***	IDH4	IDH1	***	IDH4	IDH1	***
IDH4	IDH2	***	IDH4	IDH2	***	IDH4	IDH2	***
IDH4	IDH3	***	IDH4	IDH3	***	IDH4	IDH3	***
ALDH1	IDH2	*	ALDH2	ALDH1	***	ALDH2	ALDH1	***
ALDH2	IDH2	*	ALDH3	ALDH2	***	ALDH3	ALDH1	**
ALDH2	ALDH1	***	ALDH4	ALDH3	***	ALDH3	ALDH2	**
ALDH3	IDH1	**	dhn1	dhn2	*	ALDH4	IDH1	*
ALDH3	IDH2	**	APX4.1	ALDH3	*	ALDH4	IDH4	*
ALDH3	IDH4	**	APX1.1	IDH3	**	ALDH4	ALDH3	***
ALDH3	ALDH2	***	APX1.2	APX1.1	***	dhn1	dhn2	*
ALDH4	IDH1	*	PhyB	APX4.1	*	APX1.2	dhn1	*
ALDH4	IDH2	**	ERD	APX1.1	*	APX1.2	APX1.1	***
ALDH4	IDH4	*				ERD	IDH2	*
ALDH4	ALDH2	**				ERD	dhn1	*
ALDH4	ALDH3	***				ERD	APX4.1	*
dhn1	dhn2	*						
APX1.1	IDH2	*						

Calvörde sand (continued)			Bad Grund seedlings			Unterlüß sand		
SNP1	SNP2	Significance value	SNP1	SNP2	Significance value	SNP1	SNP2	Significance value
APX1.1	ALDH2	**	IDH2	IDH1	***	IDH2	IDH1	***
APX4.2	ALDH1	**	IDH3	IDH1	***	IDH3	IDH1	***
APX4.2	dhn1	*	IDH4	IDH1	***	IDH4	IDH1	***
APX1.2	IDH2	**	IDH4	IDH2	***	IDH4	IDH2	***
APX1.2	ALDH2	*	IDH4	IDH3	***	IDH4	IDH3	***
APX1.2	ALDH3	*	ALDH2	ALDH1	***	ALDH2	ALDH1	***
APX1.2	APX1.1	***	ALDH3	ALDH2	***	ALDH3	ALDH2	***
PhyB	IDH1	*	ALDH4	ALDH2	*	ALDH4	ALDH2	**
PhyB	IDH3	*	ALDH4	ALDH3	***	ALDH4	ALDH3	***
PhyB	IDH4	*	APX4.1	dhn2	*	dhn2	IDH2	*
ERD	dhn2	*	APX1.2	APX1.1	***	dhn1	dhn2	**
			ERD	APX4.2	**	APX1.1	dhn1	*
			GPX	APX4.2	*	APX4.2	IDH1	*
						APX4.2	IDH4	*
						APX1.2	IDH3	**
						APX1.2	APX1.1	***
						APX1.2	APX4.2	*
						GPX	IDH3	*

*= p<0.05, **= p<0.01, ***= p<0.001

Appendix 5a: Chi-square test for all SNPs for the populations along the precipitation gradient.

	Calvörde sand – Unterlüß sand				Göhrde sand – Unterlüß sand		
SNP	Chi-square	Degrees of freedom	Significance value	SNP	Chi-square	Degrees of freedom	Significance value
ALDH 1	4.427	2	ns	*ALDH* 1	7.145	2	ns
ALDH 2	4.965	2	ns	*ALDH* 2	2.446	2	ns
ALDH 3	3.729	2	ns	*ALDH* 3	2.869	2	ns
ALDH 4	1.728	2	ns	*ALDH* 4	0.351	1	ns
IDH 1	37.768	2	***	*IDH* 1	19.31	2	***
IDH 2	10.520	2	**	*IDH* 2	1.635	1	ns
IDH 3	8.643	2	*	*IDH* 3	1.351	2	ns
IDH 4	34.883	2	***	*IDH* 4	19.31	2	***
*APX*1.1	11.966	2	**	*APX*1.1	8.702	2	*
*APX*1.2	20.306	2	**	*APX*1.2	3.222	2	ns
*APX*4.1	0.004	2	ns	*APX*4.1	1.872	2	ns
*APX*4.2	15.771	2	***	*APX*4.2	3.133	2	ns
dhn 1	9.377	2	**	*dhn* 1	3.365	2	ns
dhn 2	1.256	2	ns	*dhn* 2	1.994	2	ns
ERD	1.045	2	ns	*ERD*	5.107	2	ns
GPX	3.971	1	*	*GPX*	1.642	1	ns
PhyB	16.858	2	***	*PhyB*	17.348	2	***

Appendix 5b: Chi-square test for all SNPs for the populations along the precipitation gradient and the populations from the drought stress experiment comparing damaged with undamaged/slightly damaged seedlings.

