UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL & ENVIRONMENTAL SCIENCES

School of Biological Sciences

Evolutionary and Genetic Basis of Morphological Variation in *Populus nigra* (European black poplar)

by

Jennifer DeWoody, BSc, MSc

Thesis for the degree of Doctor of Philosophy

February 2011

Abstract

Changes in precipitation over the next century may impact the distribution of species, particularly in southern Europe, where droughts are predicted to increase in frequency. In forest trees, intraspecific variation in leaf size, branching architecture, and growth rate among populations are considered adaptive and likely related to climatic differences between sites. A previous common garden study of *Populus nigra* L. showed morphological variation to be highly heritable and significantly differentiated among populations, indicating phenotypic differences may be adaptive.

This project studied the evolutionary processes that have contributed to the morphological differentiation observed in *P. nigra*. Examining scales ranging from landscape-level patterns of variation to cellular differences within developing leaves identified historic and developmental processes contributing to the phenotypic differences in this species. Both isolation by distance, where migrants do not move equally across the landscape, and isolation by adaptation, where genetically divergence varies with morphological differences, have influenced differentiation among populations of *P. nigra* in western Europe. These patterns broadly correspond to the recolonization routes following the most recent glacial event, indicating that historic vicariance and not just adaptive divergence influence phenotypic variation. Identification of quantitative trait loci (QTL) for insect herbivory in a common garden study of hybrid poplar indicated that leaf morphology might also be influenced by insect preference.

Among individuals, differences in leaf size corresponded to variation in cell number and not cell size, indicating natural selection may have influenced the regulation of cell division. Further, variation in gene regulation across the developing leaf identified differences across the leaf lamina. Finally, simulations of demographic, genetic, and adaptive processes among populations revealed that a lack of correspondence between the optimal phenotype of colonists and the optimal phenotype and newly colonized populations significantly affects levels of phenotypic differentiation among populations. In addition, changes in phenotypic optima, as may occur due to climate change, impacted the level of genetic variance, and thus the future adaptive potential of populations. Together, these results provide insight into the evolution of phenotype in *P. nigra*, and contribute information for management efforts in the context of a changing climate.

Contents

List of Tables

List of Figures

Declaration of Authorship

I, **Jennifer DeWoody,** declare that the thesis entitled:

Evolutionary and Genetic Basis of Morphological Variation in *Populus nigra* (European black poplar)

and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research, with the exceptions below.

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;
    - specifically, the morphological data presented in Chapter 2 was collected as part of the PhD work done by Harreit Trewin, from which I conducted a novel statistical analysis;
- 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,
    - in particular, M. Goolsby aided with data collection in the greenhouse experiments described in Chapters 3 and 4; and
    - M. Viger and S. Travers shared equally in data collection for Chapter 5;
- none of this work has been published before submission.

Signed: Jennifer DeWoody

Date: 27 May 2011

Acknowledgements

I am grateful to Prof. Gail Taylor for the opportunity to work in the European community studying *P. nigra* and for guidance developing and interpreting these studies.

I thank Dr. Patrick Doncaster, Dr. Donna Clarke, and Rebecca Rowe for discussion and advice on various statistical analyses.

I am grateful to Dr. J.E. Cottrell, Dr. M.J.M. Smulders, Dr. S. Bordács, and Dr. G. Szulcsán for insightful discussion on the ecology and management of *P. nigra* across Europe.

Thanks as well to Dr. S. Bordács, Dr. G. Szulcsán and Prof. Dalibor Ballian for sharing cuttings of *P. nigra* from Hungary and Bosnia & Herzegovina.

I am indebted to Prof. Antoine Kremer, Dr. Pauline Garnier-Géré, and Dr. Frédéric Raspail for their patience, guidance, and collaboration on the METAPOP simulation study.

I am grateful to M. Nelson, M. Viger, S. Milner, A. Payne, H. Smith, Y. Lin, Dr. M. Tallis, and Dr. M. Totten for assistance in the laboratory and greenhouse (many hands make light work). M. Cotton provided excellent greenhouse support. Many thanks to M. Viger for the instruction and S. Travers and M. Viger for their work in the field assessing leaf damage as reported in Chapter 5. It has been a joy to work in this group. I thank M. Goolsby for assistance with ArcGIS and with data collection in the greenhouse. I am also thankful to M. Viger for ongoing discussion of the physiology, genetics, and genomics of *P. nigra*.

I am ever grateful to my sweetie and my family for their continued encouragement. My adventures and achievements are only possible because of their support.

This work is funded through the EVOLTREE Network of Excellence and an Overseas Research Student Awards Scheme (ORSAS) grant.

Definitions and Abbreviations

Abbreviation	Definition
A	Mean alleles per locus
A_e	Effective alleles per locus, standardized for sample size
AFLP	Amplified fragment length polymorphism
bp	Basepairs (nucleotide)
CA	Cell area, mm^2
cM	Centi-Morgans, a statistical measure of distance between loci
CN	Cell number, estimated as (leaf area)/(cell area)
cpDNA	Chloroplast DNA
DCA	Discriminant classification analysis
Diam	Stem diameter, mm
EST	Expressed sequence tags
F_{IS} or F	Fixation of alleles within individuals relative to the subpopulation
F_{ST}	Fixation of alleles within subpopulations relative to the total population
G0	Gap phase where the cell has exited the cell cycle
G1	Postmitotic gap phase of the cell cycle
G2	Premitotic gap phase of the cell cycle
GLM	General linear model
G_{ST}	Gene differentiation among populations
h^2	Heritability of a phenotypic trait
H_e	Expected frequency of heterozygotes at Hardy-Weinberg equilibrium
H_o	Observed frequency of heterozygotes in sample
H_S	Genetic diversity at molecular loci

Abbreviation	**Definition**
HWE	Hardy-Weinberg equilibrium
IBD	Isolation by distance
IBA	Isolation by adaptation
K	Number of genetic clusters in a sample
kb	Kilo-basepairs (nucleotide)
LA	Leaf area, mm^2
LD	Linkage disequilibrium, or the non-random association of alleles
LG	Linkage group
LL	Leaf length, mm
LOD	Log of odds likelihood ratio
LW	Leaf width, mm
m	Rate of migration between subpopulations within the total population
M	Mitotic phase of the cell cycle
MAPpt	Mean annual total precipitation, mm
MATemp	Mean annual temperature, °C
MaxDay	Maximum day length on the summer solstice, hours
MaxPpt	Mean precipitation of the wettest month, mm
MaxTemp	Maximum temperature of the warmest month, °C
Mb	Mega-basepairs (nucleotide)
MinPpt	Mean precipitation of the driest month, mm
MinTemp	Minimum temperature of the coolest month, °C
N_e	The effective population size
N_em	The effective number of migrants between subpopulations per generation
NStomata	Number of stomata

Abbreviation	Definition
p	Frequency of one allele in population genetic theory
P	Proportion of polymorphic loci
PCoA	Principle coordinates analysis
q	Frequency of an alternate allele in population genetic theory
Q_{ST}	Differentiation of phenotypes between subpopulations relative to the total population
QTL	Quantitative trait loci
RAPD	Random amplified polymorphic DNA
RG	Relative growth, calculated as the log of the ratio of the final measure to the beginning measure
RFLP	Restricted fragment length polymorphism
RER	Relative expansion rate, calculated as the diffirence of initial and final areas divided by the number of days
S	Synthesis phase of the cell cycle
SAM	Shoot apical meristem
SD	Stomatal density, estimated as the number of stomata per unit area
SDTemp	Standard deviation of the mean monthly temperature
SI	Stomatal index, estimated as the ratio of stomata to total cell count
SLA	Specific leaf area, the ratio of leaf area to dry mass, mm^2/mg
SNP	Single nucleotide polymorphism
SRS	Stress response syndrome, the convergent evolution of plant traits in stressful environments
V_{A}	Additive genetic variance
V_{B}	The between-population component of variance for a phenotypic trait
V_{w}	The within-population component of variance for a phenotypic trait

Abbreviation	Definition
VarPpt	Variation in mean monthly precipitation
V_{Zopt}	Variance phenotypic optima among populations
w^2	Intensity of stabilizing selection within a population
θ_B	A measure of linkage disequilibrium between allelic states between populations
θ_W	A measure of linkage disequilibrium between allelic states within populations
Φ_{PT}	Genetic differentiation of populations relative to the total collection, as assessed via analysis of molecular variance
σ^2_b	The between-population component of variance, either for traits or allele frequencies
σ^2_t	The total variance observed in either traits or allele frequencies
σ^2_w	The within-population component of variance, either for traits or allele frequencies

> Realistic discussion of the structure of natural populations requires simultaneous consideration of all processes by which gene frequencies may change.
>
> - Sewall Wright, *The Genetical Structure of Populations*, 1951

Chapter 1. Introduction

1.0 Overview

Understanding the genetic variation controlling morphological variation and the potential for populations to adapt to climate change is critical for the protection of natural and agricultural germplasms over the next century. The most effective management efforts will include data on the predicted climate changes for the species of interest, the historic structure of populations, the extent that natural selection has influenced morphological differences, and the level of adaptive variation available for future changes. In this study, the relationship between putatively adaptive traits and both neutral and adaptive genetic variation are assessed for a forest tree, *Populus nigra*. The studies investigated both large (landscape) and small (within leaves) scale variation in order to identify evolutionary processes that might have influenced phenotypic differentiation. Traits related to leaf morphology, in particular leaf size, were studied due to the adaptive nature of the variation. The results provide insights into developmental and adaptive mechanisms that may be applied to conservation or management goals. The adaptive nature of morphological variation was assessed first by quantifying both genetic and phenotypic differentiation and estimating to what extent the variation results from historic demographic processes or from adaptation to local conditions. Then, focusing on leaf size, the developmental mechanism underlying variation was examined in *P. nigra.* In order to examine biotic factors influencing leaf size, quantitative trait loci for insect damage were compared to 'hot-spots' for leaf development in hybrid poplar, indicating that insect preference may depend in part on the size of leaves. Finally, in order to examine the potential for populations to adapt to changing conditions, simulation studies described the effect of changing phenotypic optima on population fitness and survival. Together, these studies provide evidence that historic events may affect the adaptive potential of extant populations and should be considered when mitigating changing climatic conditions.

1.1 Earth's changing climate

The earth's surface temperature is rising. Over the past century, the global mean temperature has risen by approximately 0.74°C, with rate of warming increasing over the past fifty years (Trenberth et al. 2007). While cyclical climate changes have been documented, the rate and magnitude of the current warming is unprecedented in the last millennium (Ohlemuller et al. 2006; Raible et al. 2006), and a consensus has emerged that anthropogenic factors have contributed to the current trends (Gessler et al. 2007). Simultaneous changes in the amount and seasonal pattern of precipitation have also been recorded that are likely related to changes in temperature (Trenberth et al. 2007).

A variety of models using historic temperature and precipitation data have been developed to predict future climatic changes. While predictions can be highly variable between models (Blenkinsop and Fowler 2007), some consistent trends have emerged. The mean global temperature is expected to rise 1.5 to 5.8°C above the 1990 measures by the end of the twenty-first century (Gessler et al. 2007). In Europe, the mean temperature is expected to increase by 6°C over the next century (Raible et al. 2006; Gessler et al. 2007). A large cline of increases is predicted across latitudes, with Mediterranean areas expected to warm by 2°C, but northern areas susceptible to increases of 10°C or more (Gessler et al. 2007). In addition, rates and distribution of precipitation are also expected to change. Models have predicted that summer precipitation would decrease over much of Europe, with levels increasing only in the most northern latitudes (Blenkinsop and Fowler 2007). By contrast, winter precipitation may increase in most areas, and the largest changes occurring in northwest Europe and Scandinavia (Blenkinsop and Fowler 2007). Drought events are predicted to increase in frequency (Blenkinsop and Fowler 2007; Penuelas et al. 2007), while long droughts may decrease in frequency in northern parts of the continent but increase in the southern, Mediterranean areas (Blenkinsop and Fowler 2007). These changes in temperature and precipitation will change the distribution of climate types across Europe. A study modelling the distribution and abundance of climates analogous to those observed in 1945 (i.e. pre-warming) determined that most habitat types would decrease in size, increase in fragmentation, and generally shift to the northeast (Ohlemuller et al. 2006). In some areas of Europe, particularly Scandinavia and regions of the Mediterranean coast, novel climatic conditions without pre-warming analogues may be seen (Ohlemuller et al. 2006).

Every organism has a range of environmental conditions, such as temperature or moisture, in which it can survive. Should a species' environment exceed these conditions in some parameter, such as minimum rainfall, populations will be forced either to move to areas having suitable conditions or to adapt to the new climate in order to survive. Range shifts are expected to account for the majority of species local extinctions as environmental conditions change over the next century (Ohlemuller et al. 2006). A meta-analysis of species ranges indicates that some species are already shifting poleward by a mean distance of 6.1 km per decade as temperatures rise (Parmesan and Yohe 2003). The unprecedented rate of the current climate change may prevent natural systems (i.e. species) from adapting as occurred in previous, slower warming events. Understanding the plasticity of response to climatic conditions (i.e. phenotypic variation) together with the existing adaptive potential of wild populations (i.e. the levels of variation in functional genes) will be critical to most efficiently manage ecologically and economically important species.

1.2 Inferring non-adaptive evolutionary processes from selectively neutral loci

Conservation efforts primarily require information about traits that affect fitness, the ability of an organism to survive and reproduce. However, selectively neutral processes may impact the effects of natural selection on phenotypic traits, and thus must be included in studies concerning population adaptation to climate change. These processes include the breeding system of the species (inbreeding) and the movement of alleles among populations (gene flow and genetic drift), and serve as a null hypothesis of the phenotypic differentiation expected in the absence of natural selection. Decades of theoretical and experimental study have produced models predicting how evolutionary and demographic processes affect the genetic structure of natural populations. Experimentally, the genetic structure of populations can be quantified by assessing allele frequency variation at genetic markers. The allele frequency variation can be compared to the variation in a morphological trait as a first test of whether a trait may be under natural selection. Thus, in order to determine the extent to which a trait is adaptive, the neutral genetic structure of populations must be well described.

The effects of evolutionary processes on the genetic structure of a population are determined by comparing observed patterns to those predicted for an idealized population undergoing no evolutionary change. The theoretical characteristics and requirements of this ideal population were put forth simultaneously but independently by Hardy (1908) and Weinberg (1908), and have formed the basis for understanding how perturbations such as genetic drift and non-random mating will change the distribution of alleles within and among individuals. The ideal population of the Hardy-Wienberg model is defined by demographic characteristics, including being infinitely large and closed to immigrants, with the sexes occurring in equal ratios, reproduction occurring via completely random mating, having discrete generations, and having an absence of natural selection or mutation (Hardy 1908; Lynch and Walsh 1998). In this ideal population, the frequency of the alleles in the population will remain unchanged across generations.

Violations of any of the conditions of the Hardy-Weinberg model will perturb the relationship between allele frequencies and genotypic frequencies. A wealth of theoretical and experimental work in a variety of systems has resulted in detailed models describing the changes in genotypic frequencies expected from violations of the Hardy-Weinberg model. In his seminal work, "The genetical structure of populations," Sewall Wright (1951) described the changes in allele frequencies and heterozygosity expected following violations of Hardy-Weinberg equilibrium (HWE). In particular, he studied how deviations from random mating within and migration between finite populations change the distribution of homologous alleles among individuals within subpopulations and among subpopulations within the total population.

In order to describe the effect of migration of alleles between populations on the allele frequencies within each, Wright (1951) first described the simplest form of population structure, the island model. In this ideal model, populations are large, with a size of N individuals, and uniformly distributed across the neighbourhood, yet sufficiently isolated from each other so that migration is restricted to some finite rate, m. Over generations of restricted migration, the fixation of alleles within a subpopulation relative to the total population will be related to the population size and rate of migration such that:

$$F_{ST} = \frac{1}{4Nm + 1} \qquad (1.1)$$

(Wright 1951). This fixation index ranges from zero, indicating that allele frequencies are equivalent in each subpopulation, to unity, indicating that subpopulations are fixed for alternate alleles. Greater levels of migration will prevent the genetic differentiation of subpopulations, and will maintain equivalent allele frequencies across the landscape.

In practice, F_{ST} is estimated as the proportion of variance in allele frequencies occurring among populations, such that:

$$F_{ST} = \frac{\sigma_b^2}{\sigma_b^2 + 2\sigma_w^2}, \qquad (1.2)$$

where σ^2_{w} is the within-population component and σ^2_{b} is the between-population component of variance (Peakall et al. 1995). Using a frequency-based estimate of F_{ST}, an estimate of the effective number of migrants per population, $N_{\mathrm{e}}m$, may be calculated using equation (1.1) above (Wright 1951; Slatkin 1993). Given the difficulty if not impossibility of observing the migration of individuals or germplasm in forest trees, this metric provides an efficient description of the genetic connectivity and a relative measure of the rate of migration between subpopulations.

The most popular technique to investigate the genetic structure of populations involves assaying variation at putatively neutral molecular markers in the nuclear genome. Numerous marker systems have been developed over the past 40 years, both assessing protein variants (isozymes) and variation at the DNA level, including restriction fragment length polymorphisms (RFLPs) and amplified fragment length polymorphisms (AFLPs, Newton et al. 1999). Currently, the most popular marker for studies of intraspecific variation and structure are microsatellites, or simple sequence repeats (SSRs). Microsatellites refer to segments of DNA composed of tandem repeats of short sequences, typically two to six base pairs. The secondary structure of these repetitive elements is thought to reduce the accuracy of DNA replication, with repeats being added or removed with a greater frequency than nucleotide substitutions in the genome as a whole (microsatellite mutation rates range from approx. 10^{-2} in bacteria to 10^{-3} in humans and 10^{-4} to 10^{-5} in yeast; Burg et al. 2007). This hyper-variation provides greater power to resolve structure amongst closely related individuals or populations, though homoplasy may complicate inferences of distantly related or reproductively isolated demes (O'Reilly et al. 2004). Databases of microsatellite loci are available for *Populus* species (van der Schoot et al. 2000; Tuskan et al. 2004).

As nuclear microsatellites are biparentally inherited, their diversity and diversification are influenced by a number of demographic and genetic factors. Differentiation at nuclear markers is influenced by gene flow through both the maternal (seed) and paternal (pollen) gametes. In addition, recombination across the nuclear genome results in the independent assortment of alleles for loci outside the range of linkage disequilibrium (LD).

Models describing the change in gene frequencies under neutral scenarios provide insight into processes influencing the genetic structure of populations. The differentiation observed with neutral markers may be used as a null hypothesis against which to test for evidence of natural selection acting on phenotypic traits (Kohn et al. 2006). If the pattern of morphological differentiation does not correspond to the neutral genetic differentiation, other evolutionary processes than those of gene flow and genetic drift are required to explain the phenotypic variation. In this study, the neutral genetic structure among populations will be used to assess the evolutionary processes contributing to variation in *P. nigra*.

1.3 Genetic basis of quantitative traits

While neutral genetic markers provide insight into demographic evolutionary processes, the effects of natural selection are expected to influence the patterns of phenotypic diversity in populations. Adaptive evolutionary processes may result in a genetic structure of loci controlling phenotypic traits that varies from the neutral genetic structure of populations. Over the past century the field of quantitative genetics has dissected the genetic basis of continuously distributed phenotypic traits (Roff 2007). The complexity of this task lies in the polygenic nature of the traits of interest, whereby the interaction between the multiple loci invokes complex patterns of inheritance. With the availability of molecular genetic methods and genomic information, two of the approaches developed to identify the loci associated with variation in phenotype are identification of quantitative trait loci (QTL) and association or linkage disequilibrium (LD) mapping (Flint and Mott 2001; Morgante and Salamini 2003; Gibbs and Singleton 2006; Kohn et al. 2006; Ochieng et al. 2007).

The classical method to estimate the number of genomic regions involved in a continuous trait is through quantitative genetic studies (Roff 2007). In a typical three-generation (or ‘grandparent’) design, two individuals highly inbred or diverged for the trait of interest are mated to produce a first filial generation (F1) heterozygous for all genes involved in the trait. The F1 generation can then be mated against itself, so that due to recombination and random assortment of alleles, the second filial generation (F2) will display a range of values for the phenotype of interest (Erickson et al. 2004). Alternately, the F1 generation may be back-crossed to either parent, resulting in a back-cross pedigree that may reduce the amount of variation, but increase the resolution, in the subsequent generations. The segregation of phenotypes may provide sufficient information to estimate the number of genetic regions or loci involved in the development of the trait, but no information as to the location of the genes involved in the genome, nor the genetic mechanism involved in creating the phenotype are provided (Roff 2007).

With the advent of molecular genetic techniques, it became possible to map the allelic differences between the parental individuals of quantitative studies (Roff 2007). Initially these linkage maps were built using restriction fragment length polymorphisms (RFLP, Bradshaw et al. 1994; Jermstad et al. 1998), but currently linkage maps typically use amplified fragment length polymorphisms (AFLP) and microsatellite markers (Cervera et al. 2001; Gaudet et al. 2008). The use of single nucleotide polymorphisms (SNP) is becoming more prevalent as the number of species with genomic information available continues to increase (Gibbs and Singleton 2006; Freimer et al. 2007; Gaudet et al. 2008). Once phenotypic and genotypic data are collected on the progeny array, various statistical methods can then be employed to test for the non-random association of genetic states with observed phenotype (Kearsey and Farquhar 1998; Erickson et al. 2004; Roff 2007), thus identifying loci associated with the trait of interest (i.e., QTL).

Association, or linkage disequilibrium (LD), studies, an alternative method to identify genomic regions associated with phenotypic traits, takes advantage of historic levels of recombination by assaying natural populations (Morgante and Salamini 2003; Ehrenreich and Purugganan 2006; Neale 2007). The larger number of recombination events results in a smaller section of the genome being associated with the trait of interest, though the size of the region is a function of the rate of decay of linkage disequilibrium, or the non-random association of alleles, which varies across species and across regions of the genome (Morgante and Salamini 2003). In association studies, as in QTL studies, genetic variance in phenotype is assessed for the samples in a common environment. Unlike QTL studies, however, association studies do not require a genetic linkage map, and instead assess genotypic variation at a set of loci selected *a priori*, termed candidate genes. The selection of candidate genes can be a complex process drawing on a range of previous studies. QTL studies may identify several areas of the genome involved in the trait of interest, but each region may include hundreds of genes or gene models (Ehrenreich and Purugganan 2006), which is cost-prohibitive for most labs. An alternate way to select candidate genes involves using information on gene function and physiological pathways from model organisms to choose homologues for analysis (Ingvarsson et al. 2006; Ehrenreich et al. 2007). The number of genes assessed is expected to be a fraction of that identified from QTL studies, but are unlikely to be a complete list of genes involved in the trait of interest. As a reminder of the genetic complexity underlying quantitative traits, significant candidate genes and QTL for a single trait may not map to the same genetic region (Ehrenreich et al. 2007).

Whether by assessing the variation at markers linked to adaptively significant loci, thus potentially reflecting the effects of natural selection due to genetic hitch-hiking, or assaying sequence variation in candidate genes, these methods provide insight into the amount and distribution of genetic variation at loci under natural selection. The genetic differentiation between individuals and populations can be quantified from measures of variance in the same manner as for neutral loci ($F_{\mathrm{ST(adaptive)}}$). Just as neutral theory describes the effects of inbreeding and gene flow, theoretical models have been developed that describe the effect of selection on genetic variation, from the pattern of nucleotide substitution within genes to the distribution of haplotypes among populations (Storz 2005; Vasemägi and Primmer 2005). While neutral processes provide a background level of genetic differentiation against which to test for evidence of natural selection, these models provide tests to determine the type and intensity of selection acting on these loci. The differentiation at quantitative loci has been compared to the differentiation in the quantitative trait of interest in order to assess the relative effects of adaptive processes on each metric. This comparison requires the phenotypic differentiation to be quantified among populations.

1.4 Phenotypic variance and adaptation

Natural selection across various and varying environments has resulted in interspecific variation among taxa, and intraspecific variation among individuals. Variation in phenotype (Johannsen 1911), such as morphological differences among individuals of a species is expected to correspond to differences in fitness, or the ability of an individual to survive and reproduce. Quantification of the phenotypic differences between populations is necessary before the genetic basis or adaptive significance of such variation may be identified.

In order to quantify the differences in a trait observed among populations, the total phenotypic variance observed in a sample of individuals from multiple populations may be partitioned into the within-population component (σ'^2_w) and the between-population component (σ'^2_b), similar to the genetic variance described in Equation (1.2). Spitze (1993) showed that when the variance of a trait is described in this manner, the ratio of the variance between populations to the total variance provides a meaningful statistic analogous to the allele frequency variation described above (F_{ST}), which he denoted:

$$Q_{ST} = \frac{\sigma_b'^2}{\sigma_b'^2 + 2\sigma_w'^2}. \tag{1.3}$$

In order to estimate the phenotypic differentiation in this manner, samples from a number of populations must be assessed in a controlled environment, typically a common garden for forest species, following a randomized experimental design, and careful measurements taken of the trait of interest.

The simplest interpretation of Q_{ST} provides quantification of differences between populations relative to the total phenotypic variation observed within the collection. Yet quantifying phenotypic differences using Q_{ST} may also provide insight to the strength and type of natural selection affecting the trait, should a baseline of population differentiation be available for comparison. Typically, that baseline measure is inferred from the distribution of neutral genetic variation within and among populations, frequently summarized as F_{ST} (Wright 1951; Peakall et al. 1995), providing a mechanism to compare morphological and genetic differentiation.

While Q_{ST} provides a single summary of differentiation among populations, it does not qualitatively describe which populations are similar or divergent, nor does it describe any geographic patterns. Further, as Q_{ST} is calculated on a per-trait basis, it does not resolve the overall morphological differentiation among populations (Rogers and Harpending 1983). Multivariate statistical approaches are designed to summarize variation measured as multiple traits into a small number of synthetic variables, allowing interpretation of large data sets (Jombart et al. 2009). One type of multivariate analysis used for data reduction and description is discriminant analysis (DA) (Williams 1983). DA estimates summary variables, or function scores, for each sample that are orthogonal and thus independent (Williams 1983). Assessing the correlation between the raw data and each function score provides insight as to the variables (traits) that best resolve differences between samples. Finally, plotting the distribution of samples based on their function scores provides visual depictions of any clustering or pattern related to other factors, such as population of origin. In this manner the pattern of differentiation among samples and populations can be assessed on a pairwise basis. Multivariate analyses will be used to assess the overall morphological variation and pattern of phenotypic distribution across the landscape of populations of *P. nigra*.

1.5 Inferring evolutionary processes from neutral and adaptive genetic structure

Phenotypic, adaptive genetic, and neutral genetic variation are expected to differ among individuals and populations due to the distinct evolutionary forces acting on each, as discussed above. Accurate estimates of morphological and genetic differentiation are required to determine the relative importance of each evolutionary process. Although different methods are available, most studies apply a single metric to quantify the differentiation at each level: in phenotypic traits (Q_{ST}), at adaptively significant loci ($F_{ST(adaptive)}$), and at selectively neutral loci ($F_{ST(neutral)}$). The theoretical models describing each parameter are well established (Wright 1951; Spitze 1993) and empirical tests comparing differentiation at neutral markers to that observed in phenotypes are available (Podolsky and Holtsford 1995; Bonnin et al. 1996; Kuittinen et al. 1997; Waldmann and Andersson 1998; Lynch et al. 1999; Gonzalez-Martinez et al. 2002), but the theoretical work relating differentiation at phenotypic traits to that at underlying genetic loci has only been developed in the last decade (Latta 1998; LeCorre and Kremer 2003). Simulation studies provide a mechanism to investigate the effect of and interaction of evolutionary processes on the distribution of genetic variation.

Latta (1998) described the model decoupling variation at a phenotypic trait from that at its underlying quantitative trait loci (QTL). If a quantitative trait is selectively neutral, then its within-population (σ_w^2), between-population (σ_b^2) and total (σ_t^2) genetic variances will be:

$$\begin{aligned} \sigma_w^2 &= (1 - F_{st})\sigma_o^2 , \\ \sigma_b^2 &= 2F_{st}\sigma_o^2 \quad , \\ \sigma_t^2 &= (1 + F_{st})\sigma_o^2 , \end{aligned} \tag{1.4}$$

where σ_0^2 is the genetic variance for the trait in a single panmictic population (Latta 1998). Further, for a single trait influenced by multiple loci each having two alleles, the genetic variance for the trait, z, is:

$$\sigma_z^2 = \sum_i \sigma_i^2 + \sum_i \sum_{j \neq i} \mathrm{cov}_{(i,j)} . \tag{1.5}$$

From these equations, Latta (1998) showed that the differentiation observed at a quantitative trait under stabilizing selection toward a locally optimal value (z^*) is related to the covariance of the underlying QTL allele frequencies, and not the allele frequency differentiation at these loci, such that:

$$\sum_i \sum_{j \neq i} \mathrm{cov}_{(i,j)} = \sigma_{z*}^2 - 2F_{st}\sigma_0^2 , \tag{1.6}$$

where $\mathrm{cov}_{(i,j)}$ is the covariance of allelic effects at the i^{th} and j^{th} loci. This equation can be used to predict the covariance between alleles at QTL under different scenarios of natural selection. For the case of poplar trees, which tend to have high levels of migration due to long-distance pollen and seed distribution (thus, a low value for $2F_{ST}\sigma_0^2$), but are under diversifying selection for drought tolerance (a high σ_{z*}^2), allele frequencies at QTL across environments are expected to display positive covariance (Latta 1998). This model shows how phenotypes may show greater variation and differentiation than their underlying QTL. The genetic basis for the range of traits lies not in allele frequency differences between populations or environments, but in the covariance of alleles across many loci. Further, this model indicates that for traits controlled by many loci of equal (or small) effect, loci will not be detected using tests for significant differentiation, as are typically used in pedigree-based QTL studies, since the alleles are not expected to show significant differentiation. Rather, tests for linkage disequilibrium will be required to detect covariance among allelic states at many loci (Latta 1998; LeCorre and Kremer 2003), although these methods have yet to be developed or applied.

LeCorre and Kremer (2003) further explored the relationship between phenotypic and genotypic differentiation. In their model expanding on Latta's (1998) work, LeCorre and Kremer (2003) relate Q_{ST} to F_{ST} directly as:

$$\begin{aligned} Q_{ST} &= V_B(1+F)/(V_B(1+F)+2V_W) \\ &= (1+\theta_B)F_{ST}/[(\theta_B-\theta_W)F_{ST}+1+\theta_W] \end{aligned} \tag{1.7}$$

where V_{W} and V_{B} are the within and between population variance for the trait, respectively, F is the fixation due to inbreeding assuming it results solely from the mating system and is equivalent across loci, and θ_{W} and θ_{B} is a measure of the linkage disequilibrium among allelic states within and between populations, respectively, due to selection. This relationship shows that exactly two cases exist where $Q_{\mathrm{ST}} = F_{\mathrm{ST}}$. The first is the case when all QTL are in linkage equilibrium ($\theta_{\mathrm{W}} = \theta_{\mathrm{B}} = 0$), and the second, when linkage disequilibrium is present but equal at the within and between population components ($\theta_{\mathrm{W}} = \theta_{\mathrm{B}}$). This model also predicts the relationship between Q_{ST} and F_{ST} under different forms of selection. When a trait is under unifying selection, high levels of selective intensity, or is subject to high levels of genetic drift, Q_{ST} is expected to be smaller than F_{ST}. Conversely, when a trait is under diversifying selection, as for species occupying a range of environmental conditions, Q_{ST} will be greater than F_{ST} (LeCorre and Kremer 2003). Again, this results from the decoupling of the variance of the trait and that of its underlying QTL due to the positive covariance of allelic states among loci (Latta 1998).

Simulations based on these models resolve the relationship between Q_{ST}, $F_{ST(adaptive)}$ and $F_{ST(neutral)}$ in natural populations under a range of selective forces, mating systems, or levels of migration (LeCorre and Kremer 2003). As shown in Figure 1.1, for a highly outcrossing species (100% allogamy) experiencing high gene flow (Nm = 10) and diversifying selection (VZopt=5 or 10), the genetic diversity at QTL and the heritability of the trait are expected to decrease with increasing stabilizing selection within populations (w^2), while genetic diversity at neutral markers remain unchanged.

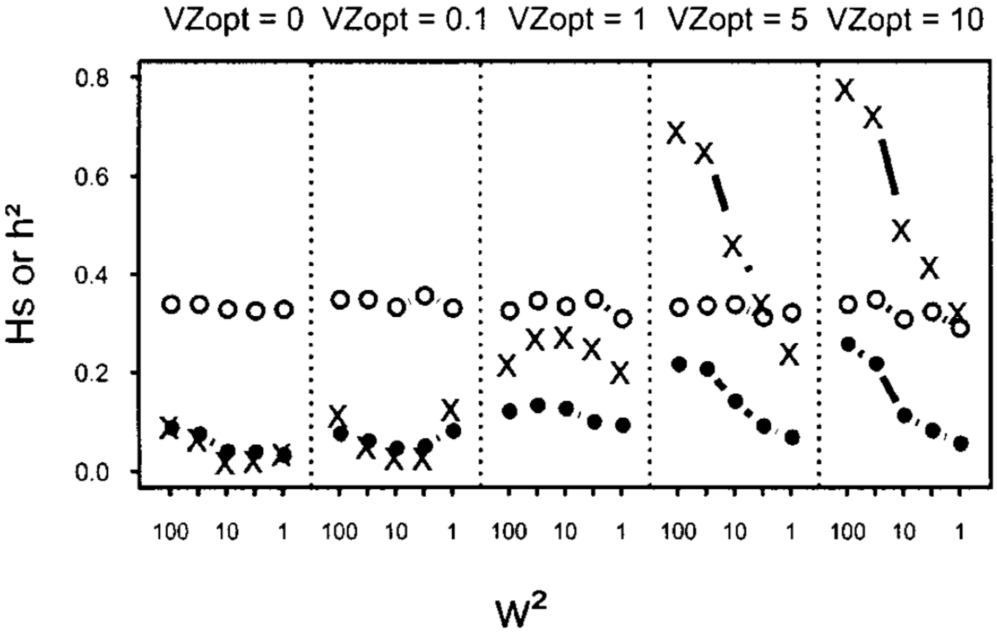


Figure 1.1 Effect of intensity (w^2) and type (V_{Zopt}) of diversifying selection on the heritability (h^2) of a trait (X), the genetic diversity (Hs) of the QTL underlying the trait (●), and the diversity at neutral markers unlinked to the trait (○). Results are for a fully-outcrossing species experiencing high levels of gene flow (Nm = 10). From LeCorre and Kremer (2003).

