Adaptive genetic structure in ecologically marginal populations of European Silver Fir (*Abies alba* MILL) at the south-western Mediterranean pre-Alps of France.

Landscape genetics using novel genetic markers to infer local adaptation and divergent selection in natural forest populations.

by

Anna Maria Roschanski

to

Department of Conservation Biology Philipps University of Marburg

in partial fulfillment of the requirements for the degree of doctor of natural science in the subject of biology

Philipps University of Marburg

Hessen, Germany

Submitted in February 2016

Vom Fachbereich 17, Biologie der Philipps-Universität Marburg als Dissertation am 01.02.2016 angenommen.

Erstgutachter: Frau Prof. Dr. Birgit Ziegenhagen
Zweitgutachter: Herr Dr. habil. Bruno Fady

Tag der mündlichen Prüfung am 15.04.2016

Table of content

Publications and manuscripts of the cumulative part

Chapter I

Roschanski AM, Fady B, Ziegenhagen B, Liepelt S (2013) **Annotation and Re-Sequencing of Genes from De Novo Transcriptome Assembly of *Abies alba* (Pinaceae).** *Applications in Plant Sciences*, 1, 1–8.

Chapter II

Anna M. Roschanski, Bruno Fady, Birgit Ziegenhagen and Sascha Liepelt (2015) **Transcriptome sequencing, SNP calling and assayed genotyping in *Abies alba* Mill.** (unpublished manuscript, Jan-2016)

Chapter III

Anna M. Roschanski, Katalin Csilléry, Sascha Liepelt, Sylvie Oddou-Muratorio, Birgit Ziegenhagen, Frédéric Huard, Kristian K. Ullrich, Dragos Postolache, Giovanni G. Vendramin and Bruno Fady (2016) **Evidence of divergent selection for drought and cold tolerance at landscape and local scales in *Abies alba* Mill. in the French Mediterranean Alps**. *Molecular Ecology*, 25.3, 776-794

List of abbreviations and units with explanations in parentheses

Abbreviations and units used in synopsis

%	percentage (ratio as a fraction of one hundred)
€	Euro (European currency)
°C	degrees Celsius (unit of temperature)
°E	degrees East (unit of longitude)
°N	degrees North (unit of latitude)
AMOVA	Analysis of MOlecular VAriance
ATP	Adenosine triphosphate (co-enzyme, donor of chemical energy)
BLAST	Basic Local Alignment Search Tool (bioinformatic algorithm to detect sequence similarities)
bp	basepair (two complement nucleotides of the DNA molecule)
ca.	circa (approximately)
DNA	Deoxyribonucleic acid (major molecule of the genome)
e.g.	for example (*exempli gratia* in Latin)
EU-EvolTree	Network of scientific excellence that addresses global issues of European forestry
et al.	and others (et alii in Latin)
etc	and so forth (et cetera in Latin)
FAM-dye	Fluorescein amidite (fluorescent lable used in DNA sequencing)
FAO	Food and Agriculture Organization of the United Nations
FRET	Fluorescence Resonant Energy Transfer (quenched cassette that emits fluorescence light when quencher is released, used in KASP genotyping)
Fst	Fixation index (measure of genetic population differentiation)
Gbp	Gigabasepairs (one billion basepairs, see also: base pairs)
GBS	Genotyping By Sequencing (genotyping method where genomic DNA is discriminatorily fragmented using methylation-sensitive restriction endonucleases, sequenced and subjected to SNP calling)
HEX-dye	Hexachloro-fluorescein (fluorescent lable used in DNA sequencing)
i.e.	that is (id est in Latin)
IBD	Isolation By Distance
IBE	Isolation By Environment
k	kilo (one thousand)
KASP	PCR-based genotyping assay (trade mark of LGC Genomics (Middlesex, UK)
km	kilometer (distance)
km2	square kilometer (area)
LFMM	Latent Factor Mixed Model (statictical algorithm for gene-environment associations)
Mb	Megabasepairs (one million basepairs, see also: base pairs)
MACE	Massive Analysis of cDNA Ends (sequencing technology where only 3'ends of cDNAs are sequenced so that RNA complexity is reduced)
Ne	effective population size (number of individuals of a population that contribute to the offspring generation)
PC	Principal Component (orthogonal eigenvector of the covariance matrix of a set of observations)
PCR	Polymerase Chain Reaction (method of DNA amplification)
pg 2C-1	picogram DNA in a diploid cell nucleus

Qst	genetic differentiation of a quantitative trait among populations
QTL	Quantitative Trait Locus (section of DNA correlating with variation at phenotypic, quantitative trait)
QTN	Quantitative Trait Nucleotides (allele correlating with variation at phenotypic, quantitative trait)
RAD-seq	restriction-site-associated sequencing (sequencing method where genomic DNA is discriminatorily fragmented using a single restriction endonuclease)
RCP	Representative Concentration Pathways (greenhouse gas concentration used for climate models in the fifth IPCC Assessment Report)
SNP	Single Nucleotide Polymorphism (is a variation in a single nucleotide occuring at a specific position in the genome)
UNECE	United Nations Economic Commission for Europe
USD	United States Dollar (currency of the United States of America)

Abbreviations and units used in chapter II

BLAST2GO	see BLAST; searches the gene ontology database where nucleotide sequences are assigned to functional annotations using a nucleotide query
BLASTx	see BLAST; searches a protein databases using a translated nucleotide query
cDNA	complementary DNA (double-stranded DNA reversely transcribed from messenger RNA, the total cDNA in a cell represents the transcriptome of the cell)
EST	Expressed Sequence Tag (fragment of a cDNA sequence)
GO	Gene Ontology (standardized, functional information about gene sequences)
KEGG	Kyoto Encyclopedia of Genes and Genomes (reference resource for gene and protein annotation)
N50 contig size	minimum contig size in sequence assemblies in which half of all sequenced bases are assembled
sd	standard deviation quantifies the amount of variation of a data set
tBLASTn	see BLAST; searches a translated nucleotide databases using a protein query

Zusammenfassung

Zu der Gruppe der Koniferen gehören einige ökonomisch und ökologisch wichtige Forstbäume. Trotzdem ist diese Gruppe genomisch nur wenig untersucht. Gründe dafür sind die langen Generationszyklen, das hohe Maß an Heterozygosität sowie die großen Genome von Koniferen, die den Zugang zu genomischen Ressourcen erschweren. Die Weißtanne (*Abies alba* MILL) ist weit über montane Habitate des temperaten und mediterranen Europas verbreitet. Die Art hat aufgrund ihres hochwertigen Holzes und ihrer bestands-stabilisierenden Funktion eine hohe Bedeutung in der Forstwirtschaft. Mittels next-generation-sequencing wurde das erste Weißtannentranskriptom sequenziert und annotiert. Das Annotationsprotokoll führte zur Identifizierung von 460 Kandidatengenen, denen eine Rolle in der Stressreaktion zukommt. Zusätzlich wurden die Transkriptome zwölf weiterer Tannen sequenziert und einem SNP-calling unterzogen, bei dem durchschnittlich 24580 SNPs pro Individuum detektiert wurden. Für diejenigen SNPs, die sich auf den annotierten Kandidatengenen befanden, wurde ein KASP Genotypisierungswerkzeug entworfen.

Im mediterranen Becken kommt die Tanne in isolierten, vergleichsweise kleinen Populationen vor. Diese sind aktuell stark von Mortalität betroffen, was auf ausgeprägte Trockenperioden zurückgeführt wird. Im Zuge des Klimawandels wird eine noch stärkere Zunahme extremer Trockenheit erwartet, die die Existenz mediterraner Tannenpopulationen bedrohen könnte. Eben diese Populationen sind jedoch von besonderer Bedeutung für das adaptive Potential der gesamten Art, da sie schon jetzt über genetische Anpassungen an Trockenheit verfügen könnten. Zur Zeit entspannt sich eine Debatte über die Praktikabilität der sog. ‚assisted migration', d. h. über die Möglichkeit und Sinnhaftigkeit der Transplantation trockenadaptierter Tannen in nördlichere Populationen, in denen sie als Allele-Donoren fungieren sollen. Zur Erhellung der Debatte bedarf es erweiterter Kenntnisse über die tatsächliche lokale Anpassung mediterraner Tannen und ihrer molekularen Grundlage sowie über die Art der Umweltfaktoren, druch die die lokale Anpassung hervorgerufen wird.

Mit den Methoden der Landschaftsgenetik wurden 376 Bäume an 267 SNP loci untersucht, die über 175 Kandidatengene verteilt waren. Die Bäume stammten von hoch und tief gelegenen Standorten, welche sich über vier Berge der süd-westlichen Voralpen in Frakreich entlang eines 170 km langen ost-west gerichteten Transekts erstreckten. Das Ergebnis einer Bayesischen Clusteranalyse zeigte, dass das übergeordnete genetische Muster einer isolation-by-distance Struktur folgt, welche wahrscheinlich von der Rekolonisationsbewegung der Art im Postglazial herrührt. F_{st}-outlier Tests förderten 16 SNPs zu Tage, die unter bidirektionaler Selektion stehen. Acht SNP loci waren signifikant mit Wintertrockenheit assoziiert. Korrelationstests der genetischen Distanz einiger Outlierloci und der Distanz von Umweltparametern ergaben, dass isolation-by-environment zwischen hoch- und tiefgelegenen Standorten auftritt. Die vergleichende Analyse von Q_{st} und F_{st} Werten, zeigte das Auftreten von adaptiver Divergenz in Blattaustrieb und Wachstumsrate am kältesten und feuchtesten Standort.

Die Arbeit beschreibt das komplexe Muster von Populationen an ökologischen Marginalstandorten, gleichsam geformt durch Demographie und Adaption. Die Arbeit betont, dass lokale Adaption im Bezug auf die unter Selektion stehenden Gene und die wirksamen Umweltfaktoren für jeden Standort spezifisch ist. So muss die Frage, ob eine Population zur Nutzung für assisted migration geeignet ist, für jeden Einzelfall neu beantwortet werden. Dieses ist in höchstem Maße unpraktikabel. Stattdessen erscheint der in-situ Schutz von Marginalpopulationen als die beste Möglichkeit die Fähigkeit der Tanne zur evolutiven Weiterentwicklung zu sichern und so ihr Potential sich den Herausforderungen des Klimawandels anzupassen zu stärken.

Abstract

Coniferous trees are of major economic and ecologic importance. Yet, they are genomically poorly studied, because their long generation times, their high levels of heterozygosity and their huge genomes impede the access to conifer genomic resources. The European silver fir (*Abies alba* MILL) is widely distributed across diverse habitats in mountainous areas of temperate Europe. The species is of high value for forestry due to its high quality timber and its ecosystem functions such as erosion control and landslide prevention. Using next-generation sequencing the first silver fir transcriptome was sequenced and annotated. The annotation protocol lead to the identification of 460 candidate genes that are putatively involved in environmental stress resistance. Additionally, the transcriptomes of twelve more individuals were sequenced. Subsequent SNP detection produced on average 24580 SNPs per individual. For the SNPs that lie on the annotated candidate genes, a KASP genotyping assay was designed that allows for the fast and cost efficient genotyping of many *A. alba* trees.

In the Mediterranean Basin, *A. alba* occurs in isolated and comparatively small populations. Recently, strong die-off events were observed in these populations, likely as a result of severe droughts. In the course of climate change, droughts are predicted to increase further in frequency and severity endangering the persistence of Mediterranean *A. alba* populations. Particularly these populations may, however, be of special value for the adaptive potential of the whole species since they may harbor genetic resources pre-adapted to relatively xeric environments. There is an ongoing debate about assisted migration, *i.e.* the transplantation of Mediterranean *A. alba* populations to more northern population in order to use them as donors of drought adapted alleles. To elucidate the question whether the idea of assisted migration is practical, knowledge about the state of local adaptation, its molecular basis and its environmental drivers is needed.

Here, 376 individuals from high and low elevational populations which are located along a 170 km east-west transect across four mountains in the south-western Mediterranean pre-Alps of France were investigated at 267 SNP loci distributed across 175 candidate genes. The methodological framework of landscape genetics was applied. Results of Bayesian cluster analysis showed that the dominant genetic pattern coincides with an east-west isolation-by-distance genetic structure, likely a result of post-glacial recolonization movement. F_{st}-outlier tests revealed 16 SNPs under divergent selection. Significant associations of allele frequencies of eight SNPs and winter drought were detected suggesting that this variable is a strong driver of local adaptation. Correlations tests of genetic distance of a subset of outlier loci and environmental distance demonstrated that isolation-by-environment occurs along altitude. Re-analysis of published common garden phenotypic data using Q_{st}-F_{st} comparisons revealed adaptive divergence of bud flush and growth rate at the coldest and most humid site.

The thesis illustrates a complex pattern of genetic structure in ecologically marginal populations that is shaped by demography and natural selection. It highlights that local adaptation is site specific in terms of both the genes under selection and the environmental drivers. Hence, whether a population is suitable to be the subject of assisted migration would have to be answered individually. Since this is highly impractical, the thesis states that *in-situ* conservation of marginal populations is the best option to ensure the capacity of evolutionary change in *A. alba* and to support the species' ability to cope with the challenges of a rapidly changing climate.

Synopsis

1. Introduction

1.1. Threats of forests: Climate change and habitat fragmentation

Forests cover 45 % of the total terrestrial area in Europe. They hold 80 % of the earth's total biomass (Kindermann *et al.* 2008). Wood, pulp and paper industries contribute 1 % (127 Billion €) to the European Gross Domestic Product. In addition, the value of non-wood goods such as mushrooms, nuts, game meat, honey, cork, resin etc. sum up to over 2.7 Million €. Equally important are the environmental services such as the preservation of soil erosion, hydrological regulation and the provision of unique habitats for many species. The function of forests as carbon sinks may be most important of all ecosystem services. Between 2005 and 2010 the average annual absorption of carbon was estimated to be 870 Million tons which accounts for 10 % of the greenhouse gas emissions of European countries (FOREST EUROPE, UNECE and FAO 2011). Forest trees harbor a great amount of genetic resources not only at species level, as a habitat provider, but also because they constitute rich genetic resources themselves. The distribution areas of many European tree species cover a wide range of different climatic and edaphic conditions where they adapt and subdivide into populations meeting the challenges of their respective environments (Aitken *et al.* 2008). The rich genetic pools of distinct populations are essential for the evolutionary change of the whole species (Hughes *et al.* 1997; Hobbs & Mooney 1998).

Ramankutty & Foley (1999) estimate that between 1700 and 1992 the global net loss of forests/woodland was 11.4 million km^2. For the period 1990 – 2005, around 0.6 million km^2 of forest were lost globally (Lindquist *et al.* 2012). Formerly continuous distribution ranges of tree species are today occurring in fragmented habitat islands which results in a reduced population connectivity (Didham 2010). In principle, isolated population patches within a landscape matrix of unsuitable habitats have an increased risk of losing genetic diversity through drift, inbreeding and reduced inter-population gene flow. Above that, reduction of allelic richness in fragmented habitats may decrease the capability of populations to respond to changes of their environments (Young *et al.* 1996).

Before the end of the 21st century the global surface temperature is likely to increase by more than 2 °C relative to the time period of 1850 to 1900 assuming a medium scenario of increased greenhouse gas emissions (RCP 6.0) (Meinshausen *et al.* 2011; Stocker *et al.* 2013). In Europe, terrestrial surface temperature has increased by 1.3 °C. By 2071 to 2100 the projected increase of land temperature in Europe ranges between 2.5 and 4.0 °C. Overall, southern European ecosystems will experience increased water stress: Precipitation has decreased in southern Europe compared to pre-industrial times. Projected decreasing summer precipitation will lead to longer and more frequent heat waves. Fires in forest

ecosystems are expected to become more frequent and forest growth is predicted to decrease (European Environment Agency 2012).

The fates of forests facing these threats may be (i) on-site adaptation through phenotypic plasticity and/ or evolution, (ii) migration to more suitable habitats or (iii) extinction (Aitken *et al.* 2008; Bussotti *et al.* 2015). Optimistic opinions predict that tree species will shift their distribution ranges and track their ecological niches (Fisichelli *et al.* 2014) or that their adaptive capacity will enable them to locally persist (Carlson *et al.* 2014). Indeed, reactions of the European ecosystems to climate change are already observable in *e.g.* advanced leaf unfolding and pollen release. Also, northwards and upwards shifts of distribution ranges of plants have been monitored (Parmesan & Yohe 2003). Yet, pessimistic scenarios describe widespread local extinction events (Davis & Shaw 2001), because the rate of the recent environmental change exceeds those known from the past (Shaw & Etterson 2012). Migration velocity will likely not be sufficient to track suitable habitats, because of low recruitment and seedling survival (Zhu et al. 2012; Nathan et al. 2011). Recent drought induced mortality (Bréda *et al.* 2006; Allen *et al.* 2010) reveals the risk that tree species may not be able to cope with the current environmental changes and even go extinct when their habitat niches are narrow (Thomas *et al.* 2004).

A controversial debate as broken out among scientist and forest managers whether human assisted range shifts of trees from more southern regions can mitigate the impact of climate change to forest trees: Proponents argue that the limitation of dispersal requires human assistance to prevent species loss and disruption of species communities (Iverson *et al.* 2013). Opponents of assisted migration policy claim that to extensively translocate species beyond their native range, sound knowledge about their future habitat is necessary and that the accuracy of this very knowledge is doubtful for biological interactions such as competition, pollination, herbivory, parasitism and symbiosis are absent from prediction models (Guisan & Thuiller 2005). Translocated individuals may not be adapted to their new environments and may even cause maladaptation in native individuals when they introduce their alleles. Further, assisted migration may bear the risk that introduced species turn into pests (McLachlan *et al.* 2007).

1.2. Theory of local adaptation

The thesis focuses on adaptive evolution as one way to cope with climate change. To study local adaptation, it is necessary to understand the conceptual basis of evolution, *i.e.* the role of variation and the mechanisms of selection as well as the effect of the latter on the genepool of populations. The concepts go back to the works of Charles Darwin and Gregor Mendel. Darwin (1850) recognized that species change

over time results from the discriminative transition of variable traits from one generation to the next driven by the surrounding environment. In 1865, Mendel postulated that the trait of an organism is governed by two alleles, a maternal and a paternal and that these alleles are inherited and recombined independently in the offspring generation. The synthesis of the two achievements was not laid down till 1930 by Fisher in his book 'The Genetical Theory of Natural Selection' that is regarded as one of the cornerstones of the modern evolutionary synthesis. The following two paragraphs explain the basic concept of the evolutionary synthesis and pay special attention to their meaning for tree populations.

The *sine qua non* for evolutionary change is the variation among organisms. Biological variation occurs at the level of the genotype and the phenotype. Sources of genetic variation in a population are mutations (*i.e*. 'errors' during DNA replication), introduced alleles from other populations (migration) and the re-shuffling of alleles in meiosis during sexual reproduction. The greater the number of reproducing individuals in a population is (effective population size, N_e), the higher is the amount of genetic variation within the population (standing genetic variation). Thus, in widespread, outbreeding, mostly wind pollinated and long living tree species with high N_e's and with the capability to release gametes over decades, gene flow and standing genetic variation are generally high (Hamrick *et al.* 1992; Hamrick & Godt 1996; Nybom & Bartish 2000; Hampe & Petit 2005; Porth & El-Kassaby 2014). At phenotype level, variation is defined as the observable differentiation of organisms as a product of their genotype and the environment. The contribution of the environment on the observed phenotype is termed phenotypic plasticity.

Natural selection is the differential contribution of genotypes to the next generation due to differences in survival and reproduction. Its operational unit is the phenotype, its target the underlying genotype. Assuming Hardy-Weinberg conditions except for the presence of selection, balancing selection occurs when the heterozygote state of a locus causes higher fitness than the homozygote (overdominance). It maintains genetic variation in a population. Directional or positive selection occurs when one allele at a given locus causes a fitness advantage over the other. In the course of evolutionary change directional selection leads to the fixation of the advantageous allele and the mean population value of the trait coded by that locus will shift. Divergent or disruptive selection, where the different alleles are favored in local environments, leads to the fixation of the advantageous allele in the respective environment and the population will sub-divide (Allendorf & Luikart 2007). Population subdivision is measured as the ratio of heterozygosity within a subpopulation and total heterozygosity (F_{st}) (Wright 1950). Divergent selection causes divergence of phenotypes in their respective local environments, a process known as local adaptation (Le Corre & Kremer 2003; Kawecki & Ebert 2004).

1.3. Studying local adaptation

Phenotypic variation is a result of both the heritable genotype and plasticity shaped by the environment. Thus, observed differences of individuals cannot readily be taken as a proof of local adaptation. *Sensu stricto*, the existence of local adaptation is proven when native demes show higher fitness than alien demes in a given environment (Kawecki & Ebert 2004). To eliminate the environmental effect, individuals or populations can be grown in the same environment, in so called common garden trials or reciprocal transplantation experiments. There, traits among families can be measured as Q_{st} (Spitze 1993; Prout & Barker 1993), *i.e* the amount of genetic variance among populations relative to the total genetic variance (Whitlock & Guillaume 2009) and compared to the distribution of neutral F_{st}. If Q_{st} is greater than F_{st}, this can be regarded as evidence for local adaptation (McKay & Latta 2002; Whitlock 2008). Evidence from reciprocal transplantation experiments suggest that local adaptation is common across species: De Kort *et al.* (2012) report Q_{st} values that exceed F_{st} values in 71 % of 401 meta-analyzed cases. Also, Hereford (2009) found evidence for local adaptation in even 71 % of the cases.

In forestry, transplantation experiments are often referred to as provenance trials (Mátyás 1996). They were originally established to select the best growing deme for forestry usage in a given environment. Today, they provide empirical data by which local adaptation can be demonstrated. For instance, the re-analysis of a provenance trial of *Pinus sylvestris* showed that survival, relative height and total fitness of central populations decrease with increasing latitudinal transfer distance (Savolainen *et al.* 2007). Provenance trials also provide valuable information of quantitative genetics and population differentiation (Alberto *et al.* 2013) at phenotypic level, namely variation in quantitative traits along with varying environments. Examples for this are the date of bud flush in *Quercus petraea* (Ducousso *et al.* 1996), differences in frost tolerance, in *Pinus sylvestris* (Andersson & Fedorkov 2004) or in critical night length for growth cessation in *Picea abies* (Viherä-Aarnio *et al.* 2005).

1.4. Small populations, isolated populations

The distribution areas of European tree species typically become increasingly scattered towards their margins, where spatially isolated population exist. Here, species often show more contrasting stand densities and lower population sizes (Restoux *et al.* 2008). In such populations, the importance of genetic drift increases. Genetic drift is the random loss of alleles during meiotic cell division. In the absence of mutations and migration, the smaller the effective population size is, the greater is the probability of random loss of alleles by genetic drift. This lessens the effect of directional selection, because rare

advantageous alleles maybe lost. Above that, small populations have an increased risk that disadvantageous alleles become fixed (Lynch & Gabriel 1990; Gabriel & Bürger 1994; Lande 1995). The more isolated a population is, the smaller is the probability that new alleles are introduced by gene flow.

Northern marginal populations mark the leading edge of a species distribution, because the 'take the lead' in the northwards movement of the species that is either due to re-colonization after Quaternary glaciation or to a shift of the distribution range in the course of the recent climate change. Southern marginal populations are termed as trailing edge populations for being the 'last ones' of the migrational movement (Hampe & Petit 2005). Under the assumption that species density decreases towards the distribution margins, higher amounts of gene flow are expected from the distribution center to the periphery than vice versa (Garcia-Ramos & Kirkpatrick 1997) which results in higher maladaptation in marginal populations, because non-adapted alleles are introduced (genetic swamping) (Lenormand 2002; Kawecki 2008). However, Alleaume-Benharira *et al.* (2006) found that moderate levels of gene flow result in higher range wide fitness ('genetic rescue' Richards & Ritland 2000). Genetic swamping would not occur at moderate levels of gene flow because of the counteracting effect of natural selection that purges non-adapted alleles from the populations' genepool. Indeed, empirical evidence of maladaptation and reduced genetic diversity is often missing (Hoffmann & Blows 1994; Sexton *et al.* 2009). For European trees, this may be because the distribution areas of trees often represent several thousand kilometers in size and are rather shaped by Holocene climate dynamics (Hewitt 2004) than by the interplay of gene flow and drift.

In special cases, southern, trailing edge population constitute relicts of refugial populations where the species has persisted during the glaciation period ('stable rear edge population' Hampe & Petit 2005). Although these populations exhibit reduced within population genetic diversity (Petit *et al.* 2003), their between population differentiation is high (Comps *et al.* 2001; Hampe *et al.* 2003; Petit *et al.* 2003; Martin *et al.* 2004; Eckert *et al.* 2008). Due to their uniqueness, such populations can serve as sources of future evolutionary diversity (Lesica & Allendorf 1999). Southern, trailing edge populations of tree species whose core distribution is in temperate areas may function as a donor of climate adaptive alleles (Hampe & Petit 2005), because they presently experience a climate that is predicted for more northern regions. For instance, southern rear edge provenances of *P.sylvestris* show a better response to increased temperature than northern ones (Rehfeldt *et al.* 2002). Thus trailing edge populations are particularly valuable for the mitigation of climate change impacts. Forestry practices should acknowledge their vulnerability and their demand for increased conservation efforts (Lefèvre *et al.* 2013). To guide decision-making on which populations are most likely to contribute to the species genetic rescue, case specific knowledge about the state of 'adaptedness' of a local population and its adaptive capacity is needed.

1.5. Mediterranean forests

The trailing edge of many temperate European tree species is located in the Mediterranean climate zone. The Mediterranean Basin, stretching between 25° and 45 ° of northern latitude, has early on been recognized as a global biodiversity hot spot (Myers *et al.* 2000). It is a fragile mosaic of varied habitats such as mountains, deserts, forests and island which provide ecological niches for about 25 000 plant species (Cowling *et al.* 1996; Vié *et al.* 2009). In comparison, only 6000 flowering plant species inhabit central and northern Europe, although this area is around four times larger in size (Scarascia-Mugnozza *et al.* 2000). High spatial heterogeneity, complex paleo-geography (Fady & Medail 2004), low nutrient availability and the typical summer drought coupled with moderate levels of disturbance by fires are interpreted as the determinants of the high species diversity (Huston 1994). Further, the alternating glaciation- and inter-glaciation periods during the Pleistocene caused a recurrent southwards vegetation expansion into Mediterranean areas, where zones of secondary contact of Mediterranean type and temperate type vegetation where established. This led to a reshuffling of genetic material, as introgressive gene flow gave rise to new endemic taxa, *e.g*. in the case of *Pinus brutia* and *Pinus halepensis* (Roussis *et al.* 1995).

Forests cover ~ 850 000 km^2 of the European Mediterranean area (FAO 2013). Ecologically, their protective function of watersheds and soils are of major importance. Until the 19^{th} century timber overuse for shipbuilding, clearing for arable land and overgrazing had led to a severe degradation of Mediterranean forests. Today they are re-expanding in south-western European countries (Merlo & Paiero 2005). Mediterranean forests contain ca. 250 arborescent species (Quezel *et al.* 1999), while only 30 tree species live in Central Europe. In the course of climate change, the predicted decrease in summer precipitation (Gao *et al.* 2006) and increase in temperature are expected to be particularly pronounced (Stocker *et al.* 2013). The Mediterranean climate zone is considered as highly vulnerable to climate change (Scarascia-Mugnozza *et al.* 2000). At the same time, Mediterranean forest systems are ideal to study local adaptation in natural populations for they are characterized by the presence of a heterogeneous, stressful environment and fragmented populations exhibiting genetic distinctiveness (see above).

1.6. Abies alba

The European silver fir (*Abies alba* MILL) is a conifer species of immense ecological and economical importance. Its southern trailing edge is located in the Mediterranean climate zone where its occurrence is highly fragmented. There, (i) populations show higher within species diversity than other conifers

worldwide (Fady-Welterlen 2005), (ii) a high inter-population variation (Vendramin *et al.* 1999; Piovani *et al.* 2010) and (iii) are severely affected by the current climate change (*e g*. Cailleret *et al.* 2014). Thus, Mediterranean silver fir populations are highly suitable to study local adaptation.

The species is ecologically confined to temperate and Mediterranean mountain ranges. Its distribution area ranges from 52 to 40 °N and from 5 to 27 °E (Wolf 2003). Evidence from pollen and macrofossils, alloenzyme and organelle DNA draws a coherent picture that silver fir survived the Quaternary glaciation periods in four refugias, namely the Apennines, the Balkans, the Pyrenees and the Calabrian Mountains (reviewed in Liepelt *et al.* 2009). While the latter two were identified as isolated and inactive, *i.e.* did not contribute to the present distribution of the species, the Apennines and the Balkans are regarded as starting points for two linages of post glacial re-colonization. From the Apennines the species extended northwards over the Alps and westwards passing the Maritime Alps towards the Massif Central. From the Balkan Peninsula the species reached the Carpathian Mountains and the eastern Alps. The limits of ecological tolerance that define the niche of *A. alba* are not totally clear. The species is described as highly shade tolerant, but sensitive to summer drought, late spring frosts and severe winter frosts (Ellenberg 1988). New findings of paleo-ecological niche modeling suggest that, around 6000 years ago, the species was capable of growing at 5 -7 °C higher July temperatures than today as long as the precipitation did not fall below 700 – 800 mm a-1 (Tinner *et al.* 2013) and that it current occurrence is mainly due to habitat loss and not to climatic restrictions. Indeed, today, the realized distribution of the species is strongly diminished due to human activities (Falk & Mellert 2011). The evergreen *A. alba* can become up to 600 years old, 60 m tall and reach over 3 m of stem diameter at breast height. The typical habitus of older trees is characterized by a flattened top of the crown that develops through reduced height and ongoing lateral growth of branches (Wolf 2003). Leaves are rhombic to aliform with stomata arranged in two rows that appear as whitish lines at the lower side of the leave (Schütt 2008). *A. alba* reaches sexual maturity at the age of 25 – 70. It is monoecius, dominantly outcrossing and wind pollinated. Pollen grains exhibit air pockets. Cones are erect, greenish-blue and around 15 cm in size. They carry aliferous seeds that are anechorously dispersed (Wolf 2003; Schütt 2008). The genome of *A. alba* consist of 2 n = 24 chromosomes. Estimates of DNA content lie between 34.58 pg 2C-1 (Puizina *et al.* 2008) and 16.55 pg C-1 (Roth *et al.* 1997), which accords to 16000 Mb and 17000 Mb and thus, lies well within the range of other conifer genomes that range from ~ 6500 Mb to 37000 Mb (Ahuja & Neale 2005).

Summer drought negatively affects tree ring growths in *A. alba* (Macias *et al.* 2006; Battipaglia *et al.* 2009; Carrer *et al.* 2010; Mazza 2013) highlighting the endangerment of the species at its rear edge. Indeed, drought induced die-offs have repeatedly been reported for Mediterranean populations (Potočić *et al.* 2005; Oliva & Colinas 2007; Cailleret & Davi 2011; Linares & Camarero 2012; Cailleret *et al.* 2014;

Camarero *et al.* 2015). Restoux *et al.* (2008) observed higher selfing rates in low-density Mediterranean conifer stands compared to high-density Mediterranean and non-Mediterranean stands. The low density stands produced less viable seeds which, however, were more resistant to water stress. This suggests a strong impact of diversifying selection on Mediterranean tree populations. Isolated *A. alba* populations in the Northern Apennines show strong inter individual variability for stomatal conductance (Piovani *et al.* 2011) and population in the Pyrenees differ in photosynthetic efficiency at more xeric and more mesic sites (Peguero-Pina *et al.* 2007) Also, climate –growth relationship of Italian fir populations appear not be stable over time and across populations (Carrer *et al.* 2010; Rita *et al.* 2014; Gazol *et al.* 2015). This illustrates the adaptive capacity of the species at the physiological level.

Studying local adaptation in trailing edge populations of *A. alba* can elucidate the question of whether they will be able to cope with climate change, whether they may be a subject of assisted migration or whether they may naturally fulfill the role of an allele-donor in the sense of Petit and Hampe (2005). For such a study, landscape features must be integrated in population genetic analyses. Doing so constitutes a modern research field termed as landscape genetics (Manel *et al.* 2003; Holderegger & Wagner 2008; Schoville *et al.* 2012; Sork *et al.* 2013; Rellstab *et al.* 2015). Landscape genetic studies have demonstrated that genetic structure of *A. alba* in the Apennines and the Alps is driven by geographical distance and by the local environment (Mosca *et al.* 2013).

Here, the landscape genetic approach is applied to marginal populations of *A. alba* to (i) describe patterns of genetic variation at regional and local scale in south-western marginal habitats (ii) determine the relative contributions of demography and selection shaping those patterns (iii) identify genes that are targets of natural selection and (iv) highlight the environmental variables that drive selection.

2. Methods

Landscape genetics encompasses a broad spectrum of analyses using genetic data alone and in combination with phenotypic, environmental and spatial data. Each partial result adds information to a coherent picture elucidating the objectives above.

