THE UNIVERSITY
of EDINBURGH

University of Edinburgh

Doctoral Thesis

Molecular and Adaptive Variation in the Caledonian Pine, *Pinus sylvestris* (L.)

Author:
Kevin Donnelly

Supervisors:
Dr. Stephen Cavers
Dr. Joan Cottrell
Prof. Richard Ennos

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

in the

School of Biological Sciences

May 27, 2015

Declaration of Authorship

I, Kevin Donnelly, declare that this thesis titled, 'Molecular and Adaptive Variation in the Caledonian Pine, *Pinus sylvestris* (L.)' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: _____

Date: _____

"Sticking feathers up your butt does not make you a chicken"

Tyler Durden

Lay Summary

Different plant species are adapted to different environments, and some can thrive where others do not. For example, cacti may survive in the desert where water is scarce, whereas others may perform well in conditions that are cold, wet, or saline. However, these adaptive differences may be found within as well as between species, when individuals perform better on their native sites than those imported from elsewhere: this is termed 'local adaptation'. Understanding local adaptation has important implications for management and conservation, as it can inform our choice of seed source for planting at any given location. This thesis explores local adaptation in Scotland within the Scots Pine (*Pinus sylvestris*). Across Scotland, rainfall varies dramatically, and on the west coast precipitation can be in excess of double that received on the east. It is therefore of particular interest as to whether trees from wet or dry pinewoods show evidence of adaptation to these conditions.

The characteristics of saplings grown from the seed of a number of native pinewoods around Scotland were examined when grown together under identical conditions. Firstly, variability among pine needles was investigated. The majority of needle properties exhibited limited variation between pinewoods, however, the number of stomatal rows was found to decrease with rainfall. Stomata are the pores through which plants breathe, and through which water is lost. The reduction in stomatal rows with rainfall was perceived as a possible adaptation to conserve water in areas where there is less available. The carbon isotope composition of plant tissue can provide us with information about the behaviour of plant stomata over time, and the efficiency of water usage. A study of carbon isotopes in needle tissue, revealed a trend from west to east, suggesting that water use is more conservative on average among individuals from the drier eastern pinewoods. An experiment was also conducted in which trees were exposed to flooding for a period of almost a year. During that period a number of morphological and physiological responses were recorded. Flooding was found to impede growth in trees from all pinewoods, as well as delay the beginning of growth in spring. The efficiency of photosynthesis was also found to be reduced by flooding, but more so for trees originating from drier pinewoods; this would suggest some degree of adaptation to local conditions on the basis of flooding or water availability.

In addition to the study of local adaptation, modern sequencing methods were employed to search of differences in the DNA of Scots Pine and three closely related species from different areas of Europe. It is hoped that the differences uncovered using this strategy will be valuable in determining the historical movement of Scots Pine in Europe following the last glaciation.

UNIVERSITY OF EDINBURGH

Abstract

College of Science and Engineering
School of Biological Sciences

Doctor of Philosophy

Molecular and Adaptive Variation in the Caledonian Pine, *Pinus sylvestris* (L.)

by Kevin Donnelly

The remnants of the Caledonian Pine Forest represent the north western boundary of the Eurasian *Pinus sylvestris* (L.) distribution. Remnant populations occupy a diverse range of environments within Scotland, subject to a steep rainfall gradient, and previous investigations have found evidence of local adaptation. Additionally, studies of biochemical and molecular markers have indicated that Scotland's native pinewoods originated from more than one glacial refugium.

Whole-genome-shotgun (WGS) sequencing was employed for the discovery of mitochondrial (mt) variants that may provide further insight into the origins of *P. sylvestris* populations both in Scotland and mainland Europe. DNA extractions were performed on megagametophyte tissue from Scottish, Finnish, and Spanish populations. Three members of the closely related *P. mugo* species complex were also sequenced. Using similarity-based approach, 160kbp of putative mitochondrial sequence was recovered by comparison of de novo assembled contigs with the mtgenome of the gymnosperm Cycas taitungensis. In total, 16 novel variants were identified among samples, which may be used in future phylogeographic studies.

A study of needle characters was performed for eight native populations of *P. sylvestris* in an outdoor provenance/progeny trial of 192 saplings. A negative correlation was detected between longitude and the number of stomatal rows present on needle surfaces. It was posited that this may be an adaptive response to lower water availability in eastern pinewoods, possibly in conjunction with increasing altitude.

The west coast of Scotland is one of the wettest regions in Europe: western pinewoods may receive in excess of 3,000mm of rainfall in a year, compared with an average of 800mm eastern sites. To determine whether native pinewoods are differentially adapted to waterlogging, a glasshouse based provenance/progeny trial of 432 saplings from nine

native populations was undertaken, in which 50% were subject to a long-term waterlogging treatment, and the remainder used as a control. Two studies were then conducted. In the first, responses to the treatment were assessed in terms of phenological and growth traits. Bud flush was delayed in response to waterlogging, and growth was impeded relative to the control. Although population differences were observed, treatment × population interactions were not detected. In the second study physiological traits known to be sensitive to plant stress and water balance were measured at intervals throughout the experiment. Prior to the commencement of the treatment needle δ^{13}C was found to exhibit interpopulation differentiation, and was positively correlated with longitude. This seems likely to represent differential selection for water use efficiency between eastern and western pinewoods. Photochemical efficiency and stomatal conductance were found to be reduced by waterlogging, and needle δ^{13}C was increased. After generalising populations into 'high' and 'low' rainfall groups (monthly averages of 214.9mm and 72.8mm, respectively), high rainfall populations were observed to maintain consistently higher photochemical efficiency under waterlogging the low rainfall populations. In addition, the low rainfall group exhibited greater variability in response to flooding (in terms of phenotypic and additive genetic variance) which may be indicative of a lack of past selection pressure.

Acknowledgements

This project was conducted at the CEH Edinburgh, and at the University of Edinburgh, and was funded by NERC in conjunction with Forest Research. I'd firstly like to thank my supervisors, Stephen Cavers, Joan Cottrell, and Richard Ennos, for entrusting me with the project, and providing me with guidance ever since. Science is hard, but a little encouragement goes a long way. You've afforded me more far more than a little, and for that I am very grateful.

I'd like to thank those at CEH Edinburgh, in particular Annika Perry, for her hard work and assistance in maintaining the trials, and Witold Wachowiak, for his efforts in testing the mitochondrial variants. I wouldn't have gotten far without any trees, and so a big thanks to the those who collected the seed: Dave Sim, Joan Beaton, and Ben Moore at the Macaulay Institute in Aberdeen. I'm also grateful to Mike Perks and Colin McEvoy at Forest Research Edinburgh, for their advice, and for kindly allowing me to borrow various interesting technologies. Thanks to Helen Grant at CEH Lancaster, who processed our carbon isotope samples, and a mention also goes to Bob Fleming at the Roslin Institute for allowing us the benefit of his time and expertise when we experimented with fluorescent probes.

I'd also like to express my gratitude to Beppe Vendramin at the Institute of Plant Genetics in Firenze who supported the mitochondria project, and to Sara Pinosio at the Institute of Applied Genomics in Udine, who carried out the sequencing work. I was very fortunate to receive a grant from the Procogen project, which allowed me to visit Umeå University in Sweden. Many thanks to Pär Ingvarsson for hosting me at his lab, and to Doug Scofield, who shared his experience with me during my brief time there. Additionally, I express my appreciation to Adam Butler, who showed great patience and resilience when subjected my relentless salvo of statistical questions.

I've been privileged to meet many totally excellent people during my time here in Edinburgh, firstly during my masters, and then during my PhD. I've made friends whom intend to keep for life if they'll let me, and I've felt very at home here (despite hailing from Glasgow). In particular, for diversions intellectual and otherwise (mainly otherwise), I should like to thank Richard Allen, Lucy Carter, Gytis Dudas, Hannah Froy, Cheryl Gibbons, Emma Hodcroft, Sam Lewis, Maarit Mäenpää, Sarah Mattey, Reuben Nowell, Darren Parker, Manon Ragonnet, and Nicolas Valeyrie.

I wish to thank my family, and especially my mother who has always supported me in my decisions, and to whom I greatly indebted. I'd never have gotten anywhere without you.

Finally, I'd like to thank Bryony, who over the last couple of years has had to put up with me more than any other. Thanks to you, in spite of the innumerable reasons there would always be to stay at work, there was always a very good reason to come home.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

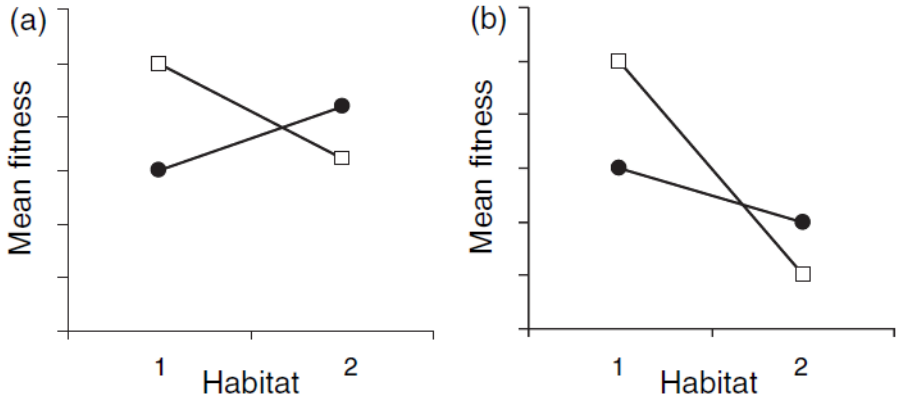
1.1 An Introduction to Local Adaptation

Species distributions are finite: even the most versatile organisms are constrained to boundaries beyond which they are unable to expand, either because they lack the requisite adaptations to survive in those environments, or are outcompeted by other species. Survival of species within their respective niches is made possible by distinct suites of adaptive traits. However, this environmental specificity occurs not only between, but also within species. Spatial variation is ubiquitous, and the natural distribution of a species may encompass a diverse range of environments, characterised by, *inter alia*, differing climates, levels of competition, predation, or pathogen prevalence. Conditions ordinarily experienced by conspecific individuals in one location may therefore be comparatively novel to those from another. When selection for differing traits is sufficiently strong relative to the gene flow between populations, local adaptation can arise. Local adaptation is widespread in nature, having been described in plants (Turner et al., 2010), vertebrates (Berven, 1982), insects (Mopper and Strauss, 1998), and bacteria (Belotte et al., 2003); and has important consequences for conservation practice, resource production, and the resilience of species to climate change.

Local adaptation can be said to have occurred when individuals exhibit greater Darwinian fitness on average at home sites than those which are introduced (Kawecki and Ebert, 2004)(fig. 1.1). Instances where members of a given population exhibit superior fitness at both their home sites and those of others, would not be recognised as local adaptation under this definition, as the disparity in fitness between populations could be attributed to causes other than differential selection. The existence of a genotype $\times$ environment interaction is therefore a fundamental requirement for local adaptation, such that no one genotype is optimal across all environments. The extent of adaptive

divergence will then be determined primarily by the extent of gene flow between populations, the magnitude of differential selection pressures, and the availability of genetic variation upon which selection may act.

Figure 1.1: Local adaptation schematic. Here, the population native to habitat 1 is represented by a square, and habitat 2, a circle. In (a) populations exhibit greater mean fitness on home sites; in (b) mean fitness is greater for both populations on habitat 1, but relative fitness is greater at home. Either scenario would be regarded as local adaptation. (Adapted from Kawecki and Ebert, 2004)



1.1.1 Prerequisites and Constraints

1.1.1.1 Genotype × Environment Interactions

Traits which enhance fitness in one environment may be disadvantageous in another: for example, some physiological adaptations come at a cost rendering them disadvantageous when not needed; trade-offs may also derive from pleiotropic effects or linkage of the genes under selection (Kawecki and Ebert, 2004). The circumstance in which an allele has conflicting effects on fitness in different environments is known as *antagonistic pleiotropy*, and provides a mechanism by which polymorphic loci can be maintained (Levene, 1953; Hedrick et al., 1976; Mitchell-Olds et al., 2007). An alternative, but similar scenario, whereby an allele is beneficial in one environment but selectively neutral in another, is termed *conditional neutrality* (Fry, 1996; Hall et al., 2010; Anderson et al., 2013). Without changes in the direction of selection, local adaptation may arise when alleles are subject to differing magnitudes of selection across environments (Kawecki and Ebert, 2004; Mitchell-Olds et al., 2007). Complex traits are determined by the action of genes at multiple loci: it is expected that alleles with large effects are more likely to contribute to local adaptation, as they are more likely to experience strong local selection and are less susceptible to loss via migration or drift (Yeaman and Whitlock, 2011).

When gene flow is limited, there is the potential for locally adapted gene complexes to develop. Conversely, an increase in population admixture acts to breakdown adaptive haplotypes via recombination, resulting in generalism rather than local adaptation (Anderson et al., 2013). Therefore while genotype × environment interactions are a necessary precursor for local adaptation, whether or not it is realised is strongly dependent on the extent of gene flow.

1.1.1.2 Gene Flow

Local adaptation can be described in terms of a balance between gene flow and selection: specifically, it occurs only if differential selective pressures are of sufficient strength to overcome the homogenising effects of gene flow (Kawecki and Ebert, 2004). The potential for adaptive differentiation is therefore anticipated to be greater among populations where dispersal is limited, and selection is strong (e.g. Brown and Pavlovic, 1992; Day, 2000). One obvious way in which gene flow between populations might be restricted, is when they occur further apart; this is 'isolation by distance' Wright (1943), and has been observed repeatedly in studies of local adaptation (e.g. Galloway and Fenster, 2000; Becker et al., 2006). The means of dispersal are, of course, an important factor in determining the spatial scale of the isolation by distance effect. Tree pollen, for example, has been reported to travel distances of almost 3000km (Campbell et al., 1999), and for many wind-pollinated plant species dispersal distances are highly variable and contingent on meteorological conditions. Gene flow can also be temporally restricted: should phenological differences in flowering exist (which may in themselves be a product of local adaptation), populations may experience a greater or lesser degree of reproductive isolation (e.g. Schuster et al., 1989).

As populations vary in size, and reproductive output (perhaps due to the varying quality of habitat), gene flow between them is liable to be asymmetric (Kawecki and Ebert, 2004). This is a particular issue for populations that exist on range margins where environments are likely to differ, but adaptation is constrained by continuous input of maladapted genes from the centre of the distribution (García-Ramos and Kirkpatrick, 1997). A consequence of this *migration load*, is that while some members of a metapopulation may be locally adapted, others may not be.

1.1.1.3 Sources of Variation

As with all adaptive evolution, local adaptation requires a source of heritable variation upon which to act. When environments are diverse, and gene flow is limited, local adaptation may still be constrained by small effective population sizes (Savolainen, 2011).

In cases such as these, increased gene flow may actually enhance adaptive potential, by supplementing the limited variation available as well as alleviating negative effects potentially caused by inbreeding depression (Kremer et al., 2012). Furthermore, fitness optima may shift over time; one predominant driver of this is ongoing climate change. It has been proposed that one way in which adaption to a changing environment may be accelerated, is by the flow of 'pre-adapted' genes from parts of the distribution where the conditions being introduced are not novel. This is often discussed in the context of increasing temperatures at higher latitudes, whereby gene flow from southern populations may be beneficial (e.g. Aitken et al., 2008).

Genetic variation observed between populations may not necessarily be the result of adaptive divergence, as differences may be the result of drift, migration, or population history (Kawecki and Ebert, 2004). Populations that exchange genes in the present, may have in the past been isolated from one another. The last glacial maximum (LGM) occurred some 20,000 years ago (Clark et al., 2009), during which time much of the contemporary range of temperate species would have been rendered inhospitable. The ancestors of present-day populations would then have survived in the remaining habitable refugia, potentially in the absence of gene flow. Glacial periods are much longer than the interglacial, and the mutations accumulated by populations in refugia may enable us to trace their origins in the present. Regardless of whether genetic structure has since been broken down, population history remains an important aspect of local adaptation, as the standing variation available for selection may be the product of contributions from multiple refugia (De Carvalho et al., 2010).

1.1.2 Studying Local Adaptation

1.1.2.1 Common Garden Experiments

The means by which local adaptation is investigated depends of course on the nature of the organism, and of the questions being asked. Phenotypic observations made on populations *in situ* may be of value for a variety of reasons, however, they cannot provide evidence for local adaptation as differences between sites could be attributable to phenotypic plasticity (although these studies may generate some interesting hypothesis to test). Common garden experiments are a mainstay of local adaptation studies, operating on the simple principle that when individuals are kept under identical conditions, environmental variation is removed, and mean differences in phenotype between populations should have a genetic basis. Reciprocal transplants, whereby common gardens are replicated at the home sites of each of the populations involved, are particularly effective as fitness can be assessed for each population in all environments including those of the

home sites. However, if the researcher is interested in adaptation to a particular aspect of the environment (e.g. temperature, salinity), reciprocal transplants are unsuitable as individuals are exposed to the full complement of site-specific environmental effects simultaneously. For these purposes, lab-based trials are preferable as the parameter of interest can be manipulated whilst others are controlled; albeit with the caveat that some populations may by chance be inherently better adapted to lab conditions than others (Kawecki and Ebert, 2004).

1.1.2.2 Molecular Methods

A number of techniques have been developed to investigate the molecular basis of local adaptation (see Savolainen et al. (2013) for review). When pedigree information is available, a panel of genomic markers may be employed in conjunction with phenotypic data in order to determine the location of quantitative trait loci (QTL); however, many crosses may be required for this to be effective, and it could be prohibitive in species with long generation times (Jain and Minocha, 1999). At the population level, association mapping approaches may be adopted (Balding, 2006). These require a relatively dense set of markers, typically single nucleotide polymorphisms (SNPs), distributed throughout the genome: variants that display a correlation with phenotypic traits may be in linkage with loci responsible. The efficacy of this approach is dependent on the extent of linkage disequilibrium within study populations, which in turn influences the density of markers required to achieve accuracy. In the absence of phenotypic data it is still possible to draw inferences about the loci underlying locally adapted traits by surveying which locations throughout the genome exhibit increased genetic structure (i.e. the Wright fixation index, F_{ST}), and may therefore be subject to differential selection (Lewontin and Krakauer, 1973; De Mita et al., 2013). Each of the methods described requires a set of molecular markers; however, given the increasing availability of sequence data, it is becoming possible to apply these approaches to a growing number of species.

1.1.2.3 Why Study Local Adaptation?

Locally adapted populations are of significant interest to evolutionary biologists: they provide systems in which traits of adaptive importance can be identified, and predictions regarding the environmental drivers of trait variation can be tested (Kawecki and Ebert, 2004). In some cases, local adaptation may even be a precursor to sympatric speciation (Sobel et al., 2010). Understanding the nature and extent of adaptive divergence is of vital importance to the conservation of genetic resources, with a view to maintaining the capacity of species to adapt in the face of climate change. Similarly, this knowledge

can be applied to the restoration of natural populations, in order to alleviate inbreeding without introducing maladapted genotypes (McKay et al., 2005). Furthermore, research into local adaptation may provide economic benefits, and can, for example, be employed to maximise crop yield by matching genotypes and environments (Takeda and Matsuoka, 2008).

1.2 Scots Pine: Adaptation in the Highlands of Scotland

The Scots pine, *Pinus sylvestris* (L.), is the most widely distributed of all pine species (Savolainen et al., 2007). Its range encompasses many degrees of latitude and longitude, extending from within the Arctic Circle in Scandinavia, southward to the mountainous regions of Turkey and Spain, and from the British Isles at the western extreme, through Siberia and ultimately into eastern Asia. Throughout this distribution the species encounters a diverse range of environments, which exhibit marked variation in temperature, moisture availability, and edaphic conditions. Numerous transplant experiments have demonstrated that this versatility cannot be attributed solely to phenotypic plasticity (e.g. Rehfeldt et al., 2002), and that populations have undergone varying degrees of adaptive divergence.

In Scotland, the remnants of the Caledonian Pine Forest represent the north western boundary of the species distribution, separated by a distance of least 500km from populations in mainland Europe (Wachowiak et al., 2011). Once widespread across the Scottish highlands, the forest has been subject to severe decline over the past 5000 years, and the fragmented stands that remain are estimated to cover only 1% of the area once occupied (McVean and Ratcliffe, 1962). Contemporary interest in the pinewoods was stimulated by Steven and Carlisle's 1959 book, *The Native Pinewoods of Scotland*, in which they identify and describe 35 sites understood to be relicts of the original forest. Since then, that number has been expanded to 84, covering an area of around 17,900 ha (Mason et al., 2004). In addition to their ecological importance, the pinewood remnants are recognised as nationally iconic, and are of significant cultural value. Continued research into the population history and adaptive characteristics of the remnants is not only of biological interest, but will be beneficial to their future management and conservation Salmela et al. (2010).

1.2.1 A Brief History of the Caledonian Pine Forest

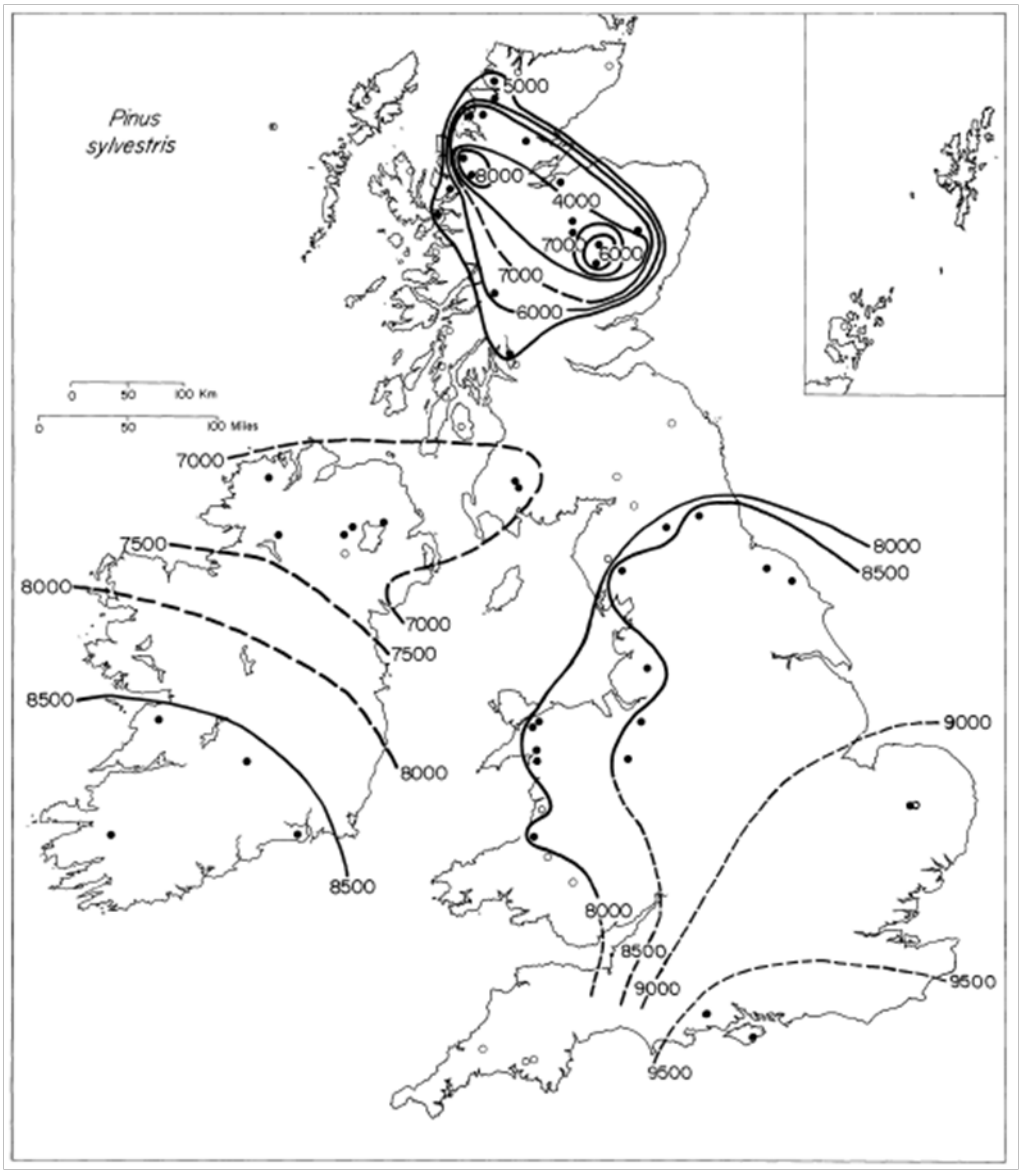
During the last glacial maximum, between 18,000–25,000 BP (years before present) (Clark et al., 2009), much of northern Europe was rendered inhospitable, and the distributions of many species were confined to refugia at lower latitudes (Willis and van Andel, 2004; Pyhäjärvi et al., 2008). *Pinus sylvestris* is believed to have begun post-glacial colonisation of what are today the British Isles by around 10,000 BP, at which time the earliest radiocarbon-dated remains appear in the south east of England (Bennett, 1984). Notably, remains in southern and western Ireland predate expansion of the mainland distribution into Wales (Craig, 1978), and may therefore be of independent origin (Birks, 1989). By 8,000 BP Scots pine was present in the north of England, but also the northwest of Scotland (Birks, 1972), despite remaining absent from the rest of the Scottish highlands.

This early occurrence of pine in the northwest of Scotland could be attributed to long-distance dispersal from the south, or alternatively, colonisation from a second refugium located in the west (Huntley and Birks, 1983a). Conceivably, these populations may have originated from the Irish distribution (Forrest, 1982; Bennett, 1984), although it has also been posited that some individuals may have survived glaciation more locally within small habitable areas (Huntley and Birks, 1983b; Kinloch et al., 1986), often referred to as 'cryptic refugia'. These hypotheses are difficult either to confirm or entirely refute: no native pinewoods are thought to remain in Ireland to permit genetic comparison, while the detection of highly localised refugia by means of fossil evidence presents an enormous challenge. However, the former does at least seem possible in view of the fact that a number of plant species present today in the south west of Scotland did indeed arrive via Ireland rather than England (Birks, 1989).

The expansion of pine in Scotland continued northward and eastward, becoming established in Upper Speyside between 8,000–7,500 BP, the Cairngorms by 6,000 BP (Dubois and Ferguson, 1985), and reaching its maximum extent at around 5,000 BP when it became present in the northernmost regions of the mainland (Birks, 1989).

By around 4,000 years ago Scots pine had begun to decline and had disappeared from the majority of the British Isles, with the exception of the Scottish highlands following the arrival of deciduous tree species (such as the Alder, *Alnus glutinosa*) which were able to outcompete *P. sylvestris* on favourable sites (Bennett, 1984). The reasons for the decline of the Caledonian Pine Forest are not wholly understood, however, it is thought to have occurred primarily via the action of regional climatic change, as many areas which once hosted Scots pine gave way to peat bogs. Ongoing soil deterioration, or an increase in pathogen prevalence, may also have contributed to the decline (Bennett, 1984). Human

Figure 1.2: An isochrone map based upon radiocarbon-dating of *P. sylvestris* fossil pollen. The sites sampled are represented by circles (pollen evidence was found only those which are filled), and numeric values indicate radiocarbon years BP (Taken from Birks, 1989)



impacts on natural populations are likely to have become significant following the introduction of agricultural practices, around 5,000 BP (i.e. 3,000 BC), whereby clearance of forest for the purpose of crops or grazing animals would have begun (Bennet, 1995). Exploitation of the pinewoods for farmland or for timber has continued throughout history, and destruction of the remaining forest has been considerable (Mason et al., 2004). Following the publication of Stephen & Carlisle's seminal book in 1959, there has been renewed interest in the Caledonian Pine Forest, and a number of initiatives developed with a view to protection and restoration of the remaining fragments.

1.3 Evidence for Multiple Refugia

1.3.1 Biochemical Markers

Prior to the widespread use of molecular markers, the concentrations of different monoterpenes (compounds found within plant resin), were commonly used as genetic markers in tree species (White et al., 2007). The relative concentrations of these monoterpenes are believed to be largely independent of the environment, and can be determined by gas chromatography. Forrest (1980) carried out a survey of almost 7,000 individuals from across the Scottish remnant distribution, to determine the extent of genetic structure between the remaining pinewoods. Each of the populations studied exhibited an unexpectedly large variability in monoterpene composition, indicative of a high level of genetic diversity. Nevertheless, it was possible to cluster populations into regions of biochemical similarity, of which six were delineated (these were used as the basis for seed zones, see below). The Wester Ross region was the most strongly differentiated of these groups, and of the sites within this region it was the westernmost, Shieldaig, that was the most distinctive.

Following on from this work, Kinloch et al. (1986) performed a second study of native populations, genotyping individuals at a number of monoterpene and allozyme loci. It was again remarked upon that in spite of severe range contraction, genetic variation was substantial; in fact, 'almost the highest of any plant species studied'. However, no meaningful geographic patterns could be discerned, with the exception of the Wester Ross populations which once again formed a distinct cluster. Shieldaig, whilst more similar in allele frequency to the other Wester Ross populations, was nonetheless at a remove from all others. Based on this, Kinloch et al. (1986) put forward a case for a second refugium, as it was unclear how two lineages may have spread north through England whilst remaining distinct.

1.3.1.1 Pinewood Management: Seed Zones

The patterns of variation identified by Forrest (1980) form the basis of the seed zones used in the management of Scots pine forests today (fig. 1.3). Under this policy, Scotland is divided into seven regions of biochemical similarity, two of which are defined as 'exclusion zones'. In order to qualify for grant aid, only material from approved native sources within these regions may be used in the regeneration of remnant pinewoods; any material planted for this purpose in the exclusion zones must originate from within that zone (Mason et al., 2004). As these regions have been designed around variation which

is not known to be of any adaptive significance, and do not reflect climatic variability, it has been recommended that a more robust and effective policy be developed based upon knowledge of adaptive divergence and gene flow between populations (Salmela et al., 2010).

Figure 1.3: Designated seed zones based upon the study of monoterpene variation by Forrest (1980). The north west and south west have been designated as exclusion zones, in which only material from the same zone can be used for pinewood regeneration.

1.3.2 Molecular Markers

1.3.2.1 In Scotland

The first direct molecular evidence in support of multiple refugia, came from a study of mitochondrial variation by Sinclair et al. (1998). Mitochondrial DNA (mtDNA) is subject to maternal inheritance in pines (Neale and Sederoff, 1988), and is therefore restricted to dispersal via seed. In contrast to nuclear or chloroplast markers, which are transmitted either partially or entirely via pollen, the spatial structure of mitochondrial variants is not readily broken down, making them well-suited for use in studies of populations history. Sinclair et al. (1998) discovered a restriction fragment length polymorphism (RFLP) at the *cox*1 locus, whereby the majority of native pinewoods sampled were fixed for 'mitotype a', with the exception of three west coast sites which were polymorphic for 'mitotype b' (fig. 1.4).

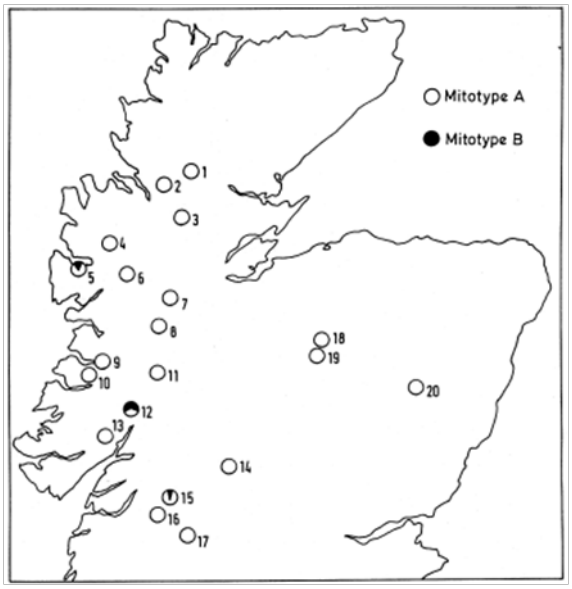


Figure 1.4: RFLP variation at mitochondrial locus *cox*3. The majority of populations were found to be fixed for mitotype a, with the exception a small number of west coast sites where mitotype b was present. (Taken from Sinclair et al. (1998))

A survey of paternally inherited chloroplast microsatellite variation was published shortly afterward by Provan et al. (1998), which included samples from both native Scottish sites and mainland Europe. In total, 13 polymorphic loci were identified, from which 174 distinct haplotypes were described within Scottish populations. Of these, 124 were unique: present in solitary individuals. By analysis of molecular variance (AMOVA), it was shown that almost 97% of variation in Scotland lay within pinewoods; furthermore, there was no significant difference between Scottish and mainland European populations. A rare allele resulting from a nine base-pair duplication was detected in 13 of the 330 Scottish individuals sampled, of which ten came from Wester Ross. The authors concluded that this was either a relatively recent mutation, or once again that this may be indicative of an alternative origin for west coast populations.

1.3.2.2 Across Europe

Several studies of post-glacial migration in *P. sylvestris* have been conducted in mainland Europe. Sinclair et al. (1999) extended sampling at the *cox*1 locus to incorporate a number of continental populations. Both of the previously identified RFLP variants were observed, in addition to a new variant present only in Spain (fig. 1.5). Although seemingly diverse, Spain was not considered to have contributed to the recolonisation of mainland Europe, given the strong differentiation from other European populations detected using nuclear markers (e.g. Prus-Glowacki and Stephan, 1994). Instead it was thought that the Spanish pinewoods were relicts that may have survived multiple glaciation events, and that mitochondrial differentiation may have occurred as populations are isolated from one another at high altitude.

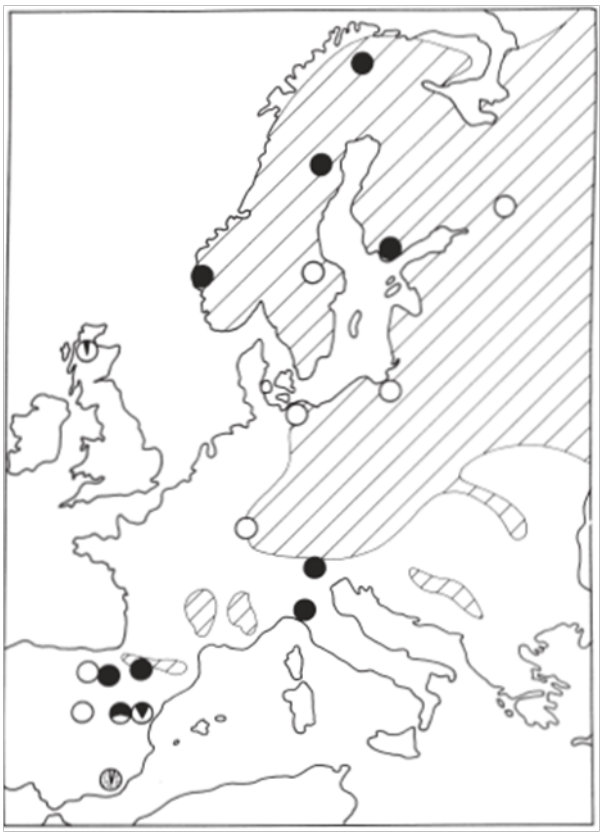


Figure 1.5: RFLP variation at mitochondrial locus *cox*3 in Europe. The greatest diversity was found in Spain, however, it was not believed to have contributed to the post-glacial colonisation of western Europe. (Taken from Sinclair et al., 1999)

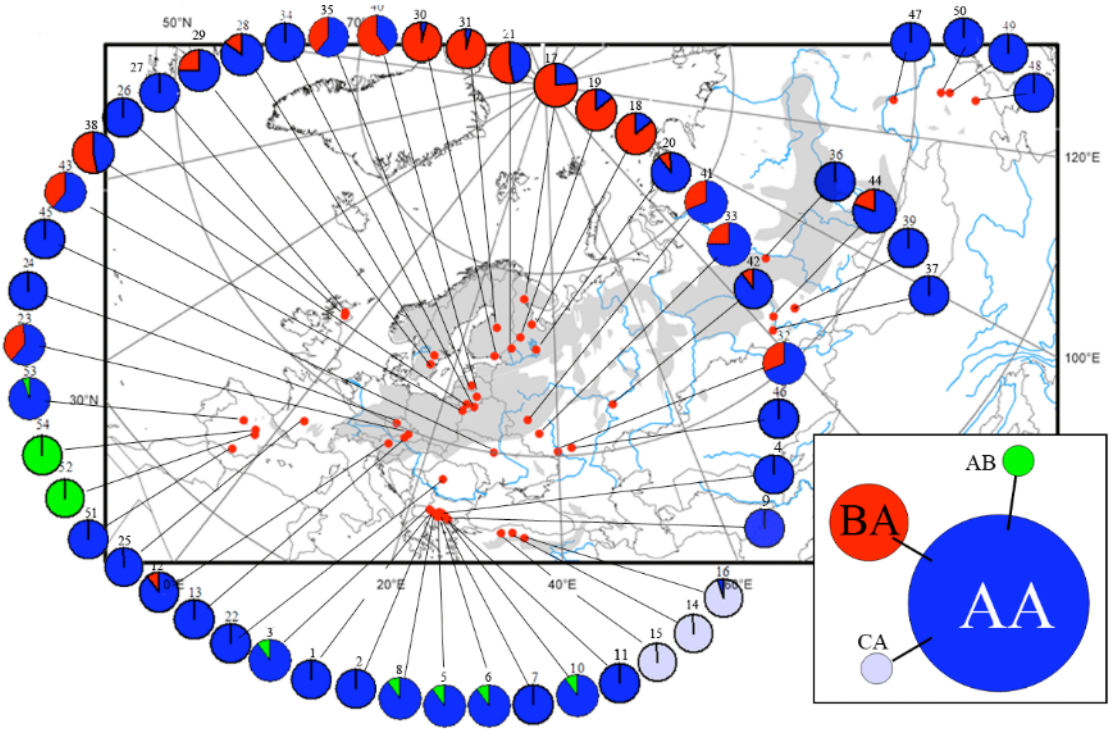
The first mitochondrial sequence variant was reported by Soranzo et al. (2000), who identified a 31bp insertion within an intron of the *nad*1 gene. This study was comprised predominantly of Spanish populations, however, a number of other European populations were also included, several of which were Scottish (including three from Wester Ross). Both alleles were found throughout Spain, but with the exception of a solitary Polish individual, all remaining populations were fixed for the allele which lacked the insertion. An additional sequence variant was later discovered at *nad*7, comprising three alleles: an ancestral sequence and two indel variants (Naydenov et al., 2007; Pyhäjärvi et al., 2008). The *nad*1 and *nad*7 loci were then screened in tandem to describe patterns of variation across Europe (fig. 1.6). These studies revealed a secondary front between two lineages in Scandinavia, and provided evidence that Scots pine may have survived in refugia further north than previously thought.

1.4 Are Pinewood Remnants Locally Adapted?

1.4.1 Environmental Variation

In order for local adaptation to occur, populations must be exposed to differential selection pressures. The Scottish pinewoods represent an outlier of a vast species range,

Figure 1.6: Distribution of mitotypes as determined by screening of *nad*1 and *nad*7 loci across Europe. Four multi-locus haplotypes were identified, designated: AA, AB, BA, and CA. (Adapted from Naydenov et al., 2007)



separated by at least 500km from the mainland distribution. The Caledonian remnants occupy a comparatively very limited geographic area; nevertheless it is comprised of a diverse range of environments which exhibit marked differences in terms of climate, altitude, and substratum.

Scotland is unusually warm given its latitude due to the influence of the Gulf Stream (Murphy et al., 2009), and the west coast in particular is characterised by a mild, oceanic climate. Mean temperatures become progressively cooler toward the east of the native pinewood distribution, which lies within the Cairngorm mountain range: the highest and coldest region in the British Isles. Pinewoods may be found at sea level, up to an altitude of around 500–520m; the few individuals that survive above the tree line tending to adopt a distorted, or 'krummholz' morphology (Grace, 1997). Although edaphic conditions are variable within pinewoods, sites on the west coast are more often characterised by fragmented stands growing upon a patchwork of suitably drained soils interspersed by peatland; whereas in the east, the water table is lower, and allows larger, more uniform forests to develop on freely draining mineral soils (Mason et al., 2004). Arguably, the most striking environmental cline is that of the rainfall gradient, which decreases sharply from west to east. The west coast of Scotland is one of the wettest regions in Europe, receiving upwards of 3,000mm each year, whereas the east receives

only ∼800mm on average.