In addition, Q_{ST} is expected to be greater than either measure of F_{ST} (Figure 1.2). As selective intensity increases, Q_{ST} and F_{ST} for QTL are expected to increase at a faster rate than F_{ST} for neutral markers (Figure 1.2). These simulations may be used as null and alternate hypotheses for empirical tests of population structure, and identify which combination of selective and demographic parameters would result in an observed pattern of differentiation for a trait, its QTL, and neutral loci.

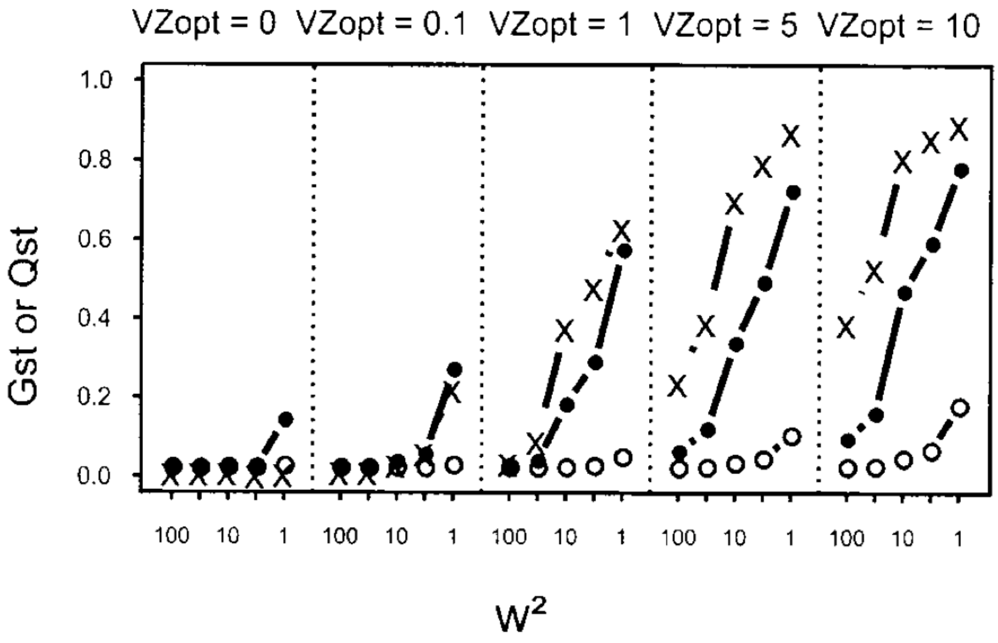


Figure 1.2 Effect of intensity (w^2) and type (VZopt) of diversifying selection on the differentiation of a trait (X), the genetic differentiation of the QTL underlying the trait (●), and the differentiation at neutral markers unlinked to the trait (○). Results are for a fully-outcrossing species experiencing high levels of gene flow (Nm = 10). From LeCorre and Kremer (2003).

A variety of studies relating (neutral) genetic variation to phenotypic variation are available, although few exist which compare these levels of variation to functional genetic variation (QTL), in part because the molecular genetic information identifying the sequences and candidate genes associated with QTL have only been developed recently and in a few model systems. In the first such study in *Populus*, the relationship between variation in phenology, differentiation in candidate genes and microsatellites linked to QTL associated with bud break, and differentiation among loci unlinked to these QTL was assessed in *Populus tremula* (Hall et al. 2007). Samples from 12 locations across a latitudinal gradient displayed significant levels of phenotypic variation for traits related to phenology (Q_{ST} ranging from 0.175 to over 0.6). Yet the genetic differentiation in candidate genes and microsatellites linked to QTL associated with these traits was similar to that observed at neutral markers ($F_{STsnp} = 0.016$; $F_{STqtl} = 0.022$; $F_{STneutral} = 0.012$), all of which were lower than any measure of phenotypic differentiation. These observations are consistent with the theoretical predictions by Latta (1998) and LeCorre and Kremer (2003) that differentiation at traits does not require genetic differentiation at the underlying QTL. Further, these results, on the surface contrary to theory, have provoked debate in the literature as to the mechanisms creating the low level of differentiation at loci under natural selection.

Meta-analyses reviewing studies relating neutral variation to quantitative trait variation disagree on the significance of the correlation between Q_{ST} and $F_{ST(neutral)}$. Merila and Crnokrak (2001) reviewed 18 studies reporting both Q_{ST} for phenotypic traits and F_{ST} for neutral genetic markers. The meta-analysis used the mean standardized difference between Q_{ST} and F_{ST} to describe the magnitude of difference between these parameters (Merila and Crnokrak 2001). Results indicate Q_{ST} tends to be greater than F_{ST}, but the significance of the difference depends on the class of traits measured (morphological versus life history traits) and the type of markers employed (isozyme versus DNA-based, Merila and Crnokrak 2001). The authors concluded, however, that available evidence supports a strong positive correlation between Q_{ST} and F_{ST} ($P = 0.05$ for microsatellite studies, $P = 0.0005$ for isozyme studies), whereby species that show strong phenotypic differentiation also display high levels of allele frequency variation. Similarly, Lynch et al. (1999) found that Q_{ST} tends to be greater than F_{ST}, and that the two values are positively correlated. Such a relationship indicates that levels of differentiation at neutral genetic markers may sufficiently reflect differentiation in traits.

A second review of the literature drew different conclusions, however. McKay and Latta (2002) reviewed studies of 29 species and, consistent with Merila and Crnokrak (2001), Q_{ST} tends to be larger than F_{ST}, indicating that neutral genetic variation would be insufficient to predict differentiation in traits. Although a positive correlation was observed between Q_{ST} and F_{ST}, the trend was marginally significant ($P = 0.049$). Latta and McKay (2002) reiterated that, based on simulations of differentiation between neutral markers and QTL under selection, no relationship is expected between Q_{ST} and F_{ST}. In contrast to

Merila and Crnokrak (2001), Latta and McKay (2002) conclude that as theory predicts, levels of differentiation at neutral markers does not predict differentiation observed in adaptively-significant traits.

The lack of concordance in these conclusions may be the result of differences in interpretation by the authors rather than contradictions in the primary literature. The direction and degree of correlation between Q_{ST} and F_{ST} were similar in both analyses, which is expected as both included a number of the same studies (Merila and Crnokrak 2001; McKay and Latta 2002). The correlation between these parameters is not consistent with that expected from simulation studies (Latta and McKay 2002). Additional reports in the primary literature comparing Q_{ST} and F_{ST} will serve to clarify the relationship between these two parameters in natural systems.

The comparison of genetic and morphological differentiation will be applied to two portions of this work. First, the pattern of genetic and morphological differentiation will be compared to identify evolutionary processes controlling phenotypic variation in *P. nigra* from western Europe. Then, simulation studies will be used to describe the effect of asymmetrical patterns of phenotypic optima across a landscape on the quantitative and genetic differentiation among populations and explore the effect of changing phenotypic optima (selective pressure) on the fitness of populations.

1.6 Evolutionary and genetic basis of leaf morphology

1.6.1 Evolutionary significance of leaf size

Leaf size is determined by a complex developmental process, which is influenced by both genetic and environmental factors. From an evolutionary perspective, much attention has been given to the broad correlation between leaf size and environmental conditions, with theoretical models and morphological surveys being used to ask whether an optimal leaf size exists for a particular environment. As the photosynthetic organ of the plant, leaf shape can be considered the most important feature influencing plant growth (Tsukaya 2005). In addition, leaf size must be optimized to maximize light capture but minimize water loss. The changes to climatic conditions predicted for the next century will present many plant species with novel environmental conditions. Although phenotypic plasticity may provide sufficient acclimation for plants to survive in the short-term, should the environment change too drastically variation in the genetic control of leaf development may be required for a species to adapt. Understanding the level of natural genetic variance in these mechanisms will be critical to predict the functional response of natural populations to changing conditions.

Potential correlations between leaf traits and habitat have long been of interest to botanists and evolutionary biologists (Bailey and Sinnott 1916; Ackerly and Reich 1999). Attempts to correlate leaf size and function with environmental variables have proved difficult, partly due to the extreme variation in leaf traits observed within some environments (Bailey and Sinnott 1916; Parkhurst and Loucks 1972), and partly because of the complications presented by the phylogenetic relationships of species in broad-scale comparisons (Ackerly and Reich 1999). Various studies of single species have demonstrated that smaller leaf size conveys an adaptive advantage in xeric environments (Dudley 1996a, b; Picotte et al. 2007). Thus, variation in leaf size in forest trees may correspond to adaptations to various environments.

In order to understand which physical factors limit leaf size, theoretical models have been employed to predict leaf size for different environments (Parkhurst and Loucks 1972). Based on the principle of optimal design, which holds that natural selection will favour forms and functions that provide an organism with the greatest growth and reproduction in a given environment, Parkhurst and Loucks (1972) modeled how the size of a leaf might change (increase or decrease) in order to maximize water use efficiency (WUE) in different environments. By modelling how the physical properties of a leaf (e.g., absorbed radiation, stomatal resistance, mesophyll resistance, and leaf temperature) interact in different environmental conditions (i.e. light intensity, CO_2 availability, temperature, and relative humidity), Parkhurst and Loucks (1972) were able to predict which leaf size would optimize WUE in various environments. Results of a series of tests using these complex models can be generalized in two trends. First, large leaves are predicted to optimize WUE only in warm or hot temperatures in deep shade, such as a tropical forest understory. Second, cooler conditions and environments with high light intensity are predicted to favour small leaves (Parkhurst and Loucks 1972). These trends are undoubtedly overgeneralizations, however, as individual trees display remarkable phenotypic plasticity in response to microclimate conditions. For instance, *Populus euphratica* displays two distinct leaf morphologies and corresponding variation in photosynthetic and WUE capacity in different areas of the crown (Wang et al. 1997). In addition, recent manipulative experiments indicate that traits of developing leaves are directly influenced by the climatic conditions (vapour pressure, light intensity, and CO_2 concentration) experienced by mature leaves on the same plant (Miyazawa et al. 2006).

Despite the generalizations, these models serve as basis for further functional experiments. For instance, as stomatal conductance is predicted to be a significant factor in determining leaf size, divergent populations displaying significantly different leaf sizes would be expected to display corresponding differences in this trait. In *Populus nigra*, significant variation in leaf size could thus correspond to significant differences in stomatal conductance, and thus WUE. While these theoretical studies provide insights into the possible physical factors limiting leaf size, they do not provide information into the genetic and developmental mechanisms controlling leaf development.

1.6.2 Developmental control of leaf size

Although a casual observer may consider leaves to have a simple structure, the typical shapes result from a complex developmental process. If leaf size is an adaptive trait, patterns of genetic variation in the pathways controlling variation in leaf size should reflect the selective processes affecting natural populations.

Leaf development has been the focus of abundant research in order to identify the genetic and developmental mechanisms controlling leaf morphology. Leaf shape is defined through controlled growth along the two main axes (proximo-distal and lateral, Figure 1.3), resulting in the genetically defined but plastic leaf shape. An overview of the pathways, mechanisms, and genes described in the development of *Arabidopsis* leaves, the best-described model species, follows. Homologues of many of the genes described below have been identified in *Populus* species.

Leaf development begins in the shoot apical meristem (SAM), the region of undeterminant cell division in the plant stem (Tsukaya 2005). Meristematic growth is maintained by the action of *KNOTTED 1-like homeobox* (class I *KNOX*) genes, which are highly expressed in the tip of the meristem (Tsukaya 2005). The function of the *KNOX* gene *SHOOT MERISTEMLESS (STM)* is down-regulated by two genes which interact to repress *KNOX* activity and promote leaf development, *ASYMMETRIC LEAVES1* (*AS1*) and *ASYMMETRIC LEAVES2* (*AS2*) (Tsukaya 2005; Hay et al. 2006). In addition, the plant hormone auxin collects into local optima at the points of leaf initiation, and appears to be required for the repression of some *KNOX* genes, in particular *BREVIPEDICELLUS* (*BP,* Hay et al. 2006). The interaction of these factors serves to promote the development of a conical initial leaf primordium (Tsukaya 2005). Dorsoventrality in the leaf is established at this early developmental stage by the activation of a short-lived marginal meristem which converts the rod-shaped initial leaf primordia into spatula-shaped structures (Tsukaya 2005). This process defines the adaxial (top or meristem side) and abaxial (bottom or stem side) of the leaf through the activation of cell division on opposite poles of the initial primordial (Tsukaya 2005).

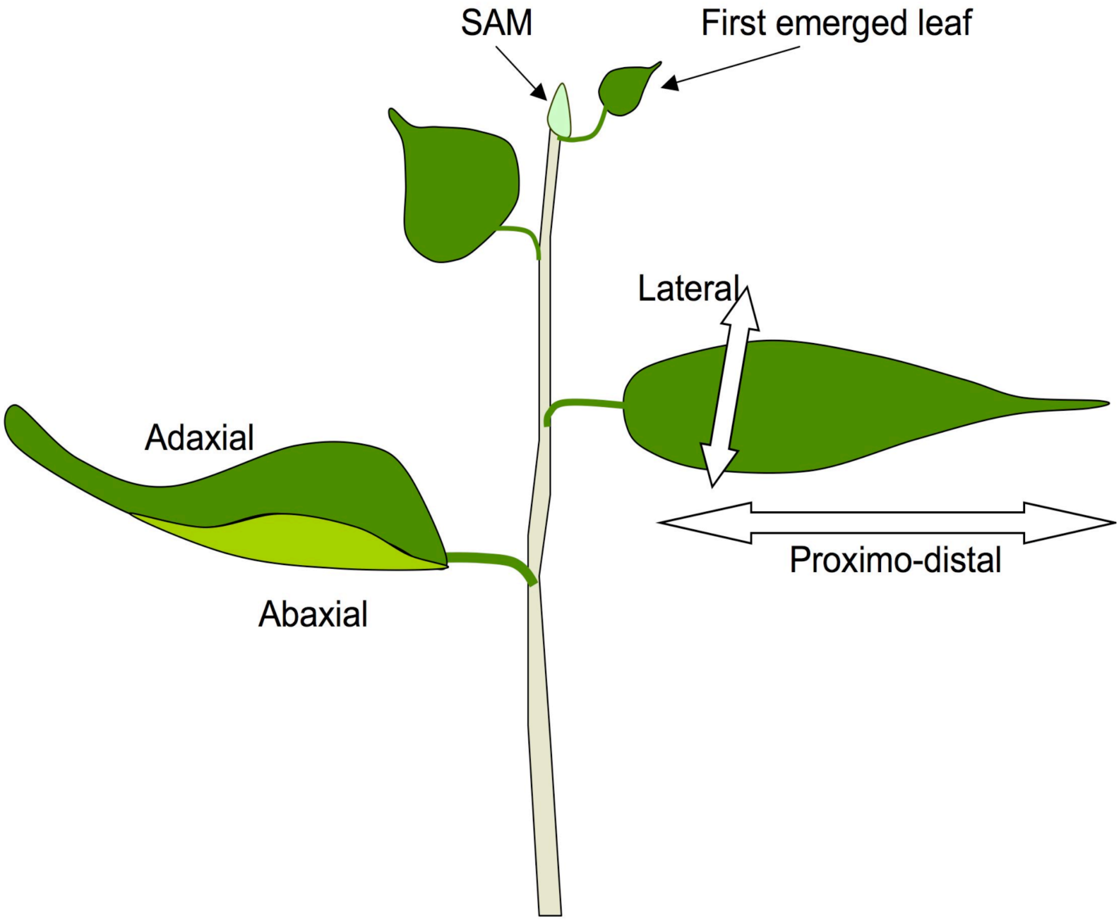


Figure 1.3 Leaf initiation occurs in the shoot apical meristem (SAM), from which developing leaves emerge and begin to expand. Leaf expansion is controlled along the proxim-distal (leaf length) and lateral (leaf width) axes to produce a leaf shape characteristic of a species. Dorsoventrality of the leaf results from the determination of abaxial and adaxial surfaces during leaf initiation in the SAM.

A distinct set of genes are involved in the development of the two surfaces of a leaf, consistent with the differences in form and function displayed by each. Abaxilization, or development of the bottom surface of the leaf, is controlled by class III HD-ZIP genes including *REVOLUTA*, *PHABULOSA*, and *PHAVOLUTA*, which are in turn regulated by micro RNAs (miRNA, Tsukaya 2005). In contrast, adaxialization, or development of the top surface of the leaf, involves members of the *KANDI* and *YABBI* gene families (Tsukaya 2005).

Ultimately, the characteristic flat surface of the leaf is formed by growth along the proximo-distal and lateral axes. Although these processes are less well-defined, it appears distinct genetic mechanisms control growth in the two directions. Leaf length is determined by growth along the proximo-distal axis, and is controlled in part by the gene *ROTUNDIFOLIA4* (*ROT4*). *ROT4* appears to function as a growth regulator via suppression of cell division, as overexpression of the *ROT4* functional domain results in shorter leaf length with no change in cell size (Tsukaya 2005). Further, loss of function mutants display no change in phenotype, likely due to the redundant nature of this developmental path. Genetic control of leaf width is not as well described, however. Though loss of function in the gene *ANGUSTIFOLIA3* has been shown to result in narrow leaves containing fewer cells, the exact function of this gene is not known (Tsukaya 2005).

Populus homologues of two *Arabidopsis* genes have been characterized through gene expression and mutant phenotype studies (Groover et al. 2006a; Du et al. 2009). The findings indicate that the genes controlling leaf initiation in *Arabidopsis* have dual roles influencing leaf initiation and secondary growth in *Populus*. The *Populus* homolog of *STM*, *ARBORKNOX1* (*ARK1*) has been implicated in the regulation of both the shoot apical meristem and vascular cambium (Groover et al. 2006a). The *Populus* homolog of *KNAT1*, *ARBORKNOX2* (*ARK2*), is involved in both the shoot apical meristem and vascular cambium, but is primarily expressed in lignified tissues (Du et al. 2009). Manipulation of expression levels of *ARK2*, however, affected internode and leaf phenotypes in transgenic plants (Du et al. 2009). Direct evidence of interaction between *AS1* and *ARK1* has not been reported in Populus as it has between *AS1* and *KNAT1* in Arabidopsis (Byrne et al. 2000). However, given the phenotypic evidence from mutant and over-expression analyses, it is likely these homologous gene models play similar roles in leaf initiation and development in *Populus* as they do in *Arabidopsis*.

1.6.3 Cellular control of leaf morphology

As should be apparent from even this brief review, a complex set of genetic pathways regulates leaf size. Leaf size is ultimately regulated by the rate and extent of cell growth and division during organ development. As a result, those genes and pathways affecting the cell cycle and cellular differentiation may contribute to the genetic determination of leaf size. The plant cell cycle follows the basic four-phase model of all eukaryotes, and several of the genes involved in the progression of the cell cycle have been identified in *Arabidopsis* and other model plant systems. In addition, leaf development also involves a number of genetic pathways unique to plants, such as those governing cell wall development and stomatal patterning.

Cell division is controlled by the cell cycle, a four-phase model common to proliferating cells (Figure 1.4; Francis 2007). Diploid cells in the postmitotic gap (G0/G1) phase contain a single copy of each chromosome, and undergo RNA and protein synthesis required for cell growth. The G0/G1 phase is maintained through the suppression of DNA replication, with the G0 distinction given cells either temporarily or permanently arrested from cell division, and the G1 phase marking the normal preparation for progression to the synthesis phase. The synthesis (S) phase is marked by DNA replication, which begins at predetermined replication origins throughout the genome (Francis 2007). In animal and yeast systems, the G1/S transition marks a commitment to complete the cell cycle, producing two daughter cells which re-enter the G1 phase, but it is still unclear whether a similar stringent control is present in plant systems (Shen 2001). Following DNA replication, cells enter the premitotic gap (G2) phase, which is marked by additional growth and protein synthesis in preparation for mitosis and cell division. Cells may arrest in G2, as they might in the G0/G1 phase (Inze 2005). Finally, when the cell reaches a sufficient size for the current environment and stage of development, it passes into the mitotic (M) phase, whereby the nucleus divides in two, each containing single copies of the genome. Mitosis is typically followed by cytokinesis, the division of the cell into two daughter cells (Francis 2007). Several plant tissues, including endosperm or cotyledons, undergo repeated rounds of DNA synthesis without undergoing mitosis. This process, called endoreduplication, results in cells with varying levels of ploidy, and is associated with large cell size (Shen 2001).

An intriguing pattern of growth has been observed in several mutant lines of *Arabidopsis* whereby a reduction in cell number is accompanied by an increase in cell size. Termed compensation, this phenomenon produces a leaf size that is not remarkably smaller than wild-type plants, especially when compared to the small-leaved plants described above (Horiguchi et al. 2006). Mutations conferring compensatory growth have been detected in the genes *an3/gif1*, the water-use regulator and cell division factor *erecta*, and *more axillary branches2/oresara9* (*max2/ore9*, Horiguchi et al. 2006). Initial models to explain this phenomenon involve either a decoupling of cell growth and cell division, or whole-organ regulation of cell growth via some manner of cell-cell communication (Horiguchi et al. 2006). Two mutants of interest are *an3-4*, which displays a severe reduction in cell number and compensatory growth, and *atgrf5-1*, which displays a mild reduction in cell number and no change in cell size (Horiguchi et al. 2006). These contrasting phenotypes have been cited as evidence of a potential threshold effect, such that compensatory growth is only triggered in response to some minimum number of cells (Horiguchi et al. 2006). However, although the two gene products interact, it is possible that the two mutations affect different regulatory paths or may be differentially redundant. Ultimately, the complexity and redundancy of pathways involved in leaf development may make such models premature.

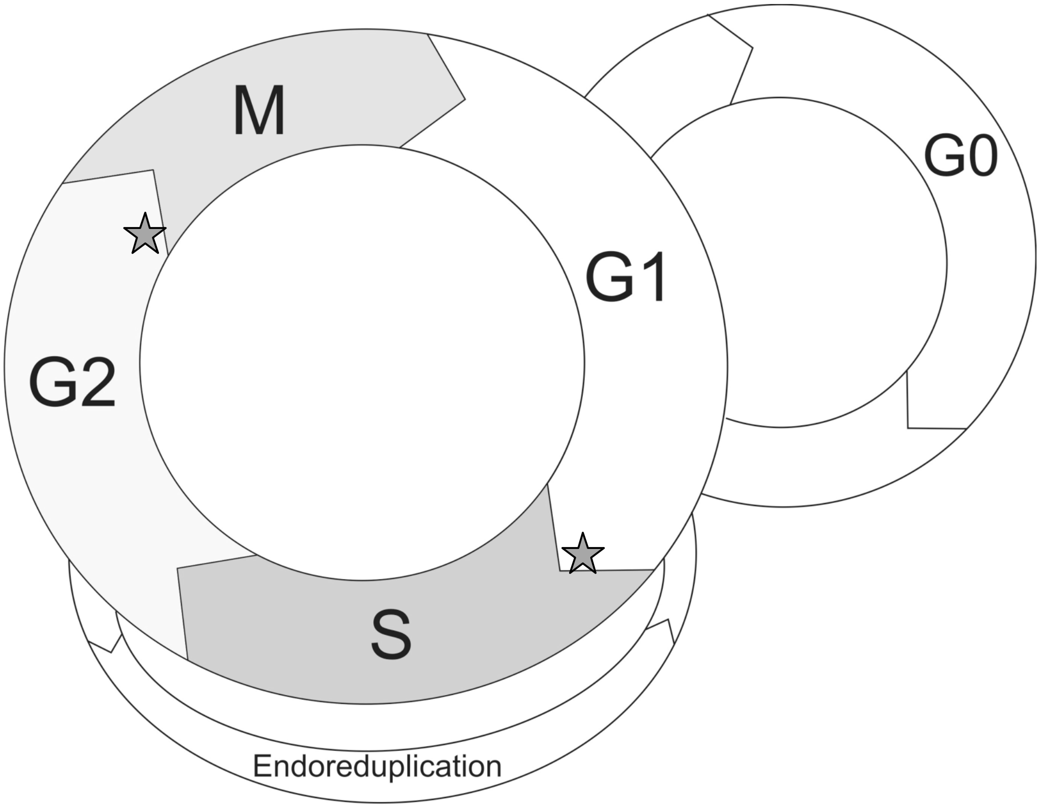


Figure 1.4 The cell cycle governs the growth and division of cells. The four phases of the cycle include postmitotic interphase (G1), DNA replication or synthesis (S), premitotic interphase (G2), and nuclear mitosis typically followed by cytokinesis (M). Progression through the cell cycle is regulated at the G1/S and G2/M checkpoints (denoted by ★). Cells may exit the cell cycle either temporarily or permanently (G0). Endoreduplication, or progression through DNA synthesis without accompanying mitosis and cytokinesis, results in cells containing elevated ploidy levels.

While the importance of compensatory growth has not been examined in morphological variation in non-mutant plants, a pattern consistent with compensation in cell size was described in a recent study of *Populus nigra* (Trewin 2008). In this common garden experiment, leaf area and epidermal cell size were significantly negatively correlated, indicating that smaller leaves were composed of a smaller number of larger cells. This pattern indicates that further examination of the role of cell division in leaf development in *P. nigra* is warranted. In addition, given the genetic pathways described for leaf development in *Arabidopsis* and the homologous genes described in *Populus* spp., examination of leaf development in *P. nigra* may provide insight into the genetic control of morphological variation in natural populations.

1.8 Populus as a model system

Populus has proven to be a powerful model for molecular genetic as well as quantitative genetic studies for a variety of reasons. Approximately 30 species of *Populus* occur throughout the Northern Hemisphere, and inhabit a range of ecological conditions (Eckenwalder 1996). Cultivation and breeding programs designed to increase yield in poplar have provided a wealth of physiological information on the genus (Bradshaw et al. 2000; Taylor 2002). Species are fast-growing (Zsuffa et al. 1996), phenotypically diverse (Eckenwalder 1996; Taylor 2002), readily propagate vegetatively (Bradshaw 1996), reach reproductive maturity at a young age (4 to 6 years, Tuskan et al. 2006), The dioecious breeding system common to the genus facilitates controlled crosses within and between species of *Populus*, although this trait does preclude self-pollinations (Bradshaw et al. 2000). Female plants are copious seed producers (hundreds of seed per plant) and, although seed is fragile and viability is short lived, establish well in greenhouse environment (Bradshaw et al. 2000). Together, these characteristics provide for powerful experimental manipulation and replication, as genets become potentially immortal through successive vegetative cuttings (Bradshaw et al. 2000; Taylor 2002). Further, species readily hybridize, producing fertile hybrids which can in turn be vegetatively propagated, allowing for experimental manipulation and comparison with previous and subsequent hybrid generations.

Cytogenetic studies have established that members of *Populus* are diploid with 19 linkage groups (Bradshaw et al. 1994; Cervera et al. 2001; Tuskan et al. 2006). The genome size of poplar is small relative to other forest trees (ca. 485 Mb; 50 times smaller than *Pinus* spp.; (Tuskan et al. 2006), making it larger than rice (389 Mb; International Rice Genome Sequencing Project 2005), but smaller than other agronomic species (e.g. the maize genome having 2300 Mb; Wei et al. 2007). Previous studies have estimated the physical/genetic distance ratio in poplar to be similar to *Arabidopsis* (200 kb/cM; Bradshaw et al. 2000), and have confirmed a rapid decay of linkage disequilibrium in *P. tremula* (Ingvarsson 2005), while slightly higher in *P. alba* in a natural hybrid zone (Lexer et al. 2007). While the high rate of decay of LD indicates that association studies based on neutral markers will require a dense coverage of the genome, the vast genomic resources available in this species make such studies feasible.

A wide range of genetic markers have been used to assess the genetic variation within and among species of *Populus*, including isozymes (Cheliak and Dancik 1982; Rajora 1990; Gallo and Geburek 1991; Liu and Furnier 1993a; Hipkins and Kitzmiller 2004; Braatne et al. 2006), AFLPs (Cervera et al. 2001; Fossati et al. 2004), microsatellites (Tuskan et al. 2004; Chen et al. 2007), RAPDs (Tuskan et al. 1996; Stevens et al. 1999; Yin et al. 2002), and RFLPs (Liu and Furnier 1993a; Bradshaw et al. 1994; Holderegger et al. 2005). Results have indicated *Populus* species may be the most genetically variable of any tree species (Cole 2005; Tuskan et al. 2006), and highly heterozygous (Kelleher et al. 2007).

The high level of allelic diversity within and ready hybridization between *Populus* species has resulted in the development of numerous independent linkage maps for the genus. As genetic marker technology progresses, the linkage maps have been either reassessed or improved with additional loci. Linkage maps are available for *P. deltoides* (Wu et al. 2000; Cervera et al. 2001), *P. trichocarpa* (Cervera et al. 2001; Kelleher et al. 2007), *P. nigra* (Cervera et al. 2001; Gaudet et al. 2008), *P. tremuloides* (Liu and Furnier 1993b), and various hybrid lines (Bradshaw et al. 1994; Yin et al. 2002; Yin et al. 2004). Many of these linkage maps have served as the basis for studies to identify quantitative trait loci (QTL), including leaf morphology (Wu et al. 1997; Rae et al. 2006), height or biomass (Bradshaw and Stettler 1995; Wu et al. 1998; Rae et al. 2008), branching architecture (Wu et al. 1998), phenology (Bradshaw and Stettler 1995; Frewen et al. 2000; Chen et al. 2002), osmotic potential (Tschaplinski et al. 2006), morphological response to elevated carbon dioxide concentrations (Ferris et al. 2002; Rae et al. 2006; Rae et al. 2007), and metabolite processes (Morreel et al. 2006). In addition to the genetic resources available, *Populus* can be efficiently transformed (Han et al. 1996; Groover et al. 2006b), allowing direct manipulation of gene expression and the assessment of functional roles of individual genes.

In confirmation of *Populus*' status as the model tree species, an international consortium selected a poplar species as the first tree genome to be sequenced. In 2006 the complete genomic sequence of *Populus trichocarpa* was released (Tuskan et al. 2006) and updated regularly (Kelleher et al. 2007; http://www.phytozome.net/poplar). In the process of completing this achievement, a number of genomic resources were developed and made available to the poplar research community. A database of microsatellite markers is available (http://www.ornl.gov/sci/ipgc/ssr_resource.htm), as is a set of over 100,000 expressed sequence tags (ESTs) representing over 11,000 unigenes (Sterky et al. 2004). In addition, two microarray platforms have been developed for gene expression studies (Sterky et al. 2004).

Outside of the laboratory, *Populus* species inhabit a range of environmental conditions, from desert to floodplain, play a vital role in many of the Northern Hemisphere's ecosystems, and thus, prove a powerful system to study ecological and evolutionary processes. A natural hybrid zone between two *Populus* species has recently been the focus of numerous studies investigating the relationship between poplar genetic variation and insect community structure in terrestrial and riparian systems (Bailey et al. 2006; Bangert et al. 2006; Wimp et al. 2007; Schweitzer et al. 2008). In terrestrial systems, the plant neutral genetic variation (an indicator of hybrid class), leaf biochemistry, and insect community structure were all correlated and related (Bangert et al. 2006). Similarly, correlations between hybrid class, leaf litter decomposition rates, and aquatic invertebrate community structure were found in a riparian zone (LeRoy et al. 2006). Together, these findings indicate that *Populus* diversity influences the larger ecological community across trophic levels. Closer investigation revealed that biochemical properties do not accurately predict the community structure of highly specialized insects in wild stands, indicating that other genetically-determined traits must be responsible for the correlation between plant genetics and insect association (Wimp et al. 2007). The relationship between plant phenology or leaf development and insect association may provide insight into the complex ecology of *Populus* species. Thus, hybrid *Populus* pedigrees provide an appropriate and likely powerful study system with which to resolve further the relationship between leaf morphology and interspecific interactions.

One species of poplar likely to be negatively impacted by climatic changes, *Populus nigra* L. (black poplar), is a pioneer tree species found throughout Central and South Europe into Central and West Asia, and into Northern Africa (Storme et al. 2004; Gaudet et al. 2008). As an early-successional species, *P. nigra* colonizes bare soil resulting from flooding. These frequent disturbance events can result in *P. nigra* persisting as a metapopulation, with frequent local extinctions and colonisations taking place (Imbert and LeFevre 2003). Seed are small, wind-pollinated, produced in copious amounts (Fossati et al. 2003). Genets (single genetic individuals) may also propagate vegetatively via broken branches or cuttings (Storme et al. 2004) or through the production of shoots from root suckers (Smulders et al. 2008b). Natural populations of *P. nigra* are increasingly fragmented and threatened by anthropogenic development and hydrological changes of their riparian habitat (Imbert and LeFevre 2003), hybridization with non-native *Populus* species (Benetka et al. 1999; Fossati et al. 2003; Holderegger et al. 2005), and displacement by fast-growing *Populus* hybrids (Gaudet et al. 2008). As a result of the decline observed over the past centuries, *P. nigra* is the subject of conservation efforts, including restoration efforts and an *ex situ* propagation programme (Storme et al. 2004).