The genetic basis of phenotypic variation can either be studied using a top-down approach that measures the differences between phenotype and correlates this with the genotype or using a bottom-up approach that identifies genetic differentiation that leads to the expression of different phenotypes. As described above, reliable phenotyping in natural tree population is only possible in controlled environments like common gardens and when kinship information of the observed individuals is available (Holderegger *et*

al. 2006). The bottom-up approach, in contrast, can be readily applied to in-situ populations without a time consuming experimental setting. Breakthroughs in DNA sequencing and genotyping technologies (see Box 1) (Shendure & Ji 2008; Metzker 2010) allow to generate sequence data that represents the entire genome ('whole genome seq' or reduced library sequencing like 'exome seq', 'restriction-site-associated' (RAD- seq) or genotyping-by-sequencing (GBS)) and subsequent screening of single nucleotide polymorphisms (SNPs) which are the most abundant type of genome variation (Morin *et al.* 2004). Typically, these approaches are performed without further information about the generated sequences and the SNPs therein. Statistical approaches that detect genes under selection from anonymous SNP surveys are based on the idea that allele frequencies of loci under selection will significantly deviate from allele frequencies of loci that are merely affected by drift (Lewontin & Krakauer 1973). Selectively favored loci should be more and loci under balancing selection should be less differentiated between populations than loci behaving neutral (Beaumont & Balding 2004). The high overall levels of standing genetic variation in trees and their undomesticated state even enhance the chances to detect loci that deviate from neutral expectation (González-Martínez *et al.* 2006). With increasingly powerful capacities for computational simulations, the Lewontin-Krakauer-Test has regained attention and a number of so called F_{st}-outlier tests have been developed. They deviate from each in the assumed models, the applied simulation algorithms to infer neutrality and in their way of consideration of background demographic structure (Lotterhos & Whitlock 2014). However, the results of simulation studies do not accord with each other in the identification of the best performing method (Pérez-Figueroa *et al.* 2010; Narum & Hess 2011; De Mita *et al.* 2013; de Villemereuil *et al.* 2014; Lotterhos & Whitlock 2014), because the outcome of F_{st}-outlier tests depends on the actual features of the dataset such as demographic history, degree of genetic differentiation between populations, genetic architecture and strength of selection (Excoffier *et al.* 2009; Le Corre & Kremer 2012). These features are often unknown in natural populations so that F_{st}-outlier tests bear the risk to erroneously display SNPs under selection which are in fact not under selection (false-positives) or to fail to identify SNPs that are under selection (false-negatives). The risk of false positives can be reduced by combining different F_{st}-outlier test. False-negatives, however, result – among other reasons – when F_{st}-outlier tests treat loci independently, although traits under selection are governed by several quantitative trait loci (QTLs) (Mackay 2001) or rather many quantitative trait nucleotides (QTNs) (Rockman 2012) and adaptive divergence is a result of covariance in allele frequency (Le Corre & Kremer 2012). This polygenic local adaptation impedes detectable shifts in allele frequencies (Storz 2005).

To lessen the risk of false F_{st}-outlier detection, the landscape genetic approach complements genetic data analysis with spatial and environmental information of local populations (Holderegger & Wagner 2008).

Spatial distance of individuals and populations is an important factor that shapes genetic differentiation within species even in the absence of heterogeneous environmental conditions and selection. To infer local adaptation from spatial patterns of genetic differentiation, the relative contributions of neutral gene flow and natural selection must be estimated. If gene flow is limited and selection is absent, this will result into an isolation-by-distance pattern (IBD). If the effect of selection on the transfer of genes to the offspring generation is stronger than the effect of gene flow, an isolation-by-environment pattern (IBE) will emerge. The latter is a strong indicator of local adaptation. Diagnostic for IBD is a significant correlation between genetic and geographic distance and a greater genetic differentiation among regions than among environments within the same region (Sexton *et al.* 2014). IBE patterns are characterized by significant correlations between and genetic and environmental distance and high genetic differentiation among environments within the same region. They have been demonstrated for *A. alba* and *Larix decidua* in contrasting alpine environments (Mosca *et al.* 2013).

Variation of environmental variables across the distribution range of a species is the basis for natural selection to shape adaptive differentiation (Linhart & Grant 1996). Principally, loci that govern the phenotypic response to environmental change can be identified by correlation analyses (Vasemägi & Primmer 2005). However, since both genetic and environmental changes are based on geography, correlation analyses must – similar to F_{st} outlier methods – take into account background population structure (Vasemägi 2006). Recent Bayesian methods implement null models derived from empirical patterns of genetic covariance (Coop *et al.* 2010) or introduce population structure via unobserved variables estimated in principle component analyses (Frichot *et al.* 2013). Using a Bayesian generalized mixed model (Eckert *et al.* 2010) identified SNPs associated to various climatic factors such as growing degree day, precipitation and aridity in a range wide study of *Pinus taeda*.

Genotyping assays that cover the entire genome are still not available for *A. alba*. In fact, whole genome information is so far only available for *Picea abies* (Nystedt *et al.* 2013), *Picea glauca* (Birol *et al.* 2013) and *Pinus taeda* (Neale *et al.* 2014; Wegrzyn *et al.* 2014). One reason for the lack of genomic data for the majority of the conifers is their huge genome size (18 Gbp to over 35 Gbp, Murray *et al.* 2012) whose representative genotyping would demand high efforts. It follows that targeted sequencing and SNP genotyping of candidate genes is often used in non-model conifers.

Candidate gene studies attempt to describe causative functional genetic polymorphisms that govern ecologically and economically important traits (Pflieger *et al.* 2001; Chhatre *et al.* 2013). They are typically used for association between genetic variants and adaptive phenotypic traits such as cold-hardiness (Eckert *et al.* 2009b), drought tolerance (Eckert *et al.* 2010) and wood quality (Dillon *et al.* 2010; Beaulieu *et al.* 2011). Also, Budde *et al.* (2014) looked for association between candidate genes and

cone serotiny at different wild fire frequencies in *Pinus pinaster*. The rational of candidate gene approaches is that – compared to anonymous sequences – their investigation increases the chances to identify polymorphisms under selection. Therefore SNP data sets from candidate gene sequencing are also analyzed using F_{st}-outlier test, tests for patterns of IBD vs IBE and genotype-environment associations.

3. Content of the thesis

For *A. alba* practically no resources appropriate for the study of adaptive variation were available until 2010 when the first transcriptome of the species was sequenced using the pyrosequencing technology of Roche 454 (in following referred to as '454 transcriptome', see Box 1) within the EU-EvolTree Network of Excellence. This data marked the starting point for adaptive genetic research in *A. alba* and a stated aim of this thesis was to generate novel genotyping tools from the resource.

Roschanski et al. 2013 (Chapter I) describe the de-novo transcriptome itself and present a protocol for candidate gene annotation, subsequent primer design and candidate gene re-sequencing. However, sequence information of at least two individuals is necessary to identify homozygous SNPs. Screening of even more individuals' sequences further increases the chance to identify even rare polymorphisms. Thus, additional to the 454 transcriptome, twelve more transcriptomes were sequenced using the dye-sequencing technology of Illumina (in the following referred to as 'Illumina transcriptomes', see Box 1). Chapter II describes the Illumina resource and how it was combined with the annotation of the 454 resource (Chapter I) to design a KASP assay (Box 1) for multiplex SNP genotyping at candidate genes.

The main part of the thesis is a landscape genetic study in the Mediterranean Alps of south-western France (Roschanski and Csilléry *et al.* 2015, Chapter III). Genetic patterns of south-western trailing edge populations of *A. alba* were studied along a 170 km east-west spanning transect of four mountain ranges.

Using the newly developed KASP assay, 376 individuals were genotyped at 276 SNPs distributed across 175 candidate genes. The genetic data was complemented with climate data obtained by combining on-site measurements and published long term data. Additionally, published data from a common garden trial (Sagnard *et al.* 2002) wherein seedlings of the studied populations had been phenotyped for growth and drought tolerance was re-analyzed.

The relative contributions of distance and environment to the genetic divergence of populations were assessed using hierarchical AMOVAs (Excoffier *et al.* 1992) and Mantel tests for correlations of genetic, geographical and environmental distance matrices. SNPs under directional selection were screened using

three different F_{st}-outlier tests. The state of each SNP was associated with each environmental variable in a latent factor mixed model (Frichot *et al.* 2013). Finally, a Q_{st}-F_{st} tests (Whitlock & Guillaume 2009) of adaptive divergence for bud flush, growth rate and response to water stress were performed on the common garden data. Chapter III demonstrates isolation-by -environment along altitude, it shows divergent selection at one site, probably driven by drought and cold and highlights that winter drought is associated with eight SNPs. However, the overall genetic structure of the studied populations was shaped by spatial distance. These key findings illustrate well the complexity of selective pressures that interfere both with each other and with demographic genetic patterns at marginal populations in the Mediterranean Alps.

Box 1 Sequencing and genotyping technologies

Excursus sequencing and genotyping technologies

Chain termination sequencing was developed in 1977 (Sanger *et al.* 1977) and has been extensively used not only in genetics but also in other fields such archeology, forensics and others (França *et al.* 2002). Not until more than twenty years later the so called next-generation-sequencing technologies have emerged and are currently transforming the field of population genetics into genomics.

Pyrosequencing '454-Sequencing'

For pyrosequencing (Ronaghi *et al.* 1998) adapter ligated DNA fragments are anchored and amplified on beads. Each bead that carries a unique single stranded sequence is placed into a micro-well plate that is loaded with the enzymes DNA polymerase, ATP sulfurylase and luciferase. The four nucleotides are then sequentially flowed over the wells. Wherever a nucleotide complements the template sequence its incorporation releases a phosphate. The luciferase consumes the phosphate and produces detectable light.

Dye-sequencing 'Illumina Sequencing'

The dye sequencing technology uses removable terminators that are base specifically labeled with fluorescent dyes. Their incorporation is detected while they are incorporated into the growing DNA strand. After incorporation, the terminators are cleaved and a new the incorporation cycle starts. This sequencing-by-synthesis steps are carried out simultaneously at several million sequence fragments which are anchored on a solid surface.

KASP-genotyping

In a KASP assay, allele specific primers competitively bind to the target DNA. They harbor unique tails that are complementary to a FRET (fluorescence resonant energy transfer) cassette that is labeled with a quenched FAM or a HEX dye. When the relevant primer elongates during PCR the corresponding FRET cassette binds to the tail, the quencher is released and the fluorescent signal of the dye is imaged.

Both pyrosequencing and dye-sequencing are sometimes confusingly termed as 'pyrosequencing'. Also 'next-generation-sequencing' and 'high-throughput-sequencing' are imprecise terms, because they refer to all new sequencing techniques that have emerged after Sanger sequencing. A more appropriate general term is 'massive-parallel-sequencing' since they both rely on this principle. The leading company for pyrosequencing is 454 Life Sciences, a Roche company (Branford, USA) and Illumina (San Diego, USA) is the leading company for dye sequencing. For simplicity, pyrosequencing is termed as 454-sequencing and dye sequencing as Illumina sequencing in this thesis. The major advantage of 454-sequencing is the length of a single read of up 1000 bp (http://www.454.com, June 2015) while maximum read length of Illumina sequencing is 250 bp (http://www.illumina.com, June 2015). However, the costs for sequencing one million base pairs are 0.07 USD using the Illumina while sequencing cost for the same output using the 454 technology are 10 USD (Liu *et al.* 2012).

4. Discussion

Human well-being requires healthy forest ecosystems. Sustaining productive and multifunctional forests is of prime importance. Forest ecosystems are endangered by threat such as fragmentation, habitat loss and climate change. Especially in small and isolated populations, at the south-western edge of tree species' distribution ranges in Europe, the risk of die-offs is high. But these very trailing edge populations may harbor genotypic and phenotypic variation crucial for future evolutionary change. This thesis contributes to the understanding of local adaptation at the south-western trailing edge of *A. alba* as it built up novel genetic resources and tools and presents their application in a landscape genetic approach revealing a complex adaptive landscape.

4.1. Genetic resources in Abies alba

In the first chapter of the thesis Roschanski *et al.* (2013) present a protocol for directed reduction of large quantities of transcriptome data to identify potentially adaptively relevant target regions. We used BLAST searches of transcriptomic sequences against the well-known and curated data bases refseq_protein (ftp://ftp.ncbi.nlm.nih.gov/refseq) and gene ontologies (Ashburner *et al.* 2000; http://geneontology.org). Complementary, proteins were downloaded from the UniProt database (http://www.uniprot.org/) that are involved in the adaptation to drought according to scientific literature. These proteins were then 'blasted' against the transcriptome. After stringent filtering, 283 candidate genes were obtained. Indeed, candidate gene genotyping is a cost efficient way to perform bottom-up approaches when investigating the local adaptation at the molecular level (Ekblom & Galindo 2011) and annotation by sequence similarity has often been used (*e.g*. Novaes *et al.* 2008; Parchman *et al.* 2010). However, often, the evolutionary distance between a gene of a model organism whose annotation is derived by mutational experiments and a gene of a non-model species is large, *e.g*. gymnosperms and angiosperms have been evolving in parallel for 300 million years. Therefore, the functional annotation can be missed or simply be wrong, because genes may lose or change function during evolution (Tiffin & Ross-Ibarra 2014). Biological interpretation of statistical evidence for local adaptation of genes annotated by sequence similarity may consequently be premature (Barrett & Hoekstra 2011). In Roschanski and Csilléry *et al.* (2015), 267 successfully genotyped SNPs were re-annotated using *Picea abies* proteins and only 187 SNPs actually appeared in coding regions which strongly hints to the fact that the annotations in Roschanski *et al.* (2013) are at least partially wrong. However, protein sequences of *Picea abies* were still not available when the annotation of Roschanski *et al.* (2013) was conducted.

A second, novel genetic resource for *A. alba* is presented in chapter II. Transcriptomes of twelve trees were sequenced and subsequent SNP calling was performed. The effort yielded on average 20 935 contigs and 24 580 SNPs per individual. The output was used as the SNP discovery panel to genotype novel candidate genes by the KASP technology. The tool was used for the main study of the thesis (chapter III).

4.2. Divergent selection in Abies alba

In the study constituting the third chapter (Roschanski and Csilléry *et al.* 2015), sampling was conducted in south-western marginal silver fir forests where the intrinsically temperate species is experiencing harsh ecologically conditions and likely subjected to selective pressures that are stronger than for central populations. Four mountain ranges were sampled along an east-west transect of 170 km Ventoux, Lure, Issole and Vesubie. At each mountain, two populations were sampled at high and low elevation, respectively. Sites differed with respect to climate, vegetation and mother bedrock. Their common refugium is Central Italy (Liepelt *et al.* 2002) from which the species re-colonized the south-western Alps in north-eastwards direction. Until today, the species only persisted at higher elevations forming rather small and isolated populations surrounded by unsuitable habitats. At Mediterranean mountain slopes environmental conditions are changing rapidly from hot and dry to cold and moist with increasing altitude (Ozenda 1975) so that populations at high and low elevations are subjected to contrasting selective pressures. The documented distinctiveness of the species' trailing edge populations (Peguero-Pina *et al.* 2007; Piovani *et al.* 2011) and the imminent threats of habitat loss and climate change make *A. alba* an excellent model for the study of local adaptation. The nested sampling allowed for the investigation of both regional and local adaptive genetic variation.

To characterize the environment at plot level up to three years of datalogger measurements of temperature and relative humidity were used to downscale long term climatic data in such a way as to create plot specific climatic data over a 40 years time period starting in 1960. The data certainly are closer to match the actual abiotic environment of the plots than gridded climate and satellite data as used in Mosca *et al.* (2013). Still, for *A. alba* being a tree the years of 1960 to 1990 may lie after the times of strongest thinning, *i.e* seedling stage, for the investigated individuals because they were sexually mature when they were genotyped, thus likely 60 years and older (Wolf 2003). Nevertheless, the data used may well represent the climate before the onset of pronounced warming and an increase of warm extremes in Europe (Klein Tank & Können 2003). Four meaningful, uncorrelated climate variables were calculated from the climate data (self-calibrating Palmer Drought Severity Index in January, February and March, relative humidity, vapour pressure deficit in March, April and May, and vapour pressure deficit in

September, October, November). Using principal component analysis of the four variables the micro climate of the population could be described. High and low elevational plots were separated along the first PC.

Genetic and genotype-environment associations were complemented with the analysis of phenotypic divergence in phenology, growth and drought resistance. Phenotyping was done from 1995 to 1999 in a common garden in Aix-en-Provence (France). The trials were established with seeds of 22 – 43 mother trees growing at the studied mountains. Although, the seeds did not originate from the very same plots as the SNP genotyped trees, sampling sites at the same mountain are likely connected via pollen flow. Traits were measured at the seedling stage, thus at the crucial moment of establishment when the mortality of tree cohorts is high. To control for kinship effects of phenotypic divergence (Q_{st}), the degree of relatedness within a population must be known. Pedigree information is hardly available in common gardens of trees where usually only a single generation is tended. Thus relatedness was inferred assuming a fixed outcrossing rate (Restoux *et al.* 2008) so that selfing of mother trees was properly accounted for.

Overall population genetic differentiation was analyzed using Bayesian clustering (Pritchard *et al.* 2000). It showed a dominant longitudinal genetic structure. Mantel tests for correlations between genetic and geographic distance revealed that this structure aligned with an isolation-by-distance pattern, likely following the east-west direction of re-colonization. Hierarchical AMOVA analysis (Excoffier *et al.* 1992) revealed that most of the variance in allele frequencies was present within populations, but that a significant genetic differentiation was detected between mountains and populations at the same mountain as well.

Genotype-environment associations were estimated using a latent factor mixed model (LFMM, Frichot *et al.* 2013). It detected association between SNPs and winter drought. Concordantly, growth decline in *A. alba* due to water limitation appears to occur temporally lagged. Retrospective dendrochronological studies have shown that in French mountains, growth is particularly effected by the conditions of the previous year (Desplanque *et al.* 1999;Rolland *et al.* 2000; Cailleret & Davi 2011; Lebourgeois 2013). For *Abies lasiocarpa*, Bigler *et al.* (2007) found a stronger correlation of growth with early seasonal than with late seasonal drought. Indeed, *A. alba* is a drought avoider (Aussenac 2002) that closes its stomata during summer, thus being relatively unaffected by actual summer drought. Stomatal conductance decreases exponentially with atmospheric vapor pressure deficit, a relationship that has long been known (Lange *et al.* 1971). High and low elevational plots of the studied mountain sites were highly contrasting with respect to vapor pressure deficit. Yet, no associations were detected with this parameter. Winter drought decreased with decreasing east-west distance between mountains and the Mediterranean Sea which is at the same time the direction of neutral genetic divergence. Under this demographic scenario the latent

factor mixed model is most powerful to detect associations (Frichot *et al.* 2015). LFMM did not detect associations between SNPs and altitude. Thus, it remains unanswered whether sampling of additional plots on mountains that do not lie on the transect may have revealed associations between climate variables that vary mainly between high and low elevation (*i.e.* vapor pressure deficit).

Nevertheless, the presence of adaptation to variables aligning with altitude could be demonstrated by another partial result of the study, namely the correlation of genetic distance between outlier SNPs of the populations and environmental distance in terms of PC1. Here, the Mantel test supported a significant IBE pattern. For Calabrian silver fir population in Swiss, Danish and Britain common garden trials, the general pattern of increased winter frost resistance and decreased growths with increasing altitude has been observed (Larsen 1986; Hansen & Larsen 2004; Kerr *et al.* 2015). Although not especially attributed to altitude, the findings are reflected by the observations of the common garden trial re-analyzed here. Individuals stemming from the coldest and most humid site, Lure, were more drought resistant, flushed earlier and grew slower than the individuals from the other sites. Marginally significant Q_{st}-F_{st} test results indicated that divergent selection for bud flush and growth is present at the site. It has commonly been observed in other conifers that, at low elevation and mild common garden conditions, individuals from high elevation sites flush earlier to maximize the length of the growing season during available frost-free periods (*e.g*. Mimura & Aitken 2010; Kreyling *et al.* 2012). Presumably, Lure individuals require lower temperature sums to initialize bud break, thus, benefitting from a longer vegetation period which can be interpreted as adaptation to cold (Vitasse 2009).

Water stress leads to carbon starvation through decreased photosynthesis and/or to hydraulic failure (McDowell 2011). According to the osmotic function hypothesis (Sala *et al.* 2012), non-structural carbon is crucial for hydraulic transport, particularly during severe stress. For instance, hydraulic failure may be associated with the loss of adequate tissue carbohydrate in *Pinus edulis* seedlings (Sevanto *et al.* 2014) and the drought resistant species *Nothofagus nitida* increases non-structural its carbon content while the drought susceptible species *Nothofagus dombeyi* decreases it under drought stress (Piper 2011). Also, non-structural carbon is likely involved in frost resistance as woody plants that experience freezing temperatures show a strong seasonal increase of soluble sugars during winter (Kozlowsky & Pallardy 1997; Ruelland *et al.* 2009). Recent common garden experiments also found that foliage cold hardiness differs between populations of *Pinus nigra* (Kreyling *et al.* 2012) and that the populations most resistant to cold were also the most drought resistant (Thiel *et al.* 2014). Thus the carbon status of a tree may be the link between drought and frost resistance. However, to what extend the capability of carbon uptake and storage is heritable and thus a target of selection rather than a plastic physiological reaction is unknown. In *Abies lasiocarpa* a trade-off between high growth rates in young trees and longevity has been observed

(Bigler & Veblen 2009). Eventually the slow growth and the high stress resistance of Lure individuals also display this trade-off.

Three different F_{st}-outlier tests were performed, each of them having particular advances with respect to the sampling design. A total of 16 unique SNP loci showed unusually high levels of differentiation relative to the null model after correcting for multiple testing. SNP 15808_snp829 was commonly detected by two methods. The SNP induces an amino acid change from leucine to methionine and the locus encodes the enzyme chalcone synthase 2. Chalcone synthase is a key enzyme of the flavonoid biosynthetic pathway. Moreover, in *Pinus sylvestris*, chalcone synthase has been significantly differentiated among Scottish populations (Wachowiak *et al.* 2011). SNP 15663_snp771 was found associated to winter drought and to be an F_{st}-outlier. It encodes the heat shock protein 83 which can be involved in various physiological responses to stress. Heat shock proteins have previously been identified as targets of balancing or divergent selection in other studies of non-model organisms using F_{st}-outlier tests (e.g. Chen *et al.* 2012; Chhatre *et al.* 2013; Mosca *et al.* 2013; Csilléry *et al.* 2014). Further, Prunier *et al.* (2013) found that the heat shock protein DNAJ was associated to height in a large scale QTL mapping study in *Picea mariana*. Additionally to heat shock protein 83, loci encoding an ubiquitin like protein and three different subunits of the proteasome were F_{st}-outliers. Notably, all are involved in protein folding and degradation. Ubiquitin is also referred to as a heat shock protein, because heat stress increases ubiquitin transcription that may reflect an increased demand for removal of damaged proteins (Vierling 1991). SNPs from ubiquitin coding genes have also been reported as F_{st}-outliers in silver fir (Mosca *et al.* 2013) and were associated to precipitation in *Picea mariana* (Prunier *et al.* 2013). Proteasome regulatory subunits have been detected as an F_{st}-outliers in other studies too, *e.g.* in *Cryptomeria japonica* (Tsumura *et al.* 2012) and in *Populus tremula* (De Carvalho *et al.* 2010). The proteasome is a multi-subunit complex which degrades ubiquitin tagged proteins and thus controls cell regulation, but has also the capacity to degrade damaged proteins (Kültz 2005). However, as already discussed, sequence annotation was determined based on homology to model species and the state of the SNPs was inferred by comparisons to protein sequences of *Picea abies* which is evolutionary quite distant to *A. alba*. Thus the possibility of false annotation cannot entirely be ruled out and the biological interpretation of the role of encoded proteins must be treated with caution (Tiffin & Ross-Ibarra 2014).

4.3. Critical remarks

As explained above, one limitation of the study was the low number of observations (eight populations from four sites and eight data points for climate variables) that most likely decreased the power of LFMM (Frichot *et al.* 2015) and Q_{st}-F_{st} analyses (Whitlock & Guillaume 2009).

The SNP markers that were used are lying on candidate genes and thus are per definition non-neutral. Conflictingly, the demographic background of the investigated populations was inferred on the basis of the same non-neutral SNP set where the assumption of neutrality was made. This bias is common to studies that seek to detect adaptation from candidate gene data using F_{st}-outlier tests (Eveno *et al.* 2008; Namroud *et al.* 2008; Tsumura *et al.* 2012) and from environmental associations (Tsumura *et al.* 2014; Eckert *et al.* 2010). A way of mitigating the bias is to divide the SNP marker set into putatively selective and putatively neutral (control) SNPs (Chen *et al.* 2012). Yet, to be stringent, candidate gene data is not appropriate for F_{st}-outlier test (Foll & Gaggiotti 2008) and the way it is used here, marks the major inconsistency of the thesis. Non-neutral markers may also inflate the neutral envelope of F_{st}-outlier tests which may decrease the chances to find loci under selection. Still, being limited by the availability of genetic resources in *A. alba*, we decided to use a candidate gene approach. In *Arabidopsis*, (Fournier-Level *et al.* 2011) found only four SNPs - out of 200k SNPs - to be associated with climate variables. Thus, we could expect the majority of the analyzed SNPs to be neutral. Since the number of loci we could genotype was rather low, we argued that a candidate gene strategy enhances the chances the detect loci under selection and is suitable for non-model trees where SNP genotyping tools that represent the whole genome are missing.

5. Further research

5.1. Genome wide patterns of selection

The new resource that was generated, thirteen transcriptomes (Chapter II) bearing more than 20 000 SNPs per individual, is far from being exhaustively exploited and offers manifold possibilities to extend adaptive genetic research in *A. alba*. The twelve Illumina transcriptomes should be re-assembled and mapped against the long reads of the 454 transcriptome applying more stringent quality controls and a more detailed annotation that includes the identification of the syn/non-syn state of the SNPs. With this, haplotype based test for signatures of selection and tests for selective sweeps can be performed (Tajima 1989; Fu & Li 1993; Fay & Wu 2000; Sabeti *et al.* 2002; Nielsen *et al.* 2005). Also, adaptive evolution can be inferred at population level, hence comparing the ratio on synonymous and non-synonymous SNPs

within and between populations using the McDonald-Kreitman test (McDonald & Kreitman 1991). When two SNPs are in LD epistatic selection (selection on two or more genes that interact with each other) may also be detected through outliers of inter-population LD variance (Ohta 1982; Csilléry *et al.* 2014). Finally, since the new resources represents the exome wide variation of twelve individuals with known geographic origin a careful examination of F_{st}-outliers could reveal genomic regions of divergence (Nosil *et al.* 2009), even though samples size is small. Alternative approaches to reduce genome complexity in non-model organisms such genotyping-by-sequencing and RAD-sequencing with subsequent SNP calling have become feasible in non-model species (Chen *et al.* 2013) and may extend genomic data in *A. alba*.

5.2. Novel candidate genes

The set of candidate genes presented in chapter I should be extended. The gene expression pattern of an organism changes under stressful conditions. Gene expression profiling is therefore a good tool to establish novel candidate genes (Prunier *et al.* 2016). In conifers, differential gene expression was reported in response to pathogens infection and wounding in *Picea sitchensis*, *Picea abies*, and *Pinus monticola* (Ralph *et al.* 2006; Arnerup *et al.* 2011; Liu *et al.* 2013) and for different tissue types in *Cryptomeria japonica* and *Cunninghamia lanceolata* (Wang *et al.* 2007; Mishima *et al.* 2014). In *A. alba*, Behringer *et al.* (2015) revealed 296 candidate genes for drought stress by comparing transcript abundancies from well-watered and drought stressed seedlings using massive analysis of cDNA ends (MACE). Many candidate transcripts did not show sequence homologies to genes of known function from model organisms and thus might hold novel insights for the study of adaptive variation.

However, observations in *Arabidopsis* have shown a common cellular response to many different types of stresses (Knight 1999; Allen *et al.* 2000). Under osmotic stress, *Arabidopsis* plants showed an at least two-fold change of expression rate at 46 % of all analyzed genes suggesting that all aspects of plant physiology are involved in stress response (Kreps *et al.* 2002). Thus, it is worth discussing whether the application of candidate gene data to infer adaptation to drought and cold is constructive, given that definable physiological responses have not been pinpointed yet. In this context, a thought provoking fact is that the physiological reasons of tree mortality are not clear (McDowell 2011; Sevanto *et al.* 2014) which hinders the interpretation whether a statistically identified gene is of biological relevance. Possibly, candidate genes should rather be subject of analyses that infer the genetic basis of phenotypes that are known to be adaptive, *e.g*. the response to the photoperiod (Källman *et al.* 2014).

5.3. Phenotyping

In chapter III genotype-environment association and F_{st}-outlier detection revealed a number of putatively adaptive SNPs. Population means of the studied populations constituted the phenotypic information used to complement the results. The study lacked phenotypic data for the individually genotyped trees. By now, phenotypic data such as diameter, tree rings, bud burst phenology, has been obtained for some of the studied trees. Since natural selection acts on the phenotype, individual phenotypes are key to understand by which genetic variant a trait is governed (Houle *et al.* 2010). Due to their low levels of intra-specific differentiation wide spread tree species are well suited for association mapping approaches that statistically link genotypes and phenotypes (Neale & Savolainen 2004). Thus, the phenotypic data that is now available should be included in further studies. Genotype-phenotype links were, for instance, established for cold hardiness traits in *Pseudotsuga menziesii* (Eckert *et al.* 2009a) using an experimental approach and seedlings in nurseries. Wood property traits, such as cellulose microfibril angle and density are also appropriate for association studies because of their high heritabilities (Baltunis *et al.* 2007). In provenance trials, wood properties were successfully associated to individual SNPs for *Picea glauca* (Beaulieu *et al.* 2011), for *Pinus radiata* (Dillon *et al.* 2010) and for *Pinus taeda* (González-Martínez *et al.* 2007). Technical achievements in non-destructive phenotyping such as the use of X-ray diffractometry to assess wood stiffness (Evans 2006) have facilitated the measurement of these traits.

In natural populations, in-situ phenotyping of trees is still challenging. Often, study populations are remote and far away from well-equipped labs (Bussotti & Pollastrini 2015). Easily measurable traits such as leaf traits, strongly vary within the crown of the same genotype depending on their exposition or, in case of evergreen conifers, on their age (Bussotti & Pollastrini 2015). Annual growth as measured from wood cores has shown to be more climate driven than gene driven (King *et al.* 2013; de Luis *et al.* 2013). Hence, phenotyping has fallen short on genotyping and there is a broad claim that advances in phenotyping are needed (Houle *et al.* 2010). Currently, new technologies in the field are emerging. An exciting new development is the sensor technology where light is pulsed on a dormant bud and its reflection is measured by a photodetector (Kleinknecht *et al.* 2015). The sensor device also measures temperature and photoperiod directly at the leaf surface, thus provides the presumably most important triggers for bud break. The technology may be superior the remote sensing attempts since the resolution of the latter is not sufficient to monitor single individuals (Schwartz *et al.* 2002).

5.4. Association studies

To test for genotype-phenotype associations, regression of single SNPs and a trait are often used (González-Martínez *et al.* 2007; Ingvarsson 2008; Eckert *et al.* 2009a; Wegrzyn *et al.* 2010; Cumbie *et al.* 2011). In these studies, the phenotypic variance explained by a single SNP did not exceed 6%. More recent approaches take all or a group of SNPs into account to predict phenotypic variables and thus accommodate polygenic control of traits. They use redundancy analysis (De Kort *et al.* 2014), Bayesian mixed models (Eckert *et al.* 2012; Budde *et al.* 2014) or the machine learning algorithm 'random forest' (Holliday *et al.* 2012). Random forest (Breiman 2001) constructs a large number of decision trees for the prediction of a trait by a random group of SNPs and then assesses the accuracy of each tree. It also allows for the testing of epistasis by backwards exclusion of the most important SNPs from the decisions tree and recalculating the importance of the remaining SNPs. When a pair of SNPs interacts the exclusion of one SNP would change the importance of the remaining SNPs.

6. Conclusions

The thesis is an example of how to elucidate landscape genetic patterns in a non-model species entirely starting from scratch, taking advantage of state of the art sequencing technologies. The main value of the study lays in (i) the well-considered sampling design (ii) the thorough characterization of the selective environment of the studied populations and (iii) the complementary methodological approach. It has shown that demography *(i.e.* migration and subsequent isolation with limited gene flow) is the dominant driver of genetic patterns in the studied populations but that phenotypic and genetic divergence driven by selection is present as well. A number of genes that are likely to be under selection were also discovered but more research is necessary for a biological interpretation of them. Maybe, the most stimulating result is the denying of unidirectional selection at regional scale. Instead, three axes of natural selection were discovered: (i) increasing winter drought from more oceanic to more continental mountains (ii) cold and humidity at Lure (iii) altitude as a multi-variable force.

Is *A. alba* capable of rapid adaptation to climate change? There neither is a simple nor a definite answer to this question. The species is ecologically confined to mountain systems. The upper tree-line in Alpine environments is strongly determined by temperature (Körner 1998). The results of chapter III suggest that that at Lure, the species is limited by cold. At this site increasing temperature could cause an upwards shift in the species distribution and ensure the persistence of highly viable populations. Above that, the distinct population at Lure may be pivotal for increased conservation efforts. Chapter III has also pinpointed

drought as the overall driver of local adaptation across all sites. At site where the species is limited by drought the resilience of the local population may not be sufficient to survive a further increase of temperature and frequency of heat waves. At Ventoux, recent die-offs and the lasting high rates of mortality (Cailleret *et al.* 2014) are obvious evidence for such a scenario.