The topography and meteorology of Scotland is such that it is difficult to separate temperature, altitudinal, or rainfall gradients which covary with one another and with longitude. This can pose something an obstacle for experimental design, as it is difficult to sample a transect of native populations which varies with respect to one environmental parameter, but not another. Future climate predictions anticipate that Scotland will become warmer, and that rainfall patterns will change, with summers potentially becoming drier, but winters wetter (Murphy et al., 2009).

1.4.2 Evidence of Local Adaptation

A great many studies have been carried out into the local adaptation of Scots pine throughout its distribution in mainland Europe (e.g. Beuker, 1994; Rehfeldt et al., 2002; Savolainen et al., 2004); those conducted on the remnants of the Caledonian Pine Forest represent a limited, but growing body of work.

In their book, Steven and Carlisle (1959) describe some of the morphological differences recorded between native pinewoods, however, as these observations were carried out *in situ*, they are unable to provide evidence of adaptive differentiation. Some early common garden trials established by the Forestry Commission in the 1920s demonstrated that material sourced from the continent typically performed poorly relative to Scottish seed (Lines and Mitchell, 1965), and therefore the native pinewoods appeared to be divergent at least at the European scale.

Perks and Mckay (1997) compared the performance of one-year old seedlings from four native provenances, a commercial British seedlot, and one Swedish population grown together in Scotland. Differences were found in growth, and in the timing and extent of frost hardening between the seedlot and the Swedish sample: however, variation within native provenances was limited, although individuals from one western population (Loch Maree), were reported to exhibit smaller stem and needle mass than the others. Later, Perks and Ennos (1999), reported on a seven year-old common garden trial consisting exclusively of native populations, and described significant differences in bud-burst, height, and diameter growth. As this trial incorporated a progeny design within populations, it was also possible to estimate genetic components. They concluded that substantial heritability existed for growth traits, and that populations should therefore be able to respond to selection.

More recently, further research has been carried out into phenological variation between provenances. In a trial conducted in parallel at two common gardens (one close to

Edinburgh, and the other near Aberdeen), timing of bud flush was assessed for 21 native populations from across Scotland (Salmela et al., 2013). Timing of bud flush was found to be correlated with altitude at site of origin (provenances from sites at low altitude flushing earlier on average), and furthermore, variability in flush date was also positively correlated with altitude. The latter was attributed to the greater temporal variability in climate experienced at higher altitudes, resulting in greater variability among individuals from those sites.

In the past common garden trials measurements have often been limited to morphological, phenological, and growth characters. Previously physiological measurements were infeasible due to the length of time required, however, indirect methods are now available that make it possible to measure adaptation for physiological traits. Stable isotope analysis is one such example: tissue samples (such as leaves) may collected from a large number of individuals, dried, and submitted for analysis via mass spectrometry. The carbon isotope ratio of leaf tissue provides indirect information on the relative rates of gas exchange and transpiration, which may be used to make inferences about plant water use efficiency (WUE) (Farquhar et al., 1989). We can use sensitive physiological measurements to determine whether exposure to defined environmental conditions impose stress on a genotype. Under local adaptation to the defined environmental conditions, we expect adapted genotypes to suffer less stress than non-adapted genotypes. One such method is chlorophyll fluorescence (Maxwell and Johnson, 2000), a surrogate for photosynthetic efficiency which is reduced when plants are under stress.

Using chlorophyll fluorescence, Salmela et al. (2011b) found that the photochemical efficiency of saplings from milder climates in Scotland was depressed relative to those originating from colder, high altitude sites when exposed to a harsh winter cold. Rapid physiological phenotyping techniques can make a valuable addition to common garden studies, especially when evaluating environmental stresses.

1.5 Thesis Objectives

The main objectives of the thesis are twofold. The first is to develop methods for studying mtDNA variation in Scots pine. The ultimate aim of this work is to obtain a more detailed understanding of the genetic structure of mtDNA variation within the Caledonian pine populations in order to infer more about the glacial history of Scots pine in Scotland and therefore understand the sources of genetic diversity that entered the population at its foundation. The second objective is to use a variety of measurements, both morphological and physiological, in the context of common environment trials, to

determine the extent and pattern of genetic variation for adaptive characters among Scots pine populations in Scotland.

Population history is of inherent importance to the patterns of genetic variability we observe in present day distributions. It can be a large determinant in the availability of variation on which selection can act, particularly as populations may have been subject to adaptive divergence in isolation prior to their reconciliation. Identification of distinct lineages may also inform management policy, as there is a strong incentive to preserve potentially unique sources of variability. In Scotland, no further mitochondrial variants have been detected since Sinclair et al. (1998), and in Europe since Naydenov et al. (2007). Undoubtedly, a large pool of variation which could be used to clarify the origins of Scots pine populations remains to be discovered. Previous studies have relied upon mtDNA gene sequences conserved between plant species which by their very nature are likely to show only low amounts of variation; by employing modern whole-genome-shotgun (WGS) sequencing, we can explore large regions of previously unobserved intergenic sequence, which is likely to exhibit more variation. In doing so, it may be possible to develop novel markers which can be used to provide greater insight into the origins of populations both in Scotland and throughout Europe. Thus in the first results chapter of the thesis WGS sequencing techniques are employed to develop new mitochondrial markers to be used for the future documentation of the mtDNA genetic structure of Scots pine in Scotland, and in the mainland European distribution.

Common garden trials are a long established and effective means by which to evaluate local adaptation among populations. I build upon existing evidence, by utilising further trials which are analysed in the context of spatial and climatic variation. As I have outlined, the major environmental gradient across Scotland is a rainfall gradient. Therefore I concentrate on looking at adaptation to differences in water relations by measuring relevant aspects of tree morphology, physiology and response of populations and families to differences in waterlogging, an environmental factor that is likely to vary from east to west across the Scottish distribution. In Scotland, the rainfall gradient is perhaps the definitive axis of climatic variability, and as such an emphasis is placed on traits which may be of adaptive significance to high or low rainfall regimes.

Steven and Carlisle (1959) make reference to differences in pine needles characters between populations, however these observations were made *in situ*. To date, there have been no investigations into the needle anatomy of native pinewoods carried out under controlled conditions. Pine needles are known to be responsive to the environmental conditions (e.g. Urbaniak et al., 2003), and may be subject to adaptive divergence. In the second results chapter an investigation is carried out into functional needle traits,

with a view to determining whether adaptive divergence has taken place, and if so what aspects of the environment might explain interpopulation differences.

The west coast of Scotland receives in excess of double the annual precipitation of the east, and western pinewoods are characterised by poorer drainage than their eastern counterparts (Mason et al., 2004). In the third results chapter an investigation into flood tolerance among native populations is undertaken, to determine whether trees from low rainfall sites are more susceptible to the stresses incurred by waterlogging than those from sites where rainfall is abundant. Conditions are manipulated within a glasshouse environment, and responses are measured in terms of phenology, growth, and maintenance of root mass.

In the fourth results chapter, I assess physiological changes in populations over time as a result of prolonged exposure to flooding, with respect to photochemical efficiency and stomatal control. Additionally, by means of carbon isotope analysis, an attempt is made to ascertain whether the water use efficiency of individuals in a common garden is determined by the conditions at their sites of origin. Finally, I test the hypothesis that populations exhibit greater heritable variation when exposed to novel conditions, due to the absence of past selection for adaptation to these environments.

Therefore the results chapters of the thesis are:

- Discovery of New Mitochondrial Variants for Use in Phylogeographic Studies via Whole-Genome-Shotgun Sequencing
- Variation in Needle Characters Among the Native Pinewoods of Scotland
- Phenological and Morphological Responses of Scotland's Native Pinewoods to Flooding
- Physiological Responses of Scotland's Native Pinewoods to Flooding

In the final chapter I consolidate this information, and discuss the implications of this research for Scotland's native pinewoods, and in the broader context.

Chapter 2

Discovery of new mitochondrial variants for use in phylogeographic studies via whole-genome-shotgun sequencing

2.1 Introduction

During the last glacial maximum (LGM) around 20,000 years ago, much of northern Europe was rendered inhospitable by the Fennoscandian ice sheet (Clark et al., 2009). During this time the distributions of many plant and animal populations were fragmented, and largely confined to habitable refugia at lower latitudes. In the Scots pine *Pinus sylvestris* (L.), evidence of glacial vicariance is based upon mitochondrial variation (e.g. Sinclair et al., 1999; Pyhäjärvi et al., 2008), which in conjunction with fossil evidence (e.g. Cheddadi et al., 2006), has been used in attempts to trace patterns of recolonisation following the retreat of the ice. Efforts to discern post-glacial migration routes, and identify the refugia of contemporary populations have however been constrained by the scarcity of mitochondrial DNA (mtDNA) markers available: in Europe all inferences from mtDNA have been based upon the discovery of a very limited pool of variants. By employing modern whole-genome-shotgun (WGS) sequencing techniques in a number of individuals from a range of locations, it should be possible to recover larger portions of the genome, and thereby increase the likelihood that novel variants are captured. The acquisition of additional markers may provide an opportunity to resolve the European phylogeography of *P. sylvestris*, and potentially hybrid subspecies, in greater detail.

In pines, the cytoplasmic genomes of the chloroplast and the mitochondrion are subject to uniparental inheritance (Neale and Sederoff, 1988). Chloroplast DNA (cpDNA) is

paternally inherited, and is therefore of limited use in studies of post-glacial migration (Provan et al., 1998) as wind-pollination facilitates gene flow over large distances, and genetic structure may be readily broken down. The suitability of mtDNA for use in phylogeography stems from its maternal mode of inheritance; as loci travel via seed (though some evidence of paternal leakage has been reported (Wagner et al., 1991)), dispersal of mitochondrial haplotypes is greatly restricted. Mitochondrial variation has been used to investigate patterns of migration in numerous *Pinus* species, including *P. ponderosa* (Johansen and Latta, 2003), *P. mariana* (Jaramillo-Correa et al., 2004), and *P. banksia* (Godbout et al., 2005); in each case haplotypes have been characterised in terms of minisatellites, indels, RFLPs, whilst point mutations have not been reported.

Studies of mitochondrial diversity within the European distribution of *P. sylvestris* began with the discovery of three RFLP variants (upon a fragment of *cox*3) throughout the Scottish native distribution (Sinclair et al., 1998), with the implication that despite their contemporary proximity, native Scottish populations contain descendants from different glacial refugia. Sampling at this locus was later extended to the mainland of western Europe (Sinclair et al., 1999) where Spain was identified as a hub of haplotypic diversity, though not thought to have contributed to the recolonisation of mainland Europe due to strong differentiation at nuclear loci (Prus-Glowacki and Stephan, 1994). Shortly afterward Soranzo et al. (2000) identified another RFLP variant, this time within an intron at *nad*1, which was attributed to a 31bp insertion. A further two indels were detected (5bp and 32bp) at *nad*7 intron 1: screening in tandem with the *nad*1 variant revealed further diversity in Scandinavia, where there was evidence for colonisation from more than one refugium (Naydenov et al., 2007; Pyhäjärvi et al., 2008); however, the authors of both manuscripts contended that a larger number of polymorphic loci would be required, in particular to resolve the origins of *P. sylvestris* in northern Europe, and to determine the extent of vicariance between populations from eastern Europe and Asia.

In contrast to their animal counterparts, the mitochondrial genomes (mtgenomes) of land plants are remarkably large: even the shortest recorded is in excess of ten times longer than that of a human mtgenome (Knoop, 2004). Land plant mtgenomes have been reported to range in length from 185kbp in liverwort (Oda et al., 1992), to over 2,750Kbp in melon (Rodríguez-Moreno et al., 2011); more recently Sloan et al. (2012) reported the mtgenome of the catchfly flower (*Silene conica*) to be 11.3Mbp in length, though as yet this represents an outlier. The base mutation rate of cpDNA is significantly lower than that of the nuclear genome, however rates of sequence evolution in mtDNA are exceptionally low (Wolfe et al., 1987). Drouin et al. (2008) found that, in gymnosperms, synonymous substitution rates of mitochondrial genes were half that of those in chloroplast, and a quarter that of nuclear genes (in angiosperms, nuclear rates

were 16 times higher than mitochondrial). Despite a low point mutation rate, recombinational activity is high and structural rearrangements such as indels, duplications, and inversions are common (Palmer and Herbon, 1988; Knoop, 2004). Gene transfer is known to have occurred from the mitochondrial to the nuclear genome (Timmis et al., 2004), however plant mtgenomes are also reported to have incorporated DNA from both the chloroplast and the nucleus (Knoop, 2004; Wang et al., 2007). Furthermore, in angiosperms, evidence of horizontal gene transfer (HGT) has been reported between the mtgenomes of distantly related species (Bergthorsson et al., 2003).

At present, there are no completed references available for either the nuclear or mitochondrial genomes of *P. sylvestris*. Taxonomically, the nearest complete mtgenome available is that of the gymnosperm *Cycas taitungensis* (Chaw et al., 2008), which is approximately 415kbp in length. Like those of other seed plants previously sequenced, the *C. taitungensis* mtgenome is non-compact: some 89.9% is comprised of non-coding sequence (introns, spacers, and pseudogenes), and relatively short (typically <2kbp) repeated sequences account for 15.1%. The small number of sequences available for *P. sylvestris* are for regions which encode genes (including introns), likely subject to a high degree of conservation between both species and individuals. There are therefore large quantities of intergenic sequence which have yet to be investigated, and which may yield useful variants.

Here, we used WGS sequencing to obtain large quantities of sequence data from the megagametophyte tissue of three individuals of *P. sylvestris* from native populations around Europe. We also extended the sample to contain one representative from each of three closely related species: *P. mugo*, *P. uliginosa*, and *P. uncinata*, all of which are members of the *P. mugo* species complex. *Pinus mugo* and *P. sylvestris* are known to form hybrids where distributions overlap (e.g. Wachowiak and Prus-Głowacki, 2008; Danusevičius et al., 2013), and it has been posited that *P. uliginosa*, and *P. uncinata* may have arisen from ancient hybridisation events (Lewandowski et al., 2000). Using a similarity-based approach, *de novo* assembled contigs were compared with published mtDNA sequences in order to recover previously unexplored regions of the mtgenome, with the objective of capturing novel variants which may be used to better resolve the phylogeography of both *P. sylvestris* and the *P. mugo* species complex throughout Europe.

2.2 Methods and Materials

2.2.1 Sample Material

Seed collections were made from a number of native woodlands across Europe (fig. 2.1, table 2.1); within these collections samples were comprised of multiple seeds per tree. *Pinus sylvestris* was represented by three samples: Glen Loy (Scotland), Punkaharju (Finland), and Valsaín (Spain). Geographically disparate sources were chosen as they were more likely to exhibit genetic differences. The other *Pinus* species, *P mugo*, *P uliginosa*, and *P uncinata*, were each included as single samples taken from Poland (Śląskie Kamienie, and the Węgliniec Reserve), and from Andorra respectively. All seed was stored at $\sim 4^{\circ}\mathsf{C}$ prior to germination.

2.2.2 Preparation and DNA Extraction

Five to six seeds per sample were placed in Petri dishes on damp tissue and allowed to germinate until the seed coats had visibly split (typically occurring within a week). Haploid megagametophyte tissue was then isolated with the aid of dissecting microscope, and DNA extractions were performed using a *Qiagen DNeasy Plant Kit* as per the manufacturer's instructions (Qiagen, Venlo, Netherlands). A single extraction therefore contained DNA from a bulked collection of megagametophyte tissue taken from the seed of a single parent tree.

2.2.3 Illumina Sequencing

Sequencing of DNA extracts was performed at the *Istituto di Genomica Applicata* (IGA) in Udine, Italy. DNA was randomly fragmented by sonication using a Bioruptor (Diagenode). Libraries were enriched by 12-cycle PCR reaction, and following electrophoresis of the PCR products on a 2% agarose gel, fragments in the approximate size range of 600bp were selected. A Genome Analyzer flowcell was prepared on the supplied cluster station and libraries were sequenced in multiplex on one lane of the Illumina HiSeq2000 platform according to the manufacturer's instructions. Images from the instrument were processed using the manufacturer's pipeline software to generate FASTQ sequence files.

2.2.4 Assembly

FASTQ files were first assessed using FastQC (Babraham Bioinformatics, Cambridge, UK) to examine per-base quality, and identify any anomalous patterns in base calling

Figure 2.1: Map of European locations from which samples were collected.

across read lengths. Using TRIMMOMATIC (Bolger et al., 2014), adapter sequences were removed, non-random G/C content was cropped from the ends of reads, and poor quality bases were trimmed; a sliding window approach was then employed to remove any reads where a substantial drop in base quality was detected.

Samples were *de novo* assembled using *fermi* (Li, 2012), an assembler based upon the FMD-index for large genomes. *fermi* was developed for use with Illumina short reads, and all assemblies were built using the default minimum overlap of 50bp.

Table 2.1: Coordinates of sampling locations, and sample abbreviations.

Species	Abbreviation	Location	Latitude	Longitude
	PSGLEN	Glen Loy, Scotland	56.91	−5.13
P. sylvestris	PSPUN	Finland (Punkaharju)	61.76	29.39
	PSVAL	Spain (Valsaín)	40.87	−4.04
P. mugo	MUGO	Poland (Śląskie Kamienie)	50.78	15.60
P. uliginosa	ULIG	Poland (Węgliniec Reserve)	51.28	15.24
P. uncinata	UNC	Andorra	42.55	1.61

2.2.5 Identification of Putative Genomic, cpDNA, and mtDNA Contigs

Assemblies were first filtered to include only contigs >500bp in length. Of the available samples, *P. sylvestris* Punkaharju contained the largest number of reads and was therefore chosen for comparison with published sequences.

NCBI's BLAST tool (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to perform nucleotide searches (blastn) of assembled contigs against a 140kbp region of *P. taeda* genomic DNA (Genbank: AC241304.1); the near-complete *P. sylvestris* cpgenome (Genbank: JN854158.1), which is ∼120kbp; the mtgenome of *C. taitungensis* (Genbank: AP009381.1); and all *P. sylvestris* mtDNA sequences >1kbp. In order to reduce the potential for genomic sequences to be included within the mitochondrial hits, searches against mtDNA were restricted to contigs >1kbp in length. In all cases only hits with an expect value (e-value) <0.0001 were recorded.

Further BLAST searches were then performed using the hits obtained above against all published nucleotide sequences on Genbank, and any contigs found to be superior matches to non-target genomes (e.g. putative mtDNA contigs bearing a greater degree of resemblance to genomic or cpDNA) were excluded. As a reference genome for *P. sylvestris* is available, chloroplast hits were filtered to include only contigs with at least 95% query cover (an exception was made for one contig which received 75% cover from *P. sylvestris*, as coverage was 100% for a number of other species). As a reference is unavailable for the mtgenome, potential mitochondrial contigs were filtered more permissively, retaining all which shared at least 500bp identity with a published plant mtDNA sequence. Summary statistics were produced for each of these contig subsets using Quast (Gurevich et al., 2013).

2.2.6 Read Mapping and Variant Detection

Having identified putative genomic, chloroplast, and mitochondrial sequences from the Punkaharju sample, reads were then mapped to a composite of these subsets using BWA (Li and Durbin, 2009); filtering and sorting of mapped reads was performed using SAMtools (Li et al., 2009). Coverage information was obtained on a per-contig basis using BEDtools (Quinlan and Hall, 2010).

Sequence from the Punkaharju sample was then used as a reference against which reads from other samples were mapped. Variant detection was performed with GATK (DePristo et al., 2011), applying a minimum phred-scaled quality threshold of 30. Variants with a read-depth <10 or >200bp were not recorded, in order to prevent inclusion of SNPs with poor support, or those with a high likelihood of being attributable to paralogues respectively. Read-mapped assemblies were visually inspected at variable loci using Tablet (Milne et al., 2013).

2.2.7 Resequencing of Polymorphic Loci

In order to verify potential variants, a pilot panel of six individuals was chosen for resequencing. Primers were designed to flank targeted polymorphic regions, and amplification was carried out via PCR at an annealing temperature of 50°C. Regions that were amplified successfully were submitted for sequencing via Sanger sequencing, and the resulting chromatograms were examined for evidence of variation. This work was carried out by Witold Wachowiak at the *Centre for Ecology and Hydrology*, Edinburgh.

2.3 Results

2.3.1 Assembly and Contig Identification

The number of paired-end reads obtained from each sample varied from 3.2×10^7 to 6.2×10^7, the greatest number being available for the Punkaharju sample. Prior to enforcing a minimum contig size of 500bp (which dramatically reduced overall assembly length), the Punkaharju assembly was also the largest at ∼109Mbp in length; the others ranging from ∼56–79Mbp. All of the following results are described following removal of short contigs. Assemblies were relatively AT-rich, with GC content ranging between 40–41%, the exception being the Valsaín sample at 49%.

Summary statistics for the contig subsets derived from the Punkaharju assembly are provided in table 2.2. BLAST hits versus the *P. taeda* genomic sample generally scored highly in terms of sequence identity and query cover, however of the 24 genomic contigs that produced a hit only five were >1kb in length, and of these query cover (i.e. the proportion of the contig identified as a match) was <500bp for all but two. Breadth of coverage for this short region of the *P. taeda* nuclear genome was therefore very limited: characterised by strong hits in terms of identity, but for a small number of relatively short contigs as represented by a low N50.

In contrast, the majority of chloroplast hits were >1kbp, and breadth of coverage was very high. Including forward and reverse contigs, and excluding any overlaps, 92% of the published *P. sylvestris* chloroplast genome was accounted for in the Punkaharju assembly.

Candidates for the mitochondrial subset were restricted to contigs >1kb in length, to reduce the likelihood of genomic sequences being erroneously included. Searches were made against the *C. taitungensis* mtgenome, and the available *P. sylvestris* mtDNA sequences; however, the latter did not yield any additional hits. Sequence length in the

Table 2.2: Summary statistics for *de novo* contig sets, based upon contigs >500bp from the *P. sylvestris* Punkaharju assembly (the mitochondrial subset contained only contigs >1kbp). Values are provided for the assembly before filtering, and for the contig subsets defined after BLAST searches against a ∼140kbp of *P. taeda* genomic sequence, the *P. sylvestris* chloroplast genome, and the *C. taitungensis* mitochondrial genome.

	Length (bp)		No Contigs		Largest			
	Total	>1000bp	Total	>1000bp	Contig (bp)	GC (%)	N50	L50
Assembly	5 354 571	3 341 982	4366	1308	20 658	40.71	1480	753
Genomic	20 265	8136	24	5	3547	38.01	759	8
Chloroplast	111 444	107 583	39	33	10 155	38.66	3756	9
Mitochondrial	159 725	159 725	46	46	20 658	47.27	4567	10

mitochondrial subset summed to ∼160kbp across 46 contigs (∼38% of the *C. taitungensis* mtgenome length), and average GC content was greater than that of the assembly as a whole at 47% compared with 41%.

2.3.2 Coverage Depth

Coverage depth was estimated for each of the contigs subsets derived from the Punkaharju assembly (fig. 2.2). The chloroplast subset received by far the greatest mean coverage at 889×, followed by the genomic and mitochondrial subsets at 228× and 61×, respectively. The coverage distribution for the genomic subset was markedly skewed, however, and median depth was substantially lower at 49×.

2.3.3 Candidate Mitochondrial Sequences

A list of putative mitochondrial contigs from the Punkaharju assembly can be found in table 2.4, together with the names of the GenBank sequences which provided the greatest query cover. Although contigs were allocated to the subset on the basis of their similarity to the *C. taitungensis* mtgenome, a number of those contigs received greater query cover from other plant species: including pines, *P. sylvestris*, *P. strobus*, and *P. monophylla*; conifers *Larix mastersiana*, *Picea smithiana*, and *Abies sachalinensis*; angiosperms *Phoenix dactylifera*, *Ricinus communis*, and *Tripsacum dactyloides*.

Query cover ranged from 6–100% (averaging at 53%), and a positive relationship was identified between contig GC content and the query cover received ($p<0.01$) (fig. 2.3). A minimum of 500bp similarity with the *Cycas* reference genome was required for inclusion in the subset, however, there was no stipulation that query cover be uninterrupted, and for 28 of the 46 contigs listed multiple hits were observed. Furthermore, in 50% of these

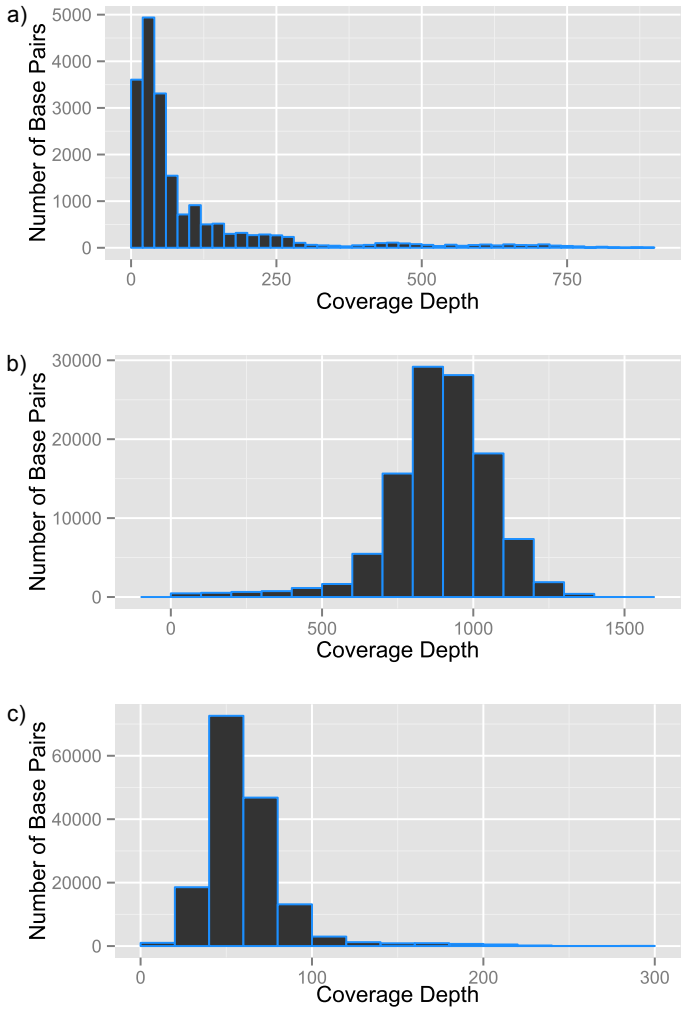


Figure 2.2: Coverage distributions after read-mapping to assembled contigs that were found to correspond to a) ~140kbp of *P. taeda* genomic sequence, b) the *P. sylvestris* chloroplast genome, and c) the *C. taitungensis* mitochondrial genome. Distributions are based upon reads and assembled contigs from the *P. sylvestris* Punkaharju sample in each case. Note that for c) reads were mapped only to contigs >1kbp in length, whereas a) and b) were generated using contigs >500bp.

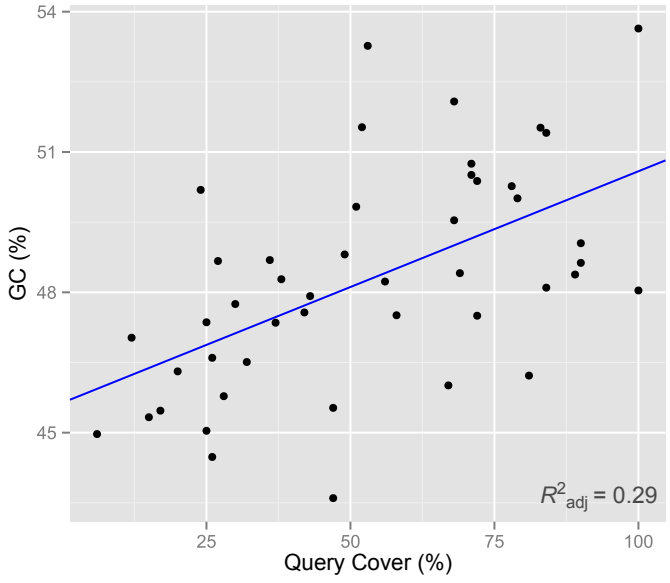


Figure 2.3: Within the mtDNA subset, contig GC content was found to increase with the proportion of each contig matched to published mtDNA ($p<0.01$), as determined via BLAST.

Table 2.3: Description of two possible indels identified on *mt27*, and *mt35* (on which a SNP (G↔T) is also visible in close proximity.)

Contig	Positions		Sample
mt27	1390–1400	TCGGGCCGCCA	Consensus
		TCG- - - - - CCA	PSVAL
mt35	138–161	GCC- - - - - - - - - - - - - - - - - - AAT	Consensus
		GCCCCTCTAAAGTAAGTAAAAAAG	PSPUN

instances hits occurred on both positive and negative strands, indicating the presence of inversions. In the small number of cases where homologous *P. sylvestris* sequences were available on Genbank, query cover was continuous, suggesting that the multiple-hit pattern observed was not due to misassembly.

2.3.4 Identification and Resequencing of Variants

Within-sample variation was observed at a number of putative SNP loci in the mt genome, whereby a proportion of reads contained an alternative allele, and the remainder were identical to the reference. SNP loci at which the proportion of reads containing the alternative allele exceeded 0.75 (in at least one sample) are described together with their current resequencing status in table 2.5. On this basis, 14 SNPs were identified in total across 11 contigs. All of the SNPs described were transversions, the most common being A↔C. In addition, two possible indels were detected: a 5bp deletion *mt27*, and an 18bp insertion on *mt35* (table 2.3), the latter of which occurred within close proximity to the SNP reported on that contig.

Following resequencing by Sanger, four SNPs were 'confirmed', in the sense that the pattern of variation observed was consistent with the previous Illumina results; two were unconfirmed, as sequencing was successful but loci were monomorphic; and four could not be evaluated as PCR amplification was unsuccessful. At time of writing, four SNPs remain to be tested. Of the indels, the 5bp deletion has been sequenced but not confirmed, and the 18bp is not yet tested.

2.4 Discussion

Much of our understanding of the post-glacial movements of European Scots pine is based upon RFLP variation at *cox*3 (Sinclair et al., 1998), and three indels within the mtgenome (Soranzo et al., 2000; Naydenov et al., 2007; Pyhäjärvi et al., 2008). The discovery of even a small number of novel markers could improve resolution, and yield new insights into the locations of glacial refugia and their contribution to the present day distribution. We employed WGS sequencing in order to recover larger regions of the mtgenome than previous studies, in an effort to capture new variants.

Table 2.4: List of candidate mitochondrial sequences taken from the Punkaharju assembly. Contigs are presented in order of increasing length, accompanied by their mean coverage depth after read-mapping, and query cover: the proportion of each contig identified as matching a published sequence as determined via BLAST. The percentage identity, species, and accession number are provided for each of these sequences, which were the best hits in terms of query cover.

Name	Length (bp)	GC (%)	Coverage	Query Cover (%)	Identity (%)	Species	Genbank Accession
mt1	1025	53.27	52.57	53	90	*C. taitungensis*	AP009381.1
mt2	1108	48.38	55.27	89	91	*C. taitungensis*	AP009381.1
mt3	1118	51.52	40.65	83	100	*P. sylvestris*	AJ223312.1
mt4	1139	53.64	180.02	100	100	*P. strobus*	AF058659.1
mt5	1176	48.04	111.13	100	94	*C. taitungensis*	AP009381.1
mt6	1183	48.10	58.97	84	100	*P. sylvestris*	KM244374.1
mt7	1238	51.53	48.21	52	87	*C. taitungensis*	AP009381.1
mt8	1374	43.60	52.82	47	92	*C. taitungensis*	AP009381.1
mt9	1476	50.27	77.58	78	91	*C. taitungensis*	AP009381.1
mt10	1588	52.08	48.22	68	86	*C. taitungensis*	AP009381.1
mt11	1594	50.75	51.94	71	99	*P. sylvestris*	AJ223312.1
mt12	1654	46.01	57.88	67	91	*C. taitungensis*	AP009381.1
mt13	1756	49.54	48.77	68	79	*C. taitungensis*	AP009381.1
mt14	1798	47.50	50.23	72	82	*C. taitungensis*	AP009381.1
mt15	1814	47.57	50.63	42	86	*C. taitungensis*	AP009381.1
mt16	1848	51.41	159.36	84	93	*T. dactyloides*	DQ984517.1
mt17	1855	48.41	66.15	69	87	*C. taitungensis*	AP009381.1
mt18	1869	47.51	64.05	58	100	*P. sylvestris*	KM244374.1
mt19	1872	47.92	48.45	43	94	*C. taitungensis*	AP009381.1
mt20	1935	48.23	50.41	56	88	*C. taitungensis*	AP009381.1
mt21	1958	46.22	99.42	81	96	*L. mastersiana*	JQ411206.1
mt22	1979	50.38	39.77	72	89	*C. taitungensis*	AP009381.1
mt23	2267	49.05	79.92	90	98	*P monophylla*	FJ824834.1
mt24	2316	47.75	44.29	30	91	*C. taitungensis*	AP009381.1
mt25	2360	48.69	75.36	36	85	*C. taitungensis*	AP009381.1
mt26	2431	45.78	57.39	28	90	*C. taitungensis*	AP009381.1
mt27	2494	48.28	50.57	38	82	*L. tulipifera*	KC821969.1
mt28	2742	50.51	63.56	71	84	*C. taitungensis*	AP009381.1
mt29	3298	44.48	55.77	26	86	*C. taitungensis*	AP009381.1
mt30	3319	49.83	73.56	51	85	*C. taitungensis*	AP009381.1
mt31	3348	48.63	74.79	90	99	*P. smithiana*	FJ824838.1
mt32	3507	50.19	50.10	24	81	*C. taitungensis*	AP009381.1
mt33	3509	46.51	58.15	32	93	*C. taitungensis*	AP009381.1
mt34	3673	50.01	55.93	79	92	*C. taitungensis*	AP009381.1
mt35	3875	47.35	51.36	37	90	*C. taitungensis*	AP009381.1
mt36	4044	48.81	50.68	49	88	*C. taitungensis*	AP009381.1
mt37	4567	46.60	45.82	26	83	*C. taitungensis*	AP009381.1
mt38	4781	48.67	76.15	27	94	*C. taitungensis*	AP009381.1
mt39	4967	47.03	64.71	12	80	*P. dactylifera*	JN375330.1
mt40	5881	45.04	45.72	25	85	*C. taitungensis*	AP009381.1
mt41	6191	45.47	65.59	17	88	*C. taitungensis*	AP009381.1
mt42	6778	45.53	48.16	47	92	*R. communis*	HQ874649.1
mt43	7410	46.31	52.78	20	88	*C. taitungensis*	AP009381.1
mt44	8851	45.33	61.23	15	92	*C. taitungensis*	AP009381.1
mt45	12 101	47.36	61.72	25	97	*A. sachalinensis*	FJ572121.1
mt46	20 658	44.97	58.65	6	87	*C. taitungensis*	AP009381.1

Table 2.5: Potential SNPs identified by mapping reads from all samples against the Punkaharju assembly. Variation was observed within samples, and SNPs are listed only where the alternate allele was present on $\geq$0.75 of reads for one or more samples ('Proportion Alternative' describes the fraction of reads containing the alternate allele). Samples are listed only if they were differentiated from the reference. At present, attempts have been made to resequence several of these loci via Sanger: those which have been tested are listed as either 'Confirmed' (SNP observed), 'Not Conf' (locus appeared to be monomorphic), or 'Not Amp' (tested, but PCR amplification unsuccessful).

Contig	Position	Sample	Punkaharju Reference	Alternative	Quality	Proportion Alternative	Depth	Status
mt13	1731	PSVAL	A	T	405	1.00	14	Untested
mt15	14	PSGLEN	A	T	203	0.66	29	Untested
mt15	14	PSVAL	A	T	209	0.65	26	Untested
mt15	14	MUGO	A	T	312	0.79	19	Untested
mt15	14	UNC	A	T	478	0.73	37	Untested
mt27	717	PSGLEN	C	A	1009	1.00	28	Not Conf
mt27	717	ULIG	C	A	1014	0.88	33	Not Conf
mt27	717	UNC	C	A	1090	0.91	35	Not Conf
mt29	1589	PSGLEN	A	C	543	1.00	14	Confirmed
mt29	1589	PSVAL	A	C	470	1.00	12	Confirmed
mt29	1589	MUGO	A	C	1343	1.00	34	Confirmed
mt29	1589	ULIG	A	C	1390	1.00	36	Confirmed
mt29	1589	UNC	A	C	1088	1.00	27	Confirmed
mt32	3446	PSGLEN	C	A	193	0.60	15	Untested
mt32	3446	PSVAL	C	A	462	0.83	18	Untested
mt32	3446	MUGO	C	A	262	0.61	18	Untested
mt32	3446	ULIG	C	A	587	0.74	27	Untested
mt32	3446	UNC	C	A	708	0.82	28	Untested
mt34	44	UNC	A	C	285	0.87	15	Not Amp
mt35	161	PSGLEN	G	T	343	1.00	11	Untested
mt35	161	PSVAL	G	T	573	1.00	18	Untested
mt35	161	MUGO	G	T	372	1.00	12	Untested
mt35	161	ULIG	G	T	409	0.88	17	Untested
mt35	161	UNC	G	T	314	0.92	12	Untested
mt35	3792	PSVAL	A	C	1300	0.97	35	Not Conf
mt36	1375	PSGLEN	T	G	1263	1.00	31	Not Amp
mt36	1375	PSVAL	T	G	876	1.00	22	Not Amp
mt36	1375	MUGO	T	G	1437	1.00	35	Not Amp
mt36	1375	ULIG	T	G	1842	1.00	45	Not Amp
mt36	1375	UNC	T	G	1190	1.00	29	Not Amp
mt43	1003	PSGLEN	G	T	1400	0.93	40	Not Amp
mt43	1003	MUGO	G	T	971	0.96	27	Not Amp
mt43	5723	PSGLEN	T	G	469	0.78	27	Confirmed
mt43	5723	MUGO	T	G	843	0.89	28	Confirmed
mt43	5723	ULIG	T	G	1280	0.80	55	Confirmed
mt44	7782	UNC	C	A	1480	0.93	45	Not Amp
mt46	14 419	PSGLEN	T	G	1164	1.00	29	Confirmed
mt46	14 419	PSVAL	T	G	1178	1.00	29	Confirmed
mt46	14 419	ULIG	T	G	2529	1.00	62	Confirmed
mt46	14 419	UNC	T	G	1881	1.00	46	Confirmed
mt46	14 419	MUGO	T	G	1877	1.00	47	Confirmed
mt46	17 759	PSVAL	A	C	894	1.00	22	Confirmed
mt46	17 759	UNC	A	C	1495	1.00	37	Confirmed

2.4.1 Assembly and Coverage

Pinus genomes are exceptionally long: the *P. taeda* draft assembly is ∼22Gbp in length (Neale et al., 2014); estimates of the *P. sylvestris* genome size from flow cytometry range from ∼21–27Gbp (Bogunic et al. (2003) and Valkonen et al. (1994), respectively). If we were to assume a genome size of 24Gbp, then with 10 million 100bp reads (similar to our data), and in the *absence* of any cytoplasmic genomes, we would anticipate a mean coverage depth of around 4.2×10^{-4}. Given that cpDNA and mtDNA sequences are expected to be present in abundance, genomic coverage is anticipated to be lower still. Being both radically shorter and present in many more copies, the cytoplasmic genomes should receive substantially greater coverage and therefore be more amenable to *de novo* assembly.

By first testing assembled contigs against an arbitrary region of *P. taeda* genomic DNA, we attempted to gauge the extent of nuclear coverage. The genomic subset was predominantly comprised of short contigs (<1kbp), and coverage breadth was poor; nevertheless, average coverage depth for those contigs was unexpectedly high. Read depth was, however, highly variable, and given the poor breadth of coverage it seems likely that a sizeable proportion of mapped reads may have originated from regions repeated throughout the genome. One caveat to this approach is that the extent of sequence homology between *P. taeda* and *P. sylvestris* is not known, and that coverage breadth may have been greater were a reference for the latter available.

Complete (or near complete) cpgenomes are available for numerous of species, and are markedly shorter than mtgenomes, typically ranging from 120–160kbp in length (Palmer, 1985). The Punkaharju assembly spanned 92% of the *P. sylvestris* reference, and at very substantial depth (approaching 900x), indicating the presence of large numbers of cpgenomes in megagametophyte tissue. Furthermore, this breadth of coverage was obtained over 39 contigs, 85% of which exceeded 1kbp, which provides some evidence that selecting for longer contigs effectively increases the bias of the assembly toward cytoplasmic genomes. The read-depth afforded to cpDNA would make WGS approaches well suited to studies of chloroplast diversity, as large numbers of samples could be multiplexed and still receive a reasonable level of coverage.