Of the six sections recognized in the *Populus* genus based on morphological analyses (Cervera et al. 2005); Table 1.1), *P. nigra* is currently placed in sect. *Aigeiros* subsect. *Euroasiaticae*, and has been shown to be morphologically similar to the American cottonwoods *P. deltoides* and *P. fremontii*, in sect. *Aigeiros* subsect. *Americanae* (Eckenwalder 1984a, b). Phylogenetic analysis of chloroplast DNA (cpDNA) and ribosomal DNA (rDNA) revealed a complex ancestry for *P. nigra*, however (Smith and Sytsma 1990; Hamzeh and Dayanandan 2004). Rather than clustering with other members of sect. *Aigeiros*, cpDNA analyses place *P. nigra* within sect. *Populus*, the most reproductively isolated and diverged section within the genus (Smith and Sytsma 1990; Hamzeh and Dayanandan 2004). Given the lack of congruence between morphological and cpDNA taxonomic classification of *P. nigra*, Smith and Sytsma (1990) suggest separate origins for the cpDNA and nuclear genomes, with *P. nigra* evolving from the hybridization of a "pre-*P. nigra*" paternal parent with a *P. alba* (or *P. alba* ancestor) maternal parent, followed by repeated backcrossing with the paternal species. Cladistic and majority rule analyses, as well as phylogenetic analysis of restriction site mutations in the chloroplast genome of 12 species of *Populus* indicate that *P. nigra* and *P. alba* (sect. *Populus* subsect. *Tomentosae*; white poplar) share a most recent common ancestor (Smith and Sytsma 1990). Subsequent and more comprehensive sequence analyses of three noncoding regions of cpDNA confirm a likely hybrid origin of *P. nigra*, but indicate that the maternal ancestor was more closely related to *P. tremula* or *P. davidiana* rather than *P. alba*, given the greater number of nucleotide substitutions observed between *P. nigra* and the latter species (Hamzeh and Dayanandan 2004).

Initial analyses of nuclear ribosomal DNA (rDNA) sequences were unable to resolve the relationship between *P. nigra*, the clade containing *P. alba*, and the clade containing both *P. deltoides* and *P. fremontii* (Smith and Sytsma 1990). Subsequent analyses have further resolved the relationship between species within sect. *Aigeiros*, but the relationships are still the subject of debate (Rajora and Dancik 1995; Cervera et al. 2005). AFLP analysis of 25 species of *Populus* revealed phylogenetic relationships consistent with the historic taxonomic sections, except that greater differentiation was observed between *P. nigra* and *P. deltoides* than predicted from morphological characteristics (Cervera et al. 2005). Analysis of the ITS1 and ITS2 rDNA sequences consistently place *P. nigra* within sect. *Aigeiros* (Hamzeh and Dayanandan 2004). However, these same rDNA analyses failed to resolve the relationship between sect. *Aigeiros* and sect. *Tacamahaca*, indicating these sections are likely polyphyletic, and may be evidence that the taxonomy should be revised and perhaps the sections collapsed into one (Hamzeh and Dayanandan 2004). Ultimately, these findings at least indicate a complex ancestry of *P. nigra*, indicating a hybrid origin of this species, rather than divergence from a common ancestor shared by *P. deltoides*. Further, results of sequence analyses may support Rajora and Dancik's (1995) nomination of a new subsection for *P. nigra*, *Nigrae*, given the topological uniqueness of *P. nigra* within the genus taxonomy.

Table 1.1 Taxonomic sections of the genus *Populus* and common species within each, in order from ancient to recent divergence, based on AFLP profiles.

Section	Species	Distribution
Abaso Ecken.	*P. mexicana* Wesm.	Mexico, Central America
Turanga Bge.	*P. euphratica* Oliv.	Southwest to Central Asia
Leucoides Spach.	*P. ciliata* Wall.	Central Asia, Himalayas
	P. wilsonii Schneid.	Eastern Asia
Leuce Duby	*P. alba* L.	Southern Europe, Central Asia
	P. davidiana Schneid.	Eastern Asia
	P. tremula L.	North, Central Europe
	P. tremuloides Michx.	North America
Tacamahaca Spach.	*P. balsamifera* L.	Northern North America
	P. koreana Rehd.	Eastern Asia
	P. maximowiczii Henry	Northeast Asia
	P. trichocarpa Torr. & Gray	Western North America
Aigeiros Duby	*P. deltoides* Marsh.	Eastern North America
	P. fremontii Wats.	Western North America
	P. nigra L.*	Central, South Europe

From Cervera et al. (2005)

*Chloroplast and nuclear DNA fragment analyses reveal *P. nigra* to be distinct from *P. deltoides* and *P. fremontii*, prompting nomination of a new section, *Nigrae* (Rajora and Dancik 1995).

Patterns of genetic diversity in biparentally-inherited (nuclear) loci indicate that populations of *P. nigra* over large geographic areas are moderately differentiated, and that significant structure may exist at smaller scales. Two works report similar levels of differentiation across central European collections (Storme et al. 2004; Smulders et al. 2008b). Storme et al. (2004) studied samples from nine gene bank collection to determine the proportion of replicate genets in each collection, and describe the genetic differentiation within and among genets, river systems, and countries. Results from isozyme, AFLP, and microsatellite analyses revealed moderate levels of genetic variation in the collections and indicates genetic diverstiy is greatest in the southern collections from Spain, France, and Italy (Storme et al. 2004). In the second study, samples were collected from 17 natural populations occurring in seven river catchments (Smulders et al. 2008b). Analyses of AFLP and microsatellite markers revealed similar levels of genetic variation as reported by Storme et al. (2004), and indicated significant genetic differentiation among populations (AFLP $F_{ST} = 0.268$, $P < 0.001$; microsatellite $F_{ST} = 0.081$, $P < 0.01$; Smulders et al. 2008b). Analysis of individual samples revealed similar geographic patterns to genetic differentiation as described by Storme et al. (2004), and indicates that demographic patterns must be accounted for in association studies of this system (Gonzalez-Martinez et al. 2006; Smulders et al. 2008b). In both studies AFLP data indicated greater differentiation than microsatellites, likely due to the greater number of alleles revealed in microsatellite analyses compared to the dominant nature of AFLP data (Storme et al. 2004; Smulders et al. 2008b).

On a smaller scale, populations located on the Drôme River in France displayed low but significant levels of differentiation ($F_{ST} = 0.046$ and 0.041 for different regions of the river), and isolation by distance, but not anisotropic gene flow (Imbert and LeFevre 2003). Patterns of genotypic diversity within populations indicate that levels of management of the river systems may influence the type of recruitment occurring within stands. Greater numbers of multi-ramet genets were observed in highly canalized and managed rivers than in systems experiencing regular flooding regimes, possibly due to inefficient seedling recruitment in the absence of disturbed habitat (Smulders et al. 2008b). Thus, changes to river hydrology, be it due to anthropogenic development or changes in precipitation, may decrease the genetic diversity of populations as stems reproduce vegetatively rather than sexually. As a species reliant on regularly flooded riparian habitat for propagation and survival, *P. nigra* may be susceptible to further decline over the next century as precipitation patterns change. Given the genomic resources available for study of its genetic structure, previous studies describing its morphology and water use efficiency, and its designation as a species of concern by national and international bodies, *P. nigra* proves a valid study system to characterize the genetic basis of evolution to differing precipitation conditions.

The studies reported herein aimed to identify the evolutionary and genetic basis of morphological variation in *Populus nigra*, the European black poplar. First, morphological variation previously reported from a common garden experiment was compared to neutral genetic differentiation among and the climate experienced by populations to determine whether local adaptation, historic vicariance or both affected patterns of phentotypic variation. Second, morphological variation in *P. nigra* from central Europe (Hungary and Bosnia & Herezegovina) was compared to trees from western Europe in a greenhouse experiment in order to resolve further the phenotypic differentiation in this species. Third, a fine-scale study of leaf expansion and gene expression was conducted to determine whether cell division or cell growth drives differences in leaf size among four distinct genets. Fourth, in order to assess the relationship between the genetic basis of leaf morphology and insect preference, a hybrid pedigree of *Populus* was assessed for leaf damage in a common garden study in order to identify QTL for insect association. Finally, simulation studies were conducted to investigate the genetic consequences of evolutionary processes resulting from divergent isolated refugia and a changing phenotypic optimum, as may be predicted in a changing climate. Together these studies provide evidence that historic vicariance events may influence patterns of phenotypic differentiation and local adaptation, that leaf size variation is driven by differences in cell division among genets, that leaf development processes may significantly influence insect preference, and that the potential adaptation by natural populations to a changing climate may be influenced by previous patterns of natural selection.

Chapter 2. Evolution of morphological differentiation in *Populus nigra*

2.0 Overview

Variation in morphological traits such as leaf size frequently results from natural selection due to the adaptive nature of these quantitative traits. *Populus nigra*, European black poplar, displays significant variation in morphology among populations sampled across western Europe. In order to infer whether natural selection affected morphological differentiation in *P. nigra*, the pattern of phenotypic variation was compared with the genetic differentiation assessed at neutral molecular markers. The process of isolation by distance significantly affected variation among populations, indicating gene flow decreases as a function of distance and may influence morphological divergence. However, a significant correlation between morphological and genetic differences indicate populations of *P. nigra* are also isolated by adaptation, where populations similar phenotypically are more similar genetically than expected at random. The pattern of morphological differences among populations did not correspond to climate variables, but did correspond to historic glacial refugia. A model of morphological evolution in western populations of *P. nigra* is proposed whereby populations in France were recolonized by the small-leaf phenotype from the Iberian peninsula, but the current climate is more similar to that associated with the large-leaf phenotype observed in more eastern populations. As a result, *P. nigra* in central France display morphology intermediate to the small-leaf Spanish samples and large-leaf phenotype typical of central Europe, likely resulting from small-leaf founder populations evolving toward the large-leaf phenotype.

2.1 Introduction

Dissecting the evolutionary processes influencing morphological diversity is a question central to evolutionary biology. The relationship between plant morphology and environment has long been of interest in the botanical sciences (Bailey and Sinnott 1916; Parkhurst and Loucks 1972). Differences in plant morphology are considered adaptive, with phenotype predicted to confer increased fitness in the habitat (Turesson 1922; Westoby and Wright 2006). Phenotypic differentiation among populations reflects a balance between natural selection in the local environment, migration of alleles via gene flow (Antonovics 1968), and, at a lower frequency, the acquisition of novel alleles through mutation. Natural selection within a population must be strong enough to overcome gene flow from morphologically divergent populations in order to maintain phenotypic differentiation.

Genetic differentiation at neutral markers, by contrast, is driven not by adaptation, but by restricted gene flow between populations, genetic drift, and mutation of novel alleles (Wright 1931). Natural selection is not expected to act on the markers directly; genetic differentiation at neutral markers reflects demographic histories, i.e. the level of inbreeding and migration within and among populations. In plant species, mating system and dispersal mechanism have significant effects on the levels of genetic differentiation observed among populations (Loveless and Hamrick 1984; Duminil et al. 2007). The ever increasing literature comparing differentiation in morphological traits and that at neutral marker loci indicate that phenotypic differentiation is typically greater than neutral genetic differentiation, implying natural selection overcomes ongoing gene flow to maintain morphological variation (Merila and Crnokrak 2001; McKay and Latta 2002).

The consequences of these adaptive and neutral processes are not mutually exclusive, however, and emerging theory holds that each process may affect differentiation in the other. The well-established and widely tested theory of isolation by distance (IBD) predicts that genetic differentiation increases as a function of the geographic distance between populations (Slatkin 1993; Rousset 1997). That is, populations that are more distant geographically will have lower rates of gene flow, and will differentiate even in the absence of divergent selection. Geographic structure may result in morphological divergence due to genetic drift alone (Dennison and Baker 1991; Eckert et al. 1996).

Similarly, divergent phenotypic selection may be sufficient to drive genetic differentiation at neutral loci. Significant differentiation is expected to occur if a marker is tightly linked to a gene under divergent selection, and even loosely linked or completely neutral loci may show increased differentiation if selection is sufficient to reduce the fitness of maladapted migrants (LeCorre and Kremer 2003; Nosil et al. 2009). Over generations, genetic markers may become differentiated as a result of selective pressures unrelated to the marker loci or neighboring genes. While correlations between genes and environment have been reported in plant species often at smaller geographic scales (e.g. Kelly et al. 2003; Mitton and Duran 2004), isolation by adaptation (IBA) - the correlation between genes and morphology while controlling for geographic separation - is rarely considered in studies across larger geographic areas. A recent literature review identified only 24 studies addressing isolation by adaptation in genetic differentiation, with only five focusing on plant species, and none on trees (Nosil et al. 2009).

Isolation by adaptation may lead to an underestimation of the actual migration rate between populations as propagules, though reaching allopatric sites, fail to contribute to the local gene pool, such that the effective number of migrants approaches zero. Accurately estimating these demographic parameters is critical to conservation, ecological and evolutionary studies in the context of an increasingly fragmented landscape and changing climate. For example, consider a species with a typical morphology occurring through the majority of its range, but an atypical morphology occurring in marginal habitat at the range perimeter. Measures of historic gene flow (e.g. N_em from genetic markers) may underestimate the potential for the typical morphology to utilize the edge habitat if the populations are isolated by adaptation. That is, migrants of the typical morphology may reach the peripherial populations but fail to contribute to the gene pool. In the future, changes in precipitation, temperature, or growth season may make the marginal habitat more suitable for the typical morphology. Future migrants may be successful, but historic gene flow estimates will predict migration to be unlikely. Conservation strategies based on the measure of historic gene flow may then be less than optimally efficient.

In addition to isolation by distance and isolation by adaptation, historic vicariance, the geographic isolation of populations, may influence the morphological and genetic structure of populations. Typically, organellar genetic markers, such as chloroplast DNA (cpDNA) variants, are used to assess the historic genetic structure of a species due to the reduced rate of evolution in the haploid genome. These methods have been successful in reconstructing the postglacial histories of oak (Petit et al. 2002), birch (Palme et al. 2003), pine (González-Martínez et al. 2004), and *Populus nigra* (Cottrell et al. 2005) in Europe. The signature of historic vicariance events may remain in nuclear bi-parentally inherited markers as well as the organellar genome. Routes of postglacial recolonization have been reconstructed from allozyme data in the Sonoran cactus (Nason et al. 2002), and from SCAR markers in *Picea abies* (Scotti et al. 2000). If historic events induced significant adaptive differentiation, residual structure may be detectable in the morphological differentiation among extant populations derived from diverged gene pools. Greater morphological differences than climatic differences between populations may indicate recolonized populations have yet to reach a phenotypic optimum for a particular site. Such processes are most likely to be observed in long-lived species having overlapping generations, such as forest trees.

This study assesses the relative importance of isolation by distance, isolation by adaptation, and admixture resulting from historic vicariance in the evolution of morphological variation in *P. nigra*, European black poplar. A study recently reported significant differentiation in leaf and biomass traits across thirteen populations in Western Europe (Trewin et al., in preparation). In this collection of trees native to Spain, France, northern Italy, western Germany and the Netherlands, adaptive traits such as leaf area and stem diameter displayed a near bimodal distribution, with populations broadly corresponding to either a large-leaf or small-leaf phenotype.

Here, multivariate analyses are used to assess the pattern of morphological and genetic (microsatellite) differentiation in the same collection. Full and partial Mantel tests are then applied to test for IBD and IBA and to test if morphological differences correspond to climatic differences between extant populations.

These analyses reveal populations of *P. nigra* from France maintain a genetic signature of admixture between two historic gene pools, and that this structure is mimicked in patterns of morphological differentiation. Results provide evidence of IBD and IBA, indicating both geographic and adaptive processes contribute to genetic differentiation. In addition, while morphological and climatic measures were correlated, the pattern of differences between populations was inconsistent with morphology reflecting strict adaptation to local climate. Taken together, the results indicate the morphology of *P. nigra* from the admixture zone in France reflects the phenotype of the ancestral small-leaf gene pool, and that these trees are likely being selected towards a larger-leaf phenotype typical of populations from central Europe. The implications of these findings are discussed in the context of land management and climate change.

2.2 Methods

2.2.1 Study system and collections

Though widely distributed across Europe and into Asia and northern Africa, *Populus nigra* L. is restricted to floodplain habitat, resulting in a patchy occurrence across the broadest landscape scale. Unlike the canonical climax forest tree species, *P. nigra* is fast-growing, early successional, and dioecious, with male and female flowers occurring on separate plants. Together, these life history traits likely increase genetic differentiation among populations and reduce the rate of genetic homogenization of admixed populations.

This work represents a novel analysis of morphological data collected by H. Trewin, who describes the common garden experiment and morphological data collection in her PhD thesis (Trewin 2008). In summary, Trewin (2008) reports that in the spring of 2004 a common garden was established consisting of approximately 500 genets of *Populus nigra* planted in six replicate blocks in a plantation located near the Institute of Forestry and Game Management near Geraardsbergen, Belgium. Hardwood cuttings were planted directly into the ground in a fully replicated randomized design. Plantings were established on a grid with 0.75 x 2.0 m spacing, surrounded by a double row of the *Populus* cultivar 'Muur' to minimize edge effects (Trewin 2008). Once established, trees were cut back in early 2005, and regrowth pruned to a single dominant stem in June 2005 (Trewin 2008). Site management, carried out by the local collaborators, included mechanical weed removal and fungicide application. No irrigation or fertilizer was provided during the experiment. Samples represented trees collected from 15 natural populations of *P. nigra* from France, Germany, Italy, the Netherlands, and Spain (Table 2.1, Figure 2.1). Two of these populations represent collections made for breeding purposes (nos. 4 and 10), and are included in Table 2.1 for comparison to Trewin (2008), but were omitted from the analyses reported here.

Table 2.1 *Populus nigra* populations sampled for phenotypic analysis in a common garden study.

Pop. No.	Country	Population Name	Latitude	Longitude	N_M	N_G
1	France	Drôme 1	44.6833	5.4000	63	31
2	France	Drôme 6	44.7500	4.9167	63	45
3	France	Durance	43.7847	5.5569	12	9
4	France	Individual clones	45.1103	2.8976	0	0
5	France	Loire East	47.3295	2.9200	26	6
6	France	Loire West	47.2571	-0.5870	21	13
7	Germany	Kuhkopf	49.8167	8.5000	56	44
8	Italy	La Zelata	45.2667	8.9833	63	41
9	Italy	Siro Negri wood	45.2000	9.0667	44	32
10	Italy	POP5 pedigree	N/A	N/A	0	0
11	The Netherlands	Ijssel, Rhine	52.2250	5.9600	31	23
12	The Netherlands	Individual Clones	51.3583	4.5500	7	5
13	The Netherlands	Waal, Maas	51.8194	5.0900	12	9
14	Spain	Ebro 1	41.9333	-1.3833	54	20
15	Spain	Ebro 2	41.5833	-1.000	60	30

Latitude and longitude provided in dd.dddd.

N_M = number of genets analyzed for morphological traits

N_G = number analyzed for genetic markers

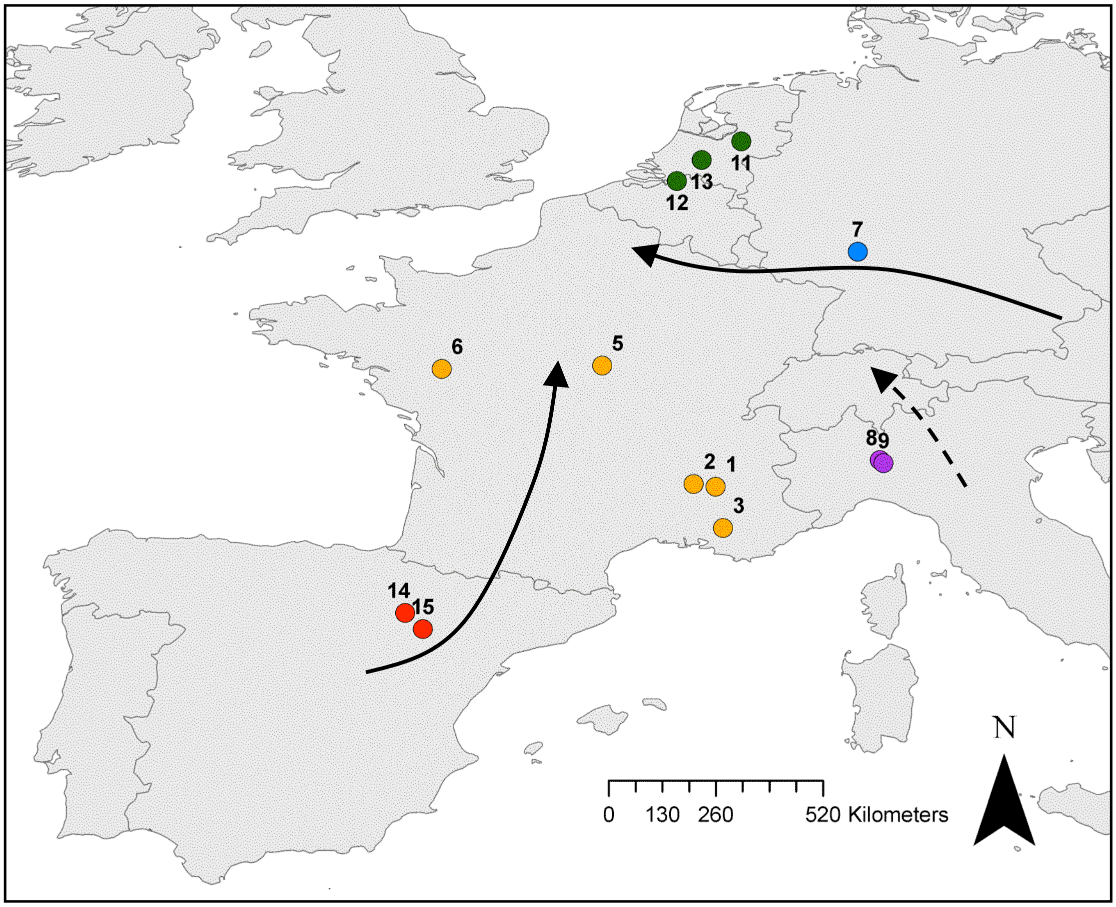


Figure 2.1 Locations of *Populus nigra* populations sampled in Spain (red), France (orange), The Netherlands (green), Germany (blue), and Italy (purple) for the common garden study (Trewin, 2008). Numbers follow Table 2.1. Arrows represent possible recolonization paths from refugia following the Pleistocene glaciation, as per Cottrell et al. (2005).

2.2.2 Phenotypic measures: quantifying differences in morphology

All phenotypic measures were conducted or overseen by H. Trewin as described in Trewin (2008). In the third year of growth (2006), H. Trewin assayed each ramet of *Populus nigra* in the common garden for a variety of morphological characteristics (Trewin 2008). Twelve traits were re-analyzed here. These include five leaf traits: leaf area, leaf length, leaf width, the leaf length-to-width ration, and specific leaf area; five cell traits: the number of stomata, stomatal density and stomatal index, mean epidermal cell area, and the number of epidermal cells per leaf all of the abaxial surface; and two biomass traits: the height of the stem and the diameter at the beginning of the growing season.

For each ramet, the youngest fully mature leaf on the leader stem was collected (N. Street, personal communication), and its image scanned at 200 DPI using an Umax Astra 6700 scanner (Trewin 2008). Digital images were assessed using the analysis software Image J (Image J.1.32j, Wayne Rasband, USA) to calculate the leaf area, leaf length from tip to petiole, and leaf width at the widest point (Trewin 2008). The ratio of leaf length to width was calculated from these measurements. This mature leaf was oven dried and the specific leaf area (SLA) then calculated as:

$$SLA = \frac{Leaf\ area\ (mm^2)}{Dry\ mass\ (mg)} \qquad (2.1)$$

In order to assess cellular characteristics, Trewin (2008) took cellular imprints from the abaxial surface of a single mature leaf from each ramet. Imprints were taken by covering an approximately 1 cm^2 area on the basal section of the leaf with clear nail varnish, allowing to dry, and then carefully peeling the varnish off with sellotape (e.g. Ferris et al. 2002). The sellotape was mounted to a microscope slide for further analysis. In the laboratory, cell imprints were captured as digital images using a Zeiss microscope at 400X magnification. Images were processed in Image J. For each field of view, the number of epidermal cells and number of stomata were counted, and the average cell area (mm^2) calculated from the areas of 10 randomly chosen epidermal cells (H. Trewin, PhD thesis). From these data the stomatal density (SD) and stomatal index (SI) of the abaxial surface and the number of cells per leaf (CN) were calculated as:

$$SD = \frac{Number\ of\ stomata}{Area\ in\ field\ of\ view} \tag{2.2}$$

$$SI = \frac{Number\ stomata}{Number\ of\ cells + Number\ of\ stomata} \times 100 \tag{2.3}$$

$$CN = \frac{Leaf\ area\,(mm^2)}{Epidermal\ cell\ area} \tag{2.4}$$

Stem height and diameter were measured at the beginning of the 2006 growing season (Trewin 2008).

2.2.3 Microsatellite analyses: resolving neutral genetic structure

Building on Trewin's (2008) work, I designed an analysis of genetic structure of the 13 native populations sampled for the common garden study. As genetic structure may be robustly estimated by as few as three samples per population (Grivet et al. 2008), a subset of the genets were analyzed for neutral nuclear markers (Table 2.1). Two populations (nos. 4 and 10) were omitted from genetic analysis as these samples are breeding individuals and were not sampled at random from natural stands.

H. Trewin collected leaves from each genet in the common garden experiment in 2004, flash frozen in liquid nitrogen and stored at -80°C until processed. In the fall of 2007, I, with the help of M. Nelson, conducted DNA extraction from these samples. Frozen leaf tissue was ground to a fine powder under liquid nitrogen using a mortar and pestle. Total genomic DNA was isolated from each sample using the Qiagen® DNeasy Mini kit (Crawley, UK) following the manufacturer's instructions. The quality and quantity of each sample was assessed using a ThermoScientific® NanoDrop 1000 spectrophotometer (Wilmington, Delaware, USA) following the manufacturer's instructions.

A panel of 10 bi-parentally inherited, co-dominant microsatellite markers was assayed for each sample. In order to minimize the potential for neutral variation to be linked to morphological variation in this collection, microsatellite loci were selected from genomic regions outside quantitative trait loci (QTL) “hotspots” for leaf traits as defined by Trewin *et al*. (in preparation). Primer sequences for each locus were obtained from the literature (Table 2.2). The forward primer of each pair was modified to incorporate a standard M13 allowing incorporation of a fluorescent M13 primer during amplification (Schuelke 2000). In addition, the reverse primer was modified with the 3’ addition of 5’-GTTTCTT-3’ in order to prevent the non-template addition of dATP by *Taq* polymerase (V. Jorge, personal communication). In order to visualize amplification products, the M13 primer included a 5’ fluorescent label suitable for analysis on ABI® capillary electrophoresis systems.

Each locus was amplified in separate polymerase chain reaction (PCR®) reactions containing 1X reaction buffer (supplied with enzyme), 1.5 mM $MgCl_2$, 2 nmoles of each dNTP, 0.5 pmole of the unlabelled locus-specific forward primer, 5 pmole of the locus-specific reverse primer, 5 pmole of the fluorescently-labelled M13 primer, and 0.2 Units Qiagen HotStarTaq® DNA polymerase in a 10.0 µL total reaction volume. Amplifications conditions included a hot start of 95°C for 15 min., followed by 42 cycles of 94°C for 20 sec., 55°C for 30 sec., and 72°C for 30 sec., with a final extension at 72°C for 5 min. Locus GCPM_2180-1 required 2.0 mM $MgCl_2$ for amplification. In order to assess amplification quality, ten samples (16%) were replicated for each locus.

Following amplification samples were diluted to optimize automated visualization. Capillary electrophorsis was carried out on an ABI3730 (Applied Biosystems, Inc., Foster City, California, USA) systems by GeneService Limited (Nottingham, UK). Results were analyzed and scores assigned using Peak Scanner v. 1.0 (Applied Biosystems, Inc., Foster City, California, USA). J. DeWoody conducted all PCR, purification, and ABI preparation in the laboratory.

2.2.4 Climatic data: quantifying differences in environment

In order to describe the environmental differences experienced by each population, a variety of meteorological data was collected for each site. Four measures of temperature were analyzed: mean annual temperature (MATemp, °C), temperature seasonality (standard deviation X 100, SDTemp), maximum temperature of the warmest month (MaxTemp, °C), minimum temperature of the coldest month (MinTemp, °C). Four measures of precipitation patterns were also analyzed: mean total annual precipitation (MAPpt, mm), precipitation seasonality (a coefficient of variation; VarPpt), precipitation of the wettest month (MaxPpt, mm), and precipitation of the driest month (MinPpt, mm). Temperature and precipitation data were taken from www.worldclim.org (Hijmans et al. 2005), a collection of data from the years 1950 to 2000. The 'WorldClim' database consists of temperature and precipitation data compiled from five sources which were interpolated using a thin-plate smoothing spline algorithm to produce climate surfaces over all land areas (except Antartica) at a 30 arc s resolution (Hijmans et al. 2005). This data set is considered at 1-km resolution because the 30 arc s grid corresponds to 0.86 km^2 at the equator and a smaller area at greater latitudes (Hijmans et al. 2005). Errors were removed by inspecting the value of each grid relative to the values in the surrounding eight grids (the neighborhood) and then iterating the algorithm (Hijmans et al. 2005). The resulting 'WorldClim' database contains temperature data from over 24,000 locations, precipitation data from over 47,000 locations, and maximum/minimimum temperature data from over 14,000 locations (Hijmans et al. 2005). The spatial scale of the climate surface is sufficient for examination of climate at scales >1 km, though variation within grids is high in mountainous areas and in areas with low coverage of data locations (Hijmans et al. 2005). Given the relatively good coverage of stations in Europe and the distance between sampling locations of *P. nigra*, the 'WorldClim' database is an appropriate source for climatic data for this study. In addition, the maximum day length on the summer solstice was included (MaxDay, h) from the U.S. Naval Observatory Astronomical Applications Department (http://aa.usno.navy.mil/data/docs/RS_OneDay.php).

In order to identify the most informative but least correlated variables from the climate data set, a discriminant factor analysis was conducted as a means of data reduction as implemented in SPSS (version 15.0, SPSS, Inc.). The correlation between variables was assessed and compared to the results of the factor analysis in order to identify a subset of variables that are minimally correlated but differentiate between climates at the sampling sites. The variance explained for each component was assessed individually and cumulatively.

Table 2.2 Linkage group, primer sequences and repeat motif of the ten microsatellite loci unlinked to leaf area QTL assessed as neutral markers in *P. nigra*. Primer sequences omit the M13 sequence added to the forward primer or stabilizing sequence added to each reverse primer described in the text.

Locus	**LG**	**Forward (5' to 3')**	**Reverse (5' to 3')**	**Motif**
WPMS_18[†]	I	CTTCACATAGGACATAGCAGCATC	CACCAGAGTCATCACCAGTTATTG	GTG
PMGC_2088[*]	II	TCACAAAAGGTTAACGACTTCG	CAGTACTCAGCTGCAGGTCC	GA
PMGC_2818[*]	II	AAGCTTCATCGTCCTGCTTG	CGTATCAATTCACGACTCTCG	GA
PMGC_486[*]	III	AGAAGTTGTTGAACCCGATGGG	GCTACAAACTTTGTTGTACCC	GA
PMGC_2879[*]	III	TTGATTCGAGCCTCACGAGC	AAACTCCAACATTTTAAGGACC	GA
WPMS_14[†]	V	CAGCCGCAGCCACTGAGAAATC	GCCTGCTGAGAAGACTGCCTTGAC	CGT
PMGC_2163[*]	X	CAATCGAAGGTAAGGTTAGTG	CGTTGGACATAGATCACACG	GA
PMGC_14[*]	XIII	TTCAGAATGTGCATGATGG	GTGATGATCTCACCGTTTG	CTT
WPMS_20[†]	XIII	GTGCGCACATCTATGACTATCG	ATCTTGTAATTCTCCGGGCATCT	TTCTGG
PMGC_2525[*]	XVIII	CGAGTCACAAGCTCCCAATAG	GCAGGCTGTCCTATCTGCG	GA

[*] *Populus* Molecular Genetics Cooperative (http://www.ornl.gov/sci/ipgc/ssr_resource.htm)

[†] (Smulders et al. 2001)

2.3 Data analysis

2.3.1 Pattern of phenotypic differentiation among populations

Phenotypic variation in each of the 12 traits measured in the common garden was previously analyzed with individual generalized linear models (Trewin 2008). All statistical analyses of phyenotypic, climatic, and genetic data described below were conceived and conducted by J. DeWoody. To assess overall morphological differentiation of *P. nigra* populations, a discriminant analysis was performed on the genet means of each trait measure. As a multivariate procedure, discriminant analysis maximizes differences between samples by reducing the variation in multiple measures into a small number of orthogonal factors. Interpretation of the correlation matrix, eigen values, and factor structure matrix produced by the discriminant analysis provides insight into which morphological measures contribute to each discriminant factor, and thus explain differences between populations. Those factors significantly contributing to differences between populations were used to visualize overall morphological differences between individuals and populations. Discriminant analyses were implemented using the Factor Analysis and Classification Analysis functions in SPSS (v. 15.0.0, SPSS, Inc., Chicago, Illinois, USA).