Still, the species is indeed capable of a physiological response to a changing environment, *e.g*. elevated atmospheric CO_2 (Bert *et al.* 1997) and edaphic nitrogen concentrations (Pinto *et al.* 2007) have led to increased growth rates during the last century. The species shows also variability in photosynthetic efficiency (Peguero-Pina *et al.* 2007) and growth in relation to climate (Carrer *et al.* 2010; Rita *et al.* 2014; Gazol *et al.* 2015).

The populations studied in chapter III although being rather small, isolated and ecologically marginal exhibit low F_{st}. Genetic structure is dominantly shaped by neutral processes, but selection also is detectable as a driver of differentiation. Thus, even if climate change and habitat fragmentation may drive local populations of *A. alba* to extinction, the Mediterranean occurrence of the species is not immediately threatened since standing genetic variation is sufficient for evolutionary change. A number of genes were suggested to be under selection, however, the molecular basis of adaptation is not understood. The environmental drivers of adaptation are location-specific. Considering the lack of knowledge, assisted migration policy is not a practical option to mitigate climate change impacts in *A. alba* forests. Instead, *in-situ* conservation of the standing genetic variation of the species may support the long term persistence of the species. This thesis helps to highlight populations of special interest for *in-situ* conservation, namely the site Lure that is characterized by genetic and phenotypic distinctiveness is worth to be considered as an evolutionary significant unit.

7. References

Ahuja MR, Neale DB (2005) Evolution of genome size in conifers. *Silvae Genetica*, **54**, 126–137.

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.

Alberto FJ, Aitken SN, Alia R *et al.* (2013) Potential for evolutionary responses to climate change evidence from tree populations. *Global Change Biology*, **19**, 1645–1661.

Alleaume-Benharira M, Pen IR, Ronce O (2006) Geographical patterns of adaptation within a species' range: interactions between drift and gene flow. *Journal of Evolutionary Biology*, **19**, 203–215.

Allen GJ, Chu SP, Schumacher K *et al.* (2000) Alteration of Stimulus-Specific Guard Cell Calcium Oscillations and Stomatal Closing in Arabidopsis det3 Mutant. *Science*, **289**, 2338–2342.

Allendorf FW, Luikart G (2007) *Conservation and the genetics of populations*. Blackwell Pub, Malden, MA.

Allen CD, Macalady AK, Chenchouni H *et al.* (2010) A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*, **259**, 660–684.

Andersson B, Fedorkov A (2004) Longitudinal differences in Scots pine frost hardiness. *Silvae Genetica*, **53**, 76–79.

Arnerup J, Lind M, Olson Å, Stenlid J, Elfstrand M (2011) The pathogenic white-rot fungus Heterobasidion parviporum triggers non-specific defence responses in the bark of Norway spruce. *Tree Physiology*, **31**, 1262–1272.

Ashburner M, Ball CA, Blake JA *et al.* (2000) Gene Ontology: tool for the unification of biology. *Nature genetics*, **25**, 25–29.

Aussenac G (2002) Ecology and ecophysiology of circum-Mediterranean firs in the context of climate change. *Annals of Forest Science*, **59**, 823–832.

Baltunis BS, Wu HX, Powell MB (2007) Inheritance of density, microfibril angle, and modulus of elasticity in juvenile wood of Pinus radiata at two locations in Australia. *Canadian Journal of Forest Research*, **37**, 2164–2174.

Barakat A, DiLoreto D, Zhang Y *et al.* (2009) Comparison of the transcriptomes of American chestnut (Castanea dentata) and Chinese chestnut (Castanea mollissima) in response to the chestnut blight infection. *BMC Plant Biology*, **9**, 51.

Barrett RDH, Hoekstra HE (2011) Molecular spandrels: tests of adaptation at the genetic level. *Nature Reviews Genetics*, **12**, 767–780.

Battipaglia G, Saurer M, Cherubini P, Siegwolf RTW, Cotrufo MF (2009) Tree rings indicate different drought resistance of a native (Abies alba Mill.) and a nonnative (Picea abies (L.) Karst.) species co-occurring at a dry site in Southern Italy. *Forest Ecology and Management*, **257**, 820–828.

Beaulieu J, Doerksen T, Boyle B *et al.* (2011) Association Genetics of Wood Physical Traits in the Conifer White Spruce and Relationships With Gene Expression. *Genetics*, **188**, 197 –214.

Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.

Behringer D, Zimmermann H, Ziegenhagen B, Liepelt S (2015) Differential Gene Expression Reveals Candidate Genes for Drought Stress Response in Abies alba (Pinaceae). *PloS One*, **10**, e0124564 doi:10.1371

Bert D, Leavitt SW, Dupouey J-L (1997) Variations Of Wood δ13C and Water-Use Efficiency Of Abies Alba during the last Century. *Ecology*, **78**, 1588–1596.

Bigler C, Gavin DG, Gunning C, Veblen TT (2007) Drought induces lagged tree mortality in a subalpine forest in the Rocky Mountains. *Oikos*, **116**, 1983–1994.

Bigler C, Veblen TT (2009) Increased early growth rates decrease longevities of conifers in subalpine forests. *Oikos*, **118**, 1130–1138.

Birol I, Raymond A, Jackman SD *et al.* (2013) Assembling the 20 Gb white spruce (Picea glauca) genome from whole-genome shotgun sequencing data. *Bioinformatics*, **29**, 1492-7

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.

Breiman L (2001) Random Forests. *Machine Learning*, **45**, 5–32.

Budde KB, Heuertz M, Hernández-Serrano A *et al.* (2014) In situ genetic association for serotiny, a fire-related trait, in Mediterranean maritime pine (Pinus pinaster). *New Phytologist*, **201**, 230–241.

Burley J, Youngquist J, Evans J (Eds.) (2004) *Encyclopedia of forest sciences*. Elsevier, Oxford.

Bussotti F, Pollastrini M (2015) Evaluation of leaf features in forest trees: Methods, techniques, obtainable information and limits. *Ecological Indicators*, **52**, 219–230.

Bussotti F, Pollastrini M, Holland V, Brueggemann W (2015) Functional traits and adaptive capacity of European forests to climate change. *Environmental and Experimental Botany*, **111**, 91–113.

Cailleret M, Davi H (2011) Effects of climate on diameter growth of co-occurring Fagus sylvatica and Abies alba along an altitudinal gradient. *Trees*, **25**, 265–276.

Cailleret M, Nourtier M, Amm A, Durand-Gillmann M, Davi H (2014) Drought-induced decline and mortality of silver fir differ among three sites in Southern France. *Annals of Forest Science*, **71**, 643–657.

Camarero JJ, Gazol A, Sangüesa-Barreda G, Oliva J, Vicente-Serrano SM (2015) To die or not to die: early warnings of tree dieback in response to a severe drought. *Journal of Ecology*, **103**, 44–57.

Carlson SM, Cunningham CJ, Westley PAH (2014) Evolutionary rescue in a changing world. *Trends in Ecology & Evolution*, **29**, 521–530.

Carrer M, Nola P, Motta R, Urbinati C (2010) Contrasting tree-ring growth to climate responses of Abies alba toward the southern limit of its distribution area. *Oikos*, **119**, 1515–1525.

Chen J, Källman T, Ma X *et al.* (2012) Disentangling the Roles of History and Local Selection in Shaping Clinal Variation of Allele Frequencies and Gene Expression in Norway Spruce (Picea abies). *Genetics*, **191**, 865–881.

Chen C, Mitchell SE, Elshire RJ, Buckler ES, El-Kassaby YA (2013) Mining conifers' mega-genome using rapid and efficient multiplexed high-throughput genotyping-by-sequencing (GBS) SNP discovery platform. *Tree Genetics & Genomes*, **9**, 1537–1544.

Chhatre VE, Byram TD, Neale DB, Wegrzyn JL, Krutovsky KV (2013) Genetic structure and association mapping of adaptive and selective traits in the east Texas loblolly pine (Pinus taeda L.) breeding populations. *Tree Genetics & Genomes*, **9**, 1161–1178.

Comps B, Gömöry D, Letouzey J, Thiébaut B, Petit RJ (2001) Diverging Trends Between Heterozygosity and Allelic Richness During Postglacial Colonization in the European Beech. *Genetics*, **157**, 389–397.

Coop G, Witonsky D, Rienzo AD, Pritchard JK (2010) Using Environmental Correlations to Identify Loci Underlying Local Adaptation. *Genetics*, **185**, 1411–1423.

Corre VL, Kremer A (2003) Genetic Variability at Neutral Markers, Quantitative Trait Loci and Trait in a Subdivided Population Under Selection. *Genetics*, **164**, 1205–1219.

Cowling RM, Rundel PW, Lamont BB, Kalin Arroyo M, Arianoutsou M (1996) Plant diversity in mediterranean-climate regions. *Trends in Ecology & Evolution*, **11**, 362–366.

Csilléry K, Lalagüe H, Vendramin GG *et al.* (2014) Detecting short spatial scale local adaptation and epistatic selection in climate-related candidate genes in European beech (Fagus sylvatica) populations. *Molecular Ecology*, **23**, 4696–4708.

Cumbie WP, Eckert A, Wegrzyn J *et al.* (2011) Association genetics of carbon isotope discrimination, height and foliar nitrogen in a natural population of Pinus taeda L. *Heredity*, **107**, 105–114.

Davis MB, Shaw RG (2001) Range Shifts and Adaptive Responses to Quaternary Climate Change. *Science*, **292**, 673–679.

De Carvalho D, Ingvarsson PK, Joseph J *et al.* (2010) Admixture facilitates adaptation from standing variation in the European aspen (Populus tremula L.), a widespread forest tree. *Molecular Ecology*, **19**, 1638–1650.

De Kort H, Vandepitte K, Bruun HH *et al.* (2014) Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree species Alnus glutinosa. *Molecular Ecology*, **23**, 4709–4721.

De Kort H, Vandepitte K, Honnay O (2012) A meta-analysis of the effects of plant traits and geographical scale on the magnitude of adaptive differentiation as measured by the difference between QST and FST. *Evolutionary Ecology*, **27**, 1081–1097.

De Mita S, Thuillet A-C, Gay L *et al.* (2013) Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, **22**, 1383–1399.

Desplanque C, Rolland C, Schweingruber FH (1999) Influence of species and abiotic factors on extreme tree ring modulation: *Trees*, **13**, 218–227.

Didham RK (2010) Ecological Consequences of Habitat Fragmentation. In: *Encyclopedia of Life Sciences* (ed John Wiley & Sons, Ltd). John Wiley & Sons, Ltd, Chichester, UK.

Dillon SK, Nolan M, Li W *et al.* (2010) Allelic Variation in Cell Wall Candidate Genes Affecting Solid Wood Properties in Natural Populations and Land Races of Pinus radiata. *Genetics*, **185**, 1477 – 1487.

Ducousso A, Guyon J, Krémer A (1996) Latitudinal and altitudinal variation of bud burst in western populations of sessile oak (Quercus petraea (Matt) Liebl). *Annales des Sciences Forestières*, **53**, 775–782.

Eckert AJ, Bower AD, Wegrzyn JL *et al.* (2009a) Association Genetics of Coastal Douglas Fir (Pseudotsuga menziesii var. menziesii, Pinaceae). I. Cold-Hardiness Related Traits. *Genetics*, **182**, 1289 –1302.

Eckert AJ, van Heerwaarden J, Wegrzyn JL *et al.* (2010) Patterns of Population Structure and Environmental Associations to Aridity Across the Range of Loblolly Pine (Pinus taeda L., Pinaceae). *Genetics*, **185**, 969 –982.

Eckert CG, Samis KE, Lougheed SC (2008) Genetic variation across species' geographical ranges: the central–marginal hypothesis and beyond. *Molecular Ecology*, **17**, 1170–1188.

Eckert AJ, Wegrzyn JL, Cumbie WP *et al.* (2012) Association genetics of the loblolly pine (Pinus taeda, Pinaceae) metabolome. *New Phytologist*, **193**, 890–902.

Eckert AJ, Wegrzyn JL, Pande B *et al.* (2009b) Multilocus Patterns of Nucleotide Diversity and Divergence Reveal Positive Selection at Candidate Genes Related to Cold Hardiness in Coastal Douglas Fir (Pseudotsuga menziesii var. menziesii). *Genetics*, **183**, 289–298.

Ekblom R, Galindo J (2011) Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, **107**, 1–15.

Ellenberg H (1988) Vegetation Ecology of Central Europe. *Cambridge University Press.*

European Environment Agency (2012) Climate change, impacts and vulnerability in Europe 2012 An indicator-based report.

Evans R (2006) Wood Stiffness by X-Ray Diffractometry. In: *Characterization of the Cellulosic Cell Wall* (eds Stokke DD, Groom LH), pp. 138–146. Blackwell Publishing Professional.

Eveno E, Collada C, Guevara MA *et al.* (2008) Contrasting Patterns of Selection at Pinus pinaster Ait. Drought Stress Candidate Genes as Revealed by Genetic Differentiation Analyses. *Molecular Biology and Evolution*, **25**, 417 –437.

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

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Fady-Welterlen B (2005) Is There Really More Biodiversity in Mediterranean Forest Ecosystems? *Taxon*, **54**, 905–910.

Falk W, Mellert KH (2011) Species distribution models as a tool for forest management planning under climate change: risk evaluation of Abies alba in Bavaria. *Journal of Vegetation Science*, **22**, 621–634.

Fay JC, Wu CI (2000) Hitchhiking under positive Darwinian selection. *Genetics*, **155**, 1405–1413.

Fisichelli NA, Frelich LE, Reich PB (2014) Temperate tree expansion into adjacent boreal forest patches facilitated by warmer temperatures. *Ecography*, **37**, 152–161.

Foll M, Gaggiotti O (2008) A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics*, **180**, 977–993.

Food and Agriculture Organization of the United Nations (2013) State of Mediterranean forests 2013. *FAO*, [Rome].

FOREST EUROPE, UNECE and FAO (2011) State of Europe's forests 2011: status & trends in sustainable forest management in Europe *(Ministerial Conference on the Protection of Forests in Europe, Ed,). Ministerial Conference on the Protection of Forests in Europe,* Oslo.

Fournier-Level A, Korte A, Cooper MD *et al.* (2011) A Map of Local Adaptation in Arabidopsis thaliana. *Science*, **334**, 86–89.

França LTC, Carrilho E, Kist TBL (2002) A review of DNA sequencing techniques. *Quarterly Reviews of Biophysics*, **35**, 169–200.

Frichot E, Schoville SD, Bouchard G, François O (2013) Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, **30**, 1687-1699.

Frichot E, Schoville SD, de Villemereuil P, Gaggiotti OE, François O (2015) Detecting adaptive evolution based on association with ecological gradients: Orientation matters! *Heredity*, **115,** 22–28.

Fu YX, Li WH (1993) Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

Gabriel W, Bürger R (1994) Extinction risk by mutational meltdown: Synergistic effects between population regulation and genetic drift. In: *Conservation Genetics* EXS. (eds Loeschcke DV, Jain DSK, Tomiuk DJ), pp. 69–84. Birkhäuser Basel.

Gao X, Pal JS, Giorgi F (2006) Projected changes in mean and extreme precipitation over the Mediterranean region from a high resolution double nested RCM simulation. *Geophysical Research Letters*, **33**, L03706.

Garcia-Ramos G, Kirkpatrick M (1997) Genetic Models of Adaptation and Gene Flow in Peripheral Populations. *Evolution*, **51**, 21–28.

Gazol A, Camarero JJ, Gutiérrez E *et al.* (2015) Distinct effects of climate warming on populations of silver fir (Abies alba) across Europe. *Journal of Biogeography*, **42**, 1150-1162.

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

González-Martínez SC, Wheeler NC, Ersoz E, Nelson CD, Neale DB (2007) Association Genetics in Pinus taeda L. I. Wood Property Traits. *Genetics*, **175**, 399 –409.

Guisan A, Thuiller W (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.

Hampe A, Arroyo J, Jordano P, Petit RJ (2003) Rangewide phylogeography of a bird-dispersed Eurasian shrub: contrasting Mediterranean and temperate glacial refugia. *Molecular Ecology*, **12**, 3415–3426.

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, **8**, 461–467.

Hamrick J, Godt M (1996) Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **351**, 1291–1298.

Hamrick JL, Godt MJW, Sherman-Broyles SL (1992) Factors influencing levels of genetic diversity in woody plant species. *New Forests*, **6**, 95–124.

Hansen JK, Larsen JB (2004) European silver fir (Abies alba Mill.) provenances from Calabria, southern Italy: 15-year results from Danish provenance field trials. *European Journal of Forest Research*, **123**, 127–138.

Hereford J (2009) A Quantitative Survey of Local Adaptation and Fitness Trade-Offs. *The American Naturalist*, **173**, 579–588.

Hewitt GM (2004) Genetic consequences of climatic oscillations in the Quaternary. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, **359**, 183–195.

Hobbs RJ, Mooney HA (1998) Broadening the Extinction Debate: Population Deletions and Additions in California and Western Australia. *Conservation Biology*, **12**, 271–283.

Hoffmann AA, Blows MW (1994) Species borders: ecological and evolutionary perspectives. *Trends in Ecology & Evolution*, **9**, 223–227.

Holderegger R, Kamm U, Gugerli F (2006) Adaptive vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology*, **21**, 797–807.

Holderegger R, Wagner HH (2008) Landscape Genetics. *BioScience*, **58**, 199–207.

Holliday JA, Wang T, Aitken S (2012) Predicting Adaptive Phenotypes From Multilocus Genotypes in Sitka Spruce (Picea sitchensis) Using Random Forest. *G3: Genes Genomes Genetics*, **2**, 1085–1093.

Houle D, Govindaraju DR, Omholt S (2010) Phenomics: the next challenge. *Nature Reviews Genetics*, **11**, 855–866.

Hughes JB, Daily GC, Ehrlich PR (1997) Population Diversity: Its Extent and Extinction. *Science*, **278**, 689–692.

Huston MA (1994) *Biological diversity: the coexistence of species on changing landscapes*. Cambridge University Press, Cambridge ; New York, NY, USA.

Ingvarsson PK (2008) Multilocus Patterns of Nucleotide Polymorphism and the Demographic History of Populus tremula. *Genetics*, **180**, 329 –340.

Iverson LR, Peters MP, Matthews S, Prasad A In: Browning J and Palacios P (2013) An overview of some concepts, potentials, issues, and realities of assisted migration for climate change adaptation in forests. *Proceedings of the 60th Annual Western International Forest Disease Work Conference; 2012 Oct. 8-12; Tahoe City, CA,* 25-34

Källman T, De Mita S, Larsson H *et al.* (2014) Patterns of Nucleotide Diversity at Photoperiod Related Genes in Norway Spruce [Picea abies (L.) Karst.]. *PLoS ONE*, **9**, e95306. doi:10.1371

Kawecki TJ (2008) Adaptation to Marginal Habitats. *Annual Review of Ecology, Evolution, and Systematics*, **39**, 321–342.

Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225–1241.

Kerr G, Stokes V, Peace A, Jinks R (2015) Effects of provenance on the survival, growth and stem form of European silver fir (Abies alba Mill.) in Britain. *European Journal of Forest Research*, **134**, 349–363.

Kindermann GE, McCallum I, Fritz S, Obersteiner M (2008) A global forest growing stock, biomass and carbon map based on FAO statistics. *Silva Fennica* **42**, 387–396

King GM, Gugerli F, Fonti P, Frank DC (2013) Tree growth response along an elevational gradient: climate or genetics? *Oecologia*, **173**, 1587–1600.

Kleinknecht GJ, Lintz HE, Kruger A *et al.* (2015) Introducing a sensor to measure budburst and its environmental drivers. *Frontiers in Plant Science*, **6**, 123.

Klein Tank AMG, Können GP (2003) Trends in Indices of Daily Temperature and Precipitation Extremes in Europe, 1946–99. *Journal of Climate*, **16**, 3665–3680.

Knight H (1999) Calcium Signaling during Abiotic Stress in Plants. In: *International Review of Cytology* (ed Jeon KW), pp. 269–324. Academic Press.

Körner C (1998) A re-assessment of high elevation treeline positions and their explanation. *Oecologia*, **115**, 445–459.

Kozlowsky TT, Pallardy SG (1997) Physiology of woody plants. *Biologia Plantarum*.40, 168–168

Kreps JA, Wu Y, Chang H-S *et al.* (2002) Transcriptome Changes for Arabidopsis in Response to Salt, Osmotic, and Cold Stress. *Plant Physiology*, **130**, 2129–2141.

Kreyling J, Thiel D, Nagy L *et al.* (2011) Late frost sensitivity of juvenile Fagus sylvatica L. differs between southern Germany and Bulgaria and depends on preceding air temperature. *European Journal of Forest Research*, **131**, 717–725.

Kreyling J, Wiesenberg GLB, Thiel D *et al.* (2012) Cold hardiness of Pinus nigra Arnold as influenced by geographic origin, warming, and extreme summer drought. *Environmental and Experimental Botany*, **78**, 99–108.

Kültz D (2005) Molecular and Evolutionary Basis of the Cellular Stress Response. *Annual Review of Physiology*, **67**, 225–257.

Lande R (1995) Mutation and Conservation. *Conservation Biology*, **9**, 782–791.

Lange OL, Lösch R, Schulze E-D, Kappen L (1971) Responses of stomata to changes in humidity. *Planta*, **100**, 76–86.

Larsen JB (1986) Die geographische Variation der Weißtanne (Abies alba Mill.) Wachstumsentwicklung und Frostresistenz. *Forstwissenschaftliches Centralblatt*, **105**, 396–406.

Lebourgeois F (2013) Climatic signal in annual growth variation of silver fir (Abies alba Mill.) and spruce (Picea abies Karst.) from the French Permanent Plot Network (RENECOFOR). *Annals of Forest Science*, **64**, 333–343.

Le Corre V, Kremer A (2012) The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*, **21**, 1548–1566.

Lefèvre F, Boivin T, Bontemps A *et al.* (2013) Considering evolutionary processes in adaptive forestry. *Annals of Forest Science*, **71**, 723–739.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, **17**, 183–189.

Lesica P, Allendorf FW (1999) Ecological Genetics and the Restoration of Plant Communities: Mix or Match? *Restoration Ecology*, **7**, 42–50.

Lewontin RC, Krakauer J (1973) Distribution of Gene Frequency as a Test of the Theory of the Selective Neutrality of Polymorphisms. *Genetics*, **74**, 175–195.

Liepelt S, Bialozyt R, Ziegenhagen B (2002) Wind-dispersed pollen mediates postglacial gene flow among refugia. *Proceedings of the National Academy of Sciences*, **99**, 14590–14594.

Liepelt S, Cheddadi R, de Beaulieu JL *et al.* (2009) Postglacial range expansion and its genetic imprints in Abies alba (Mill.)—A synthesis from palaeobotanic and genetic data. *Review of Palaeobotany and palynology*, **153**, 139–149.

Linares JC, Camarero JJ (2012) Growth patterns and sensitivity to climate predict silver fir decline in the Spanish Pyrenees. *European Journal of Forest Research*, **131**, 1001–1012.

Lindquist EJ, Food and Agriculture Organization of the United Nations, European Commission (Eds.) (2012) *Global forest land-use change, 1990-2005*. Food and Agriculture Organization of the United Nations, Rome.

Linhart YB, Grant MC (1996) Evolutionary Significance of Local Genetic Differentiation in Plants. *Annual Review of Ecology and Systematics*, **27**, 237–277.

Liu L, Li Y, Li S *et al.* (2012) Comparison of Next-Generation Sequencing Systems. *BioMed Research International*, e251364 doi:10.1155/2012/251364

Liu J-J, Sturrock RN, Benton R (2013) Transcriptome analysis of Pinus monticola primary needles by RNA-seq provides novel insight into host resistance to Cronartium ribicola. *BMC Genomics*, **14**, 884.

Lotterhos KE, Whitlock MC (2014) Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, **23**, 2178–2192.

de Luis M, Čufar K, Di Filippo A *et al.* (2013) Plasticity in Dendroclimatic Response across the Distribution Range of Aleppo Pine (Pinus halepensis). *PLoS ONE*, **8**, e83550 doi:10.1371

Lynch M, Gabriel W (1990) Mutation Load and the Survival of Small Populations. *Evolution*, **44**, 1725–1737.

Macias M, Andreu L, Bosch O, Camarero JJ, Gutiérrez E (2006) Increasing Aridity is Enhancing Silver Fir Abies Alba Mill.) Water Stress in its South-Western Distribution Limit. *Climatic Change*, **79**, 289–313.

Mackay TFC (2001) The genetic architecture of quantitative traits. *Annual review of genetics*, **35**, 303–339.

Manel S, Schwartz MK, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, **18**, 189–197.

Martin PR, McKay JK, Noor M (2004) Latitudinal variation in genetic divergence of populations and the potential for future speciation. *Evolution*, **58**, 938–945.

Mátyás C (1996) Climatic adaptation of trees: rediscovering provenance tests. *Euphytica*, **92**, 45–54.

Mazza G (2013) Tree-Ring Growth Trends of Abies alba Mill: Possible Adaptations to Climate Change in Marginal Populations of Central Italy. *The Open Forest Science Journal*, **6**, 46–49.

McDonald JH, Kreitman M (1991) Adaptive protein evolution at the Adh locus in Drosophila. *Nature*, **351**, 652–654.

McDowell NG (2011) Mechanisms Linking Drought, Hydraulics, Carbon Metabolism, and Vegetation Mortality. *Plant Physiology*, **155**, 1051–1059.

McKay JK, Latta RG (2002) Adaptive population divergence: markers, QTL and traits. *Trends in Ecology & Evolution*, **17**, 285–291.

McLachlan JS, Hellmann JJ, Schwartz MW (2007) A Framework for Debate of Assisted Migration in an Era of Climate Change. *Conservation Biology*, **21**, 297–302.

Meinshausen M, Smith SJ, Calvin K *et al.* (2011) The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, **109**, 213–241.

Merlo M, Paiero P In: Merlo M, Croitoru L (Eds.) (2005) Valuing Mediterranean forests: towards total economic value. *CABI Pub*, Wallingford, Oxfordshire, UK ; Cambridge, MA.

Metzker ML (2010) Sequencing technologies - the next generation. *Nature Reviews. Genetics*, **11**, 31–46.

Mimura M, Aitken SN (2010) Local adaptation at the range peripheries of Sitka spruce. *Journal of Evolutionary Biology*, **23**, 249–258.

Mishima K, Fujiwara T, Iki T *et al.* (2014) Transcriptome sequencing and profiling of expressed genes in cambial zone and differentiating xylem of Japanese cedar (Cryptomeria japonica). *BMC Genomics*, **15**, 219.

Morin PA, Luikart G, Wayne RK, group the S workshop (2004) SNPs in ecology, evolution and conservation. *Trends in Ecology & Evolution*, **19**, 208–216.

Mosca E, González-Martínez SC, Neale DB (2013) Environmental versus geographical determinants of genetic structure in two subalpine conifers. *New Phytologist*, **201**, 180-192

Murray B, Leitch I, Bennett M (2012) Gymnosperm DNA C-values database. (release 5.0, Dec 2012). [WWW document] URL http://www.kew.org/cvalues [accessed December 2015]

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. *Nature*, **403**, 853–858.

Namroud M-C, Beaulieu J, Juge 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.

Narum SR, Hess JE (2011) Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, **11**, 184–194.

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

Neale DB, Wegrzyn JL, Stevens KA *et al.* (2014) Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biology*, **15**, R59 doi:10.1186

Nielsen R, Williamson S, Kim Y *et al.* (2005) Genomic scans for selective sweeps using SNP data. *Genome Research*, **15**, 1566–1575.

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

Novaes E, Drost D, Farmerie W *et al.* (2008) High-throughput gene and SNP discovery in Eucalyptus grandis, an uncharacterized genome. *BMC Genomics*, **9**, 312.

Nybom H, Bartish IV (2000) Effects of life history traits and sampling strategies on genetic diversity estimates obtained with RAPD markers in plants. *Perspectives in Plant Ecology, Evolution and Systematics*, **3**, 93–114.

Nystedt B, Street NR, Wetterbom A *et al.* (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature*, **497**, 579–584.

Ohta T (1982) Linkage Disequilibrium with the Island Model. *Genetics*, **101**, 139–155.

Oliva J, Colinas C (2007) Decline of silver fir (Abies alba Mill.) stands in the Spanish Pyrenees: Role of management, historic dynamics and pathogens. *Forest Ecology and Management*, **252**, 84–97.

Ozenda P (1975) Sur les étages de végétation dans les montagnes du bassin méditerranéen *Documents de cartographie écologique* **16**, 1-32

Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, **421**, 37–42.

Peguero-Pina JJ, Camarero JJ, Abadía A *et al.* (2007) Physiological performance of silver-fir (Abies alba Mill.) populations under contrasting climates near the south-western distribution limit of the species. *Flora - Morphology, Distribution, Functional Ecology of Plants*, **202**, 226–236.

Pérez-Figueroa A, García-Pereira MJ, Saura M, Rolán-Alvarez E, Caballero A (2010) Comparing three different methods to detect selective loci using dominant markers. *Journal of Evolutionary Biology*, **23**, 2267–2276.

Petit RJ, Aguinagalde I, De Beaulieu J-L *et al.* (2003) Glacial Refugia: Hotspots But Not Melting Pots of Genetic Diversity. *Science*, **300**, 1563–1565.

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

Pinto PE, Gégout J-C, Hervé J-C, Dhôte J-F (2007) Changes in environmental controls on the growth of Abies alba Mill. in the Vosges Mountains, north-eastern France, during the 20th century. *Global Ecology and Biogeography*, **16**, 472–484.

Piovani P, Leonardi S, Magnani F, Menozzi P (2011) Variability of stomatal conductance in a small and isolated population of silver fir (Abies alba Mill.). *Tree Physiology*, **31**, 500–507.

Piovani P, Leonardi S, Piotti A, Menozzi P (2010) Conservation genetics of small relic populations of silver fir (Abies alba Mill.) in the northern Apennines. *Plant Biosystems - An International Journal Dealing with all Aspects of Plant Biology*, **144**, 683–691.

Piper FI (2011) Drought induces opposite changes in the concentration of non-structural carbohydrates of two evergreen Nothofagus species of differential drought resistance. *Annals of Forest Science*, **68**, 415–424.

Porth I, El-Kassaby YA (2014) Assessment of the Genetic Diversity in Forest Tree Populations Using Molecular Markers. *Diversity*, **6**, 283–295.

Potočić N, Ćosić T, Pilaš I (2005) The influence of climate and soil properties on calcium nutrition and vitality of silver fir (Abies alba Mill.). *Environmental Pollution*, **137**, 596–602.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Prout T, Barker JS (1993) F statistics in Drosophila buzzatii: selection, population size and inbreeding. *Genetics*, **134**, 369–375.

Prunier J, Pelgas B, Gagnon F *et al.* (2013) The genomic architecture and association genetics of adaptive characters using a candidate SNP approach in boreal black spruce. *BMC Genomics*, **14**, 368.

Prunier J, Verta J-P, MacKay JJ (2016) Conifer genomics and adaptation: at the crossroads of genetic diversity and genome function. *New Phytologist*, **209**, 44–62.

Puizina J, Sviben T, Krajačić-Sokol I *et al.* (2008) Cytogenetic and molecular characterization of the Abies alba genome and its relationship with other members of the Pinaceae. *Plant Biology*, **10**, 256–267.

Quezel P, Medail F, Loisel R, Barbero M (1999) Biodiversity and conservation of forest species in the Mediterranean basin. *Unasylva (FAO)*.

Ralph SG, Yueh H, Friedmann M *et al.* (2006) Conifer defence against insects: microarray gene expression profiling of Sitka spruce (Picea sitchensis) induced by mechanical wounding or feeding by spruce budworms (Choristoneura occidentalis) or white pine weevils (Pissodes strobi) reveals large-scale changes of the host transcriptome. *Plant, Cell & Environment*, **29**, 1545–1570.

Ramankutty N, Foley JA (1999) Estimating historical changes in global land cover: Croplands from 1700 to 1992. *Global Biogeochemical Cycles*, **13**, 997–1027.

Rehfeldt GE, Tchebakova NM, Parfenova YI *et al.* (2002) Intraspecific responses to climate in Pinus sylvestris. *Global Change Biology*, **8**, 912–929.

Rellstab C, Gugerli F, Eckert AJ, Hancock AM, Holderegger R (2015) A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, **24**, 4348–4370.

Restoux G, Silva DE, Sagnard F *et al.* (2008) Life at the margin: the mating system of Mediterranean conifers. *Web Ecology*, **8**, 94–102.

Richards CM, Ritland AEK (2000) Inbreeding Depression and Genetic Rescue in a Plant Metapopulation. *The American Naturalist*, **155**, 383–394.

Rita A, Gentilesca T, Ripullone F, Todaro L, Borghetti M (2014) Differential climate–growth relationships in Abies alba Mill. and Fagus sylvatica L. in Mediterranean mountain forests. *Dendrochronologia*, **32**, 220–229.

Rockman MV (2012) The Qtn Program and the Alleles That Matter for Evolution: All That's Gold Does Not Glitter. *Evolution*, **66**, 1–17.

Rolland C, Desplanque C, Michalet R, Schweingruber FH (2000) Extreme Tree Rings in Spruce (Picea abies [L.] Karst.) and Fir (Abies alba Mill.) Stands in Relation to Climate, Site, and Space in the Southern French and Italian Alps. *Arctic, Antarctic, and Alpine Research*, **32**, 1–13.

Ronaghi M, Uhlén M, Nyrén P (1998) A Sequencing Method Based on Real-Time Pyrophosphate. *Science*, **281**, 363–365.