Putative mitochondrial contigs were identified by their similarity to the *C. taitungensis* mtgenome, excluding any <1kbp in length. Query cover (i.e. the proportion of a contig matched to the reference) varied greatly, and multiple hits were often observed across the length of an individual contig. Although taxonomically the closest available reference, *C. taitungensis* appeared nevertheless to be quite dissimilar. Interestingly, inversions

were apparently frequent, and if genuine, may be of benefit for use in phylogenetic inference: indeed, structural changes in mtDNA have previously been used to resolve plant phylogenies (Manhart and Palmer, 1990; Dombrovska and Qiu, 2004). After searching the Genbank database using the subset as a query, around a third of contigs found stronger hits with other plant species; in particular *Pinus*, where available. A positive relationship was observed between contig GC content, and the percentage query cover received ($p<0.01$): this may be due the propensity of regions containing genes to be higher in GC content (Bernardi, 1989; Oliver and Marín, 1996), and therefore subject to greater conservation on average. Plant mtgenomes have been shown to incorporate DNA from both chloroplast and nuclear genomes (Knoop, 2004; Wang et al., 2007): chloroplast DNA can easily be detected on account of the availability of references, however the possibility that genomic hybrid sequences may have been assembled cannot be discounted. Overall, however, the mean GC content of the mitochondrial subset (47%) was markedly differentiated from the 'background' of the assembly (41%), and from the genomic and cpDNA subsets (38 and 39%), but was virtually equal to the *C. taitungensis* mtgenome, and comparable to those of other plant species.

In the absence of suitable reference sequences, assembly quality is difficult to vet. The commonly used N50 statistic provides an indicator of contig size distribution (naturally, this is directly altered by size filtering), but while larger contigs are desirable, this is only true if they have not been misassembled. The only reference genome available for *P. sylvestris* was that of the chloroplast, against which *de novo* contigs were found to be highly accurate. Although the mtgenome is structurally dissimilar and may present alternative difficulties in assembly, the comprehensive coverage and fidelity provided by the cpDNA subset is encouraging.

2.4.2 mtDNA Polymorphisms

A number of potential SNPs in the mtgenome were identified between samples, however many of these also exhibited within-sample variation. There are two possible reasons for this, which are not in themselves mutually exclusive: heteroplasmy, whereby an individual may possess a number of mitochondrial haplotypes, or the occurrence of paralogous sequences not represented in the assembly.

Heteroplasmy is believed to be common among plant mitochondria (Kmiec et al., 2006), and in species where mitochondria are maternally inherited, may arise from leakage of the paternal mtgenome. This can in turn facilitate the establishment of new recombinant haplotypes (Städler and Delph, 2002; Pearl et al., 2009), such that emergence of variation is not solely dependent on mutation. Paternal leakage has previously been

reported to occur in other *Pinus* species (Wagner et al., 1991), and if widespread has implications for phylogeographic studies, as markers need not segregate exclusively via seed. Nevertheless, strong spatial differentiation has been observed in such studies, although it is feasible that the methods used to genotype individuals were insufficiently sensitive to detect low frequency haplotypes.

In our study, SNPs were described if they had alternate allele frequency $\geq$0.75, as sites closer to fixation are more desirable for use in phylogeography, as well as better suited for genotyping via Sanger sequencing. This threshold could of course be lowered, to expose intermediate frequency alleles, which may also be useful as markers as some populations may exhibit heteroplasmy at a given locus where others do not. Two of the variants described were indels: in these instances gaps were observed in coverage when mapping to the reference, however, at both sites these gaps were partially encroached by one or more individual reads, suggesting that the reference haplotype may also have been present at low frequency. It is also worth reiterating that our samples were not of individual gametophytes, but represented bulked collections of individual gametophytes from single open-pollinated maternal families. Therefore whilst it is assumed that progeny from a common mother should share a common mitochondrial haplotype, it is possible that there may have been some contribution from paternal leakage.

Plant mtgenomes are known to have acquired DNA from both plastid and nuclear genomes (Knoop, 2004; Wang et al., 2007), the latter having similarly assimilated mtDNA (Timmis et al., 2004). In the absence of complete reference genomes, it is possible that some reads were erroneously mapped to paralogues, resulting in false variant calls (this is less likely for reads of plastid origin, as reads were simultaneously mapped to cpDNA and mtDNA subsets). Variants at sites with unusually high read depth were omitted, though suitable references would be required to improve confidence further.

Overall, 14 putative SNPs were identified, and of these four have so far been verified by Sanger resequencing. All of those described were transversions, although a small number of potential transitions were observed at intermediate alternative allele frequencies (not shown). The rarity of transitions seems unusual, but is not without precedent: Wolfe et al. (1987) found that transitions comprised less than 50% of SNPs occurring between the plant mitochondrial genes, and in a study of date palm cultivars, $\sim$70% of the mitochondrial substitutions reported by Sabir et al. (2014) were transversions, in this case found overwhelmingly within intergenic sequences.

The pattern of variation observed did not support delineation between *P. sylvestris* and the *P. mugo* species complex. There is evidence that these species can readily hybridise (Wachowiak and Prus-Głowacki, 2008; Danusevičius et al., 2013), and conceivably may have occupied the same glacial refugia. They may therefore share a common pool of

mitochondrial haplotypes, which could be valuable for use in phylogeographic studies across lineages. Being (predominantly) maternally inherited, these markers may also be useful in studies of hybrid swarms, as the direction of hybridisation could be inferred.

2.4.3 Conclusions

A number of new variants were discovered in the study, which if authentic, will greatly increase the number of markers available for use in phylogeographic studies, and may therefore provide a substantial improvement in both resolution and our understanding of post glacial migration. WGS sequencing permits us to investigate large portions of previously unexplored non-coding regions, which are likely to be a far richer source of variation than those conserved between plant species. The contigs we identified were broadly consistent with the profile of mtDNA, however, their identities could not be confirmed with certainty. A prudent next step would therefore be to evaluate segregation in the offspring of controlled crosses, and confirm whether the variants described follow expected patterns of mitochondrial inheritance.

Our data suggest that pines exhibit mitochondrial heteroplasmy: if this is the case, the concept of haplotype frequency would be relevant at the individual as well as the population level. This has implications for studies of migration, as mitochondria may be able to traverse large distances via pollen. The impact of this is dependent on the frequency at which this occurs, as well as the rate at which 'foreign' haplotypes may become established within a population. Nevertheless, as chloroplast and nuclear markers are transmitted consistently via pollen, mitochondrial markers remain favourable in discerning geographic structure.

A major obstacle in screening WGS data for variants is the absence of suitable reference sequences, in this case those of the mitochondrial and nuclear genomes. Our assembly was provisionally screened against the mtgenome of *C. taitungensis*, however, despite being taxonomically the closest reference available, the breadth of coverage attained was very limited. This may be largely due to the large proportions of intergenic sequence present, which are likely subject to poor conservation. Should a completed *P. sylvestris* mtgenome become available, many new variants could be uncovered using these methods.

Chapter 3

Variation in needle characters among the native pinewoods of Scotland

3.1 Introduction

Scotland's native populations of *Pinus sylvestris* are the remnants of the Caledonian Pine Forest, now fragmented across the Highlands; an area the remainder shares with around ten times as much planted material. It is the mandate of Forestry Commission Scotland that the 'genetic integrity' of the native pinewoods be preserved (Mason et al., 2004), retaining the adaptive potential of the remaining pinewoods as much as possible. However, evidence on the extent of local adaptation in Scotland remains limited, and there is currently none available with regards to leaf morphology.

Leaves are the principal photosynthetic organs of plants, providing the supply of carbohydrates required for construction other plant tissues and re-investment in new leaves (Wright et al., 2004; Pallardy, 2008). Leaf architecture is also a compromise between maximal photosynthetic output, and the constraints imposed by the environment such as limited water availability, exposure, fungal attack, or herbivory. Within the range of a single species environments can vary substantially, and exert differential selection pressures such that no one design is optimal. Phenotypic variation in leaf morphology may be attributable to plasticity, genetic variation, or both. Within a common garden, systematic differences among populations may provide evidence for local adaptation; particularly if traits can be shown to covary with environmental parameters at the sites of origin. Despite the comparably small geographic scale, altitude and rainfall vary substantially throughout the Scottish Highlands, and both have well-documented influences on leaf morphology and physiology (e.g. Cunningham et al., 1999; Fonseca et al., 2000; Scheepens et al., 2010).

Stomata are pores on leaf surfaces via which gas exchange and transpiration take place, and are the sites where the majority of plant water loss occurs. Stomata are highly plastic, and stomatal apertures are responsive to light, water availability, and ambient CO_2 concentration (Pallardy, 2008). Despite the importance of stomata in plant water relations, consistent trends relating stomatal traits to the environment have proved somewhat elusive (Dillen et al., 2008). Lower stomatal densities have been reported in drier environments; for example in Canadian hardwoods (Carpenter and Smith, 1975), and in members of the Rubiaceae family (Hogan et al., 1994). However, stomatal densities have also been reported to increase with decreasing rainfall, such as (Dunlap and Stettler, 2001) who recorded higher stomatal densities in *Populus* clones from low altitude xeric sites, or similarly Limin et al. (2007) who found that density increased across a rainfall gradient in the grass *Leymus chinensis*. In *P. sylvestris*, stomata occur in longitudinal rows the number of which can vary between populations *in situ* (Bobowicz and Korczyk, 1994; Urbaniak et al., 2003; Androsiuk and Urbaniak, 2006). Native pinewoods in Scotland occupy a steep rainfall gradient, which decreases from west to east; as such it would be of interest to investigate whether stomatal characters exhibit any degree of covariation.

Needle length has been reported to differ between populations measured *in situ* in *P. sylvestris* and other *Pinus* species (Bobowicz and Korczyk, 1994; Donahue and Upton, 1996; Boulli et al., 2001). Donahue and Upton (1996) observed that individuals from drier northern provenances possessed shorter needles on average compared to those from the south in a study of *Pinus greggi* in Mexico. Tiwari et al. (2013) observed a decrease in the needle length of *Pinus roxburghii* with increasing altitude, and James et al. (1994) had previously reported a similar trend in Scotland.

Resin canals or 'ducts' are common to conifer species, and consist of longitudinal tube-like structures within needles, functioning as reservoirs for terpenoids and other antipathogenic compounds; serving to defend the plant against insect, bacterial, or fungal attack (Christiansen et al., 1987; Schroeder, 1990). The number of resin canals per needle can vary within species, and has repeatedly been shown to differ between provenances of *P. sylvestris* and between those of other pines (Matziris, 1982; Bobowicz and Radziejewska, 1989; Bobowicz and Korczyk, 1994; Donahue and Upton, 1996). Resin canals occupy leaf space at the expense of photosynthetic tissue, and may therefore be perceived as an adaptive trade-off between defence and production. In previous trials, material transplanted to the west of Scotland has performed poorly in comparison with local material (Mason et al., 2004), and it has been posited that this may be due to increased pathogen prevalence in the milder and damper conditions. In their book, *The Native Pinewoods of Scotland*, Steven and Carlisle (1959) describe differences in the

number of resin canals between native remnants; again, measurements were made *in situ* and so do not permit inferences on local adaptation.

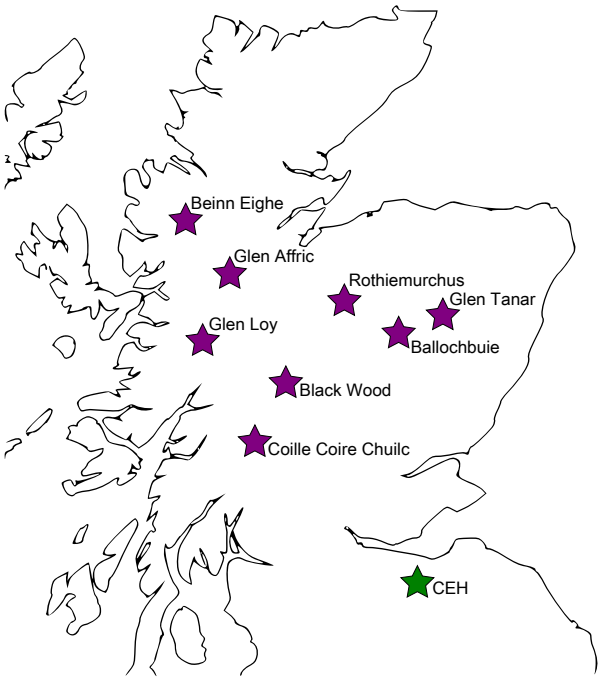
The study of functional leaf traits has increased in popularity, particularly as a means to generalise the 'resource-use' strategies of plant species. This is centred on the premise that plant species tend to exhibit a common suite of adaptations to environmental stresses (Grime, 1977; Chapin, 1991; Reich et al., 2003), although such generalisation is not without opposition (e.g. Grubb, 1985, 1998). Nevertheless, the field describes some candidate traits which may provide good targets for studies of local adaptation.

One such trait is specific leaf area (SLA), and is simply the ratio of leaf surface area to dry mass. Leaves with high SLA values tend to be broader and thinner relative to those with low SLA. This is perceived as a trade-off, with higher SLA values permitting more optimal photosynthetic activity per unit of mass, contrasted against the increased durability and longevity possible at lower SLA values e.g. Reich et al. (1991). Leaves with low SLA have higher construction costs per unit of surface area, and operate at lower photosynthetic rates per unit of dry mass, but may also last longer owing to the larger proportion of compounds and structural tissues that serve a function other than photosynthesis (Reich et al., 1992; Cunningham et al., 1999). SLA values have been observed to decrease in conjunction with rainfall and soil nutrient gradients for a variety of species, being predominantly lower when availability to either becomes limited (Givnish, 1987; Cunningham et al., 1999; Warren et al., 2005).

Analogous to SLA, leaf dry matter content (LDMC) is also considered to be indicative of leaf durability, and is the dry mass of a leaf expressed as a proportion of its original fresh (or water-saturated) mass. LDMC reflects leaf tissue density, and higher values indicate a larger proportion of structural compounds. LDMC has been shown to be positively correlated with leaf lifespan in grasses (e.g Ryser, 1996). Increased tissue density can reduce leaf palatability (Coley, 1988), and provide increased protection against abiotic stress (Pammenter et al., 1986). Wilson et al. (1999) advocate the use of LDMC in favour of SLA, due to difficulties in consistent estimation of the latter.

Here, we aim to determine whether native populations of *P. sylvestris* in Scotland show evidence of local adaptation for needle traits, by anatomical assessment of a five year old family/provenance common garden trial that includes eight populations sampled from across the natural range in Scotland. We begin by estimating genetic variance components across populations as a whole, to determine what proportion of the variance we observe is heritable, and to gauge the potential for leaf traits to adapt under selection. Secondly, we assess whether systematic differences in trait means can be observed between individuals from different provenances. Although needle characters have previously been reported to covary with site conditions, few studies report observations

Figure 3.1: Map of native provenances sampled for use in the study, including site of common garden trial (green symbol).



from a common environment, and so cannot provide evidence for local adaptation. In view of this, we assess whether interpopulation variation can be explained by covariation with the environment, and if so how this could be interpreted in the context of adaptive divergence.

3.2 Methods and Materials

3.2.1 Study Provenances

The trial consisted of material collected from eight native provenances from sites across the Scottish Highlands, which experience different average meteorological conditions (fig. 3.1 & table 3.1). Within each provenance seed was collected from four open-pollinated mother trees, chosen across an altitudinal gradient at each location. The collection therefore consisted of 32 families in total (8 Populations $\times$ 4 Families), with six offspring per family, and individuals within a given family were regarded as half-sibs (i.e. seeds with a common mother, but an unknown set of fathers).

Table 3.1: Population differentiation in trait values should be evaluated in the context of environmental variation between sites.

Provenance	Latitude	Longitude	Altitudinal Range (m a.s.l.)	Mean Monthly Temp (°C)	Mean Monthly Rainfall (mm)
Ballochbuie	56.99	−3.30	421–524	6.44	79.85
Beinn Eighe	57.63	−5.35	17–91	8.53	183.51
Black Wood	56.67	−4.32	250–307	7.06	96.61
Coille Coire Chuilc	56.41	−4.71	222–298	6.46	257.64
Glen Affric	57.27	−4.92	205–274	7.02	153.11
Glen Loy	56.91	−5.13	136–197	7.96	152.02
Glen Tanar	57.05	−2.86	293–422	7.40	70.91
Rothiemurchus	57.15	−3.77	306–329	6.82	67.66
CEH	55.86	−3.21	190	7.08	77.73

3.2.2 Experimental Design

The seed collected was used to establish a common garden trial at the CEH, located ∼10km south of Edinburgh. The trial site receives similar levels of precipitation to the easternmost provenances, but is at lower altitude and latitude (fig. 3.1). In June 2007, seed was sown on trays containing 3:1 John Innes compost (type 1) to sand, and maintained in the glasshouse under natural lighting conditions. Following germination, seedlings were transferred to pots 11cm diameter × 9.6cm depth, and watered two to three times per week throughout the growing season. In spring 2008, seedlings were re-potted into 11 × 11 × 12cm pots and relocated to outdoor benches where plants were watered only during periods of unusually low precipitation to prevent droughting.

Altogether, the outdoor trial consisted of 192 individuals, incorporating six representatives from each of the 32 families, arranged over six blocks, each consisting of 32 plants (one member from each half-sib family), which were randomised within blocks. Plants were re-potted once more in spring 2011 into 13 × 13 × 13cm pots.

3.2.3 Sampling protocol

Needles were harvested for analysis during summer 2012. In order to ensure that only fully expanded leaves were measured, all needles were drawn from previous-year whorl branches. In *P. sylvestris* a fascicle consists of a pair of similarly sized needles. Five fascicles were collected from each individual: one needle from each pair was used for destructive measurements, while its partner was weighed. After removal from the plant, needles were placed immediately onto damp tissue and stored in Petri dishes for no longer than two hours prior to dissection or weight recording.

3.2.4 Anatomical Measurements

Following collection, all needles were scanned using a *Canon Lide 210*flatbed scanner to provide images for length estimation. Needles subject to anatomical assessment were first viewed using a *Leica Wild M3Z*stereo microscope. The number of stomatal rows on the adaxial (upper) and abaxial (lower) surfaces of the needle were counted after which the needles were cut in half and sectioned by hand using a sharp razor blade to provide transverse sections (TS) (fig. 3.2). As internal and external characters may vary throughout the length of a needle, both stomatal row counts and cross sections were obtained from the approximate centre of each. Each TS was stained with 0.05% Aniline Blue, and digitally photographed using a *Leica DFC290* camera attached to a *Leica DM2500* light microscope (x10 objective). All measurements on captured images were made via *MacBiophotonics ImageJ* (Abramoff et al. (2004)). Fresh mass was obtained from the second needle in each pair using an *Ohaus Analytical Plus* balance, and dry mass was recorded after needles had spent a minimum of three days in a desiccating oven at ∼60°C. In summary, the following traits were recorded from intact needles: length (mm), number of stomatal rows (adaxial and abaxial), fresh mass and dry mass (mg); and the remaining from transverse cross-sections: TS area mm^2, TS perimeter (mm), vascular bundle (VB) area mm^2, VB perimeter (mm), width (mm), depth (mm), and number of resin canals.

3.2.5 Estimation of SLA and LDMC

Specific leaf area (SLA) is defined as the ratio of leaf surface area to dry mass. For the purposes of estimating area needles were envisaged as open semi-cylinders, and SLA ($\mathsf{mm}^2\,\mathsf{mg}^{-1}$) was approximated by the following:-

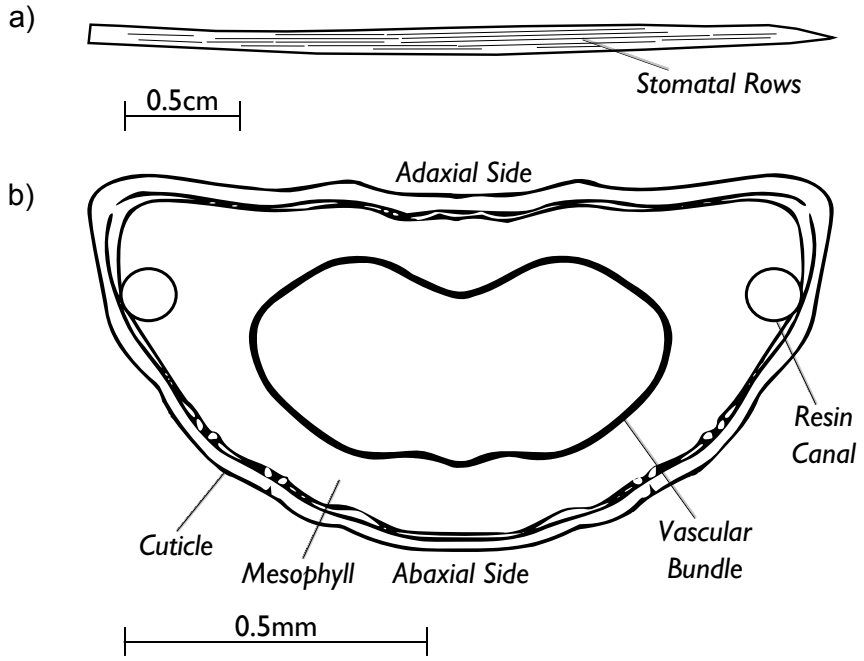
$$SLA = (\pi rl + 2rl)/m \tag{3.1}$$

where r (mm) is the radius (i.e. depth) of a needle, l (mm) is length, and m (mg) is dry mass. Leaf dry matter content (LDMC) was expressed as the ratio of leaf dry mass to fresh mass, and was logit transformed prior to analyses.

3.2.6 Analysis

All analyses were performed in *R* (R Development Core Team, 2012). Mixed-effects models were produced via the 'lme4' (Bates et al., 2013) package, and compared using 'MuMIn' (Barton, 2013). Figures were generated with 'ggplot2' (Wickham, 2009). Prior

Figure 3.2: Schematics of a) an intact pine needle (stomatal rows are present on both sides), and b) a transverse cross-section. *Pinus sylvestris* needles are typically curved to some extent, and are often twisted; they possess a minimum of two resin canals, but commonly have more distributed around the needle perimeter. Scale bars are approximate and based upon average values.



to further analyses, phenotypic correlations were estimated between all traits using family means. Morphological traits typically exhibit a high degree of correlation: in order to avoid repetitive analysis of multiple near-equivalent traits, a smaller number was chosen to be representative of the data set as a whole.

3.2.7 Model Fitting Strategy

Mixed effects models are not new, but have only relatively recently become more widely used in biology, representing something of a watershed statistical methodology. Currently there is some contention on how best to apply these models, and how parameters under consideration might be deemed meaningful (or 'significant' in the traditional sense), in that they contribute some measurable explanatory power to the model. One of the most popular packages currently used for mixed effects modelling is the 'lme4' package (Bates et al., 2013) available for use in *R*. Models produced by lme4 do not provide p-values for any parameter estimates, as it is contended that F-statistics cannot be estimated effectively using current methodology. As p-values have long been the mainstay of biological statistics, this is problematic; however, alternative 'information theoretic' (IT) approaches are becoming increasingly adopted.

In this and subsequent chapters, fixed effects are evaluated by comparing them against a null, or base model (typically containing only an intercept, without fixed effects) and comparing the Akaike Information Criteria (AIC) scores (Akaike, 1974). Models generated for the purpose of comparison were fitted using maximum likelihood (ML), however restricted maximum likelihood (REML) was employed once . AIC is a measure of the relative explanatory power of a given model, based upon the maximised likelihood function of the model penalised by the number of parameters used. The objective of AIC is therefore to find the 'simplest best' model, such that useful parameters are retained but overfitting is avoided by excluding additional parameters which do not sufficiently contribute to explanatory power. For a given set of models, those with lower AIC scores are deemed 'better' than the alternative candidates in this context. Strictly speaking, we employed 'corrected AIC' (AICc) during model fitting, which contains a correction for finite sample sizes (Sugiura, 1978). As sample size increases, AICc converges to AIC, making it generally preferable to the latter regardless of sample size (Burnham and Anderson, 2004).

It is often the case that more than one fixed effect is of interest (e.g. longitude, rainfall, and altitude), and in such cases it is possible to generate a set of candidate models representing each possible combination of fixed effects. For example, in the case of two parameters, four models would be compared: a null model (intercept only), a model for each fixed effect individually, and a model containing both. Generally, a given number of fixed effects, i, will require 2^i candidate models for assessment of all possible combinations, however we may exclude some candidates *a priori*, in order to restrict the maximum number of parameters. This process can be streamlined via another *R* package, 'MuMIn'. The output for MuMIn provides AICc scores for each of the models tested, and in addition assigns each a ΔAICc value, which represents the difference in AICc between the given model and the best model, and an 'Akaike Weight'. The Akaike weight is calculated from a model's ΔAICc relative to those of the other candidate models (the weights of all candidate models sum to one), and represents the probability of that model being the best of those under consideration. AIC scores are convenient in that they can be used to rate the comparative merits of a large set of models simultaneously, and those models need not be nested. Models generated for the purpose of comparison via AICc were necessarily fitted using maximum likelihood (ML), however restricted maximum likelihood (REML) was employed to obtain provide estimates ultimately reported.

The above describes how fixed effects were evaluated in our model testing framework, but does not address the random effects. In our analyses random effects were used to account for variance between blocks or to control for pseudoreplication owing to the non-independence of observations within a group (i.e. each population group consists of a number of maternal families, and is not a random sample of the population as a

whole). For this purpose, random effects were not tested, and were retained within all candidate models irrespective of what proportion of the total variance they were able to account for. We also used random effects to estimate genetic variance components, such as V_{A} and V_{POP}. In these instances, estimation of random effects was the purpose of the model, rather than evaluation of fixed effects.

3.2.8 Meteorological Data

To characterize the environments occupied by all 84 existing remnants of the Caledonian Pinewood resource, data were extracted for each sites from a Met Office dataset containing meteorological data between the years 1961–1990 (Perry and Hollis, 2005), using coordinates published by the Forestry Commission (Mason et al., 2004) (NB: with corrected coordinates for Beinn Eighe as those in the database were found to be incorrect). The Met Office dataset provides information on 13 parameters recorded daily for all locations within the UK landmass as represented by 5 × 5km grid squares; however, due to the incomplete coverage of monitoring stations many values have been estimated by regression and interpolation.

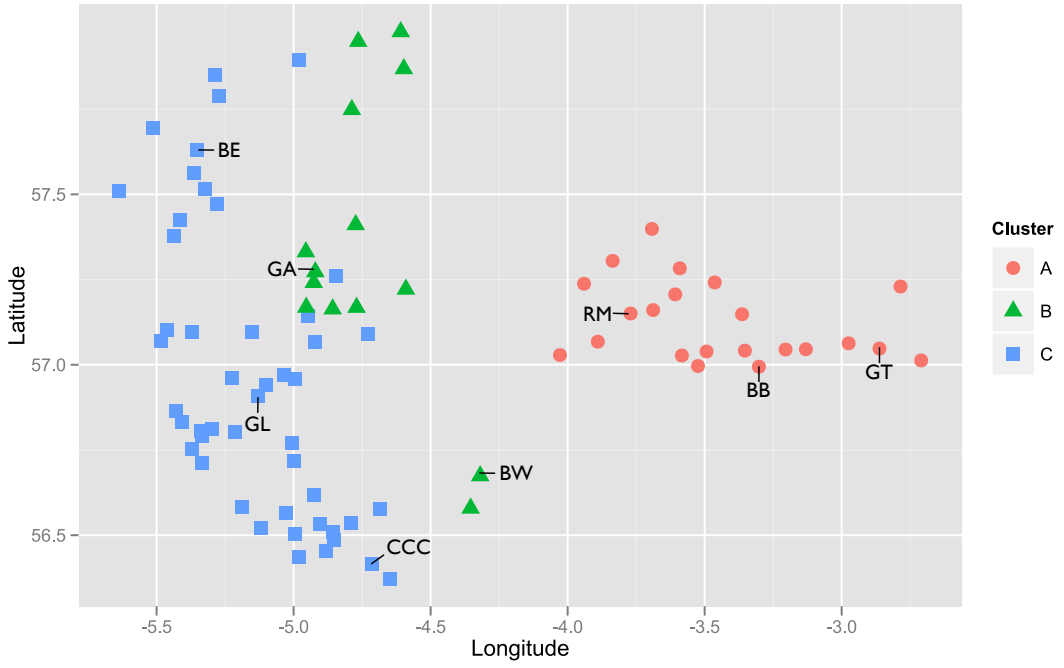
For each of the sites, the mean monthly values were calculated for all 13 of the meteorological parameters evaluated. Many of these parameters are highly correlated with one another, and the *R* package 'FactoMineR' (Hussonm et al., 2013) was used to perform a principal components analysis (PCA) in order to represent the data as a smaller number of uncorrelated components. The first two principal components accounted for ∼85% of the total variation (55% and 30% respectively); PC1 was primarily representative of temperature-based variation, while the largest contributors to PC2 pertained to precipitation and cloud cover. A complete list of parameters and their proportional contributions to the principal components is provided in table 3.2.

Sites were then clustered based upon similarity of PC1 and PC2 values using the package 'mclust' (Fraley et al., 2012). The optimal number of groups was found to be three, representing climates with low (A), intermediate (B), and high (C) values for both principal components (in real terms, cooler/drier to warmer/wetter). All 84 provenances are represented in fig. 3.3, coloured by cluster: group A contains pinewoods in the Cairngorm Mountains, group B forms a belt dividing east and west, and group C is a collection of sites adjoining or approaching the west coast. The clustering outlined regions of climatic similarity between the native pinewoods, and this might be a useful consideration in future common garden design.

Parameter	Contribution PC1	PC2
Maximum Temperature (°C)	8.29	8.79
Minimum Temperature (°C)	13.27	0.59
Temperature (°C)	11.34	4.55
Days of Air Frost	12.74	0.49
Days of Ground Frost	12.31	0.39
Sunshine Duration (hours/day)	1.64	16.75
Precipitation (mm)	5.68	11.81
Days of Rainfall ≥ 1(mm)	3.93	17.71
Days of Rainfall ≥ 10(mm)	4.72	14.41
Sea-level Pressure (hPa)	0.85	0.16
Relative humidity (%)*	9.07	3.70
Vapour Pressure (hPa)	12.70	1.95
Cloud Cover (%)*	3.45	18.69

Table 3.2: Meteorological parameters used to generate principal components, and their proportional contributions to PC1 and PC2. Contributions for each component sum to one (those presented have been rounded).*logit transformed prior to PCA.

Figure 3.3: The 84 provenances currently recognised as native pinewoods. Sites are coloured based on similarity of PC1 and PC2 values (determined via *mclust*), and those represented in the needle anatomy study have been labelled.



3.2.9 Estimation of Pooled Genetic Components

Genetic variance components were estimated pooling across populations by first fitting a model to each trait in which Population, Family (within Population), and Block were incorporated as random effects with random intercepts:-

$$Trait = \mu + Population_i + Family_{j(i)} + Block_k + \epsilon_{ijk} \tag{3.2}$$

Pooled narrow-sense heritabilities (h^2) were estimated as follows using data from all populations:-

$$h^2 = \frac{V_{\mathrm{A}}}{V_{\mathrm{P}}} = \frac{4V_{\mathrm{fam}}}{V_{\mathrm{fam}} + V_{\mathrm{blk}} + V_{\mathrm{res}}} \tag{3.3}$$

where V_{A} is additive genetic variance, V_{P} is phenotypic variance, V_{fam} and V_{blk} are the between-family and between-block components, and V_{res} is the residual variance (in a half-sib design, V_{A} may be estimated as four times V_{fam}). To control for possible inter-population differences the population component, V_{pop}, was excluded from the estimate of V_{P}. Standard errors for the heritability were estimated using the method proposed by (Visscher, 1998) as:-

$$\mathrm{SE}_{h^2} = 4\sqrt{\frac{2\left(1-\frac{h^2}{4}\right)^2\left[1+(s-1)\frac{h^2}{4}\right]^2}{s(s-1)(f-1)}} \tag{3.4}$$

where s is the number of offspring per family, and f is the number of families. The genetic coefficient of variation $\mathrm{CV_A}$ (Houle, 1992) is a measure of additive genetic variability normalised by the trait mean, and was also estimated pooling across populations by the following:-

$$\mathrm{CV_A} = \frac{\sqrt{V_{\mathrm{A}}}}{\mu_{\mathrm{trait}}} \times 100 \tag{3.5}$$

where μ_{trait} is the mean of the given trait.

3.2.10 Interpopulation Differences in Trait Means

Interpopulation differences were evaluated by comparison of model pairs by AICc for each trait. Each model pair consisted of a null model containing no fixed effects other than the intercept, to an alternative model including Population as a fixed effect:-

$$Trait = \mu + Family_j + Block_k + \epsilon_{jk} \tag{3.6}$$

$$Trait = \mu + Population_i + Family_{j(i)} + Block_k + \epsilon_{ijk} \tag{3.7}$$

where Family and Block are considered random effects, and Family is nested within Population (when included).

Additionally, Q_{ST}, a measure of the extent of differentiation for quantitative traits (Spitze, 1993), was derived as:-

$$Q_{\mathrm{ST}} = \frac{V_{\mathrm{pop}}}{V_{\mathrm{pop}} + 2(V_{\mathrm{A}})} \tag{3.8}$$

using the variance components obtained from eq. (3.2).

3.2.11 Environmental Covariation

To investigate the explanatory power of environmental parameters, trait values were regressed against longitude, latitude, and meteorological components PC1 and PC2. The landscape and climate of Scotland are such that a considerable degree of correlation exists between longitude, PC1, and PC2 (population altitude was excluded on account of a particularly strong relationship with PC1); in order to determine which factors, if any, best explain variation across sites, a global model was first specified for each of the traits examined using a similar format as eq. (3.2), but with the four environmental parameters included as fixed effects. Model subsets were generated from the global model using MuMIn, but were restricted to a maximum of two fixed effects (11 possible models including the null model), and the best models were identified by AICc.

Estimation of the coefficient of determination (R^2) is non-trivial for mixed models, and R^2 estimates were obtained in two ways. Firstly, using the method proposed by Nakagawa and Schielzeth (2013), whereby marginal R^2 ($R^2_{\mathrm{GLMM(m)}}$) is defined as the variance accounted for by fixed effects divided by the total variance (including random effects and residuals). An alternative (R^2_{POPMEANS}), was estimated via simple linear regression of population means against the parameter of interest, and is reported as the *adjusted-*R^2 estimate produced by the model.

3.3 Results

Phenotypic correlations between family trait means were commonly high (>0.5), particularly those which describe the dimensions of a needle cross-section, that were typically >0.9 (table 3.3). In addition, all traits were positively correlated with the exception of SLA; as the surface area of a needle increases at a lower rate than mass, SLA decreases with increasing needle size. The following traits were chosen to be representative of the

data:- TS area, SLA, length, number of resin canals, number of stomatal rows (adaxial & abaxial), dry mass, and LDMC.

3.3.1 Pooled Genetic Components

Narrow sense heritability estimates ranged from 0.23–0.73 (table 3.3) between traits, although SEs were substantial (table 3.4). Number of resin canals was the only trait for which the 95% CI did not include zero.

CV_A ranged between 6.90–16.70 (table 3.3): the lowest values were for functional leaf traits SLA (6.90), and LDMC (negligible); the highest being the number of resin canals (16.70). For TS area CV_A was estimated as 11.37, however it should be noted that surface areas usually produce higher values than linear measurements (Lande, 1977). For comparison, the h^2 and CV_A for the highly correlated trait, TS perimeter, were 0.35 and 5.74 respectively.

The proportion of total variance attributable to between-block differences was generally low ($< 5\%$), with the notable exceptions of SLA and LDMC (25.32 and 42.03%). These traits appeared to be highly sensitive to the small environmental variability present within the trial, which may have been caused by inadvertent shading from nearby foliage.

Table 3.3: Pearson's correlation coefficient between family means for all needle traits examined. TS Area, TS Perim, Width, and Depth are transverse cross-section parameters, while VB Area, and VB Perim relate to the area and perimeter of the vascular bundle within a transverse cross-section. F Mass and D Mass are the fresh and dry masses of needles.

	TS Perim	VB Area	VB Perim	Width	Depth	SLA	Length	No Canals	Ad Rows	Ab Rows	F Mass	D Mass	LDMC
TS Area	0.987	0.930	0.927	0.962	0.955	−0.708	0.331	0.475	0.741	0.711	0.786	0.770	0.130
TS Perim		0.910	0.926	0.988	0.918	−0.713	0.280	0.527	0.765	0.691	0.805	0.787	0.127
VB Area			0.981	0.876	0.875	−0.683	0.415	0.623	0.656	0.666	0.864	0.841	0.094
VB Perim				0.911	0.839	−0.708	0.418	0.669	0.668	0.664	0.885	0.863	0.109
Width					0.866	−0.702	0.221	0.549	0.775	0.664	0.798	0.772	0.081
Depth						−0.673	0.333	0.296	0.727	0.765	0.659	0.657	0.180
SLA							−0.167	−0.565	−0.713	−0.676	−0.777	−0.829	−0.562
Length								0.212	0.344	0.227	0.889	0.877	0.134
No Canals									0.141	0.158	0.313	0.334	0.155
Ad Rows										0.758	0.599	0.596	0.191
Ab Rows											0.536	0.525	0.146
F Mass												0.985	0.174
D Mass													0.330

*LDMC is a proportion, and was logit transformed prior to analysis.

Table 3.4: Narrow-sense heritabilities (h^2) with their associated standard errors, the coefficient of genetic variation (CV_A), and the proportions of variance attributable to family and block effects pooled across populations. CV_A was not estimated for LDMC, as negative and positive values were present following logit transformation.

Trait	h^2 (SE)	CV_A	Variance due to Pop (%)	Variance due to Fam (%)	Variance due to Block(%)
TS Area	0.41 (0.25)	11.37	1.49	10.16	2.45
SLA	0.25 (0.23)	6.89	1.79	6.18	25.32
Length	0.26 (0.23)	11.10	7.18	6.03	3.42
No Canals (log_{10})	0.73 (0.29)	16.70	0.00	18.27	3.32
No Ad Rows	0.37 (0.25)	10.23	8.33	8.57	3.05
No Ab Rows	0.39 (0.25)	9.42	1.99	9.44	1.74
Dry Mass	0.23 (0.23)	16.59	5.04	5.52	4.89
LDMC (logit)	0.09 (0.20)	-	0.00	2.19	42.03

3.3.2 Interpopulation Differentiation

Generally, little variation could be explained by interpopulation differences. Models including population as a fixed effect had superior AICc scores compared with their null counterparts for two of the traits examined: length, and number of stomatal rows (adaxial) (table 3.5). However, ΔAICc values were <2, which suggests that the improvement was not substantial.

Needles ranged from 14.11–54.91mm in length, with a mean of 32.37($\pm$1.06)mm; the longest needles on average were from Rothiemurchus at 37.42($\pm$1.68)mm, followed by Glen Affric at 34.82($\pm$2.96)mm, and the shortest were from Beinn Eighe at 29.66($\pm$1.16)mm (though other provenances were not dissimilar)(fig. 3.4a). Needles possessed between 3–12 stomatal rows on the adaxial surface (on average 6.66($\pm$0.17)); individuals from Rothiemurchus had the largest number on average (7.12$\pm$0.12), and those from Glen Tanar the fewest (5.86$\pm$0.37) (fig. 3.4c). Stomatal rows on the abaxial surface were positively correlated with those on the adaxial surface and occupied a similar range (4–12).

Individually, needle dry masses ranged from 1.53–25.03mg, the heaviest being around 16$\times$ that of the lightest; average dry mass was 7.98($\pm$0.41)mg. Similarly to length, the heaviest needles by dry mass were from Rothiemurchus followed by Glen Affric (9.72($\pm$0.70) and 8.96($\pm$0.96)mg) (fig. 3.4b). Mean values among the other provenances were highly similar; the lightest belonged to Ballochbuie at 7.10($\pm$0.51)mg. No evidence was found for differences in SLA between populations (fig. 3.4d), which averaged 14.18($\pm$0.44)mm^2mg^{-1}, nor for the number of resin canals which ranged between 2–8 per needle TS section (averaging at 3.78($\pm$0.12)), or furthermore for LDMC which was typically $\sim$46%.

Table 3.5: Comparisons of models with and without population as a fixed effect by AICc and Q_{ST} estimates for each trait. ΔAICc values represent the difference in AICc between the null model and the best (a value of zero indicates that the null model was the most parsimonious), and Akaike Weight represents the probability of that model being the best of the two considered.