2.3.2 Pattern of genetic differentiation among populations

Six standard measures of genetic diversity were assessed for each population: percentage of polymorphic loci (P), mean alleles per locus (A), effective number of alleles (A_e), observed (H_o) and expected (H_e) heterozygosity, and the fixation index over loci (F), as implemented by GenAlEx v6 (Peakall and Smouse 2006). The presence of null alleles was assessed using the program MICROCHECKER (van Oosterhout et al. 2004) as described in DeWoody et al. (2006). Evidence of a recent genetic bottleneck was assessed for each population using the Wilcoxon sign-rank test under the two-phased model of mutation as implemented by the program BOTTLENECK (Luikart and Cornuet 1998; Luikart et al. 1998). Due to insufficient sample sizes, populations 3, 12 and 14 were omitted from the MICROCHECKER analsyis and populations 3, 5, and 12 from the BOTTLENECK analysis. Overall genetic differentiation was quantified as Φ_{PT} through analysis of molecular variance (AMOVA) as implemented by GenAlEx v6 (Peakall and Smouse 2006).

Admixture among populations was assessed using STRUCTURE v2.1 (Pritchard et al. 2000). The number of genetically similar clusters was estimated using the MCMC procedures and Bayesian likelihood methods. As this model was designed to quantify admixture in populations based on multilocus marker data for individuals, this is considered an *ad hoc* procedure for estimating the number of genetic groups (K) in a sample (Pritchard et al. 2000). The probability that the number of genetic clusters was estimated for each K in the set $K=\{1:13\}$, corresponding to a minimum of a single genetic unit to a maximum equal to the number of populations sampled. The likelihood of the observed data given K [ln(Pr(X|K))] was estimated over five simulation runs, with each run consisting of a burn in period of 50,000 replications with data collected for 100,000 MCMC reps. Correlation in allele frequencies was allowed among populations (allowing for variation in F_{ST} among subpopulations), and the remaining parameters were set to the default values. Assuming equal priors, the mean value of the likelihood statistic for each K was used to calculate the test statistic:

$$dK_i \approx \frac{e^{K_i}}{\sum_K e^{K_i}}. \tag{4.1}$$

The K with the greatest dK value (closest to 1) may be interpreted as the most likely number of genetic clusters in the sample set (Pritchard et al. 2000). For the K with the greatest probability, the proportion of each genotype assigned to each genetic cluster was assessed for geographic patterning of genetic similarity among populations.

In order to assess the phylogenetic relationship between the populations, a population phenogram was built from (Nei 1972) genetic distance for each pair of populations, using Neighbor Joining methods as implemented by PHYLIP (Felsenstein 2005). Significance was determined using extended majority rule, and is indicated as the percentage of times a node was observed in 1000 bootstrap replicates.

In addition, to compare the genetic differences between populations to the morphological differences summarized by the discriminant classification analysis, the microsatellite data were subjected to a principal coordinates analysis (PCoA) based on pairwise genetic distance (Peakall et al. 1995; Smouse and Peakall 1999) as implemented by GenAlEx v6 (Peakall and Smouse 2006). Similar to the discriminant classification analysis, this PCoA reduces variation in multilocus genotypic data to a small number of orthogonal coordinates based on the genetic distance matrix.

2.3.3 Tests for isolation by distance, isolation by adaptation, and correlation between morphological variation and climate

Correspondence between overall morphological variation, genetic differentiation, and climatic conditions was assessed using simple and partial Mantel tests. The Mantel is a form of regression analysis testing for correspondence between values in distance matrices by randomizing the position of the values in one matrix over multiple permutations and estimating the probability of observing the original data pattern (Smouse et al. 1986; Manly 1997). A significant correspondence between the matrix elements is described by a regression analysis of the matrix elements, with one matrix considered a dependent variable for the other, independent matrix. Partial Mantel tests assess correlation between two matrices while controlling for variation in a third indicator matrix. More specifically, the partial Mantel analyzes matrices built from the residuals of simple Mantel tests of both the dependent and independent distance matrices with a third indicator (or factor) matrix (Smouse et al. 1986). Here, the partial Mantel procedure was used to test for correspondence between genetic and geographic distance when controlling for morphological differences (a test for IBD), and to test for correspondence between genetic and morphological differences when controlling for geographic distance (a test for IBA).

Four matrices were built to describe differences between populations of *P. nigra*. First, morphological differences were assessed as the Euclidian distance between the population means for the first two factors of the discriminant classification analysis. Second, genetic differences were calculated as the Euclidian distance between the population means for the first two coordinates of the PCoA. Third, geographic distance was calculated from the latitude and longitude of the site of sample collection using GenAlEx (Peakall and Smouse 2006). Fourth, climatic differences between populations were summarized as the Euclidian distance between the MaxTemp and MinPpt for each population site.

Three simple Mantel tests were used to test for correlation between (1) genetic differences and geographic distance (Gen, Geo), (2) morphological and genetic differences (Morpho, Gen), and (3) morphological and climatic (Morpho, Clim) differences between populations. Then, two partial Mantel tests were performed. The first tested for isolation by distance (IBD) as the correlation between genetic and geographic differences when controlling for morphological differences (Gen, Geo | Morpho). The second tested for isolation by adaptation (IBA) as the correlation between morphological and genetic differences when controlling for geographic distances (Gen, Morpho | Geo).

Significant correlation between (Gen, Geo | Morpho) will indicate that gene flow is restricted across the landscape (IBD), independent of the morphological variation observed among populations. A significant partial correlation between (Gen, Morpho | Geo) will indicate that morphological divergence has influenced the genetic differentiation at these neutral loci through IBA. Where significant IBA was observed, regression analysis was used to describe the relationship between the residual morphological values and residual genetic values (the partial regression plots) for all pairs of populations, comparisons of small-leaf to small-leaf populations, large-leaf to large-leaf populations, and small-leaf to large-leaf populations (Moya-Larano and Corcobado 2008). All regression analyses were preformed in SPSS v17 (SPSS, Inc., Chicago, Illinois).

2.4 Results

2.4.1 Phenotypic differentiation among populations

Correlation between morphological measures varied within and among trait classes (leaf traits, cell traits and biomass traits; Table 2.3). Measures of leaf size were significantly correlated, and varied with biomass traits. Cell traits displayed lower but significant levels of correlation, with cell area varying inversely with leaf and biomass traits.

Table 2.3 Correlation matrix for 12 morphological traits measured in a common garden study of *Populus nigra*. Pairwise correlations are provided above the diagonal, and P-values from a one-tailed test of significance below the diagonal.

	Leaf area	**Leaf length**	**Leaf width**	**Leaf length:width**	**Specific leaf area**	**Cell area**	**Cell number**	**Stomatal number**	**Stomatal density**	**Stomatal index**	**Height**	**Diam.**
Leaf area		0.962	0.964	-0.236	0.913	-0.195	0.894	0.363	0.321	0.123	0.736	0.793
Leaf length	<0.001		0.938	-0.201	0.893	-0.200	0.872	0.351	0.316	0.121	0.733	0.790
Leaf width	<0.001	<0.001		-0.280	0.899	-0.209	0.859	0.339	0.277	0.085	0.744	0.815
Length:width	<0.001	<0.001	<0.001		-0.235	-0.206	-0.170	-0.075	-0.159	-0.161	-0.312	-0.234
Specific leaf area	<0.001	<0.001	<0.001	<0.001		-0.197	0.836	0.363	0.315	0.144	0.769	0.839
Cell area	<0.001	<0.001	<0.001	<0.001	<0.001		-0.404	-0.455	-0.297	-0.090	-0.089	-0.192
Cell number	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001		0.483	0.449	0.192	0.686	0.732
Stomatal number	<0.001	<0.001	<0.001	0.064	<0.001	<0.001	<0.001		0.724	0.402	0.282	0.320
Stomatal density	<0.001	<0.001	<0.001	0.001	<0.001	<0.001	<0.001	<0.001		0.579	0.300	0.279
Stomatal index	0.006	0.007	0.043	0.001	0.002	0.034	<0.001	<0.001	<0.001		0.124	0.110
Height	<0.001	<0.001	<0.001	<0.001	<0.001	0.036	<0.001	<0.001	<0.001	0.006		0.880
Diameter	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.013	<0.001	

The first two components of the discrimant factor analysis described 55% and 16% of the variation in the data, for a cumulative total of 71%. The most informative traits to distinguish between populations were leaf area, leaf length, leaf width, specific leaf area, cell number per leaf, stem height, and stem diameter on the first component (Table 2.4).

The discriminant cluster analysis revealed geographic patterning in morphology. Samples from Spain were morphologically distinct from those from Italy, Germany and the Netherlands, with the French populations intermediate (Figure 2.2). Overall, population means fell distinctly on opposite sides of the origin of the first factor, supporting the observation of a small-leaf and large-leaf distinction between populations. This pattern of morphological variation may be due to reduced gene flow between to the two groups of populations (large-leaf and small-leaf) or due to adaptive divergence, in which case the morphological variation would be expected to correspond to climatic differences between population sites.

Table 2.4. Discriminant factor analysis revealed leaf and biomass traits to distinguish between populations of *Populus nigra*.

Trait	**Component 1**	**Component 2**
Leaf area	0.948	-0.179
Leaf length	0.935	-0.178
Leaf width	0.938	-0.218
Leaf length:width	0.931	-0.168
Specific leaf area	0.520	0.697
Cell area	0.487	0.742
Cell number	0.248	0.660
Stomatal number	0.923	0.044
Stomatal density	0.832	-0.205
Stomatal index	0.878	-0.198
Height	-0.288	0.108
Diameter	-0.307	-0.467
% variation	54.8	16.0

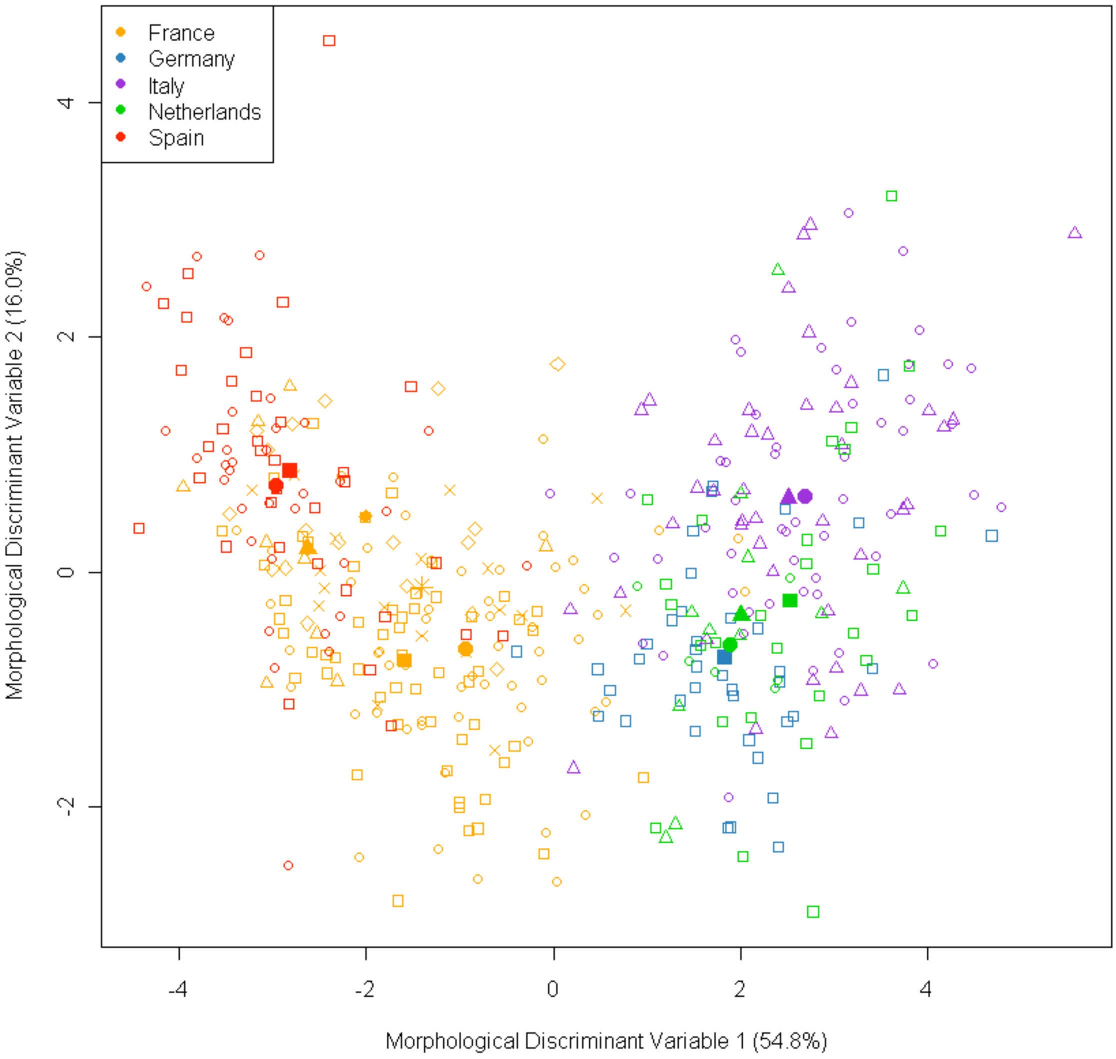


Figure 2.2. Discriminant classification analysis of 12 traits reveals geographic structuring to phenotypic variation in *Populus nigra*. Open symbols represent a single genet, with color corresponding to country of origin. Symbols sharing colors represent different populations within each country. Closed symbols represent population means.

2.4.2 Patterns of neutral genetic structure in *P. nigra*

All populations assessed for microsatellite variation displayed high levels of polymorphism and moderate levels of allelic diversity and heterozygosity, with no evidence of allele fixation (Table 2.7, Appendix 2). A null allele was detected for the majority of populations at one locus, PMGC_2525; this locus was omitted from further analyses. In addition, null alleles were detected in three populations for a second locus (PMGC_2088) and at low frequency in a single population for WPMS_14 and WPMS_20 (Appendix 2). Low rates of mismatches between repeated samples resulted in low error rates: 3.3% per allele or 4.6% per reaction.

An excess of heterozygosity compared to that expected at mutation-drift equilibrium was observed for populations 9 ($P<0.01$), 13 and 14 ($P<0.05$), indicating these populations have undergone a recent population bottleneck. Results for populations 13 and 14 may be influenced by the small sample size in this analysis.

Overall genetic differentiation was significant ($\Phi_{PT} = 0.120$, $P<0.001$), indicating gene flow among populations of *P. nigra* is restricted across Western Europe.

Table 2.5 Genetic diversity observed at nine microsatellite loci in 12 populations of *Populus nigra*.

Population	N	A^{*}	$A_{e}^{\dagger}$	$H_{o}^{\ddagger}$	$H_{e}^{\S}$	$F^{\wedge}$
1	30.2	9.2	4.9	0.774	0.770	-0.006
2	42.4	11.2	6.0	0.799	0.808	0.012
3	5.9	5.4	4.0	0.727	0.708	-0.057
5	6.0	5.9	4.2	0.796	0.736	-0.084
6	12.1	7.0	5.0	0.755	0.751	-0.012
7	42.0	9.0	4.3	0.767	0.743	-0.026
8	39.0	10.7	5.6	0.787	0.797	0.019
9	30.4	9.6	6.1	0.817	0.802	-0.021
11	22.4	8.2	4.8	0.813	0.777	-0.049
12	5.0	5.2	4.2	0.867	0.749	-0.168
13	8.7	5.0	3.5	0.735	0.693	-0.054
14	15.3	6.1	4.2	0.640	0.719	0.071
15	27.6	8.4	4.8	0.734	0.738	0.005
Overall	22.1	7.8	4.7	0.770	0.753	-0.029

[*] Mean alleles per locus

[†] Effective number of alleles per locus

[‡] Observed heterozygosity

[§] Expected heterozygosity

[ˆ] Fixation index

Nine genetic clusters were identified in the *P. nigra* microsatellite data set (Figure 2.3). While smaller than the number of populations sampled for these data, this number may be large for biological interpretation in the context of a wind-dispersed, obligate outcrossing forest tree species. Patterns of genotype assignment to each of the genetic clusters reveals geographic patterning, indicating the genetic groups may correspond to broad gene pools resulting from restricted gene flow among river systems across the landscape (Figure 2.4). In general, samples from Spain are distinct from those from the large-leaf populations. The larger-leaf samples display further genetic structuring between countries (Germany, Italy and the Netherlands).

The observed value of K is not strictly consistent with the genetic structure being solely determined from admixture between ancestral gene pools, which would predict a K of 2 or 3 and samples in the admixture zone (France) partitioned between the genetic clusters. The observation of distinct genetic clusters in the French populations indicates these samples to be differentiated from either historic gene pool, though some French populations display genetic similarity to the Spanish genotypes, and some with the Italian genotypes (Figure 2.4). The level of admixture in individual genets (trees) differed significantly across populations ($F_{12,295} = 2.23$, P=0.011), with trees from Span and Italy, and to a lesser extent the Netherlands, displaying less admixture (individuals assigned to a fewer number of genetic clusters) than the French collections (Figure 2.5).

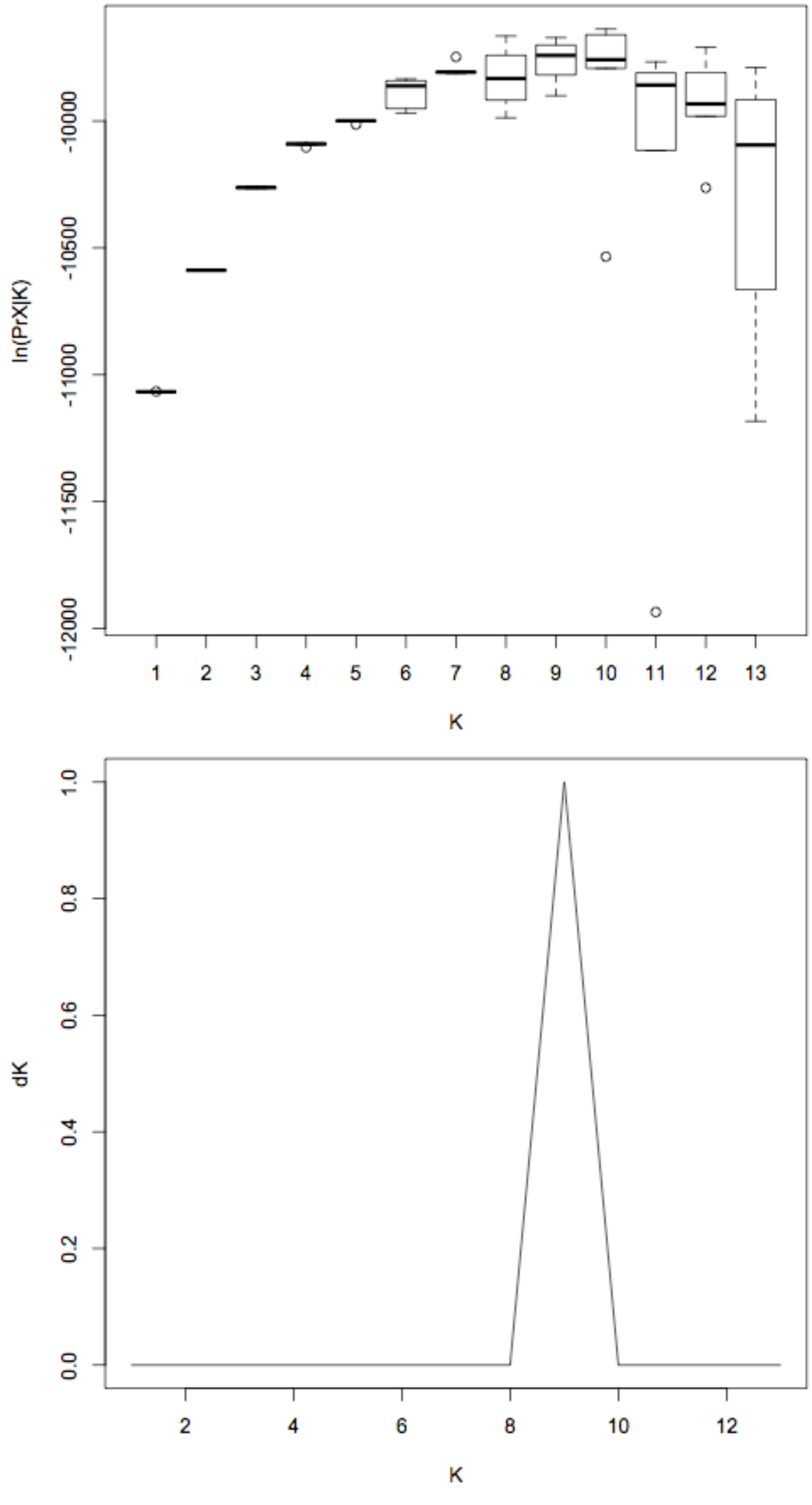


Figure 2.3. (A) Distribution of ln-likelihood values for each number of genetic clusters, *K*, simulated over five replications using the MCMC methods implemented by the software Structure (Pritchard et al. 2000). (B) Estimation of the most likely number of genetic clusters from the test statistic *dK*,as described by Pritchard *et al.* (2000).

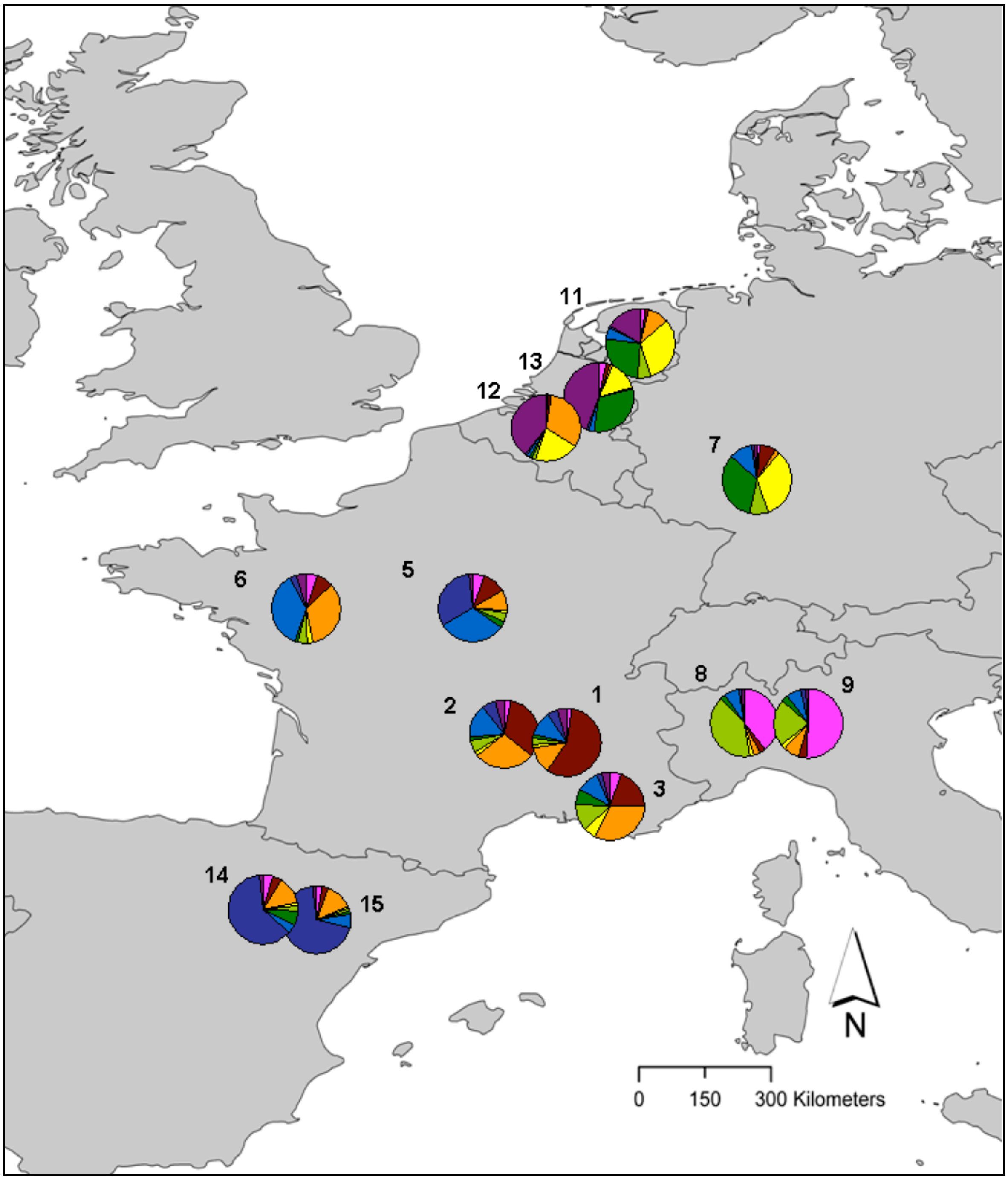


Figure 2.4 Admixture analysis reveals geographic patterning in the proportion of genotypes in each population assigned to each anonymous genetic cluster (colored pie wedges). Placement of pie charts corresponds to the location of the sampled population of *P. nigra*, which are labeled according to Table 2.1. Colors assigned to each genetic cluster (wedge) are consistent across pie charts.

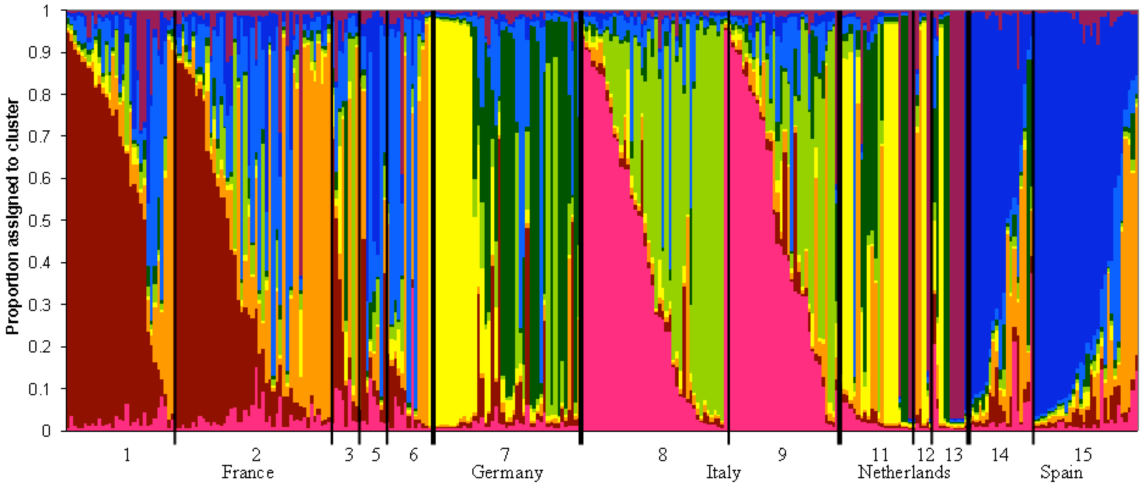


Figure 2.5. *P. nigra* samples display variable levels of admixture. Each bar represents one genet (sample), with colors corresponding to genetic clusters. Colors are consistent with Fig. 2.6.

The historic relationship among populations described using a Neighbor-Joining phenogram is consistent with the genetic differentiation observed in the admixture test (Figure 2.6). Samples from Germany and the Netherlands are distinct from Italy, but these large-leaf trees are distant from the small-leaf Spanish samples. Populations from France are intermediate, with no significant differentiation resolved in the bootstrap consensus phenogram. Only one population from central France (no. 5) consistently clustered with the Spanish populations. This pattern is consistent with the samples from France constituting an admixture zone, as yet undifferentiated among local populations, though more closely related to the Spanish small-leaf samples than the large-leaf gene pools.

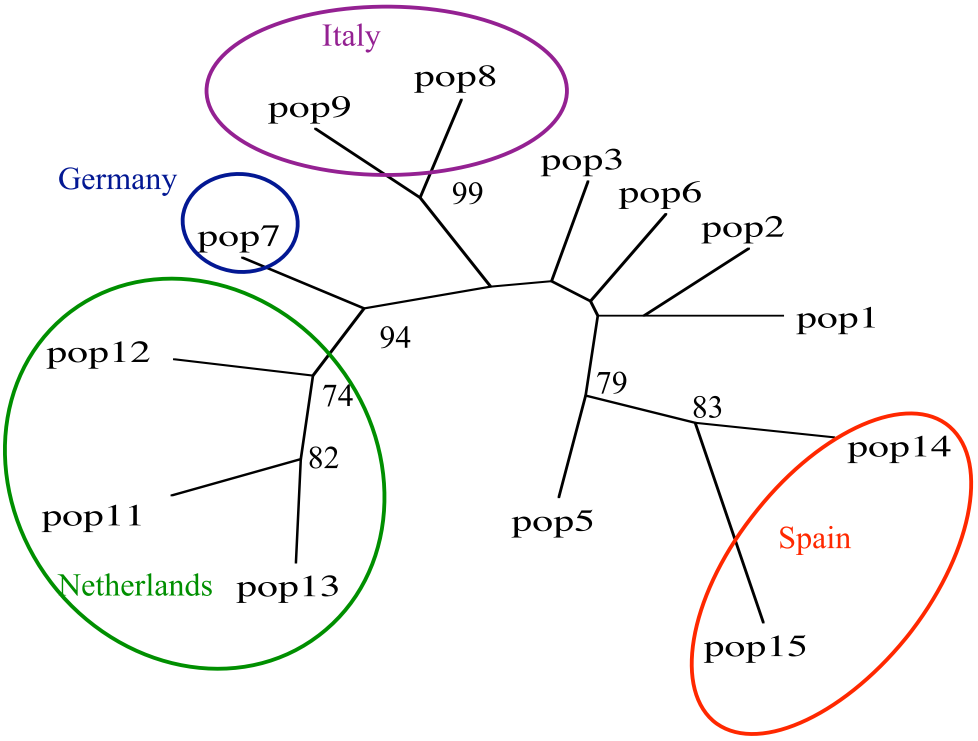


Figure 2.6. *P. nigra* populations group into distinct clusters in a Neighbor-Joining phenogram. Population numbers follow Table 2.1. Numbers represent the percent the node was represented in 1000 bootstrap replicates.

Finally, the principal coordinate analysis (PCoA) reduced the genetic differences among populations to two dimensions, which can then be compared to the pattern of morphological differences among populations (Figure 2.7). The overall differences displayed by the PCoA were of a smaller magnitude than those observed for the morphological differences, as the first two components explained 23.9% and 20.1% of the variation in the data set, for a cumulative total of 44.0%. The pattern of genetic structure was concordant with that observed in the admixture and phenogram analyses: the large-leaf populations from Germany, Italy and the Netherlands were distinct from the Spanish samples, with the samples from France intermediate. Greater overlap was observed between the French populations and the large-leaf samples in the genetic data than in the morphological data. As in the other analyses, the large-leaf populations were not genetically homogeneous, and the Italian and Dutch populations displayed greater genetic similarity to the French samples than did the German samples.

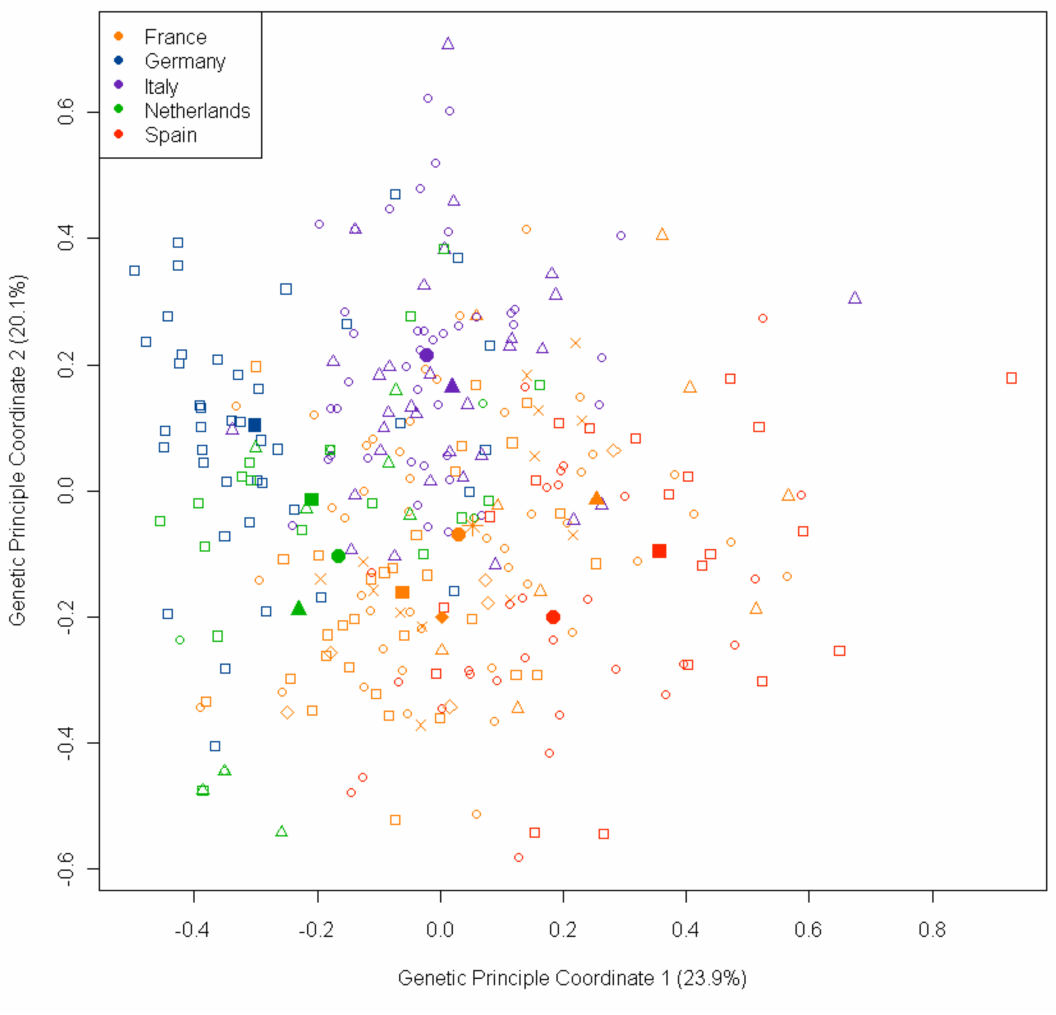


Figure 2.7 Principal coordinates analysis of microsatellite data reveals geographic patterning of differentiation among populations of *Populus nigra*. Open symbols represent individual genets, with symbols representing different populations. Closed symbols represent population means.