Roschanski AM, Csilléry K, Liepelt S *et al.* (2015) Evidence of divergent selection for drought and cold tolerance at landscape and local scales in Abies alba Mill. in the French Mediterranean Alps. *Molecular Ecology*, doi: 10.1111

Roschanski AM, Fady B, Ziegenhagen B, Liepelt S (2013) Annotation and Re-Sequencing of Genes from De Novo Transcriptome Assembly of *Abies alba* (Pinaceae). *Applications in Plant Sciences*, **1**, 1–8.

Roth R, Ebert I, Schmidt J (1997) Trisomy associated with loss of maturation capacity in a long-term embryogenic culture of Abies alba. *Theoretical and Applied Genetics*, **95**, 353–358.

Roussis V, Petrakis PV, Ortiz A, Mazomenos BE (1995) Volatile constituents of needles of five Pinus species grown in Greece. *Phytochemistry*, **39**, 357–361.

Ruelland E, Vaultier M-N, Zachowski A, Hurry V (2009) Chapter 2 Cold Signalling and Cold Acclimation in Plants. In: (ed Research B-A in B), pp. 35–150. Academic Press.

Sabeti PC, Reich DE, Higgins JM *et al.* (2002) Detecting recent positive selection in the human genome from haplotype structure. *Nature*, **419**, 832–837.

Sagnard F, Barberot C, Fady B (2002) Structure of Genetic diversity in Abies alba Mill. from southwestern Alps: multivariate analysis of adaptive and non-adaptive traits for conservation in France. *Forest Ecology and Management*, **157**, 175–189.

Sala A, Woodruff DR, Meinzer FC (2012) Carbon dynamics in trees: feast or famine? *Tree Physiology*, **32**, 764–775.

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.

Savolainen O, Pyhäjärvi T, Knürr T (2007) Gene Flow and Local Adaptation in Trees. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 595–619.

Scarascia-Mugnozza G, Oswald H, Piussi P, Radoglou K (2000) Forests of the Mediterranean region: gaps in knowledge and research needs. *Forest Ecology and Management*, **132**, 97–109.

Schoville SD, Bonin A, François O *et al.* (2012) Adaptive Genetic Variation on the Landscape: Methods and Cases. *Annual Review of Ecology, Evolution, and Systematics*, **43**, 23–43.

Schütt P (Ed.) (2008) Lexikon der Nadelbäume: die große Enzyklopädie mit über 800 Farbfotos unter Mitwirkung von 30 Experten. *Nikol, Hamburg.*

Schwartz MD, Reed BC, White MA (2002) Assessing satellite-derived start-of-season measures in the conterminous USA. *International Journal of Climatology*, **22**, 1793–1805.

Sevanto S, Mcdowell NG, Dickman LT, Pangle R, Pockman WT (2014) How do trees die? A test of the hydraulic failure and carbon starvation hypotheses. *Plant, Cell & Environment*, **37**, 153–161.

Sexton JP, Hangartner SB, Hoffmann AA (2014) Genetic Isolation by Environment or Distance: Which Pattern of Gene Flow Is Most Common? *Evolution*, **68**, 1–15.

Sexton JP, McIntyre PJ, Angert AL, Rice KJ (2009) Evolution and Ecology of Species Range Limits. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 415–436.

Shaw RG, Etterson JR (2012) Rapid climate change and the rate of adaptation: insight from experimental quantitative genetics. *New Phytologist*, **195**, 752–765.

Shendure J, Ji H (2008) Next-generation DNA sequencing. *Nat Biotech*, **26**, 1135–1145.

Sork VL, Aitken SN, Dyer RJ *et al.* (2013) Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes*, **9**, 901–911.

Spitze K (1993) Population structure in Daphnia obtusa: quantitative genetic and allozymic variation. *Genetics*, **135**, 367–374.

Stocker, T. F., Qin, D., Plattner, G. K. *et al.* (2013) IPCC, 2013: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14**, 671–688.

Tajima F (1989) Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics*, **123**, 585 –595.

Thiel D, Kreyling J, Backhaus S *et al.* (2014) Different reactions of central and marginal provenances of Fagus sylvatica to experimental drought. *European Journal of Forest Research*, **133**, 247–260.

Thomas CD, Cameron A, Green RE *et al.* (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.

Tiffin P, Ross-Ibarra J (2014) Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology & Evolution*, **29**, 673–680.

Tinner W, Colombaroli D, Heiri O *et al.* (2013) The past ecology of Abies alba provides new perspectives on future responses of silver fir forests to global warming. *Ecological Monographs*, **83**, 419–439.

Tsumura Y, Uchiyama K, Moriguchi Y *et al.* (2014) Genetic Differentiation and Evolutionary Adaptation in Cryptomeria japonica. *G3: Genes|Genomes|Genetics*, **4**, 2389–2402.

Tsumura Y, Uchiyama K, Moriguchi Y, Ueno S, Ihara-Ujino T (2012) Genome scanning for detecting adaptive genes along environmental gradients in the Japanese conifer, Cryptomeria japonica. *Heredity*, **109**, 349–360.

Vasemägi A (2006) The Adaptive Hypothesis of Clinal Variation Revisited: Single-Locus Clines as a Result of Spatially Restricted Gene Flow. *Genetics*, **173**, 2411–2414.

Vasemägi A, Primmer CR (2005) Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Molecular Ecology*, **14**, 3623–3642.

Vendramin GG, Degen B, Petit RJ *et al.* (1999) High level of variation at Abies alba chloroplast microsatellite loci in Europe. *Molecular Ecology*, **8**, 1117–1126.

Vié J-C, Hilton-Taylor C, Stuart SN, IUCN--The World Conservation Union, IUCN Species Survival Commission (Eds.) (2009) *Wildlife in a changing world: an analysis of the 2008 IUCN red list of threatened species*. IUCN ; Lynx Edicions, Gland, Switzerland : Barcelona, Spain.

Vierling E (1991) The Roles of Heat Shock Proteins in Plants. *Annual Review of Plant Physiology and Plant Molecular Biology*, **42**, 579–620.

Viherä-Aarnio A, Häkkinen R, Partanen J, Luomajoki A, Koski V (2005) Effects of seed origin and sowing time on timing of height growth cessation of Betula pendula seedlings. *Tree Physiology*, **25**, 101–108.

de Villemereuil P, Frichot É, Bazin É, François O, Gaggiotti OE (2014) Genome scan methods against more complex models: when and how much should we trust them? *Molecular Ecology*, **23**, 2006–2019.

Wachowiak W, Salmela MJ, Ennos RA, Iason G, Cavers S (2011) High genetic diversity at the extreme range edge: nucleotide variation at nuclear loci in Scots pine (Pinus sylvestris L.) in Scotland. *Heredity*, **106**, 775–787.

Wang G, Gao Y, Yang L, Shi J (2007) Identification and analysis of differentially expressed genes in differentiating xylem of Chinese fir (Cunninghamia lanceolata) by suppression subtractive hybridization. *Genome / National Research Council Canada,* **50**, 1141–1155.

Wegrzyn JL, Eckert AJ, Choi M *et al.* (2010) Association genetics of traits controlling lignin and cellulose biosynthesis in black cottonwood (Populus trichocarpa, Salicaceae) secondary xylem. *New Phytologist*, **188**, 515–532.

Wegrzyn JL, Liechty JD, Stevens KA *et al.* (2014) Unique Features of the Loblolly Pine (Pinus taeda L.) Megagenome Revealed Through Sequence Annotation. *Genetics*, **196**, 891–909.

Whitlock MC (2008) Evolutionary inference from QST. *Molecular Ecology*, **17**, 1885–1896.

Whitlock MC, Guillaume F (2009) Testing for Spatially Divergent Selection: Comparing QST to FST. *Genetics*, **183**, 1055–1063.

Wolf H (2003) EUFORGEN Technical Guidelines for genetic conservation and use for for silver fir (Abies alba). *Bioversity International.*

Wright S (1950) Genetical structure of populations. *Nature*, **166**, 247–249.

Young A, Boyle T, Brown T (1996) The population genetic consequences of habitat fragmentation for plants. *Trends in Ecology & Evolution*, **11**, 413–418.

Publications and manuscripts of the cumulative part

Chapter I

Roschanski AM, Fady B, Ziegenhagen B, Liepelt S (2013) **Annotation and Re-Sequencing of Genes from De Novo Transcriptome Assembly of *Abies alba* (Pinaceae).** *Applications in Plant Sciences*, 1, 1–8.

Applications in Plant Sciences

Applications in Plant Sciences 2013 1(1): 1200179

Protocol Note

Annotation and re-sequencing of genes from de novo transcriptome assembly of *Abies alba* (Pinaceae)[1]

Anna M. Roschanski[2,4], Bruno Fady[3], Birgit Ziegenhagen[2], and Sascha Liepelt[2]

[2]University of Marburg, Faculty of Biology, Conservation Biology, Karl-von-Frisch-Strasse 35032 Marburg, Germany; and
[3]INRA, UR629, Ecologie des Forêts Méditerranéennes (URFM), 84914 Avignon, France

- *Premise of the study:* We present a protocol for the annotation of transcriptome sequence data and the identification of candidate genes therein using the example of the nonmodel conifer *Abies alba*.
- *Methods and Results:* A normalized cDNA library was built from an *A. alba* seedling. The sequencing on a 454 platform yielded more than 1.5 million reads that were de novo assembled into 25 149 contigs. Two complementary approaches were applied to annotate gene fragments that code for (1) well-known proteins and (2) proteins that are potentially adaptively relevant. Primer development and testing yielded 88 amplicons that could successfully be resequenced from genomic DNA.
- *Conclusions:* The annotation workflow offers an efficient way to identify potential adaptively relevant genes from the large quantity of transcriptome sequence data. The primer set presented should be prioritized for single-nucleotide polymorphism detection in adaptively relevant genes in *A. alba*.

Key words: *Abies alba*; adaptation; annotation; candidate genes; de novo sequencing; Pinaceae.

To gain insights into the molecular level of adaptation, attention has turned to the investigation of adaptively relevant genes (candidate genes). For nonmodel organisms, access to candidate genes is limited and the transfer of primers, e.g., from expressed sequence tag (EST) libraries, if available, requires high labor costs. For instance, the resequencing of 800 genes selected from more than 7000 ESTs from *Pinus taeda* L. yielded only 70 candidate genes for *Abies alba* Mill. (Mosca et al., 2012). Because sequencing costs are decreasing rapidly, de novo sequencing in nonmodel organisms is now achievable. For the identification of candidate genes in de novo–sequenced organisms, the use of differential expression profiling (e.g., Street et al., 2006; Huang et al., 2012) can be performed, but it requires the sequencing of several samples. The sequencing of a single transcriptome, in contrast, is very cost-effective. However, the reduction of the data remains challenging. Blasting against available databases is the standard method, which results in outputs of large quantities and is therefore mainly used for annotation only (e.g., Parchman et al., 2010). Here, we present a protocol for the efficient reduction of transcriptomic data down to 283 candidate gene sequences that were used for immediate primer development. The protocol is applicable for species that lack genomic resources. It combines a standard and a specific annotation approach and led to the resequencing of 88 gene fragments in *A. alba*.

[1]Manuscript received 13 April 2012; revision accepted 30 June 2012.

The authors thank E. Utesch for technical assistance. We further thank the collaborators from the LinkTree Project and G. Müller-Starck. This work was funded by the ERA-net BiodivERsA LinkTree Project and by the EU Network of Excellence EvolTree.

[4]Author for correspondence: anna.roschanski@biologie.uni-marburg.de

doi:10.3732/apps.1200179

METHODS AND RESULTS

A normalized transcriptome of a 1-yr-old *A. alba* seedling from the Black Forest (Forest District Calw, Germany; voucher MB-P-001007, Herbarium Marburgense, University of Marburg) was sequenced on a 454 GS FLX Titanium platform (cDNA library preparation: Vertis Biotechnology AG, Freising, Germany; sequencing: Genoscreen, Lille, France). The 454 run yielded 1 521 698 reads with an average length of 359 nucleotides (nt). Trimming and de novo assembly of the raw reads into contigs using Newbler software version 2.3 (454 Life Sciences, Branford, Connecticut, USA) resulted in 25 149 contigs consisting of 381 808 complete and 619 615 partially assembled reads. The contig length was between 100 nt and 2394 nt, with an average length of 498 nt. A total of 484 576 reads remained as singletons (Table 1). Contigs were submitted to the Transcriptome Shotgun Assembly database (TSA) at the National Center for Biotechnology Information (NCBI) (accession no.: JV134525–JV157085).

In the specific approach (Fig. 1), we tested a novel annotation protocol: After a literature survey with key words "adaptation," "candidates," "drought," "evolution," "RT-PCR," and "selection" in various combinations using the Web of Science database, we selected 5349 unique proteins and downloaded them from UniProt or NCBI (downloaded in November 2011). The proteins were subsequently searched against the contigs coming from the de novo transcriptome sequencing that were formatted as the reference database using the BLAST+ 2.2.24 toolkit (tBLASTn parameters: softmasking = threshold 15 max_target_seqs 10 000). To increase reliability of alignments and to avoid too-short amplicons, only alignments with a length of at least 100 amino acids and an identity of at least 90% were considered further. From the contigs that passed the filter, 157 were selected for primer design. In the standard approach (Fig. 1), contigs were searched against the refseq_protein database (downloaded from NCBI 14 June 2011) with strict BLAST-settings (BLASTx parameters: threshold 999, window-size 4, gapopen 32767, gapextend 32767, *E*-value $1e^{-20}$) (Altschul et al., 1990). Gene ontologies (Ashburner et al., 2000) were assigned to contig-protein hits using Blast2GO 2.5.0 (Conesa et al., 2005) and subsequently filtered as described above. To select for well-described proteins, contig sequences were used for primer design if they could be assigned to enzyme IDs with the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Ogata et al., 1999) in the final annotation step. Primers were developed specifying the amplified range according to the contig-protein alignment boundaries using default standard PCR settings of PerlPrimer (version 1.1.12; Marshall, 2004). Primers were tested in a 30 µL PCR reaction with 17.28 µL double-distilled water, 3 µL 10× PCR buffer with $MgCl_2$ (20 mM), 1.2 µL $MgCl_2$ (25 mM), 3 µL Primermix (forward and reverse each 2 µM), 1.44 µL dNTPs (each 5 mM),

TABLE 1. Statistics of the 454 transcriptome sequencing run and metrics of the Newbler assembly software.

Sequence type	Number	%	Nucleotides	Average (nt)	Size (nt) in quantiles				
					0%	25%	50%	75%	100%
Reads trimmed	1 521 698	100	546 346 058	359.0	<21	<303	<395	<444	<1088
Reads assembled	381 808	25.1							
Reads partial	619 615	40.7							
Reads singleton	484 576	31.8	175 198 711	361.6	<50	<307	< 397	< 443	<876
Reads repeat	1617	0.1							
Reads outlier	20 389	1.3							
Reads too short	13 693	0.9							
Contigs	25 149		12 511 848	498	<100	<365	<468	<601	<2394
N50 Contig[a]			704						

[a] Half of all bases are assembled in contigs of this size or longer.

0.24 µL bovine serum albumin (BSA) (20 mg mL^{-1}), 0.24 µL Dream *Taq* polymerase (5 U µL^{-1}, Fermentas, St. Leon-Rot, Germany), and 3.6 µL DNA (10 ng µL^{-1}). The PCR was performed with 5 min initial denaturation at 94°C followed by 35 repetitions of 45 s denaturation at 94°C, 45 s annealing at 52–59°C, 45 s elongation at 72°C, and a 10 min final elongation at 72°C. For the amplification test, four samples were randomly chosen for each gene from a set of 80 different silver fir trees that were sampled in May 2011 in Mont Ventoux (44°10′44.35″N, 5°14′32.29″E, France). Amplification was evaluated by electrophoresis in 1% agarose gels. When amplification was too weak, the volume of $MgCl_2$ was increased to 1.8 µL. When faint ancillary bands appeared, no additional magnesium was added to the mastermix. If PCR products occurred as a single band, one sample was chosen for sequence analysis in each case to ensure that the region of interest was amplified (LGC Genomics GmbH, Berlin, Germany). Gene sequences were aligned to the corresponding contigs using the CodonCode Aligner software (default large gap settings) to reveal the location of the introns. The gene sequences were searched against the nr nucleotide database of NCBI (default discontiguous megaBLAST settings, web application).

In the specific approach, tBLASTn and subsequent sorting led to 321 contigs. For primer development, 185 contigs were picked. In the standard approach, the initial number of contigs was decreased to one third after the BLASTx step. Approximately half of the hits could be further annotated with Gene Ontologies. After filtering, 126 contigs were successfully assigned to enzyme-IDs and used for primer design (Fig. 1). In combination, 283 different contigs were annotated and only 28 were annotated with both approaches. Primer testing and sequencing resulted in 88 gene sequences (Table 2). Fifty-seven genes were annotated using the specific approach, and 42 using the standard approach. Eleven were annotated by both approaches. The assembly of the gene sequences and the corresponding cDNA contigs revealed 43 introns in 26 genes. The length of the gene sequences ranged from 262 to 1486 nt. All gene sequences aligned to sequences from the nr nucleotide database (NCBI) where the highest *E*-value was 5.00e^{-32}. Twelve gene sequences hit organelle DNA (10 chloroplast, one mitochondrial, and one ribosomal). The remaining 76 are involved in the biosynthesis of different compounds (21), regulation (20), primary metabolism (14), growth (11), stress response (8), and water transport (2). In the biosynthesis group, enzymes from the auxin pathways, the phenylpropanoid pathways, and the tetrapyrrol pathways were dominant. With the exception of the primary metabolism group, all groups included candidates for the analysis of adaptation at gene level that had been investigated in previous studies of conifers (e.g., González-Martínez et al., 2006).

CONCLUSIONS

The two approaches of the workflow are complementary, each contributing approximately half of the annotations in the final set of sequences. The standard approach can be run rapidly, but targets only well-known genes. The specific approach based on a review of the relevant literature is novel and provided a substantial amount of nonredundant annotations. As an advantage, it can be easily adjusted and extended freely to the researcher's interest. The quality-tested primers can be used for assessing the degree of gene polymorphism in ecological genetics studies.

LITERATURE CITED

ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS, AND D. J. LIPMAN. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215: 403–410.

ASHBURNER, M., C. A. BALL, J. A. BLAKE, D. BOTSTEIN, H. BUTLER, J. M. CHERRY, A. P. DAVIS, ET AL. 2000. Gene Ontology: Tool for the unification of biology. *Nature Genetics* 25: 25–29.

CONESA, A., S. GÖTZ, J. M. GARCÍA-GÓMEZ, J. TEROL, M. TALÓN, AND M. ROBLES. 2005. Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics (Oxford, England)* 21: 3674–3676.

GONZÁLEZ-MARTÍNEZ, S. C., E. ERSOZ, G. R. BROWN, N. C. WHEELER, AND D. B. NEALE. 2006. 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.

HUANG, H.-R., P.-C. YAN, M. LASCOUX, AND X.-J. GE. 2012. Flowering time and transcriptome variation in *Capsella bursa pastoris* (Brassicaceae). *New Phytologist* 194: 676–689.

MARSHALL, O. 2004. PerlPrimer: Cross-platform, graphical primer design for standard, bisulphite, and real-time PCR. *Bioinformatics* 20: 2471–2472.

MOSCA, E., A. J. ECKERT, J. D. LIECHTY, J. L. WEGRZYN, N. LA PORTA, G. G. VENDRAMIN, AND D. B. NEALE. 2012. Contrasting patterns of nucleotide diversity for four conifers of Alpine European forests. *Evolutionary Applications* 5: 762–775..

OGATA, H., S. GOTO, K. SATO, W. FUJIBUCHI, H. BONO, AND M. KANEHISA. 1999. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research* 27: 29–34.

PARCHMAN, T. L., K. S. GEIST, J. A. GRAHNEN, C. W. BANKMAN, AND C. A. BUERKLE. 2010. Transcriptome sequencing of an ecologically important tree species: Assembly, annotation, and marker discovery. *BMC Genomics* 11: 180.

STREET, N. R., O. SKOGSTRÖM, A. SJÖDIN, J. TUCKER, M. RODRÍGUEZ-ACOSTA, P. NILSSON, S. JANSSON, AND G. TAYLOR. 2006. The genetics and genomics of the drought response in *Populus*. *Plant Journal* 48: 321–341.

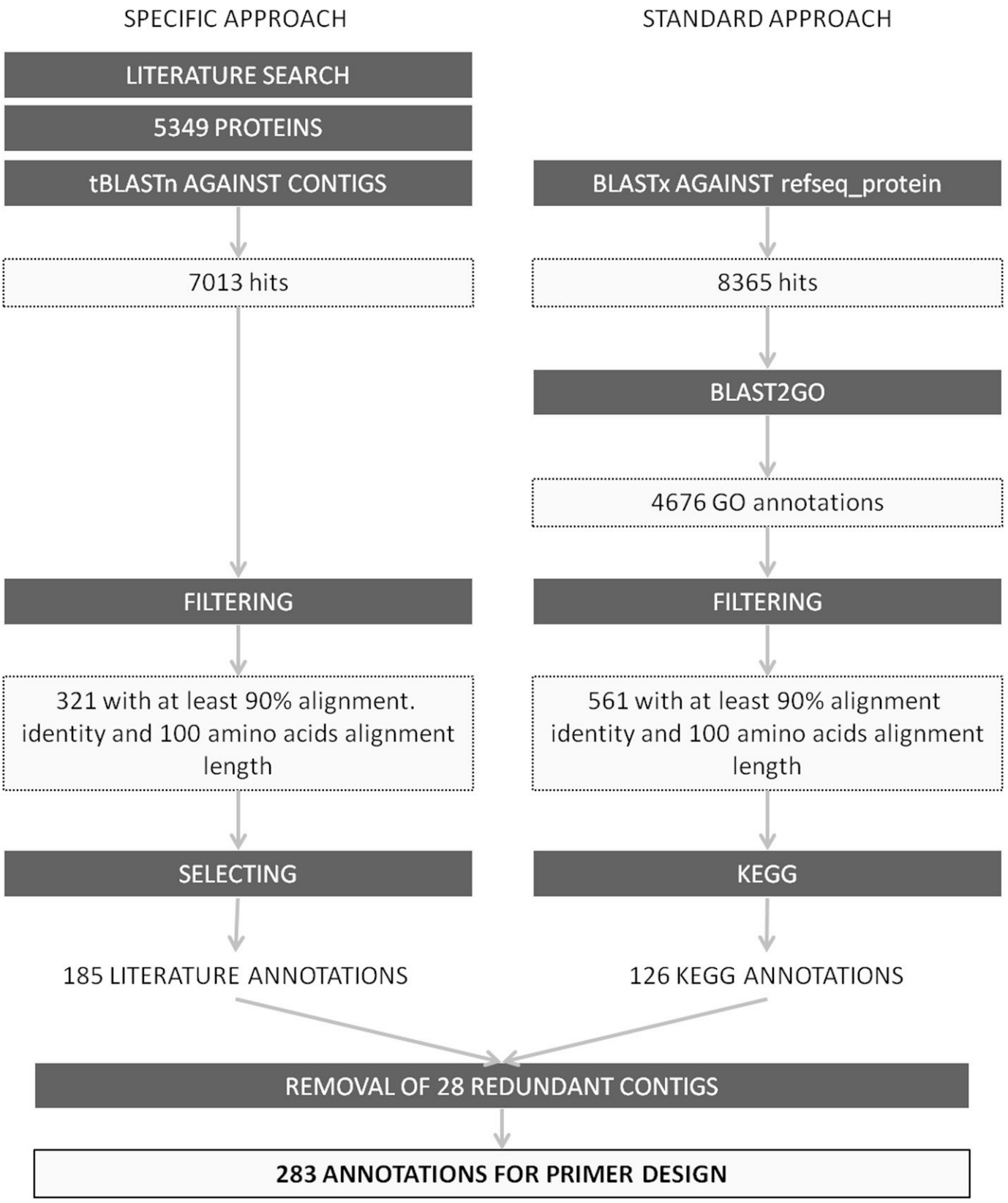

Fig. 1. Workflow of the annotation protocol. Numbers of the output after each step are given. The standard approach starts with 25 149 contigs. The specific approach uses them as the reference database for the tBLASTn step.

TABLE 2. Primers for resequencing of annotated gene fragments in *Abies alba*.[a]

Gene Locus ID	Primer sequences (5′–3′)	T_a (°C)	No. of introns	Intron length (nt)	Total length (nt)	Annotation approach	BLASTn of gene sequences against nr nucleotide database (*E*-value)	GO-ID biological process
95	F: ACAGAAACTAAAGCTAGTGTCG R: CCTTAATTTCACCCGTCTCAG	57	0	—	696	1	*Keteleeria davidiana* chloroplast DNA, complete sequence (0)	reductive pentose-phosphate cycle, photorespiration, oxidation reduction
215	F: CCAAGGACTCTGATCGAATCC R: GAAGCCAGCATTCAAAGACTC	56	2	411	1486	2	*Abies firma* clone 1 4-coumarate:CoA ligase (4CL) gene, partial cds (0)	response to UV, response to wounding, phenylpropanoid metabolic process, response to fungus
241	F: AACGTCCGTTAATACTTCGG R: AGTAAGTGTAGCCCTTCACG	56	3	256	1370	2	*Arabidopsis thaliana* fructose-bisphosphate aldolase, class I (FBA1) mRNA, complete cds (1E-125)	glycolysis
323	F: AAGCAAGCTTCTGAAATTCC R: TGGTAGAGTCTACCAAATGAG	53	2	278	804	1	*A. thaliana* plasma membrane H+-ATPase gene, complete cds (1E-90)	auxin biosynthetic process, ATP biosynthetic process, proton transport
1362	F: GAAGAGGTAGCTGCATTGGT R: GGGCTTATACCGTAAATATACCCA	59	0	—	871	1	*Ricinus communis* processing-splicing factor, putative, mRNA (0.0)	response to hypoxia, sucrose biosynthetic process, nuclear mRNA splicing, via spliceosome
1704	F: CAACTACTTCAGAGACAGAC R: AAAGATTCCCTCCAAATCAG	52	2	327	858	2	*Pinus taeda* mitogen-activated protein kinase 13 (MAPK13) mRNA, complete cds (2E-84)	
2387	F: TAAATGGCTCAATTCCTCCTACTG R: GTTCCAAGCTTCCACAATACTC	61	1	128	624	1	*Medicago truncatula* Alpha-1,4-galacturonosyltransferase (MTR_7g075840) mRNA, complete cds (8E-99)	
2565	F: GTGTCTGGAAGGGAATACAAGG R: CCTTGACTCCTTCATGGATCAG	58	0	—	432	1	PREDICTED: *Vitis vinifera* adenosylhomocysteinase-like, transcript variant 1 (LOC100253872), mRNA (1E-109)	embryonic development ending in seed dormancy, one-carbon metabolic process, posttranscriptional gene silencing, methylation-dependent chromatin silencing
2774	F: GTTACAGGAAGCCTTTCTGG R: GCGGGATGAATTATCTTGTC	55	0	—	502	2	*Citrus sinensis* pectinesterase mRNA, complete cds (5E-32)	cell wall modification
2937	F: TGAGCTGATTGCTAATGCGG R: GGACATGGTGGTCATTGAGG	58	0	—	622	2	*Solanum tuberosum* clone 154D06 fructose-bisphosphate aldolase-like mRNA, complete cds (5E-120)	glycolysis
2986	F: CTGTCTGTGACGGATCTAGC R: TGAGGATGGCTTACAACACG	57	0	—	355	1	*Populus trichocarpa* arogenate/prephenate dehydratase (PDT1), mRNA (1E-52)	L-phenylalanine biosynthetic process
3421	F: CTCATCTCTGCCAGAAAGAC R: GTAGAGCTTCATCTACGAGG	55	0	—	324	2	*Picea sitchensis* isolate CR201 phenylalanine ammonia lyase-like protein mRNA, partial cds (0.0)	phenylpropanoid metabolic process, biosynthetic process, L-phenylalanine catabolic process
3593	F: AGGACCTGAAATACCTTGCT R: TCCGTGTTTATCTCACAGGT	56	0	—	337	2	*Abies firma* chloroplast, partial genome (6E-170)	transport, respiratory electron transport chain, photosynthesis
3689	F: CGATTGCATCTCTGTACGCC R: GCTCTTGAGCCTCTTGACAC	58	0	—	619	2	*Pseudotsuga menziesii* var. *menziesii* haplotype Pm-TBE_412m2 thiazole biosynthetic enzyme (TBE) gene, complete cds (0.0)	thiamin biosynthetic process
3918	F: TTCCAAGGTCTTCTCAAGGT R: TGAAGAGTAGGAGTTTCGGT	55	0	—	400	2	*Pinus taeda* cellulose synthase catalytic subunit (CesA1) mRNA, complete cds (0.0)	cellulose biosynthetic process, cellular cell wall organization, secondary cell wall biogenesis, rhamnogalacturonan I side chain metabolic process
3942	F: GTATGATACCGATGTGACGA R: TTTGTAATGGATGCACTCGG	55	0	—	273	2	*Ricinus communis* proteasome subunit alpha type, putative, mRNA (8E-48)	ubiquitin-dependent protein catabolic process
3981	F: GGAGAAGTCTACAGTTCCAG R: ATAGTCCAGTGTCTTGAACTC	54	0	—	918	1	*Pinus radiata* UDP-glucose dehydrogenase gene, partial sequence (0.0)	oxidation reduction
4103	F: ATGGCCACCTTACTAAGAAGC R: CCACTTAAGGACCTTTACAGTCTC	57	0	—	841	1	*Pinus pinaster* mRNA for *S*-adenosylmethionine synthase 1 (sams1 gene) (0.0)	auxin biosynthetic process, one-carbon metabolic process
4492	F: TGGGTGCAACTGAAGATAGAG R: TTTCTACAACTAGCAAGCCTGAG	57	0	—	698	1	*Medicago truncatula* magnesium-chelatase subunit chlI (MTR_2g015390) mRNA, complete cds (4E-160)	auxin biosynthetic process, chlorophyll biosynthetic process, photosynthesis

TABLE 2. Continued.

Gene Locus ID	Primer sequences (5′–3′)	T_a (°C)	No. of introns	Intron length (nt)	Total length (nt)	Annotation approach	BLASTn of gene sequences against nr nucleotide database (*E*-value)	GO-ID biological process
4921	F: GAAGGTCGGCTATATCAGGT R: AGCTTAGACAGAGACTCAGG	56	0	—	664	2	PREDICTED: *Glycine max* proteasome subunit alpha type-4-like, transcript variant 1 (LOC100786457), mRNA (2E-147)	response to cadmium ion, ubiquitin-dependent protein catabolic process
5004	F: CAGATGTGAGCCATTACTTTGAC R: CAACCTCTGAATATAGCTGCCT	57	0	—	461	1	*Picea sitchensis* isolate VD401 magnesium chelatase H-like protein mRNA, partial cds (0)	chlorophyll biosynthetic process
5823	F: TGCTTGATATACGTCCTGGG R: CTAGACAGTGTTGCTCCACG	57	0	—	293	2	*Picea sitchensis* isolate VD401 phytochrome A-like protein mRNA, partial cds (0)	regulation of transcription, photomorphogenesis, tryptophan biosynthetic process
5945	F: CTGTCACTCAGATCTTCAGC R: AGATGATCAGCGAGATTCTC	55	0	—	339	2	*P. abies* (L.) Karst. Lhcb1*2-2 mRNA for light-harvesting chlorophyll a/b-binding protein (0.0)	photosynthesis, light harvesting, protein-chromophore linkage
6119	F: AGAGGATGTTGGGCATTATGG R: CATCACATGGTATCTCATCCGA	57	0	—	567	1	*Picea mariana* pyruvate dehydrogenase E1 beta subunit (Sb68) mRNA, partial cds (0.0)	pollen tube development, oxidation reduction
6594	F: TGGCTTTATCTTGGAGACTTCAC R: GAATAAGGTCATAGCCTGCCG	58	1	348	712	1	*Ricinus communis* phosphatidylinositol 4-kinase, putative, mRNA (5E-51)	phosphoinositide biosynthetic process, phosphoinositide phosphorylation, signal transduction, phosphoinositide-mediated signaling
6757	F: TATCATGCCCTGAAAGCGTC R: ACTTCCACAAGCAAGACACTC	58	5	177	939	1	*Arabidopsis thaliana* ribonucleoside-diphosphate reductase subunit M1 (RNR1) mRNA, complete cds (1E-39)	
7098	F: CTTTACTGTTGGAGGTAGATCAG R: GTTTGTTTGTCTTTGTACTCCC	55	0	—	782	1	*Arabidopsis thaliana* UDP-glucuronic acid decarboxylase (AUD1) mRNA, complete cds (1E-153)	dTDP-rhamnose biosynthetic process, D-xylose metabolic process
7208	F: GTTACATTCGTAAGTAGCTTGG R: AAATGGTCGAGAAGTCTACTG	54	0	—	326	1	*Pinus thunbergii* NADH dehydrogenase subunit 5 (nad5) gene, partial cds; mitochondrial (0)	transport, ATP synthesis coupled electron transport
7324	F: ATTGGAGATGGAGCCATGAC R: TCTCTGCATATGGGTAACCC	57	0	—	471	1	*Picea abies* 1-deoxy-D-xylulose 5-phosphate synthase type I (DXS1) mRNA, complete cds (0)	terpenoid biosynthetic process, thiamin biosynthetic process
8248	F: CAAGTATTCCGAAAGGCAGC R: ACAAAGGTGCCCACAATCTC	57	1	601	1128	2	*P. abies* mRNA for porin Mip1 (3E-154)	response to water deprivation, water transport, transmembrane transport, response to salt stress
8583	F: TCTCCTACATTGACGATCCC R: CCATCCAAGCACTTGAAGAG	56	0	—	393	2	*Picea sitchensis* isolate VA301 phenylalanine ammonia lyase-like protein mRNA, partial cds (5E-162)	phenylpropanoid metabolic process, biosynthetic process, L-phenylalanine catabolic process
8855	F: TATTTGCTGGTCGGGATTCG R: CTGCACTAGGTTCTCGAACG	58	2	275	926	2	*P. sylvestris* Lhca4*1-2 mRNA encoding Lhca4 protein (type 4 protein of light-harvesting complex of photosystem I) (partial) (7E-179)	photosynthesis, light harvesting
9366	F: AGTGAAAGCAACAACTTAGG R: TCTGGCTTCATTGATTTGTC	53	0	—	598	1	*Tamarix hispida* peroxiredoxin 2 (Prx2) mRNA, complete cds (5E-139)	cell redox homeostasis, oxidation reduction
9512	F: GTACTGGAGTAGCTGCACGA R: TACAAAGTGCTGCACAGCAG	59	1	99	415	2	*Cycas revoluta* class III HD-Zip protein HDZ32 gene, partial cds (4E-52)	regulation of transcription, DNA-dependent
9652	F: TGCAAAGAAAGTCAAGGCGA R: CCCATACGGTGTTAATGGCT	58	2	418	913	2	*Pinus pinaster* COBRA-like protein gene, partial cds (0)	
11301	F: GATGTTGTTCGTGCAAAGAC R: GCGAACTTAATTCCCTTCTC	54	0	—	490	2	*Pinus pinaster* mRNA for malate dehydrogenase (MDH gene) (0.0)	malate metabolic process, oxidation reduction, tricarboxylic acid cycle, glycolysis
13329	F: GATATGTGCCCAAGAACATTCTG R: CCTTGCATGCTTCAAGAAGG	57	0	—	350	1	PREDICTED: *Glycine max* probable rhamnose biosynthetic enzyme 1-like (LOC100789909), mRNA (7E-87)	
13536	F: CTGCTGATTCTGATCAGTCC R: TCCACAATGCAAACATAGGC	56	0	—	368	2	*Pinus thunbergii* PtANTL1 mRNA for AINTEGUMENTA-like protein, complete cds (0.0)	regulation of transcription, DNA-dependent

TABLE 2. Continued.