Trait	Intercept	Population as Fixed Effect	df	ΔAICc	Akaike Weight	Q_{ST}
TS Area (mm^2)	0.62	-	4	0.00	0.69	0.02
SLA (mm^2mg^{-1})	14.18	-	4	0.00	0.91	0.04
Length (mm)	30.96	+	11	1.43	0.67	0.13
No Canals (log_{10})	0.56	-	4	0.00	0.99	0.00
No of Stom Rows (adaxial)	6.29	+	11	1.48	0.68	0.11
No of Stom Rows (abaxial)	7.62	-	4	0.00	0.94	0.03
Dry Mass (mg)	7.98	-	4	0.00	0.60	0.10
LDMC (logit)	−0.14	-	4	0.00	0.99	0.00

Q_{ST} values were generally very low, and broadly reflected the model testing results described above. The traits needle length, and number of stomatal rows (adaxial) possessed somewhat higher values of Q_{ST} in the region of 0.11–0.13, than the others which were consistently <0.1.

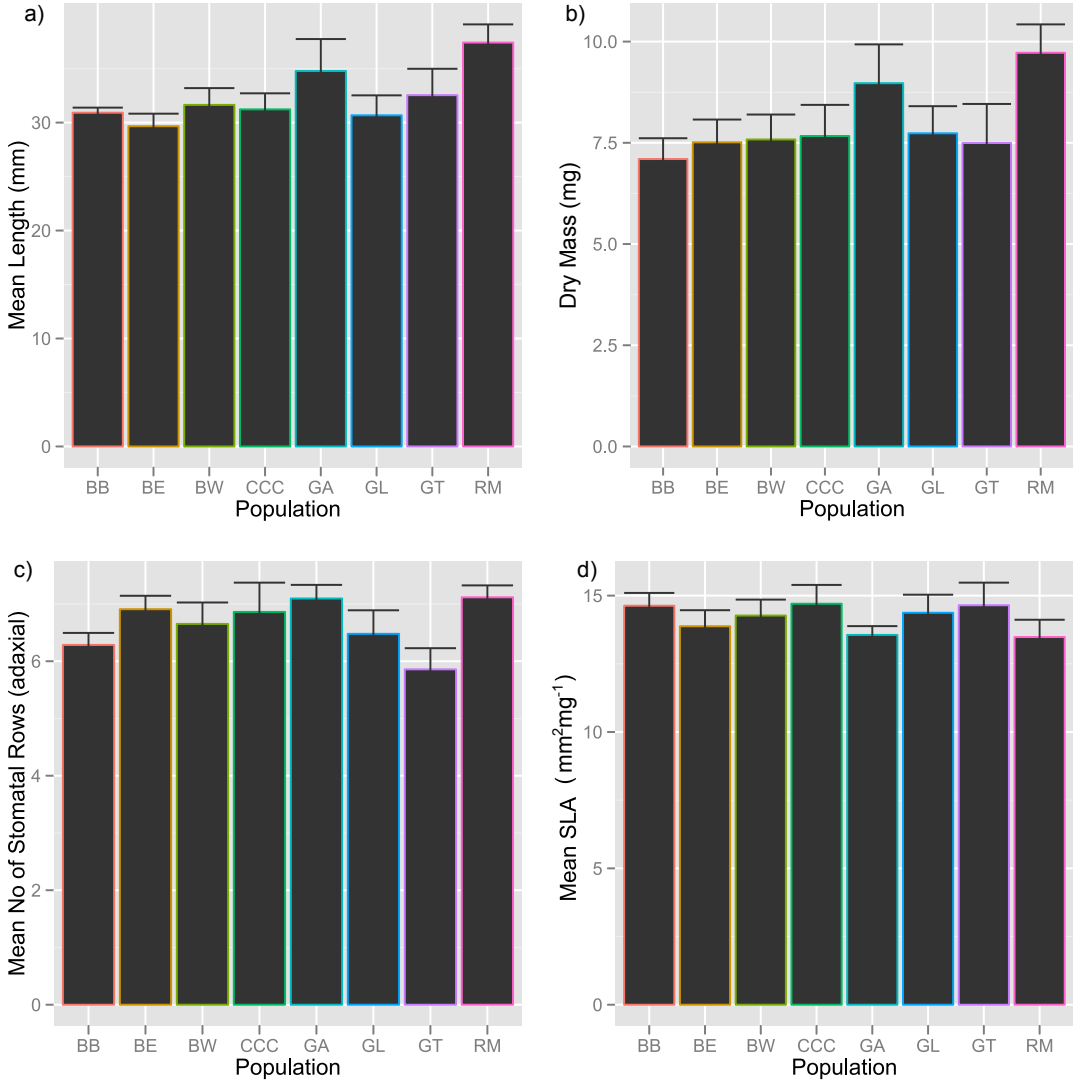
3.3.3 Environmental Covariation

A summary of the exploratory analysis of environmental predictors is presented in table 3.6. A negative relationship was detected between the number of stomatal rows on the abaxial surface of needles and longitude (fig. 3.5), by which the average number decreases from west to east. The longitude model was found to be the best overall, and a substantial improvement over the base model (ΔAICc=4.41).

The coefficient of determination was very low ($R^2_{\mathrm{GLMM(m)}}$=0.05), suggesting that longitude alone is a poor predictor of the number abaxial rows at the level of individual trees (regressing through population means, R^2_{POPMEANS}=0.62). The numbers of stomatal rows on opposite sides of the needle are positively correlated with one another, and results suggested the number of the rows on the adaxial surface are also negatively correlated with longitude (ΔAICc=1.82); but in this case the additional inclusion of a negative relationship with PC1 was found to improve AICc further (ΔAICc=4.55). As with results describing the abaxial surface, the model had relatively poor explanatory power at the individual level,($R^2_{\mathrm{GLMM(m)}} = 0.09$, R^2_{POPMEANS}=0.58).

There was some evidence of environmental covariation among the other traits evaluated, however it was not substantial (ΔAICc<2), and simple linear regressions through population mean values were not significant.

Figure 3.4: Barplots of selected traits by population: a) Length, b) Dry Mass, c) Stomatal Rows (adaxial), and d) SLA. Error bars represent ±1 SE, and were estimated by modelling each population separately using REML.



3.4 Discussion

In this study, genetic variation for in leaf traits in native Scots pine stands was assessed within and among provenances, and in relation to environmental variation at sites of origin.

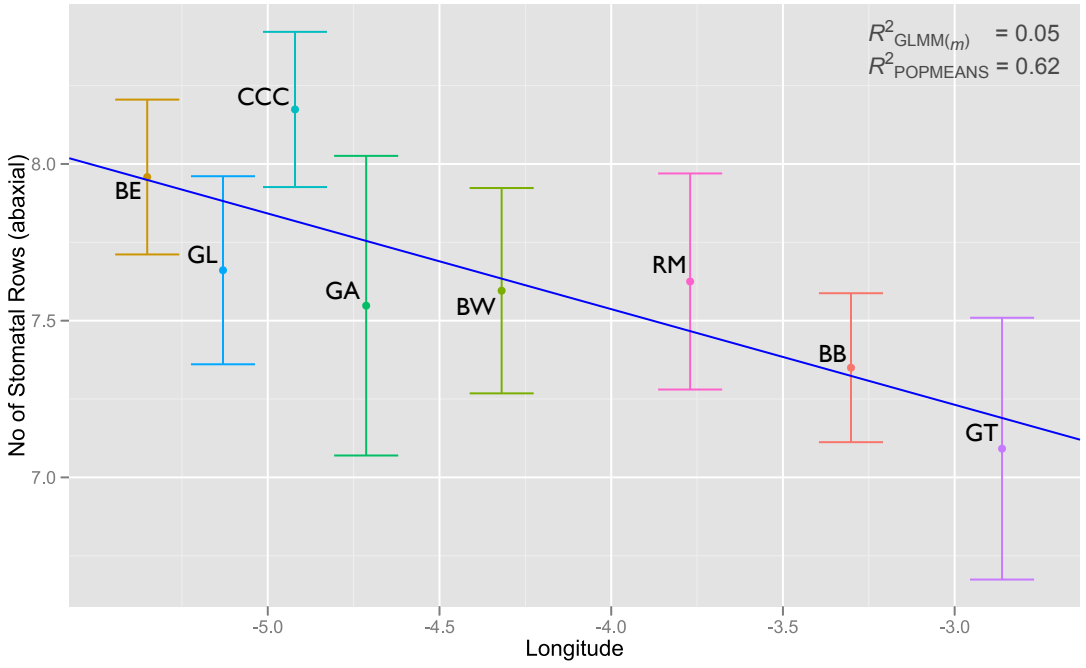
3.4.1 Genetic Basis of Leaf Trait Variations

Heritabilities and genetic coefficients of variability were estimated to provide an indication of additive variation remaining across Scottish populations, and the potential for

Table 3.6: Summary of model tests for environmental parameters. Traits were tested for relationships with environmental covariates (see methods), and regression coefficients for models that were the best by AICc are presented. ΔAICc values represent the difference in AICc between the null model and the best (a value of zero indicates that the null model was the most parsimonious). The proportion of variance explained by fixed effects, $R^2_{\text{GLMM(m)}}$, is provided for the best models.

	Coefficients of Best Models					
Trait	Latitude	Longitude	PC1	PC2	ΔAICc	$R^2_{\text{GLMM(m)}}$
TS Area (mm^2)	-	-0.02	-	-	0.95	0.02
SLA ($mm^2\,mg^{-1}$)	-0.82	-	-	-	1.20	0.02
Length (mm)	-	-	-0.59	-	0.20	0.03
No Canals (log_{10})	-	-	-	-	0.00	-
No Ad Rows	-	-0.73	-0.20	-	4.55	0.09
No Ab Rows	-	-0.30	-	-	4.41	0.05
Dry Mass (mg)	-	-	-	-	0.00	-
LDMC (logit)	0.04	-	-	-	0.20	0.01

Figure 3.5: Negative relationships between population means and longitude for the number of stomatal rows (abaxial). Coefficients were estimated via REML; error bars represent ±1 SE, and were estimated for each population individually.



leaf traits to respond to selection.

With the exception of LDMC, all of the traits produced heritability estimates which were at the least moderate (>0.20), however estimates were subject to large standard errors and high accuracy is unlikely to be achieved without a very large increase in sample size. The number of resin canals found within TS sections of needles was also highly heritable (0.73), with a large proportion of phenotypic variance attributable to differences among families. A large h^2 is a characteristic of fluctuating selection (Bell, 2010); as resin canals are understood to perform a defensive function, we might speculate that the threat from pathogens or herbivorous insects has varied considerably through time, or alternately that selection varies across small spatial scales. However, as the lower bound 95%CI for resin canal h^2 exists at 0.16, it cannot be said with certainty that this estimate is unusually high.

In contrast, additive genetic variance for LDMC was negligible (at least within the confines of our data). Notably, both SLA and LDMC were subject to large block variation (∼26 & 42% respectively), which could be attributable to shading across part of the trial from a nearby tree, as SLA is known to be responsive to shade (e.g. Lombardini et al. (2009)). The sensitivity of these traits to environmental variation within a single site underscores the difficulty of making inferences on adaptation from data obtained *in situ*.

Genetic components are valid only for the environment in which they are estimated, and furthermore estimates of h^2 tend to be higher in single site common garden trials than reciprocal transplants (White et al., 2007). Nevertheless, despite the losses in population size suffered by the Caledonian Pine Forest, our estimates suggest that native stands have retained considerable adaptive capacity for needle related traits.

3.4.2 Differences Among Populations

We found limited evidence that populations were distinguishable by virtue of their sites of origin alone. In *P. sylvestris*, needle length and stomatal row number have been reported to show sizeable variation between sites *in situ* (Bobowicz and Korczyk, 1994; Urbaniak et al., 2003; Androsiuk and Urbaniak, 2006). Both of these traits displayed interpopulation variation in the common garden trial, although evidence was not strong, and only the number of stomatal rows on the adaxial (upper) side of needles was differentiated between sites.

Resin canal numbers have also been previously demonstrated regional differences (Matziris, 1982; Bobowicz and Radziejewska, 1989; Bobowicz and Korczyk, 1994; Donahue and

Upton, 1996), however, although ostensibly heritable at the metapopulation level, canal numbers per TS needle section in our trial were highly similar between populations. From this it might be deduced that resin canal numbers are not subject to differential selection in the trial provenances, or alternatively that selective differences are not sufficiently strong as to offset gene flow among populations.

Similarly, the measure of genetic differentiation, Q_{ST}, was low for all traits examined (<0.15), which suggests that the genetic variability within populations is high relative to the variation between. As our trial was restricted to a single location, we were unable to detect adaptive variation that may be concealed by genotype $\times$ environment interactions. It is possible that some interpopulation differences may only become manifest in the presence of specific environmental cues; however, with the exception of needle length and stomatal row numbers, we did not find evidence that needle traits differed among native pinewoods.

3.4.3 Environmental Trends

Population differentiation in trait values should be evaluated in the context of environmental variation between sites. Even against a background of low interpopulation variability, trends may become apparent with the inclusion of environmental covariates.

The numbers of stomatal rows on abaxial needle surfaces were found to be negatively correlated with longitude, becoming lower on average from west to east. Similarly, the number of stomatal rows on the adaxial surface decreased with longitude, in conjunction with decreasing PC1 (which corresponds to lowering temperatures and increasing altitude). As stomata are the structures through which the majority of plant water loss takes place, we could interpret this as a response to the reduction in water availability with increasing longitude. Stomatal density has previously been observed to be lower for plants inhabiting drier environments (e.g. Carpenter and Smith, 1975; Hogan et al., 1994). Larger, wider needles also have a greater number of stomatal rows on average, so an alternative possibility is that some individuals possess more by virtue of selection for increasing size. There was some evidence that needle TS area also decreased toward the east, however, there was little improvement over the null model in terms of AICc. Longitudinal variation is not in itself causative, but in Scotland it does correlate with a number of environmental variables which conceivably could be.

In a field trial of *P. sylvestris* in Scotland, James et al. (1994) observed a decrease in needle length with increasing altitude. Within our common garden needle length did show some correspondence with PC1, which is a good proxy for altitude, but the relationship received weak support in terms of AICc.

SLA has been reported to decrease with declining water availability (e.g. Cunningham et al., 1999, however we found no evidence of a relationship between SLA and precipitation at home sites. It may be that while eastern provenances can be said to be more xeric than western, they are insufficiently dry to result in selection for reduced SLA. Furthermore, SLA is known to be highly plastic in plants generally, and Scots pine is no exception (Xiao et al., 2006): this was also reflected in our results which displayed large inter-block variation. It is possible that in Scots pine, SLA may be determined predominantly by the prevailing local conditions, which in turn would make it a poor candidate for local adaptation.

3.4.4 Conclusions

The results demonstrate that significant heritable variation exists for at least one of the leaf traits examined, and therefore that some capacity to respond to selection. A notably large narrow-sense heritability was attributed to the number of resin canals per TS of needles (albeit accompanied by a large standard error), which could be an indicator of previous fluctuating selection for this trait. Within the common garden environment, few traits exhibited interpopulation differentiation, and in those that did differences were subtle. Functional leaf traits SLA, and LDMC, which are often used to characterise plant resource-use strategies (normally between species), were found to vary more dramatically between blocks than populations: this environmental sensitivity highlights the need for caution when interpreting data collected *in situ*, given the potential for uncontrolled variation between and within individual trees.

Despite the limited variability found between populations, it was possible to identify trends using the geographical coordinates and meteorological parameters of individuals' home sites. Specifically, populations in the east possessed fewer stomatal rows on average, which may indicate adaptation to lower water availability. It would be of some interest to discover if the same pattern could be observed in trees inhabiting native sites, to confirm whether the variation observed in the common garden is representative.

In this study we examined morphological variation between native Scots pine stands. The main finding is that populations differ in number of stomatal rows in a manner that suggests adaptation to water availability at the site of origin. If this is the case, then we would expect the populations to behave differently in terms of their water use efficiency under common garden conditions. Measures of carbon isotope discrimination are an accepted indicator of water use efficiency in plants (Farquhar et al., 1989). It would therefore be of particular interest to measure carbon isotope discrimination of

Scots pine populations in the common garden environment to confirm the adaptive importance of the differences in leaf anatomy that have been detected in this study.

Chapter 4

Phenological and morphological responses of Scotland's native pinewoods to flooding

4.1 Introduction

Phenological studies describe patterns of recurring events in nature, and the interplay between the timing of those events and conditions in the surrounding environment (Lieth, 1974). In boreal forests temperature is the predominant driver of phenological events (Murray et al., 1989; Kramer et al., 2000), and because of this phenological variation is often described in terms of latitude or altitude. Common garden trials are employed to evaluate genetic variation in phenology and growth within species; however, locally adapted variation may only become discernible when the environment is manipulated to impose stress on populations (Hoffmann and Merilä, 1999; Charmantier and Garant, 2005). In this study, we apply a flood treatment lasting several months to a common garden trial of native *Pinus sylvestris* (L.) populations sampled across a strong rainfall gradient in Scotland, each comprised of four half-sib families, and assess potential adaptive differences in flooding tolerance in terms of phenology and growth.

Ongoing climate change has stimulated considerable interest in forest phenology, with respect to how survival and productivity will be affected as global temperatures continue to increase (Rehfeldt et al., 2002). Growth should be optimal when the timings of events such as bud-flush and frost-hardening are in synchrony with local seasonality, enabling trees to capitalise on favourable conditions throughout the growing season without incurring damage during winter. Consequently, the importance of latitude is

widely recognised in temperate phenology: survival of material from southern provenances is reduced when transferred northwards, while in the south northern material performs poorly relative to native trees (Eriksson et al., 1980; Beuker, 1994; Persson, 1994). Elevation is similarly recognised as an important determinant in phenology, on account of its strong correspondence with temperature and growing season length (Ziello et al., 2009; Inouye and Wielgolaski, 2013).

Scottish populations of *P. sylvestris* are distributed across a relatively narrow range of latitudes (56–58°), but within this, sites can be found from sea level to around 500m in the Cairngorm Mountains. Salmela et al. (2013), have previously shown differences timing of in bud flush between native provenances under common garden conditions, and that intrapopulation variation tends to increase with altitude, possibly as a reflection of the greater climatic variability. Scotland is also subject to a highly pronounced rainfall gradient, whereby pinewoods on the west coast may receive in excess of 3,000mm annually compared with ∼850mm in eastern Cairngorm Mountains (Perry and Hollis, 2005). Coupled with this disparity in rainfall, western sites are often characterised by peaty soils with poor drainage, and are more susceptible to waterlogging.

Individuals within species may vary in their response to flooding (Kozlowski, 1997), and in crop plants genetic variation has been studied with a view to improving tolerance (e.g. Daugherty and Musgrave, 1994; Setter and Waters, 2003; Collaku and Harrison, 2005). Abiotic stresses can have a direct impact upon genetic variation, and stress exposure has been reported both to increase and decrease trait heritabilities (Hoffmann and Merilä, 1999; Charmantier and Garant, 2005). This may have relevance for tree populations in the context of climate change, as a presently 'untapped' adaptive capacity may exist within standing variation. Climate change models predict that Scotland will become warmer throughout the year, but as winters are expected to become wetter (Murphy et al., 2009), an understanding of flood-tolerance in native plants would be advantageous.

In pines, growth and shoot extension are known to be sensitive to water availability (Dougherty et al., 1994). Sands and Rutter (1959) were able to demonstrate that relatively small increases in soil moisture tension inhibited shoot and needle growth of *P. sylvestris* saplings during the growing season; similarly, stem diameter and root growth are strongly influenced by soil moisture and temperature (Bassett, 1964; Kaufmann, 1968; Linder et al., 1987; Cregg et al., 1988; Vapaavuori et al., 1992) . The term 'water-stress' is typically used in plant physiology to describe the dehydration induced by drought; perhaps counterintuitively an over-abundance of water can produce similar effects (Aroca et al., 2011), albeit tolerance to waterlogged conditions requires an alternative suite of adaptations (Kozlowski, 1997).

Flood-tolerance may be defined as a plant's ability to continue growing under flooded conditions, which is largely dependent on the maintenance of a functional root system. Oxygen diffuses around 10,000 times more slowly in water than in air (Colmer, 2003), and during flooding air pockets in soil are displaced resulting in hypoxia. Survival of root tissue is adversely affected by a lack of aeration, and furthermore by the toxic products of anaerobiosis (Ponnamperuma, 1972; Armstrong et al., 1994; Kozlowski, 1997). Because of this, adaptations that serve to increase conductivity of oxygen to root tissues may improve their resilience to flooding. 'Tissue porosity' describes the ratio of air space to plant tissue: the potential for gas transport within plants is greater among those with higher tissue porosities, a property determined during development (Jeffree et al., 1986). This capacity for gas transport may be augmented by the formation of aerenchyma, a tissue type characterised by large intercellular spaces (Pallardy, 2008); these tissues have been shown to form in response to an increase in ethylene production, stimulated by oxygen deficits in flooded plants (Kawase, 1979). Aerenchyma are not reported to occur in *P. sylvestris*, however Armstrong and Read (1972) found evidence of gas transport from the aerial parts of seedlings to the roots when submerged in a deoxygenised medium. Ethylene production is increased in flooded plants (Kozlowski, 1984), and in *Pinus densiflora*, Yamamoto and Kozlowski (1987) found that a high ethylene content coincided with increased stem diameter growth.

Other morphological responses to flooding include the formation of hypertrophied lenticels or adventitious roots, also associated with an increase in the production of ethylene (Tang and Kozlowski, 1984). Lenticels forming on submerged stems act as pores to enable gas exchange with flood water (Hook et al., 1971; Kozlowski, 1997), and may serve to release toxic compounds produced by anaerobiosis (Chirkova and Gutman, 1972). Adventitious roots replace some of the lost function caused by damage to the original root system (Kozlowski, 1984), and develop on submerged stems or on older roots; although observed in Scots pine (e.g. Hansen and Ernstsen, 1982), adventitious roots have not been reported in response to flooding.

The capacity of populations to adapt is underpinned by the availability of genetic variation on which selection may act; however, estimates of genetic components are valid only for the environments in which they are estimated (Falconer and Mackay, 1996). Cryptic genetic variation is so-named to describe the diversity that may exist within the genepool, but is ordinarily unobservable (see Gibson and Dworkin (2004) for review). It has been posited by some authors that stress or environmental novelty may serve to 'release' this cryptic variation (Rutherford and Lindquist, 1998; Schlichting, 2008), enabling adaptive evolution to proceed. However, the application of stress has been found to both increase and decrease genetic variability (Hoffmann and Merilä, 1999), and the

direction of the effect cannot be regarded as separable from the nature of the stress or the trait under consideration.

In this chapter, local adaptation in native Scots pine populations is evaluated in the context of plants' ability to sustain shoot, needle, diameter, and root growth under flooded conditions, and whether bud flush is inhibited or delayed in individuals from drier provenances. In view of ongoing environmental change, we also attempt to ascertain whether flood stress serves to increase or otherwise effect heritable genetic variation, and therefore the potential for selection.

4.2 Methods and Materials

4.2.1 Study Provenances

Seed was collected from nine of the native Scottish pinewoods across a strong rainfall and altitudinal gradient (fig. 4.1 & table 4.1). Within each site seed was collected from four open-pollinated mother trees, taken across an altitudinal gradient at each location. The collection therefore consisted of 36 families in total (9 Populations $\times$ 4 Families), and individuals within families were regarded as half-sibs.

4.2.2 Experimental Design

Seed was sown in June 2007, and saplings were around 3.5 years old when phenological data collection began in spring 2011, having recently been transferred to 1.5l pots. The trial consisted of 432 individuals in total, incorporating 12 individuals from each of the 36 families: plants were distributed across 12 plastic basins (hereafter 'plots'), such that each held 36 individuals (one member from each family). Half of the plots were assigned to the flooding treatment, and the remainder to the control; the placement of treatment and control plots was alternated throughout the glasshouse.

Small overflow holes were made at intervals around the perimeter of all plots: in the control plots these holes were made $\sim$1cm from the base of the containers to permit a shallow maximum water depth, while in the flooded plots these perforations were made substantially higher to maintain a water depth of $\sim$2cm below soil surface. The waterlogging regime began on the 5th of July 2011, following which the water levels across all plots were maintained at around their respective maximums by an automated watering system.

Figure 4.1: Map of native provenances sampled for use in the study, including site of common garden trial.

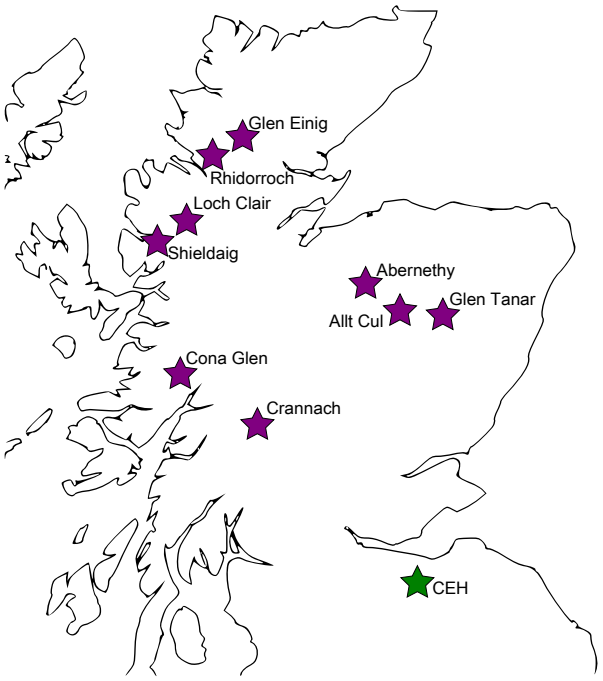


Table 4.1: Coordinates of trial populations, the range of altitudes at which the mother trees were sampled, alongside mean monthly temperature and rainfall taken from Met Office estimates.

Provenance	Latitude	Longitude	Altitudinal Range (m a.s.l.)	Mean Monthly Temp (°C)	Mean Monthly Rainfall (mm)
Abernethy	57.21	−3.61	365–363	6.82	67.66
Allt Cul	57.04	−3.35	435–512	6.44	79.85
Cona Glen	56.79	−5.33	89–180	8.45	181.73
Crannach	56.58	−4.68	258–338	6.10	211.07
Glen Einig	57.95	−4.76	45–69	7.04	111.18
Glen Tanar	57.05	−2.86	293–422	7.40	70.91
Loch Clair	57.56	−5.36	102–166	8.07	234.94
Rhidorroch	57.89	−4.98	138–220	8.45	139.39
Shieldaig	57.51	−5.64	44–132	7.60	198.71
CEH	55.86	−3.21	190	7.08	77.73

4.2.3 Phenological Recording

Throughout spring 2011 (prior to flooding), plants were scored on a weekly basis and individuals were identified as having burst bud when needle tips had visibly emerged from their sheaths on the terminal apical bud. If observation was not possible on a given day, scores recorded on the prior or following day were substituted. The trial was scored a second time during spring 2012, by which time the flood treatment had been in effect for around eight months.

For the purpose of analysis, phenology scores were adjusted to a common date: in both years scores were expressed relative to the 1st of April, which was assigned a value of '1'.

4.2.4 Morphological Variation

4.2.4.1 Needle Length

During June-July 2011, estimates of needle length were made for all individuals on previous-year whorls (i.e. growth that developed in 2010). Needle length was determined by scanning five needles taken from each plant, and performing measurements via *MacBiophotonics ImageJ* (Abramoff et al. (2004)) to obtain mean values.

Needle length was estimated a second time in November 2012 on current-year needles (those which emerged and extended during the flood trial) on a subsample of 216 individuals.

4.2.4.2 Apical Extension and Diameter Growth

In Autumn 2012, new apical shoot growth was measured to the nearest 0.5cm, and stem diameter was recorded using digital calipers at two points on the main stem: 'upper diameter' immediately above the emergence of the most recent whorls (2012), and 'basal diameter', above whorls produced in 2010 (the lowest point accessible across all plants due to sporadic branching close to the soil surface).

4.2.4.3 Destructive Sampling

A subsample consisting of 72 individuals (50% Control, 50% Flooded) was harvested in Autumn 2012 to enable destructive measurements to be carried out. Six of the nine populations were chosen (Abernethy, Allt Cul, Crannach, Glen Tanar, Loch Clair, and Shieldaig), and within each two half-sib families were selected, from which six plants were collected ($6 \times 2 \times 6 = 72$).

Plant roots were rinsed, and root collar diameter was recorded at the lowest point on the stem before root branching occurred. Individuals were then dried in a desiccating oven at $\sim 80°\mathsf{C}$ for a period of at least three days. Remaining excess soil was removed from root material, needles were detached from the remainder of the plant, allowing individual dry mass estimates to be obtained for roots, shoots, and needles. The whorls and apical shoots that emerged in 2012 and extended during the flood treatment were

weighed separately, and are referred to in the results as 'new' needle and shoot mass. Prior to drying, harvested plants were kept in cold storage at ∼4°C.

4.2.5 Analysis of Treatment and Population Differences

The importance of treatment and population effects were evaluated via mixed models. The following global model was fitted via ML:-

$$\begin{aligned} Trait = \mu + Treatment_i + Population_j + Treatment : Population_{ij} \\ + Family_{k(j)} + Plot_{l(i)} + \epsilon_{ijkl} \end{aligned} \tag{4.1}$$

where Treatment, Population, and their interaction were considered as fixed effects; Family within Population, and Plot within Treatment were random effects. Model subsets were generated via 'MuMIn' (Barton, 2013), and the best models were defined as those possessing the lowest AICc. Linear regression was performed by including relevant predictors as fixed effects, and including Population as a random effect.

4.2.6 Estimation of Genetic Components

The genetic coefficient of variation $\mathrm{CV_A}$ (Houle, 1992) is a measure of additive genetic variability normalised by the trait mean, and was also estimated pooling across populations within treatment groups by the following:-

$$\mathrm{CV_A} = \frac{\sqrt{V_\mathrm{A}}}{\mu_\mathrm{trait}} \times 100 \tag{4.2}$$

V_A is additive genetic variance (for half-sibs this is four times the family variance), and μtrait is the mean trait value. $\mathrm{CV_A}$ was estimated for all traits excluding those measured via destructive sampling due to the limited sample size.

4.3 Results

A summary of model testing results can be found in table 4.2. All ΔAICc values quoted in the text describe the AICc values of the null models relative to those with the lowest (i.e. 'best') values.

4.3.1 Phenology

Phenological data was collected in both 2011 (when all plants were untreated), and 2012 after half of the trial had been exposed to prolonged flooding. The earliest bud-burst recorded in 2011 was on April 22, and the latest on June 24. Mean scores were highly similar between the treatment and control groups (May 21±1.20) , however interpopulation differences were observed in both (fig. 4.2). As these population differences were inconsistent between treatment groups, a model incorporating a treatment by population interaction term received the strongest support by AICc (ΔAICc=14.96);the 'next-best' model in terms of AICc included Population, but not Treatment as a fixed effect. Although interpopulation differences were detected, there was no evidence for a relationship with latitude, longitude, or altitude.

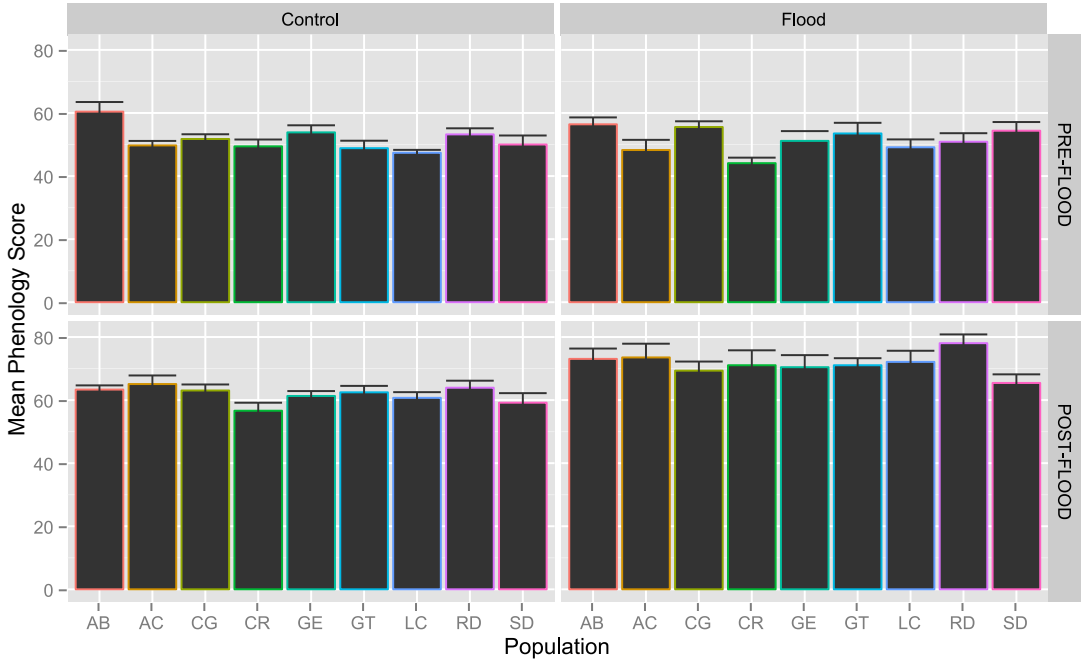
Phenology data was collected a second time during spring 2012, after the flooding treatment had been in effect for around eight months. The first bud-burst was recorded on March 29, however this was an outlier, and bud-burst occurred later on average than 2011 in both control and treatment groups (fig. 4.2). Mean bud-burst dates were the May 31(±0.91) and the June 10(±1.15) for control and treatment groups respectively. Model testing provided very strong evidence for a treatment effect (ΔAICc=25.71), but models containing population and interaction terms were not competitive. The latest bud-burst events were recorded on July 25 for three individuals, all of which belonged to the flood group. By spring 2012 two individuals had died (i.e. foliage had become brown), and a further eight failed to burst bud. All of these individuals belonged to the flood group; although too few to permit statistical evaluation, both of the dead plants and three of the bud-burst failures originated from the Abernethy provenance.

4.3.2 Morphology

4.3.2.1 Needle Length

No evidence was found for treatment or population effects in the 2011 pre-flood sample; individual means ranged from 23.40–101.48mm with an overall mean of 50.07(±1.31)mm fig. 4.3. Post-flooding (2012) both control and treatment groups possessed needles substantially shorter than those collected in 2011, with means of 27.1(±0.69) and 18.2(±0.84)mm respectively. Model testing results found strong support for the treatment effect with ΔAICc=20.76, but no evidence was found for a population effect.

Figure 4.2: Mean phenology scores in 2011 and 2012 (prior to and following initiation of the flood treatment, respectively). In 2011, bud-burst dates were found to vary between populations and a treatment × population interaction was identified as these differences were not consistent between treatment groups (though treatment means were highly similar). Post-flooding, the control group burst-bud earlier on average than the flooded group, however no population effects were detected. Bars are coloured by population and errorbars represent ±1 SE, and were estimated for individual treatment groups by REML.



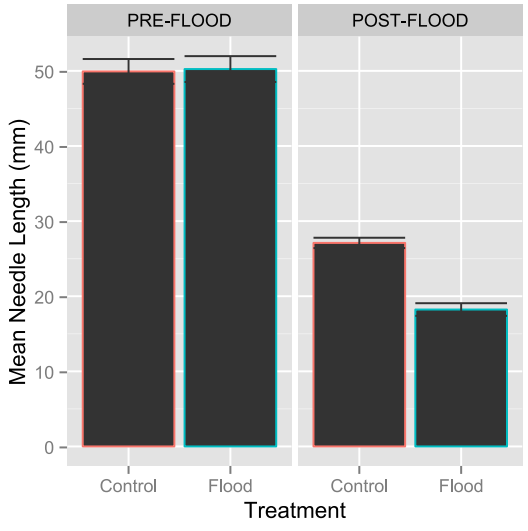


Figure 4.3: Mean length of needles collected prior to (2011) and following (2012) the flood treatment for a subsample of 216 individuals. Overall mean length was lower for needles which developed during flood trial, and growth in the flooded group was impeded relative to the control. Errorbars represent ±1 SE, and were estimated for individual treatment groups by REML.

4.3.2.2 Apical Extension and Diameter Growth

Excluding those plants which failed to burst bud, apical growth varied substantially from 3.00 to 38.50cm. Growth of individuals in the treatment group was impeded, with a mean growth of 13.93(±0.53)cm compared to 22.24(±0.74)cm in the control (fig. 4.5a). Model testing indicated support for a treatment effect only (ΔAICc=26.93) (table 4.2).

Stem diameter was measured at two positions on each plant: one 'basal', above the 2010 whorl, and one 'upper' above the 2012 whorl (produced during the treatment). Stem diameter between positions was positively correlated (β =0.31(±0.02), Akaike weight =1)(fig. 4.4), and the diameter of new stem growth was greater in individuals which were broader at the base. Evidence was found for a population effect at the basal position (ΔAICc=3.82); those with the greatest average diameter belonging to individuals from Allt Cull and Abernethy (12.23±0.55 and 12.18±0.53mm, respectively), and the narrowest from Crannach, Rhidorroch, and Loch Clair (10.4–10.8mm). Both treatment and population effects were detected at the upper position (ΔAICc=18.71)(fig. 4.5c). Mean apical stem diameter was 5.12(±0.13)mm in control plants compared with 4.20(±0.09)mm in the flooded group; averaged across treatment groups, plants from Allt Cul possessed the highest diameter (5.26±0.42mm), and Rhidorroch the lowest (4.42±0.17mm).

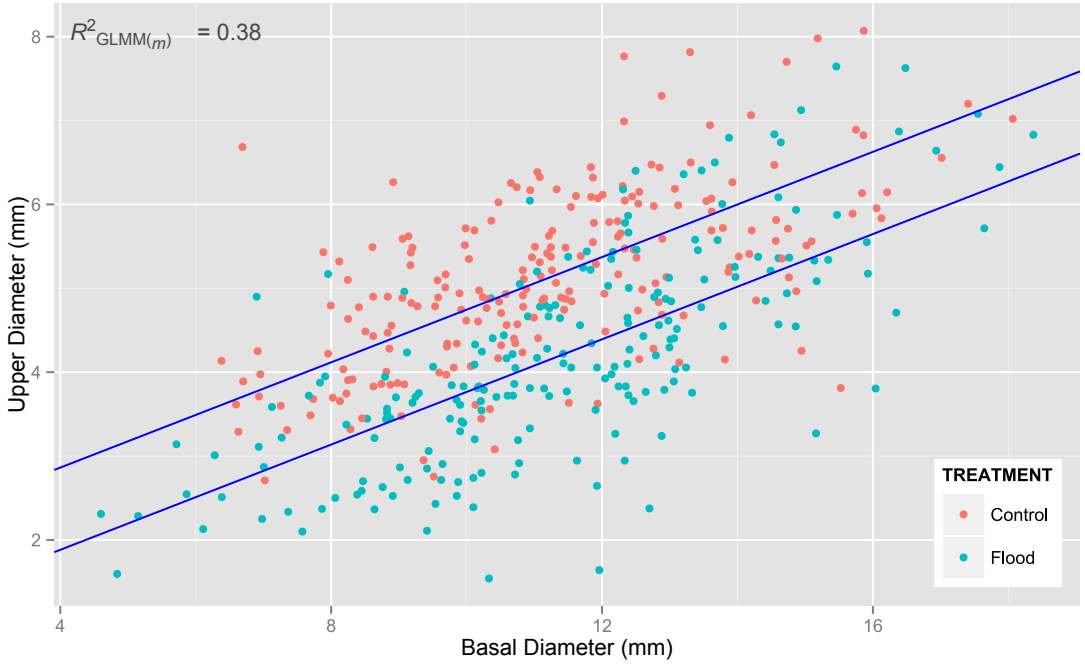
Table 4.2: Summary of best ANOVA candidate models by AICc. Model subsets were generated and compared by AICc to assess the importance of Treatment, Population, and their interaction as fixed effects. The best models were determined by AICc, and the presence or absence (+/-) of fixed effects are indicated for each of the traits listed. ΔAICc represents the difference in AICc between the best model and the null model (a value of zero indicates the null model was the best). The proportion of total variance due to Family and Plot effects are shown for each.