2.4.3 Climatic differences among sampling sites

The sites sampled for *Populus nigra* varied in temperature, precipitation and solstice day length (Table 2.6). Several climatic variables were significantly correlated (Table 2.7). The factor analysis of the climatic variables revealed that all nine measures loaded onto the first two principal components (Table 2.8). These two components cumulatively accounted for 84.9% of the variance in the measures. MaxTemp and MinPpt are uncorrelated measures which differentiate between local climates at the sampling sites.

Differences in MaxTemp and MinPpt between sampling sites indicate that most French populations experience a climate more similar to that of the large-leaf populations than the Spanish populations (Figure 2.8). One population, no. 3 from the Durance River in southern France, is similar to the Spanish sites in temperature and climate.

Table 2.6 Environmental conditions vary across 13 sites sampled for *Populus nigra*. Four measures of precipitation (mean annual precipitation, MAPpt; precipitation seasonality, VarPpt; precipitation of the wettest month, MaxPpt; precipitation of the driest month, MinPpt; all in mm), four measures of temperature (mean annual temperature, MATemp; seasonality of temperature, SDTemp; maximum temperature of warmest month, MaxTemp; minimum temperature of coldest month, MinTemp; all in °C), and the day length on summer solstice (MaxDay, in hours) were assessed for each site.

Pop. No.	**MAPpt (mm)**	**VarPpt**	**MaxPpt (mm)**	**MinPpt (mm)**	**MATemp (°C)**	**SDTemp**	**MaxTemp (°C)**	**MinTemp (°C)**	**MaxD (h)**
1	890.0	15.0	83.0	48.0	10.3	6259	25.7	-2.4	15.
2	840.0	20.0	95.0	41.0	12.4	6357	28.1	0.0	15.
3	688.1	26.4	84.7	23.6	12.8	5945	27.4	0.6	15.
5	706.8	11.3	72.0	47.7	11.0	5708	24.8	-0.1	16.
6	710.6	17.5	77.5	44.3	11.6	5145	24.4	1.6	16.
7	589.9	19.0	65.2	36.1	9.9	6525	24.5	-1.8	16.
8	982.0	23.0	122.0	55.0	13.0	7248	29.0	-1.0	15.
9	966.0	23.0	121.0	55.0	13.0	7296	29.0	-0.9	15.
11	774.2	15.3	76.8	47.4	9.3	5451	21.4	-0.7	16.
12	801.8	14.8	81.2	50.3	9.7	5359	21.1	0.2	16.
13	790.6	14.3	77.1	46.9	9.5	53839	21.4	-0.3	16.
14	439.0	26.0	56.0	20.0	14.1	6097	29.7	1.8	15.
15	365.0	31.0	53.0	17.0	13.7	6243	29.5	1.3	15.

Table 2.7 Measures of environmental variation from 13 *Populus nigra* sites are highly correlated, but MinPpt and MaxTemp provide uncorrelated measures which distinguish differences in climate at sampling sites (bold). Correlations between measures are provided above the diagonal; P-values from a one-tailed test of significance are provided below the diagonal. Abbreviations of data follow Table 2.6.

	MAPpt	**VarPpt**	**MaxPpt**	**MinPpt**	**MATemp**	**SDTemp**	**MaxTemp**	**MinTemp**	**MaxDay**
MAPpt		-0.462	0.878	0.881	-0.291	0.270	-0.185	-0.604	0.298
VarPpt	0.048		-0.028	-0.716	0.803	0.479	0.745	0.429	-0.653
MaxPpt	<0.001	0.462		0.665	0.131	0.582	0.205	-0.411	0.003
MinPpt	<0.001	0.002	0.005		-0.503	0.074	**-0.416**	-0.584	0.513
MATemp	0.156	<0.001	0.328	0.033		0.524	0.939	0.515	-0.885
SDTemp	0.175	0.041	0.014	0.401	0.027		0.724	-0.405	-0.490
MaxTemp	0.263	0.001	0.241	**0.070**	<0.001	0.002		0.217	-0.925
MinTemp	0.011	0.063	0.072	0.014	0.030	0.076	0.228		-0.307
MaxDay	0.150	0.006	0.496	0.030	<0.001	0.038	<0.001	0.143	

Table 2.8 Factor analysis of climate data for 13 *Populus nigra* sites indicates the variation of nine measures reduces to two principal components. Variable abbreviations follow Table 2.6.

	Component	
Variable	1	2
MAPpt	-0.616	0.730
VarPpt	0.887	0.120
MaxPpt	-0.226	0.905
MinPpt	-0.797	0.510
MATemp	0.914	0.309
SDTemp	0.374	0.837
MaxTemp	0.856	0.484
MinTemp	0.573	-0.522
MaxDay	-0.862	-0.282
% Var.	51.5	33.4
Cum. %	51.5	84.9

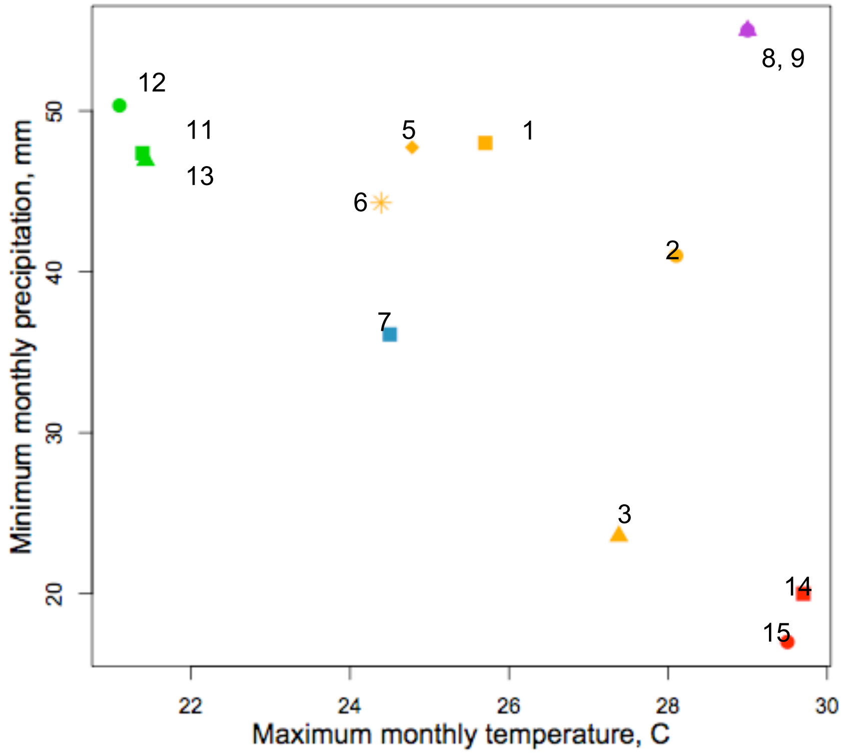


Figure 2.8 Variation in climate at *Populus nigra* sampling locations. Points are labelled by population number following Table 2.1 and coloured according to country of origin: France (orange); Germany (blue); Italy (purple); The Netherlands (green); Spain (red).

2.4.4 Tests for isolation by distance, isolation by adaptation, and correlation between morphological variation and climate

Simple and partial Mantel tests were conducted to test for correlation between genetic, geographic, morphological, and climatic differences between populations of *P. nigra* (Table 2.9). These tests were conducted on four distance matrices summarizing the data presented above.

Significant correlation between genetic and geographic distance independent of morphological divergence (Gen, Geo|Morpho) indicates IBD influences the genetic structure of *P. nigra* in Western Europe. Genetic differences increased as a function of distance, consistent with the genetic pattern resolved by the admixture and phenogram analyses. Tests using pairwise F_{ST}, (Rousset 1997) measure of genetic difference, and log-transformed data were concordant with these results (data not presented).

Isolation by adaptation (IBA) was also found to contribute to the genetic structure of populations, with genetic similarity increasing with phenotypic similarity, even when controlling for geographic structure among populations (partial Mantel tests, Table 2.9). Analyses of log-transformed data were consistent (data not presented). This result indicates that genetic differentiation at microsatellite loci is influenced by reduced gene flow between morphologically distinct populations, independent of geographic distance between sites. Significant correlation between morphological and climatic differences was observed (Morpho, Clim), as would be expected if adaptation is driving the divergence of phenotype. However, the climatic differences between large-leaf and small-leaf populations is not consistent with the bimodal distribution of morphology.

Table 2.9. Correlations between genetic, geographic, morphological, and climatic differences among 13 populations of *P. nigra* were tested with simple and partial Mantel tests.

Comparison	***r***	***P***
Simple Mantel tests		
Genetic, Geographic	0.605	<0.001
Morphological, Genetic	0.482	<0.001
Morphological, Climate	0.416	0.002
Partial Mantel tests		
Genetic, Geographic| Morphological (IBD)	0.456	<0.001
Morphological, Genetic|Geographic (IBA)	0.201	0.03

While IBD explains a significant portion of the genetic structure of this collection of *P. nigra*, partial regression plots of the genetic and morphological differences between populations when controlling for geographic distance allow further dissection of the significant IBA (Figure 2.9). Linear regression over all pairs of populations identified a significant positive relationship between morphological differences and genetic divergence when controlling for geographic distance ($F_{1,76}$=28.1, P<0.001; slope = 0.038, 95% C.I. 0.023 to 0.052). Categorical comparisons of small-leaf and large-leaf pairs revealed the trend to be driven by differences among the small-leaf populations. Significant positive relationships were observed for small-leaf to small-leaf pairs ($F_{1,19}$=17.7, P<0.001; slope = 0.07, 95% C.I. 0.035 to 0.105), and small-leaf to large-leaf pairs ($F_{1,40}$=26.5, P<0.001; slope = 0.053, 95% C.I. 0.032 to 0.074), but not large-leaf to large-leaf comparisons ($F_{1,13}$=3.52, P=0.083; slope = 0.02, 95% C.I. -0.003 to 0.043), though the number of pairs for the last analysis was small. This trend indicates the genetic and morphological variation observed in the large-leaf populations is sufficiently explained by isolation by distance, with gene flow decreasing as a function of distance. The genetic and morphological structure in small-leaf populations, and their relationship to the large-leaf phenotypes, however, results from isolation by adaptation as well. Morphologically differentiated populations are more genetically dissimilar than their geographic distance would predict.

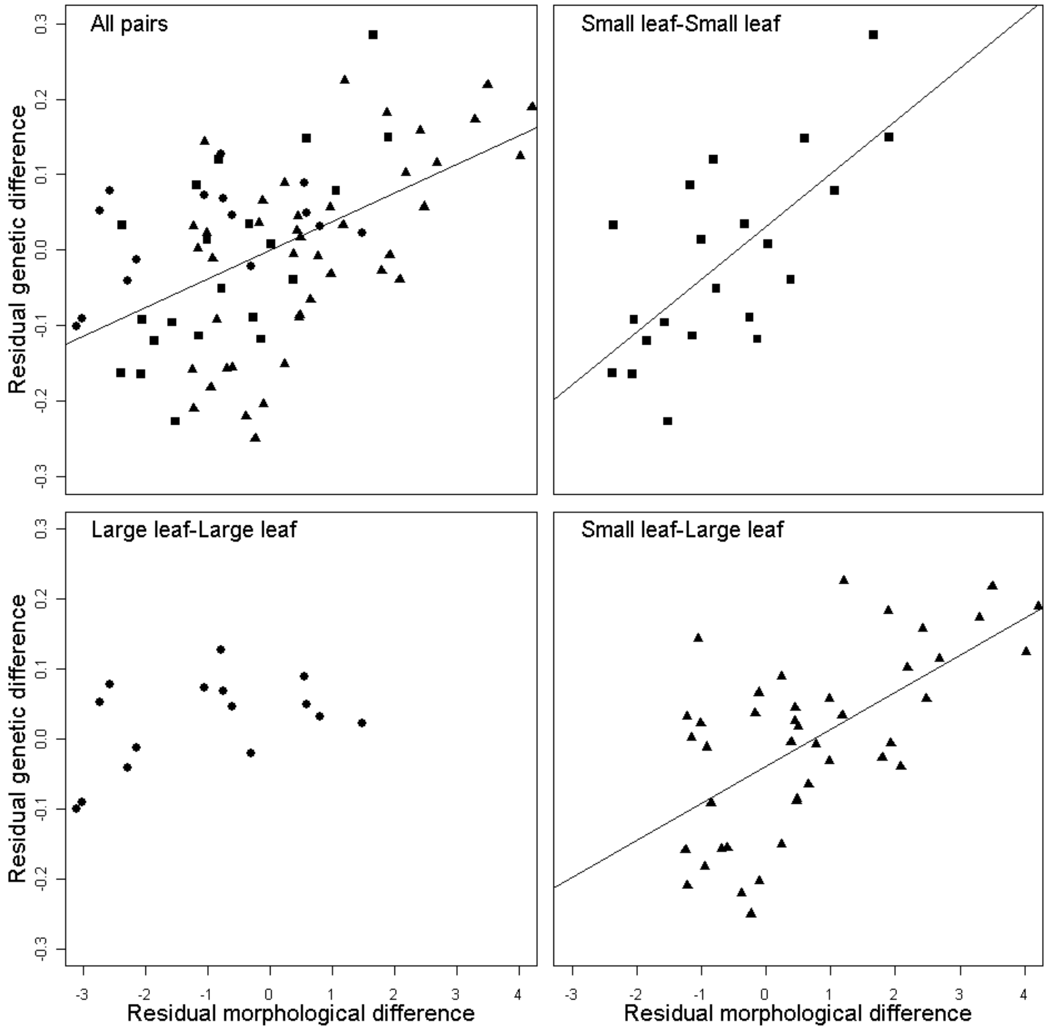


Figure 2.9. Partial regression plots illustrating correlated genetic and morphological differences between populations of *P. nigra* after controlling for the effects of geographic separation, indicating populations are isolated by adaptation (IBA). Points represent the residuals of the regression of morphological (x-axes) and genetic (y-axes) distances against geographic distance. Lines represent significant regressions of the residuals. Comparisons of large-leaf to large-leaf populations revealed no significant relationship between residuals.

If the morphological differentiation observed in *P. nigra* represents adaptation to local climate, the significant correlation between morphological and climatic differences detected by the Mantel test (Table 2.9) should reflect low climatic differences between populations of like phenotype (small-leaf to small-leaf and large-leaf to large-leaf), and greater climatic differences between pairs of contrasting phenotypes (small-leaf to large-leaf). Pairwise plots of morphological and climatic differences (not the residuals) indicate that comparisons of population pairs displaying similar phenotypes experience climatic conditions as diverse as many of the pairs with contrasting phenotypes (Figure 2.10). Although nonparametric correlations between the first morphological factor and the minimum monthly precipitation reveal a positive correlation over population means (Kendall's τ=0.581, P=0.006), most of the populations from France experience precipitation comparable to the large-leaf populations (Figure 2.10), yet are intermediate to the Spanish and large-leaf phenotype. This pattern is consistent with contemporary evolution driving populations from central France from the ancestral small-leaf morphology toward the large-leaf phenotype.

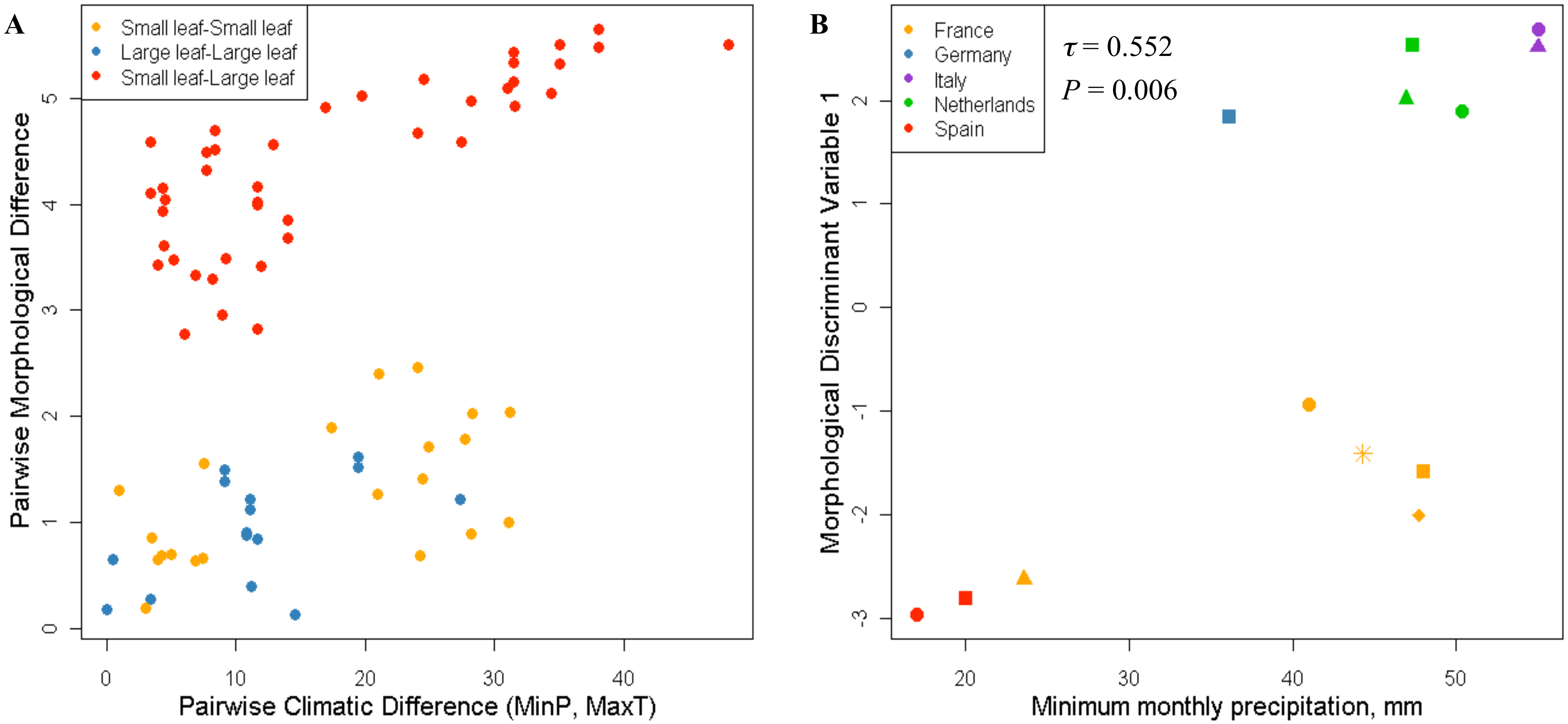


Figure 2.10 Pairwise differences in morphology are not concordant with the climate experienced by populations of *P. nigra*. (A) Pairs of populations with contrasting phenotypes (calculated as the difference of mean population scores for the first discriminant factor, broadly distinguishing large-leaf and small-leaf phenotypes) experience climates as similar as pairs of populations with similar morphologies. (B) *P. nigra* from central France display morphologies more similar to the small-leaf Spanish populations but experience minimum precipitation equivalent to the large-leaf populations. No significant correlation was observed between morphology and maximum temperature.

2.5 Discussion

2.5.1 Historic vicariance, genetic differentiation, and admixture in *P. nigra*

This is the first study to indicate that historic vicariance influences the extant morphological variation in a forest tree, independent of geographic distance between populations. The pattern of genetic, morphological, and climatic differences between populations supports the hypothesis that *P. nigra* trees from France display an ancestral small-leaf phenotype inconsistent with the climatic conditions (minimum precipitation and maximum temperature) experienced at their sites of origin.

These conclusions are founded in part on the findings of Cottrell et al. (2005), who showed that variation in chloroplast DNA fragments corresponded with three glacial refugia during the most recent glacial maximum: the Iberian Peninsula, the Italian Peninsula, and a region east of Italy, likely in the Balkan Peninsula. In general, haplotype diversity was greater in the southern areas than in the northern populations, consistent with the recolonization of northern populations from southern refugia (Cottrell et al. 2005). In addition, a greater portion of samples from France displayed haplotypes more similar to the Spanish samples than either of the other refugia, indicating that populations in France were likely derived from seed flow from the Iberian refugia.

The microsatellite data presented here revealed a similar pattern of genetic differences between populations as reported for cpDNA markers, though the magnitude of differentiation was lower, as would be expected for biparentally inherited nuclear loci. The large-leaf populations from Germany, the Netherlands and Italy did not constitute a single homogeneous gene pool, yet all were distinct from the Spanish collection. Trees from France display intermediate genetic characteristics consistent with the populations being more recently derived from (as depicted in the population phenogram) and admixed between (indicated in the PCoA) the other gene pools. The genetic distinctiveness of the French populations in the admixture analysis may result from the genetic bottleneck or founder effect that likely occurred as *P. nigra* recolonized France following the most recent glacial maximum (Smulders et al. 2008b). This severe form of genetic drift may have resulted in allele frequencies at nuclear loci in recolonized populations being significantly different from those in the source gene pool. The pattern of differentiation observed in this study is consistent with a previous study of nuclear markers in naturally-occurring *P. nigra* (Smulders et al. 2008b). While

Smulders et al. (2008b) sampled a larger range of *P. nigra* natural populations, only the Drôme river system was sampled from France, preventing a comparison of the relative admixture across the landscape.

Anthropogenic factors may have contributed to the genetic structure of *P. nigra*, particularly in the genetic bottlenecks identified in populations 9 and 14. A previous study reported the depletion of the natural poplar stands in the Ticino river region in northern Italy (pops. 8 and 9) during the Second World War (Fossati et al. 2003). This episode of deforestation may have contributed to the genetic bottleneck identified from the microsatellite data.

2.5.2 Geographic isolation influences both genetic and morphological differentiation in *P. nigra*

The most significant process influencing the extant genetic structure of *P. nigra* was isolation by distance (IBD), indicating gene flow is limited across the landscape. These results are consistent with previous studies of population structure in similar collections. The overall level of genetic differentiation observed in *P. nigra* (Φ_{PT} = 0.125 in this study; F_{ST} = 0.081 for microsatellites in Smulders et al. (2008b); F_{ST} = 0.113 in Storme et al. (2004) is higher than that reported in other forest trees (Mediterranean pine F_{ST} = 0.048 in Gonzalez-Martinez et al. (2002); sessile oak F_{ST} = 0.024 in LeCorre et al. (1997a); Douglas-fir F_{ST} = 0.004 in Krutovsky et al. (2009). Three characteristics of *P. nigra* make it distinct from the canonical forest tree. First, *P. nigra* is restricted to river floodplain habitat (Smulders et al. 2008b). While pollen and seed may be wind-dispersed, appropriate habitat in the interriver zones may be rare enough to restrict dispersal between river systems, thus increasing the allele frequency variation between populations. Second, *P. nigra*, like all *Populus* species, is dioecious, with male and female flowers occurring on different plants. This mating system is expected to increase the genetic variation within and decrease differentiation among populations (Loveless and Hamrick 1984), which is counter to the significant differentiation observed here. However, dioecy may result in increased genetic drift due to separate male and female gene pools, especially if sex ratios are unequal. In subdivided populations, such as *P. nigra* among river systems, subdivision will decrease the effective size of the population, especially if fitness varies among subpopulations (Wang and Caballero 1999). Thus, the significant differentiation observed indicates the obligate outcrossing maintained by the dioecious mating system in *P. nigra* is counteracted by the effects of restricted gene flow and variable reproductive rates among populations, making dioecy insufficient to prevent differentiation across river systems. In addition, the dioecious mating system likely drives the lack of significant fixation (inbreeding) in each population. Third, *P. nigra* is an early-successional, fast-growing, colonizing species adapted to floodplain habitats having high frequency of disturbance events (Imbert and LeFevre 2003). The canonical forest tree is typically monecious, occurring in a climax community, with zoochorous seed dispersal. Together, these demographic characteristics likely contribute to the significant differentiation and isolation by distance observed among populations of *P. nigra*.

The significant correlation between genetic and geographic distance (IBD) indicates that genetic isolation may explain some portion of the morphological differentiation observed between populations. Even when controlling for morphological similarity (i.e. adaptation) among populations, gene flow decreases as a function of distance. Thus, populations that share common environmental and thus selective pressures may diverge both genetically and morphologically due to the neutral process of random genetic drift. Limited gene flow across the landscape may reduce the ability of *P. nigra* to migrate to suitable habitat or adapt to a changing environment over the coming century (discussed below).

2.5.3 Isolation by adaptation in a forest tree

Geographic distance alone does not account for genetic differences between pairs of *P. nigra* populations. Morphologic and genetic differences between pairs of populations are significantly correlated, independent of geographic separation, indicating that *P. nigra* ecotypes are isolated by adaptation (IBA). Two mechanisms may account for this pattern of genetic and morphological differentiation. First, the correlation may result from natural selection, i.e., the pattern reflects the canonical IBA where genes track diversifying selection across landscape in small-leaf and large-leaf phenotypes. Alternately, the correlation may be residual IBA from the historic vicariance event, leaving an “imprint” (González-Martínez et al. 2004) of historic differentiation in both morphological and genetic structure. Given the discordance between morphology and climate observed, residual IBA is the more likely mechanism. That is, the morphological and genetic structure of *P. nigra* tracks the historic vicariance between glacial refugia. The partial regression plots further dissect the relationship between morphological and genetic differences while controlling for geographic distance between populations. No significant relationship was observed between the residuals for large-leaf to large-leaf comparisons, indicating that genes do not vary with morphology in the samples from Italy, Germany, and the Netherlands. Both partial regressions involving the small-leaf populations were significant, however, indicating that genes vary with morphology in both small-leaf to small-leaf and small-leaf to large-leaf comparisons. The bimodal distribution of extant phenotypes indicates that admixture between two (or three) historic gene pools has not been sufficient to normalize the distribution of quantitative traits among populations. The significant correlation between individual traits (Trewin et al., in preparation) may be due to co-adapted gene complexes that control early stages of development (biomass and leaf production) or have a role in the basic cellular function (Edmands et al. 2009). Putative genetic association between biomass traits and a gene involved in leaf initiation, *ASYMMETRIC LEAVES1* has been reported in this collection (Trewin 2008). Additional work is underway to identify additional genetic mechanisms underlying the phenotypc diversity observed in *P. nigra*.

One other study has reported a correlation between maternal lineage (i.e. historic vicariance) and phenotypic and nuclear differentiation. González-Martínez et al. (2004) described mitochondrial, nuclear and quantitative differentiation in maritime pine, and identified congruent phenotypic and genetic structure as an “imprint” of historic vicariance on the genome. The correlations were not tested in a pairwise manner and did not control for geographic distance, however. While this omission does not necessarily invalidate the evolutionary significance of the findings, the potential for IBD to drive such correlations cannot be discounted.

A study of *Quercus petraea*, however, did consider geographic distance when comparing historic vicariance and extant morphological and genetic structure among populations. Kremer et al. (2002) did not find significant correlation between phenotypic differentiation and maternal lineage when controlling for geographic distance, but did report significant cytoplasmic-nuclear disequilibrium as nuclear and cpDNA differentiation were correlated independent of geographic distance. The authors interpreted the cytoplasmic-nuclear disequilibrium to be the signature of historic vicariance remaining in the neutral markers despite admixture following recolonization of northern Europe. In contrast to the neutral genetic signature, natural selection had erased any phenotypic-cytoplasmic disequilibrium that might have arisen during the most recent glacial maximum (Kremer et al. 2002). Tests of the partial correlation between phenotypic divergence and nuclear differentiation while controlling for geographic distance between populations were not reported, preventing comparison of IBA in these populations.

Here, the correlation between morphological and genetic differences independent of geographic distance provides a more conservative test than cpDNA analyses reported in other studies due to the recombination and higher mutation rate of the nuclear markers. Further, the congruent pattern is driven by variation in the small-leaf ecotypes. The genetic patterning and the climatic differences between small-leaf populations from France are consistent with admixture and contemporary evolution driving phenotypic variation in *P. nigra*.

Analysis of additional populations, especially from the Iberian and Italian Peninsulas, would further resolve the relationship between the admixture and adaptation in *P. nigra*. The Ebro River system may constitute a distinct refuge among other areas of the Iberian Peninsula. A study of a Tertiary relic, *Ramonda myconi*, identified the Ebro River populations as distinctive compared to populations from across the species range (Dubreuil et al. 2008). The pattern of differentiation may vary with the addition of other populations if *P. nigra* varies across the Iberian Peninsula. In addition, trees from southern Italy display reduced leaf size similar to the Spanish populations (Regier et al. 2009), and may further resolve the relationship between genetic and morphological structure among distinct refugia.

2.5.4 Implications for species survival in a changing climate

The morphological divergence described in this study must be considered in context of the climate of the common garden site in Belgium. The growth data reflect genotype performance in a cool, moist climate typical of central Europe. This climate category is predicted to decrease in frequency over the coming century (Blenkinsop and Fowler 2007; Penuelas et al. 2007). A different pattern of growth and survival may be expected if this collection of *P. nigra* was grown in an arid environment more similar to Spain or southern France, where the adaptive benefit of the small-leaf ecotype may be apparent in tree survival and growth. A recent greenhouse experiment testing the growth and morphology of genets from Spain, France, Italy and the Netherlands under moderate drought indicates the small-leaf samples from Spain display a more rapid response to drought stress (Viger et al., in preparation). The large-leaf trees from Italy and the Netherlands displayed divergent response patterns in the drought experiment, though both were more severe than the small-leaf genotypes and included mortality

(Viger et al., in preparation), indicating the genetic differences at neutral loci may correspond to physiological differences among large-leaf trees.

The small-leaf and large-leaf ecotypes described here may broadly correspond to differences in intrinsic growth rates among populations of *P. nigra*. The intrinsic growth rate of an organism is highly adaptive. In plants, rapid growth rate is associated with low stress conditions, or as a response to factors favoring large size, such as the length of the growing season (Arendt 1997). Slow growth may not confer the selective advantage directly, but instead may be a consequence of adaptations maximizing nutrient use (Poorter 1989). If differentiation in traits related to growth (height, stem diameter, SLA) in *P. nigra* is related to adaptation to environmental stressors, the pattern observed should follow the pattern of climatic conditions for each site. The pattern of differences among populations were not congruent, with French trees displaying morphology more similar to Spanish trees while experiencing precipitation and temperature more similar to central Europe. This would imply that the trees in the admixture zone are not optimally adapted to their local climate. It is possible that other biotic (e.g. herbivore) or abiotic (e.g. soil characteristics) not measured here may drive the phenotypic variation among populations. However, given the canonical model of leaf area and growth form being associated with water use efficiency, which involves both water availability (precipitation) and transpiration (affected by temperature) (Chapin et al. 1993), it is reasonable to expect overall morphology to relate to these measures of climate in this floodplain tree.

If the trees display morphology not optimal for growth in the current climate, how might climate change affect population and species survival over the coming decades? While climate models vary in precise predictions, general trends indicate a warmer Europe with more frequent drought episodes in all but the highest latitudes (Blenkinsop and Fowler 2007). If droughts become more frequent, the migration of drought-tolerant genotypes from Spain and the admixture zone (Viger et al., in preparation) may become important to the survival of populations in the currently cool, moist climates (the large-leaf populations). The potential for *P. nigra* to migrate to suitable habitat or adapt to novel conditions may be compromised by the limited movement of genes among river systems and across the landscape.

The two ecotypes described here (large-leaf and small-leaf) each display phenotypic qualities of interest to land managers and tree breeders. The high biomass and fast growth rate of the large-leaf phenotypes from Italy, Germany and the

Netherlands are valued for coppice and breeding strategies in areas of moderate climate. Trees displaying the small-leaf phenotype may be better suited to survive water deficit (Viger et al., in preparation), which may be of interest to land managers and conservation priorities as described above. In addition, the small-leaf trees produce a higher number of sylleptic branches (personal observation), another trait of interest to breeders (Taylor 2002). In the hunt for the ‘ideal poplar,’ the natural admixture zone of the French populations may prove a powerful natural experiment for association tests.

In anticipation of the utility of trees from the admixture region, work is currently underway to investigate growth of an extended collection of *P. nigra* in common garden experiments planted in replicate in cool moist and arid Mediterranean conditions (G. Taylor, personal communication). Parallel work to resequence the *P. nigra* genome will provide a large resource of single nucleotide polymorphisms (SNPs) for use in admixture analyses and association tests (G. Zaina, personal communication). Together, these projects provide a powerful opportunity to unravel the genetic basis of complex morphological traits in this important tree species.