	Calvörde sand – Göhrde sand			Damaged – undamaged/slightly damaged			
SNP	Chi-square	Degrees of freedom	Significance value	SNP	Chi-square	Degrees of freedom	Significance value
ALDH 1	1.818	2	ns	*ALDH* 1	0.75	2	ns
ALDH 2	1.523	2	ns	*ALDH* 2	0.355	2	ns
ALDH 3	0.891	2	ns	*ALDH* 3	1.006	2	ns
ALDH 4	3.115	2	ns	*ALDH* 4	0.272	1	ns
IDH 1	4.772	2	ns	*IDH* 1	0.843	2	ns
IDH 2	7.628	2	*	*IDH* 2	4.276	2	ns
IDH 3	4.179	2	ns	*IDH* 3	2.162	2	ns
IDH 4	3.570	2	ns	*IDH* 4	0.567	2	ns
*APX*1.1	0.678	2	ns	*APX*1.1	1.426	2	ns
*APX*1.2	2.108	2	ns	*APX*1.2	0.933	2	ns
*APX*4.1	1.830	2	ns	*APX*4.1	6.173	2	*
*APX*4.2	7.551	2	*	*APX*4.2	0.847	2	ns
dhn 1	2.800	2	ns	*dhn* 1	0.849	2	ns
dhn 2	5.169	2	ns	*dhn* 2	2.023	2	ns
ERD	1.800	2	ns	*ERD*	0.526	2	ns
GPX	0.547	1	ns	*GPX*	3.047	2	ns
PhyB	0.135	2	ns	*PhyB*	1.255	2	ns

Appendix 6: Allele and genotype frequencies of all SNPs for the damaged and not damaged/slightly damaged seedlings (N= sample size by populations; *and slightly damaged).

SNP	Allele/ genotype	Damaged	Not damaged*	SNP	Allele/ genotype	Damaged	Not damaged*
ALDH 1	N	100	100	*IDH* 2	N	100	100
	T	0.290	0.315		A	0.155	0.175
	C	0.710	0.685		G	0.845	0.825
	TT	0.110	0.100		AA	0.000	0.040
	TC	0.370	0.430		AG	0.310	0.270
	CC	0.520	0.470		GG	0.690	0.690
ALDH 2	N	100	100	*IDH* 3	N	100	100
	A	0.380	0.400		C	0.830	0.845
	C	0.620	0.600		G	0.170	0.155
	AA	0.140	0.170		CC	0.730	0.720
	AC	0.480	0.470		CG	0.200	0.250
	CC	0.380	0.360		GG	0.070	0.030
ALDH 3	N	100	100	*IDH* 4	N	100	100
	C	0.910	0.900		A	0.470	0.485
	G	0.090	0.100		G	0.530	0.515
	CC	0.820	0.810		AA	0.260	0.250
	CG	0.180	0.180		AG	0.420	0.470
	GG	0.000	0.010		GG	0.320	0.280
ALDH 4	N	100	100	*APX*1.1	N	100	100
	A	0.955	0.965		T	0.785	0.820
	T	0.045	0.035		C	0.215	0.180
	AA	0.910	0.930		TT	0.620	0.660
	AT	0.090	0.070		CT	0.330	0.320
	TT	0.000	0.000		CC	0.050	0.020
IDH 1	N	100	100	*APX*1.2	N	100	100
	T	0.470	0.490		T	0.235	0.195
	C	0.530	0.510		G	0.765	0.805
	TT	0.260	0.250		TT	0.060	0.040
	TC	0.420	0.480		TG	0.350	0.310
	CC	0.320	0.270		GG	0.590	0.650