Trait	Mean Control	Mean Flood	Treatment	Population	Treat × Pop	df	ΔAICc	Akaike Weight	Variance due to Family (%)	Variance due to Plot (%)
Phen Score 2011	60.38	56.47	+	+	+	21	14.96	0.89	9.28	2.74
Phen Score 2012	61.76	71.41	+	-	-	5	25.71	0.76	7.50	0.00
Needle Length 2010 (mm)*	49.93	50.29	-	-	-	4	0.00	0.72	18.48	0.00
Needle Length 2012 (mm)*	27.10	18.23	+	-	-	5	20.76	0.99	2.98	0.00
Apical Extension (cm)	22.24	13.93	+	-	-	5	26.93	0.91	1.40	1.21
Stem Diameter Low (mm)	11.21	11.39	-	+	-	12	3.82	0.59	2.59	0.08
Stem Diameter High (mm)	5.12	4.20	+	+	-	13	18.71	0.53	1.26	1.23
New Needle Dry Mass (g)†	4.99	2.70	+	-	-	5	8.21	0.85	0.00	0.00
New Shoot Dry Mass (g)†	4.09	2.84	+	-	-	5	1.35	0.64	0.00	1.32
Root Collar Diameter (mm)†	12.52	13.79	+	-	-	5	1.10	0.58	0.52	0.00
Root Dry Mass (g)†	11.70	10.25	-	-	-	4	0.00	0.60	3.95	0.00

*subset of 216 individuals

†subset of 72 individuals (destructive measurements)

Figure 4.4: Apical stem diameter of growth produced during flood treatment in 2012 (high), regressed upon stem diameter at a lower point on the plant (above the whorl produced during 2010). Stem diameters were positively correlated (β=0.31), and a treatment effect was detected in the diameter growth produced during flooding.



4.3.2.3 Destructive Measurements

A strong treatment effect was observed in the new needle mass developed between treatment groups (ΔAICc= 8.21), and flooded individuals accumulated 2.26($\pm$0.55)g less dry mass on average (fig. 4.6a). Similarly, new dry shoot mass was an average of 1.24($\pm$0.66)g lower in the flooded group (fig. 4.6b), though evidence for the treatment effect was less strong (ΔAICc= 1.35).

There was some evidence for a small increase in root collar diameter in the treatment group relative to the control (ΔAICc= 1.10); flooded plants had mean of 13.79($\pm$0.71)mm, compared with 12.52($\pm$0.49)mm in the control (fig. 4.6c). Dry root mass was 11.00($\pm$0.69)g on average: mass in the treatment group was lower than the control (fig. 4.6d), however the effect was insufficient to be supported by AICc (the null model being more parsimonious). In the destructive sample, no evidence was found for population or interaction effects.

Figure 4.5: Treatment effects on apical extension and stem diameter traits. Evidence was found for a treatment effect for apical extension (a), and upper diameter (c), but not basal diameter (b); however stem diameters at both heights varied by population. Errorbars represent ±1 SE, and were estimated for individual treatment groups by REML.

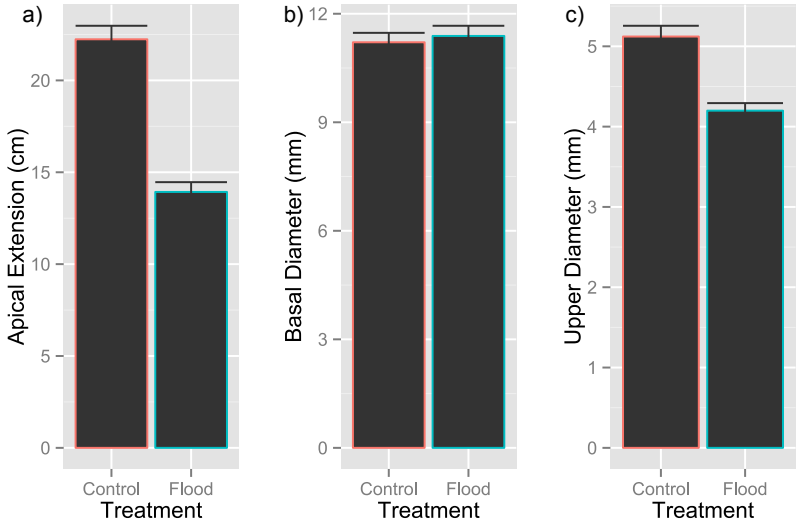


Figure 4.6: Treatment effects for traits measured on harvested plants. Model testing by AICc found evidence of a treatment effect for leaf dry mass (a), shoot dry mass (b), and root collar diameter (c), but not for root dry mass (d). Errorbars represent ±1 SE, and were estimated for individual treatment groups by REML.

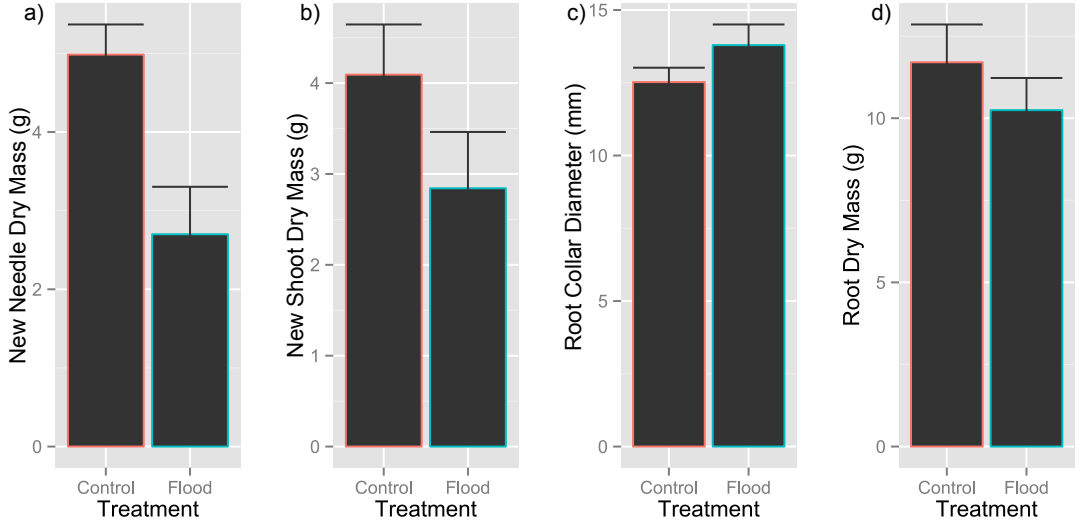
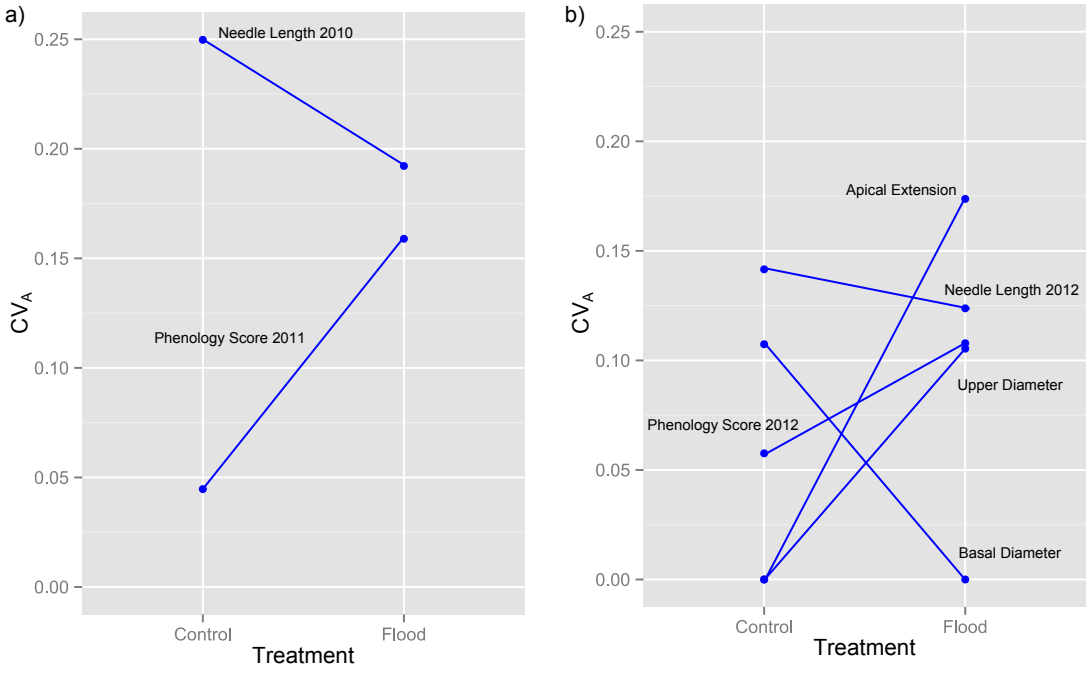


Figure 4.7: Treatment group estimates of CV_A for traits measured prior to (a), and after the flood treatment had been in effect for a prolonged period (b). Estimates between groups differed prior to flooding, and there was no consensus in the direction of the effect post-flooding.



4.3.3 Genetic Variation

Family variances typically accounted for <5% of the total variation for each trait, with a few exceptions (table 4.2). The family component accounted for a similar proportion of the total variance in phenology scores across both years 9.28% in 2010, and 7.50% in 2011; and a larger proportion, 18.48%, for needle length. For all traits examined, the majority of variation was unexplained.

Differences in CV_A between control and treatment groups are presented in fig. 4.7. CV_A values ranged from around 0 to 25%, and there was no consistent shift up or downward between flooded and non-flooded plants. Despite measurements being collected prior to the commencement of flooding, treatment groups differed between 5 to 11%. Post flooding, the traits which describe new vertical growth, apical extension and upper diameter, were 10 to 17% higher in the treatment group than in the control (though basal diameter was lower); phenology scores and needle lengths followed the same trends as pre-treatment measurements, but the differences between treatment groups were less pronounced.

4.4 Discussion

4.4.1 Phenological Variation

Evidence of local adaptation is often found in the timings of phenological events, and a relationship between the spring phenology native Scots pine populations and the length of the growing season at home sites has been previously demonstrated (Salmela et al., 2013). Bud flush was first recorded in 2011, prior to the commencement of the flood treatment, and yet rather paradoxically strong support was found for a treatment by population interaction. The timings of bud flush were highly comparable between treatment groups, and this effect seems attributable to interpopulation differences which were distributed differently in each group. It is unclear as to why this occurred, given that treatment groups were interspersed throughout the glasshouse under identical conditions, and the variance attributable to plots within treatments was very low. Differences between populations could not be attributed to environmental variability between home sites; it's possible that a larger sample of populations is required to observe a trend, or that trait expression within the glasshouse is not necessarily consistent with that on native sites.

Following application of the treatment in spring 2012, average bud flush occurred later overall than in 2011, and was delayed in the flooded group relative to the control. However, unlike 2011, no differences were observed between populations. It may be that flood stress at this level of severity serves to suppress interpopulation variability, at least in terms of phenological differences, however there was also an absence of differentiation within the control. Prior to the treatment plants rested upon irrigated matting, before being transferred to containers allow water depth to be manipulated; although water depth was maintained at a low level in the control plots, the increase in soil moisture may nevertheless have had an effect on phenotype.

4.4.2 Morphological Variation

The needles produced by flooded plants were significantly shorter than those in the control; it seems likely this was due to reduced water availability, which has previously been shown to impact upon needle length in *P radiata* (Linder et al., 1987), and possibly also due to a delay incurred in growth initiation. There was also a large disparity in needle lengths between 2010 and 2012. The cause of this reduction in length between years is unclear: it may be that the difference in watering conditions remarked upon above imposed some degree of stress upon the control plants (albeit substantially less

than those flooded); alternatively, the available soil nutrition may have been more limited relative to the plant biomass in 2012 than in 2010, thereby restricting needle growth potential. It's also possible that the difference in needle length between years was due to an age-effect.

The sensitivity of plant growth to water-stress is well-known (Kozlowski, 1997), and flooding markedly inhibited terminal shoot extension. Height growth was on average 36% lower in the flooded group, likely as a result of root damage incurred by prolonged waterlogging. Despite this pronounced treatment effect, Scottish populations did not vary substantially in their responses, and all were similarly inhibited. Cregg et al. (1988) found that rates of basal diameter growth were closely related to soil water potential in *P taeda*, and in our trial there was evidence that stem diameter growth was affected by flooding. Similarly to shoot extension, diameter growth of new shoots was inhibited by flooding, however there was no apparent effect on diameter growth at the lower point of the stem; possibly as any differences that may have been caused by the treatment were buffered by previous years of growth. Although no interactions were identified between treatment and population terms, stem diameters differed between populations at both heights measured. In a previous trial incorporating four native pinewoods Perks and Mckay (1997), found that individuals from Loch Maree (on the west coast) allocated on average less mass to woody structures than other populations. In our trial, individuals from Loch Clair and Rhidorroch exhibited among the lowest stem diameters: these sites are nearby Loch Maree, and so this may reflect a genetic difference in stem form.

Flooded individuals accumulated less mass than the control group in terms of new needle and stem growth, but there was no evidence of a difference in root mass between groups. This seems rather surprising, given that it is the roots that are directly exposed to flooding, and that responses observed throughout the rest of the plant are likely attributable to root damage. The removal of compacted soil from roots resulted in the inevitable loss of fine root material, which would have been most susceptible to flood damage; mass measurements were therefore taken from the remaining (and more resilient) woody material which had accumulated for several years prior to the treatment. Roots of flooded plants were more readily freed from the soil, and mean root mass was lower in the flooded group; statistical support for a difference between groups may have been lacking on account of the restrictive sample size. Lenticels were relatively ubiquitous on the roots of both flooded and control plants, but adventitious roots were not observed. The difference in root collar diameter between treatment groups is of interest as, unlike the other growth parameters, diameters were greater on average among the flooded group. In some plants flooding is known to promote the formation of internal air spaces for the purposes of gas exchange (Yamamoto et al., 1987); although the effect

was relatively weak, it's conceivable that this is also true in *P. sylvestris* and that the differences observed were the results of a plastic morphological response.

4.4.3 Impacts of Stress on Genetic Variability

Exposure to stress has been put forward as a mechanism by which cryptic genetic variation may be released (Hoffmann and Merilä, 1999; Badyaev, 2005; Gibson and Dworkin, 2004), though the evidence base for this has been questioned (McGuigan and Sgrò, 2009). It is difficult to draw direct conclusions from our results as, in the absence of any kind of manipulation, CV_A estimates were observed to differ between groups. In these cases group means were similar, but variance structures were not; this could have been either due to a genuine difference between groups (of unknown cause), or because sample sizes were not sufficiently large to produce accurate estimates for the traits evaluated. Post-flooding there was no consensus in the direction of the effects of flood-stress on CV_A. The traits describing new vertical growth (apical extension and upper diameter) displayed an increase in heritable variation under flooding: in trees height growth is often employed as a proxy for fitness, and we might tentatively infer that evolvability in growth is increased as a result of flood stress, however, these traits are non-independent and must in any case be interpreted with caution given the variation in the pre-treatment results. The CV_A for basal stem diameter was lower in the flooded group than the control, but unlike the latter traits no evidence was found for a mean difference in basal diameter between groups.

The responses of genetic variability to stress are often made by contrasting extreme environments (Hoffmann and Merilä, 1999); whereas a continuum of increasingly stressful conditions would be more desirable, as conceivably potential effects on variability might be greater at low or moderate stress levels. Our trial incorporated only two treatment levels, with the flooding effect intended to be relatively severe; the inclusion of further groups in a similar trial may require a prohibitively large sample size.

4.4.4 Conclusions

Overall, flooding had a pronounced effect on Scots pine saplings; delaying bud flush and reducing growth on average. Nevertheless, native populations were not found to respond differentially to flooding for any of the traits examined, however some differences in spring phenology and stem diameter were observed independently of the treatment. The effects of flood stress on genetic variation were not easy to interpret, especially given the disparity between groups in the absence of any kind of treatment; similar experiments may benefit from a pre-treatment analyses of variance components in order

to avoid drawing potentially spurious conclusions from post-treatment results, particularly if sample sizes are not large. There were signs that heritable variation for apical growth may increase as a consequence of flooding, but this was against a background of conflicting results.

What is notable is that given the relatively severe and prolonged exposure to flooding, there was very little mortality; in fact, despite sizeable average differences between treatment groups many flooded plants showed no conspicuous ill-effects. *P. sylvestris* saplings then, appear to be relatively resilient to soil waterlogging over a long period of time, at least in the absence of other stressors. The phenological and morphometric data examined are useful and straightforward to obtain, but nevertheless limits research to surface-level traits. In the following chapter we attempt to develop a greater insight into the effects of flood-stress on native pinewoods by tracking underlying physiological responses, and probe deeper for evidence of local adaptation.

Chapter 5

Differential responses of Scotland's native pinewoods to flooding detected via rapid physiological phenotyping

5.1 Introduction

The remnants of the Caledonian Pine Forest represent the westernmost frontier of the Eurasian distribution of *Pinus sylvestris* (L.), separated from mainland populations by a distance of at least 500km Wachowiak et al. (2011). Subject to the warming influence of the Gulf Stream, the climate of Scotland is atypically mild in comparison with other countries at similar latitudes. Within Scotland, however, the climate varies substantially: due to a steep rainfall gradient native pinewoods on the west coast receive several times the annual precipitation of those in the Cairngorm mountains to the east. Local adaptation may occur when genotypes are differentially selected across environments (Kawecki and Ebert, 2004); here we investigate several sensitive physiological traits for a number of populations within a common garden, to assess whether water use efficiency (WUE) covaries with precipitation rates at sites of origin, and determine whether tolerance to soil waterlogging differs among Scotland's native pinewoods.

An insight into the effects of stress on plant photochemistry may be obtained by measurement of chlorophyll fluorescence. Photosynthesis is not 100% efficient: a proportion of the light energy received by chlorophyll molecules may be re-emitted as light at a longer wavelength (fluorescence), and the remainder is dissipated as heat. By measuring fluorescence yield it is possible to estimate efficiency of Photosystem II (PSII) (Genty et al., 1989), and gauge the extent to which photoinhibition has been incurred in response to stress (Maxwell and Johnson, 2000). This 'quantum efficiency' is often described in

terms of the fluorescence parameter *Fv/Fm*, and has been used in numerous studies to evaluate the effects of a variety of stresses on photochemical efficiency, including high (Gamon and Pearcy, 1989) and low (Groom and Baker, 1992) temperatures, freezing injury (Lindgren and Hällgren, 1993), and chemical pollutants (Strand, 1993). Tolerance to water stress has also been investigated by means of chlorophyll fluorescence, chiefly in terms of drought (Demmig et al., 1988; Huang and Gao, 1999; Garg et al., 2002; Salmela et al., 2011a), but also waterlogging (Pearson et al., 2013); in all cases a decline in photochemical efficiency was reported in response to stress exposure. As a technique, chlorophyll fluorescence is particularly appropriate for use in monitoring common garden trials, as a large number of samples may be processed non-destructively in a short time.

The majority of plant water loss occurs via stomata, and exposure to stresses which affect water relations can impact upon stomatal behaviour. Rapid closure of stomata in response to flooding has been reported in many species, following which photosynthesis is impeded (e.g. Bradford and Hsiao, 1982; Kozlowski, 1997; Yordanova et al., 2005). The degree of stomatal closure can be ascertained by measuring stomatal conductance (g_s); essentially the rate with which water passes via the stomata to the environment. When flooding is prolonged, species intolerant to waterlogging may lack the facility to reopen stomata following initial closure (Pallardy, 2008); stomata of the birch *Betula papyrifera* failed to reopen after 14 days of exposure to flooding (Tang and Kozlowski, 1982), whereas those of the red ash *Fraxinus pennsylvanica* were able to do so following a period of closure (Gomes and Kozlowski, 1980). A study investigating the effects of waterlogging on two *Pinus* species, *P. caribaea* and the more flood tolerant *P. elliottii*, Lewty (1990) found that the latter maintained higher stomatal conductance under flooded conditions, and the onset of stomatal closure was delayed relative to *P. caribaea*. By measuring stomatal conductance it may be possible to assess a plant's ability to maintain productivity under flooded conditions, as well as any subsequent acclimation or recovery following initial exposure.

Evidence of local adaptation with respect to water usage may also be found by analyses of isotopic signatures. The ratio between the two stable isotopes of carbon, ^{12}C and ^{13}C, present in leaf or other tissues is understood to be an effective proxy for WUE, and has been used routinely in selection for improved WUE in crop plants (e.g. Ehdaie et al., 1991; Condon and Richards, 1992; Impa et al., 2005). In atmospheric CO_2 the ratio of ^{12}C:^{13}C is approximately 99:1, however in plant tissue ^{13}C is present in lower abundance (this proportion is usually reported relative to a standard as $\delta^{13}C$). The reasons for this are twofold: firstly the heavier $^{13}CO_2$ diffuses more slowly through stomata and into tissues, and secondly the carboxylating enzyme Rubisco preferentially fixes $^{12}CO_2$ (Hall et al., 1993). A reduction in stomatal aperture reduces water loss via

transpiration more so than it impedes CO_2 transfer, and as such WUE can be increased. One consequence of this lower respiration rate is the enrichment of internal spaces with $^{13}CO_2$, and subsequently increased uptake (Pallardy, 2008); plants operating at higher WUE over time are therefore expected to be relatively less depleted in ^{13}C (represented by a less negative $\delta^{13}C$). It should be noted that adaptation to low water environments may not always necessitate high a WUE: for example, avoidance of water stress might be achieved via a higher root/shoot ratio Warren et al. (2001). Nevertheless, trends in $\delta^{13}C$ have been observed across rainfall and soil moisture gradients *in situ* (Ehleringer and Cooper, 1988; Stewart et al., 1995; MIller et al., 2001), whereby plants in areas which experience lower water availability exhibit higher $\delta^{13}C$ values (and by proxy, greater WUE). Average $\delta^{13}C$ is also reported to increase with altitude (Körner et al., 1988; Hultine and Marshall, 2000), possibly in response to reduced humidity (Panek and Waring, 1997), barometric pressure (Marshall and Zhang, 1994), or in conjunction with other differences in leaf morphology and composition which show a correspondence with altitude (Vitousek et al., 1990). In Scotland, the native pinewoods that experience the lowest mean rainfall are also those at the greatest elevation: we anticipate that these may exhibit exhibit less negative $\delta^{13}C$ than those inhabiting the opposite end of these gradients. Previously, Brendel et al. (2002) measured wood cellulose $\delta^{13}C$ for five native Scottish pinewoods *in situ*, and identified two groups: a west coast group with more negative $\delta^{13}C$, and a second group composed of the remaining populations with less negative $\delta^{13}C$. Additionally, $\delta^{13}C$ was found to covary with potential water deficit: pinewoods which were likely to experience greater precipitation than evaporation on average exhibited more negative $\delta^{13}C$, than those in which evaporation may exceed precipitation.

The application of stress has been reported both to increase and decrease heritable variation (e.g. Hoffmann and Merilä, 1999), and therefore the potential for selection. The direction of the effect (if any) seems likely to be dependent on the trait in question, as well as the nature and severity of the stress applied. Should heritability be influenced by environmental manipulation, we predict that populations frequently subject to a given stress may not only possess greater tolerance, but also exhibit lower variability in response due to the action of past selection. Conversely, the potential for selection may be greater in populations for which the conditions imposed represent greater novelty.

In this trial, we expose a common garden trial of Scots pine saplings derived from the seed of native pinewoods from sites around Scotland to a prolonged flood treatment, and gauge responses over time in terms of photochemical efficiency and stomatal conductance, both of which are known to be sensitive to plant water balance. We also evaluate intrinsic WUE by means of carbon isotope discrimination both prior to and following the flood treatment, and determine whether any potential differences observed

may be explained by environmental variation between home sites. We hypothesise that material from the western, more maritime sites is genetically more predisposed to sustain productivity under flooded conditions relative to that originating from the drier, high altitude sites in the east. In addition, we estimate genetic variance components, test how exposure to the treatment influences the potential for selection, and address whether heritable variation is higher among populations where flooded conditions are less common.

5.2 Methods and Materials

5.2.1 Study Provenances

All material was grown from seed within the glasshouse, and consisted of samples taken from nine native Scottish populations of *P. sylvestris* across a strong rainfall gradient (fig. 5.1, table 5.1). At each site seed was collected from four open-pollinated trees at varying altitudes; population samples were therefore comprised of four maternal families, and within each individuals were regarded as half-sibs.

5.2.2 Experimental Design

Seed was sown in June 2007 and saplings were around four years old at the time treatment commenced. The trial consisted of 432 individuals in total, incorporating 12 individuals from each of 36 families: plants were contained within 1.5l pots which were distributed across 12 plastic basins (hereafter 'plots'), such that each held 36 individuals (one member from each family). Half of the plots were assigned to the flooding treatment, and the remainder to the control; the placement of treatment and control plots was alternated throughout the glasshouse.

Small overflow holes were made at intervals around the perimeter of all plots: in the control plots these holes were made ∼1cm from the base of the containers to permit a shallow maximum water depth, while in the flooded plots these perforations were made substantially higher to maintain a water depth of ∼2cm below soil surface. The waterlogging regime began on the 5th of July 2011, following which the water levels across all plots were maintained close to their respective maxima by an automated watering system.

Figure 5.1: Map of native provenances sampled for use in the study, including site of common garden trial. Sites are coloured with respect to the group to which they were assigned in the cluster analysis of meteorological data described in Chapter 3.

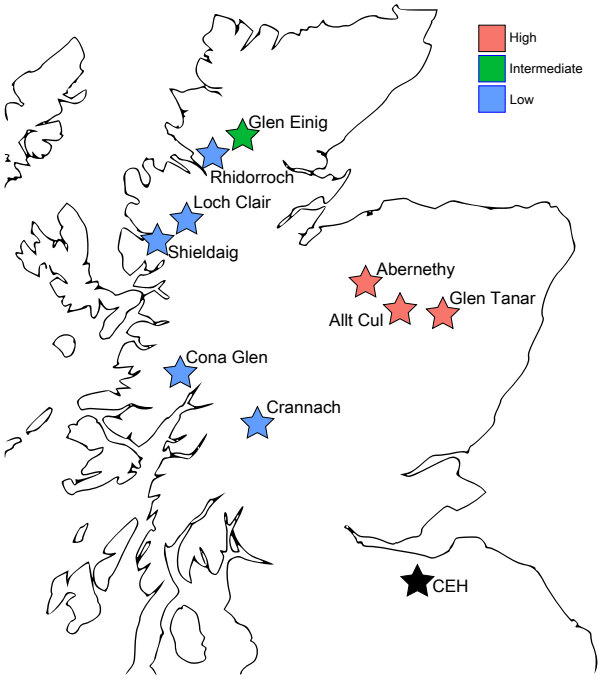


Table 5.1: Coordinates of trial populations, the range of altitudes at which the mother trees were sampled, alongside mean monthly temperature and rainfall taken from Met Office estimates.

Provenance	Latitude	Longitude	Altitudinal Range (m a.s.l.)	Mean Monthly Temp (°C)	Mean Monthly Rainfall (mm)
Abernethy	57.21	−3.61	365–363	6.82	67.66
Allt Cul	57.04	−3.35	435–512	6.44	79.85
Cona Glen	56.79	−5.33	89–180	8.45	181.73
Crannach	56.58	−4.68	258–338	6.10	211.07
Glen Einig	57.95	−4.76	45–69	7.04	111.18
Glen Tanar	57.05	−2.86	293–422	7.40	70.91
Loch Clair	57.56	−5.36	102–166	8.07	234.94
Rhidorroch	57.89	−4.98	138–220	8.45	139.39
Shieldaig	57.51	−5.64	44–132	7.60	198.71
CEH	55.86	−3.21	190	7.08	77.73

5.2.3 Chlorophyll Fluorescence

Measurements of dark-adapted chlorophyll fluorescence are based upon observations made by Kautsky et al. (1960), describing the rapid change in fluorescence output upon transferring chloroplasts from darkness into the light. A brief period of darkness is required in order to allow PSII reaction centres to reset to an 'open' state, i.e. in a position to absorb actinic light. A short, high intensity pulse of light is then given with the effect of closing all available PSII reaction centres; it is during this period that the maximum fluorescence will be emitted, termed F_m. In conjunction with a measurement of the minimum fluorescence yield in the absence of actinic light, F_o, the 'variable fluorescence', or F_v, may be calculated as $F_v = F_m - F_o$. It is the ratio F_v/F_m that provides a sensitive measure of the maximum efficiency of PSII, shown to be ∼0.8 across a wide range of plants (Björkman and Demmig, 1987).

F_v/F_m was measured for all individuals at intervals throughout the trial using a *Hansatech Instruments Handy Pea* chlorophyll fluorimeter. Data sets were collected on a total of 22 occasions: the first on the 7th of July 2011, three days after the initiation of the flooding treatment, and the last on the 8th of October 2012. Individual needles were detached at random from the main stem of each plant and dark-adapted (using clips provided with the fluorimeter) for a period of at least 10 minutes before fluorescence measurements were taken. Measurements were obtained in mid-afternoon; control and treatment plots were measured alternately, the process typically taking around 2.5–3h to complete.

Previous-year (2010) needles were used initially, as current-year needles had not yet matured. Following the 15th of September 2011 (after which data had been collected six times), all measurements were taken from needles which had emerged in spring 2011. Prior to analyses, F_v/F_m was odds transformed to improve the normality of the residuals (eq. (5.1)), and data are reported in terms of transformed values.

$$\frac{(Fv/Fm)}{1-(Fv/Fm)} \tag{5.1}$$

5.2.4 Stomatal Conductance

Stomatal conductance was recorded on 29 occasions for a subset of 48 individuals using a *Licor LI-1600* steady state porometer. Data were first collected on the 5th of July 2011, and the final collection was made on the 12th of September 2012. As above, treatment and control plots were measured alternately; measurements began in mid-morning, and typically took around 2.5h to complete. The subsample consisted of a

western and eastern population (Glen Einig and Allt Cul), each respectively comprised of 24 individuals sampled across four half-sib families.

Measurements were made on needle pairs from whorl branches, which were not detached. Previous-year (2010) needles were used until mid-September 2011 (15 rounds of data collection), after which all measurements were made on 2012 needles. As needle pairs were too narrow to cover the sensor head aperture (necessary for operation), a microscope coverslip was affixed to form a seal. The porometer was configured to an arbitrary leaf area of $1cm^2$, and measurements were adjusted using a function provided in the manufacturer's instructions:-

$$C = \frac{1}{(\frac{1}{C_d} + 0.15)(\frac{A_t}{A}) - R_b} \tag{5.2}$$

where C is the 'true conductance', C_d is the displayed conductance, A_t is the exposed needle area (estimated as 0.39^2 from Chapter 3), A is the area value entered into the porometer, and R_b is boundary layer resistance (0.15 used by default).

5.2.5 Carbon Isotope Discrimination

In June 2011, prior to commencement of the flood treatment, whorl needles were sampled from 50% of the plants in the trial (216 individuals); this collection was repeated on the same individuals at the end of the trial in November 2012, after a prolonged period of flooding. The pre-treatment collection consisted of whorl needles that had emerged in 2010, and the post-treatment group from 2012: needles that had developed while the treatment was in effect. In each case a small number, typically 5–6, needles were randomly sampled from each plant (a larger number if needles were very short or underdeveloped), and dried in a desiccating oven for a period of no less than three days. Samples were then homogenised using a mechanical grinder, and submitted to the Lancaster *Centre for Ecology and Hydrology* for analysis.

1mg of each sample was combusted in a Carlo Erba NA1500 elemental analyser coupled to a Dennis Leigh Technology Isotope Ratio Mass-Spectrometer, and results were expressed in δ notation, whereby:-

$$\delta^{13}(‰) = (\frac{R_{sample}}{R_{standard}} - 1) \times 1000 \tag{5.3}$$

where R represents the ratio of ^{13}C to ^{12}C in the sample and the international standard Pee Dee Belemnite respectively.

5.2.6 Analysis of Treatment and Population Differences

The importance of treatment and population effects was evaluated via ANOVA. The following global model was fitted:-

$$\begin{aligned} Trait = \mu + Treatment_i + Population_j + Treatment : Population_{ij} \\ + Family_{k(j)} + Plot_{l(i)} + \epsilon_{ijkl} \end{aligned} \tag{5.4}$$

where Treatment, Population, and their interaction were considered as fixed effects; Family within Population, and Plot within Treatment were random effects. Model subsets were generated via *MuMIn* (Barton, 2013), and the best models were defined as those possessing the lowest AICc. When required, linear regression was performed by including relevant predictors as fixed effects, and including Population as a random effect.

5.2.7 Analysis of Differences by Climate

By assigning individuals to climate groups based on the conditions at their home sites, potential differences in average response can be evaluated. Individuals were grouped according to the cluster analysis of meteorological variables described in Chapter 3 (fig. 5.1). Six of the nine available populations were included to represent the two extremes of the rainfall gradient: three forming the 'low rainfall' group (Abernethy, Allt Cul, and Glen Tanar) with a mean rainfall of 72.8mm per month, and three forming the 'high rainfall' group (Crannach, Loch Clair, and Shieldaig) with a mean rainfall of 214.9mm per month. The following model was fitted:-

$$\begin{aligned} Trait = \mu + Treatment_i + Climate\ Group_j + Treatment : ClimateGroup_{ij} \\ + Population_{k(j)} + Family_{l(k(j))} + Plot_{m(i)} + \epsilon_{ijklm} \end{aligned} \tag{5.5}$$

this time incorporating climate group as a fixed effect, and population within group as a random effect. This method of analysis was applied only for chlorophyll fluorescence, the largest of the data sets.

5.2.8 Estimation of Genetic Components

Trait variances were first partitioned within treatment groups via the following model fitted by ML:-

$$Trait = \mu + Population_i + Family_{j(i)} + Plot_k + \epsilon_{ijk} \tag{5.6}$$

in which the intercept is the treatment mean, and all other parameters are considered to be random effects. Stomatal conductance data were not analysed here due to very limited sample size.

The coefficient of phenotypic variation $\mathrm{CV_P}$, was estimated by pooling across populations within treatment groups:-

$$\mathrm{CV_P} = \frac{\sqrt{V_\mathrm{P}}}{\mu_\mathrm{trait}} \times 100 \tag{5.7}$$

where V_P is phenotypic variance (total variance, excluding population and plot components), and μ_trait is the mean trait value. In addition, the coefficient of additive genetic variation ($\mathrm{CV_A}$), as defined by Houle (1992), was estimated using the above equation by substituting V_P for V_A (for half-sibs this is four times the family variance).

5.3 Results

5.3.1 Chlorophyll Fluorescence

Following initiation of the flood treatment, Fv/Fm values in the treatment group began to fall appreciably relative to the control, producing evidence of an effect within 30 days (table 5.2, fig. 5.2). After two months this effect had increased in magnitude, and for the remainder of the trial all of the best models incorporated a treatment term. The mean values of both treatment groups varied substantially over time, but exhibited some degree of covariance: using data from the 2012 needles (day 86 onward), treatment means were found to be significantly correlated ($p < 0.01$, $AdjR^2 = 0.37$). Mean Fv/Fm (untransformed) ranged between 0.71–0.80 in the control and 0.57–0.78 in the flood group, the lowest values in each case being recorded on the last day of measurement.

Pooling across treatments, a population effect was first detected on day 93 (table 5.2), however differentiation was detected consistently between days 128–357 (10/11/2011–26/06/2012). Analysis of populations within treatments (see electronic supplement for model tests within treatment groups) suggests that this effect was largely attributable to variation within the control, as under flooding group evidence was found for a population effect on only three occasions throughout the trial (days 30, 160, and 357). A treatment by population interaction was detected on a single occasion on day 357.

Ranking populations relative to one another (beginning from day 30, when a treatment effect was first identified), the three populations exhibiting the highest Fv/Fm in the control were Loch Clair and Crannach in the west, and Allt Cul from the Cairngorms mountains in the east (fig. 5.4). These three populations also ranked highly under flooding, in which the most westerly site, Shieldaig, increased in rank relative to the control.

The two sites which experience the lowest rainfall, Glen Tanar and Abernethy, were intermediate in the control but typically exhibited the lowest mean Fv/Fm values under flooding, being ranked last 45% and 30% respectively over the 20 dates evaluated. Eastern and western populations exhibited some degree of segregation under flooding, although this was not well defined: Cona Glen experiences the fourth highest precipitation rate, but frequently occupied the lowest positions in both the control and under flooding; conversely, the third driest population (in terms of mean rainfall), performed strongly in both groups.

By assigning populations to high and low rainfall climate groups (at site of origin), statistical power to detect differentiation was increased. On the initial six measurements performed on previous year (2010) needles, climate group means fluctuated within treatment groups and no effect of climate was detected either within or pooling across treatments (table 5.3, fig. 5.3). Following the transition to current-year (2011) needles on day 86, the mean Fv/Fm of the high rainfall group was consistently higher than the low rainfall group within the flood treatment until the final day of measurement; evidence was found for a climate effect under flooding on nine occasions, firstly on day 86, and then on all measurements taken between day 160 and 386 (12/12/2010–25/07/2012). Within the control, climate group means were generally highly similar, with the exception of a period during winter when mean Fv/Fm was greater in the high rainfall group; during this time evidence was found for a climate at site of origin effect on days 191 and 225 (12/01/2012 and 15/02/2012). Modelling both treatment groups together, there was evidence for treatment by climate interactions on four occasions: the first on day 86, the remaining three occurring toward the later stages of the trial on days 290, 357, and 407 (between 20/04/2012 and 15/08/2012).

Figure 5.2: Mean Fv/Fm (odds transformed) by treatment throughout the duration of the trial. Evidence was found for a treatment effect from day 30 onwards (table 5.2). Instances where population differences were identified by ANOVA within treatment groups are highlighted with a star. The transition between 2010 and 2011 needles is indicated by a line break. Errorbars represent ± 1 SE, and were estimated for individual treatment groups by REML.

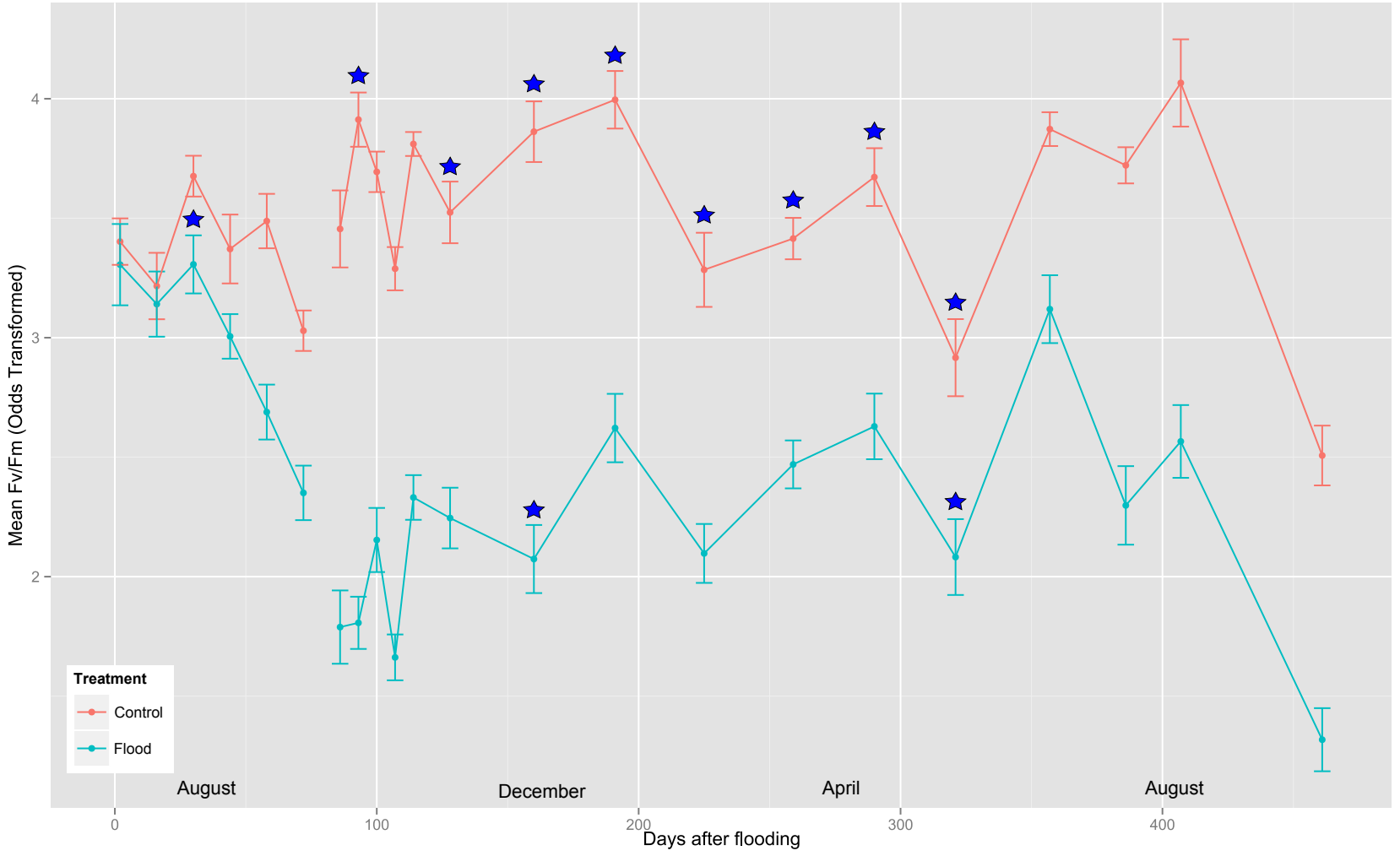


Table 5.2: Summary of chlorophyll fluorescence analyses by treatment and population. Model subsets were generated and compared by AICc to assess the importance of Treatment, Population, and their interaction as fixed effects. The best models were determined by AICc, and the presence or absence (+/-) of fixed effects are indicated for each of the traits listed. ΔAICc values represent the difference in AICc between the null model and the best (a value of zero indicates that the null model was the most parsimonious).