Gene Locus ID	Primer sequences (5′–3′)	T_a (°C)	No. of introns	Intron length (nt)	Total length (nt)	Annotation approach	BLASTn of gene sequences against nr nucleotide database (*E*-value)	GO-ID biological process
14455	F: GAACAAGATCGACTACTGCC R: TTTGATGGCCTTGAAAGCAG	56	0	—	834	2	*Pinus taeda* mRNA for alpha-1, 6-xylosyltransferase (x34.1 gene) (0.0)	root hair elongation, xyloglucan biosynthetic process
14479	F: CCACTCCCAAGTACTCAAAGG R: CAAGTGTGGCAATCCAACAC	57	0	—	588	1	*Picea abies* mRNA for translation elongation factor-1 alpha, partial (0.0)	translational elongation
14514	F: GGGTTCTGATTCTCCAAAGG R: CTGCATACTTGGCCAAAGTG	56	0	—	322	2	*Metasequoia glyptostroboides* fructose-1,6-diphosphate aldolase mRNA, complete cds (2E-74)	pentose-phosphate shunt, response to salt stress, glycolysis, response to cadmium ion
14585	F: TCTTGAATTCTTCCTATGTCCCAG R: AATTGCACATCTGCACAAACTC	57	1	193	915	1	PREDICTED: *Vitis vinifera* galacturonosyltransferase 8-like (LOC100258818), mRNA (6E-119)	homogalacturonan biosynthetic process
14887	F: GGTTAGACCAGTTCATAACC R: GTCTTCAAACTCTGACAAGG	53	0	—	1156	2	PREDICTED: *Glycine max* elongation factor 2-like (LOC100788357), mRNA (0.0)	
15135	F: TTGCAGGACTTCTTTAATGG R: TCTTCTTGTCAGATGGATCC	53	0	—	657	2	*Ricinus communis* heat shock protein, putative, mRNA (0.0)	oxidation reduction, response to stress, auxin biosynthetic process
15337	F: TTTATTGTATTCCTCCTAGGCCAG R: CACAATCTAAGCCACATTCTTCC	57	1	232	1086	1	*Picea glauca* isolate D8411049-162 cellulose synthase family protein gene, partial sequence (0.0)	cellulose biosynthetic process, cellular cell wall organization
15484	F: TTCGACGCCAACGTTATCTG R: GGCCCAGAGAATTGACATCC	58	0	—	663	2	*Pinus pinaster* phenylalanine ammonia-lyase (pal2) mRNA, complete cds (0.0)	phenylpropanoid metabolic process, biosynthetic process, L-phenylalanine catabolic process
15727	F: CACTGAAGGTTGTGGACGAG R: GTTCAGAAGGGCTGTGTAGG	58	0	—	325	2	*Pinus pinaster* mRNA for cytosolic serine hydroxymethyltransferase (cshmt gene) (2E-138)	L-serine metabolic process, one-carbon metabolic process, glycine metabolic process
15811	F: TTCGAGATCATCTGGACTGC R: CGACTGTTTCGACAGTGAGG	57	0	—	438	2	*Abies alba* genotype Lamacce 1 chalcone synthase (CHS) gene, CHS-A8 allele, complete cds (0)	biosynthetic process
15969	F: GGAAACCTTCTTGTTCACATCTG R: CTTGTCTGGAATCCTCCCTG	57	0	—	990	1	*Pinus contorta S*-adenosylmethionine synthetase (sams2) mRNA, complete cds (0.0)	auxin biosynthetic process, one-carbon metabolic process
16727	F: GGTGACTGTGAAGGCAATGG R: TCCCACATTTCTTTCCAGCT	58	0	—	331	2	*Populus* EST from severe drought-stressed opposite wood (0.000000003)	lipid transport
16816	F: CATCTTGGCTTCGTGATTGTC R: TGCAATTTGGCGTAATCGAC	57	3	132	562	2	*Pseudotsuga menziesii* class III homeodomain-leucine zipper (C3HDZ1) gene, complete cds (0)	
16883	F: CTCACAGAGGTCAGAAAGAATGG R: CTGCTTCAAAGGTTTGACAATCTC	58	0	—	710	1	*Pinus contorta S*-adenosylmethionine synthetase (sams2) mRNA, complete cds (0.0)	auxin biosynthetic process, one-carbon metabolic process
16979	F: CCTGGATAGTGAAATTGGAGG R: ATCCTTCTCTGAATGAGTTTCG	55	0	—	535	1	*Keteleeria davidiana* chloroplast DNA, complete sequence (0)	auxin biosynthetic process
17340	F: CTTGGTTAATTTCCGTCCTG R: CAGCTCCTACATTTAAACCC	54	0	—	281	2	*Abies firma* chloroplast, partial genome (0)	transport, photosynthesis, electron transport chain
17637	F: TGCTGAGAAAGTTGATTCTTCC R: GTATTCGAGGTGTAGATTGCTG	56	0	—	424	1	*Ricinus communis* transferase, transferring glycosyl groups, putative, mRNA (2E-69)	
17975	F: CAAACATTGCTGCAAAGCTC R: CCTATTCCAGCAACCAATATGTC	56	2	94	547	1	*Ricinus communis* cysteine synthase, putative, mRNA (1E-65)	cysteine biosynthetic process from serine
18135	F: GAGACTTTGGATTCGATCCC R: AGAAGGCCGCAAATATAGTG	55	1	132	683	2	*Picea abies* mRNA for putative chlorophyll A-B binding protein, (pPA0001 gene) (0)	photosynthesis, light harvesting in photosystem I
18444	F: ATTAATCTTTGCAGGGAAGC R: AGACGAGATGAAGTGTAGAC	54	0	—	313	2	*P. sylvestris* mRNA for polyubiquitin (3E-116)	
18599	F: GGAATGCATGATCCATTTCTG R: TACCTGAATTGTTCTTGCGA	55	0	—	678	1	*S. tuberosum* mRNA for NADH dehydrogenase, NADH-binding subunit (complex I) (0.0)	oxidation reduction
18680	F: CTGCGATGGATAAACTACCT R: GCTAGTGTTGCTATTGTGGG	55	1	214	465	2	*Picea glauca* isolate D761009-28 myb family protein gene, partial sequence (1E-140)	regulation of transcription

Table 2. Continued.

Gene Locus ID	Primer sequences (5′–3′)	T_a (°C)	No. of introns	Intron length (nt)	Total length (nt)	Annotation approach	BLASTn of gene sequences against nr nucleotide database (*E*-value)	GO-ID biological process
19005	F: GGAGATTGAGCAACGAAGAG R: TTTGAATCCCTGAAATCCTGG	56	0	—	368	1	*Abies firma* chloroplast, partial genome (0)	auxin biosynthetic process
19173	F: AGAACCAATCCCTGTTACAC R: GATCAGTTCCAATCACACCT	55	0	—	343	2	*Ricinus communis* proteasome subunit alpha type, putative, mRNA (2E-84)	defense response to bacterium, ubiquitin-dependent protein catabolic process, response to zinc ion
19540	F: ACCAATTCTCTTGTTCTCGG R: CGAACCATGTAAAGATCATTCC	55	0	—	634	1	*Cedrus deodara* chloroplast DNA, complete sequence (0)	plasma membrane ATP synthesis coupled proton transport, auxin biosynthetic process
20156	F: ATGGATCCCTGGAATTTATGC R: ATACTCTACCTACTACAGAATCCC	55	0	—	386	1	*Picea sitchensis* isolate VD401 magnesium chelatase H-like protein mRNA, partial cds (3E-110)	RNA processing, chlorophyll biosynthetic process
20318	F: ACAGCTCCCATTAATCTGAC R: CCAGAATTGTTCATTTCTCCAC	55	0	—	356	1	PREDICTED: *Glycine max* cellulose synthase-like protein D3-like (LOC100785985), mRNA (6E-69)	root hair elongation, cellulose biosynthetic process, response to cold, cellular cell wall organization, plant-type cell wall biogenesis
20694	F: GTCGAACAATGAAGACGAGG R: TGTGAGCGAAGAAACAAACC	56	0	—	346	1	PREDICTED: *Vitis vinifera* zinc finger CCCH domain-containing protein 49-like (LOC100259323), mRNA (6E-43)	cell wall modification, regulation of transcription
21136	F: AGACTGGTGTTACATTTGCGT R: CCAACAAGCTTCTCACTAATTTCC	57	1	229	535	1	*P. taeda* gene for protochlorophyllide reductase (3E-168)	oxidation reduction, chlorophyll biosynthetic process, photosynthesis
21165	F: ATGCACGATGTTCTTGATGC R: GGTGTCATGTTTATATGACAGTGG	57	2	204	644	1	PREDICTED: *Glycine max* premRNA-processing-splicing factor 8-like (LOC100804026), mRNA (4E-46)	response to hypoxia, sucrose biosynthetic process
21173	F: ACATTGTTGCTAACGATCCG R: AGACGAGGTAGAGATTGAGC	56	0	—	333	2	*Picea sitchensis* isolate VA100 basic endochitinase-like protein mRNA, partial cds (1E-138)	cell wall macromolecule catabolic process, chitin catabolic process
21890	F: GAAAGCTTACAGGGAAGCAG R: ACGATATCCAAGCATCATCC	55	1	358	607	2	*Picea sitchensis* isolate VD401 SWAP domain-containing protein-like protein mRNA, partial cds (2E-116)	RNA processing
21957	F: AACAACTTCACAGTTTCTCC R: GGAATCGGTAAATCAACGAC	54	0	—	292	2	*Abies firma* chloroplast, partial genome (2E-157)	auxin biosynthetic process, chlorophyll biosynthetic process, oxidation reduction, photosynthesis, dark reaction
22174	F: GATGATCCGGTTCGAATACC R: AAACGTAAGATACAAGTGGGTG	55	0	—	334	1	*Abies firma* chloroplast, partial genome (6E-157)	regulation of apoptosis, transcription, DNA-dependent
23660	F: AGGAAGATGTTAGGCTCGGG R: GAAGCCCTTCACAACTCCAG	58	1	781	1232	2	*P. abies* mRNA for porin Mip1 (6E-157)	response to water deprivation, water transport, transmembrane transport
23809	F: ATGCGCTCTATGTTAGAACG R: AATCTCAAGACGTTTACCGA	55	0	—	1058	2	*Abies firma* chloroplast, partial genome (9E-168)	oxidation reduction, chlorophyll biosynthetic process, photosynthesis, dark reaction
23850	F: GAAGATTTATTCGGCAACTG R: ATCTGATCCTCTGTTAAGGT	52	1	449	695	2	*Pinus taeda* mitogen-activated protein kinase 6 (MAPK6) mRNA, complete cds (4E-65)	auxin biosynthetic process, protein amino acid phosphorylation, conjugation, mitosis, cell division
23982	F: TGAGACTTGCTTGGGAAGAG R: AGCCCATTGTAAACGAAGGA	57	2	586	921	1	*Pisum sativum* nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase (gapN) gene, complete cds (4E-33)	metabolic process
24523	F: TTCAGACTCGAACGTTTGCA R: AAGCTTTCATTCCCAGACGG	58	0	—	449	2	*Ginkgo biloba* catalase mRNA, complete cds (3E-98)	hydrogen peroxide catabolic process, oxidation reduction
24699	F: AAGATAAGCAGTTTGCTGCA R: AACATTCTTCTCGCCAACAG	56	0	—	262	2	*Ageratina adenophora* heat shock protein 70.58 mRNA, complete cds (2E-81)	auxin biosynthetic process, response to stress
24902	F: CCCTCTCAATCTTGAGGATGC R: CAGATGGACCTGTAATTTGAACCT	58	1	240	662	1	*Arabidopsis thaliana* ferredoxin–NADP+ reductase (RFNR2) mRNA, complete cds (3E-54)	electron transport chain
25060	F: CTGCAAGATACTTCAAAGATGCAC R: ATTTGGTGTAGAGAACATCTTCCC	58	2	163	624	1	PREDICTED: *Glycine max* ATP-citrate synthase beta chain protein 1-like (LOC100800904), mRNA (2E-74)	acetyl-CoA biosynthetic process, cellular carbohydrate metabolic process
26089	F: GATTATTGATTCTACCACCGGA R: TTTCTCAACAGCCTTGATGAC	55	1	281	1233	1	*Rosa multiflora* elongation factor 1-alpha mRNA, complete cds (0.0)	translational elongation

TABLE 2. Continued.

Gene Locus ID	Primer sequences (5′–3′)	T_a (°C)	No. of introns	Intron length (nt)	Total length (nt)	Annotation approach	BLASTn of gene sequences against nr nucleotide database (*E*-value)	GO-ID biological process
26764	F: GGGAATTGGCTCGTATCTGG R: GTTCTGCTTAGCAATCTTTGTCC	58	0	—	359	1	*Cucumis sativus* 6-phosphogluconate dehydrogenase (6PGDH) mRNA, complete cds (7E-55)	response to glucose stimulus, response to sucrose stimulus, response to fructose stimulus, response to salt stress, pentose-phosphate shunt, oxidation reduction, response to cadmium ion
27033	F: TTTACTCCACCATTACGAGG R: TTCGCAATGATAGGATTGCA	55	0	—	948	2	*Medicago truncatula* heat shock protein (MTR_7g024390) mRNA, complete cds (0)	response to virus, auxin biosynthetic process, protein folding, response to heat, response to bacterium, response to cadmium ion, response to high light intensity, response to hydrogen peroxide, protein amino acid phosphorylation
27963	F: TAGGCCCATAGCTAACAAACC R: TCGAATTGTTTCATCCTCCCA	57	0	—	318	1	*Keteleeria davidiana* chloroplast DNA, complete sequence (0)	transcription, DNA-dependent
28203	F: TGTGGACGAGGAGATATTCG R: TTCAGAAAGGGCTGTGTAGG	56	0	—	315	2	*Pinus pinaster* mRNA for cytosolic serine hydroxymethyltransferase (cshmt gene) (1E-128)	L-serine metabolic process, one-carbon metabolic process, glycine metabolic process
28456	F: GATTTCGAGAGCTGGTATCCC R: AGCTGTCGGTTGATGTTCTG	58	0	—	853	1	*Ricinus communis* oligosaccharyl transferase, putative, mRNA (7E-169)	protein amino acid glycosylation
28639	F: GTAGAATAAGTGGGAGCCGT R: ATAGGAAGAGCCGACATCGA	57	0	—	438	2	*Abies fabri* 26S ribosomal RNA gene, partial sequence (0)	
29437	F: CTTCAGGTGCTCGATATCGT R: TCAACTGGAAACGTTAGCTC	56	0	—	403	2	*Populus trichocarpa* argonaute protein group (AGO911), mRNA (2E-99)	

Note: — = not available; T_a = annealing temperature.

[a] Values are based on the sequence of one sample randomly chosen from a sample set of 80 trees from a population at Mont Ventoux (France).

Chapter II

Anna M. Roschanski, Bruno Fady, Birgit Ziegenhagen and Sascha Liepelt (2015)

Transcriptome sequencing, SNP calling and assayed genotyping in *Abies alba* Mill.

(unpublished manuscript, Jan-2016)

Transcriptome sequencing, SNP calling and assayed genotyping in *Abies alba* MILL.

Authors: Anna M. Roschanski, Bruno Fady, Birgit Ziegenhagen and Sascha Liepelt
Unpublished Manuscript January 2016

Abstract

Transcriptome sequencing is a cost efficient way to access large and complex genomes, such as those of conifers. We sequenced the transcriptomes of twelve individuals of the European silver fir (*Abies alba* MILL) using the Illumina HiSeq2000 platform. Raw reads were assembled using a published *A. alba* transcriptome as the reference wherein stress related candidate genes have previously been annotated. SNP detection between the twelve newly sequenced transcriptomes and the reference transcriptome yielded on average 24 580 SNPs per individual. We extended the sequence annotation and designed a KASP assay for multiplex genotyping of 763 SNPs distributed across 316 candidate genes. The new tool yielded 267 successfully genotyped SNPs lying on 175 candidate genes. It can now be used for fast genotyping of many *A. alba* trees at low costs.

Introduction

Conifers are of outstanding economic and ecological importance. Of the world's forests the share of coniferous forests is approximately 39 % (Armenise *et al.* 2012). Notwithstanding their importance, conifers are, unlike other important plants such as crops, genetically poorly studied. The reasons for this are that they are characterized by long generation times, high levels of heterozygosity and huge genomes that comprise long stretches of repeat rich regions (Torre *et al.* 2014), thus by features opposite to the suitability for genome studies (Uddenberg *et al.* 2015). To date, three draft genomes exist, namely those of *Picea abies* (Nystedt *et al.* 2013), *Picea glauca* (Birol *et al.* 2013) and *Pinus taeda* (Wegrzyn *et al.* 2014; Neale *et al.* 2014). A finding resulting from these projects is that the number and size of coding regions within conifer genomes is not proportionally larger than in angiosperms, but that conifer genome size is due to very long introns (Torre *et al.* 2014), often in the form of transposable elements (Morse *et al.* 2009). Therefore, transcriptome (exome) sequencing is a promising strategy to reduce sequencing costs and sequence complexity (Ekblom & Galindo 2011; Mackay & Dean 2011). It allows for the investigation

of gene expression and function. For instance, comparative transcriptome characterization was applied to study differences in the mRNA composition of different tissues (Canas *et al.* 2014), developmental stages (Elbl *et al.* 2014), treatments (Provost *et al.* 2013) or related species or ecotypes (Liu *et al.* 2014). Further whole exome sequences can be used to develop dense marker sets that can help to elucidate the roles of evolutionary processes when studied simultaneously in population genomic approaches (Luikart *et al.* 2003).

For over a decade, sequencing of expressed sequence tags (ESTs) has produced over 900 000 exome sequences of conifers (Mackay & Dean 2011). With the advent of next-generation sequencing (Shendure & Ji 2008; Metzker 2010) the number of accessible conifer transcriptomes has rapidly increased to at least twelve species (Lorenz *et al.* 2012).

Here, we present twelve novel transcriptomes of the European silver fir (*Abies alba* MILL) using the Illumina HiqhSeq2000 platform. *A. alba* is distributed over a wide range of mountainous ecosystems of temperate and Mediterranean areas of Europe. It is increasingly appreciated in forestry, among other things, because of its high quality timber (Kerr *et al.* 2015) and its positive effect on stand stability (Hansen & Larsen 2004). In the course of the current rapid climate change, the species may become particularly important for forest conservation and management, because Mediterranean population might show adaptations to more xeric conditions (Hampe & Petit 2005). The genome of *A. alba* consists of 16 – 17 Gb (Roth *et al.* 1997; Puizina *et al.* 2008) in 2n = 24 chromosomes. Using Roche 454-sequencing Roschanski *et al.* (2013), presented the first de novo transcriptome of *A alba* and a set of stress related candidate genes. We continue with the development of genomic resources in *A alba*. Six of the newly sequenced individuals originate from distribution-center populations in Germany and six from distribution-periphery populations in France. Transcriptome assembly was based on the 454 transcriptome of Roschanski *et al.* (2013) as the reference. Subsequent detection of single nucleotide polymorphisms (SNPs) was performed between all thirteen transcriptomes. Combing the new SNP information and the annotation of Roschanski *et al.* (2013), a KASP genotyping assay was designed and tested on 400 adult *A. alba* trees from wild populations of southern France.

Materials and Methods

Sites and sampling

Twelve trees were sampled at two sites, one in the Middle Black Forest (District Oberharmersbach) in Germany, the other on Mont Ventoux in the south-western pre-Alps of France (Table 1). Both sites lie within the same post-glacial re-colonization lineage that starts in the Apeninnes and extends north-westwards whereby it crosses the western Alps including its south-westernmost areas (Fady *et al.* 1999; Liepelt *et al.* 2009). At each site, we sampled three to five fully flushed, sun exposed leaves of six sexually mature individuals. Three trees at each site were sample at high and three at low elevation. Since flushing occurred earlier at Mt Ventoux current needles were sample here, while the leave generation of the previous year was sample in the Black Forest. After dispatching from the tree, leaves were immediately frozen in liquid nitrogen.

Table 1: Sampling material, location and date. F: France, G: Germany. _1/_2 indicates different elevation at each location

Individual	Leaf Generation	Location	Sampling Date
F_1.1	2011	France, Mt Ventoux (44°10'39"N; 5°14'32"E; 1350 m a.s.l.)	18-May-11
F_1.2	2011		18-May-11
F_1.3	2011		18-May-11
F_2.1	2011	France, Mt Ventoux (44°11'02"N; 5°14'16"E; 1100 m a.s.l.)	18-May-11
F_2.2	2011		18-May-11
F_2.3	2011		18-May-11
G_1.1	2010	Germany, Black Forest (48°22'40"N; 8°06'12"E; 540 m. a.s.l.)	8-Jun-11
G_1.2	2010		8-Jun-11
G_1.3	2010		8-Jun-11
G_2.1	2010	Germany, Black Forest (48°24'39"N; 8°10'12"E; 590 m. a.s.l.)	8-Jun-11
G_2.2	2010		8-Jun-11
G_2.3	2010		8-Jun-11

Sequencing and SNP calling

Twelve cDNA libraries were constructed from the leaf material, individually barcoded and sequenced on an Illumina HiSeq2000 platform at IGA Technology Services (Udine, Italy). Raw reads were assembled using the annotated transcriptome of Roschanski *et al.* (2013) as the reference transcriptome. SNP detection was subsequently performed between all thirteen transcriptomes using an in-house pipeline at IGA.

Annotation and design of KASP assay

Candidate genes annotated by the specific and the standard approach of Roschanski et al. 2013 were used (Figure 1). Additionally, 4676 genes assigned to gene ontology (GO) terms (Ashburner *et al.* 2000) were screened for drought stress related GOs using a custom perl script (available upon request, see: data availability section of the thesis). This yielded 764 candidate genes of which 119 passed the filter of 90 % alignment identity to the annotated database entry and a minimum length of 100 amino acids. In total, 460 genes were identified as being stress related. Amongst them, 331 bore 3391 SNPs as detected between the thirteen transcriptomes. Their flanking regions were checked for primer designability using an in-house pipeline of LGC Genomics (LGC Genomics, Middlesex, United Kingdom). A number of 1379 SNPs of 316 candidate genes passed the quality check and thus, were suitable for assayed genotyping. Form them, 763 SNPs were manually selected applying the criteria of balanced distribution across candidate genes and highest designability scores. Genotyping was done using a KASP assay (LGC Genomics, Middlesex, United Kingdom).

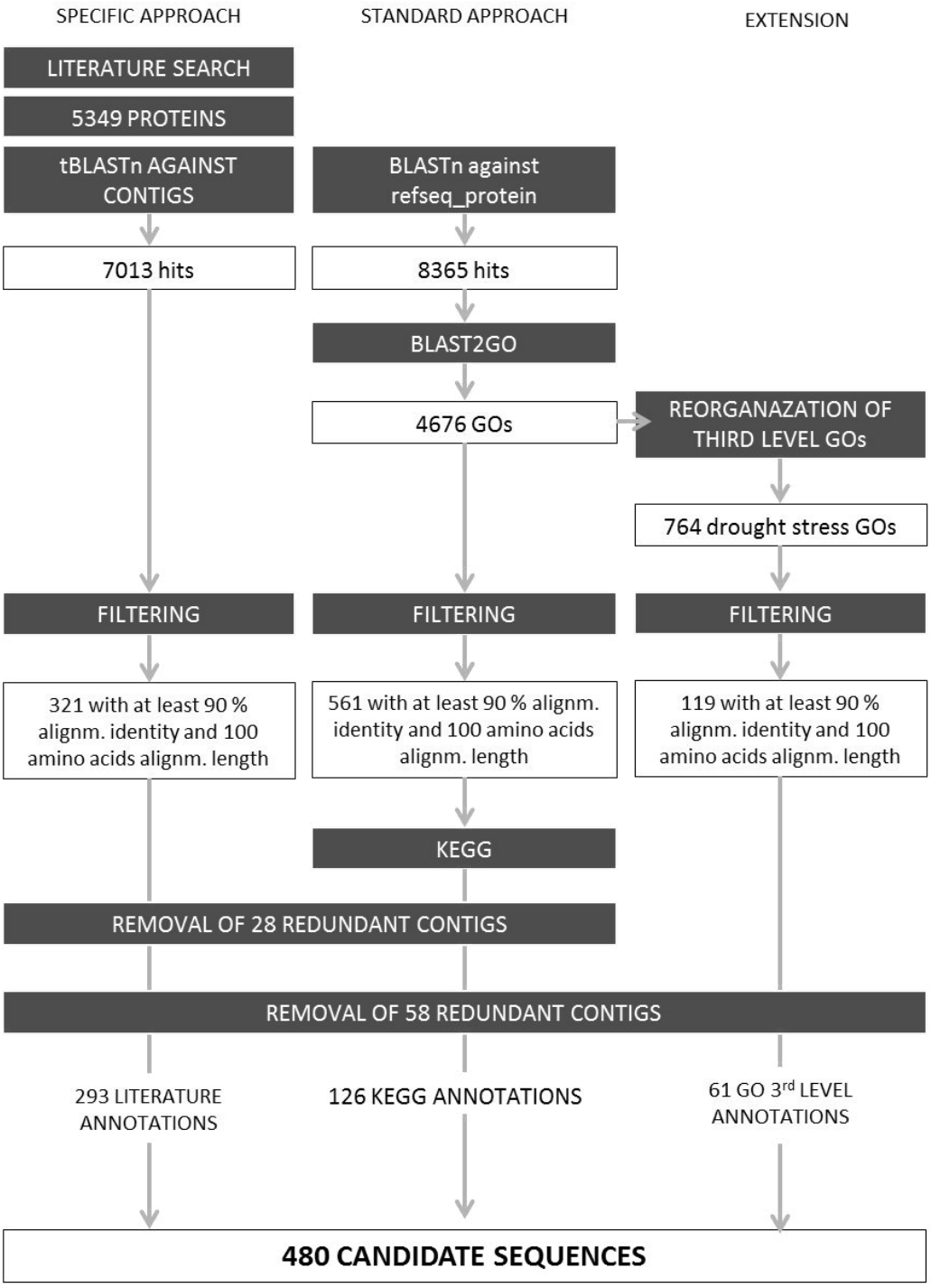

Figure 1: Workflow of the annotation protocol of Roschanski et al (2013). Numbers of the output after each step are given. The standard approach starts with 25 149 contigs. The specific approach uses them as the reference database for the tBLASTn step. Extension of the published annotation protocol is shown in the third column. tBLASTn: searches a translated nucleotide databases using a protein query; BLASTx: searches a protein databases using a translated nucleotide query; BLAST2GO searches the gene ontology database where nucleotide sequences are assigned to functional annotations using a nucleotide query; alignm.: alignment; KEGG: Kyoto Encyclopedia of Genes and Genomes

Results

Sequencing

The sum of sequence output ranges from 9.13E+06 bases for F_1.1 to 1.001E+07 bases for F_1.2 (mean: 9.78E+06; sd: 2.48E+05) (Table 2). The mean number of contigs per individual was 20 935 (sd: 468). N50 contig size, defined as the contig site whereto 50 % of all bases were assembled, lay between 437 bases for G_1.1 and G_1.3 and 441 bases for F_2.3 and F1_1.2. Mean N50 contig size was 466.9 bases (sd: 2.39). Between 78 % (F_1.1) and 85 % (G_1.2) of contigs obtained per individual could be mapped against the contig sequences of the 454-reference assembly. The total number of polymorphic contigs was 4268. Therein, between 19 089 SNPs for F_1.3 and 28 622 SNPs for F_1.2 were detected. On average 24 580 SNPs per individual were found.

KASP genotyping assay

Of 763 SNPs in total, 406 (53 %) were successfully genotyped and of these 267 were polymorphic. Thus the final conversion rate was 35 % with respect to the number of SNPs. However, from the 316 genes submitted to genotyping 175 bore analyzable results, *i.e*. conversion in term of genes was 55 %. The analysis of the genotyping SNP of 400 trees is presented in a further study (Roschanski and Csilléry *et al.* 2015).

Table 2: Output of sequencing, assembly and SNP detection of twelve individual *A. alba* transcriptomes and of the reference transcriptome published in Roschanski *et al.* (2013).

	Reference 454	G_1.1	G_1.2	G_1.3	G_2.1	G_2.2	G_2.2	F_2.1	F_2.2	F_2.3	F_1.1	F_1.2	F_1.3
contigs	25149	21094	20989	21149	20716	21407	20378	20790	21222	21352	19796	21364	20972
nr.missed contigs	-	4055	4160	4000	4433	3742	4771	4359	3927	3797	5353	3785	4177
% mapped contigs	-	83.88	83.46	84.09	82.37	85.12	81.03	82.67	84.39	84.90	78.71	84.95	83.39
Total sequence length (bases)	1,25E+07	9,89E+06	9,82E+06	9,92E+06	9,63E+06	9,98E+06	9,48E+06	1,000E+07	9,92E+06	9,71E+06	9,81E+06	1,001E+07	9,13E+06
Max contig size	2394	2373	2375	2373	2373	2375	2368	2372	2375	2375	2372	2376	2373
Contig size (N50) [a]	704	440	440	440	437	438	437	439	440	441	433	441	440
SNPs in 4268 contigs		24673	23520	25367	22516	25078	21973	27259	26584	24398	25885	28622	19089
mean nr.SNPs / contig		5.78	5.51	5.94	5.28	5.88	5.15	6.39	6.23	5.72	6.06	6.71	4.47

[a] Half of all bases are assembled in contigs of this size or longer

Discussion

Here, we present the sequencing of twelve transcriptomes of *A. alba* and the subsequent transcriptome wide SNP detection. Combining the results with the annotation work of Roschanski *et al.* (2013) we designed genotyping assay, a high value tool for further population genetic analyzes in the non-model conifer.

The conversion rate of the genotyping assay was 35 % and thus lower than for other conifers where it was 60 % or higher (Pavy *et al.* 2008; Eckert *et al.* 2009; Budde *et al.* 2014). However, in contrast to this work the cited studies developed genotyping assays at the basis of EST libraries or combinations of EST and next-generation sequencing data and thus presumably based on more extensive transcriptomic resources for primer design. Indeed, Karam *et al.* (2015) who exclusively derived SNPs from de novo sequencing report a lower conversion rate of 50 % and Cullingham *et al.* (2013) who exclusively performed in silico development of SNP loci merely achieved a conversion rate of 30 %. The most probable reason for the poor the success of our assay is that alignment errors have led to SNP calling errors. Particularly indicative for this is the high rate of monomorphic loci in the assay. A more thorough quality assessment of the assembly may have led to a higher number of analyzable SNPs. Nevertheless the KASP assay presented here, is the first multiplex genotyping tool in *A. alba* and opens the door for research on selective markers. Further, the thirteen transcriptomes provide the basis for exome wide analyses in the species.