Chapter 3. Phenotypic variation in *Populus nigra* from Central Europe: a glasshouse study

3.0 Overvivew

Populus nigra is an ecologically and economically important riparian forest species endemic to Europe, Asia and northern Africa. Previous assessments of phenotypic variation in this species have focused on populations from western Europe, a relatively small portion of the native range. This study expands the assessment of morphological variation in *P. nigra* by assessing populations from central Europe (Hungary and Bosnia and Herzegovina), representing the third glacial refugium. Analysis of 23 leaf, biomass, and cell traits identified significant morphological differentiation among populations, with most populations from central Europe displaying similar morphology to the large-leaf samples from Italy. One population of the Mediterranean hairy black poplar (*P. nigra* ssp. *caudina*) displayed small leaf size, increased branching, and higher stomatal index than the typical morphologies. Comparisons between *P. nigra* ssp. *caudina* and a small-leaf sample from Spain determined the overall morphologies of these trees to be distinct, despite each having a reduced leaf size. Comparison of morphological and climatic differences between populations indicates the small-leaf phenotypes to be adaptations to distinct climatic variables, while no correlation between morphology and climate were detected for the large-leaf populations. Differences in cell number, and not cell size, appear to drive differences in leaf area, identifying a putative developmental mechanism for the phenotypic diversity observed in this species. Together, these findings inform the potential to move germplasm among populations in central Europe as part of conservation or management efforts.

3.1 Introduction

The morphological variation observed in populations of *P. nigra* collected from Western Europe is likely the consequence of historic vicariance between glacial refugia, ongoing adaptation to local environments, and genetic differentiation due to the restricted gene flow among river systems (Chapter 2). Given the evidence that extant populations have been recolonized from two or three glacial refugia in the Iberian, Italian, and Balkan peninsulas (Cottrell et al. 2005), a complete understanding the evolution of leaf traits in this species requires investigation of populations from the third refugium in central Europe. This study reports an analysis of morphological variation in *P. nigra* sampled from five populations from Hungary and Bosnia and Herzegovina, representing the third, eastern-most glacial refugium proposed by Cottrell et al. (2005).

Leaf morphology varies with climatic variables such as precipitation and temperature (Barboni et al. 2004) due to its role in light interception, water regulation, and interaction with pest species (Bright and Rausher 2008; Park et al. 2008). In order to quantify morphological differentiation between populations of *P. nigra* and the relationship between morphology and climate, traits related to leaf shape and petiole morphology were studied. *Populus nigra* leaves vary not only in size (Chapter 2) but also shape. The base shape varies from cordate (heart-shaped with indentation at the petiole), to deltoid (triangular with a flat base), to rhomboid (less than 180° angle at the petiole, nearing a diamond shaped) in field-grown leaves (Figure 3.1). As this variation in leaf morphology may have adaptive consequences if the shape is related to water use efficiency (the ratio of carbon capture to water loss) or light capture, understanding the level of genetic variance in this trait will be important to management and breeding efforts.

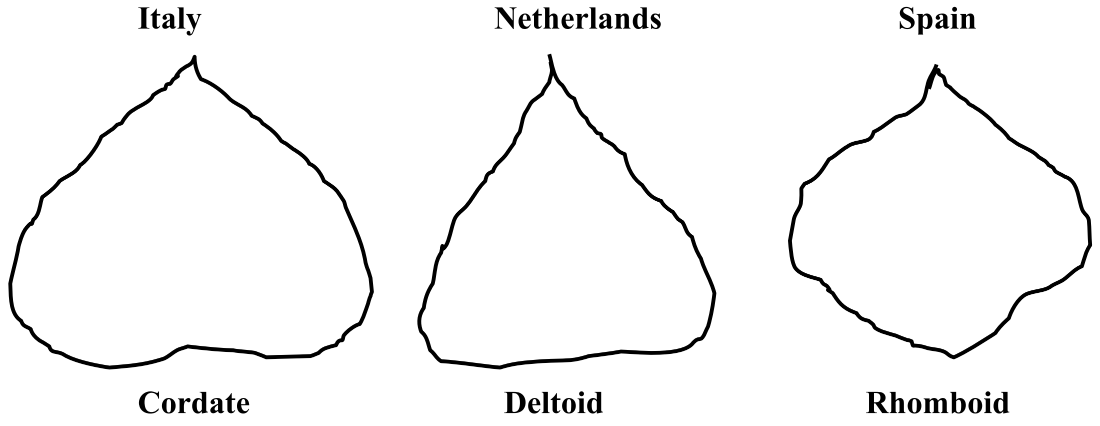


Figure 3.1 The shape of the leaf base varies in *Populus nigra* trees grown in a common garden experiment as described in Chapter 2. Traced from photos collected by H. Trewin (2008). Three leaf shapes can be distinguished: heart-shaped cordate bases, flat deltoid bases, and angular rhomboid bases.

The petiole influences leaf physiology as the organ attaching the leaf blade (the site of photosynthesis and gas exchange) to the plant stem. Petiole morphology determines spatial patterning of leaves and provides hydraulic connection of the leaf to the stem sap (Poorter and Rozendaal 2008). Petioles may display phenotypic plasticity in response to abiotic stresses such as flooding, providing a potential adaptive advantage (Colmer and Voesenek 2009). The interaction of internode spacing, petiole length and leaf size may also contribute to community interactions across trophic levels as predators forage on phytophagous insects (Park et al. 2008). In addition, some of the same genes involved in leaf development have been shown to influence petiole morphology (Theodoris et al. 2003).

A previous assessment of the evolution of morphological differentiation in *P. nigra* from Western Europe found the morphological and genetic structure of populations in France to be consistent with the region representing an admix zone between the glacial refugia in the Iberian Peninsula the Italian and Balkan refugia (Chapter 2). The large-leaf populations from Central Europe displayed significant genetic differentiation, indicating the large-leaf phenotype does not represent a homogeneous gene pool (Chapter 2). Here, the phenotypic comparison of large-leaf samples from Northern Italy with samples from Hungary and Bosnia & Herzegovina is the first study to compare the morphological differentiation of populations from the Italian and Balkan refugia.

In addition, one population from Bosnia & Herzegovina consists of *P. nigra* ssp. *caudina*, a subspecies distinct to the common European black poplar that forms the basis of the large association collection. *P. nigra* ssp. *caudina*, also known as Mediterranean hairy European black poplar, is thought to be a glacial relict displaying morphology predicted to convey an adaptive advantage in Mediterranean climates characterized by extended hot, arid summers (Kajba et al. 2004). Once thought to have occupied a continuous range from Greece to Herzegovinia, the distribution is now fragmented, and *P. nigra* ssp. *caudina* inhabits a small number of isolated populations in the Balkan region, including Greece, Macedonia, Herzegovina, and reportedly Albania and Turkey (Kajba et al. 2004). Populations of the two subspecies are allopatric, with neither co-occurrence nor hybridization of the two types being reported. Phenotypically, the *caudina* subspecies is distinguished by dense pubescence on the petioles, young stems, and reproductive structures of the trees, as well as an overall smaller stature, smaller leaf blade, and sparser crown when compared to the typial *nigra* subspecies (Kajba et al. 2004). The pubescence is the distinguishing feature in this subspecies. A morphological study of field-collected leaves was conducted with samples of one population of *P. nigra* ssp. *caudina* from Bosnia and Herzegovina and two populations of *P. nigra* ssp. *nigra* from adjacent Croatia revealed overlapping morphology in the two subspecies in five leaf traits unrelated to pubescence (Kajba et al. 2004). Understanding the morphological variation of *P. nigra* ssp. *caudina* compared to other populations of the common subspecies and to other populations displaying adaptation to arid conditions (e.g. Spain) will improve estimates of the ability of this forest tree to adapt to climate changes predicted for the next century.

This study of *P. nigra* from western and central Europe, including the Mediterranean hairy black poplar, addressed three questions. First, what is the pattern of phenotypic divergence between populations from France, Italy, Bosnia and Herzegovina, and Hungary? If the large-leaf phenotype is derived from the eastern (Balkan) refugium, the Central Europe populations should display low interpopulation variation, and the Italian population should be more similar to the central collection than the French samples. Second, how do the phenotypes of the small-leaf sample from Spain and the *P. nigra* ssp. *caudina* differ? If the two populations display convergent evolution consistent with adaptation to similar climatic conditions, measures of leaf area, growth rate, and cell patterning should be consistent among samples. Third, does phenotype vary with climate in this collection? If the small-leaf and hairy phenotypes area adaptive to arid climates, and the large-leaf phenotype is adaptive to cooler, moister climates, morphological diversity should vary with climatic differences among populations.

3.2 Methods

3.2.1 Phenotypic variance in *P. nigra* from Central Europe and the Balkan Peninsula

Samples of *P. nigra* from three countries (Spain, France and Italy) were collected for a previous common garden study (Trewin 2008). Samples from two populations of *P. nigra* from Hungary were generously shared by Mr. Gábor Szulcsán (KEFAG, Inc., pop. 24) and Mr. Sándor Bényi (pop. 25) through collaboration with Dr. Sándor Bordács (Central Agricultural Office, Department of Forest and Biomass Reproductive Material, Hungary). Samples from Bosnia & Herzegovina (including *P. nigra* ssp. *caudina*) were generously shared by Dr. Dalibor Balian (University of Sarajevo, Faculty of Forestry, Bosnia and Herzegovina).

Three cuttings (replicates or ramets) of five genets (genotypes) from each of seven populations from France, Italy, Bosnia & Herzegovina and Hungary (Table 3.1, Figure 3.2) were grown in a three block replicated design in glasshouse conditions. The position of each genet was randomized within each block on the same center bench of the glasshouse, with one row of guard trees grown between blocks and around the entire design. The samples from France and Italy were previously described in a common garden study and re-analyzed in Chapter 2. A single sample from Spain was included for comparison to the small-leaf phenotype, but was omitted from statistical analyses. Population 17 from Bosnia & Herzegovina is *P. nigra* ssp. *caudina* (the hairy ecotype). One month prior to the start of the experiment each replicate was pruned to a single main stem and staked for support. Glasshouse conditions provided supplemental light from 14 hours of the day (06:00 to 20:00). Vents and fans provided cooling, resulting in a mean temperature during the study period of 20.2°C (range 12.6 to 38.5°C). Plants were watered at least once daily and fertilized intermittently.

In order to describe overall phenotypic variation, measures of leaf, cell, and biomass traits were made for a 1-month growth period. On the first day of the experiment (13 July 2009), the first completely unfurled leaf on each stem was marked with a piece of thin string to indicate the youngest leaf on the tree. The diameter (mm) of the main stem 2 cm above its branch-point from the stump was measured with digiMax digital callipers. The height of each stem (cm) was recorded on 15 July 2009.

Table 3.1 Sampling details of *Populus nigra* samples from Hungary and Bosnia & Herzegovina that were compared to genets from France, Italy, and Spain previously assessed as part of a larger study (Chapter 2).

Pop. No.[*]	**Country**	**Population**	**Latitude**[+]	**Longitude**	***N***
2	France	Drôme 6	44.7500	4.9167	5
8	Italy	La Zelata	45.2667	8.9833	5
15	Spain	Ebro 2	41.5833	-1.0000	1
17[‡]	Bosnia & Herzegovina	Capljina	43.1447	17.7297	5
21	Bosnia & Herzegovina	Tegare	44.1147	19.4889	5
23	Bosnia & Herzegovina	Banja Luka	44.7692	17.5014	5
24	Hungary	Duna-Tisza interriver	46.4011	19.3351	5
25	Hungary	Fadd	46.4657	18.8410	5

[*] Population numbers correspond with those presented in Table 2.1 and therefore are not sequential.

[+] Latitude and longitude are provided in dd.dddd.

[‡] *P. nigra* ssp. *caudina*, the Mediterranean hairy black poplar

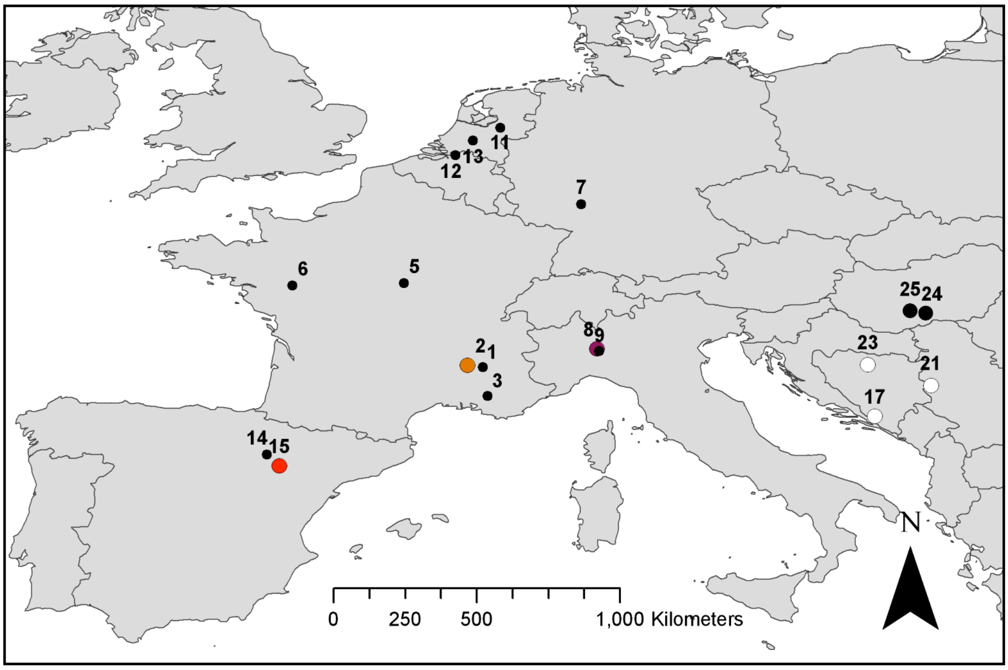


Figure 3.2 Populations from Spain (red circle), France (orange), Italy (purple), Hungary (large black) and Bosnia & Herzegovina (white) were analyzed in a glasshouse experiment. Population numbers follow Table 3.1. Additional populations described in Chapter 2 are marked with small black points.

On the last day of the experiment (8 August 2009), overall plant morphology was assessed through a total of 23 leaf, biomass and cell traits. M. Goolsby assisted with data collection on this day. The mature leaf marked on the first day of the experiment was removed from the stem and petiole dissected from the leaf blade. Sampling a single leaf per ramet did not allow for intra-plant variation in leaf size to be assessed, but minimized the environmental effects on leaf size by assessing only leaves that developed under the same conditions (temperature, day length, etc.). The petiole length, basal width (at the point of attachment to the stem), and height (at point of attachment to the leaf blade) were measured in mm. This leaf was photographed on a 5 mm grid background using a Canon PowerShot a710is digital camera, and the image imported into ArcGIS v. 9.2 (ESRI, Inc., 2006) and geocorrected using bilinear interpolation in the Rectify feature of the Georeferencing extension. M. Goolsby assisted with setting up the ArcGIS application. The Distance tool was then used to measure leaf area (mm^2), leaf length (mm), leaf width (mm), and the distance from leaf base to widest point (mm). The leaf length to width ratio was calculated from these measures. The basal shape of the leaf blade was then categorized as deltoid (D), cordate (C), or rhomboid (R). The leaf blade was oven-dried at 80°C for 36 hours and then weighed. This dry mass was used to calculate specific leaf area (SLA) as described in Section 2.2.2.

Each ramet was assessed for the presence and density of hairs on the main stem, petiole, and abaxial leaf surface, with the score categorized as dense (D), sparse (S) or absent (A) for each.

The final stem height (cm) and diameter (mm) of each replicate was assessed following the same procedure as used the first day. From the starting and final measures a transformed measure of relative grow rate, assuming a linear rate over time, was calculated as $RG = \log(final/start)$ for both height and diameter. In order to assess rates of sylleptic branch formation, the proportion of branching nodes was calculated as the ratio of the number of sylleptic branches to the total number of buds along the main stem, which was then arcsine transformed before statistical analysis. In addition, the number of new unfurled leaves, beginning with the leaf marked on the first day of the experiment, was counted for each ramet.

Cell imprints were taken from the third interveinal regions (mid-leaf) of each mature leaf collected (prior to drying) following the procedure described in Section 2.2.2. Imprints were imaged in three randomly-selected regions and video-photographed using Metamorph software on a Zeiss light microscope using a DAPI filter with bright field microscopy using a 40x/0.75 air objective. Images were saved as 8-bit JPEG format with a 10 μm scale bar on at least one of the three images. To calculate cell number and cell size, the highest-quality image was imported into Image-J software v1.43u (Wayne Rasband, NIH USA, available for download at rsb.info.nih.gov/ij). Three measures were made of each image: the number of whole stomata photographed, the number of whole epidermal (pavement) cells, and the area of 10 randomly-chosen epidermal cells. Four cell traits were calculated from these measures. The average cell area of epidermal (pavement) cells (μm^2) was calculated as the mean of the 10 measures. The number of cells per leaf was calculated as

$$\frac{CN_i}{Area_i} \times LA, \tag{3.1}$$

where CN_i is the number of cells in the image field of view, $Area_i$ is the area of the field of view in the image converted to mm^2, and LA is the are of the mature leaf (mm^2).

Stomatal density was calculated as the number of stomata per mm^2, and stomatal index as the number of stomata per 100 cells.

3.2.2 Climatic variation among sampling locations

In order to assess the local climate at the native sites where these *P. nigra* were collected, the minimum monthly precipitation (mm) and maximum monthly temperature (°C) was determined from www.worldclim.org (Hijmans et al. 2005) as described in Chapter 2.

3.2.3 Data analysis

Significant morphological variation between populations (excluding no. 15) was tested using a general linear model for each trait:

$$y = a_i + b_{ij} + c_k + a_i c_k, \tag{3.2}$$

where a_i represents the fixed factor of population level of variance, b_{ij} the random factor genet (within population) variance, c_k the random factor block variance, and $a_i c_k$ the interaction between population and block. Although the three blocks were situated on the same greenhouse bench, the block factor was included to test for any environmental variance across the greenhouse setting. Population was designated a fixed factor and genet and block random factors. The GLM was implemented in Minitab v16 (Coventry, UK).

Difference between populations in the four categorical traits (hairs on stem, petiole, or abaxial leaf surface and the shape of the leaf base) was tested using χ^2 tests.

Overall morphological differences using the mean value of each trait for each genet (including pop. 15) were described using discriminant analysis as described in Section 2.3.1. A total of 21 traits were included: 11 leaf traits (leaf area, leaf length, leaf width, leaf length:width ratio, distance from the widest point to the base of the leaf, specific leaf area, petiole length, petiole width, petiole height, proportion of leaves having a cordate or deltoid leaf base), two traits related to pubescence (absence or density on the petiole or abaxial leaf surface), four biomass traits (relative change in stem height and diameter, proportion of branching nodes, number of unfurled leaves), and four cell traits (average cell area, average number of cells per leaf, stomatal density and stomatal index). Two measures (proportion of rhomboid leaf bases and stem pubescence) were omitted from the analyses in order to maintain a positive definite correlation matrix.

Previous studies have reported significant correlation between leaf, biomass, and cell traits in *Populus* (Rae et al. 2004, Trewin et al. in preparation). Specific consideration of the corrlation matrix from the factor analysis was made for four pairs of traits: leaf area and cell area; cell area and cell number per leaf; leaf area and relative change in diameter; and leaf area and relative change in height. Correlation between these traits may provide selection indices for breeding and management purposes.

Correspondence between overall morphological differentiation and climatic differences at sampling sites was tested using Kendall's τ as implemented in the package R (v 2.10.0, R Foundation for Statistical Computing). Correlations were tested between four morphologic factors and the minimum monthly precipitation and maximum monthly temperature at each site, with experiment-wide error rates controlled through sequential Bonferroni procedures (Rice 1989). Curve-fitting algorithms implemented by SPSS (v17.0, Chicago, Illinois) were applied to describe trends between variables that did not fit a linear relationship.

3.3 Results

P. nigra from western (France and Italy) and central Europe (Hungary and Bosnia and Herzegovina) display significant morphological variation when grown in a common greenhouse setting. 11 of the 17 of the quantitative traits measured and four of the six categorical traits scored varied significantly among populations (data range and tests of the population main effect are provided in Figures 3.3 and 3.4). Significant differences were observed in the frequency of cordate ($\chi^2_6 = 31.4$, $P<0.001$) and rhomboid ($\chi^2_6 = 30.8$, $P<0.001$) but not deltoid ($\chi^2_6 = 10.9$, $P=0.09$) leaf bases. Rhomboid leaf bases were most frequenct in population 17 (ssp. *caudina*), but absent from the other populations from Bosnia and Herzegovina (nos. 21 and 23). Leaf size, including petiole size, varied among the populations, with the largest leaves occurring in trees from population 23 (Bosnia and Herzegovina) and the smallest in population 17 (*P. nigra* ssp. *caudina*), though the single sample from Spain was smallest overall.

Post hoc comparisons were conducted for traits displaying non-significant population main effects when either the population x block (P*B') or genet within population [G'(P)] main effects were non-significant at the $P \geq 0.25$ level (Doncaster & Davey, 2007). Relative growth in height displayed $P \geq 0.25$ for both P*B' and G'(P) main effects, resulting in a pooled population factor ratio of $F_{6, 40} = 0.36$, $P>0.05$, confirming that growth in height did not vary among populations. Relative growth in diameter displayed $P \geq 0.25$ for the G'(P) main effect, and post hoc pooling resulted in a pooled population factor ratio of $F_{6, 12} = 1.45$, $P>0.05$, providing no evidence of significant differentiation in diameter growth rate among populations. Two traits (SLA and distance between base and widest point of leaf) displayed non-significant P*B' effects of $P \geq 0.25$. For the distance between leaf base and width, the post hoc pooled population factor ratio of $F_{6,} 28 = 1.00$, $P>0.05$, indicating populations do not differ in this leaf trait. Post hoc pooling for SLA resulted in a pooled population factor ratio of $F_{6, 28} = 2.51$, $P = 0.045$, providing marginal evidence of population differentiation in leaf thickness.

Significant block effects were detected for number of unfurled leaves ($F_{2,12} = 6.7$, $P=0.011$) and specific leaf area ($F_{2,12} = 17.0$, $P<0.001$), though no population by block interactions were observed for these traits ($F_{12,53} = 1.09$, $P=0.389$ for leaf number; $F_{12,56} = 1.03$, $P=0.434$ for SLA), indicating genet means were sufficient for downstream analyses. Although the number of biological replicates (three) approached the minimum required for such analyses, these results indicate high levels of genetic variance for leaf morphology exist among natural populations of *P. nigra* in central Europe.

Overall there was no evidence of significant differences in biomass between populations. Post hoc comparisons confirmed the lack of differences in relative growth in height and diameter (above). The number of unfurled leaves produced during the experiment did vary significantly among populations, with the French population and *P. nigra* ssp. *caudina* producing a greater number of leaves than the large-leaf trees from Italy and Hungary. Differences in branching corresponded to the number of leaves, with *P. nigra* ssp. *caudina* and the single Spanish sample displaying the greatest proportion of sylleptic branches.

Only the *P. nigra* ssp. *caudina* samples (population 17) displayed hairy stems (80% of samples dense, 20% sparse), petioles (80% of samples dense, 20% sparse), or abaxial surfaces (20% sparse, 80% absent). Significant differences were observed in the frequency of hairs on the petiole ($\chi^2_6 = 168$, $P<0.001$) and hairs on the stem ($\chi^2_6 = 162$, $P<0.001$), but not hairs on the abaxial surface ($\chi^2_6 = 6$, $P=0.423$). These results are consistent with dense pubescence being a key trait of the Mediterranean hairy black poplar.

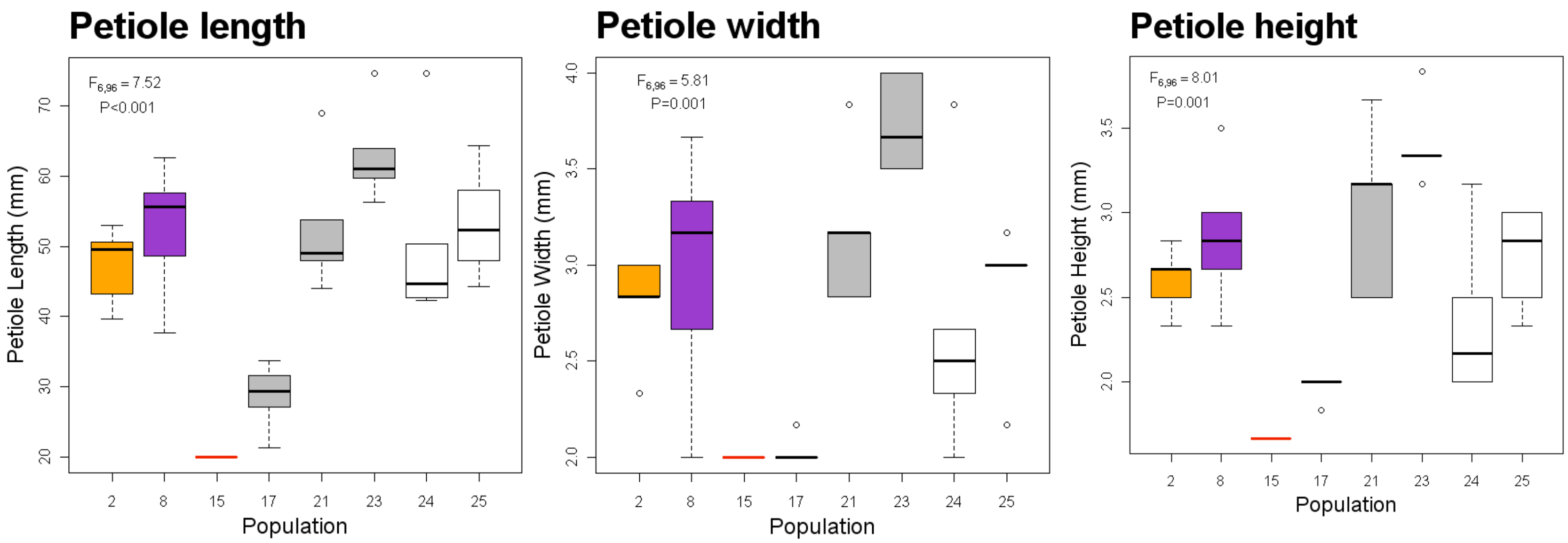


Figure 3.3 Morphological variation in seven populations of *Populus nigra*. Boxplots display the range of values observed in the five genets from each population. A general linear model was used to test for differences between population means, with the test statistic presented in each plot. Populations 2 (from France) and 8 (from Italy) and a single sample from a Spanish population (15) were included in the experiment to provide comparison to the collection described in Chapter 2. Population 17 is *P. nigra* ssp.*caulina*. Colours represent the country of origin: France (orange), Italy (purple), red (Spain), gray (Bosnia & Herzegovina) and white (Hungary).

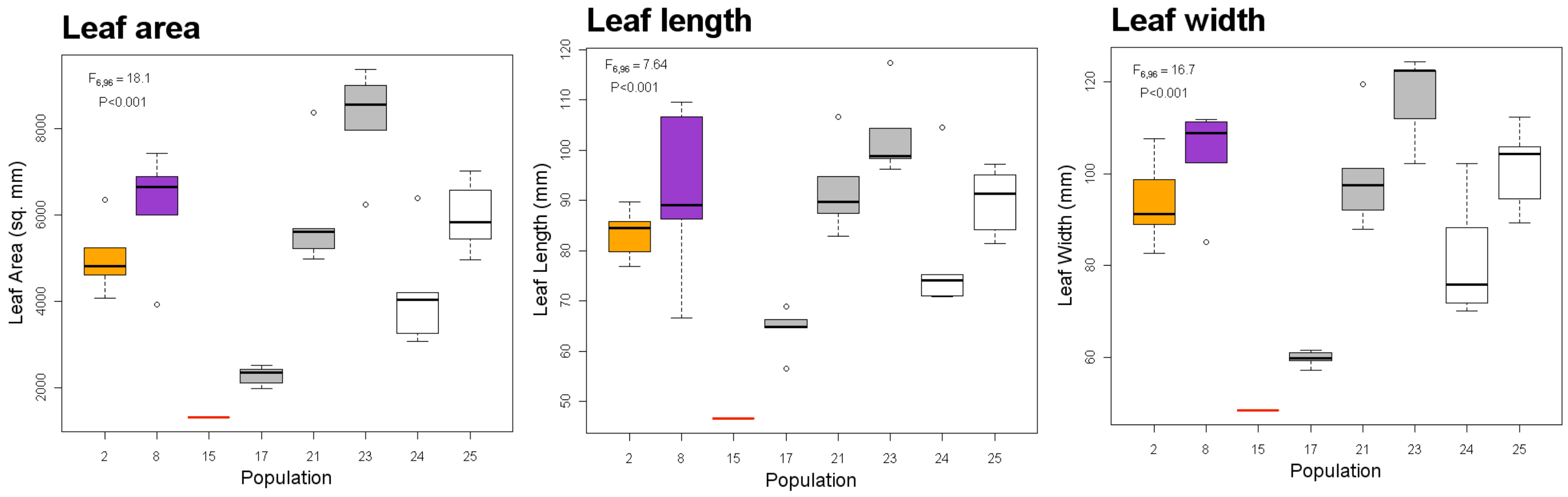
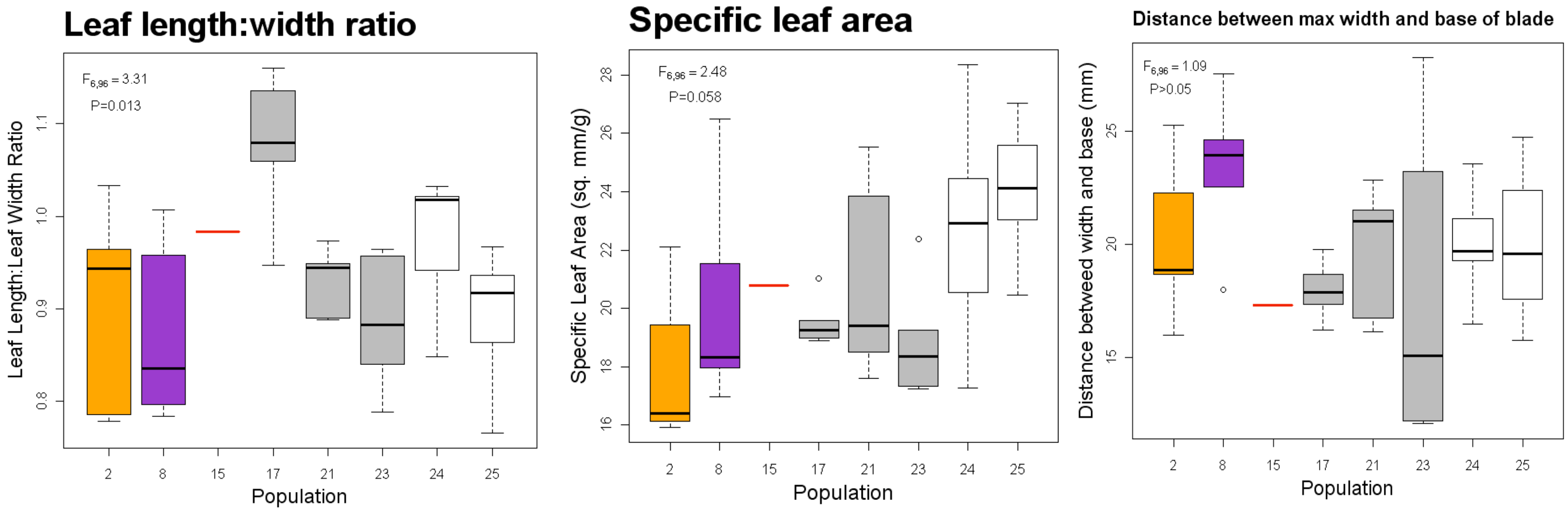


Figure 3.3 Continued. Colours represent the country of origin: France (orange), Italy (purple), red (Spain), gray (Bosnia & Herzegovina) and white (Hungary).


Leaf length:width ratio

Specific leaf area

Distance between max width and base of blade

Figure 3.3 Continued. Colours represent the country of origin: France (orange), Italy (purple), red (Spain), gray (Bosnia & Herzegovina) and white (Hungary).

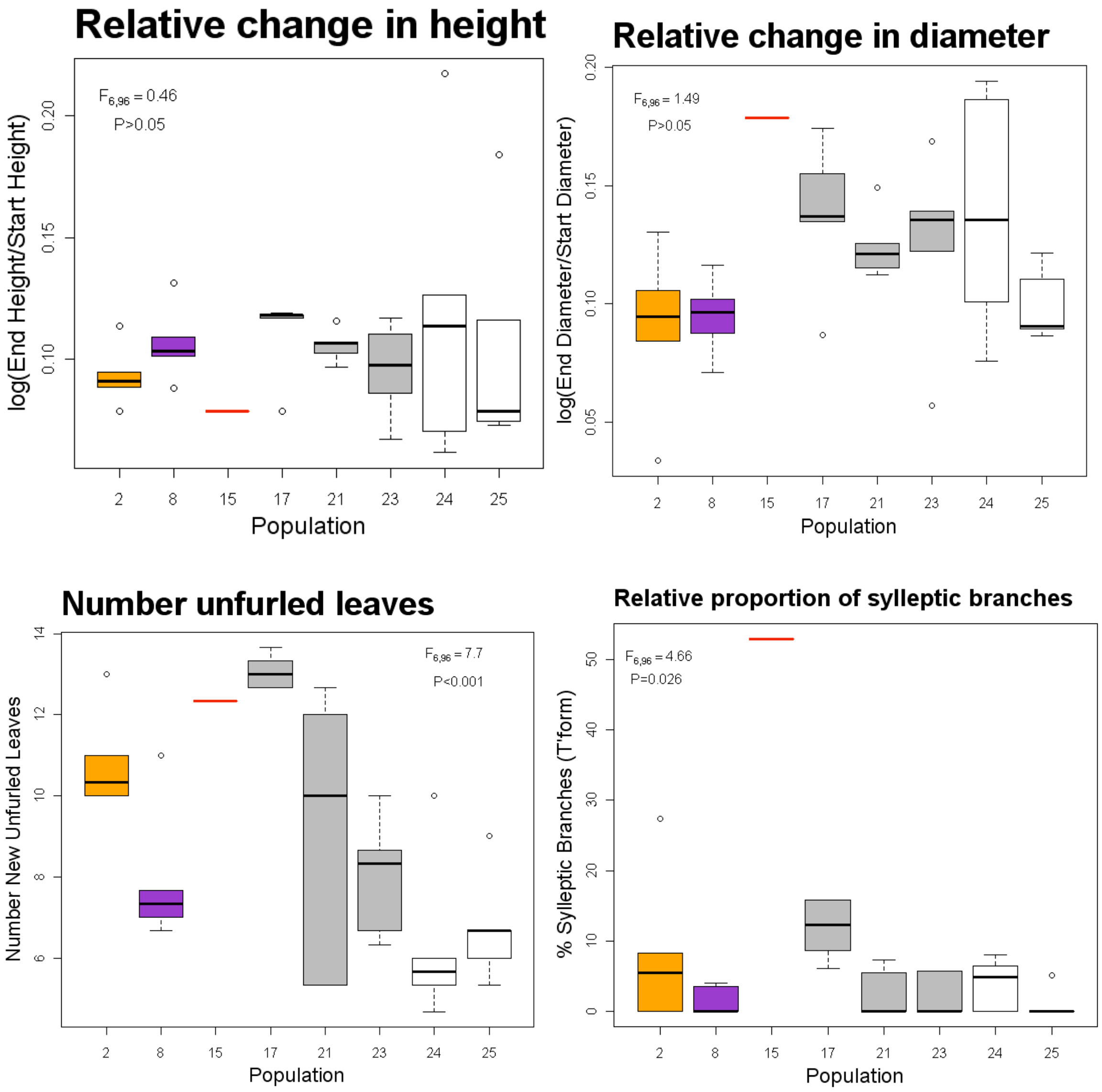


Figure 3.3 Continued. Colours represent the country of origin: France (orange), Italy (purple), red (Spain), gray (Bosnia & Herzegovina) and white (Hungary).