Days After Flooding	Mean Fv/Fm* Control	Flood	Treatment	Population	Treat × Pop	df	ΔAICc	Akaike Weight	Variance due to Family (%)	Variance due to Plot (%)
2	3.40	3.30	-	-	-	4	0.00	0.69	0.17	8.93
16	3.22	3.14	-	-	-	4	0.00	0.60	2.10	5.28
30	3.68	3.31	+	-	-	5	6.49	0.92	0.00	0.15
44	3.37	3.01	+	-	-	5	3.64	0.86	3.89	1.21
58	3.49	2.69	+	-	-	5	14.85	0.91	1.42	0.84
72	3.03	2.35	+	-	-	5	13.01	1.00	0.00	1.22
86	3.46	1.79	+	-	-	5	22.15	0.81	2.57	5.09
93	3.91	1.81	+	+	-	13	45.10	0.61	2.36	0.00
100	3.69	2.16	+	-	-	5	30.02	0.93	3.40	1.57
107	3.29	1.66	+	-	-	5	36.30	0.88	3.99	0.86
114	3.81	2.33	+	-	-	5	40.98	0.91	3.71	0.00
128	3.53	2.24	+	+	-	13	29.26	0.90	2.67	2.74
160	3.86	2.08	+	+	-	13	40.68	0.96	3.75	1.13
191	4.00	2.62	+	+	-	13	35.59	0.95	5.31	0.81
225	3.29	2.10	+	+	-	13	26.78	0.79	7.33	2.06
259	3.41	2.47	+	+	-	13	28.54	0.81	3.54	0.94
290	3.67	2.63	+	+	-	13	21.08	0.60	1.42	2.70
321	2.92	2.08	+	+	-	13	29.64	0.99	1.81	4.28
357	3.87	3.12	+	+	+	21	15.73	0.41	0.00	3.86
386	3.72	2.30	+	-	-	5	22.50	0.92	2.14	3.91
407	4.07	2.57	+	-	-	5	18.51	0.94	3.86	5.61
461	2.51	1.30	+	-	-	5	16.49	0.96	4.53	3.50

*Odds transformed

Figure 5.3: Mean Fv/Fm (odds transformed) for climate groups within treatments. As with the full data, evidence was found for a treatment effect from day 30 onwards (table 5.3). The transition between 2010 and 2011 needles is indicated by a line break. Differences between climate groups within treatment groups are marked with a star.

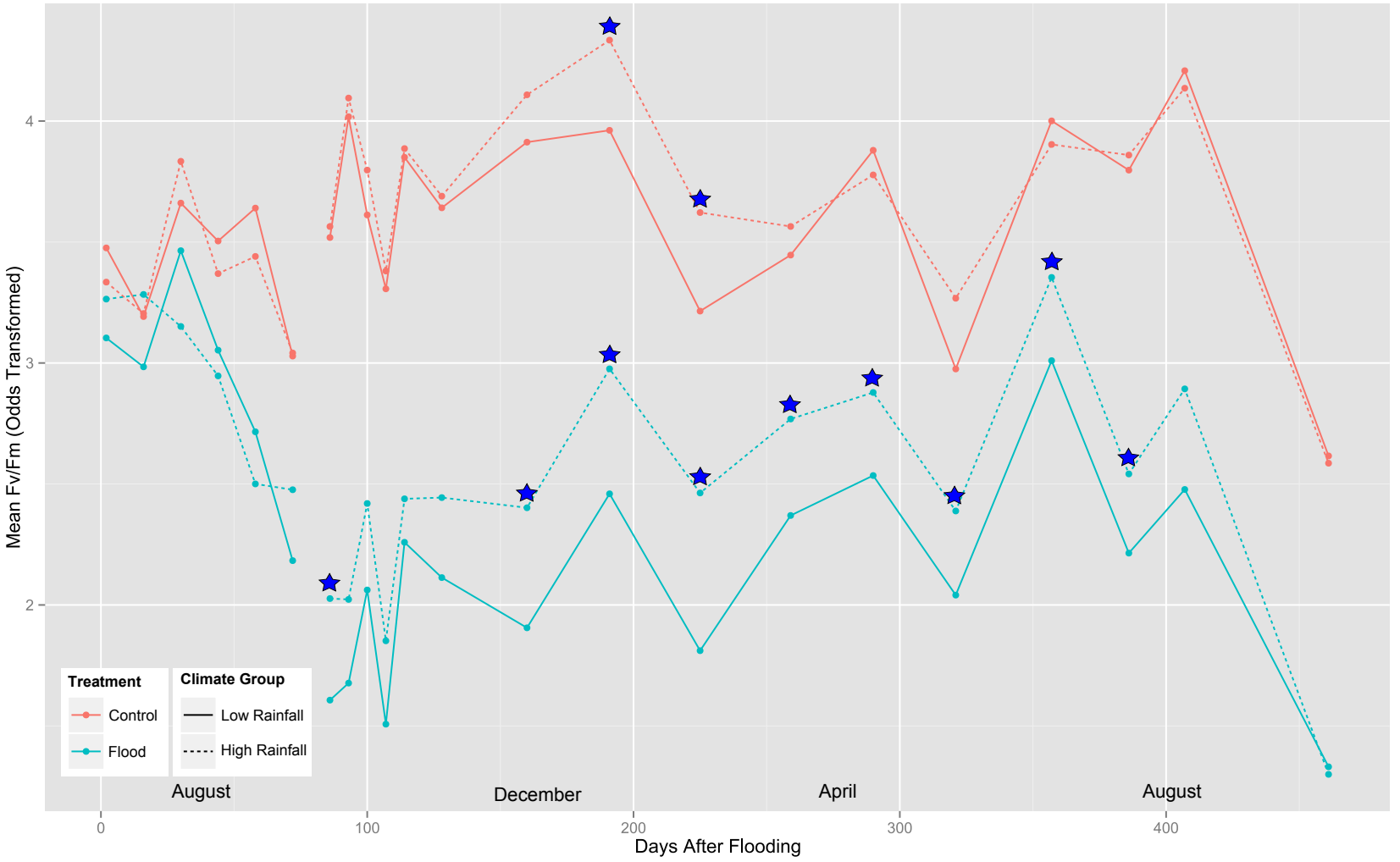
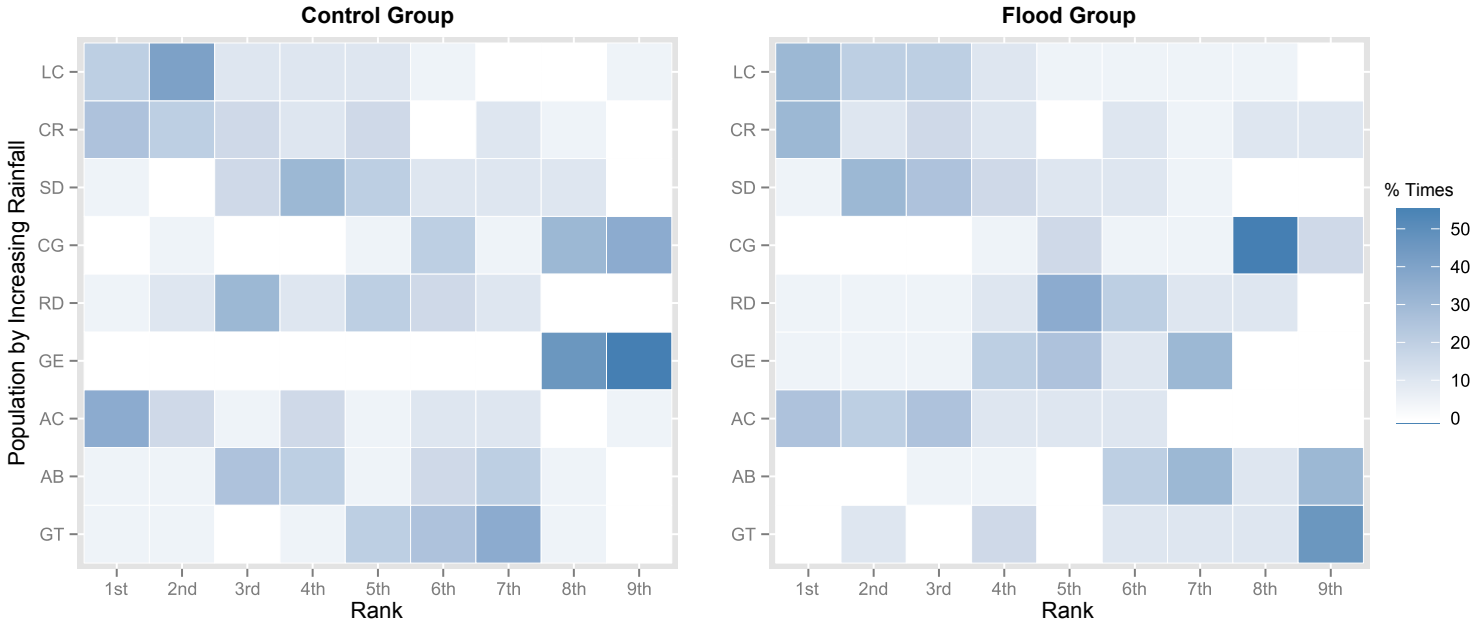


Table 5.3: Summary of chlorophyll fluorescence analyses by treatment and climate group. Model subsets were generated and compared by AICc to assess the importance of Treatment, Climate Group, and their interaction as fixed effects. The best models were determined by AICc, and the presence or absence (+/-) of fixed effects are indicated for each of the traits listed. ΔAICc values represent the difference in AICc between the null model and the best (a value of zero indicates that the null model was the most parsimonious).

Days After Flooding	Mean Fv/Fm* Control	Flood	Treatment	Clim Group	Treat × Clm Grp	df	ΔAICc	Akaike Weight	Variance due to Population (%)	Variance due to Family (%)	Variance due to Plot (%)
2	3.41	3.18	-	-	-	5	0.00	0.41	0.00	0.00	9.52
16	3.20	3.13	-	-	-	5	0.00	0.46	0.00	0.00	5.55
30	3.75	3.31	+	-	-	6	5.96	0.47	2.34	0.28	0.64
44	3.44	3.00	+	-	-	6	3.26	0.53	0.76	2.52	1.92
58	3.54	2.61	+	-	-	6	15.53	0.52	0.00	0.00	0.05
72	3.03	2.33	+	-	-	6	13.82	0.50	1.81	0.00	0.00
86	3.54	1.82	+	+	+	8	20.31	0.38	0.00	2.40	6.34
93	4.06	1.85	+	-	-	6	37.82	0.53	0.58	2.77	0.00
100	3.70	2.24	+	+	-	7	24.26	0.47	2.59	4.12	2.51
107	3.34	1.68	+	-	-	6	29.72	0.40	0.00	3.96	1.86
114	3.87	2.35	+	-	-	6	35.71	0.62	1.35	5.35	0.00
128	3.67	2.28	+	-	-	6	24.95	0.47	1.11	5.58	2.38
160	4.01	2.15	+	-	-	6	31.04	0.42	1.38	5.35	1.20
191	4.15	2.72	+	+	-	7	25.50	0.55	4.69	6.42	1.74
225	3.42	2.14	+	+	-	7	29.40	0.58	1.89	7.84	1.35
259	3.51	2.57	+	+	-	7	20.11	0.38	0.00	6.84	1.88
290	3.83	2.71	+	+	+	8	17.58	0.41	0.57	0.00	3.46
321	3.12	2.22	+	+	-	7	15.75	0.51	2.12	3.26	3.23
357	3.95	3.18	+	+	+	8	18.34	0.64	2.89	0.00	2.37
386	3.83	2.38	+	+	-	7	19.33	0.39	0.55	1.41	5.49
407	4.17	2.69	+	+	+	8	18.29	0.43	0.00	4.01	5.18
461	2.60	1.32	+	-	-	6	18.22	0.68	0.00	3.10	3.53

*Odds transformed

Figure 5.4: Relative rankings of populations over time in terms of Fv/Fm. Over the 20 dates considered (from day 30 onward, when a treatment effect was first identified), populations are ranked in terms of mean Fv/Fm (highest to lowest) and the proportion of instances each was found to be at a given position is indicated. Populations are listed here in order of decreasing rainfall (top to bottom).



5.3.2 Stomatal Conductance

Values of needle g_s ranged from 0.00 to a maximum of 0.83 $\mathsf{cm\ s^{-1}}$, although the latter was an outlier, the second highest value being 0.32 $\mathsf{cm\ s^{-1}}$.

Mean needle g_s was consistently higher in the control group than under flooding (with one exception on day 196) (table 5.4). However, treatment differences were detected on thirteen occasions only (table 5.4). No evidence was found for interpopulation differentiation, though given the greatly restricted sample size (48 individuals), substantial differences would have been necessary to produce an effect. Plot variance was very high compared with the other traits measured; commonly above 20% of the total, and as much as 58% in one instance. As with the chlorophyll fluorescence data, significant covariation was observed between the means of the treatment groups ($p < 0.01$, $AdjR^2 = 0.50$).

5.3.3 Carbon Isotope Discrimination

Prior to the commencement of flooding, treatment groups (essentially replicates at this stage) exhibited highly similar mean values of δ^{13}C (table 5.5, fig. 5.7), with an overall mean of -29.01(±0.11) ‰; individual observations ranged between -30.84 and -26.63 ‰. In the absence of treatment, there was strong evidence of population differentiation (ΔAICc=11.20); this variation was explainable in terms of longitude, or rainfall at home sites (fig. 5.6 a,b), with somewhat stronger support for the former (ΔAICc=13.7 and -9.6, respectively). Individuals from eastern, drier provenances were less depleted in δ^{13}C on average than those from the west, which indicates greater WUE in the former.

Following the second collection of needles, which took place after approximately 17 months of flooding, strong evidence was found for a treatment effect (ΔAICc=9.8). Flooded plants were less depleted in ^{13}C relative to the controls: average δ^{13}C was -30.90(±0.08) and -29.63(±0.23) ‰for control and flood groups, respectively. However, there was no evidence for a population effect, nor any relationship with longitude or rainfall. In the post-flood sample both treatment groups exhibited substantially lower values δ^{13}C than in the pre-flood, and a wider range of values was observed (-32.68 to -22.12 ‰).

Both pre and post flooding, family and plot variances were estimated to be very low or negligible when pooling across treatment groups.

Figure 5.5: Mean stomatal conductance (g_s) by treatment. Evidence for a treatment effect was first found on day 10, but the effect was not detected consistently thereafter (table 5.4). No population differences were detected within or pooling across treatments. The transition between 2010 and 2011 needles is indicated by a line break. Errorbars represent ± 1 SE, and were estimated for individual treatment groups by REML.

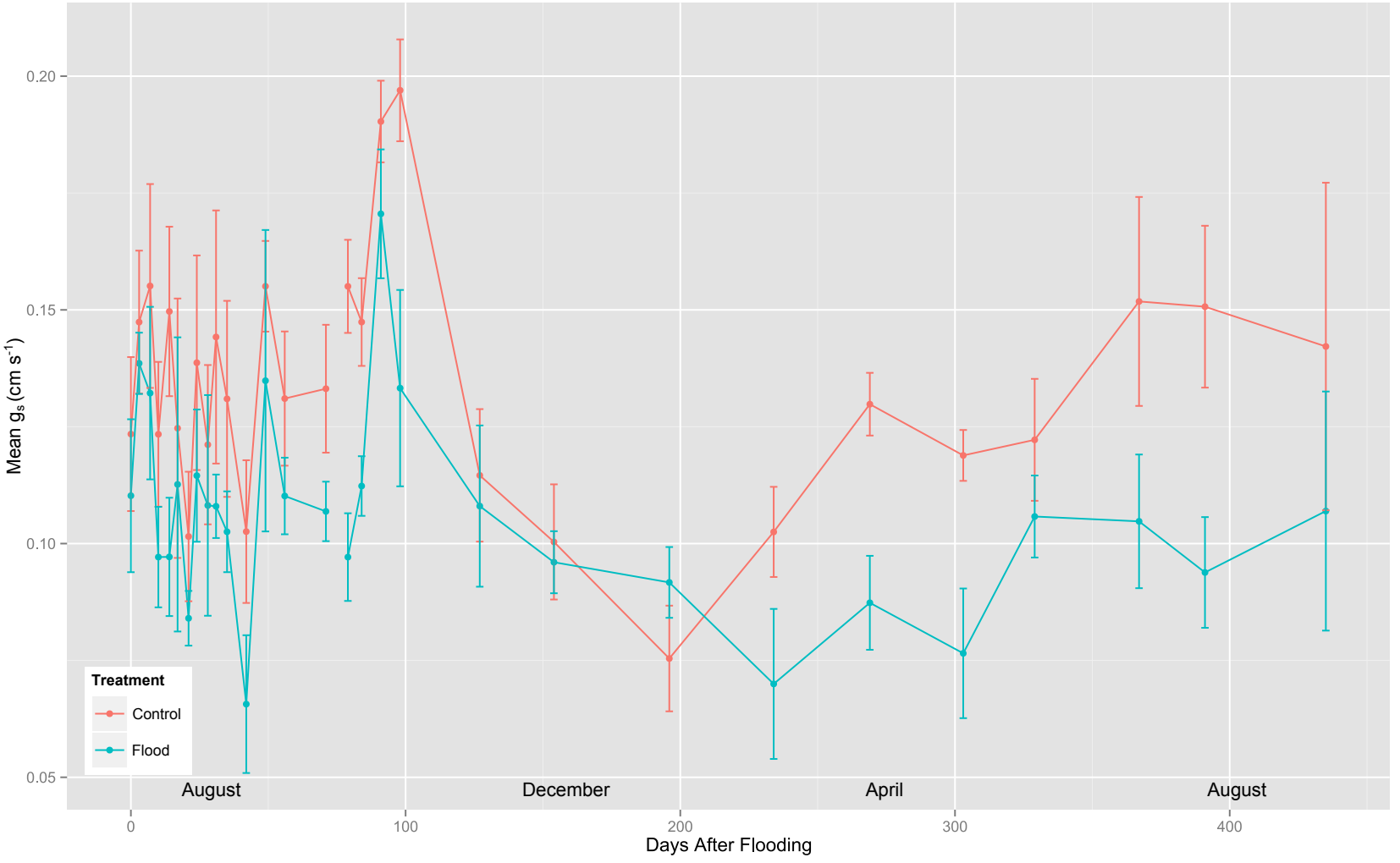


Table 5.4: Summary of stomatal conductance (g_s) analyses by treatment and population. Model subsets were generated and compared by AICc to assess the importance of Treatment, Population, and their interaction as fixed effects. The best models were determined by AICc, and the presence or absence (+/-) of fixed effects are indicated for each of the traits listed. ΔAICc values represent the difference in AICc between the null model and the best (a value of zero indicates that the null model was the most parsimonious).

Days After Flooding	Mean g_s (cm s^{-1}) Control	Flood	Treatment	Population	Treat × Pop	df	ΔAICc	Akaike Weight	Variance due to Family (%)	Variance due to Plot (%)
0	0.12	0.11	-	-	-	4	0.00	0.49	0.00	8.43
3	0.15	0.14	-	-	-	4	0.00	0.44	8.58	9.78
7	0.16	0.13	-	-	-	4	0.00	0.51	32.70	25.23
10	0.12	0.10	+	-	-	5	0.28	0.40	4.38	10.69
14	0.15	0.10	+	-	-	5	3.21	0.59	3.59	14.53
17	0.12	0.11	-	-	-	4	0.00	0.54	0.24	32.08
21	0.10	0.08	-	-	-	4	0.00	0.41	15.89	14.04
24	0.14	0.11	-	-	-	4	0.00	0.41	7.63	36.76
28	0.12	0.11	-	-	-	4	0.00	0.41	0.00	28.39
31	0.14	0.11	-	-	-	4	0.00	0.41	3.26	49.85
35	0.13	0.10	-	-	-	4	0.00	0.40	12.10	45.23
42	0.10	0.07	+	-	-	5	0.95	0.45	1.11	32.53
49	0.16	0.13	-	-	-	4	0.00	0.44	0.00	0.00
56	0.13	0.11	+	-	-	5	0.03	0.37	0.06	5.71
71	0.13	0.11	+	-	-	5	1.41	0.48	7.25	10.77
79	0.16	0.10	+	-	-	5	9.11	0.67	0.00	4.43
84	0.15	0.11	+	-	-	5	6.81	0.71	0.00	0.00
91	0.19	0.17	-	-	-	4	0.00	0.40	0.00	14.25
98	0.20	0.13	+	-	-	5	4.28	0.64	0.00	15.70
127	0.11	0.11	-	-	-	4	0.00	0.58	0.00	24.86
154	0.10	0.10	-	-	-	4	0.00	0.51	4.97	5.33
196	0.08	0.09	-	-	-	4	0.00	0.36	0.00	7.29
234	0.10	0.07	+	-	-	5	1.28	0.30	3.14	33.63
269	0.13	0.09	+	-	-	5	8.39	0.47	0.00	0.00
303	0.12	0.08	+	-	-	5	6.46	0.50	1.21	3.68
329	0.12	0.11	-	-	-	4	0.00	0.39	2.25	11.71
367	0.15	0.10	+	-	-	5	1.03	0.43	0.00	36.41
391	0.15	0.09	+	-	-	5	4.08	0.65	0.00	23.08
435	0.14	0.11	-	-	-	4	0.00	0.44	0.00	58.46

Table 5.5: Summary of carbon isotope analyses by treatment and population. Model subsets were generated and compared by AICc to assess the importance of Treatment, Population, and their interaction as fixed effects. The best models were determined by AICc, and the presence or absence (+/-) of fixed effects are indicated for each of the traits listed. ΔAICc values represent the difference in AICc between the null model and the best (a value of zero indicates that the null model was the most parsimonious). Pre-flood results are presented for the sample collected in June 2011 before the start of treatment, and post-flood for the sample collected in November 2012 after approximately 17 months of sustained flooding.

Days After Flooding	Mean δ^{13}C(‰) Control	Flood	Treatment	Population	Treat × Pop	df	AICc	ΔAICc	Akaike Weight	Variance due to Family (%)	Variance due to Plot (%)
Pre-Flood	-28.98	-29.03	-	+	-	12	498.8	11.2	0.73	0.00	0.00
Post-Flood	-30.83	-29.63	+	-	-	5	822.7	9.8	0.99	2.75	0.70

5.3.4 Effects of Flooding on Phenotypic and Heritable Variation

Estimates of CV_P and CV_A were made within treatment groups for all dates on which Fv/Fm was recorded, and for $\delta^{13}C$ pre and post flooding.

In the Fv/Fm data set CV_P was found to be consistently higher in flooded plants than in the control from day 44 (fig. 5.8), peaking shortly after the transfer to 2011 needles. CV_A were highly variable over time, particularly in the control, and often restricted to zero which seems likely due to limitations on sample size. As with CV_P, following transition to 2011 needles CV_A was higher in the treatment group with the exception of two occasions (days 321 and 357), when family variance estimates were restricted to zero.

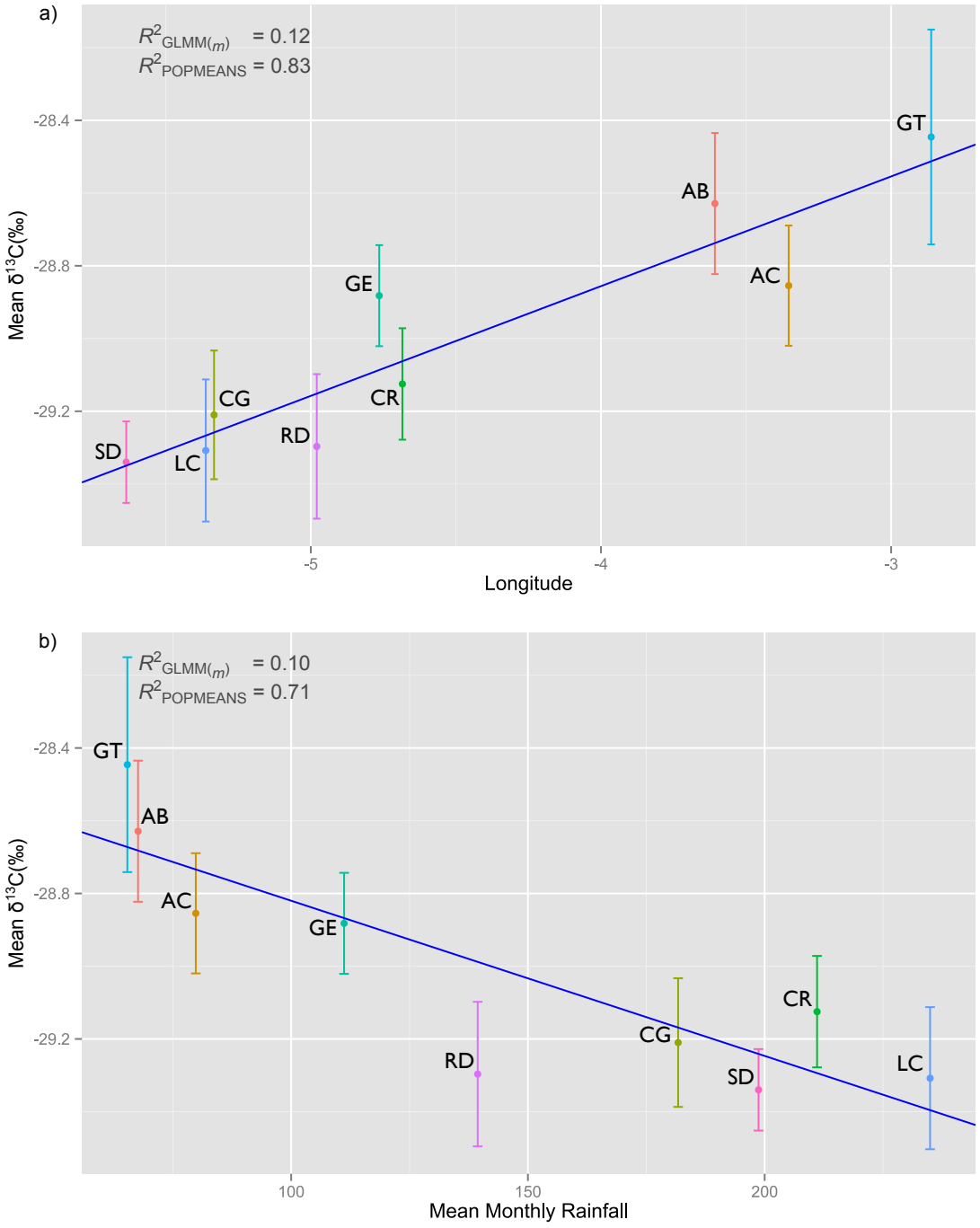
When considered in terms of high and low rainfall groups, CV_P estimates within the control remained highly similar throughout the trial (fig. 5.9a); under flooding, neither group was consistently higher or lower prior to the change to 2011 needles, after which CV_P estimates were greater in the low rainfall group until the final date of measurement. CV_A estimates were more erratic in both treatment groups, reflecting the reduced sample sizes on which they were based (fig. 5.9b). In the control, family variances were frequently restricted to zero and rainfall groups were not clearly differentiated; within the flooded group CV_A estimates were typically higher for the low rainfall group, however both groups were subject to large variability.

For carbon isotope data, all CV estimates were relatively low. CV_P was 2.92 and 2.69% for control and flood groups respectively prior to flooding; 2.47 and 7.38% post flooding. CV_A estimates were especially low: 0.00% and 1.10% prior to flooding, 2.03% and 0.00% post-flooding.

5.4 Discussion

Scotland's native *P. sylvestris* populations occupy a diversity of environments with respect to meteorological, altitudinal, and edaphic conditions. In this study we examined the performance of *P. sylvestris* populations sampled across the native Scottish range under flooded conditions, expected to be more common on western, high rainfall sites. Three physiological measures were recorded: chlorophyll fluorescence, stomatal conductance (for a small subsample), and carbon isotope discrimination. We hypothesised that high rainfall sites would exhibit lower stress, and lower variation in response as a result of past selection.

Figure 5.6: Mean δ^{13}C(‰) by population regressed onto a)longitude, and b)mean monthly rainfall at home sites. Of these longitude provided moderately greater explanatory power. Errorbars represent ± 1 SE, and were estimated for individual treatment groups by REML.



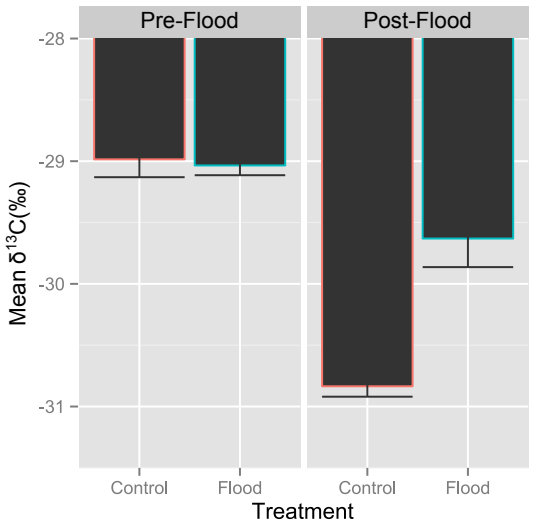


Figure 5.7: Mean treatment δ^{13}C(‰) pre and post-flooding. Pre-flooding, treatment means were highly similar; after 17 months of prolonged flooding both groups were relatively more depleted in ^{13}C, however there was also a marked treatment effect.

Figure 5.8: Estimated CV_P and CV_A pooled across populations within treatments for odds transformed Fv/Fm data. Both values were generally higher within the flooded group, however CV_A estimates were less stable, probably as a result of sample size limitations. The transition between 2010 and 2011 needles is indicated by a line break.

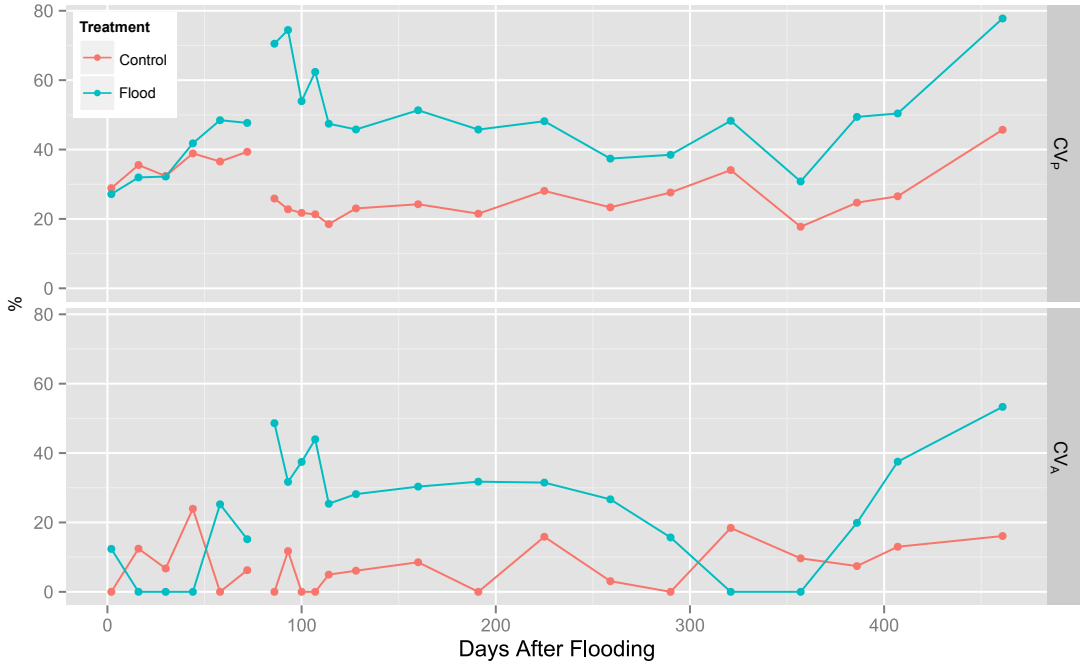
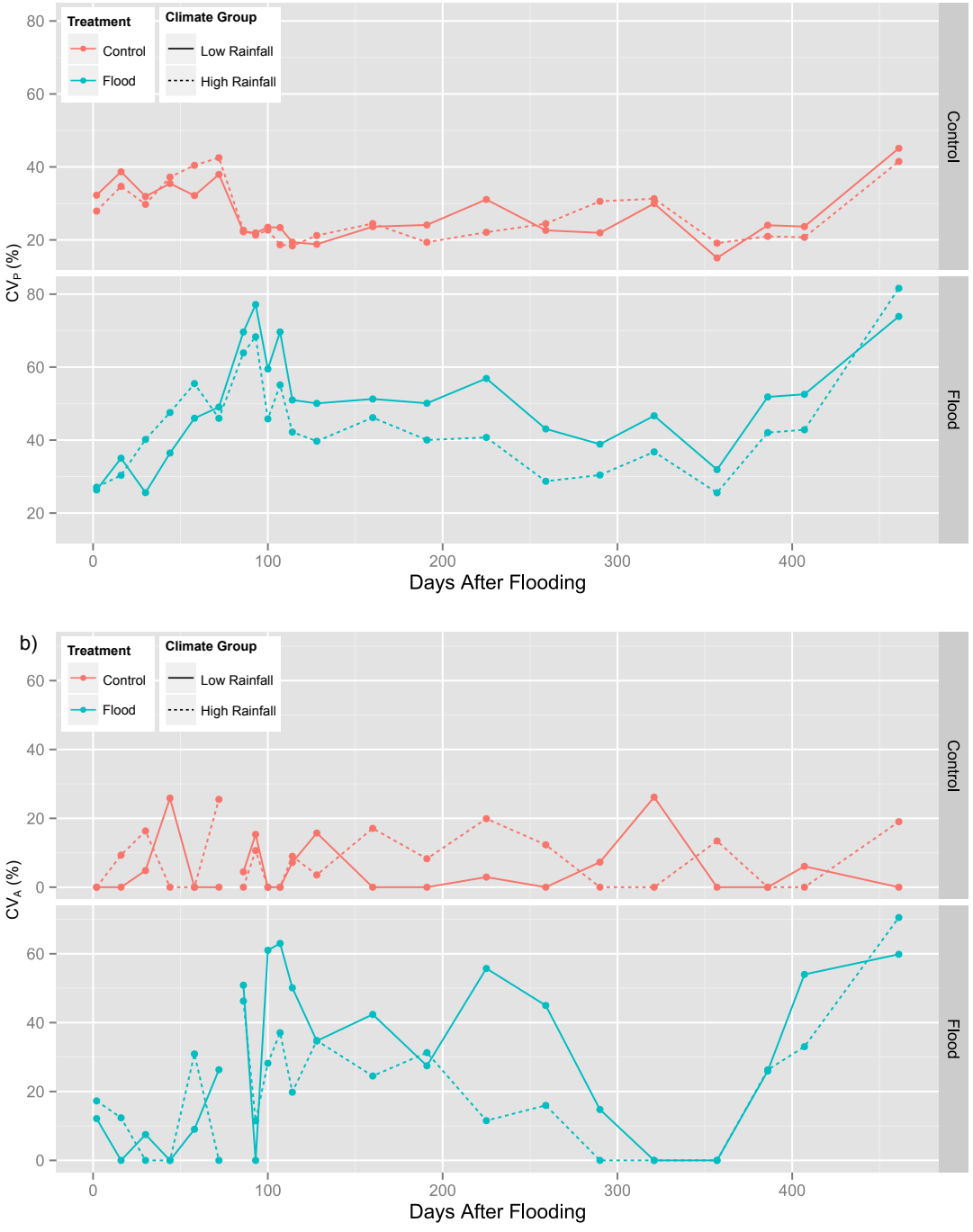


Figure 5.9: Estimated CV_P (a) and CV_A (b) within climate groups for odds transformed Fv/Fm data. Under flooding, the low rainfall group typically exhibited greater variance than the high rainfall. Subdivision of the sample into climate groups resulted in increasingly erratic estimates of CV_A. The transition between 2010 and 2011 needles is indicated by a line break.



5.4.1 Chlorophyll Fluorescence

Chlorophyll fluorescence provides a rapid and means by which to gauge the efficiency of plant photosynthetic apparatus, and has been shown to be responsive to low and high water regimes (e.g. Demmig et al., 1988; Pearson et al., 2013). In the first month following initiation of flooding, there was a discernible drop in Fv/Fm in the treatment group relative to the control. The treatment was sustained for over a year: during that time the disparity between groups peaked around four months after flooding began, but was variable throughout the subsequent period; the treatment group exhibited neither a progressive decline nor recovery. This is consistent with the results of Pearson et al. (2013), who performed a flooding experiment on *P. sylvestris* seedlings over the course of a single growing season. Given a longer period it seems likely flooded plants would have deteriorated further: shoot and needle growth were visibly impaired (as described in the previous chapter), and fluorescence measurements would at some point have had to be transferred to a younger, and probably less vigorous, age-class of needles. Nevertheless, preceding the initial effect, there was no appreciable trend in photoinhibition relative to the control group. In this regard *P. sylvestris* saplings were found to be relatively robust to waterlogging, at least in the short to medium term.

Throughout the trial we tracked the performance of saplings from nine native pinewoods around Scotland, with a view to determining whether those from areas more susceptible to flooding would outperform those from sites where such conditions are uncommon. Interpopulation differences were detected on numerous occasions within the control group, and occurred on all measurement dates between November 2011 and June 2012. In contrast, few interpopulation differences were detected under flooding.

Examining the relative performances of populations within treatments, it was apparent that some were able to maintain higher Fv/Fm over time than others. Loch Clair and Crannach are sites of origin that experience the greatest mean rainfall, and exhibited relatively high Fv/Fm in both treatment groups, while mean Fv/Fm of the third highest in terms of rainfall, Shieldaig, appeared to increase in relative terms under flooded conditions. Arguably, the most conspicuous difference between treatments was the relative drop in performance of the two populations which experience the lowest rainfall, Abernethy and Glen Tanar, when exposed to flooding. Nevertheless, a straightforward relationship between rainfall at home sites and Fv/Fm was lacking, particularly as Allt Cul (among the driest) performed strongly irrespective of conditions, whereas Cona Glen, which experiences relatively high rainfall, was consistently among the lowest ranked in both the control and under flooding.

Assigning populations to high or low rainfall groups reduced the number of parameters, and increased statistical power to detect potential phenotypic differences on the basis of precipitation at home sites. In the control, climate groups performed similarly to one another, with the exception of a period during winter 2011/12 during which the high rainfall group exhibited higher mean Fv/Fm. Why this occurred is unclear: under exception winter cold, Salmela et al. (2011b) found that the Fv/Fm of western maritime populations deteriorated relative to those from drier, high altitude sites. This is in contrast to our result, however, the trial described was conducted outdoors, whereas our experiment was carried out under glasshouse conditions where temperatures were maintained above 4°C. A very shallow water depth was maintained within control plots, and as a result, whilst unflooded, soils were damp due to capillary uptake through pot bases. It is possible that this elevated moisture regime may have imposed a mild degree of stress in itself, exposing variation within the control. Following the switch to current-year needles and until the final day of measurement, the populations comprising the high rainfall group consistently maintained a higher mean photochemical efficiency under flooding; differences between group means were significant on the majority of occasions. This would suggest that on average western, maritime sites were better able to maintain photosynthetic capacity under flooding than the drier, high altitude sites in the east.