SNP	Allele/genotype	Damaged	Not damaged*	SNP	Allele/genotype	Damaged	Not damaged*
*APX*4.1	N	100	100	*ERD*	N	100	100
	T	0.830	0.905		A	0.640	0.655
	G	0.170	0.095		G	0.360	0.345
	TT	0.700	0.840		AA	0.430	0.470
	TG	0.270	0.130		AG	0.420	0.370
	GG	0.030	0.030		GG	0.150	0.160
*APX*4.2	N	99	99	*GPX*	N	100	100
	C	0.768	0.803		T	0.890	0.900
	G	0.232	0.197		C	0.110	0.100
	CC	0.606	0.667		TT	0.780	0.820
	CG	0.313	0.273		CT	0.220	0.160
	GG	0.081	0.061		CC	0.000	0.020
dhn 1	N	100	100	*PhyB*	N	100	100
	C	0.365	0.365		A	0.435	0.380
	G	0.635	0.635		G	0.565	0.620
	CC	0.140	0.110		AA	0.190	0.150
	CG	0.450	0.510		AG	0.490	0.460
	GG	0.410	0.380		GG	0.320	0.390
dhn 2	N	100	100				
	C	0.935	0.915				
	G	0.065	0.085				
	CC	0.870	0.850				
	CG	0.130	0.130				
	GG	0.000	0.020				

Curriculum Vitae

Sarah Seifert

Personal Data

Date of birth   28.12.1983

Place of birth   Northeim, Germany

Contact   sarah-seifert@gmx.de; sseifer@gwdg.de

Education

1997 - 2003	Gymnasium Corvinianum Northeim
	Abitur, overall grade „2,0"
Oct. 2003 - Nov. 2008	Georg-August-University Göttingen, Faculty of Biology
	Major subjects: Anthropology, Microbiology, Psychology
	Diploma thesis: "Cleaning of contaminated material relevant to biomolecular analysis without loss of substance with the aid of ultrasound"
Dec. 2008 - Jan. 2009	Research assistant in the section "Forest Genetics and Forest Tree Breeding", Georg-August-University Göttingen
Feb. 2009 – Jan. 2012	PhD student in the section "Forest Genetics and Forest Tree Breeding", Georg-August-University Göttingen

Conference abstracts and posters

K.C. R, **Seifert S**, Gailing O, Finkeldey R (2011) Forest management does not reduce genetic diversity of adult beech (Fagus sylvatica L.) stands but changes the spatial genetic structure. The 9th International Beech Symposium: Ecology and Silviculture of Beech (Dresden, 12.09-17.09.2011)

Seifert S, Müller M, Vornam B, Finkeldey R (2011) Adaptability of trees to climate change – genetic variation of beech (Fagus sylvatica L.) along an environmental gradient. Tagung "Veränderungen genetischer Variation in Raum und Zeit – Anpassungsprozesse im Klimawandel" (Teisendorf, 30.05-01.06.2011)

Seifert S, Vornam B, Finkeldey R (2010) Genetic variation of beech (*Fagus sylvatica* L.) along an environmental gradient. International workshop: Adaptation of perennial plants to episodic drought: impact on ecosystems (Göttingen, 06.10-08.10.2010). poster

Seifert S, Vornam B, Gailing O, Finkeldey R (2010) Charakterisierung adaptiver genetischer Variation bei Fagaceae. In Strunk M (ed.) Forstwissenschaftliche Tagung 2010, Forstwissenschaften: Grundlage nachhaltiger Waldbewirtschaftung. Georg-August-Universität Göttingen, Germany, 22 – 24 September 2010. Göttingen, Niedersachs : Cuvillier, E. Pp. 17. ISBN: 978-3-86955-482-2

K.C. R, **Seifert** S, Prinz K, Gailing O, Finkeldey R (2010) Räumliche Dynamik intraspezifischer Variation bei der Rotbuche *(Fagus sylvatica* L.). In Strunk M (ed.) Forstwissenschaftliche Tagung 2010, Forstwissenschaften: Grundlage nachhaltiger Waldbewirtschaftung. Georg-August-Universität Göttingen, Germany, 22 – 24 September 2010. Göttingen, Niedersachs : Cuvillier, E. Pp. 170. ISBN: 978-3-86955-482-2. poster

Language skills

German (native language), english