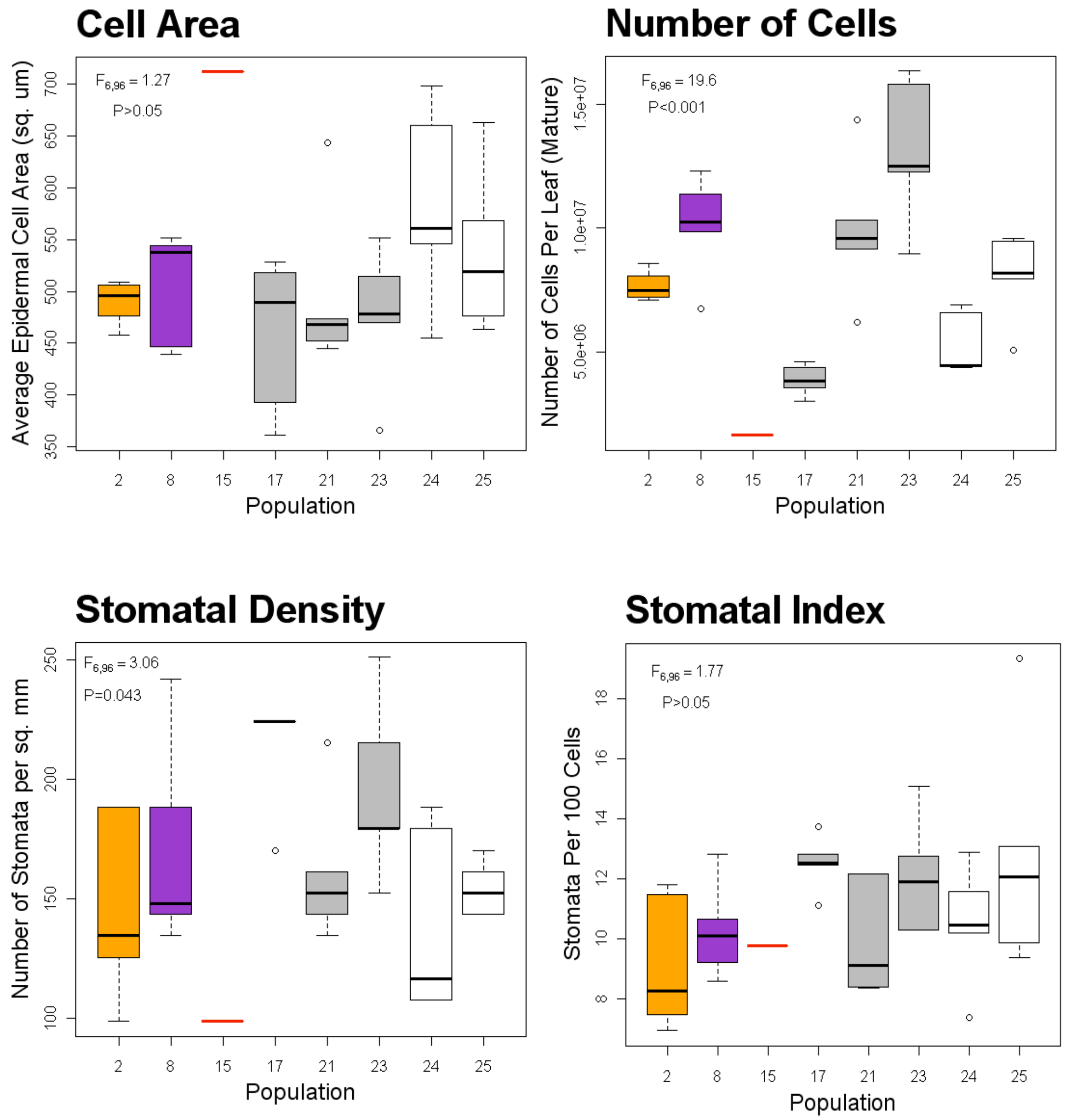


Figure 3.3 Continued. Colours represent the country of origin: France (orange), Italy (purple), red (Spain), gray (Bosnia & Herzegovina) and white (Hungary).

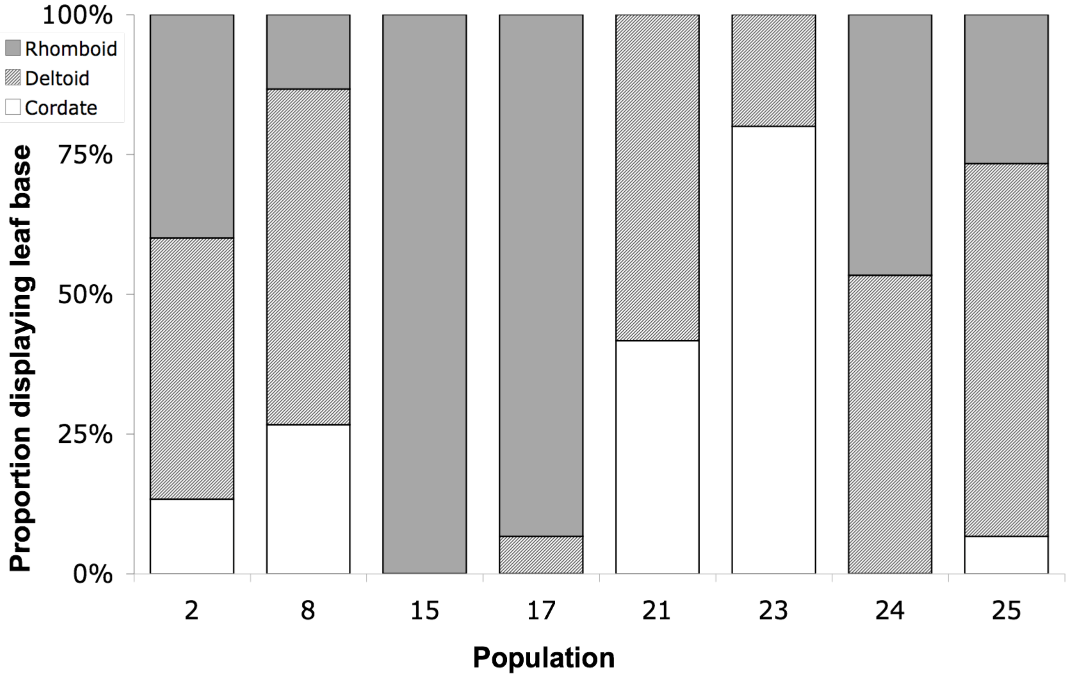


Figure 3.4 Morphological variation in the shape of leaf base in seven populations of *Populus nigra*. Barplots display the proportion of genets displaying different leaf base shapes. Population 2 is from France, population 8 from Italy, population 15 is a single sample is from Spain, populations 17 (*P. nigra* ssp.*caulina*), 21, and 23 are from Bosnia & Herzegovina, and populations 24 and 25 are from Hungary.

Only marginal evidence was observed for differences in cell morphology among populations. Cell area displayed a population by block interaction ($F_{12,54}$=2.29, P=0.019), but no population effect (P>0.05), indicating cell size displays phenotypic plasticity influenced by genetic variance among populations. Estimates of cell number per leaf varied significantly among populations, but as this trait is a product of measures of cell and leaf area, this variation is likely driven by leaf size. No difference in stomatal patterning (stomatal index) was detected among populations, though population 17 (*P. nigra* ssp. *caudina*) displayed slightly greater values than the other collections. This higher stomatal frequency is inconsistent with predicted adaptation to water deficit in this subspecies. The small-leaf sample from Spain and *P. nigra* ssp. *caudina* did not display congruent cell traits. The Spanish sample displayed larger cell area than the other samples, and a lower stomata index than *P. nigra* ssp. *caudina*. These observations are inconsistent with the predicted convergent evolution of cell traits in response to water deficit.

The correlation matrix for the discriminant factor analysis revealed significant correlation among and between biomass, leaf, pubescence, and cell traits (Table 3.2). Leaf and petiole traits were highly correlated, with larger leaves having longer petioles. Petiole and abaxial leaf surface pubescence was negatively correlated with leaf size, consistent with *P. nigra* ssp. *caudina* having small leaves. Leaf base was moderately correlated with leaf size, and indicated the cordate shape is associated with larger leaves containing a greater number of cells. Branching pattern was significantly negatively correlated with leaf size, indicating trees with more sylleptic branches display smaller leaves. Leaf production was negatively correlated with specific leaf area (Figure 3.5), but not leaf size, indicating trees with greater leaf production had thicker leaves but not necessarily larger leaf area.

Although correlation among biomass traits was generally low, including a non-significant relationship between the relative growth in height and diameter, greater branching was correlated with a higher relative growth rate for diameter.

Cell size and stomatal patterning were uncorrelated with leaf size. However, cell area was positively correlated with specific leaf area and diameter growth rate, consistent with larger cells producing thicker leaves and greater biomass (Figure 3.5). Cell area was negatively correlated with the number of cells per leaf, indicating smaller leaves are composed of larger cells (Figure 3.5). Stomatal patterning was moderately positively correlated with petiole pubescence, indicating a greater density of hairs on the petiole is associated with a higher stomatal index, which is not consistent with the pattern predicted if *P. nigra* ssp. *caudina* morphology represents an adaptation to water deficit.

In summary, no evidence was found that leaf area serves as an indicator for biomass in this collection of *P. nigra*. Four pairs of traits previously shown to be correlated in *Populus* trees displayed no significant correlation in this collection: leaf area and cell area, leaf area and relative change in both height and diameter, and the number of unfurled leaves and relative change in height. In contrast, significant correlation between cell area and three traits, the number of cells per leaf, specific leaf area and the relative growth in diameter, indicates that differences in cell size may underpin variation in overall morphology in *P. nigra*.

Table 3.2 Correlation among 21 traits scored in a glasshouse experiment of seven populations of *P. nigra*. Two traits (proportion with rhomboid leaf bases and stem pubescense) were omitted in order to obtain a positive definite correlation matrix. Correlation coefficients are provided above the diagonal and 1-tail tests of significance below the diagonal.

	Petiole length	Petiole width	Petiole height	Leaf area	Leaf length	Leaf width	Length:width ratio
Petiole length		0.819	0.809	0.885	0.913	0.872	-0.336
Petiole width	<0.001		0.908	0.867	0.867	0.854	-0.356
Petiole height	<0.001	<0.001		0.918	0.904	0.919	-0.431
Leaf area	<0.001	<0.001	<0.001		0.920	0.978	-0.526
Leaf length	<0.001	<0.001	<0.001	<0.001		0.894	-0.230
Leaf width	<0.001	<0.001	<0.001	<0.001	<0.001		-0.631
Leaf length:width ratio	0.023	0.016	0.004	<0.001	0.089	<0.001	
Specific leaf area	0.118	0.017	0.030	0.143	0.055	0.153	0.193
Distance leaf width:base	0.067	0.220	0.129	0.096	0.038	0.056	0.322
Cordate	0.004	0.001	0.003	<0.001	0.004	0.001	0.020
Deltoid	0.014	0.019	0.002	0.023	0.008	0.004	0.035
Change in height	0.453	0.448	0.456	0.422	0.334	0.439	0.111
Change in diameter	0.064	0.135	0.092	0.075	0.050	0.028	0.081
No. unfurled leaves	0.012	0.165	0.187	0.047	0.082	0.025	0.014
Branching	0.001	0.005	0.002	0.001	<0.001	<0.001	0.036
Hairs on petiole	<0.001	<0.001	0.001	<0.001	0.001	<0.001	<0.001
Hairs on abaxial	0.018	0.086	0.064	0.054	0.036	0.055	0.451
Cell area	0.339	0.355	0.248	0.255	0.225	0.258	0.434
No. cells per leaf	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	0.001
Stomatal density	0.225	0.463	0.393	0.352	0.314	0.306	0.296
Stomatal index	0.213	0.334	0.339	0.303	0.279	0.248	0.285

Table 3.2 Continued.

	Specific leaf area	Distance leaf width:base	Cordate	Deltoid	Change height	Change diameter	No. unfurled leaves
Petiole length	-0.203	0.254	0.440	0.367	0.021	-0.258	-0.375
Petiole width	-0.354	0.133	0.513	0.348	-0.022	-0.189	-0.167
Petiole height	-0.318	0.194	0.453	0.460	0.019	-0.227	-0.153
Leaf area	-0.183	0.223	0.585	0.335	-0.034	-0.245	-0.284
Leaf length	-0.270	0.300	0.430	0.396	-0.074	-0.278	-0.238
Leaf width	-0.176	0.270	0.513	0.436	0.027	-0.322	-0.330
Leaf length:width ratio	-0.149	-0.080	-0.345	-0.305	-0.209	0.238	0.367
Specific leaf area		-0.046	-0.120	-0.016	0.252	0.093	-0.537
Distance leaf width:base	0.396		-0.292	0.261	0.031	0.001	-0.105
Cordate	0.242	0.042		-0.358	-0.079	-0.080	-0.173
Deltoid	0.462	0.062	0.016		0.196	-0.292	-0.201
Change in height	0.069	0.429	0.323	0.126		-0.191	-0.068
Change in diameter	0.295	0.497	0.321	0.042	0.132		0.114
No. unfurled leaves	<0.001	0.271	0.156	0.120	0.346	0.254	
Branching	0.199	0.081	0.090	0.051	0.348	0.024	0.001
Hairs on petiole	0.226	0.150	0.070	0.008	0.337	0.060	<0.001
Hairs on abaxial	0.455	0.385	0.267	0.132	0.323	0.058	0.048
Cell area	0.024	0.103	0.102	0.254	0.358	0.014	0.059
No. cells per leaf	0.047	0.146	<0.001	0.123	0.401	0.046	0.170
Stomatal density	0.078	0.036	0.067	0.029	0.233	0.222	0.137
Stomatal index	0.234	0.016	0.293	0.214	0.105	0.492	0.339

Table 3.2 Continued.

	Branching	Hairs on petiole	Hairs on abaxial	Cell area	No. cells per leaf	Stomatal density	Stomatal index
Petiole length	-0.515	-0.619	-0.352	-0.072	0.782	-0.130	-0.137
Petiole width	-0.425	-0.532	-0.233	-0.064	0.769	-0.016	-0.074
Petiole height	-0.480	-0.514	-0.259	-0.117	0.826	-0.047	-0.072
Leaf area	-0.512	-0.583	-0.272	-0.113	0.906	-0.066	-0.089
Leaf length	-0.542	-0.525	-0.304	-0.130	0.821	-0.083	-0.101
Leaf width	-0.561	-0.640	-0.271	-0.112	0.883	-0.087	-0.117
Leaf length:width ratio	0.303	0.569	0.021	-0.029	-0.490	0.092	0.098
Specific leaf area	-0.145	-0.129	0.019	0.333	-0.283	-0.242	0.125
Distance leaf width:base	-0.238	-0.178	-0.050	0.216	0.180	-0.303	-0.359
Cordate	-0.228	-0.251	-0.107	-0.217	0.615	0.255	0.094
Deltoid	-0.277	-0.401	-0.191	0.114	0.199	-0.319	-0.136
Change in height	-0.067	0.073	0.079	-0.063	-0.043	0.126	0.214
Change in diameter	0.331	0.263	0.267	0.366	-0.285	-0.132	0.003
No. unfurled leaves	0.507	0.582	0.282	-0.265	-0.164	0.187	-0.072
Branching		0.246	0.110	0.146	-0.426	-0.087	-0.013
Hairs on petiole	0.074		0.450	-0.239	-0.484	0.444	0.280
Hairs on abaxial	0.261	0.003		-0.046	-0.218	0.237	0.108
Cell area	0.197	0.080	0.394		-0.405	-0.622	0.015
No. cells per leaf	0.005	0.001	0.101	0.007		0.194	-0.153
Stomatal density	0.308	0.003	0.082	<0.001	0.128		0.570
Stomatal index	0.470	0.049	0.266	0.465	0.186	<0.001	

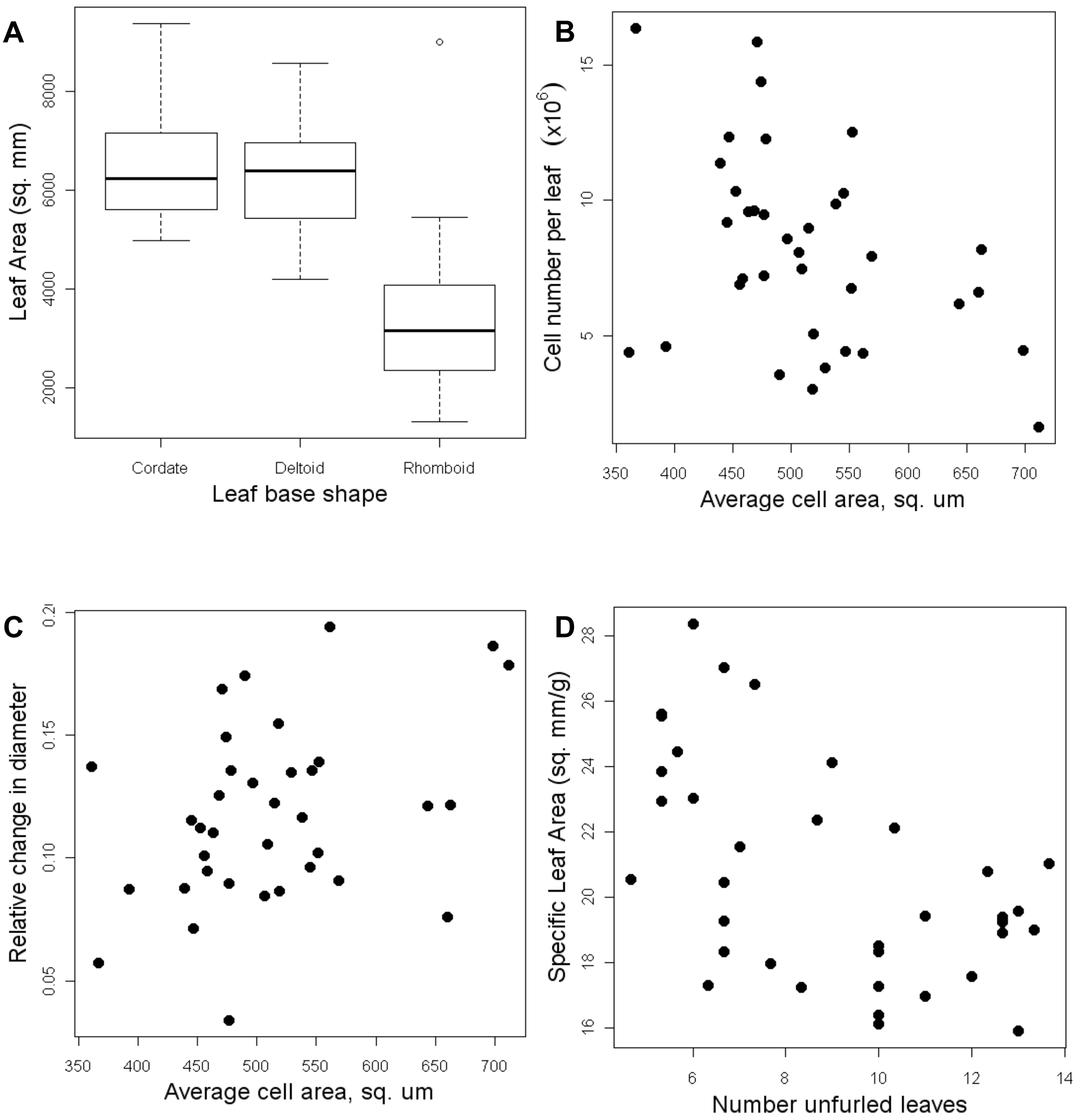


Figure 3.5 Discriminant analysis revealed significant correlation among traits measured in a greenhouse study of *P. nigra*, including leaf are and leaf shape (A), cell area and the number of cells per leaf (B) and the relative change in diameter (C), and specific leaf area and the number of unfurled leaves (D).

Table 3.3 Tests of discriminant dimensions resulting from the discriminant analysis of 21 traits identify three functions that significantly differentiate between populations of *P. nigra*.

Function	Canonical Correlation	X^2	df	Sig.
1	0.994	264.2	147	<0.001
2	0.962	172.8	120	0.001
3	0.941	119.7	95	0.044

The discriminant classification analysis partitioned the variation in the morphological traits into three signficant factors (Table 3.3), accounting for a total of 93.9% of the observed variation. The first two factors explain 87% of the variation. The structure matrix showed the presence of hairs to be highly correlated with the first factor and branching pattern to be correlated with the second. The third factor correlated with leaf morphology generally related to leaf size, including the number of cells per leaf (which was calculated from leaf area) and petiole size.

The discriminant classification analysis also indicated the *P. nigra* ssp. *caudina* (pop. 17) and the sample from Spain (population 15) to be distinct from the other five populations (Figure 3.6). The first factor, which was correlated with petiole pubescence, distinguishes the Mediterranean hairy black poplar from the common subspecies, consistent with the presence of hairs being a key trait in identifying this taxon. The second factor, which was correlated with the frequency of sylleptic branches, separates the Spanish sample from the other populations, indicating the plant architecture in this sample to be distinct. Qualitative observation of other Spanish samples in a glasshouse setting is consistent with sylleptic branches being typical of trees from this population. The third factor resolves differences between the typical *P. nigra* populations, and is correlated to measures of leaf size. The pattern of population differences reveals the French population to be distinct from those collected in Bosnia and Herzegovina, with samples from Italy ad Hungary intermediate (Figure 3.6).

Table 3.4 Correlation between morphological variation in 21 traits and three standardized canonical discriminant functions (factors) that significantly differentiate between seven populations of *P. nigra*.

	Factor		
Trait	1	2	3
Hairs on petiole	-0.696*	-0.197	0.038
Hairs on abaxial	-0.050*	-0.014	0.003
Branching	-0.069	0.452*	-0.165
No. cells per leaf	0.117	-0.183	0.340*
Leaf area	0.149	-0.209	0.339*
Cordate	0.046	-0.030	0.306*
Petiole height	0.109	-0.151	0.282*
Leaf width	0.174	-0.257	0.267*
Leaf length	0.109	-0.198	0.233*
Petiole length	0.120	-0.150	0.214*
Stomatal density	-0.060	-0.117	0.171*
Change in diameter	-0.037	0.142	0.117
Leaf lenth:width ratio	-0.088	0.059	-0.009
Stomatal index	-0.032	-0.060	0.122
Cell area	0.029	0.156	-0.064
Distance leaf width:base	0.024	-0.041	-0.067
No. unfurled leaves	-0.106	0.073	-0.117
Specific leaf area	0.018	-0.014	-0.014
Petiole width	0.098	-0.086	0.261
Deltoid	0.053	-0.070	-0.084
Change in height	-0.010	-0.028	0.006
% of Variance	76.1	10.9	6.8
Cumulative %	76.1	87.0	93.9

*Largest absolute correlation between the trait and any function.

†No. cells per leaf is calculated from and thus correlated with leaf area.

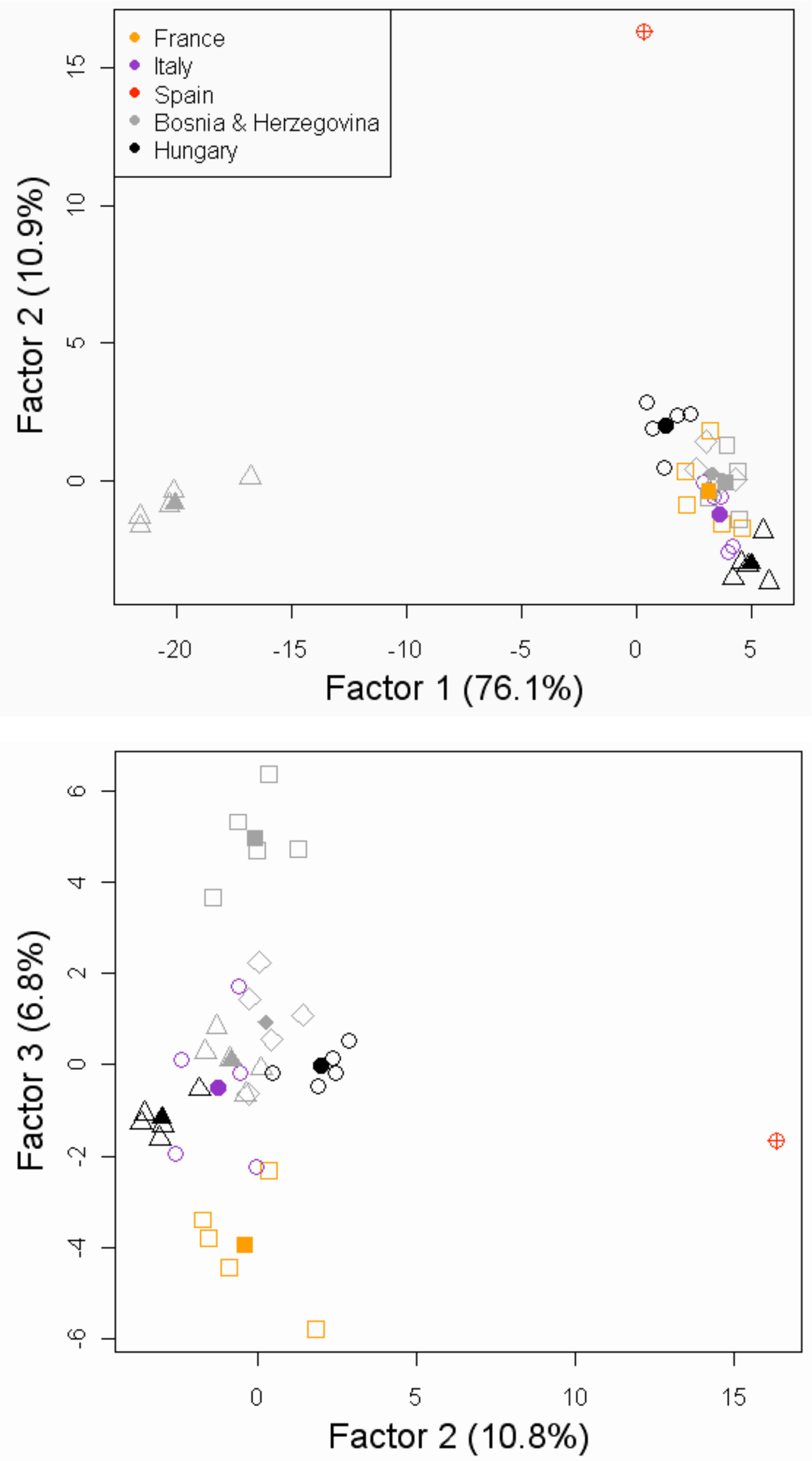


Figure 3.6 Discriminant analysis of 21 traits differentiates between populations of *Populus nigra*. Samples from France, Italy, and a single sample from Spain were previously described in a common garden experiment (Chapter 2); one population (gray triangles) is *P. nigra* ssp. *caudina*. Open symbols represent a single genet while closed symbols represent population means. Different symbols sharing colours represent populations within each country.

Non-parametric tests for simple correlation (Kendall's τ) were used to explore the relationship between morphology (population centroids for the three morphological factors) and climate (minimum monthly precipitation and maximum monthly temperature). All tests of simple correlation were non-significant ($P>0.05$), indicating that either no relationship exists between the measures or a simple correlation is insufficient to describe one. Visual inspection of the relationship between each morphological factor and the two measures of climate provide insight into the lack of linear relationship between the variables (Figure 3.7). The first morphological factor separates *P. nigra* ssp. *caudina* (pop. 17, gray triangles) from the other populations. The site of pop. 17 differs from other climates in maximum monthly temperature but not minimum monthly precipitation (Figure 3.7), indicating the morphological adaptation in *P. nigra* ssp. *caudina* may be in response to temperature and not water deficit. Indeed, curve-fitting analysis of the pattern observed in Figure 3.7(A) identified a significant quadratic relationship between the first morphologic factor and maximum monthly temperature ($y = -4.722x^2 + 266.3x - 3749$, regression ANOVA $F_{2,5}=64.4$, $P<0.001$; adj. $R^2 = 0.948$). Similarly, the simple correlation between morphological factor 2 and minimum monthly precipitation was non-significant [Figure3.7(D)], but curve-fitting analyses identify a significant quadratic relationship between the variables ($y = 0.021x^2 - 1.948x + 42.82$, regression ANOVA $F_{2,5} = 29.2$, $P=0.002$; adj. $R^2 = 0.890$). While higher-power relationships are difficult to interpret biologically, the trend of these to observations indicate that the small-leaf trees from Spain and hairy subspecies from the Balkan Peninsula may be adaptations to different climatic extremes rather than convergent evolution. The small number of sampling points (a single population of *P. nigra* ssp. *caudina* and a single genet from Spain) makes conclusions tentative at best, and additional sampling will be required to explain fully these observations.

No significant regression analyses were identified between morphological factor 3 and either measure of climate, indicating the phenotypic differentiation observed is non-adaptive or reflects environmental variables not assessed here. Additional sampling points will be required to resolve fully the relationship between morphological variation and climate in *P. nigra* from central Europe.

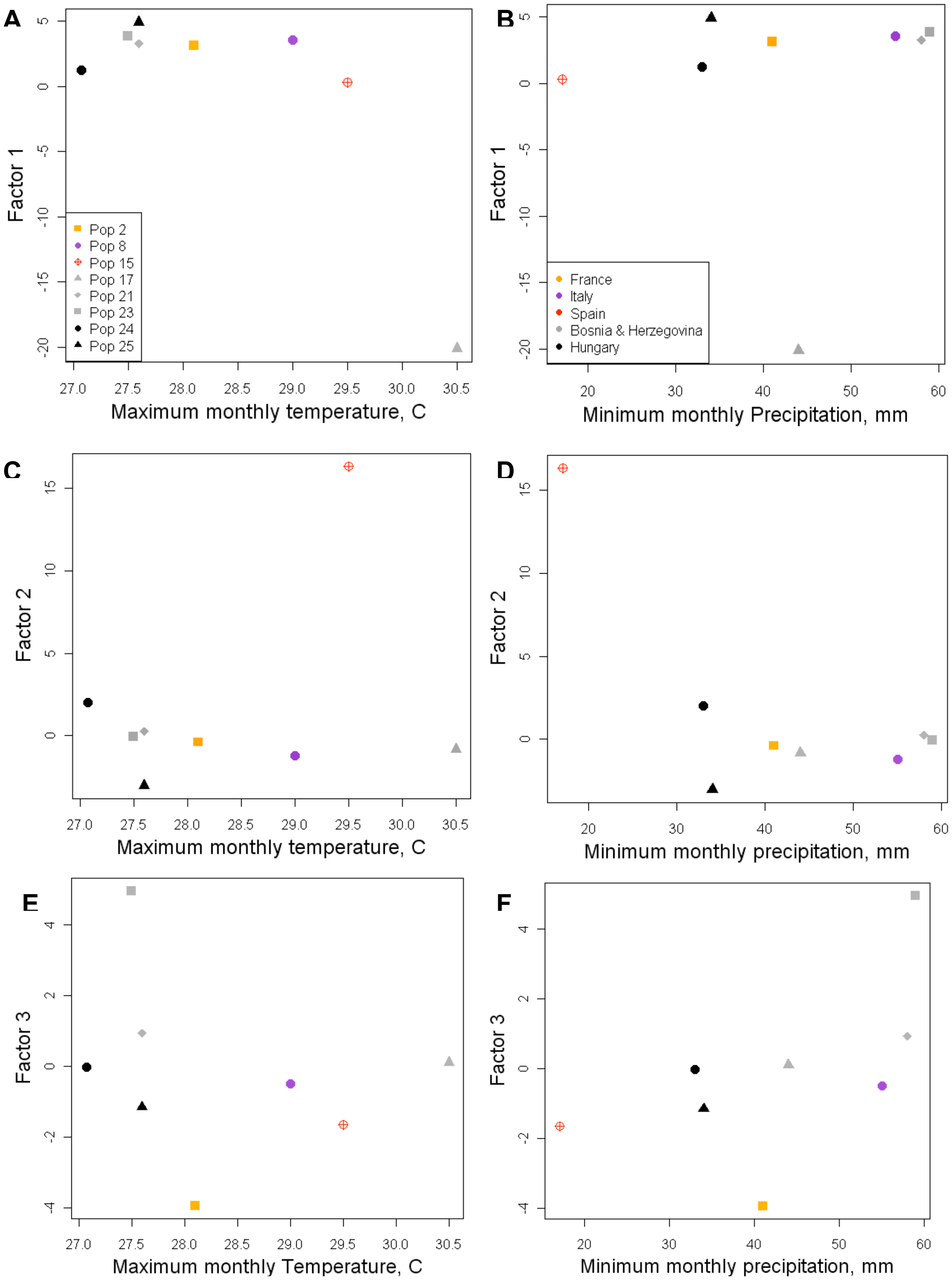


Figure 3.7 Morphological variation in *P. nigra* samples from Spain (red dot) and *P. nigra* ssp. *caudina* from Bosnia & Herzegovina (gray triangle) do not correspond to the same climatic variables. Tests of simple correlation were non-significant but quadratic relationships were detected for (A) and (D).