5.4.2 Stomatal Conductance

Stomata are highly responsive to the stresses induced both by drought (Radin and Ackerson, 1981), and flooding (Bradford and Hsiao, 1982). In the latter tolerant plants have been shown to maintain higher conductances under flooding (Lewty, 1990), or re-open following an initial period of closure (Gomes and Kozlowski, 1980). Throughout the trial the treatment had a measurable effect on needle g_s, which was depressed in the flooded group relative to the control; however, although there were a few individual instances, at no point did mean g_s approach zero, which would indicate complete stomatal closure. This may be testament to the inherent tolerance of *P. sylvestris* to waterlogging, at least to the level where soil is not entirely submerged. Due to the time-intensive nature of the measurements, needle g_s was recorded on only a small subset of individuals, and statistical power was therefore limited. The variance attributable to plots within treatments was often very high; likely this reflects the responsiveness of stomatal behaviour to the changeable weather conditions, as well as the diurnal progression experienced over the course of a single data collection. Although porometry provides valuable information on plant responses to water stress, the large variability, coupled with the restriction to small

sample sizes, may render it unsuited to the detection of potentially subtle intraspecific differences within a common garden.

5.4.3 Carbon Isotope Discrimination

Carbon isotope discrimination provides an indirect measure of the degree to which transpiration and gas exchange are restricted via stomata, and may be employed in selection for WUE (Farquhar et al., 1989). As tissue samples are desiccated immediately after collection, δ^{13}C can be measured effectively simultaneously for a large number of individuals. With the exception of those involving crop species, studies of carbon isotope discrimination are very often carried out *in situ*; a strategy which may provide physiological or ecological insights, but is genetically uninformative. By analysing samples from a common environment we have much more power to estimate the component of variation due to genetic differences.

Strong evidence of interpopulation differentiation was found in the isotopic composition of needles collected prior to the commencement of the treatment. Furthermore, there was evidence for a positive relationship between δ^{13}C and longitude, and with average rainfall, where δ^{13}C was greater for populations which experienced reduced precipitation. It is possible that the larger δ^{13}C values observed for eastern populations reflects past selection for WUE, however, δ^{13}C is also reported to be correlated with altitude (Körner et al., 1988; Hultine and Marshall, 2000), which also increases toward the east of the distribution where pinewoods are situated within the Cairngorm mountains. Due to the collinearity between altitude and rainfall gradients in our trial populations, their effects are not easily separable.

Previous investigations have provided evidence that δ^{13}C varies within plant species, as result of both phenotypic plasticity and local adaptation. Conducting an *in situ* study of the evergreen tree *Metrosideros polymorpha*, Meinzer et al. (1992) found the least negative δ^{13}C values were recorded for individuals growing in sites where water availability was considered to be the most limited. Anderson et al. (1996) demonstrated that isotopic discrimination increased with precipitation at home sites in a common garden trial of Australian *Eucalyptus* species. This seems likely to reflect selection for increasing WUE when water is less abundant, though conversely it could be posited that divergent selection occurs for lower WUE in regions where water is not a limiting factor, due to the tradeoff with net carbon assimilation. Among conifers, Korol et al. (1999) observed that isotope discrimination was greater during years in which trees experienced increased rainfall, analysing cores taken from *P. radiata*. For *P. strobus*, McNulty and Swank (1995) reported that variation in annual wood δ^{13}C, could be explained by soil

water potential during the growing season. A study of a Spanish common garden trial of *P. pinaster* identified differences among populations that could be be explained by climatic variation at sites of origin, finding in this case that temperature was more effective than annual precipitation (Correia et al., 2008). Warren et al. (2001) provide a review of carbon isotope discrimination in conifers in relation to water availability, and concluding that the ratio of precipitation to transpiration could explain a significant proportion of variation in δ^{13}C, but only when precipitation was relatively low relative to transpiration.

In Scotland, Brendel et al. (2002) previously reported that wood cellulose δ^{13}C differed among Scottish pinewoods measured *in situ*, and covaried with potential moisture deficit. As part of the same study measurements were also repeated for a clone bank located in Elgin, northern Scotland, where each group of clones was grown from shoots taken from individual mother trees sampled from throughout the Scottish distribution. Unlike the natural populations, variation in δ^{13}C did not correspond to climate variables at sites of origin (although some degree of geographic clustering was observed). The authors concluded that either the differences among populations were too small to be detectable with the experimental design chosen, or that interpopulation differences were entirely attributable to plasticity. The results from the present study would suggest that, although plasticity may be important, local adaptation has taken place.

A number of studies have shown that carbon isotope discrimination is reduced (i.e. δ^{13}C increased) under droughted conditions (Hall et al., 1990; Meinzer et al., 1990; Sayre et al., 1995). Given that flooding reduces g_s, and causes soil anoxia which may in turn result in root damage, flooding is predicted to influence δ^{13}C in the same direction as drought. Post-flooding, we found mean δ^{13}C to be markedly higher in the treatment group than in the control, implying that WUE was greater in flooded plants. However, population effects were absent and no environmental trends were identified, which seems surprising given the extent of differentiation observed prior to flooding. Statistical power to detect population differences may have been diminished by division of the sample into treatment groups, but changes in trial conditions seem likely to be of more importance. Needles collected from within the treatment group had spent the entirety of their development under flooding, and as described above, control plots were exposed to different soil conditions relative to pre-treatment. Furthermore, δ^{13}C values are subject to seasonal variability (Voltas et al., 2006), and post-treatment needles were collected at an earlier stage of their life cycle than the original sample.

5.4.4 Effects of Flooding on Genetic Variation

It has been posited that the application of stress, may serve to 'release' cryptic variation (Rutherford and Lindquist, 1998; Schlichting, 2008). We estimated coefficients of phenotypic and additive genetic variation for the chlorophyll fluorescence and carbon isotope datasets. The former supported the hypothesis, at least for the level of flooding in the trial: shortly after initiation of the treatment CV_P increased relative to the control, and remained consistently higher for the trial duration, with the exception of the final measurement. CV_A estimates followed a similar pattern, but were considerably more erratic, particularly as V_A was frequently constrained to zero; likely due to sample size limitations. Variation was also significantly larger for regions where flooding was less common: following the transfer of measurements to a younger age-class of needles CV_P was consistently higher for the low rainfall group, however CV_A was more ambiguous due to the reduction in group sample sizes.

Following exposure to prolonged flooding, CV_P for $\delta^{13}C$ values was increased in the treatment group relative to the control. The corresponding CV_A was estimated to be zero, though again this seems likely to reflect the sample size limitations, particularly as $\delta^{13}C$ values were measured on half of the trial (216 individuals).

Overall, exposure to flooding appeared to increase variability, and at least in terms of photochemical efficiency, this effect was more pronounced in populations from drier provenances. This may be a reflection of greater past selection pressure for adaptation to flooded conditions in western, high rainfall sites, compared with drier sites in the east.

5.4.5 Conclusions

Prolonged flooding of *P. sylvestris* saplings produced a strongly discernible effect in all three of the physiological traits examined: photochemical efficiency and stomatal conductance were reduced, and WUE was increased. Chlorophyll fluorescence data demonstrated that individuals from the wettest climates are able to maintain higher Fv/Fm under flooding than those from the driest; however, there was some ambiguity at the population level, and performance could not be expressed as a simple function of rainfall or longitude. While population level variance should not be neglected, experimental designs that assign populations to larger groups, when there is a rationale for doing so, may be more effective at detecting local adaptation.

In the absence of any manipulation, a direct relationship was identified between the needle $\delta^{13}C$ of individuals in the glasshouse, and their sites of origin. Those from areas

which receive lower rainfall were less depleted in ^{13}C, indicating greater WUE. Exposure to prolonged waterlogging was found to increase WUE, but potentially diminish population variability; however, this was true in both the treatment and control groups, suggesting that alternative factors such as changes in trial conditions or needle age class may have been important.

In the chlorophyll fluorescence data, both phenotypic and additive genetic variances were increased in response to flooding, and this effect was greater on average in individuals from regions of low rainfall. This would suggest that the stress induced by flooding served to expose cryptic variation, and that the potential for adaptation may be greater in populations from drier sites, where soil waterlogging is of greater novelty. This effect would be interesting to evaluate in the context of other environmental extremes.

Scots pine saplings demonstrated considerable resilience to the flooding treatment. It is pertinent that the effects described were observed during a period of more than a year, and had the trial been shorter, the conclusions drawn would likely have been rather different.

Chapter 6

Discussion

Local adaptation is a process by which populations adapt differentially to their local environments in the presence of gene flow (Kawecki and Ebert, 2004). Natural selections operates to erode genetic variation (Falconer and Mackay, 1996); however, variation may be preserved by means of local adaptation, as genotypes are differentially favoured across environments. Because of this, locally adapted systems are of importance not only to evolutionary biology, but also to the management and conservation of genetic resources, in order to maintain the adaptive capacity of natural populations faced with a changing and unpredictable climate (McKay et al., 2005; Kremer et al., 2012).

The Caledonian Pinewoods represent the north-western frontier of the vast Eurasian distribution of *Pinus sylvestris* (L.), and are relicts of what was once an extensively forested region. Since the publication of Stephen and Carlisle's *The Native Pinewoods of Scotland* in 1959, there has been a growing interest in the conservation and restoration of the remaining fragments, which are the last known native stands of *P. sylvestris* in the British Isles. The remaining populations inhabit a geographically limited area, but one which is highly diverse in terms of climate and topography (Mason et al., 2004). Scots pine is wind-pollinated, and gene flow can occur over many tens of kilometers (Varis et al., 2009). Genetic structure among pinewoods for neutral markers is very low Provan et al. (1998), and despite varying degrees of spatial separation, there is as yet no evidence that populations are unable to exchange genes. The remnants of the Caledonian Pine Forest therefore present an excellent opportunity to study local adaptation, an increasing understanding of which will also contribute to their future management (Salmela et al., 2010).

In this thesis I set out to discover new mitochondrial markers, which may be used to provide insights into the origins and postglacial history of *P. sylvestris* both in Scotland, and throughout its distribution. Within a common garden I evaluated evidence

for local adaptation in needle characters, and sought to identify which environmental factors are responsible. In view of the disparity in rainfall between native populations, a glasshouse-based experiment was undertaken to determine whether populations differed in their response to flooding, and whether drier populations exhibited greater variation in response due to a lack of past selection.

6.1 Molecular Work

On a number of occasions, previous studies have identified genetic variation, in terms of biochemical or molecular markers, that point toward a distinct origin for pinewoods in Wester Ross (Forrest, 1980; Kinloch et al., 1986; Sinclair et al., 1998; Provan et al., 1998). Sinclair et al. (1998) was the first of these studies to employ mitochondrial markers, which are particularly well-suited to studies of population history in pine due to their (at least predominantly) maternal mode of inheritance (Neale and Sederoff, 1988), which restricts dispersal to seed thereby preserving geographic structure. Following Sinclair et al. (1998), several studies have employed mitochondrial markers in efforts to deduce postglacial colonisation routes of *P. sylvestris* throughout Europe (Sinclair et al., 1999; Soranzo et al., 2000; Naydenov et al., 2007; Pyhäjärvi et al., 2008). These have been relatively successful, and have provided evidence for multiple origins of Scots pine in Europe, however, more markers are needed to increase resolution.

We employed modern whole-genome-shotgun (WGS) sequencing in order to search previously unexplored regions of the mitochondrial genome for novel variants. A major obstacle in this regard, was the absence of a reference genome; the closest taxonomically available being that of *Cycas taitungensis*. Altogether, 46 putative mitochondrial contigs were identified ranging from ∼1–20kbp in length. A number of polymorphic loci were observed among the samples sequenced which, as well as *P. sylvestris*, included members of the closely related *P. mugo* species complex. There was no clear separation between species groups based upon the pattern of variants observed, suggesting that both may share a common pool of haplotypes. In the absence of a reference genome, however, it cannot at this point guaranteed that the contigs identified were of mitochondrial origin. A prudent next step would therefore be to conduct controlled crosses between individuals which are polymorphic at the relevant loci, and then genotype the offspring to determine whether alleles exhibit a maternal mode of inheritance.

One of the main findings of this study was the apparent heteroplasmy at the majority of polymorphic loci discovered. Mitochondrial heteroplasmy has been observed before in *Pinus* (Wagner et al., 1991), and is reportedly widespread in plants (Kmiec et al., 2006). One possible cause is occasional transmission of the paternal allele: so-called ‘paternal

leakage'. This has repercussions for the use of mitochondrial variants in studies of population history, as haplotypes may be transported over large distances via pollen, violating the assumption that dispersal occurs via seed alone.

WGS has an advantage of Sanger sequencing in that heteroplasmy can be quantified based on the proportion of reads at a given loci which feature the alternate allele. When screening for variants in future, it may be appropriate to evaluate the proportions of each haplotype present at the level of the individual, as well as the population as a whole. Quantitative PCR (qPCR) may provide a means by which this could be achieved, and has previously been used to quantify SNP heteroplasmy in mammals (e.g. Niederstätter et al., 2006; Burgstaller et al., 2007). Using a similar methodology, it may also be possible to estimate the rate at which paternal leakage occurs.

The new markers developed here will require validation; however, if authenticated, they represent a substantial increase in the number of variants currently available for phylogeographic studies, and may provide new insights into postglacial history.

6.2 Common Gardens Trials

6.2.1 Intrapopulation Variability

Previous work has shown that a very large proportion of variability lies within pinewoods: this has been observed for biochemical makers (Forrest, 1980), monoterpene and allozyme loci (Kinloch et al., 1986), chloroplast microsatellites (Provan et al., 1998), nuclear markers (Wachowiak et al., 2011), as well as quantitative growth and phenological traits (Perks and Ennos, 1999; Salmela et al., 2013). Our findings were consistent with this pattern, and intrapopulation variances remained high irrespective of whether environmental trends or interpopulation differences were detected.

Long-lived woody plant are expected to harbour large amounts of diversity within populations, symptomatic of the high levels of gene flow between them (Hamrick et al., 1992). The survey of polymorphic nuclear loci conducted by Wachowiak et al. (2011), found that genetic diversity within Scotland was comparable to that of mainland European populations. This is encouraging for the future prospects of the native pinewoods, in that high levels of diversity appear to have been maintained in spite of a drastic contraction in range. Although gene flow may restrain local adaptation, under strong selective pressure adaptive divergence may still take place (Savolainen et al., 2007), and the work presented here adds to the growing body of evidence that this has happened in Scotland.

6.2.2 Environmental Clines

On two occasions quantitative trait variation was observed within common garden environments that could be explained by at least in part by conditions at home sites via regression analysis. In Chapter 3, the number of stomatal rows present on needle surfaces was found to decrease from west to east; and in Chapter 5, needle δ^{13}C was (conversely) observed to increase from west to east. These traits were measured in separate trials, which were comprised of different population samples. Nevertheless, given that both traits are related to plant water use, it's conceivable these observations are interrelated, and possibly the result of similar selection pressures.

The majority of water loss takes place via the stomata (Pallardy, 2008), and the average number of stomatal rows was found to be lower in the east where rainfall is substantially lower than the west. Studies describing variation in stomatal rows have often been carried out *in situ* (e.g. Androsiuk and Urbaniak, 2006; Bobowicz and Korczyk, 1994; Urbaniak et al., 2003), and as populations were not typically sampled across environmental gradients, it is difficult to find a direct precedent for this result. Carbon isotope ratios are known to exhibit a strong relationship with water use efficiency (WUE) (Farquhar et al., 1989), and trends in δ^{13}C have repeatedly been observed across moisture gradients Ehleringer and Cooper (1988); Stewart et al. (1995); MIller et al. (2001). Notably, δ^{13}C has also been documented to increase with altitude (Körner et al., 1988; Hultine and Marshall, 2000). Across the native pinewood distribution, altitude and mean rainfall are positively and negatively correlated with longitude respectively, making their effects difficult to delineate.

In a study of Chinese *Ginkgo* populations, Sun et al. (2003) made measurements of δ^{13}C and stomatal density on leaves collected from several locations subject to varying degrees of rainfall. Both carbon isotope discrimination and stomatal density were found to decrease in relation to decreasing rainfall. This is consistent with our results, albeit stomatal density is not directly equivalent to the number of stomatal rows (a conifer-specific trait). Hultine and Marshall (2000) examined variation in δ^{13}C and stomatal density in four conifer species across 1800m of altitude in the north-central Rockies, USA. They reported that δ^{13}C increased with altitude, while the number of stomata (per gram of leaf mass) decreased.

It seems likely that the trends observed with respect to stomatal rows and δ^{13}C are the result of differential selection in response to the moisture gradient across Scotland; possibly in conjunction with altitude, as these drivers need not be mutually exclusive. Common-garden experiments are required to detect local adaptation, however, given

these results it would be of interest to sample a longitudinal transect *in situ*, and determine whether the result can be reproduced on home sites.

6.2.3 Responses to Flooding

The area occupied by native pinewoods is subject to a steep rainfall gradient, and western sites may receive several times the annual precipitation of those in the east (Mason et al., 2004). Soil flooding is understood to have a variety of negative impacts on plant growth, inducing stomatal closure and a decrease in the rate of photosynthesis; however, flood tolerance is recognised to vary both between and within species (Kozlowski, 1997). We undertook two studies examining the effects of prolonged waterlogging on the performance of native Scots pine populations, in order to evaluate the evidence for local adaptation. The first study was concerned with phenological and growth traits, whereas in the second I addressed a number of physiological traits.

6.2.3.1 Treatment Effects

Waterlogging had a pronounced effect on phenology, and almost all of the morphological traits measured. Bud flush was delayed, and growth was very noticeably impeded relative to the control group.

In the physiology trial, chlorophyll fluorescence was used to gauge changes in photochemical efficiency (Fv/Fm) in response to flooding. In previous studies, Fv/Fm has been shown to decrease in response to both droughting and waterlogging (Demmig et al., 1988; Huang and Gao, 1999; Garg et al., 2002; Pearson et al., 2013), and a similar effect was observed in this study. Although Fv/Fm in the flood group was depressed relative to the control, they were observed to covary over time, and within the period of the experiment the disparity between treatment groups did not increase indefinitely over time. Following prolonged exposure to flooding, δ^{13}C was elevated in the treatment group relative to the control: this is indicative of increased WUE via stomatal restriction, and is a similar response to that induced via drought (e.g. Meinzer et al., 1990).

Despite prolonged exposure to relatively severe waterlogging (in excess of a year), mortality was very low, and Scots pine saplings were found to be largely very durable to the treatment.

6.2.3.2 Site of Origin Effects

Although the treatment effect was marked for growth and bud flush traits, no treatment × population interactions were detected.

Chlorophyll fluorescence recording were taken on many occasions throughout the flood trial. In the initial analysis, where population was treated as a fixed effect, interpopulation differences were detected in several instances predominantly due to variation within the control group. However, after subsequently generalising populations into low and high rainfall groups, the latter was consistently observed to maintain a higher mean Fv/Fm under flooded conditions. This would suggest that some degree of adaptive divergence has taken place with respect to flood tolerance; although the effect was made readily detectable by restricting analysis to populations inhabiting extremes of the rainfall gradient.

Prior to initiation of the flood treatment, a longitudinal trend was observed for δ^{13}C, as described above. However, neither population differences nor environmental relationships were apparent in the measurements made post-flooding; possibly as they were either obscured, or mitigated by the treatment.

6.2.3.3 Cryptic Variation: Exposure to Novel Environments

Exposure to stress has been reported to both increase and decrease heritable variation (Hoffmann and Merilä, 1999). As natural selection reduces heritable variation (Falconer and Mackay, 1996), exposure of populations to novel environmental conditions may expose 'cryptic' variation (Rutherford and Lindquist, 1998; Schlichting, 2008), present due to a lack of previous selection pressure. Based upon this, I hypothesised that western populations may not only differ from eastern in terms of trait means under flooded conditions, but that the latter may also exhibit greater variation in response.

In terms of photochemical efficiency, this does indeed seem to be the case, as when flooded the low rainfall group exhibited consistently higher variability in response than the low rainfall group. This was not true within the control, where neither group exhibited consistently higher variability. This is interesting as it suggests that although populations may be less well adapted to a particular environment than those local to it, they may also possess a greater heritability. Should future conditions change therefore, there could be considerable adaptive potential within sites.

There is major concern regarding the impact of rapid climate change upon the distributions of plant and animal species worldwide Parmesan and Yohe (2003); Root et al.

(2003). The ability of forest trees to track changing optima will be in part due to the latent adaptive capacity of within populations (Rehfeldt et al., 2002), and measuring trait variation under novel environments provides a way in which we might measure that capacity. Flow of pre-adapted genes is an important mechanism by which the rate of adaptation may be increased (Kremer et al., 2012). This is usually discussed in the context of increasing temperatures, and the potential flow of pre-adapted genes from south to north. The Scottish pinewoods lie at least 500km from the mainland distribution, and may be effectively isolated. Average temperatures are predicted to increase, however, rainfall patterns are also anticipated to change; leading potentially to drier summers, but wetter winters (Murphy et al., 2009). Depending on the severity of these changes, longitudinal flow of genes pre-adapted to different moisture regimes may be of greater importance than latitudinal.

6.2.4 Implications for Scottish Seed Zones

The current delineation of pinewood seed zones is based upon research performed by Forrest (1980), which describes regions of biochemical similarity in terms of monoterpenes and isozymes found in leaves. This strategy been subject to criticism, firstly as the biochemical markers are not known to have an adaptive function, and secondly as the zones are not representative of environmental variation (Salmela et al., 2010). The purpose of this thesis is not to redefine the Scottish seed zones, however, the results presented do not provide support for the present boundaries. Trait differences among populations, where detected, could be explained either in terms of a longitudinal gradient (likely attributable to corresponding variation in rainfall, temperature, and altitude), or alternately by assigning populations to eastern and western extremes (although within population variances were very high as is typical with Scots Pine).

It is possible that today's seed zones may be able to explain some portion of trait variation; however, as populations within a zone tend to be geographically proximate, and therefore often subject to similar selection pressures, it would be surprising if this were not the case. In an idealised scenario, zones would represent groups of populations which possess a common set of adaptive traits. Common garden and reciprocal transplants continue to contribute to our understanding of variation in Scotland, and as a SNP-chip is currently in development (based on a large panel of transcriptome variants), it may in the near future be possible to begin identification of loci which are of adaptive significance. A zoning policy which incorporated adaptive markers would be a natural improvement.

In the absence of high-resolution data on adaptation, climate data might be used in its stead. 'Climate matching', whereby planting sites are paired with seed lots from areas which are climatically similar, is currently being promoted elsewhere, for example in the US (http://www.fs.usda.gov/ccrc/tools/seedlot-selection-tool). Zoning on the basis of climate data alone, however, carries with it the implicit assumption that seed sources *are* locally adapted, which is not necessarily true. Further to this, environmental data may be insufficiently accurate or complete: two sites which appear near identical may in reality exhibit subtle yet important differences. In the UK, the Forestry Commission has produced an Ecological Site Classification (ESC) tool (http://www.forestry.gov.uk/esc) designed to assist foresters in selecting ecologically suited species for planting sites based upon climatic and soil data (the latter of which must be provided by the user). It's conceivable that a tool such as this could be repurposed for use in assisting in choice of seed sources for Scots Pine. As the climate is changing, there remains the question of whether trees should be planted on the basis of their suitability to a predicted future climate rather than the present: so-called 'assisted migration'. This is a major topic of debate in forestry more generally (e.g. McLachlan et al., 2007), as all options possess an element of risk.

In Chapter 3 I presented a cluster analysis performed on meteorological data from the UK Met Office, based up Scotland's 84 recognised native pinewood sites. Native populations were divided into three groups: western, central, and eastern (specifically, the Cairngorm Mountains). This analysis is not intended as a guide to zoning policy; nevertheless, I anticipate that a future policy would likely define fewer, larger groups, and would at the least separate the higher altitude Cairngorm populations from the remainder of the distribution. Given the climatic distinctiveness of the west coast (both in Scotland and Europe), and the potentially disparate ancestry of the populations occupying this region, it may also be appropriate to distinguish these populations from the remainder of the distribution.

6.2.4.1 Aspects of Experimental Design: Covariates and Groups

Some of the most compelling evidence for local adaptation may be obtain via models which incorporate covariates: persuasive arguments can be made if traits can be demonstrated to covary with environmental conditions at their sites of origin. Nevertheless, it is beneficial to test for differences among populations as they may not be distributed across a continuous environmental gradient, or accurate data on site conditions may be unavailable. Additionally, genotype $\times$ environment effects may obscure relationships occurring between traits and the conditions at home sites, such that these trends are not observed within a common garden environment. By testing for group differences it

may be possible to detect interpopulation variation which has a genetic basis, even if the causes of that variation remains unclear.

When performing analysis with covariates, it is desirable to incorporate a larger number of populations as these will provide more data points upon which to fit a slope, providing a more robust result. However, if populations are considered as levels of a fixed effect, statistical power may be lost due to the large number of degrees of freedom required. AIC scores are calculated as a function of the number of model parameters: to avoid overfitting, models receive a penalty for each additional parameter incorporated (Akaike, 1974). Therefore, when large numbers of populations are included as a fixed effect, power is lost to detect small effects. This can be exacerbated if the researcher wishes to test for interactions between population and another fixed effect, as this will result in a substantial increase the number of parameters required.

One way to overcome to this problem, is by assigning populations to a smaller number of meaningful groups. For example, in Chapter 3 I perform a cluster analysis on Met Office data (Perry and Hollis, 2005) for the 84 remaining pinewood sites, assigning each to one of three climate groups. Interpopulation variance may then be accounted for as a random effect nested within climate group. This model contains considerably fewer parameters, and we are more likely to detect differences between climate groups if they exist. In addition, designs such as these may better address the hypothesis relating to adaptation and the environment, as it could be more appropriate to consider populations living in proximity under similar conditions as one, rather than several entities. Interpopulation variance should not be disregarded, however, as it may be substantial within groups.

Depending on the nature of the question, populations might be clustered into groups that reflect differences in climate, soil type, or some other aspect of the environment as long as there is a rationale for doing so. The statistical power to detect group differences can be increased, while still retaining the facility to treat populations as independent data points to model relationships with covariates.

6.3 Possibilities for Future Work

Local adaptation is said to have occurred when individuals exhibit greater average fitness on their home sites, than those introduced from other populations (Kawecki and Ebert, 2004). Long-term reciprocal transplants among Scotland's native pinewoods would therefore be highly desirable in continuing to assess the evidence for adaptive divergence. As sites may be subject to substantial temporal variation from year to year, short-term trials can provide misleading results, as maladapted individuals may perform

well until the occurrence of an intermittent event (e.g. unseasonal frost). Of particular interest would be not only differences in trait means between populations at transplant sites, but also differences in the variation of response. Do populations exhibit greater variability on foreign sites, if so for which traits, and could the magnitude of variability be explained by the disparity in environmental conditions?

Scottish populations exhibit very little geographic structuring for neutral markers (Provan et al., 1998), which would suggest that gene flow is, or has been, extensive. However, a common garden trial has shown that the phenology of native pinewoods is determined at least to some extent by altitude (Salmela et al., 2013). It would be interesting to determine to what extent the phenology of native pinewoods varies *in situ*, as this may indicate to what extent populations experience reproductive isolation in the present day. This was recommended by Salmela et al. (2011a), and at an earlier stage of this thesis there was some experimentation with on-site phenology cameras. In this case, however, the image quality and field of view was too limited to permit analysis. Nevertheless, remote sensing phenology via camera is sound in principal (e.g. Ide and Oguma, 2010), and a small network of cameras may be a good investment.

In recent years there has been a large proliferation in the availability of genome data, thanks to advent of modern sequencing technology. Among the objectives of the Procogen project (http://www.procogen.eu/), is the development of genomic resources for a number of conifers including *P. sylvestris*. In the near future we should benefit from a substantial panel of nuclear markers, and not only be able to examine the genetic structure of the native pinewoods at far greater resolution, but also take advantage of molecular approaches (e.g. association studies, Fst Scans) designed to detect loci under differential selection. In developing a deeper understanding molecular variation throughout the pinewood distribution, these techniques may also make a valuable contribution to the redesign of Scotland's seed zones.

Scotland is a small country, and the native pinewoods are restricted to a smaller area still, but it was nevertheless able to detect signatures of local adaptation. By continuing to characterise the variation among and within pinewoods, we not only gain insights into evolutionary processes, but maybe able to make better decisions in the future to preserve and restore what remains of the Caledonian Pine Forest for future generations.

Bibliography

Abramoff, M., Magalhaes, P., and Ram, S. (2004). Image processing with imagej. *Biophotonics International*, 11:36–42.

Aitken, S. N., Yeaman, S., Holliday, J. A., Wang, T., and Curtis-McLane, S. (2008). Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, 1(1):95–111.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19:716–723.

Anderson, J., Williams, J., Kriedemann, P., Austin, M., and Farquhar, G. (1996). Correlations Between Carbon Isotope Discrimination and Climate of Native Habitats for Diverse Eucalypt Taxa Growing in a Common Garden. *Functional Plant Biology*, 23(3):311–320.

Anderson, J. T., Lee, C.-R., Rushworth, C. A., Colautti, R. I., and Mitchell-Olds, T. (2013). Genetic trade-offs and conditional neutrality contribute to local adaptation. *Molecular Ecology*, 22(3):699–708.

Androsiuk, P. and Urbaniak, L. (2006). Differentiation of scots pine (pinus sylvestris l.) populations in the tatra mountains based on needle morphological traits. *Biodiversity Research and Conservation*, 3–4:227–231.

Armstrong, W., Brändle, R., and Jackson, M. B. (1994). Mechanisms of flood tolerance in plants. *Acta Botanica Neerlandica*, 43(4):307–358.

Armstrong, W. and Read, D. J. (1972). Some Observations on Oxygen Transport in Conifer Seedlings. *New Phytologist*, 71(1):55–62.

Aroca, R., Porcel, R., and Ruiz-Lozano, J. M. (2011). Regulation of root water uptake under abiotic stress conditions. *Journal of Experimental Botany*, page err266.

Badyaev, A. V. (2005). Stress-induced variation in evolution: from behavioural plasticity to genetic assimilation. *Proceedings of the Royal Society B: Biological Sciences*, 272(1566):877–886.

Balding, D. J. (2006). A tutorial on statistical methods for population association studies. *Nature Reviews Genetics*, 7(10):781–791.

Barton, K. (2013). *MuMIn: Multi-model inference*. R package version 1.9.5.

Bassett, J. R. (1964). Diameter Growth of Loblolly Pine Trees as Affected by Soil-Moisture Availibility. *New Orleans, LA: U.S. Department of Agriculture, Forest Service, Southern Forest Experiment Station*, (SO-9):7.

Bates, D., Maechler, N., and Bolker, B. (2013). *lme4: Linear mixed-effects models using S4 classes*. R package version 0.999999-2.

Becker, U., Colling, G., Dostal, P., Jakobsson, A., and Matthies, D. (2006). Local adaptation in the monocarpic perennial Carlina vulgaris at different spatial scales across Europe. *Oecologia*, 150(3):506–518.

Bell, G. (2010). Fluctuating selection: the perpetual renewal of adaptation in variable environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1537):87–97.

Belotte, D., Curien, J. B., Maclean, R. C., and Bell, G. (2003). An Experimental Test of Local Adaptation in Soil Bacteria. *Evolution*, 57(1):27–36.

Bennet, K. (1995). Post-glacial dynamics of pine *Pinus sylvestris*(l.) and pinewoods in scotland. In Aldhous, J., editor, *Our pinewood heritage*. Forestry Commission, Royal Society for the Protection of Birds, and Scottish Natural Heritage, Edinburgh.

Bennett, K. D. (1984). The post-glacial history of Pinus sylvestris in the British Isles. *Quaternary Science Reviews*, 3(2–3):133–155.

Bergthorsson, U., Adams, K. L., Thomason, B., and Palmer, J. D. (2003). Widespread horizontal transfer of mitochondrial genes in flowering plants. *Nature*, 424(6945):197–201.

Bernardi, G. (1989). The Isochore Organization of the Human Genome. *Annual Review of Genetics*, 23(1):637–659.

Berven, K. A. (1982). The Genetic Basis of Altitudinal Variation in the Wood Frog Rana sylvatica. I. An Experimental Analysis of Life History Traits. *Evolution*, 36(5):962.

Beuker, E. (1994). Adaptation to climatic changes of the timing of bud burst in populations of Pinus sylvestris L. and Picea abies (L.) Karst. *Tree Physiology*, 14(7-8-9):961–970.

Birks, H. H. (1972). Studies in the Vegetational History of Scotland: II. Two Pollen Diagrams from the Galloway Hills, Kirkcudbrightshire. *Journal of Ecology*, 60(1):183–217.

Birks, H. J. B. (1989). Holocene Isochrone Maps and Patterns of Tree-Spreading in the British Isles. *Journal of Biogeography*, 16(6):503–540.

Björkman, O. and Demmig, B. (1987). Photon yield of O2 evolution and chlorophyll fluorescence characteristics at 77 K among vascular plants of diverse origins. *Planta*, 170(4):489–504.

Bobowicz, M. A. and Korczyk, A. F. (1994). Interpopulational variability of Pinus sylvestris L. in eight Polish localities expressed in morphological and anatomical traits of needles. *Acta Societatis Botanicorum Poloniae*, 63(1):67–76.

Bobowicz, M. A. and Radziejewska, A. (1989). The variability of Scots pine from Piekielna Góra as expressed by morphological and anatomical traits of needles. *Acta Societatis Botanicorum Poloniae*, 58(3):375–384.

Bogunic, F., Muratovic, E., Brown, S. C., and Siljak-Yakovlev, S. (2003). Genome size and base composition of five Pinus species from the Balkan region. *Plant Cell Reports*, 22(1):59–63.

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics (Oxford, England)*, 30(15):2114–2120.

Boulli, A., Baaziz, M., and M'Hirit, O. (2001). Polymorphism of natural populations of Pinus halepensis Mill. in Morocco as revealed by morphological characters. *Euphytica*, 119(3):309–316.

Bradford, K. J. and Hsiao, T. C. (1982). Stomatal Behavior and Water Relations of Waterlogged Tomato Plants. *Plant Physiology*, 70(5):1508–1513.

Brendel, O., Handley, L., and Griffiths, H. (2002). Differences in δ^{13}c and diameter growth among remnant Scots pine populations in Scotland. *Tree Physiology*, 22(14):983–992.

Brown, J. S. and Pavlovic, N. B. (1992). Evolution in heterogeneous environments: Effects of migration on habitat specialization. *Evolutionary Ecology*, 6(5):360–382.

Burgstaller, J. P., Schinogl, P., Dinnyes, A., Müller, M., and Steinborn, R. (2007). Mitochondrial DNA heteroplasmy in ovine fetuses and sheep cloned by somatic cell nuclear transfer. *BMC Developmental Biology*, 7(1):141.

Burnham, K. and Anderson, D. (2004). Multimodel inference: understanding aic and bic in model selection. *Sociological Methods and Research*, 33:261—304.

Campbell, I. D., McDonald, K., Flannigan, M. D., and Kringayark, J. (1999). Long-distance transport of pollen into the Arctic. *Nature*, 399(6731):29–30.

Carpenter, S. B. and Smith, N. D. (1975). Stomatal distribution and size in southern Appalachian hardwoods. *Canadian Journal of Botany*, 53(11):1153–1156.

Chapin, III, F. S. (1991). Integrated Responses of Plants to Stress. *BioScience*, 41(1):29–36.

Charmantier, A. and Garant, D. (2005). Environmental quality and evolutionary potential: lessons from wild populations. *Proceedings of the Royal Society B: Biological Sciences*, 272(1571):1415–1425.

Chaw, S.-M., Shih, A. C.-C., Wang, D., Wu, Y.-W., Liu, S.-M., and Chou, T.-Y. (2008). The mitochondrial genome of the gymnosperm Cycas taitungensis contains a novel family of short interspersed elements, Bpu sequences, and abundant RNA editing sites. *Molecular biology and evolution*, 25(3):603–615.

Cheddadi, R., Vendramin, G. G., Litt, T., François, L., Kageyama, M., Lorentz, S., Laurent, J.-M., De Beaulieu, J.-L., Sadori, L., Jost, A., and Lunt, D. (2006). Imprints of glacial refugia in the modern genetic diversity of Pinus sylvestris. *Global Ecology and Biogeography*, 15(3):271–282.

Chirkova, T. and Gutman, T. (1972). Physiological role of branch lenticels in willow and poplar under conditions of root anaerobiosis. *Sov.Plant Physiol. (Eng.Transl.*, 19:289–295.

Christiansen, E., Waring, R. H., and Berryman, A. A. (1987). Resistance of conifers to bark beetle attack: Searching for general relationships. *Forest Ecology and Management*, 22(1–2):89–106.

Clark, P. U., Dyke, A. S., Shakun, J. D., Carlson, A. E., Clark, J., Wohlfarth, B., Mitrovica, J. X., Hostetler, S. W., and McCabe, A. M. (2009). The Last Glacial Maximum. *Science*, 325(5941):710–714.

Coley, P. D. (1988). Effects of plant growth rate and leaf lifetime on the amount and type of anti-herbivore defense. *Oecologia*, 74(4):531–536.

Collaku, A. and Harrison, S. A. (2005). Heritability of Waterlogging Tolerance in Wheat. *Crop Science*, 45(2):722.

Colmer, T. D. (2003). Long-distance transport of gases in plants: a perspective on internal aeration and radial oxygen loss from roots. *Plant, Cell & Environment*, 26(1):17–36.

Condon, A. and Richards, R. (1992). Broad sense heritability and genotype × environment interaction for carbon isotope discrimination in field-grown wheat. *Australian Journal of Agricultural Research*, 43(5):921–934.

Correia, I., Almeida, M. H., Aguiar, A., Alía, R., David, T. S., and Pereira, J. S. (2008). Variations in growth, survival and carbon isotope composition (δ^{13}c) among Pinus pinaster populations of different geographic origins. *Tree Physiology*, 28(10):1545–1552.

Craig, A. J. (1978). Pollen Percentage and Influx Analyses in South-East Ireland: A Contribution to the Ecological History of the Late-Glacial Period. *Journal of Ecology*, 66(1):297–324.

Cregg, B. M., Dougherty, P. M., and Hennessey, T. C. (1988). Growth and wood quality of young loblolly pine trees in relation to stand density and climatic factors. *Canadian Journal of Forest Research*, 18(7):851–858.

Cunningham, S. A., Summerhayes, B., and Westoby, M. (1999). Evolutionary Divergences in Leaf Structure and Chemistry, Comparing Rainfall and Soil Nutrient Gradients. *Ecological Monographs*, 69(4):569–588. ArticleType: research-article / Full publication date: Nov., 1999 / Copyright © 1999 Ecological Society of America.

Danusevičius, D., Buchovska, J., Stanys, V., Šikšnianienė, J. B., Marozas, V., and Bendokas, V. (2013). DNA marker based identification of spontaneous hybrids between Pinus mugo and P. sylvestris at the Lithuanian sea-side. *Nordic Journal of Botany*, 31(3):344–352.

Daugherty, C. J. and Musgrave, M. E. (1994). Characterization of populations of rapid-cycling Brassica rapa L. selected for differential waterlogging tolerance. *Journal of Experimental Botany*, 45(3):385–392.

Day, T. (2000). Competition and the Effect of Spatial Resource Heterogeneity on Evolutionary Diversification. *The American Naturalist*, 155(6):790–803.

De Carvalho, D., Ingvarsson, P. K., Joseph, J., Suter, L., Sedivy, C., Macaya-Sanz, D., Cottrell, J., Heinze, B., Schanzer, I., and Lexer, C. (2010). Admixture facilitates adaptation from standing variation in the European aspen (Populus tremula L.), a widespread forest tree. *Molecular Ecology*, 19(8):1638–1650.

De Mita, S., Thuillet, A.-C., Gay, L., Ahmadi, N., Manel, S., Ronfort, J., and Vigouroux, Y. (2013). Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, 22(5):1383–1399.

Demmig, B., Winter, K., Krüger, A., and Czygan, F.-C. (1988). Zeaxanthin and the Heat Dissipation of Excess Light Energy in Nerium oleander Exposed to a Combination of High Light and Water Stress. *Plant Physiology*, 87(1):17–24.

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., Philippakis, A. A., del Angel, G., Rivas, M. A., Hanna, M., McKenna, A., Fennell, T. J., Kernytsky, A. M., Sivachenko, A. Y., Cibulskis, K., Gabriel, S. B., Altshuler, D., and Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, 43(5):491–498.

Dillen, S. Y., Marron, N., Koch, B., and Ceulemans, R. (2008). Genetic Variation of Stomatal Traits and Carbon Isotope Discrimination in Two Hybrid Poplar Families (Populus deltoides 'S9-2' x P. nigra 'Ghoy' and P. deltoides 'S9-2' x P. trichocarpa 'V24'). *Annals of Botany*, 102(3):399–407.