References

Armenise L, Simeone MC, Piredda R, Schirone B (2012) Validation of DNA barcoding as an efficient tool for taxon identification and detection of species diversity in Italian conifers. *European Journal of Forest Research*, **131**, 1337–1353.

Ashburner M, Ball CA, Blake JA *et al.* (2000) Gene Ontology: tool for the unification of biology. *Nature genetics*, **25**, 25–29.

Birol I, Raymond A, Jackman SD *et al.* (2013) Assembling the 20 Gb white spruce (Picea glauca) genome from whole-genome shotgun sequencing data. *Bioinformatics*, **29**, 1492–1497.

Budde KB, Heuertz M, Hernández-Serrano A *et al.* (2014) In situ genetic association for serotiny, a fire-related trait, in Mediterranean maritime pine (Pinus pinaster). *New Phytologist*, **201**, 230–241.

Canas RA, Canales J, Gomez-Maldonado J, Avila C, Canovas FM (2014) Transcriptome analysis in maritime pine using laser capture microdissection and 454 pyrosequencing. *Tree Physiology*, **34**, 1278–1288.

Cullingham CI, Cooke JEK, Coltman DW (2013) Effects of introgression on the genetic population structure of two ecologically and economically important conifer species: lodgepole pine (Pinus contorta var. latifolia) and jack pine (Pinus banksiana). *Genome*, **56**, 577–585.

Eckert AJ, Wegrzyn JL, Pande B *et al.* (2009) Multilocus Patterns of Nucleotide Diversity and Divergence Reveal Positive Selection at Candidate Genes Related to Cold Hardiness in Coastal Douglas Fir (Pseudotsuga menziesii var. menziesii). *Genetics*, **183**, 289–298.

Ekblom R, Galindo J (2011) Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, **107**, 1–15.

Elbl P, Lira BS, Andrade SCS *et al.* (2014) Comparative transcriptome analysis of early somatic embryo formation and seed development in Brazilian pine, Araucaria angustifolia (Bertol.) Kuntze. *Plant Cell, Tissue and Organ Culture (PCTOC)*, **120**, 903–915.

Fady B, Forest I, Hochu I *et al.* (1999) Genetic differentiation in Abies alba Mill. populations from Southeastern France. *Forest Genetics (Slovak Republic).*

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, **8**, 461–467.

Hansen JK, Larsen JB (2004) European silver fir (Abies alba Mill.) provenances from Calabria, southern Italy: 15-year results from Danish provenance field trials. *European Journal of Forest Research*, **123**, 127–138.

Mackay JJ, Dean JFD (2011) Transcriptomics. In: *Genetics, Genomics and Breeding of Conifers*. Science Publishers.

Karam M-J, Lefèvre F, Dagher-Kharrat MB, Pinosio S, Vendramin GG (2015) Genomic exploration and molecular marker development in a large and complex conifer genome using RADseq and mRNAseq. *Molecular Ecology Resources*, **15**, 601–612.

Kerr G, Stokes V, Peace A, Jinks R (2015) Effects of provenance on the survival, growth and stem form of European silver fir (Abies alba Mill.) in Britain. *European Journal of Forest Research*, **134**, 349–363.

Liepelt S, Cheddadi R, de Beaulieu JL *et al.* (2009) Postglacial range expansion and its genetic imprints in Abies alba (Mill.)—A synthesis from palaeobotanic and genetic data. *Review of Palaeobotany and palynology*, **153**, 139–149.

Liu J-J, Sniezko RA, Sturrock RN, Chen H (2014) Western white pine SNP discovery and high-throughput genotyping for breeding and conservation applications. *BMC Plant Biology*, **14**, 380.

Lorenz WW, Ayyampalayam S, Bordeaux JM *et al.* (2012) Conifer DBMagic: a database housing multiple de novo transcriptome assemblies for 12 diverse conifer species. *Tree Genetics & Genomes*, **8**, 1477–1485.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981–994.

Metzker ML (2010) Sequencing technologies - the next generation. *Nature Reviews. Genetics*, **11**, 31–46.

Morse AM, Peterson DG, Islam-Faridi MN *et al.* (2009) Evolution of Genome Size and Complexity in Pinus. *PLoS ONE*, **4**, e4332 doi: 10.1371

Neale DB, Wegrzyn JL, Stevens KA *et al.* (2014) Decoding the massive genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biology*, **15**, R59.

Nystedt B, Street NR, Wetterbom A *et al.* (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature*, **497**, 579–584.

Pavy N, Pelgas B, Beauseigle S *et al.* (2008) Enhancing genetic mapping of complex genomes through the design of highly-multiplexed SNP arrays: application to the large and unsequenced genomes of white spruce and black spruce. *BMC genomics*, **9**, 21.

Provost GL, Domergue F, Lalanne C *et al.* (2013) Soil water stress affects both cuticular wax content and cuticle-related gene expression in young saplings of maritime pine (Pinus pinaster Ait). *BMC Plant Biology*, **13**, 95.

Puizina J, Sviben T, Krajačić-Sokol I *et al.* (2008) Cytogenetic and molecular characterization of the Abies alba genome and its relationship with other members of the Pinaceae. *Plant Biology*, **10**, 256–267.

Roschanski AM, Csilléry K, Liepelt S *et al.* (2015) Evidence of divergent selection for drought and cold tolerance at landscape and local scales in Abies alba Mill. in the French Mediterranean Alps. *Molecular Ecology*, doi: 10.1111

Roschanski AM, Fady B, Ziegenhagen B, Liepelt S (2013) Annotation and Re-Sequencing of Genes from De Novo Transcriptome Assembly of Abies alba (Pinaceae). *Applications in Plant Sciences*, **1**, 1–8.

Roth R, Ebert I, Schmidt J (1997) Trisomy associated with loss of maturation capacity in a long-term embryogenic culture of Abies alba. *Theoretical and Applied Genetics*, **95**, 353–358.

Shendure J, Ji H (2008) Next-generation DNA sequencing. *Nat Biotech*, **26**, 1135–1145.

Torre A, Birol I, Bousquet J *et al.* (2014) Insights into Conifer Giga-Genomes. *Plant Physiology*, **166**, 1724–1732.

Uddenberg D, Akhter S, Ramachandran P, Sundström JF, Carlsbecker A (2015) Sequenced genomes and rapidly emerging technologies pave the way for conifer evolutionary developmental biology. *Frontiers in Plant Science*, **6,** 970 doi: 10.3389

Wegrzyn JL, Liechty JD, Stevens KA *et al.* (2014) Unique Features of the Loblolly Pine (Pinus taeda L.) Megagenome Revealed Through Sequence Annotation. *Genetics*, **196**, 891–909.

Chapter III

Anna M. Roschanski, Katalin Csilléry, Sascha Liepelt, Sylvie Oddou-Muratorio, Birgit Ziegenhagen, Frédéric Huard, Kristian K. Ullrich, Dragos Postolache, Giovanni G. Vendramin and Bruno Fady (2016) **Evidence of divergent selection for drought and cold tolerance at landscape and local scales in *Abies alba* Mill. in the French Mediterranean Alps**. *Molecular Ecology*, 25.3, 776-794

MOLECULAR ECOLOGY

Molecular Ecology (2016) **25**, 776–794 doi: 10.1111/mec.13516

Evidence of divergent selection for drought and cold tolerance at landscape and local scales in *Abies alba* Mill. in the French Mediterranean Alps

ANNA M. ROSCHANSKI,*[1,3] KATALIN CSILLÉRY,†[2,3] SASCHA LIEPELT,* SYLVIE ODDOU-MURATORIO,† BIRGIT ZIEGENHAGEN,* FRÉDÉRIC HUARD,‡ KRISTIAN K. ULLRICH,§ DRAGOS POSTOLACHE,¶**†† GIOVANNI G. VENDRAMIN** and BRUNO FADY†

*Conservation Biology, Faculty of Biology, University of Marburg, Karl-von-Frisch-Straße, Marburg 35032, Germany, †INRA, UR629, Ecologie des Forêts Méditerranéennes (URFM), Avignon Cedex 9 84914, France, ‡INRA, AGROCLIM, Avignon Cedex 9, 84914 France, §Cell Biology, Faculty of Biology, University of Marburg, Karl-von-Frisch-Straße, Marburg 35032, Germany, ¶Scuola Superiore Sant'Anna, Piazza Martiri della Libertà 33, Pisa 56127, Italy, **Institute of Biosciences and Bioresources, National Research Council (CNR), Via Madonna del Piano 10, Sesto Fiorentino (Firenze) 50019, Italy, ††National Institute of Forest Research and Development (INCDS), Simeria Research Station, Str. Biscaria 1, Simeria 335900, Romania*

Abstract

Understanding local adaptation in forest trees is currently a key research and societal priority. Geographically and ecologically marginal populations provide ideal case studies, because environmental stress along with reduced gene flow can facilitate the establishment of locally adapted populations. We sampled European silver fir (*Abies alba* Mill.) trees in the French Mediterranean Alps, along the margin of its distribution range, from pairs of high- and low-elevation plots on four different mountains situated along a 170-km east–west transect. The analysis of 267 SNP loci from 175 candidate genes suggested a neutral pattern of east–west isolation by distance among mountain sites. F_{ST} outlier tests revealed 16 SNPs that showed patterns of divergent selection. Plot climate was characterized using both in situ measurements and gridded data that revealed marked differences between and within mountains with different trends depending on the season. Association between allelic frequencies and bioclimatic variables revealed eight genes that contained candidate SNPs, of which two were also detected using F_{ST} outlier methods. All SNPs were associated with winter drought, and one of them showed strong evidence of selection with respect to elevation. Q_{ST}–F_{ST} tests for fitness-related traits measured in a common garden suggested adaptive divergence for the date of bud flush and for growth rate. Overall, our results suggest a complex adaptive picture for *A. alba* in the southern French Alps where, during the east-to-west Holocene recolonization, locally advantageous genetic variants established at both the landscape and local scales.

Keywords: candidate gene, European silver fir, isolation by environment, landscape genomics, microclimate, selection

Received 22 April 2015; revision received 20 November 2015; accepted 9 December 2015

Correspondence: Sascha Liepelt, Fax: +49-6421-2826588; E-mail: liepelt@biologie.uni-marburg.de

[1]Present address: Department Genebank, Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Satellite Collections North, Inselstr. 9, Malchow (Insel Poel), 23999, Germany

[2]Present address: ETH Zürich, Adaptation to a Changing Environment, Universitaetstrasse 16, Zurich, 8092, Switzerland

[3]These two authors contributed equally.

Introduction

The capacity of organisms to adapt to changing environments, for example during range expansions or environmental change, is crucial for their persistence. Long-lived, sessile organisms, such as forest trees, have to rely mainly on standing genetic variation to produce

locally adaptive phenotypes (Petit & Hampe 2006; Aitken *et al.* 2008). Although tree populations often show substantial levels of genetic differentiation in adaptive traits, suggesting a high adaptive capacity (e.g. Savolainen *et al.* 2007; Alberto *et al.* 2013), ongoing and predicted climate change may exceed their evolutionary potential. Indeed, the velocity of environmental change must not be too fast and population size and standing genetic variation must not be too low for rapid genetic adaptation to eventually rescue a local population (Carlson *et al.* 2014). Tree populations may also be capable of fast migration; nevertheless, estimates of possible migration rates are much debated (e.g. Feurdean *et al.* 2013; Tzedakis *et al.* 2013), and it is widely agreed that ongoing and predicted climate change imposes migration rates that most tree populations cannot keep up with, especially due to their low recruitment and seedling survival (Nathan *et al.* 2011; Zhu *et al.* 2012). Overall, the severe, mainly, drought-induced forest tree die-off events that have already been documented around the world (e.g. Allen *et al.* 2010; Martínez-Vilalta *et al.* 2011) suggest that many tree populations are facing an environmental change within one or two generations for which their phenotypic plasticity and adaptive potential is insufficient to cope with. Because trees provide several key ecological, economical and social ecosystem services, the assessment of the adaptive potential of forest trees is one of the most pressing research goals of evolutionary ecology.

Tree populations at the margins of the distribution range can be particularly useful case studies for understanding the processes by which populations adapt or not to novel environmental conditions (e.g. Hampe & Petit 2005; Restoux *et al.* 2008; Hampe & Jump 2011). Adaptation in marginal habitats plays a major role in the evolution of species ranges, a subject not entirely free from contradictions (e.g. Kawecki 2008; Sexton *et al.* 2009). Species range edges are characterized by small population sizes, increased isolation, thus reduced gene flow and genetic variability (Hampe & Petit 2005). Theory often predicts that adaptation is prevented under such conditions, particularly by maladaptive gene flow from central populations (e.g. Kirkpatrick & Barton 1997). Many technical constrains limit empirical studies for comparing the adaptive capacity of central versus marginal populations, such as differing population size and the lack of replicates; nevertheless, empirical evidence often suggests a lack of decreased abundance or fitness at range edges (Sexton *et al.* 2009). This apparent contradiction can, at least partly, arise from the fact that many marginal populations are relicts of populations that survived during the cold stages of the Quaternary, and despite their reduced genetic diversity at neutral marker loci, they contain key genetic variants in ecologically relevant traits (Crawford 2008; Kawecki 2008; Hampe & Jump 2011). This may be particularly true for populations at the southern margins that can potentially be preadapted to climatic conditions expected in future in more northerly central populations. For example, in the case of *Fagus sylvatica*, Thiel *et al.* (2014) demonstrated that southern populations (Spain) were more resistant to drought than more northerly central populations (Germany) in a common garden setting. Thus, local adaptations at the southern rear edge limit of tree populations can be highly relevant for implementing conservation strategies under global change or habitat fragmentation, for example, by providing sources for assisted migration (e.g. Iverson *et al.* 2013).

Here, we study the patterns of divergent selection in the European silver fir (*Abies alba* Mill.) across four mountain ranges in the French Mediterranean Alps, at the southwestern edge of the species distribution range. Studies based on palynological, palaeobotanical, isoenzyme and organelle DNA marker data suggest that silver fir in Europe recolonized its current habitats after the last glacial maximum from five refugia, and populations from the French Mediterranean Alps most likely originate from a refugium in Central Italy (Liepelt *et al.* 2009). Since recolonization, both land use and climate change have contributed to the present fragmented distribution of silver fir in the French Mediterranean Alps. As a result, current Mediterranean silver fir populations often occupy isolated mountains, where climatic conditions may rapidly change from hot, thermo-Mediterranean to colder and moister oro-Mediterranean (Ozenda 1975; Sagnard *et al.* 2002), that is as passing from the bottom to the top of a mountain slope. Such contrasted climatic conditions are likely to impose a strong selective pressure on Mediterranean silver fir populations. Accordingly, severe die-off events have already been demonstrated on Mediterranean *A. alba* populations, especially at low elevations (e.g. Aussenac 2002; Linares & Camarero 2012; Cailleret *et al.* 2014). Climate change is expected to further contribute to the selective pressure on silver fir in the Mediterranean, where climate models concordantly predict a pronounced decrease in precipitation and an increase in temperature, principally during the warm season (Giorgi & Lionello 2008). In conclusion, we argue that peripheral Mediterranean silver fir populations provide an excellent system to study adaptation to drought (especially in low-altitude populations) and to frost (e.g. in high-altitude or continental populations), as such abiotic stress conditions are likely to become more common in large parts of *A. alba*'s range during the 21st century.

Studying adaptation to climate change in natural populations can be extremely challenging due to the

complexity of the genetic, physiological and environmental processes involved (Sork *et al.* 2013). The 'genomic revolution' in forest tree species holds great promise for detecting locally adapted tree populations (Neale & Kremer 2011) and to better understand the genetic architecture of key adaptive traits (Ekblom & Galindo 2011). In recent years, a wealth of new statistical methods aimed at detecting local adaptations have been developed (e.g. Vitti *et al.* 2013). Nevertheless, challenges remain, the most important one being disentangling demographic and adaptive processes (Tiffin & Ross-Ibarra 2014). Today, it is widely accepted that genetic data analysis alone, such as an F_{ST} genome scan (Luikart *et al.* 2003; Foll & Gaggiotti 2008; Bonhomme *et al.* 2010), is not sufficient to detect the signature of selection and is best applied in combination with environmental and/or phenotypic data (Le Corre & Kremer 2012). One of the promising approaches for combining data types is landscape genomics, where the objective is to identify the environmental and/or ecological drivers of local adaptation by correlating allele frequency variation with abiotic or biotic factors at the landscape scale (e.g. reviewed in Schoville *et al.* 2012; Sork *et al.* 2013). Phenotypic data can also be used and, in forest trees, may come from on-site observations, provenance trials, common garden experiments or reciprocal transplantations (reviewed in Savolainen *et al.* 2007). Traits measured in common gardens may be analysed using classic quantitative genetics approaches to provide complementary evidence for genome scan methods (e.g. De Kort *et al.* 2014; methods discussed in Lepais & Bacles 2014). Another way to exploit phenotypic data from common gardens is using so-called Q_{ST}–F_{ST} tests, where quantitative trait differentiation (Q_{ST}) is compared with genetic divergence inferred from neutral markers (F_{ST}) (reviewed in Whitlock 2008 and Leinonen *et al.* 2013).

Here, we used an integrative approach to study the adaptive response of silver fir across a range of environmental conditions at both the local and the landscape scale. We used genotypic and environmental data to detect signatures of divergent selection that may suggest local adaptation by using both classic F_{ST} outlier methods and landscape genomics approaches. For this purpose, we employed a nested sampling design with pairs of study plots on four mountain sites situated along an east–west transect in the Mediterranean French Alps spanning widely contrasting ecological conditions. Further, using published phenotypic data from corresponding populations studied in a common garden experiment (Sagnard *et al.* 2002), we were also able to test for adaptive divergence between populations.

Materials and methods

Species and study sites

Silver fir has a wide and discontinuous distribution following the mountain ranges of southern and central Europe. We studied silver fir on four mountains (subsequently sites; namely Ventoux (Vtx), Lure (Lur), Issole (Iss) and Vesubie (Ves); see Fig. 1A) situated in the French Mediterranean Alps, along a 170-km east–west transect stretching from 5°14′31.7895″ E to 7°21″37.1173″ E. Sites belong to different phytoecological groups, Vtx and Lur to the intermediate Alps, Iss to the sub-Mediterranean Alps and Ves to the Ligurian Alps (Sagnard *et al.* 2002). The four studied mountains have contrasted climate (see details below). Bedrock composition is sandstone schist in Ves, calcareous clay in Iss and Lur and limestone in Vtx. The studied populations are autochthonous stands, naturally regenerated using a classical uneven-aged management without planting. There is no reason to believe that the studied forests have been planted in the past because *A. alba* has rarely, if ever, been planted in the French Mediterranean and available genetic data do not demonstrate the introduction of exotic material (Fady *et al.* 1999).

At each site, we collected leaf samples from low (L) and high (H) elevation. Subsequently, we will refer to these sampling points as plots and to the trees therein as individuals within populations. The average horizontal distance between high and low plots was ~1500 m, and their average distance in elevation was 365 m. Between 39 and 55 trees were sampled per plot (44/39 at Vtx_L/Vtx_H, 55/55 at Lur_L/Lur_H, 49/47 at Iss_L/Iss_H, 43/44 at Ves_L/Ves_H). All sampled trees were dominant and sexually mature. The average distance between sampled trees within a plot was 37 m, which is sufficiently large to avoid competition for light between trees. Sampled trees were georeferenced using two methods: (1) using a Trimble Pro XR and PathPower GPS when satellite signals were adequate (at least seven satellites in range) or, alternatively (2), when signals were less than adequate, using a Vertex laser dendrometer (which measures distance and angle) and a compass using method (1) GPS trees as reference points. Measurements carried out using (2) were later recalculated as GPS coordinates.

Candidate gene data

Discovery of 316 candidate genes was conducted using the protocol of Roschanski *et al.* (2013) that we extended by assigning additional transcriptome sequences to drought-related gene ontologies using BLAST2GO, version 2.8 (Conesa *et al.* 2005). Twelve additional transcrip-

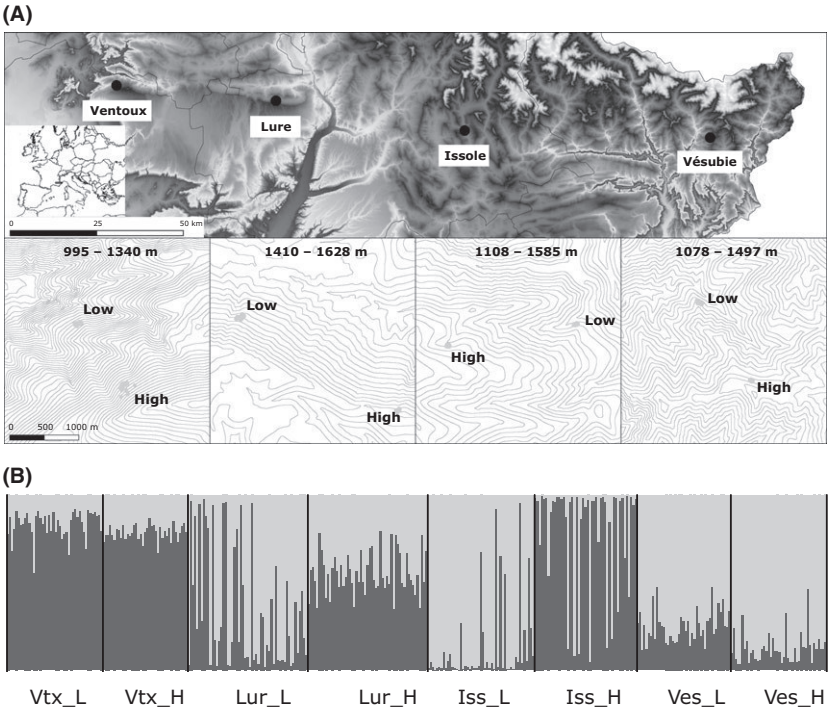

Fig. 1 A: Sampling locations. Upper part: overview map of sampled mountains located along an east-west transect in the French Mediterranean Alps. Darker colours indicate higher elevation. Location of sampling sites within Europe is indicated by the black bar in the small map. Lower part, detailed map of sampling plots at high and low elevations on each mountain. B: population membership coefficient of all sampled trees in a geographic order that corresponds to part A obtained with STRUCTURE for $K = 2$ based on 267 SNP loci.

tomes originating from Ventoux and the Black Forest (District Oberharmersbach, Germany) sequenced on an Illumina HiSeq2000 at IGA Technology Services (Udine, Italy) were mapped against the annotated transcriptome of Roschanski *et al.* (2013) (accession no.: JV134525–JV157085 TSA NCBI) as reference. Based on these mappings, a SNP detection was subsequently performed for the 12 sequenced transcriptomes and the candidate genes using an in-house pipeline at IGA Technology Services (Udine, Italy). Note that the 12 transcriptomes were only used for SNP calling and that no transcriptome assembly was performed. We do not expect any significant ascertainment bias between our discovery panel and our genotyped samples because they originate from the same evolutionary lineage (Liepelt *et al.* 2009).

A set of KASP assays manufactured by LGC Genomics (Middlesex, United Kingdom) was used for genotyping, which yielded 406 SNPs, of which 267 were polymorphic. They were distributed across 175 candidate genes, 104 genes carrying a single SNP, 50 carrying two and 21 three SNPs. Linkage disequilibrium (LD) between all pairs of loci was tested based on allele frequencies (D) using the R package GENETICS (R Core Team 2013; Warnes 2013). Based on STRUCTURE coefficients of population membership, two populations were defined and LD was determined for each population. Candidate genes were functionally annotated using BLAST2GO, version 2.8 (Conesa *et al.* 2005).

We determined the synonymous or nonsynonymous state of SNPs using a strategy modified from TransPipe (Barker *et al.* 2010). Briefly, the contigs containing the selected candidate genes were compared using BLASTX (Altschul *et al.* 1990) against annotated *Picea abies* (Nystedt *et al.* 2013) proteins. Subsequently using the best BLAST hit based on the top e-value, we generated a protein–nucleotide alignment between *Picea abies* (protein) and the *Abies alba* contig (nucleotide) using the exonerate software (Slater & Birney 2005). The resulting protein–nucleotide alignment was then used to extract the matching coding sequences for the *Abies alba* candidate genes followed by a translation into protein sequences. The resulting protein sequences were screened for protein domains using INTERPROSCAN, version 5 (Jones *et al.* 2014). The state of SNPs occurring on coding sequences was classified into synonymous or nonsynonymous by custom R scripts (R Development Core Team 2013).

We used PHASE, version 2.1 (Stephens & Donnelly 2003), to impute the missing genotypes on each gene fragment that contained at least two SNPs, which reduced the original missing data rate of 6.61% to 4.31%. We used model MR0 with varying recombination rate (Li & Stephens 2003) and probability thresholds of 0.95 both for missing alleles and for missing phase information. We ran five independent Markov chains of length 10^4, with a thinning interval of 10 and a burn-in period of 10^4.

Environmental data

We combined both on-site observations and published climatic data to obtain long-term local microclimatic

time series for the reference period of 1960–1990 for each plot. Several climatic and bioclimatic variables were tested as candidates to best describe the selective environment between plots. We found that drought, that is the self-calibrating Palmer Drought Severity Index (scPDSI) following Palmer (1965) and Wells *et al.* (2004), in January, February and March (scPDSI_jfm), relative humidity (RH), vapour pressure deficit in March, April and May (VPD_mam) and vapour pressure deficit in September, October, November (VPD_son) were the best and also sufficient variables to describe the climatic differences between plots (Fig. 2A–D; details on variable selection in Appendix S1, Supporting information).

In addition, to help describe the microclimate of the eight plots, and to derive even more concise environmental variables, we performed a principal component analysis (PCA) of these four variables (Fig. 2E). The first two PC axes explained most of the variability (85.74%). PC1 principally separated low- and high-elevation plots (Fig. 2E), where low-elevation plots were characterized by higher temperatures, that is longer growing seasons (Fig. S1, Supporting information), and generally lower RH (Fig. 2B), that is more drought during the growing season (Fig. 2C,D). In contrast, PC2 ordered sites. At one extreme, Lur plots were the coldest and most humid (Fig. S1, Supporting information), with most of the precipitation falling in winter and limited drought (Fig. 2C,D). In the middle, Vtx_L, Iss_L, Ves_L were the driest sites that suffered from the most severe drought in the vegetative period (Fig. 2C,D). Finally, at the other extreme, Ves_H/Iss_H sites had little winter precipitation and a fair amount of summer precipitation; thus, they were drought stressed only in winter (Fig. 2A).

Population structure

We used the Bayesian clustering algorithm implemented in STRUCTURE, version 2.3.4 (Pritchard *et al.*

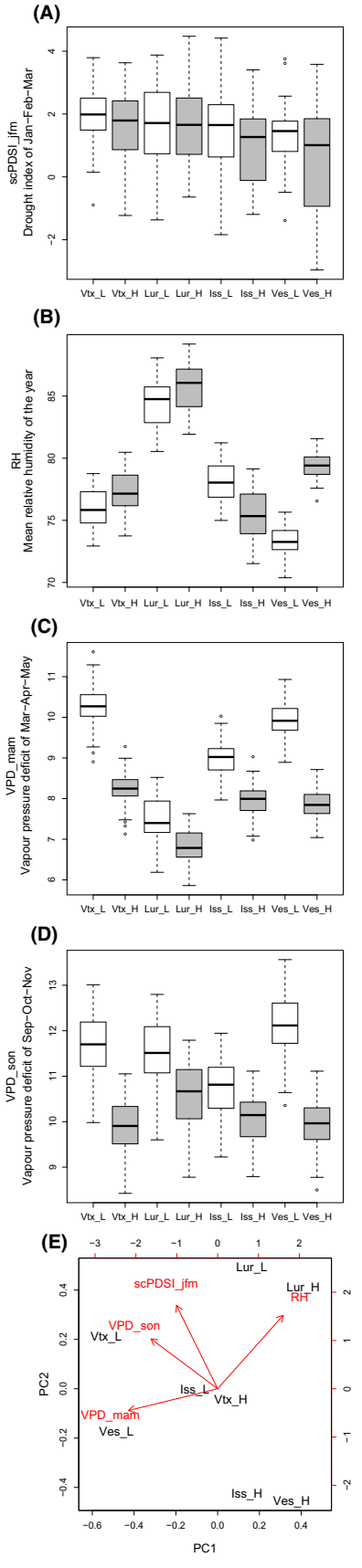

Fig. 2 (A–D) Interannual variation in four quasi-independent bioclimatic variables across the eight sampling plots for the period of 1960–1990. White boxes indicate low-elevation sampling plots (i.e. Vtx_L, Lur_L, Iss_L and Ves_L stand for Ventoux, Lure, Issole and Vesubie low plots, respectively), while grey ones indicate high-elevation sampling plots (also indicated by _H in the labels). scPDSI_jfm is the mean self-calibrating Palmer Drought Severity Index of January, February and March. Relative humidity (RH) is the mean annual RH in percentages. VPD_mam is the vapour pressure deficit in March, April and May, and VPD_son is the vapour pressure deficit in September, October and November. E: PC1 versus PC2 of a principal component analysis of the four bioclimatic variables.

2000), to cluster individuals. We used the admixture model with correlated allele frequencies and included sampling location information to improve clustering performance (Hubisz *et al.* 2009). *K* values were tested from 1 to 9 assuming that it is unrealistic to find more clusters than plots, but still considering the extreme, but unlikely, case (i.e. $K = 9$). A burn-in period of 10^5 iterations was followed by 5×10^5 iterations for the estimation of the membership coefficients. Ten independent Markov chains were run for each *K*. Admixture coefficients for $K = 2$ were averaged across 10 repeated runs using CLUMPP, version 1.1.2 (Jakobsson & Rosenberg 2007), and averaged repeated runs for $K = 2$ were visualized using an R script (R Development Core Team 2013). The choice of the true number of clusters (*K*) was based on the method of Evanno *et al.* (2005), implemented in Structure Harvester (Earl & vonHoldt 2012) (Fig. S3, Supporting information).

Isolation by distance and isolation by environment

We used a hierarchical AMOVA (Excoffier *et al.* 1992), as implemented in ARLEQUIN, version 3.5.1.2 (Excoffier & Lischer 2010), to search for a significant genetic differentiation between high- and low-elevation plots within each site, as a potential indicator of isolation by environment (IBE) in comparison with genetic differentiation between sites, which is most probably driven by isolation by distance (IBD). Tests were repeated by excluding only Vtx and Ves and only Iss and Lur, respectively. Statistical significance was evaluated using 10^4 permutations. Furthermore, we used Mantel and partial Mantel tests implemented in the R package, ECODIST (Goslee & Urban 2007), to distinguish between patterns of IBD and IBE using all loci or only the loci for which we detected evidence of directional selection with any other method (i.e. those in Table 1). Genetic distances between plots were described using Reynold's distances and were calculated using a custom-made R script available at https://qgsp.jouy.inra.fr/archives/FLK/FLK.R. Geographic distances were calculated as Euclidean physical distances between sites in terms of longitude, latitude and elevation. Environmental distances were calculated for PC1 and PC2 of the bioclimatic variables (Fig. 2 E).

Frequency-based outlier detection

Frequency-based outlier detection was carried out using hierarchical FDIST (subsequently HFDIST) (Beaumont & Nichols 1996; Excoffier *et al.* 2009), BAYESCAN, version 2.1 (Foll & Gaggiotti 2008), and FLK (Bonhomme *et al.* 2010). In the BAYESCAN and FLK analyses, each one of the eight sampling plots was considered as a population. In the HFDIST analysis, we defined six groups according to the population structure obtained by the STRUCTURE analysis (i.e. group 1: Vtx_L/H, group 2: Lur_L, group 3: Lur_H, group 4: Iss_L, group 5: Iss_H and group 6: Ves_L/H). All three approaches provide some advantages in the context of our data set, even though some of them have been criticized for high false-positive rates (e.g. Lotterhos & Whitlock 2014). Notably, HFDIST can explicitly take into account a hierarchical structure such as in our sampling design (two plots per site). The null model of BAYESCAN assumes a common migrant gene pool, which may be suited here, because the populations are likely to originate from the same refugium. Finally, FLK uses a phylogenetic estimation of the population's kinship matrix and thus accounts for the historical branching and the heterogeneity of genetic drift.

The HFDIST analysis was performed using ARLEQUIN, version 3.5.1.2 (Excoffier & Lischer 2010). Fourteen loci having negative F_{ST} values in a global AMOVA were excluded. The neutral envelope was established using 10^5 simulations. We used BAYESCAN, version 2.1, with 20 pilot runs of 5×10^3 iterations and a burn-in of 5×10^4 iterations followed by 5×10^4 iterations for the estimation of the posterior distributions with a thinning interval of 10. Because the used SNP markers are located on candidate genes, we decreased the prior odds to 2 (10 by default) following Csilléry *et al.* (2014), who argue that the probability of a SNP being under selection is higher for SNPs situated in candidate genes than for randomly chosen SNPs. For FLK, the neighbour-joining (NJ) tree of Reynold's genetic distances between plots was rooted using *A. alba* samples from Romania that were genotyped at the same set of loci. We ran 5×10^4 simulations to construct the neutral envelope.