Dombrovska, O. and Qiu, Y.-L. (2004). Distribution of introns in the mitochondrial gene nad1 in land plants: phylogenetic and molecular evolutionary implications. *Molecular Phylogenetics and Evolution*, 32(1):246–263.

Donahue, J. K. and Upton, J. L. (1996). Geographic variation in leaf, cone and seed morphology of Pinus greggii in native forests. *Forest Ecology and Management*, 82(1–3):145–157.

Dougherty, P. M., Whitehead, D., and Vose, J. M. (1994). Environmental Influences on the Phenology of Pine. *Ecological Bulletins*, (43):64–75.

Drouin, G., Daoud, H., and Xia, J. (2008). Relative rates of synonymous substitutions in the mitochondrial, chloroplast and nuclear genomes of seed plants. *Molecular Phylogenetics and Evolution*, 49(3):827–831.

Dubois, A. D. and Ferguson, D. K. (1985). The climatic history of pine in the cairngorms based on radiocarbon dates and stable isotope analysis, with an account of the events leading up to its colonization. *Review of Palaeobotany and Palynology*, 46(1–2):55–80.

Dunlap, J. M. and Stettler, R. F. (2001). Variation in leaf epidermal and stomatal traits of Populus trichocarpa from two transects across the Washington Cascades. *Canadian Journal of Botany*, 79(5):528–536.

Ehdaie, B., Hall, A. E., Farquhar, G. D., Nguyen, H. T., and Waines, J. G. (1991). Water-Use Efficiency and Carbon Isotope Discrimination in Wheat. *Crop Science*, 31(5):1282.

Ehleringer, J. R. and Cooper, T. A. (1988). Correlations between carbon isotope ratio and microhabitat in desert plants. *Oecologia*, 76(4):562–566.

Eriksson, G., Andersson, S., Eiche, V., Ifver, J., and Persson, A. (1980). *Severity index and transfer effects on survival and volume production of Pinus sylvestris in northern Sweden*. Sveriges lantbruksuniversitet.

Falconer, D. and Mackay, T. (1996). *Introduction to Quantitative Genetics (fourth edition)*. Longman Group.

Farquhar, G. D., Hubick, K. T., Condon, A. G., and Richards, R. A. (1989). Carbon Isotope Fractionation and Plant Water-Use Efficiency. In Rundel, P. W., Ehleringer, J. R., and Nagy, K. A., editors, *Stable Isotopes in Ecological Research*, number 68 in Ecological Studies, pages 21–40. Springer New York.

Fonseca, C. R., Overton, J. M., Collins, B., and Westoby, M. (2000). Shifts in trait-combinations along rainfall and phosphorus gradients. *Journal of Ecology*, 88(6):964–977.

Forrest, G. I. (1980). Genotypic Variation among Native Scots Pine Populations in Scotland based on Monoterpene Analysis. *Forestry*, 53(2):101–128.

Forrest, G. I. (1982). Relationship of Some European Scots Pine Populations to Native Scottish Woodlands Based on Monoterpene Analysis. *Forestry*, 55(1):19–37.

Fraley, C., Raftery, A., and Luca Scrucca, L. (2012). *mclust Version 4 for R: Normal Mixture Modeling for Model-Based Clustering, Classification, and Density Estimation*. R package version 2.15.

Fry, J. D. (1996). The Evolution of Host Specialization: Are Trade-Offs Overrated? *The American Naturalist*, 148:S84–S107.

Galloway, L. F. and Fenster, C. B. (2000). Population Differentiation in an Annual Legume: Local Adaptation. *Evolution*, 54(4):1173–1181.

Gamon, J. A. and Pearcy, R. W. (1989). Leaf movement, stress avoidance and photosynthesis in Vitis californica. *Oecologia*, 79(4):475–481.

García-Ramos, G. and Kirkpatrick, M. (1997). Genetic Models of Adaptation and Gene Flow in Peripheral Populations. *Evolution*, 51(1):21–28.

Garg, A. K., Kim, J.-K., Owens, T. G., Ranwala, A. P., Choi, Y. D., Kochian, L. V., and Wu, R. J. (2002). Trehalose accumulation in rice plants confers high tolerance levels to different abiotic stresses. *Proceedings of the National Academy of Sciences*, 99(25):15898–15903.

Genty, B., Briantais, J.-M., and Baker, N. R. (1989). The relationship between the quantum yield of photosynthetic electron transport and quenching of chlorophyll fluorescence. *Biochimica et Biophysica Acta (BBA) - General Subjects*, 990(1):87–92.

Gibson, G. and Dworkin, I. (2004). Uncovering cryptic genetic variation. *Nature Reviews Genetics*, 5(9):681–690.

Givnish, T. J. (1987). Comparative Studies of Leaf Form: Assessing the Relative Roles of Selective Pressures and Phylogenetic Constraints. *New Phytologist*, 106:131–160.

Godbout, J., Jaramillo-Correa, J. P., Beaulieu, J., and Bousquet, J. (2005). A mitochondrial DNA minisatellite reveals the postglacial history of jack pine (Pinus banksiana), a broad-range North American conifer. *Molecular Ecology*, 14(11):3497–3512.

Gomes, A. R. S. and Kozlowski, T. T. (1980). Growth Responses and Adaptations of Fraxinus pennsylvanica Seedlings to Flooding. *Plant Physiology*, 66(2):267–271.

Grace, J. (1997). The oceanic tree-line and the limit for tree growth in Scotland. *Botanical Journal of Scotland*, 49(2):223–236.

Grime, J. P. (1977). Evidence for the Existence of Three Primary Strategies in Plants and Its Relevance to Ecological and Evolutionary Theory. *The American Naturalist*, 111(982):1169–1194.

Groom, Q. J. and Baker, N. R. (1992). Analysis of Light-Induced Depressions of Photosynthesis in Leaves of a Wheat Crop during the Winter. *Plant Physiology*, 100(3):1217–1223.

Grubb, P. J. (1985). Plant Populations and Vegetation in Relation to Habitat, Disturbance and Competition: Problems of Generalization. In White, J., editor, *The Population Structure of Vegetation*, number 3 in Handbook of Vegetation Science, pages 595–621. Springer Netherlands.

Grubb, P. J. (1998). A reassessment of the strategies of plants which cope with shortages of resources. *Perspectives in Plant Ecology, Evolution and Systematics*, 1(1):3–31.

Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics (Oxford, England)*, 29(8):1072–1075.

Hall, A. E., Mutters, R. G., Hubick, K. T., and Farquhar, G. D. (1990). Genotypic Differences in Carbon Isotope Discrimination by Cowpea under Wet and Dry Field Conditions. *Crop Science*, 30(2):300.

Hall, D., Scurlock, J., Bolhar-Nordenkampf, H., Leegood, R., and Long, S. (1993). *Photosynthesis and Production in a Changing Environment - A Field and Laboratory Manual*. Chapman and Hall.

Hall, M. C., Lowry, D. B., and Willis, J. H. (2010). Is local adaptation in Mimulus guttatus caused by trade-offs at individual loci? *Molecular Ecology*, 19(13):2739–2753.

Hamrick, J. L., Godt, M. J. W., and Sherman-Broyles, S. L. (1992). Factors influencing levels of genetic diversity in woody plant species. *New Forests*, 6(1-4):95–124.

Hansen, J. and Ernstsen, A. (1982). Seasonal changes in adventitious root formation in hypocotyl cuttings of Pinus sylvestris: Influence of Photoperiod during stock plant growth and of indolebutyric acid treatment of cuttings. *Physiologia Plantarum*, 54(2):99–106.

Hedrick, P. W., Ginevan, M. E., and Ewing, E. P. (1976). Genetic Polymorphism in Heterogeneous Environments. *Annual Review of Ecology and Systematics*, 7:1–32.

Hoffmann, A. A. and Merilä, J. (1999). Heritable variation and evolution under favourable and unfavourable conditions. *Trends in Ecology & Evolution*, 14(3):96–101.

Hogan, K. P., Smith, A. P., Araus, J. L., and Saavedra, A. (1994). Ecotypic differentiation of gas exchange responses and leaf anatomy in a tropical forest understory shrub from areas of contrasting rainfall regimes. *Tree Physiology*, 14(7-8-9):819–831.

Hook, D. D., Brown, C. L., and Kormanik, P. P. (1971). Inductive Flood Tolerance in Swamp Tupelo (Nyssa sylvatica var. biflora (Walt.) Sarg.). *Journal of Experimental Botany*, 22(1):78–89.

Houle, D. (1992). Comparing evolvability and variability of quantitative traits. *Genetics*, 130(1):195–204.

Huang, B. and Gao, H. (1999). Physiological Responses of Diverse Tall Fescue Cultivars to Drought Stress. *HortScience*, 34(5):897–901.

Hultine, K. R. and Marshall, J. D. (2000). Altitude trends in conifer leaf morphology and stable carbon isotope composition. *Oecologia*, 123(1):32–40.

Huntley, B. and Birks, H. (1983a). An Atlas of past and present pollen maps for Europe: 0 - 13000 years ago.

Huntley, B. and Birks, H. (1983b). *An atlas of past and present pollen maps for Europe: 0-13000 years ago*. Cambridge University Press.

Hussonm, F., Josse, J., Le, S., and Mazet, J. (2013). *FactoMineR: Multivariate Exploratory Data Analysis and Data Mining with R*. R package version 1.25.

Ide, R. and Oguma, H. (2010). Use of digital cameras for phenological observations. *Ecological Informatics*, 5(5):339–347.

Impa, S. M., Nadaradjan, S., Boominathan, P., Shashidhar, G., Bindumadhava, H., and Sheshshayee, M. S. (2005). Carbon Isotope Discrimination Accurately Reflects Variability in WUE Measured at a Whole Plant Level in Rice. *Crop Science*, 45(6):2517.

Inouye, D. W. and Wielgolaski, F. E. (2013). Phenology at High Altitudes. In Schwartz, M. D., editor, *Phenology: An Integrative Environmental Science*, pages 249–272. Springer Netherlands.

Jain, S. M. and Minocha, S. C. (1999). *Molecular Biology of Woody Plants*. Springer Science & Business Media.

James, J. C., Grace, J., and Hoad, S. P. (1994). Growth and Photosynthesis of Pinus Sylvestris at its Altitudinal Limit in Scotland. *Journal of Ecology*, 82(2):297–306. ArticleType: research-article / Full publication date: Jun., 1994 / Copyright © 1994 British Ecological Society.

Jaramillo-Correa, J. P., Beaulieu, J., and Bousquet, J. (2004). Variation in mitochondrial DNA reveals multiple distant glacial refugia in black spruce (Picea mariana), a transcontinental North American conifer. *Molecular Ecology*, 13(9):2735–2747.

Jeffree, C. E., Dale, J. E., and Fry, S. C. (1986). The genesis of intercellular spaces in developing leaves ofPhaseolus vulgaris L. *Protoplasma*, 132(1-2):90–98.

Johansen, A. D. and Latta, R. G. (2003). Mitochondrial haplotype distribution, seed dispersal and patterns of postglacial expansion of ponderosa pine. *Molecular Ecology*, 12(1):293–298.

Kaufmann, M. R. (1968). Water Relations of Pine Seedlings in Relation to Root and Shoot Growth 1. *Plant Physiology*, 43(2):281–288.

Kautsky, H., Appel, W., and Amann, H. (1960). Chlorophyll fluorescenz und kohlensaure assimilation. *Biochemische Zeitschrift*, 322:277–292.

Kawase, M. (1979). Cellulase activity in waterlogged herbaceous horticultural crops. *HortScience*, 16:30–34.

Kawecki, T. J. and Ebert, D. (2004). Conceptual issues in local adaptation. *Ecology Letters*, 7(12):1225–1241.

Kinloch, B. B., Westfall, R. D., and Forrest, G. I. (1986). Caledonian Scots Pine: Origins and Genetic Structure. *New Phytologist*, 104(4):703–729.

Kmiec, B., Woloszynska, M., and Janska, H. (2006). Heteroplasmy as a common state of mitochondrial genetic information in plants and animals. *Current Genetics*, 50(3):149–159.

Knoop, V. (2004). The mitochondrial DNA of land plants: peculiarities in phylogenetic perspective. *Current Genetics*, 46(3):123–139.

Korol, R. L., Kirschbaum, M. U. F., Farquhar, G. D., and Jeffreys, M. (1999). Effects of water status and soil fertility on the C-isotope signature in Pinus radiata. *Tree Physiology*, 19(9):551–562.

Kozlowski, T. (1997). Responses of woody plants to flooding and salinity - tree physiology monograph no 1. *http://www.heronpublishing.com/tp/monograph/kozlowski.pdf*.

Kozlowski, T. T. (1984). *Flooding and Plant Growth*. Academic Press.

Kramer, K., Leinonen, I., and Loustau, D. (2000). The importance of phenology for the evaluation of impact of climate change on growth of boreal, temperate and Mediterranean forests ecosystems: an overview. *International Journal of Biometeorology*, 44(2):67–75.

Kremer, A., Ronce, O., Robledo-Arnuncio, J. J., Guillaume, F., Bohrer, G., Nathan, R., Bridle, J. R., Gomulkiewicz, R., Klein, E. K., Ritland, K., Kuparinen, A., Gerber, S., and Schueler, S. (2012). Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecology letters*.

Körner, C., Farquhar, G. D., and Roksandic, Z. (1988). A global survey of carbon isotope discrimination in plants from high altitude. *Oecologia*, 74(4):623–632.

Lande, R. (1977). On comparing coeffcients of variation. *Systematic Zoology*, 26:214–217.

Levene, H. (1953). Genetic Equilibrium When More Than One Ecological Niche is Available. *The American Naturalist*, 87(836):331–333.

Lewandowski, A., Boratyński, A., and Mejnartowicz, L. (2000). Allozyme investigations on the genetic differentiation between closely related pines —Pinus sylvestris, P. mugo,

P. uncinata, andP. uliginosa (Pinaceae). *Plant Systematics and Evolution*, 221(1-2):15–24.

Lewontin, R. C. and Krakauer, J. (1973). Distribution of Gene Frequency as a Test of the Theory of the Selective Neutrality of Polymorphisms. *Genetics*, 74(1):175–195.

Lewty, M. J. (1990). Effects of waterlogging on the growth and water relations of three Pinus taxa. *Forest Ecology and Management*, 30(1–4):189–201.

Li, H. (2012). Exploring single-sample SNP and INDEL calling with whole-genome de novo assembly. *Bioinformatics*, 28(14):1838–1844.

Li, H. and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)*, 25(14):1754–1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Subgroup, . G. P. D. P. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16):2078–2079.

Lieth, H. (1974). Phenology and Seasonality Modeling. In Lieth, H., editor, *Phenology and Seasonality Modeling*, number 8 in Ecological Studies, pages 3–19. Springer Berlin Heidelberg.

Limin, Y., Mei, H., Guangsheng, Z., and Jiandong, L. (2007). The changes in water-use efficiency and stoma density of Leymus chinensis along Northeast China Transect. *Acta Ecologica Sinica*, 27(1):16–23.

Linder, S., Benson, M. L., Myers, B. J., and Raison, R. J. (1987). Canopy dynamics and growth of Pinusradiata.: I. Effects of irrigation and fertilization during a drought. *Canadian Journal of Forest Research*, 17(10):1157–1165.

Lindgren, K. and Hällgren, J.-E. (1993). Cold acclimation of Pinus contorta and Pinus sylvestris assessed by chlorophyll fluorescence. *Tree Physiology*, 13(1):97–106.

Lines, R. and Mitchell, A. (1965). Forestry commission report on forest research 19641965: Results of some older scots pine provenance experiments. *London: HMSO*, pages 172–194.

Lombardini, L., Restrepo-Diaz, H., and Volder, A. (2009). Photosynthetic Light Response and Epidermal Characteristics of Sun and Shade Pecan Leaves. *Journal of the American Society for Horticultural Science*, 134(3):372–378.

Manhart, J. R. and Palmer, J. D. (1990). The gain of two chloroplast tRNA introns marks the green algal ancestors of land plants. *Nature*, 345(6272):268–270.

Marshall, J. D. and Zhang, J. (1994). Carbon Isotope Discrimination and Water-Use Efficiency in Native Plants of the North-Central Rockies. *Ecology*, 75(7):1887–1895. ArticleType: research-article / Full publication date: Oct., 1994 / Copyright © 1994 Ecological Society of America.

Mason, W., Hampson, A., and Edwards, C. (2004). *Managing the Pinewoods of Scotland*. Forestry Commission.

Matziris, D. (1982). Genetic variation in number of resin canals per needle of the Black Pine of Peloponnese. *Dasiki Erevna (Greece)*.

Maxwell, K. and Johnson, G. N. (2000). Chlorophyll fluorescence—a practical guide. *Journal of Experimental Botany*, 51(345):659–668.

McGuigan, K. and Sgrò, C. M. (2009). Evolutionary consequences of cryptic genetic variation. *Trends in Ecology & Evolution*, 24(6):305–311.

McKay, J. K., Christian, C. E., Harrison, S., and Rice, K. J. (2005). "How Local Is Local?"—A Review of Practical and Conceptual Issues in the Genetics of Restoration. *Restoration Ecology*, 13(3):432–440.

McLachlan, J. S., Hellmann, J. J., and Schwartz, M. W. (2007). A Framework for Debate of Assisted Migration in an Era of Climate Change. *Conservation Biology*, 21(2):297–302.

McNulty, S. G. and Swank, W. T. (1995). Wood (delta)^(13)C as a Measure of Annual Basal Area Growth and Soil Water Stress in a Pinus Strobus Forest. *Ecology*, 76(5):1581–1586.

McVean, D. N. and Ratcliffe, D. A. (1962). Plant communities of the Scottish Highlands. A study of Scottish mountain, moorland and forest vegetation. *Monographs of the Nature Conservancy, No. 1.*, page 458 pp.

Meinzer, F. C., Goldstein, G., and Grantz, D. A. (1990). Carbon Isotope Discrimination in Coffee Genotypes Grown under Limited Water Supply. *Plant Physiology*, 92(1):130–135.

Meinzer, F. C., Rundel, P. W., Goldstein, G., and Sharifi, M. R. (1992). Carbon isotope composition in relation to leaf gas exchange and environmental conditions in Hawaiian Metrosideros polymorpha populations. *Oecologia*, 91(3):305–311.

MIller, J. M., Williams, R. J., and Farquhar, G. D. (2001). Carbon isotope discrimination by a sequence of Eucalyptus species along a subcontinental rainfall gradient in Australia. *Functional Ecology*, 15(2):222–232.

Milne, I., Stephen, G., Bayer, M., Cock, P. J. A., Pritchard, L., Cardle, L., Shaw, P. D., and Marshall, D. (2013). Using Tablet for visual exploration of second-generation sequencing data. *Briefings in Bioinformatics*, 14(2):193–202.

Mitchell-Olds, T., Willis, J. H., and Goldstein, D. B. (2007). Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nature Reviews Genetics*, 8(11):845–856.

Mopper, S. and Strauss, S. Y. (1998). Genetic structure and local adaptation in natural insect populations: effects of ecology, life history, and behavior. pages xix + 449 pp.

Murphy, J., Sexton, D., Jenkins, G., Boorman, P., Booth, B., Brown, C., Clark, R., Collins, M., Harris, G., Kendon, E., Betts, R., Brown, S., Howard, T. P., Humphrey, K. A., McCarthy, M. P., McDonald, R. E., Stephens, A., Wallace, C., Warren, R., Wilby, R., and Wood, R. A. (2009). Uk climate projections science report: Climate change projections. *Met Office, Hadley Centre, Exeter*.

Murray, M. B., Cannell, M. G. R., and Smith, R. I. (1989). Date of Budburst of Fifteen Tree Species in Britain Following Climatic Warming. *The Journal of Applied Ecology*, 26(2):693.

Nakagawa, S. and Schielzeth, H. (2013). A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2):133–142.

Naydenov, K., Senneville, S., Beaulieu, J., Tremblay, F., and Bousquet, J. (2007). Glacial vicariance in Eurasia: mitochondrial DNA evidence from Scots pine for a complex heritage involving genetically distinct refugia at mid-northern latitudes and in Asia Minor. *BMC Evolutionary Biology*, 7(1):233.

Neale, D. B. and Sederoff, R. R. (1988). Inheritance and Evolution of Conifer Organelle Genomes. In Hanover, J. W., Keathley, D. E., Wilson, C. M., and Kuny, G., editors, *Genetic Manipulation of Woody Plants*, number 44 in Basic Life Sciences, pages 251–264. Springer US.

Neale, D. B., Wegrzyn, J. L., Stevens, K. A., Zimin, A. V., Puiu, D., Crepeau, M. W., Cardeno, C., Koriabine, M., Holtz-Morris, A. E., Liechty, J. D., Martínez-García, P. J., Vasquez-Gross, H. A., Lin, B. Y., Zieve, J. J., Dougherty, W. M., Fuentes-Soriano, S., Wu, L.-S., Gilbert, D., Marçais, G., Roberts, M., Holt, C., Yandell, M., Davis, J. M., Smith, K. E., Dean, J. F., Lorenz, W. W., Whetten, R. W., Sederoff, R., Wheeler, N., McGuire, P. E., Main, D., Loopstra, C. A., Mockaitis, K., deJong, P. J., Yorke, J. A., Salzberg, S. L., and Langley, C. H. (2014). Decoding the massive

genome of loblolly pine using haploid DNA and novel assembly strategies. *Genome Biology*, 15(3):R59.

Niederstätter, H., Coble, M. D., Parsons, T. J., and Parson, W. (2006). Characterization of mtDNA SNP typing using quantitative real-time PCR with special emphasis on heteroplasmy detection and mixture ratio assessment. *International Congress Series*, 1288:1–3.

Oda, K., Yamato, K., Ohta, E., Nakamura, Y., Takemura, M., Nozato, N., Akashi, K., Kanegae, T., Ogura, Y., and Kohchi, T. (1992). Gene organization deduced from the complete sequence of liverwort Marchantia polymorpha mitochondrial DNA. A primitive form of plant mitochondrial genome. *Journal of Molecular Biology*, 223(1):1–7.

Oliver, J. L. and Marín, A. (1996). A relationship between GC content and coding-sequence length. *Journal of Molecular Evolution*, 43(3):216–223.

Pallardy, S. (2008). *Physiology of Woody Plants (third edition)*. Elsevier.

Palmer, J. D. (1985). Comparative Organization of Chloroplast Genomes. *Annual Review of Genetics*, 19(1):325–354.

Palmer, J. D. and Herbon, L. A. (1988). Plant mitochondrial DNA evolved rapidly in structure, but slowly in sequence. *Journal of Molecular Evolution*, 28(1-2):87–97.

Pammenter, N. W., Drennan, P. M., and Smith, V. R. (1986). Physiological and Anatomical Aspects of Photosynthesis of Two Agrostis Species at a Sub-Antarctic Island. *New Phytologist*, 102(1):143–160.

Panek, J. A. and Waring, R. H. (1997). Stable Carbon Isotopes as Indicators of Limitations to Forest Growth Imposed By Climate Stress. *Ecological Applications*, 7(3):854–863.

Parmesan, C. and Yohe, G. (2003). A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421(6918):37–42.

Pearl, S. A., Welch, M. E., and McCauley, D. E. (2009). Mitochondrial Heteroplasmy and Paternal Leakage in Natural Populations of Silene vulgaris, a Gynodioecious Plant. *Molecular Biology and Evolution*, 26(3):537–545.

Pearson, M., Saarinen, M., Nummelin, L., Heiskanen, J., Roitto, M., Sarjala, T., and Laine, J. (2013). Tolerance of peat-grown Scots pine seedlings to waterlogging and drought: Morphological, physiological, and metabolic responses to stress. *Forest Ecology and Management*, 307:43–53.

Perks, M. and Ennos, R. (1999). Analysis of genetic variation for quantitative characters between and within four native populations of scots pine (Pinus sylvestris). *Botanical Journal of Scotland*, 51(1):103–110.

Perks, M. P. and Mckay, H. M. (1997). Morphological and physiological differences in Scots pine seedlings of six seed origins. *Forestry*, 70(3):223–232.

Perry, M. and Hollis, D. (2005). The development of a new set of long-term climate averages for the uk. *International Journal of Climatology*, 25:1023–1039.

Persson, B. (1994). Effects of provenance transfer on survival in nine experimental series with Pinus sylvestris (L). In northern Sweden. *Scandinavian Journal of Forest Research*, 9(1-4):275–287.

Ponnamperuma, F. (1972). The chemistry of submerged soils. *Adv.Agron.*, 24:29–96.

Provan, J., Soranzo, N., Wilson, N. J., McNicol, J. W., Forrest, G. I., Cottrell, J., and Powell, W. (1998). Gene–pool variation in Caledonian and European Scots pine (Pinus sylvestris L.) revealed by chloroplast simple–sequence repeats. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1407):1697–1705.

Prus-Glowacki, W. and Stephan, B. (1994). Genetic variation of *Pinus sylvestris* from spain in relation to other european populations. *Silvae Genetica*, 3:7–14.

Pyhäjärvi, T., Salmela, M. J., and Savolainen, O. (2008). Colonization routes of Pinus sylvestris inferred from distribution of mitochondrial DNA variation. *Tree Genetics & Genomes*, 4(2):247–254.

Quinlan, A. R. and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6):841–842.

R Development Core Team (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Radin, J. W. and Ackerson, R. C. (1981). Water Relations of Cotton Plants under Nitrogen Deficiency III. Stomatal Conductance, Photosynthesis, and Abscisic Acid Accumulation During Drought. *Plant Physiology*, 67(1):115–119.

Rehfeldt, G. E., Tchebakova, N. M., Parfenova, Y. I., Wykoff, W. R., Kuzmina, N. A., and Milyutin, L. I. (2002). Intraspecific responses to climate in Pinus sylvestris. *Global Change Biology*, 8(9):912–929.

Reich, P., Wright, I., Cavender-Bares, J., Craine, J., Oleksyn, J., Westoby, M., and Walters, M. (2003). The Evolution of Plant Functional Variation: Traits, Spectra, and Strategies. *International Journal of Plant Sciences*, 164(S3):S143–S164.

Reich, P. B., Uhl, C., Walters, M. B., and Ellsworth, D. S. (1991). Leaf lifespan as a determinant of leaf structure and function among 23 amazonian tree species. *Oecologia*, 86(1):16–24.

Reich, P. B., Walters, M. B., and Ellsworth, D. S. (1992). Leaf Life-Span in Relation to Leaf, Plant, and Stand Characteristics among Diverse Ecosystems. *Ecological Monographs*, 62(3):365–392.

Rodríguez-Moreno, L., González, V. M., Benjak, A., Martí, M. C., Puigdomènech, P., Aranda, M. A., and Garcia-Mas, J. (2011). Determination of the melon chloroplast and mitochondrial genome sequences reveals that the largest reported mitochondrial genome in plants contains a significant amount of DNA having a nuclear origin. *BMC Genomics*, 12(1):424.

Root, T. L., Price, J. T., Hall, K. R., Schneider, S. H., Rosenzweig, C., and Pounds, J. A. (2003). Fingerprints of global warming on wild animals and plants. *Nature*, 421(6918):57–60.

Rutherford, S. L. and Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. *Nature*, 396(6709):336–342.

Ryser, P. (1996). The Importance of Tissue Density for Growth and Life Span of Leaves and Roots: A Comparison of Five Ecologically Contrasting Grasses. *Functional Ecology*, 10(6):717.

Sabir, J. S. M., Arasappan, D., Bahieldin, A., Abo-Aba, S., Bafeel, S., Zari, T. A., Edris, S., Shokry, A. M., Gadalla, N. O., Ramadan, A. M., Atef, A., Al-Kordy, M. A., El-Domyati, F. M., and Jansen, R. K. (2014). Whole Mitochondrial and Plastid Genome SNP Analysis of Nine Date Palm Cultivars Reveals Plastid Heteroplasmy and Close Phylogenetic Relationships among Cultivars. *PLoS ONE*, 9(4):e94158.

Salmela, M., Cavers, S., Cottrell, J., and Ennos, R. (2011a). Phd thesis: "adaptive genetic variation in scots pine (*Pinus sylvestris* l.) in scotland".

Salmela, M. J., Cavers, S., Cottrell, J. E., Iason, G. R., and Ennos, R. A. (2011b). Seasonal patterns of photochemical capacity and spring phenology reveal genetic differentiation among native Scots pine (Pinus sylvestris L.) populations in Scotland. *Forest Ecology and Management*, 262(6):1020–1029.

Salmela, M. J., Cavers, S., Cottrell, J. E., Iason, G. R., and Ennos, R. A. (2013). Spring phenology shows genetic variation among and within populations in seedlings of Scots pine (Pinus sylvestris L.) in the Scottish Highlands. *Plant Ecology & Diversity*, 6(3-4):523–536.

Salmela, M. J., Cavers, S., Wachowiak, W., Cottrell, J. E., Iason, G. R., and Ennos, R. A. (2010). Understanding the evolution of native pinewoods in Scotland will benefit their future management and conservation. *Forestry*, 83(5):535–545.

Sands, K. and Rutter, A. J. (1959). Studies in the Growth of Young Plants of Pinus sylvestris L. II. The Relation of Growth to Soil Moisture Tension. *Annals of Botany*, 23(2):269–284.

Savolainen, O. (2011). The Genomic Basis of Local Climatic Adaptation. *Science*, 334(6052):49–50.

Savolainen, O., Bokma, F., Garcia-Gil, R., Komulainen, P., and Repo, T. (2004). Genetic variation in cessation of growth and frost hardiness and consequences for adaptation of Pinus sylvestris to climatic changes. *Forest Ecology and Management*, 197(1–3):79–89.

Savolainen, O., Lascoux, M., and Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, 14(11):807–820.

Savolainen, O., Pyhäjärvi, T., and Knürr, T. (2007). Gene Flow and Local Adaptation in Trees. *Annual Review of Ecology, Evolution, and Systematics*, 38(1):595–619.

Sayre, K. D., Acevedo, E., and Austin, R. B. (1995). Carbon isotope discrimination and grain yield for three bread wheat germplasm groups grown at different levels of water stress. *Field Crops Research*, 41(1):45–54.

Scheepens, J. F., Frei, E. S., and Stöcklin, J. (2010). Genotypic and environmental variation in specific leaf area in a widespread Alpine plant after transplantation to different altitudes. *Oecologia*, 164(1):141–150.

Schlichting, C. D. (2008). Hidden Reaction Norms, Cryptic Genetic Variation, and Evolvability. *Annals of the New York Academy of Sciences*, 1133(1):187–203.

Schroeder, L. M. (1990). Duct resin flow in Scots pine in relation to the attack of the bark beetle Tomicus piniperda (L.) (Col., Scolytidae). *Journal of Applied Entomology*, 109(1-5):105–112.

Schuster, W. S., Alles, D. L., and Mitton, J. B. (1989). Gene Flow in Limber Pine: Evidence from Pollination Phenology and Genetic Differentiation Along an Elevational Transect. *American Journal of Botany*, 76(9):1395–1403.

Setter, T. L. and Waters, I. (2003). Review of prospects for germplasm improvement for waterlogging tolerance in wheat, barley and oats. *Plant and Soil*, 253(1):1–34.

Sinclair, W. T., Morman, J. D., and Ennos, R. A. (1998). Multiple origins for Scots pine (Pinus sylvestris L.) in Scotland: evidence from mitochondrial DNA variation. *Heredity*, 80(2):233–240.

Sinclair, W. T., Morman, J. D., and Ennos, R. A. (1999). The postglacial history of Scots pine (Pinus sylvestris L.) in western Europe: evidence from mitochondrial DNA variation. *Molecular Ecology*, 8(1):83–88.

Sloan, D. B., Alverson, A. J., Chuckalovcak, J. P., Wu, M., McCauley, D. E., Palmer, J. D., and Taylor, D. R. (2012). Rapid Evolution of Enormous, Multichromosomal Genomes in Flowering Plant Mitochondria with Exceptionally High Mutation Rates. *PLoS Biol*, 10(1):e1001241.

Sobel, J. M., Chen, G. F., Watt, L. R., and Schemske, D. W. (2010). The Biology of Speciation. *Evolution*, 64(2):295–315.

Soranzo, N., Alia, R., Provan, J., and Powell, W. (2000). Patterns of variation at a mitochondrial sequence-tagged-site locus provides new insights into the postglacial history of European Pinus sylvestris populations. *Molecular Ecology*, 9(9):1205–1211.

Spitze, K. (1993). Population structure in Daphnia obtusa: quantitative genetic and allozymic variation. *Genetics*, 135(2):367–374.

Städler, T. and Delph, L. F. (2002). Ancient mitochondrial haplotypes and evidence for intragenic recombination in a gynodioecious plant. *Proceedings of the National Academy of Sciences*, 99(18):11730–11735.

Steven, H. and Carlisle, A. (1959). *The Native Pinewoods of Scotland*. Oliver and Boyd.

Stewart, G., Turnbull, M., Schmidt, S., and Erskine, P. (1995). 13c Natural Abundance in Plant Communities Along a Rainfall Gradient: a Biological Integrator of Water Availability. *Functional Plant Biology*, 22(1):51–55.

Strand, M. (1993). Photosynthetic Activity of Scots Pine (Pinus sylvestris L.) Needles During Winter is Affected by Exposure to SO2 and NO2 During Summer. *New Phytologist*, 123(1):133–141.

Sugiura, N. (1978). Further analysis of the data by akaike's information criterion and the finite corrections. *Communications in Statistics - Theory and Methods*, A7:13—-26.

Sun, B., Dilcher, D. L., Beerling, D. J., Zhang, C., Yan, D., and Kowalski, E. (2003). Variation in Ginkgo biloba L. leaf characters across a climatic gradient in China. *Proceedings of the National Academy of Sciences*, 100(12):7141–7146.

Takeda, S. and Matsuoka, M. (2008). Genetic approaches to crop improvement: responding to environmental and population changes. *Nature Reviews Genetics*, 9(6):444–457.

Tang, Z. C. and Kozlowski, T. T. (1982). Some physiological and growth responses of Betula papyrifera seedlings to flooding. *Physiologia Plantarum*, 55(4):415–420.

Tang, Z. C. and Kozlowski, T. T. (1984). Water relations, ethylene production, and morphological adaptation ofFraxinus pennsylvanica seedlings to flooding. *Plant and Soil*, 77(2-3):183–192.

Timmis, J. N., Ayliffe, M. A., Huang, C. Y., and Martin, W. (2004). Endosymbiotic gene transfer: organelle genomes forge eukaryotic chromosomes. *Nature Reviews Genetics*, 5(2):123–135.

Tiwari, S. P., Kumar, P., Yadav, D., and Chauhan, D. K. (2013). Comparative morphological, epidermal, and anatomical studies of *Pinus roxburghii* needles at different altitudes in the North-West Indian Himalayas. *Turkish Journal of Botany*, 37(1):65–73.

Turner, T. L., Bourne, E. C., Von Wettberg, E. J., Hu, T. T., and Nuzhdin, S. V. (2010). Population resequencing reveals local adaptation of Arabidopsis lyrata to serpentine soils. *Nature Genetics*, 42(3):260–263.

Urbaniak, L., Karliński, L., and Popielarz, R. (2003). Variation of morphological needle characters of Scots pine (Pinus sylvestris L.) populations in different habitats. *Acta Societatis Botanicorum Poloniae*, 72(1):37–44.

Valkonen, J. P. T., Nygren, M., Ylönen, A., and Mannonen, L. (1994). Nuclear DNA content of Pinus sylvestris (L.) as determined by laser flow cytometry. *Genetica*, 92(3):203–207.

Vapaavuori, E. M., Rikala, R., and Ryyppö, A. (1992). Effects of root temperature on growth and photosynthesis in conifer seedlings during shoot elongation. *Tree Physiology*, 10(3):217–230.

Varis, S., Pakkanen, A., Galofré, A., and Pulkkinen, P. (2009). The extent of south-north pollen transfer in Finnish Scots pine. *Silva Fennica*, 43(5):717–726.

Visscher, P. M. (1998). On the Sampling Variance of Intraclass Correlations and Genetic Correlations. *Genetics*, 149(3):1605–1614.

Vitousek, P. M., Field, C. B., and Matson, P. A. (1990). Variation in foliar δ13c in Hawaiian Metrosideros polymorpha: a case of internal resistance? *Oecologia*, 84(3):362–370.

Voltas, J., Serrano, L., Hernández, M., and Pemán, J. (2006). Carbon Isotope Discrimination, Gas Exchange and Stem Growth of Four Euramerican Hybrid Poplars under Different Watering Regimes. *New Forests*, 31(3):435–451.

Wachowiak, W. and Prus-Głowacki, W. (2008). Hybridisation processes in sympatric populations of pines Pinus sylvestris L., P. mugo Turra and P. uliginosa Neumann. *Plant Systematics and Evolution*, 271(1-2):29–40.

Wachowiak, W., Salmela, M. J., Ennos, R. A., Iason, G., and 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(5):775–787.

Wagner, D. B., Dong, J., Carlson, M. R., and Yanchuk, A. D. (1991). Paternal leakage of mitochondrial DNA in Pinus. *Theoretical and Applied Genetics*, 82(4):510–514.

Wang, D., Wu, Y.-W., Shih, A. C.-C., Wu, C.-S., Wang, Y.-N., and Chaw, S.-M. (2007). Transfer of Chloroplast Genomic DNA to Mitochondrial Genome Occurred At Least 300 MYA. *Molecular Biology and Evolution*, 24(9):2040–2048.

Warren, C. R., McGrath, J. F., and Adams, M. A. (2001). Water availability and carbon isotope discrimination in conifers. *Oecologia*, 127(4):476–486.

Warren, C. R., Tausz, M., and Adams, M. A. (2005). Does rainfall explain variation in leaf morphology and physiology among populations of red ironbark (Eucalyptus sideroxylon subsp. tricarpa) grown in a common garden? *Tree Physiology*, 25(11):1369–1378.

White, T., Adams, W., and Neale, D. (2007). *Forest Genetics*. CABI Publishing.

Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. Springer New York.

Willis, K. J. and van Andel, T. H. (2004). Trees or no trees? The environments of central and eastern Europe during the Last Glaciation. *Quaternary Science Reviews*, 23(23–24):2369–2387.

Wilson, P. J., Thompson, K., and Hodgson, J. G. (1999). Specific leaf area and leaf dry matter content as alternative predictors of plant strategies. *New Phytologist*, 143(1):155–162.

Wolfe, K. H., Li, W. H., and Sharp, P. M. (1987). Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences*, 84(24):9054–9058.

Wright, I. J., Reich, P. B., Westoby, M., Ackerly, D. D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J. H. C., Diemer, M., Flexas, J., Garnier, E., Groom, P. K., Gulias, J., Hikosaka, K., Lamont, B. B., Lee, T., Lee, W., Lusk, C., Midgley, J. J., Navas, M.-L., Niinemets, l., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V. I., Roumet, C., Thomas, S. C., Tjoelker, M. G.,

Veneklaas, E. J., and Villar, R. (2004). The worldwide leaf economics spectrum. *Nature*, 428(6985):821–827.

Wright, S. (1943). Isolation by Distance. *Genetics*, 28(2):114–138.

Xiao, C.-W., Janssens, I. A., Yuste, J. C., and Ceulemans, R. (2006). Variation of specific leaf area and upscaling to leaf area index in mature Scots pine. *Trees*, 20(3):304–310.

Yamamoto, F. and Kozlowski, T. T. (1987). Effects of Flooding, Tilting of Stems, and Ethrel Application on Growth, Stem Anatomy and Ethylene Production of Pinus densiflora Seedlings. *Journal of Experimental Botany*, 38(2):293–310.

Yamamoto, F., Kozlowski, T. T., and Wolter, K. E. (1987). Effect of flooding on growth, stem anatomy, and ethylene production of Pinushalepensis seedlings. *Canadian Journal of Forest Research*, 17(1):69–79.

Yeaman, S. and Whitlock, M. C. (2011). The Genetic Architecture of Adaptation Under Migration–Selection Balance. *Evolution*, 65(7):1897–1911.

Yordanova, R. Y., Uzunova, A. N., and Popova, L. P. (2005). Effects of short-term soil flooding on stomata behaviour and leaf gas exchange in barley plants. *Biologia Plantarum*, 49(2):317–319.

Ziello, C., Estrella, N., Kostova, M., Koch, E., and Menzel, A. (2009). Influence of altitude on phenology of selected plant species in the Alpine region (1971–2000). *Climate Research*, 39(3):227–234.