Environmental association-based outlier detection

The latent factor mixed model (LFMM, Frichot *et al.* 2013) is a variant of the Bayesian principal component analysis in which residual population structure is introduced via latent factors. We used a model with four latent factors as compromise between the weak observed population structure and a sufficient number of latent factors to control the number of potential false positives. Two sets of environmental variables were tested: (i) the four key bioclimatic variables (Fig. 2A–D; subsequently LFMM_env) and (ii) PC1 and PC2 (Fig. 2E; subsequently LFMM_PCA). For both tests, we ran 10^5 MCMC iterations with 5×10^4 burn-in steps with 10 replicate runs. *Z*-scores from replicate runs were combined using the Fisher–Stouffer method and adjusted using the genomic inflation factor (λ; Fig. S4, Supporting information). The genomic inflation factor measures the excess

Table 1 Summary of all loci with evidence of directional selection in decreasing order of evidence per method (ordered as BAYESCAN, FLK, LFMM), their functional annotation obtained using BLAST2GO v2.8, the amino acid change (AAC) for SNPs that code for nonsynonymous mutations and references to other studies (nonexhaustive Web of Science http://apps.webofknowledge.com search in January 2015) of forest trees that found evidence of directional selection from similar genes

Locus	BAYESCAN *q*-value	FLK *P*-value	LFMM_env (scPDSI_jfm)	Annotation	AAC	References
15808_snp829	2.12e-02	9.53e-05		Chalcone synthase 2	L->M	Pinus sylvestris (Wachowiak *et al*. 2011)
15663_snp771	2.20e-02		9.10e-05	Heat shock protein 83-like		*Pinus cembra*, *Pinus mugo* (Mosca *et al*. 2012a), Fagus sylvatica (Csilléry *et al*. 2014)
7150_snp658	3.21e-02			GTP-binding protein sar1a		
5004_snp671	3.96e-02	4.45e-11		Magnesium chelatase subunit chloroplastic-like		
6804_snp247		2.00e-38		Tubulin beta-3 chain-like		
17813_snp275		4.50e-33		ADP-ribosylation factor 1-like 2		
2161_snp64		8.89e-29		Proteasome subunit beta type-6-like	M->I	*Cryptomeria japonica* (Tsumura *et al*. 2012), *Populus tremula* (De Carvalho *et al*. 2010)
2937_snp405		1.25e-10		Fructose bisphosphate aldolase cytoplasmic isozyme		
16125_snp157		9.65e-08		Ubiquitin-like protein		*Pinus cembra* (Mosca *et al*. 2012b), *Pseudotsuga menziesii* (Howe *et al*. 2003), *Picea mariana* (Prunier *et al*. 2011)
643_snp311		1.57e-06		5-Methyltetrahydropteroyltriglutamate–homocysteine methyltransferase 1-like		
14915_snp331		2.72e-06		Predicted: uncharacterized protein C167.05-like isoform 4		
20680_snp343		4.00e-06		Probable phenylalanine–tRNA ligase alpha subunit		
24159_snp152		4.06e-05		Proteasome subunit beta type-7-b		*Cryptomeria japonica* (Tsumura *et al*. 2012), *Populus tremula* (De Carvalho *et al*. 2010)
9065_snp715		7.47e-05		Proteasome subunit beta type-1-like		*Cryptomeria japonica* (Tsumura *et al*. 2012), *Populus tremula* (De Carvalho *et al*. 2010)
1181_snp99/201		1.17e-04		Serine/threonine-protein phosphatase pp2a-2 catalytic subunit-like	A->E (snp99)	*Pinus cembra* (Mosca *et al*. 2012a)
10568_snp484*			1.40e-05	Mannose-1-phosphate guanylyltransferase 1-like		
1455_snp289			2.30e-05	Succinate dehydrogenase	S->P	
14580_snp627			3.20e-05	ATP-dependent Clp protease proteolytic subunit-related protein chloroplastic-like		
20267_snp369			1.10e-04	Actin-like isoform x1		
24145_snp82			1.14e-04	60S ribosomal protein l7-4-like isoform x2	P->H	
5004_snp249			3.85e-04	Magnesium chelatase subunit chloroplastic-like		
6757_snp154			6.35e-04	Ribonucleoside diphosphate reductase large subunit		

BAYESCAN was used with FDR <0.1, FLK with *P*-value ≤1.87e-04 and LFMM_env with FDR <0.05. For LFMM_env, adjusted *P*-values of outliers are shown for scPDSI_jfm. Locus 10568_snp484 (indicated with an *) was also an outlier for relative humidity with *P*-value = 2.40e-05, for VPD_mam with *P*-value = 1.77e-04 and also in LFMM_PCA for PC1 with *P*-value = 1.20e-05.

of the false discovery rate due to multiple testing, and it is defined as the ratio of the observed and the expected median of the distribution of the test statistic (Devlin & Roeder 1999). λ was estimated as median $(Z^2)/0.456$. We corrected for multiple testing by fixing the false discovery rate to 5%.

Reanalysis of published common garden data

Sagnard *et al.* (2002) used a common garden trial to assess phenotypic differences across 16 sites in the southwestern Alps, including the four north-facing mountain slopes that were sampled in this study. Here, we reanalysed these published data for the four mountains studied herein: VTX in Table 1 of Sagnard *et al.* (2002) is identical to Vtx, LUR to Lur, LAB to Iss and TUR to Ves. Because we did not sample the same mother trees that Sagnard *et al.* (2002) sampled in 1994 for sowing, we reconciled the two sampling efforts by redefining a comparable set of populations (see details in Appendix S2, Supporting information). As a result, we obtained five populations with both SNP and phenotype data: Vtx_L, Vtx_H, Lur_H, Iss_H and Ves (i.e. Ves_L and Ves_H pooled).

Sagnard *et al.* (2002) collected seeds from 24 (Vtx_L), 22 (Vtx_H), 32 (Lur_H), 30 (Iss_H) and 43 (Ves) mother trees with, on average, nine offspring per mother tree. Seeds were sown and planted in 1995, and seedlings were monitored until the end of 1999 in a nursery located near Aix-en-Provence, southeastern France. A balanced block design was used where half of the seedlings were exposed to a drought treatment in 1998. The state of the bud flush, the growth of control trees and the effect of the drought treatment on plant health were recorded (see Sagnard *et al.* 2002 for details). Bud flush was coded from 1 to 5 (1: dormant bud, 2: swelling bud, 3: swollen bud, leaf green tips visible, 4: bud broken, shoot elongation started and 5: shoot elongated with first unfolded leaf visible), growth was measured as height increment, and response to water stress was coded from 0 to 4 (0: no sign of stress, 1: lateral shoots start to wither, 2: terminal shoot starts to wither, 3: whole plant is withered and needles are yellowing and 4: plant died after the treatment).

We used a linear mixed-effects model with site as fixed effect and family as random effect to test whether there were significant differences among sites in any traits (using the R function lme and model fit with REML). We performed a Q_{ST}–F_{ST} test of adaptive divergence using the above-described traits and the candidate gene data. Q_{ST} was calculated for all traits using raw values for growth and log-transformed values for categorical traits (i.e. bud flush and response to stress). F_{ST} was calculated using all SNP loci, except those for which we detected evidence of directional selection with any other methods (i.e. the 23 loci listed in Table 1). Calculations and the Q_{ST}–F_{ST} tests were performed using the bootstrap procedure described in Whitlock & Guillaume (2009) and using the R package QstFstComp associated with Whitlock & Gilbert (2012). We considered a one-tailed test, because we were interested in looking for adaptive divergence only, thus F_{ST} being significantly smaller than Q_{ST}.

Silver fir is a predominantly outcrossing species. However, it may exhibit a relatively high selfing rate, especially when population density is low. Restoux *et al.* (2008) estimated the outcrossing rate of silver fir in Ventoux and found that the mean outcrossing rate was typical for mixed mating conifers (tm = 0.85), but strongly dependent on forest plot density and varied from 0.87 in high-density to 0.43 in low-density marginal stands. Thus, here we assumed that the average relatedness between progenies of the same mother trees was 0.2875 when calculating Q_{ST}, which corresponds to the mean tm. Due to the large variation in outcrossing rates, we also investigated the sensitivity of Q_{ST} to the relatedness in the true half-sib (0.25) and true full-sib (0.5) range.

Results

Population structure and IBD

The average fixation index (F) for all loci over all individuals was low (0.031). Pairwise F_{ST} between plots also revealed a weak genetic differentiation across pairs of plots: 0.017, on average, the lowest (0.009) between Lur_L and Iss_L and the highest (0.025) between Vtx_L and Iss_L. $K = 2$ had the highest support in STRUCTURE (Fig. S3, Supporting information), suggesting that Vtx and Ves represent two distinct populations (Fig. 1B), while the two geographically intermediate sites, Lur and Iss, contain a mixture of Vtx- and Ves-like individuals with some genetic differentiation between high- and low-elevation plots (Fig. 1B). Reynold's genetic distances between plots revealed the highest distance between the Vtx_L and Ves_L (0.245), while the highest intersite between-plot genetic distance was observed between the two Vtx plots (0.196) (Table S2, Supporting information). Four pairs of SNPs that lie on different contigs and 22 pairs of SNPs that lie on the same contig exhibited a significant LD (r^2 between 0.22 and 0.88 and P-value <0.05) in at least one of the two clusters identified by STRUCTURE. However, due to this low number of SNPs in LD, we decided to keep all SNPs for all further analyses and assumed their statistical independence. After Bonferroni correction, 0.636% of the SNP pairs showed a significant LD (226 SNP pairs) with r^2 rang-

ing from 0.033 to 0.991; 168 of the significant LD arose between SNPs from different contigs, and 58 between SNPs on the same contig. Given this negligible number of SNP pairs with a significant LD, we kept all SNPs for further analysis.

The hierarchical AMOVA revealed that most of the variance in allele frequencies between individuals was due to the differences within plots (97.0%, $P < 0.001$, $F_{ST} = 0.030$, Table S3, Supporting information). Nevertheless, a weak yet significant genetic differentiation was detected between sites (1.2%, $F_{CT} = 0.0119$, $P < 0.001$) and between high- and low-elevation plots within sites (1.8%, $F_{SC} = 0.0185$, $P < 0.001$, Table S3, Supporting information). In agreement with STRUCTURE results, a hierarchical AMOVA including only Vtx and Ves revealed a higher between-site variance (3.1%), while the analysis including only Lur and Iss showed nearly 0% between-site variance and a 2.4% variance between plots within sites (Table S3, Supporting information).

The Mantel test of IBD corroborated the STRUCTURE and AMOVA results showing a significant IBD pattern using all loci (Mantel $r = 0.35$ with 95% confidence interval: 0.159–0.448, P-value = 0.045, Fig. 3A). The partial Mantel test with geographic distance accounting for PC1 and PC2 of the four bioclimatic variables (Fig. 2E) resulted in a slightly stronger IBD pattern (PC1: $r = 0.35$, P-value = 0.044, PC2: $r = 0.432$, P-value = 0.022). Mantel tests with PC1 and PC2 only were not significant, indicating the absence of a genome-wide IBE pattern (Fig. 3B, grey dots; PC1: $r = -0.018$, P-value = 0.49, PC2: $r = -0.264$, P-value = 0.184).

Frequency-based outlier detection

Using three different F_{ST} outlier approaches, a total of 16 unique SNP loci showed unusually high levels of differentiation relative to a null model after correcting for multiple testing (Fig. 4, Table 1). Four SNPs showed a signal of divergent selection using BAYESCAN (Fig. 4A), and one of them coded for a nonsynonymous mutation (Table 1). Three of these SNPs also revealed evidence of divergent selection when using HFDIST (Fig. 4B), which became nonsignificant after correcting for multiple testing. FLK revealed 14 SNPs with unusually high levels of differentiation, three of them showing extremely high values in the FLK statistics: 6804_snp247, 17813_snp275 and 2161_snp64 (Fig. 4C).

Four of the FLK outliers coded for nonsynonymous mutations (Table 1). Two outlier SNPs were detected both by BAYESCAN and by FLK (5004_snp671 and 15808_snp829). Both HFDIST and FLK revealed two loci with unusually low F_{ST} values (Fig. 4), thus indicating balancing selection. However, we decided to ignore these because of the weak power of outlier tests to detect balancing selection and because of the overall low F_{ST} in our samples.

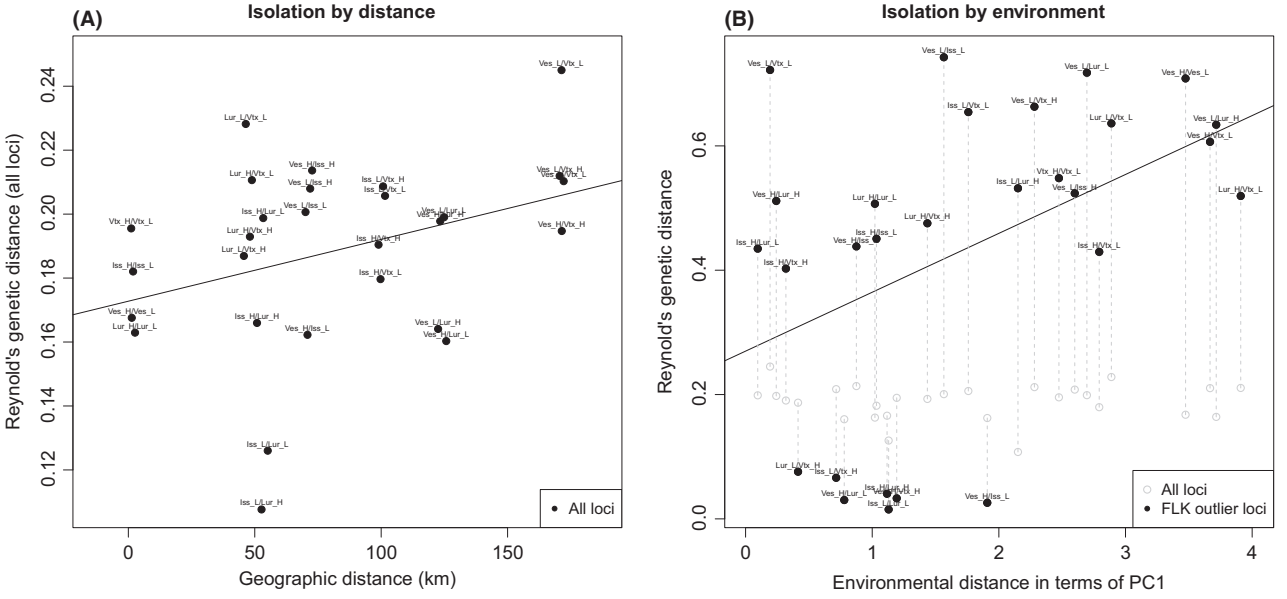

Fig. 3 A: Isolation by distance in terms of Reynold's genetic distance as a function of Euclidean geographic distance of longitude, latitude and elevation. Each point indicates a plot pair. B: Isolation by environment in terms of Reynold's genetic distance as a function of environmental distance in terms of PC1 using all SNPs (grey dots) and using FLK outlier loci only (filled black dots). Grey lines show how the estimated genetic distance changes between plot pairs from using all SNPs to using outlier SNPs.

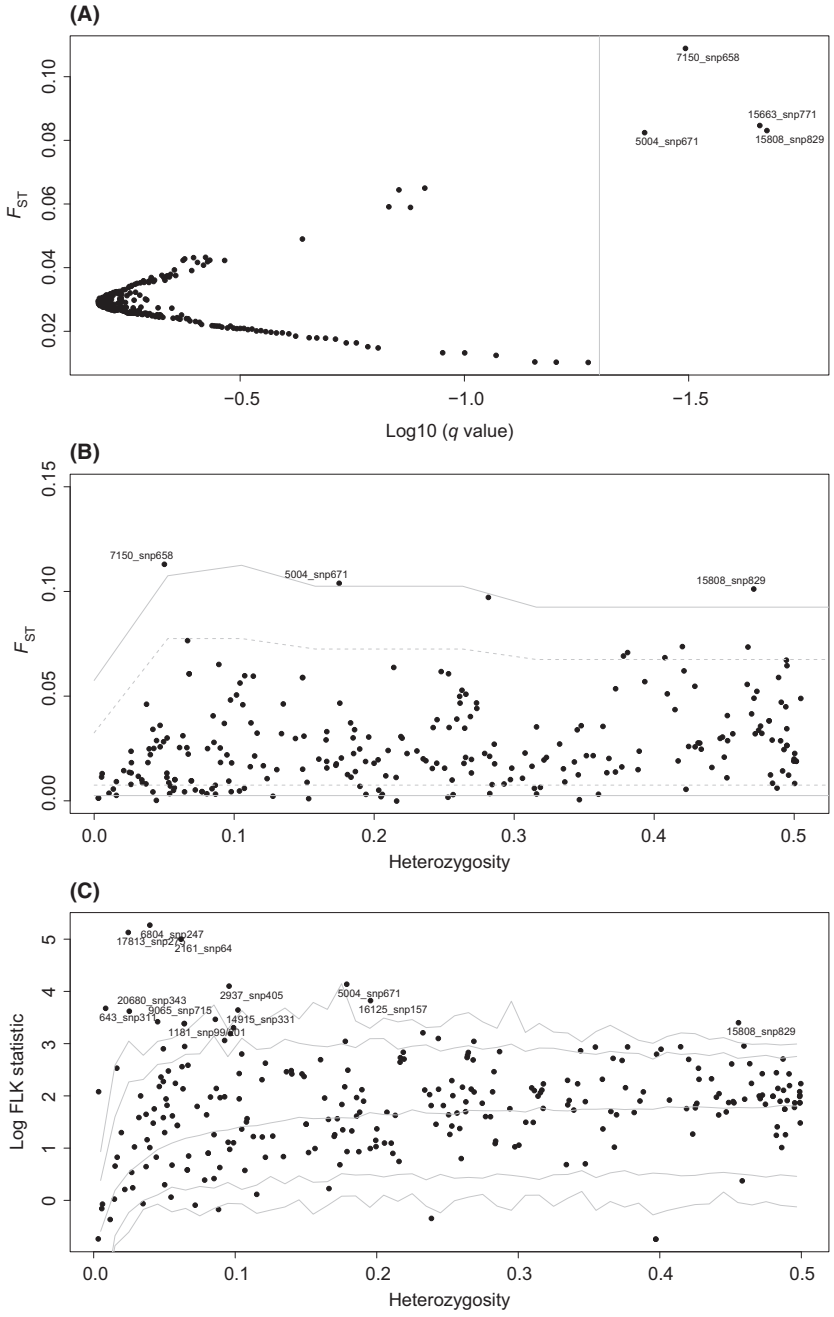

Fig. 4 Evidence for directional selection from three F_{ST} outlier tests. A: Distribution of observed F_{ST} values as a function of log10 q-values obtained using BAYESCAN. The vertical line indicates the log10 q-value that corresponds to a false discovery rate FDR = 0.1. B: Distribution of observed F_{ST} values as a function of heterozygosity obtained using HFDIST. The solid (dashed) line indicates the 1% (5%) and 99% (95%) boundaries of the simulated, neutral distribution of F_{ST} values obtained with 50 000 simulations assuming a hierarchical island model and six groups (Vtx_L and Vtx_H, Lur_L, Lur_H, Iss_L, Iss_H, Ves_L and Ves_H). C: Distribution of log FLK statistic as a function of heterozygosity. Lines depict the 0.5%, 2.5%, 50%, 97.5% and 99.5% boundaries of the empirical neutral distribution of FLK statistic obtained by forward in time simulation. On all graphs only outlier SNPs are labelled with their SNP IDs.

We observed that most FLK outlier loci exhibited highly contrasted allele frequencies between high- and low-elevation plots. To formally test whether these outliers indicate the presence of local adaptation to elevation, we repeated the Mantel test using only genetic distances calculated from the FLK outlier SNPs against PC1 distances. This is because PC1 principally separated high- and low-elevation plots. We found that IBD was no longer significant (r = 0.034, P-value = 0.401), while IBE was significant (Mantel: r = 0.44, P-value = 0.033, partial Mantel with geographic distance: r = 0.442, P-value = 0.039). These results indicate that on average, plots from similar environments (Vtx_H, Lur_L, Iss, Ves_H) are genetically more similar at outlier loci than at all loci, and plots from different environments are genetically more distant at outlier loci than at all loci (Fig. 3B).

Environmental association-based outlier detection

Using a stringent criterion for candidate SNP detection with LFMM (i.e. correcting z-scores for λ and

FDR = 0.05), LFMM_env revealed eight SNPs that were all associated with scPDSI_jfm, and one of them, 10568_snp484, was also associated with RH and VPD_mam (Table 1, Fig. 5). No outliers associated with VPD_son were detected. LFMM_PCA also revealed a significant association between 10568_snp484 and PC1 (Table 1, Fig. 5). Using a less stringent correction for multiple testing (correction for λ but adjusted *P*-value <0.001), the outlier SNP list of LFMM_env did not change, but three additional outliers appeared for PC2 in LFMM_PCA (5004_snp249, 1455_snp289, 6119_snp125), two of them already detected using LFMM_env. Genomic inflation factors (λs) were often much higher than 1 (Fig. S4, Supporting information), which is not surprising because we have a relatively small, and candidate gene-type data set, with a weak, but significant population structure. Under such conditions, we may fail to correctly estimate the variance of allele frequencies at neutral loci. Thus, correcting for λ, *P*-values were significantly increased as illustrated by change in the shape of the histograms (Fig. S4, Supporting information).

All SNPs that showed a significant environmental correlation were associated with winter drought (scPDSI_jfm, Table 1, Fig. S4, Supporting information), the environmental variable that suggested, on average across years, less winter drought from east to west (Fig. 2A). Accordingly, outlier SNPs showed an east-to-west variation in allele frequencies (Fig. 6). However, this variation in allele frequencies was not completely congruent with the gradual change in scPDSI_jfm. This is partially due to the fact that we plotted the raw data on Fig. 6, while in LFMM allele frequencies are adjusted for population structure (here with $K = 4$). Further, when only a few points are available in space, LFMM may detect SNPs along an environmental axis that is orthogonal to the first PCA axis of the SNP data (Frichot *et al.* 2015; see more details in the Discussion), even if the true underlying adaptive landscape is more complex. Finally, visual assessment of Fig. 6 suggests that it is often the geographically central sites, Lur and Iss, that show a higher (or lower) allele frequency than predicted by a linear east-to-west trend (see SNPs 20267_snp369 or 24145_snp82 as extreme examples), which is in accordance with phenotypic observations from the common garden data (see more details below and Fig. S6, Supporting information).

Locus 10568_snp484 also showed a strong evidence of selection with respect to PC1 (Figs 2E and 5), suggesting a role in adaptation to climatic differences imposed by elevation. Indeed, SNP 10568_snp484, along with 1455_snp289 and 24145_snp82, showed the highest differences in allele frequencies between high- and low-elevation plots, on average (results not shown). The locus 15663_snp771 was detected as an outlier by both LFMM_env and BAYESCAN and shows a strong east–west-orientated change in allele frequencies. An LFMM_env outlier SNP was also detected on gene 5004, where a different SNP also showed evidence of selection both using BAYESCAN and using FLK (Table 1). Thus, gene 5004 shows evidence of selection with all three methods. Two of the LFMM outliers were found to code for nonsynonymous mutations (Table 1).

Functional annotation of candidate genes and SNPs

The alignment of candidate gene sequences with *P. abies* protein sequences allowed to annotate all genes and to differentiate between synonymous and nonsynonymous SNPs (see outlier SNPs in Table 1 and all loci in Table S4, Supporting information). A total of 187 of 267 SNPs occurred within coding sequences, so that their synonymous or nonsynonymous state could be determined. Among all SNPs, 95.7% of the outlier SNPs occurred in coding regions (22 of 23 SNPs), which was significantly higher than among the nonoutlier SNPs (165 of 244 SNPs; Fisher's exact test *P*-value = 0.002). No significant enrichment was observed for the coding outlier SNPs for nonsynonymous mutations in comparison with the nonoutlier SNPs (Fisher's exact test *P*-value = 0.214).

Reanalysis of published common garden data

Descendants of mother trees originating from five of the plots and grown in common garden significantly differed in bud flush, growth and resistance to drought at several observations (Fig. S6, Supporting information). During the first-year phenological observation (01/04/1997), seedlings originating from Lure had the earliest bud flush date, followed by Issole and Ventoux, while Ves flushed the latest (Wald test: $F_{4,69} = 6.28$, *P*-value <0.001; Fig. S6A, Supporting information). Differences stayed significant during the second observation (04/15/1997, Wald test: $F_{4,69} = 4.2$, *P*-value = 0.004), but evened out during the following year (Wald test: $F_{4,69} < 2.21$, *P*-value >0.076). However, in 1999, differences became significant again and showed an east-to-west decrease in bud flush precocity (04/01/1999, Wald test: $F_{4,69} = 3.05$, *P*-value = 0.022; Fig. S6B, Supporting information). In terms of growth increment, in 1999, we observed that Lur_H lagged behind the other four populations (Wald test: $F_{4,69} = 6.65$, *P*-value <0.001; Fig. S6C, Supporting information). The water stress treatment also differentially affected the individuals from the four different sites, with Lur_H being the most resistant (most plants showed no sign of weakness) and Ves being the most fragile, where in every progeny at

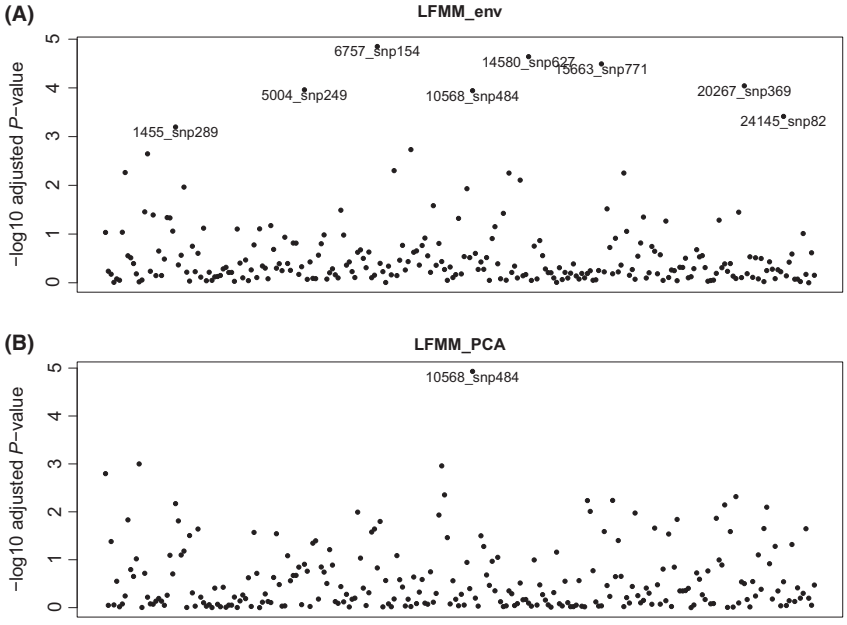

Fig. 5 Manhattan plot of LFMM_env and LFMM_PCA. *X*-axis represents an arbitrary order of the candidate genes SNPs. Only the *P*-values of significant SNPs with FDR <0.05 are labelled. A: LFMM_env -log10 *P*-values for the environmental variable scPDSI_jfm. B: LFMM_PCA -log 10 *P*-values for PC1.

least one seedling showed signs of weakening. The strongest differences among sites were observed on 06/09/1998 (Wald test: $F_{4,69} = 4.02$, P-value = 0.005; Fig. S6D, Supporting information). The maximum differences between sites were observed one and a half months after the treatment, while no differences were observed 1 year after the treatment (probably because most affected seedlings had died the year before, results not shown). When comparing phenotypic divergence with divergence at all but the outlier SNP loci (Table 1), the Q_{ST}–F_{ST} test was marginally significant for adaptive divergence ($Q_{ST} > F_{ST}$) for bud flush on 04/01/1997 and for the growth increment in 1999 (Table 2). As expected, the Q_{ST}–F_{ST} test was sensitive to the assumed relatedness between progeny within the family (Fig. S7A, Supporting information). However, even assuming the strongest possible relatedness, that is all progenies of the same mother are full-sibs, qualitatively the same results were obtained with slightly smaller *P*-values (Fig. S7B, Supporting information).

Discussion

In this study, we combined candidate gene, climatic and phenotypic data to study divergent selection at drought and cold tolerance candidate genes and phenotypic traits in peripheral populations of silver fir in the French Mediterranean Alps. We found that the genetic structure of the four studied mountains was dominated by a neutral pattern of east–west IBD (Fig. 3A), while the adaptive genetic landscape was more complex, and suggested the presence of locally advantageous genetic variants both at the landscape scale and between high and low plots within mountains. We identified at least three possible components of the selective environment that may drive the presence of locally adapted variants. First, at least eight SNPs suggested directional selection along an east–west axis that represents a winter drought gradient (Table 1, Fig. 6). Second, many of these loci show allele frequency patterns that indicate the presence of different selective forces at Lure (Figs 1 and 6), a site that can be characterized by the presence of humid and cold winters (with over 2 m of snow) and springs (Fig. 2). The signature of this second component can be detected not only at the genomic, but also at the phenotypic level. Adaptive divergence at phenotypic traits measured in the common garden revealed that seedlings from Lure flushed earlier, grew more slowly and better survived drought stress than any other provenances (Table 2, Fig. S6, Supporting information). Third, a significant IBE was detected at the FLK outlier loci along an environmental axis (PC1) that principally distinguished low- and high-elevation plots (Figs 2E and 3B), suggesting the presence of selective forces at short geographic scale despite likely gene flow (as suggested at Ventoux site by Davi *et al.* 2011).

Demographic history of silver fir in the French Mediterranean Alps

A weak but significant genetic differentiation was found among sites, with a low average F_{ST} over all loci (0.017), lower to what has been shown for Italian

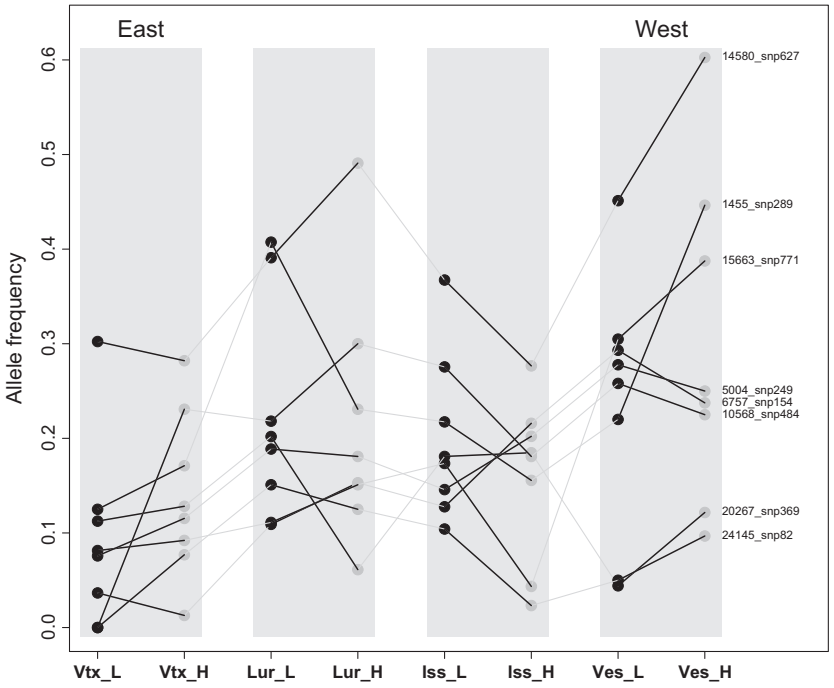

Fig. 6 Allele frequency clines of SNP loci that showed significant associations with scPDSI_jfm in LFMM_env. Grey rectangles illustrate the four study sites from east to west. Allele frequencies at low (high) elevation are indicated with black (grey) dots. Grey lines connect the allele frequencies of each SNP locus across the eight sampling plots, and SNP IDs are indicated at the right end of each line.

Table 2 Q_{ST}–F_{ST} tests for phenotypic traits measured in a common garden established in 1997 with provenances from the four studied mountains

Trait	Date of observation	Q_{ST}	Q_{ST}–F_{ST}	Lower bound	Upper bound	*P*-value
Bud flush	04/01/1997	**0.1016**	**0.0735**	**−0.0418**	**0.1064**	**0.054**
	04/15/1997	0.0604	0.0323	−0.0437	0.1128	0.181
	04/03/1998	0.0196	−0.0085	−0.0985	0.2896	0.496
	04/17/1998	0.0203	−0.0079	−0.0535	0.1762	0.4961
	04/01/1999	0.0416	0.0135	−0.0414	0.1103	0.3115
	04/13/1999	−0.0083	−0.0364	−0.0481	0.1534	0.8744
Growth rate (mm)	Summer 1999	**0.1002**	**0.0721**	**−0.0452**	**0.1294**	**0.0754**
Response to water stress	06/05/1998	0.0293	0.0012	−0.0341	0.0712	0.4003
	06/09/1998	0.0312	0.0031	−0.0352	0.0696	0.3727
	06/17/1998	0.0232	−0.0049	−0.0353	0.0711	0.4837
	06/22/1998	−0.0034	−0.0315	−0.0382	0.0831	0.8971

Five populations were used in this analysis that correspond to Vtx_L, Vtx_H, Lur_H, Iss_H and Ves (i.e. Ves_L and Ves_H pooled; see more details in Supplementary Material A). Q_{ST}'s were calculated for bud flush, growth rate and response to water stress. Bud flush and growth rate were measured on well-watered plants, while response to water stress was measured for individuals that were treated with a 'no-watering' starting 6 May 1998 and lasting all summer. Q_{ST} is reported assuming relatedness of 0.2875 between progenies of the same family, which corresponds to an outcrossing rate of 0.85. F_{ST} was calculated over all candidate gene SNPs except the outliers listed in Table 1, was 0.0281 (with lower and upper 95% confidence interval of 0.0251 and 0.0312). Q_{ST}–F_{ST} is reported with its lower and upper 95% critical values and one-tailed Monte Carlo *P*-values based on 10^4 bootstrap simulations. Lines in boldface indicate marginally significant deviations from neutrality.

populations of *A. alba* (0.037 in Mosca *et al.* 2013) and generally for conifers (e.g. 0.035–0.05 for *Picea abies*, Chen *et al.* 2012). We cannot, however, exclude a bias introduced by our choice of marker loci because we only used candidate gene loci, as is commonly suggested for nonmodel conifer species with large genome sizes (e.g. Neale & Kremer 2011). Interestingly, studies in model species with a dense genome-wide coverage of SNP loci generally only detect a very small percentage of loci exhibiting non-neutral patterns (e.g. Four-

nier-Level *et al.* 2011). Therefore, it was expected that most of our loci would also behave neutrally.

The dominant pattern of neutral genetic variation was IBD from east to west (Figs 1B and 4, Table S2, Supporting information), which is in agreement with the hypothesis that after the last glacial maximum, the French Southern Alps were recolonized from a common refugium in Central Italy (Liepelt *et al.* 2009). After this Holocene recolonization, populations in the different mountains probably became genetically isolated because suitable habitats for silver fir were restricted to high elevations (approximately above 900 m), and therefore, the lowlands between the mountains constituted physical barriers for gene flow. This isolation process may have been accentuated by the climatic and recurrent land-use changes of the past centuries (Liepelt *et al.* 2009; Lander *et al.* 2011). In contrast, the significant differentiation found between high- and low-elevation plots at Lure and Issole (Fig. 1B) may suggest gene flow between different mountains. We found that both at Lure and Issole, the high-elevation plots contained more Ventoux-like individuals than low-elevation plots. This may have resulted from pollen flow from Ventoux driven by the predominant N–NW wind of the region (the 'Mistral'). Such long-distance pollen dispersal has been well documented in conifers (Kremer *et al.* 2012).

Detecting the signature of selection using F_{ST} outliers

A growing number of simulation studies illustrate the advantages and limitations of different F_{ST} outlier methods for detecting loci under selection, with overall little concordance regarding the best performing method and sampling design (Pérez-Figueroa *et al.* 2010; Narum & Hess 2011; De Mita *et al.* 2013; Lotterhos & Whitlock 2014, 2015; de Villemereuil *et al.* 2014). This is simply because the problem of detecting loci under selection is extremely difficult due to the confounding effect of demography (Tiffin & Ross-Ibarra 2014) and also because methods are often evaluated based on different demographic scenarios. The demographic history of our model species and our nested sampling design may have contributed to finding partially concordant evidence of adaptation across methods (Table 1). Indeed, when IBD is the dominant neutral population structure, the detection of signatures of selection using F_{ST} outlier methods is generally improved (Lotterhos & Whitlock 2015). Furthermore, sampling pairs of plots from two contrasted environmental conditions, low and high elevation in our case, is an efficient sampling strategy for outlier tests (Lotterhos & Whitlock 2015). Apart from the above-mentioned statistical limitations in detecting signatures of selection, the candidate SNP-only nature of our data may have influenced our findings. As already mentioned, most of our SNPs can be assumed to behave neutrally. Nevertheless, background F_{ST} levels may be slightly higher than expected for 'truly' neutral SNPs. As a result, our data set may actually tend to be conservative for F_{ST} outlier tests.

Between 1.5 and 5.4% of the loci were detected as outliers depending on the method used. The three F_{ST} outlier tests (BAYESCAN, HFDIST and FLK) used in our study showed a rather unusual concordance: the three HFDIST outliers (before Bonferroni correction) were common with BAYESCAN, and two of the four BAYESCAN outliers were also detected using FLK. Mosca *et al.* (2013), who studied selection in Italian populations of silver fir using a candidate gene data set of similar size, detected four outliers using BAYESCAN and four outliers using HFDIST, among which only one or two passed Bonferroni correction depending on the group definitions, and they identified no overlap between the two methods. FLK has been recently reported as the most promising F_{ST} outlier approach because of its low false-positive rate in comparison with BAYESCAN and FDIST in population expansion (nonequilibrium IBD) scenarios (Lotterhos & Whitlock 2014). We also found that FLK outliers, on average, corresponded to a plausible biological pattern of adaptation to different climatic conditions imposed by elevation (Fig. 3B).

Detecting the signature of selection using gene–environment correlations

It is difficult to know what is the selective environment of any organism and landscape (Lotterhos & Whitlock 2014), and this is even more true for long-lived species such as forest trees. We chose bioclimatic variables to approximate the selective environment that are supported by other studies as likely to influence the fitness of forest trees (Kühl *et al.* 2002; Howe *et al.* 2003) and also showed markedly higher among-site differences in comparison with the interannual differences (Fig. 2). In this context, the sampling design of our study was slightly disadvantaged by having relatively few observations and by the fact that environmental variables were available only at plot level (i.e. eight unique observations). We attempted to obtain approximate estimates of the environment at the individual tree level (by interpolating temperature at each elevation), and reassuringly, we detected the same associations as in our plot-level environmental data study (Appendix S3, Supporting information).

All candidate SNPs in LFMM_env were associated with a winter drought gradient, which is indeed a plausible selective gradient for silver fir. Nevertheless, two points have to be discussed. First, the winter drought gradient is parallel to the dominant IBD pattern, thus

orthogonal to the first PC axis of the genetic data. It has been shown that maximum statistical power and the least number of false-positive rates may be expected under such conditions (Frichot *et al.* 2015). However, it does not mean that this axis is the best representation of the selective environment. Second, in fact, allele frequency patterns of the candidate SNPs were not fully congruent with the east–west cline of winter drought, but also reflected the divergence of Lure from the other sites (Fig. 6). This suggests that many of these SNPs are, in fact, associated with an environmental gradient that our other environmental variables perhaps could not grasp and/or that LFMM did not have sufficient power to detect with only a limited number of observations.

Phenotypic divergence in the common garden

In line with the results of previous meta-analyses (De Kort *et al.* 2012), the pattern emerging from this study is that Q_{ST} is generally low, but still slightly exceeds F_{ST} (Table 2). Thus, directional natural selection seems to be the most common cause for divergence in the studied traits. The strongest result that emerges from this study is that the coldest site, Lure, contains some unique variants related to adaptation to cold and drought. Seedlings with mothers from Lure flushed buds earlier, but grew more slowly than seedlings originating from the other sites (Fig. S6A–C, Supporting information). It has commonly been observed in conifers that in low-elevation, mild-condition common gardens, individuals from high-elevation sites tend to flush bud earlier to maximize the length of the growing season during available frost-free periods (e.g. Mimura & Aitken 2010; Kreyling *et al.* 2011). Similarly, populations from colder habitats tend to grow slower than those from warmer habitats, suggesting that growth potential and cold hardiness are negatively correlated (Rehfeldt *et al.* 2002; reviewed in Alberto *et al.* 2013). This could be interpreted as an adaptive strategy, where seedlings from cold environments flush as soon as they reach their required sum of temperature, which is lower than that of seedlings from warm environments (Vitasse *et al.* 2009). Lure was also the most resistant to drought (Fig. S6D, Supporting information). A correlation between drought and frost resistance has been observed previously and could possibly be explained, for example, by resistance to xylem embolism (Lens *et al.* 2013; Sengupta & Majumder 2014). Accordingly, recent common garden experiments also found that foliage cold hardiness differs between populations of *Pinus nigra* (Kreyling *et al.* 2012) and that populations most resistant to cold are also the most drought resistant (Thiel *et al.* 2014). Overall, the analysis of the common garden data corroborates our findings from the gene–environment association study and suggests winter climate as a potentially strong selective driver in climate change (Kreyling 2010).

Conclusions from combining genetic, environmental, phenotypic and functional evidence

Incorporating genetic, phenotypic and environmental data may help to uncover the mechanisms of adaptation of natural tree populations (e.g. Le Corre & Kremer 2012; Savolainen *et al.* 2013; Sork *et al.* 2013; Lotterhos & Whitlock 2015). Indeed, the most convincing results have been provided by studies that combined, for example, genomic data from natural population samples with data from common garden experiments (e.g. De Kort *et al.* 2014; Lepais & Bacles 2014). Here, we combined results from three F_{ST} outlier tests, a gene–environment association method, a common garden experiment (Sagnard *et al.* 2002), and we also functionally annotated the studied candidate genes. As a result, we were able to suggest bioclimatic elements of the selective environment, some phenotypes that express signals of adaptive divergence and also some candidate genes that may underlie local adaptations. Due to a limited number of spatial observations, we could not explore phenotype–environment correlations; nevertheless, it is perhaps not just a coincidence that Lure represents a both phenotypically and climatically extreme site.

The loss of forested areas in Europe has important economic consequences (Hanewinkel *et al.* 2013). In the Mediterranean, numerous decline and die-back events that affect forest dynamics have recently been reported (e.g. Linares & Camarero 2012; Cailleret *et al.* 2014). Our study made the first step towards identifying locally adapted marginal populations of silver fir in the French Mediterranean Alps and, as such, has some implications for conservation genetics of this species. First, as adaptive genetic differentiation seems to be present at short spatial scales along elevational gradients, there may be an added value to collect seeds from low- and high-altitude locations for afforestation purposes. Second, isolated Mediterranean mountains, where silver fir survived during the Holocene under contrasted climatic conditions, may be valuable seed sources for more northern parts of the distribution range (assisted gene flow *sensu* Aitken & Whitlock 2013), where silver fir is likely to experience similar conditions as in the Mediterranean today.

Acknowledgements

We thank Frédéric Jean, Norbert Turion and Olivier Gilg (UEFM, INRA Avignon) for help with the fieldwork, Klaus

Haslinger for providing an R code for the scPDSI calculations, Bertrand Servin for help with the interpretation of the FLK results and Marianne Correard (UEFM), Olivier Marloie and Florence Courdier (URFM, INRA Avignon) for providing climate data. The present work is part of the Ph.D. thesis of AR., funded by the German Federal Ministry of Education and Research (Grant 01LC0822A to BZ and SL). BF, SL, AR, SOM and BZ were supported by the ERAnet BiodivERsA project 'LinkTree' (ANR-08-Biodiversa-006-06). BF, SL, KC, SOM, GGV and BZ were supported by the ERAnet BiodivERsA project 'TipTree' (ANR-12-EBID-0003). GGV was also supported by a grant by the Italian MIUR project 'Biodiversitalia' (RBAP10A2T4).

References

Aitken SN, Whitlock MC (2013) Assisted gene flow to facilitate local adaptation to climate change. *Annual Review of Ecology, Evolution, and Systematics*, **44**, 367–388.

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.

Alberto FJ, Derory J, Boury C *et al.* (2013) Imprints of natural selection along environmental gradients in phenology-related genes of *Quercus petraea*. *Genetics*, **195**, 495–512.

Allen CD, Macalady AK, Chenchouni H *et al.* (2010) A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*, **259**, 660–684.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.

Aussenac G (2002) Ecology and ecophysiology of circum-Mediterranean firs in the context of climate change. *Annals of Forest Science*, **59**, 823–832.

Barker MS, Dlugosch KM, Dinh L *et al.* (2010) EvoPipes.net: bioinformatic tools for ecological and evolutionary genomics. *Evolutionary Bioinformatics Online*, **6**, 143–149.

Beaumont MA, Nichols RA (1996) Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society of London Series B: Biological Sciences*, **263**, 1619–1626.

Bonhomme M, Chevalet C, Servin B *et al.* (2010) Detecting selection in population trees: the Lewontin and Krakauer test extended. *Genetics*, **186**, 241–262.

Cailleret M, Nourtier M, Amm A, Durand-Gillmann M, Davi H (2014) Drought-induced decline and mortality of silver fir differ among three sites in Southern France. *Annals of Forest Science*, **71**, 643–657.

Carlson SM, Cunningham CJ, Westley PAH (2014) Evolutionary rescue in a changing world. *Trends in Ecology & Evolution*, **29**, 521–530.

Chen J, Källman T, Ma X *et al.* (2012) Disentangling the roles of history and local selection in shaping clinal variation of allele frequencies and gene expression in Norway spruce (*Picea abies*). *Genetics*, **191**, 865–881.

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

Crawford RMM (2008) *Plants at the Margin Ecological Limits and Climate Change*. Cambridge University Press, Cambridge, New York.

Csilléry K, Lalagüe H, Vendramin GG *et al.* (2014) Detecting short spatial scale local adaptation and epistatic selection in climate-related candidate genes in European beech (*Fagus sylvatica*) populations. *Molecular Ecology*, **23**, 4696–4708.

Davi H, Gillmann M, Ibanez T *et al.* (2011) Diversity of leaf unfolding dynamics among tree species: new insights from a study along an altitudinal gradient. *Agricultural and Forest Meteorology*, **151**, 1504–1513.

De Carvalho D, Ingvarsson PK, Joseph J *et al.* (2010) Admixture facilitates adaptation from standing variation in the European aspen (Populus tremula L.), a widespread forest tree. *Molecular Ecology*, **19**, 1638–1650.

De Kort H, Vandepitte K, Honnay O (2012) A meta-analysis of the effects of plant traits and geographical scale on the magnitude of adaptive differentiation as measured by the difference between Q_{ST} and F_{ST}. *Evolutionary Ecology*, **27**, 1081–1097.

De Kort H, Vandepitte K, Bruun HH *et al.* (2014) Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree species *Alnus glutinosa*. *Molecular Ecology*, **23**, 4709–4721.

De Mita S, Thuillet A-C, Gay L *et al.* (2013) Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, **22**, 1383–1399.

Devlin B, Roeder K (1999) Genomic control for association studies. *Biometrics*, **55**, 997–1004.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Ekblom R, Galindo J (2011) Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity*, **107**, 1–15.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

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

Fady B, Forest I, Hochu I *et al.* (1999) Genetic differentiation in Abies alba Mill. populations from Southeastern France. *Forest Genetics*, **6**, 129–138.

Feurdean A, Bhagwat SA, Willis KJ *et al.* (2013) Tree migration-rates: narrowing the gap between inferred post-glacial rates and projected rates. *PLoS One*, **8**, e71797.

Foll M, Gaggiotti O (2008) A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian Perspective. *Genetics*, **180**, 977–993.

Fournier-Level A, Korte A, Cooper MD *et al.* (2011) A map of local adaptation in *Arabidopsis thaliana*. *Science*, **334**, 86–89.

Frichot E, Schoville SD, Bouchard G, François O (2013) Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution*, **30**, 1687–1699.

Frichot E, Schoville SD, de Villemereuil P, Gaggiotti OE, François O (2015) Detecting adaptive evolution based on association with ecological gradients: orientation matters!. *Heredity*, **115**, 22–28.

Giorgi F, Lionello P (2008) Climate change projections for the Mediterranean region. *Global and Planetary Change*, **63**, 90–104.

Goslee SC, Urban DL (2007) The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, **22**, 1–19.

Hampe A, Jump AS (2011) Climate relicts: past, present, future. *Annual Review of Ecology, Evolution, and Systematics*, **42**, 313–333.

Hampe A, Petit RJ (2005) Conserving biodiversity under climate change: the rear edge matters. *Ecology Letters*, **8**, 461–467.

Hanewinkel M, Cullmann DA, Schelhaas M-J, Nabuurs G-J, Zimmermann NE (2013) Climate change may cause severe loss in the economic value of European forest land. *Nature Climate Change*, **3**, 203–207.

Howe GT, Aitken SN, Neale DB *et al.* (2003) From genotype to phenotype: unraveling the complexities of cold adaptation in forest trees. *Canadian Journal of Botany*, **81**, 1247–1266.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Iverson LR, Peters MP, Matthews S, Prasad A (2013) An overview of some concepts, potentials, issues, and realities of assisted migration for climate change adaptation in forests. In: Browning J. *Comp. Proceedings of the 60th Annual Western International Forest Disease Work Conference*, 2012, pp. 8–12.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Jones P, Binns D, Chang H-Y *et al.* (2014) InterProScan 5: genome-scale protein function classification. *Bioinformatics*, **30**, 1236–1240.

Kawecki TJ (2008) Adaptation to marginal habitats. *Annual Review of Ecology, Evolution, and Systematics*, **39**, 321–342.

Kirkpatrick M, Barton NH (1997) Evolution of a species' range. *The American Naturalist*, **150**, 1–23.

Kremer A, Ronce O, Robledo-Arnuncio JJ *et al.* (2012) Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecology Letters*, **15**, 378–392.

Kreyling J (2010) Winter climate change: a critical factor for temperate vegetation performance. *Ecology*, **91**, 1939–1948.

Kreyling J, Thiel D, Nagy L *et al.* (2011) Late frost sensitivity of juvenile *Fagus sylvatica* L. differs between southern Germany and Bulgaria and depends on preceding air temperature. *European Journal of Forest Research*, **131**, 717–725.

Kreyling J, Wiesenberg GLB, Thiel D *et al.* (2012) Cold hardiness of *Pinus nigra* Arnold as influenced by geographic origin, warming, and extreme summer drought. *Environmental and Experimental Botany*, **78**, 99–108.

Kühl N, Gebhardt C, Litt T, Hense A (2002) Probability density functions as botanical-climatological transfer functions for climate reconstruction. *Quaternary Research*, **58**, 381–392.

Lander TA, Oddou-Muratorio S, Prouillet-Leplat H, Klein EK (2011) Reconstruction of a beech population bottleneck using archival demographic information and Bayesian analysis of genetic data. *Molecular Ecology*, **20**, 5182–5196.

Le Corre V, Kremer A (2012) The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*, **21**, 1548–1566.

Leinonen T, McCairns RJS, O'Hara RB, Merilä J (2013) Q_{ST}–F_{ST} comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nature Reviews Genetics*, **14**, 179–190.

Lens F, Tixier A, Cochard H *et al.* (2013) Embolism resistance as a key mechanism to understand adaptive plant strategies. *Current Opinion in Plant Biology*, **16**, 287–292.

Lepais O, Bacles CF (2014) Two are better than one: combining landscape genomics and common gardens for detecting local adaptation in forest trees. *Molecular Ecology*, **23**, 4671–4673.

Li N, Stephens M (2003) Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics*, **165**, 2213–2233.

Liepelt S, Cheddadi R, de Beaulieu JL *et al.* (2009) Postglacial range expansion and its genetic imprints in *Abies alba* (Mill.) – a synthesis from palaeobotanic and genetic data. *Review of Palaeobotany and Palynology*, **153**, 139–149.

Linares JC, Camarero JJ (2012) Growth patterns and sensitivity to climate predict silver fir decline in the Spanish Pyrenees. *European Journal of Forest Research*, **131**, 1001–1012.

Lotterhos KE, Whitlock MC (2014) Evaluation of demographic history and neutral parameterization on the performance of F_{ST} outlier tests. *Molecular Ecology*, **23**, 2178–2192.

Lotterhos KE, Whitlock MC (2015) The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, **24**, 1031–1046.

Luikart G, England PR, Tallmon D, Jordan S, Taberlet P (2003) The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, **4**, 981–994.

Martínez-Vilalta J, Lloret F, Breshears DD (2011) Drought-induced forest decline: causes, scope and implications. *Biology Letters*, **8**, 689–691.

Mimura M, Aitken SN (2010) Local adaptation at the range peripheries of Sitka spruce. *Journal of Evolutionary Biology*, **23**, 249–258.

Mosca E, Eckert AJ, Di Pierro EA *et al.* (2012a) The geographical and environmental determinants of genetic diversity for four alpine conifers of the European Alps. *Molecular Ecology*, **21**, 5530–5545.

Mosca E, Eckert AJ, Liechty JD *et al.* (2012b) Contrasting patterns of nucleotide diversity for four conifers of Alpine European forests. *Evolutionary Applications*, **5**, 762–775.

Mosca E, González-Martínez SC, Neale DB (2013) Environmental versus geographical determinants of genetic structure in two subalpine conifers. *New Phytologist*, **201**, 180–192.

Narum SR, Hess JE (2011) Comparison of F_{ST} outlier tests for SNP loci under selection. *Molecular Ecology Resources*, **11**, 184–194.

Nathan R, Horvitz N, He Y *et al.* (2011) Spread of North American wind-dispersed trees in future environments. *Ecology Letters*, **14**, 211–219.

Neale DB, Kremer A (2011) Forest tree genomics: growing resources and applications. *Nature Reviews Genetics*, **12**, 111–122.

Nystedt B, Street NR, Wetterbom A *et al.* (2013) The Norway spruce genome sequence and conifer genome evolution. *Nature*, **497**, 579–584.

Ozenda P (1975) Sur les étages de végétation dans les montagnes du bassin méditerranéen *Documents de cartographie écologique*, 1–32.

Palmer WC (1965) *Meteorological drought, Tech. Rep*. U.S. Dep. of Commerce, Washington, DC, 45.

Pérez-Figueroa A, García-Pereira MJ, Saura M, Rolán-Alvarez E, Caballero A (2010) Comparing three different methods to detect selective loci using dominant markers. *Journal of Evolutionary Biology*, **23**, 2267–2276.

Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 187–214.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

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.

R Development Core Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rehfeldt GE, Tchebakova NM, Parfenova YI *et al.* (2002) Intraspecific responses to climate in *Pinus sylvestris*. *Global Change Biology*, **8**, 912–929.

Restoux G, Silva DE, Sagnard F *et al.* (2008) Life at the margin: the mating system of Mediterranean conifers. *Web Ecology*, **8**, 94–102.

Roschanski AM, Fady B, Ziegenhagen B, Liepelt S (2013) Annotation and re-sequencing of genes from *de novo* transcriptome assembly of *Abies alba* (Pinaceae). *Applications in Plant Sciences*, **1**, 1–8.

Sagnard F, Barberot C, Fady B (2002) Structure of genetic diversity in *Abies alba* Mill. from southwestern Alps: multivariate analysis of adaptive and non-adaptive traits for conservation in France. *Forest Ecology and Management*, **157**, 175–189.

Savolainen O, Pyhäjärvi T, Knürr T (2007) Gene flow and local adaptation in trees. *Annual Review of Ecology, Evolution, and Systematics*, **38**, 595–619.

Savolainen O, Lascoux M, Merilä J (2013) Ecological genomics of local adaptation. *Nature Reviews Genetics*, **14**, 807–820.

Schoville SD, Bonin A, François O *et al.* (2012) Adaptive genetic variation on the landscape: methods and cases. *Annual Review of Ecology, Evolution, and Systematics*, **43**, 23–43.

Sengupta S, Majumder AL (2014) Physiological and genomic basis of mechanical-functional trade-off in plant vasculature. *Frontiers in Plant Science*, **5**, 244.

Sexton JP, McIntyre PJ, Angert AL, Rice KJ (2009) Evolution and ecology of species range limits. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 415–436.

Slater GS, Birney E (2005) Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics*, **6**, 31.

Sork VL, Aitken SN, Dyer RJ *et al.* (2013) Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes*, **9**, 901–911.

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, **73**, 1162–1169.

Thiel D, Kreyling J, Backhaus S *et al.* (2014) Different reactions of central and marginal provenances of *Fagus sylvatica* to experimental drought. *European Journal of Forest Research*, **133**, 247–260.

Tiffin P, Ross-Ibarra J (2014) Advances and limits of using population genetics to understand local adaptation. *Trends in Ecology & Evolution*, **29**, 673–680.

Tsumura Y, Uchiyama K, Moriguchi Y, Ueno S, Ihara-Ujino T (2012) Genome scanning for detecting adaptive genes along environmental gradients in the Japanese conifer, *Cryptomeria japonica*. *Heredity*, **109**, 349–360.

Tzedakis PC, Emerson BC, Hewitt GM (2013) Cryptic or mystic? Glacial tree refugia in northern Europe. *Trends in Ecology & Evolution*, **28**, 696–704.

de Villemereuil P, Frichot É, Bazin É, François O, Gaggiotti OE (2014) Genome scan methods against more complex models: when and how much should we trust them? *Molecular Ecology*, **23**, 2006–2019.

Vitasse Y, Delzon S, Bresson CC, Michalet R, Kremer A (2009) Altitudinal differentiation in growth and phenology among populations of temperate-zone tree species growing in a common garden. *Canadian Journal of Forest Research*, **39**, 1259–1269.

Vitti JJ, Grossman SR, Sabeti PC (2013) Detecting natural selection in genomic data. *Annual Review of Genetics*, **47**, 97–120.

Wachowiak W, Salmela MJ, Ennos RA, Iason G, Cavers S (2011) High genetic diversity at the extreme range edge: nucleotide variation at nuclear loci in Scots pine (Pinus sylvestris L.) in Scotland. *Heredity*, **106**, 775–787.

Warnes G, with contributions from Gorjanc G, Leisch F, Man M (2013) GENETICS: Population Genetics. R Package Version 1.3.8.1. Available from: http://CRAN.R-project.org/package=genetics.

Wells N, Goddard S, Hayes MJ (2004) A self-calibrating Palmer drought severity index. *Journal of Climate*, **17**, 2335–2351.

Whitlock MC (2008) Evolutionary inference from Q_{ST}. *Molecular Ecology*, **17**, 1885–1896.

Whitlock MC, Gilbert KJ (2012) Q_{ST} in a hierarchically structured population. *Molecular Ecology Resources*, **12**, 481–483.

Whitlock MC, Guillaume F (2009) Testing for spatially divergent selection: comparing Q_{ST} to F_{ST}. *Genetics*, **183**, 1055–1063.

Zhu K, Woodall CW, Clark JS (2012) Failure to migrate: lack of tree range expansion in response to climate change. *Global Change Biology*, **18**, 1042–1052.

B.F. designed the sampling, and A.R. and B.F. collected the samples. A.R. designed the genotyping array, S.L. performed the sequencing of the transcriptomes and contributed in the SNP genotyping, and K.U. annotated the SNPs. D.P. and G.G.V. provided genotypic data for the Romanian silver fir population. K.C. and F.H. performed the treatment and downscaling of the climatic data. A.R. and K.C. performed all other data analysis and wrote the manuscript. B.F. and S.O.M. contributed in the interpretation of the results, and all authors helped improving the manuscript.

Data accessibility

All genotypic, environmental and common garden data are available either in Dryad (https://datadryad.org/resource/doi:10.5061/dryad.t671s), as supplementary files, or in NCBI (accession no.: JV134525–JV157085 TSA NCBI).

Supporting information

Additional supporting information may be found in the online version of this article.

Fig. S1 Inter-annual variation across sampling plots.

Fig. S2 Pairwise correlations between all climatic variables for each sampling plot.

Fig. S3 Evaluation of optimal number of clusters for STRUCTURE analysis.

Fig. S4 Results of LFMM_env and LFMM_PCA analysis.

Fig. S5 Allele frequencies of significant associated SNPs.

Fig. S6 Most differentiated phenotypic traits determined in common garden.

Fig. S7 Sensitivity of the Q_{ST}–F_{ST} test to the out-crossing rate.

Table S1 On-site observations for daily temperature, precipitation, and relative humidity.

Table S2 Pairwise Reynold's genetic distance among all plots.

Table S3 Hierachical AMOVA analysis for all sites; Vtx and Ves; Lur and Iss.

Table S4 Synonymous and non-synonymous status of SNP loci.

Appendix S1 Site-specific environmental variables.

Appendix S2 Reconciling the sampling for SNP genotyping and the sampling for the common garden experiment (Sagnard *et al.* 2002).

Appendix S3 Testing the sensitivity of LFMM to plot-level versus individual level environmental variables.

Appendix

Data availability

All data are available at University of Marburg, Faculty of Biology, Conservation Biology, Karl-von-Frisch-Straße, 35032 Marburg, Germany (birgit.ziegenhagen@biologie.uni-marburg.de). Data media are (i) data CD submitted together with the thesis (processed data), (ii) hard drive at Conservation Biology group of the University of Marburg (raw data) and additionally, public resource databases for published data of chapter I and III.

Data encompasses

Hard drive:

- Raw output 454 sequencing
- Raw output Illumina sequencing
- Raw output KASP genotyping

Data CD:

- ChapterI_454AllContigs_parsed_as_database; Folder
- ChapterI_primer_sequences.xlsx
- ChapterI&II_annotation_protocol_output.xlsx
- ChapterI&II_3rd_level_annotation.pl
- ChapterII_Illumina_sample_lables.xlsx
- ChapterII_Illumina_assembly_metrics.xlsx
- ChapterII_SNP_calls_table.tsv
- ChapterII_KASP_primers.xlsx
- ChapterIII_Supplement.pdf

Public resource databases:

- 454 all assembled contigs (NCBI TSA accession no.: JV134525–JV157085)
- All genotypic, environmental and common garden data of chapter III (https://datadryad.org/resource/doi:10.5061/dryad.t671s and as supplementary files of the publication)

Curriculum Vitae

Personal information

Name	Anna Maria Roschanski
Affiliation	IPK-Gatersleben Satallite Collections North| Inselstrasse 9 | D-23999 Malchow Insel Poel roschanski@ipk-gatersleben.de

Education

2004 – 2010	Ernst-Moritz-Arndt University Greifswald Department of Botany and Landscape Ecology
2010 – 2014	Philipps-University Marburg Faculty of Biology - Conservation Biology Phd-Thesis: Adaptive genetic structure in ecologically marginal populations of European Silver Fir (*Abies alba* MILL) at the south-western Mediterranean pre-Alps of France.
2015	IPK Gatersleben, Dept. Genebank (Satellite collections north) Project: GrassLandscape - Landscape genomics and quantitative genetics in *Lolium perenne*

Publications

2010	*Diploma Thesis (EMAU Greifswald*) **Geschlechterverhältnis und klonale Populationsstruktur bei Mercurialis perennis L** A.M. Roschanski
2011	*J Hered (2011) 102(4): 479-488 doi:10.1093* **Characterization of Microsatellite Loci and Reliable Genotyping in a Polyploid Plant, *Mercurialis perennis* (Euphorbiaceae)** T. Pfeiffer, A. M. Roschanski, J.R. Pannell, G. Korbecka, M. Schnittler
2013	*APPS (2013) 1 (1) doi 1200179* **Annotation and re-sequencing of genes from de novo transcriptome assembly of *Abies alba* (Pinaceae)** A.M. Roschanski, B. Fady, B. Ziegenhagen, S. Liepelt
2013	*PMBR (2014) 32:750–760 doi 10.1007* **Transcriptome versus genomic microsatellite markers: highly informative multiplexes for genotyping *Abies alba* Mill. and congeneric species** D. Postolache, C. Leonarduzzi, A. Piotti, I. Spanu, A. Roig, B. Fady, A.M. Roschanski, S. Liepelt, G.G. Vendramin
2015	*MEC-15-1017 doi 10.1111* **Evidence of divergent selection for drought and cold tolerance at landscape and local scales in *Abies alba* Mill. in the French Mediterranean Alps** A.M. Roschanski, K. Csilléry, S. Liepelt, S. Oddou-Muratorio, B. Ziegenhagen, F. Huard, K.K. Ullrich, D. Postolache, G.G. Vendramin, B. Fady

Erklärung gemäß § 10 Abs (1) b und c der Promotionsordnung der mathematisch-naturwissenschaftlichen Fachbereiche und des medizinischen Fachbereichs für seine mathematisch-naturwissenschaftlichen Fächer der Philipps Universität Marburg vom 15.07.2009

Ich erkläre, dass von mir zu keinem früheren Zeitpunkt eine Promotion versucht wurde. Weiterhin versichere ich, dass die hier vorgelegte Dissertation von mir selbst ohne fremde Hilfe verfasst wurde. Ich verwendete ausschließlich die in der Arbeit angegebenen Quellen und Hilfsmittel und kennzeichnete alle vollständig oder sinngemäß übernommenen Zitate als solche. Die vorgelegte Dissertation wurde in dieser Form noch bei keiner in-oder ausländischen Hochschule anlässlich eines Promotionsgesuches oder zu anderen Prüfungszwecken eingereicht.

Unterschrift	Ort, Datum

Acknowledgement

Sincere thanks to: Beppe Vendramin, Bertrand Servin, Birgit Ziegenhagen, Bruno Fady, Christina Mengel, David Behringer, Dragos Postolache, Elmer Utesch, Eva Friebertshäuser, Felicia Gherghel, Felix Staeps, Francois Lefevre, Frédéric Jean, Galiner Bauer, Gerhard Kost, Hadrian Lalagüe, Hans Lehmann, Heike Zimmermann, Hendrik Davi, Illona Leyer, Jorge Bozzi, Katalin Csilléry, Katharina Budde, Katharina Fettweis, Katrin Heer, Kristian Ullrich, Lars Opgenoord, Lisa Wolf, Marianne Correard, Marie Nourtier, Marie-Cluade Bouhedi, Marie-Joe Karam, Matthieu Lingrand, Michael Brauer, Natascha Hasenkamp, Norbert Touron, Oliver Gilg, Pauline Garnier-Géré, Ronald Bialozyt, Santiago Gonzáles-Martínez, Sascha Liepelt, SF BG Marburg, Stefan Rensing, Sylvie Oddou-Mouratorio, Thomas Källmann.

My position at the University of Marburg and the works for this thesis were funded by the German Federal Ministry of Education and Research (Grant 01LC0822A) within the framework of the ERAnet BiodivERsA project ‘LinkTree’ (ANR-08-Biodiversa-006-06).