3.4 Discussion

3.4.1 Possible adaptive origins of phenotypic diversity

In general, phenotypic differentiation among plant species and populations is expected to reflect adaptation to local environments (Dudley 1996a, b; Picotte et al. 2007). Previous phenotypic and genetic analyses of *P. nigra* samples from western Europe indicate historic vicariance and isolation by adaptation can explain part of the morphological differentiation observed among those populations (Chapter 2). In this greenhouse study of additional populations of *P. nigra*, the small number of populations sampled and the lack of genetic data limit interpretations regarding the evolutionary forces influencing the pattern of phenotypic differences observed. However, some conclusion can be made from the observations.

If adaptation is the primary force influencing phenotypic differentiation, the pattern of morphological differences should correspond to environmental differences. In this collection of *P. nigra*, two populations were morphologically distinct compared to the majority of the collection: *P. nigra* ssp. *caudina* (pop. 17) based on the presence of hairs on petioles, stems, and even abaxial leaf surfaces, and Spain (pop. 15) based on the frequency of sylleptic branches. The inclusion of additional samples from Spain and France in the previous study likely explains the lower overall differentiation observed in Chapter 2 compared to the exaggerated distinctiveness indicated in this analysis. The morphological distinctiveness of *P. nigra* ssp. *caudina* in this study was more pronounced than in a previous comparison with *P. nigra* ssp. *nigra* (Kajba et al. 2004). Three reasons may account for this difference. First, Kajba et al. did not include pubescence in the phenotypic assessment, and this trait had the greatest correlation with the discriminant factor distinguishing pop. 17 in this study. Second, the previous examination used field-collected leaves (Kajba et al. 2004). The glasshouse environment used here may have induced genotype by environment interactions that resolve greater differences between populations, though the controlled environment would be expected to minimize, not amplify phenotypic differences. Third, this collection of *P. nigra* ssp. *nigra* consisted of populations from eastern Bosnia and Herzegovina and central Hungary, while (Kajba et al. 2004) studied samples from neighboring Croatia. Phenotypic differences in the typical subspecies may account for the differences between these studies.

Here, the morphological differences of *P. nigra* ssp. *caudina* correlated with the maximum monthly temperature and not minimum monthly precipitation. Previous examination of two populations from Spain indicated the morphological pattern corresponded to differences in minimum monthly precipitation, not maximum monthly temperature (Chapter 2). While the Spanish sample here consisted of a single genetic individual, the pattern of morphological and environmental differences (factor 2 and minimum montly precipitation) was consistent with the previous observations. Together, these patterns indicate the Spanish and *caudina* subspecies samples represent two distinct small-leaf phenotypes that have likely evolved in response to different environmental variables: precipitation patterns in the Iberian peninsula and summer temperature in the Mediterranean hairy subspecies.

In *P. nigra*, specific leaf area was negatively correlated with leaf production (Figure 3.5), indicating high leaf production was associated with greater leaf thickness. If leaf thickness is positively correlated with leaf persistence in *P. nigra* as has been observed in other species (Diemer 1998), then this observation may indicate a difference in nitrogen and phosphorous use and photyosynthetic capacity among populations. Variation in leaf area to mass ratio has been shown to correlate with photosynthetic and nitrogen characteristics of the leaf in herbaceous and woody species, indicating thicker leaves have a lower nitrogen content and photosynthetic capability (He et al. 2009). Within environmental classes, the correlations between specific leaf area and nitrogen and phosphorus content weakened with decreased temperature and increased precipitation (He et al. 2009), indicating the relationships may reflect a response to environmental stress. Additional analyses into the relationship between environmental stress (e.g. water deficit), leaf traits, and population of origin is warranted in this species.

Historic vicariance was found to influence phenotypic differences in *P. nigra* from western Europe (i.e. residual adaptation to glacial refugia) (Chapter 2). The pattern of phenotypic differences in the 'typical' *P. nigra* populations, summarized by the third discriminant factor, provided no evidence of residual phenotypic differentiation expected if the Italian and Balkan refugia were under divergent selection during the most recent glacial maximum. These findings are consistent with a previous study of chloroplast markers suggesting the two refugia maintained some gene flow during the glacial maximum (Cottrell et al. 2005). Phenotypically, samples from Italy were similar to the typical *P. nigra* from central Europe (Hungary and Bosnia and

Herzegovina), while the samples from the genetic admixture zone of France were divergent.

In addition to a greater sampling strategy (both the number of populations and sample size within each population) and the addition of genetic data to investigate the demographic structure of these populations, a common garden or reciprocal transplant study would be informative to resolve further the adaptive differences between populations. Phenotypic plasticity may mask morphological differences in a glasshouse setting that will be critical to plant growth rate and survival in a field setting (Abrams 1994). The adaptive advantage of the distinct small-leaf phenotypes will best be examined in common garden environments experiencing differences in temperature and precipitation similar to the native stands where the samples were collected.

3.4.2 Possible hybrid origins of phenotypic diversity

As interspecific hybridization is widespread among *Populus* species (Taylor 2002; Lexer et al. 2005), it is possible that part of the phenotypic variation observed in this collection may originate in past hybridization events. *P. nigra* regularly co-occurs with horticultural and industrial plantings of exotic and hybrid poplars, typically *P. deltoides* and the (DxN) hybrid *P. × canadensis*, which are common in plantations found along major rivers. Detailed historic records of the planting of non-native species are typically lacking for large river systems, meaning the provenance of many apparently natural stands cannot be determined without genetic marker evidence.

Naturally regenerated seedlings and saplings of hybrid origin have been identified from genetic analysis most recently in the Elbe rivers in Germany (Ziegenhagen et al. 2008) and the Rhine river in the Netherlands (Smulders et al. 2008a) and in river across Austria (Heinze 2008), indicating natural hybridization with *P. × canadensis* may threaten the genetic integrity of remaining natural *P. nigra* stands. Hybrid regeneration is not ubiquitous, however, as no hybrids were detected in young stands and seeds collected from the Ticino river in Italy, despite the presence of mature hybrids in the adjacent stand (Fossati et al. 2003). Siring of *P. nigra* seed by *P. × canadensis* (or *P. deltoides*) pollen appears less likely when *P. nigra* pollen is available, however (Vanden Broeck et al. 2006; Smulders et al. 2008a), indicating the probability of hybridization to be lower in larger, mature stands.

Distinguishing between species based on morphological features is difficult in young samples (seedlings, saplings, or newly-established cuttings, as used in this study) as immature samples may vary from the mature type or lack diagnostic features relating to, for example, inflorescence morphology or canopy architecture (Smulders et al. 2008a). In addition, neoformed leaves developed during the growing season (as used in this study) may be less informative than preformed leaves emerging from buds in the spring (Kajba et al. 2004). The variation in leaf base shape observed in this study may indicate introgressed hybrids are present in the larger collection of *P. nigra* being studied across western Europe. An acute leaf base (rhomboid or cuneate, Smulders et al. 2008a), is typical for *P. nigra*, with the leaf base shape *P. deltoides* varies from truncate (deltoid) to cordate (Eckenwalder 1984a). In addition, the colour of young leaves varies among species. *P. nigra* is characterized as having green immature leaves, while *P. × canadensis* is reported to produce red to bronze colored leaves (Smulders et al. 2008a). This trait does not appear to be diagnostic, however, as a sample that was classified as pure *P. nigra* based on genetic analyses was observed to produce bronze young leaves, indicating leaf color may be variable in *P. nigra* or ongoing backcrosses to *P. nigra* have resulted in introgression of *P. × canadensis* traits into natural stands (Smulders et al. 2008a). Leaf color was not characterized in this study, but qualitatively, some bronze color was observed in young leaves in this collection. Specifically, trees from the Ticino river Italy tended to display this trait (pop. 8), which were collected in a region previously shown to have a high frequency of mature hybrid individuals in what were considered natural stands (Fossati et al. 2003). These observations indicate that tests for introgression may be warranted to identify putative hybrids in the European collections of *P. nigra*, not only to validate ongoing genotypic and phenotypic analyses but also to continue to verify the source of germplasm for restoration efforts (Holderegger et al. 2005). Without disgnostic genetic tests, a possible hybrid origin for the morphology atypical of *P. nigra* (e.g. non-rhomboid leaf shape) cannot be ruled out.

3.4.3 Possible developmental mechanism underlying phenotypic diversity

The pattern of cell size and leaf area provides clues as to the developmental mechanisms underlying variation in leaf morphology in *P. nigra*. Specifically, cell area was negatively correlated with leaf area. The relative role of cell growth and cell division in leaf size varies among *Populus* species, with *P. trichocarpa* leaf size resulting from cell expansion, but *P. deltoides* leaf size being driven by cell division (Al Afas et al. 2007). The taxonomic relationship between *P. nigra* and *P. deltoides*, which belong to the same section *Tacamahaca*, would indicate cell division and not cell expansion to be the underlying mechanism driving leaf size variation in this collection.

The lack of significant variation in cell size among populations (omitting the Spanish sample) in this collection of *P. nigra* is consistent with cell division, and not cell expansion, driving leaf size variation in this species (Figure 3.3). Inclusion of the Spanish sample in the comparison of cell size and leaf area, however, indicates cell expansion may play a role in the small-leaf phenotype from the Iberian Peninsula. Here, a negative correlation was detected between cell area and leaf size when assessed over individual genets (Figure 3.5), consistent with a previous assessment of a larger collection from western Europe that found cell size and leaf area to be negatively correlated (Trewin 2008), Trewin *et al.*, in preparation). If small-leaf trees produce larger cells than the large-leaf phenotypes, the developmental mechanism of compensation, where cell expansion increases in response to reductions in cell division in order to maintain overall organ size (Horiguchi et al. 2006), may be responsible for developmental differences between these ecotypes. A fine-scale analysis of the relative role of cell growth and cell division in leaf development among four genets of *P. nigra* was undertaken to investigate the developmental mechanism underlying this phenotypic variation and is described in Chapter 4.

3.4.4 Implications of phenotypic differentiation for conservation of *P. nigra*

There has been great interest in the study and conservation of *P. nigra* in the European forestry community in the past decade (Koskela et al. 2004; Storme et al. 2004). A number of countries have initiated local germplasm collections and analyses of the genetic structure within existing populations in France (Legionnet and Lefevre 1996; Imbert and LeFevre 2003), Germany (Ziegenhagen et al. 2008), Italy (Fossati et al. 2003), the Netherlands (Arens et al. 1998; Smulders et al. 2008a), Switzerland (Holderegger et al. 2005), and the United Kingdom (Winfield and Hughes 2002).

Collections from Hungary and Bosnia and Herzegovina described here represent populations of particular conservation or management interest. Specifically, population 24 from Hungary are mature trees comprising a relict population on an interriver floodplain where the annual flooding is now controlled through river canalization and engineering works. Trees exist as individual trees or small clusters within *Pinus* plantations and natural *Populus alba* – juniper mixed stands, with natural regeneration of *P. nigra* negligible. In this sense, the morphologic and genetic structure of the genets from this region represent a declining population that is unlikely to regenerate and persist as the decadent trees age. Similarly, the Mediterranean hairy subspecies is of particular conservation concern in Bosnia and Herzegovina given its distinct phenotype and fragmented population distribution (Kajba et al. 2004).

This study indicates that populations of the non-hairy *P. nigra* from central Europe (northern Italy, Hungary, and Bosnia and Herzegovina) display similar overall morphology in a glasshouse setting, as assessed by a discriminant analysis. The overall phenotype did not vary with conditions at their home site such as minimum monthly precipitation nor with maximum monthly temperature, indicating local adaptation (diversifying selection) may be minimal among these populations. Analysis of the genetic structure of these populations with molecular markers will be required to confirm the possibility that trees from this region represent a single evolutionary unit for conservation. Differences between the small-leaf and large-leaf phenotypes and among the small-leaf phenotypes from Spain and *P. nigra* ssp. *caudina* indicate local adaptation may justify restricting the movement of germplasm among divergent populations. Further, the potential for introgression of *P. deltoides*-like traits (lack of sylleptic branches, deltoid to cordate leaf base, young leaf color) from naturally-occurring hybrids to account for part of the observed morphological variation indicates care must be taken to screen germplasm stocks with genetic markers when initiating conservation, restoration, or breeding plans (Storme et al. 2004).

Ultimately, the threats to the survival and evolution of *P. nigra* in its native habitats (changes in flood regimes and land use along rivers, increase in mean annual temperature, greater potential for seasonal drought) will require coordination and management on a European scale rather than by individual countries. The application of novel and potentially controversial strategies such as human-mediated assisted migration (McLachlan et al. 2007) may be required to maintain healthy, self-regenerating stands of this ecologically important riparian tree.

Chapter 4. Proximo-distal variation in cell division and gene expression during leaf development

4.0 Overview

In order to determine whether cell growth (cell size) or cell division (cell number) drive the significant leaf size variation observed in *Populus nigra*, four genetic individuals (genets) were examined in a greenhouse experiment. The expansion rates and cell division patterns were assessed across the proximo-distal axis of the leaf surface. Significant variation in the leaf area, interveinal area, and relative expansion rates were identified among genets, across the leaf surface, and throughout leaf development, confirming the genetic basis to leaf morphology in this species. Cell size and cell number varied across the leaf surface and throughout leaf development, but only cell number varied among genets, indicating differences in cell division drive variation in leaf size in *P. nigra*. In order to assess the transcriptional differences underlying variation across the leaf lamina, a microarray experiment compared gene expression in three leaf segments: the leaf base, mid-lamina, and the leaf apex. Absolute differences in transcript levels were low, with ten gene models displaying at least a 4-fold change between leaf segments. Wilcoxon rank sum tests identified significant differences in regulation of genes with related function, with three patterns emerging. The first category includes genes related to energy production and substrate recovery that were upregulated in the base of the leaf. The second category included genes related to a variety of processes (e.g. glycolysis and lipid metabolism) that were down-regulated mid-leaf. Transcripts related to the light reactions of photosynthesis were the major exception to this category and were up-regulated mid-leaf. Third, genes related to carbohydrate and lipid metabolism cell division and DNA synthesis, signalling and transport were up-regulated in the leaf apex. This study demonstrates fine-scale control of cell division and compartmentalization of cellular processes that corresponds to the development of leaf size and shape in *Populus nigra.*

4.1 Introduction

Understanding the genetic control of leaf development is one of the fundamental questions of plant biology due to the organ's critical role as the site of photosynthesis and gas exchange and the link between leaf size and overall plant biomass. While leaf development is highly heritable, with plants of equivalent genetic backgrounds displaying little phenotypic variation when experiencing equal conditions (Horiguchi et al. 2006), leaf morphology is highly plastic in response to environmental variables (Taylor et al. 2003; Cookson et al. 2006; Street et al. 2006; Poorter et al. 2009). Both the genotypic and environmental components of leaf development relay adaptive components to leaf morphology, with leaf structure responding the biotic (e.g. herbivory) and abiotic (e.g. water availability) stresses.

Leaf development can be classified into two phases: leaf initiation, where cell fate is determined from pluripotent cells in the meristem, and leaf morphogenesis, where leaf structures develop and the leaf matures (Cnops et al. 2004). Leaf morphogenesis involves the coordination of three developmental processes, cell division, cell growth (expansion), and cell differentiation, in order to produce a mature leaf (Cnops et al. 2004). Three theories have been proposed to describe the role of cell growth in leaf morphogenesis. The cell theory holds that leaf size and shape is driven by cell division, with the number of cells determining leaf morphology (Fleming 2006a). Under the organismal theory of development, leaf morphology is determined at the organ-level, with cell division and cell growth coordinated to achieve a final leaf size (Tisné et al. 2008). The neo-cell theory states that both cell division and cell growth determine leaf morphology, with a partial compensation mechanism linking the two processes in order to govern leaf size (Tsukaya 2005). This compensation mechanism, where cell growth increases to compensate for reductions in cell division, has been proposed to explain the inverse relationship between cell size and cell number observed in several *Arabidopsis* leaf mutants (Horiguchi et al. 2006), and indicates that leaf morphogenesis is likely to involve coordination between cell division and cell growth (Tsukaya 2005; Tisné et al. 2008).

At the organ level, the epidermal layer of the developing leaf influences leaf morphogenesis. Chimeric tomato plants with genotypic differences between epidermal and mesophyll layers indicate that the genetic signals from epidermal cells can influence development of internal leaf tissues (Marcotrigiano 2010). The epidermal genotype of the developing leaf (or half-leaf) influenced cell division, but not cell size, of the underlying contrasting mesophyll layer (Marcotrigiano 2010). Similarly, restriction of cell elongation in the leaf margin affects leaf size and shape in *Arabidopsis* (Reinhardt et al. 2007). These results demonstrate that a short-range signaling mechanism may be involved in the organ-level control of leaf development, and that assessment of epidermal characteristics may reflect overall leaf development. Evidence of organ-level control of leaf morphogenesis does not imply that development is constant across the leaf lamina. In dicots, leaf development proceeds along the proximo-distal axis of the lamina, with cell division ceasing in the leaf apex before the base (Trainotti et al. 2004; Marcotrigiano 2010).

Most studies of fine-scale leaf development have focused on herbaceous annual species (Granier and Tardieu 1998; Trainotti et al. 2004; Tsukaya 2005). Identifying the developmental and genetic controls underlying leaf morphology in woody species is important due to the developmental differences between annual and perennial plants. *Populus* has become the model tree species due to its fast growth, ecological and economic importance, and widely available genomic tools (Taylor 2002). The relative contribution of cell division and cell growth to leaf development, and the resulting gradient across the lamina, varies among *Populus* species (Ridge et al. 1986). Assessment of leaf expansion in *Populus deltoides* failed to identify a base-tip gradient in leaf expansion (Matsubara et al. 2006), but a distinct basipetal gradient has been reported in *P. nigra* and *P.* x *euramericana* (Ferris et al. 2001). To date, investigations of the contribution of cell traits to leaf morphology have focused on differences among species or within hybrid families, not natural variation within collections of a single species.

The availability of genomic tools for *Populus* species has provided insight into the genetic mechanisms controlling leaf development in this genus. Assessment of leaves from a single genet (genotype) across the growing season and over two years indicated the transcriptome of leaf development to dominate gene expression patterns, with the progression of development tightly linked to temperature (Sjödin et al. 2008). This assessment provided a time-course of gene expression during leaf development, including markers distinguishing developmental phases (Sjödin et al. 2008). The transcriptome of leaf development in *Populus* has been further refined through the comparison of several microarray studies, which identified suites of genes associated with leaf tissue and stages of development (Street et al. 2008). Together with the availability of the *Populus trichocarpa* genomic sequence (Tuskan et al. 2006) and a microarray platform encompassing the majority of the *Populus* transcriptome (Anonymous 2005), *Populus* species prove a model system to investigate fine-scale control of leaf development and variation.

Populus nigra, the European black poplar, is of interest to breeding and management programs due to its high level of phenotypic diversity across Europe (Trewin et al., in preparation, Chapters 2 and 3). Many studies have focused on leaf development in *P. nigra* and its hybrids (e.g. *P.* x *euramericana*) in relation to bioenergy production given the link between leaf area and biomass (Al Afas et al. 2007). Relatively few studies have examined the cellular traits of *P. nigra* leaves (Ferris et al. 2001), and to date no study has reported variation in cellular characteristics among *P. nigra* collections or across the leaf lamina. Given the lack of information on leaf development in *P. nigra*, this study aimed to identify the developmental and genetic mechanisms underlying variation in leaf size by addressing three questions. First, does cell division or cell growth determine leaf size in *P. nigra*? Second, does relative expansion rates, cell size, or cell division vary across the lamina of developing leaves? Third, does the transcriptome vary across the leaf lamina, and might the pattern of gene expression correspond to developmental differences among leaf segments? These questions were assessed in four genets (genotypes) of *P. nigra* from natural populations across western and central Europe in a replicated greenhouse experiment. Results indicate that cell division, and not cell size, drives leaf morphology in *P. nigra*, and that relative growth rates vary across the leaf lamina. Assessment of transcriptional differences using Wilcoxon sign rank tests identified suites of genes differentially regulated across the lamina in newly unfurled leaves. Together, these findings provide insight into the developmental and genetic variation available for breeding and management purposes.

4.2 Methods

4.2.1 Greenhouse experimental design

In order to identify the developmental mechanism controlling variation in leaf morphology in *P. nigra*, genets from four diverse populations representing the three glacial refugia and one zone of introgression of refugial lineages were analyzed in a fine-scale leaf development experiment. Cuttings of 8 replicates of a single genet from Spain, France, Hungary and Italy were grown in the manner described in Section 3.2.1. The genets from Spain, France, and Italy were started from hardwood cuttings in 2008, while the samples from Hungary were received as hardwood cuttings in February 2009 and started in the greenhouse in March 2009. The plants were grown in a three-block design with three replicates of each genet randomized within each block, and one row of guard plants placed between blocks and around the entire design. These genets represent populations characterized in Chapter 2 and Chapter 3 (Table 4.1), and are the subject of ongoing physiological (M. Viger, PhD thesis, in preparation) and genomic resequencing (Evoltree/NovelTree/EnergyPoplar consortium) studies.

Table 4.1 Replicates of four genets (genotypes) were assessed in a glasshouse study of leaf development to assess the relative contribution of cell growth and cell division to morphological variation observed in *Populus nigra.*

Country	Population No.*	Genet ID	Glacial refugia
Spain	15	cart2	Iberian Peninsula
France	2	6J21	Admixture zone
Italy	8	N53	Italian Peninsula
Hungary	24	FTNY19	Balkan Peninsula

*Population numbers correspond to those listed in Chapters 2 and 3.

4.2.2 Quantifying differences in leaf expansion among genets

On the first day of the experiment (13 July 2009), the youngest completely unfurled leaf was marked with a piece of thin string to mark the first leaf of the study. Each morning a photograph was taken of this and all subsequent leaves against a 5 mm grid background using a Canon EOS 300D or a Canon PowerShot a710is digital camera. Photos were imported into ArcGIS v. 9.2 (ESRI, Inc., 2006) and geocorrected using bilinear interpolation in the Rectify feature of the Georeferencing extension. The Distance tool was then used to measure the overall leaf area (mm^2), and the interveinal area (mm^2) by tracing the primary and secondary viens for up to seven interveinal (i.v.) regions on one side of the leaf (Figure 4.1).

The relative expansion rate (RER) of each leaf and each interveinal area (mm^2 mm^{-2} day^{-1}) was calculated for each day as:

$$RER_i = \frac{Area_i - Area_{i-1}}{Area_{i-1}}, \tag{4.1}$$

where $Area_i$ is the area measured on day i and $Area_{i-1}$ is the area measured on day i-1. The mean daily relative expansion rate (mm^2 mm^{-2} day^{-1}) over the entire experiment was calculated for leaf area and each interveinal area as:

$$RER_{1-25} = \frac{Area_{25} - Area_1}{Area_1} \div 24, \tag{4.2}$$

where $Area_1$ is the area for day 1, $Area_{25}$ is the area measured on day 25, and the relative expansion rate is averaged over 24 days of observation. Differences between genets for interveinal area, and RER over each interveinal area were tested using the general linear model:

$$Y_{ijkl} = \alpha'_{ij}\left(\beta_j\right) | \delta_k | \lambda_l, \tag{4.3}$$

where Y_{ijk} is the measured trait, α_{ij} is the random factor of ramet i of genet j, β_j is the fixed factor of genet j, and δ_k is the fixed factor of leaf age, and λ_l is the fixed factor of interveinal region l. Due to varying leaf anatomy (the number of interveinal regions) or poor photos, data were unbalanced. Interveinal regions 6 and 7 were omitted from the analysis, and the remaining nulls (1% of data points) were interpolated for the GLM analyses. For a null in area at leaf age k, the mean of observations at ages (k-1) and

(k+1) for the same leaf was used. For a null in RER at leaf age k in interveinal region l, the mean of the other replicates (leaves) was used. General linear models were assessed using Minitab v 16 (Coventry, UK).

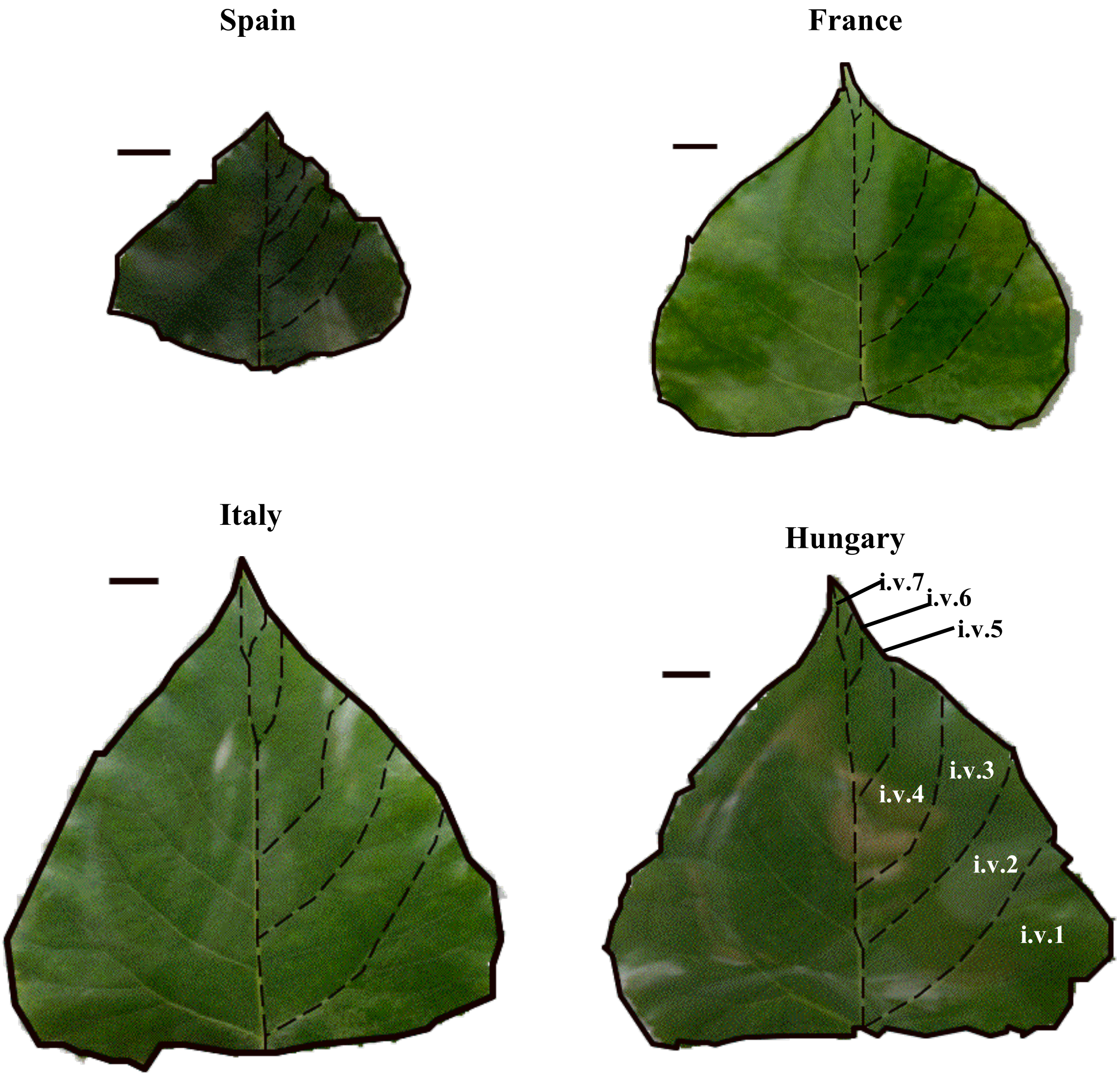


Figure 4.1 Photographs of expanding leaves taken during a greenhouse experiment were assessed to quantify overall leaf area (solid outlines) and the area of each interveinal region (dashed outlines) to determine if leaf expansion varies across the leaf surface or among genets of *P. nigra*. Interveinal regions (i.v.) were numbered from the base of the leaf to the tip (demonstrated on the Hungarian sample). Scale bars represent 1cm.

4.2.3 Quantifying differences in epidermal cell size in expanding leaves

On the final day of the experiment, cell imprints were taken from each interveinal region on two leaves on each of three ramets of each genet: the first fully unfurled leaf and the first fully mature leaf (identified by visual and physical inspection at the time of sampling). Imprints were taken and processed as described in Section 2.2.2, with the exception that imprints from the first fully unfurled leaf were imaged using oil-imersion objectives of either 63x/1.4 or 100x/1.4 magnification. Calculations of cell area and cell number per interveinal area were adjusted accordingly.

The average size of 10 epidermal (pavement) cells and number of cells per field of view were calculated for each sample. From these counts, the total numbers of cells per interveinal region and per leaf (twice the sum over interveinal regions) were calculated. Differences in each measure between genets and leaf age for the mean leaf values and each interveinal region were determined using the general linear model:

$$Y_{ijkl} = \alpha'_{ij}\left(\beta_j\right) \mid \delta_k, \tag{4.4}$$

where Y_{ijk} is the measured trait, α_{ij} is the random factor of ramet i of genet j, β_j is the fixed factor of genet j, and δ_k is the fixed factor of leaf age. Differences in cell area and cell number among interveinal regions were tested using the general linear model:

$$Y_{ijkl} = \alpha'_{ij}\left(\beta_j\right) \mid \delta_k \mid \lambda_l, \tag{4.5}$$

where Y_{ijk} is the measured trait, α_{ij} is the random factor of ramet i of genet j, β_j is the fixed factor of genet j, and δ_k is the fixed factor of leaf age, and λ_l is the fixed factor of interveinal region l. Due to the difficulty of making robust leaf imprints from small interveinal areas (leaf tips or young leaves), interveinal region 7 was omitted from the analysis, and 2% of data were missing in the final data set. As with RER data, null values in cell area and cell number at interveinal region l from leaf age k were replaced by estimates interpolated as the mean of replicate (ramet) observations for that genet. General linear models were assessed using Minitab v 16 (Coventry, UK).

4.2.5 Examination of transcriptome variation across the developing lamina

A microarray experiment was conducted to examine differences in gene expression across the lamina of a newly unfurled leaf. As the objective was to address variation within a developing leaf and not among genets, a single genet was examined in this experiment. The microarray experiment represents an initial study into variation in transcription across the leaf lamina and among large-leaf and small-leaf genets. The large-leaf genet from Italy (N53) was chosen for this experiment based on the high level of overall gene expression observed in genets from this population in previous microarry experiments (Trewin 2008, M. Viger, PhD thesis, in preparation). Further, the technical requirement of the quantity of RNA required for a microarray experiment prevented the use of the small-leaf genet from Spain. Future work will compare levels of transcription among large-leaf and small-leaf genets using qRT-PCR.

The experimental design described in Section 4.2.1 was maintained following the collection of leaf and cell data in 2009. Leaves were harvested for RNA analysis on 17 June 2010. In order to assess constitutive and developmental gene expression and not inducible gene expression resulting from environmental stresses, leaves were harvested at dawn, between 05:05 and 05:35. The youngest fully unfurled leaf was removed from the apical meristem by cutting the petiole away from the stem. While holding the leaf by the petiole, the leaf blade was then dissected into three segments from tip to base of the leaf (Figure 4.2). The first cut was made at the fifth vein, the second cut was made at the third vein, and the last cut was made at the base of the leaf to detach the remaining blade from the petiole. Immediately following each cut, segments were placed in 1.5 mL microcentrifuge tubes and flash-frozen in liquid nitrogen. Samples were stored at -70°C until analysis.

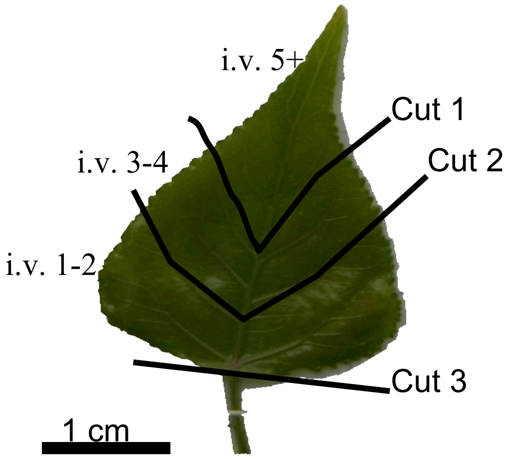


Figure 4.2 The youngest unfurled leaf of three ramets of genet N53 from Italy were dissected into three sections to assess variation in gene expression across the leaf lamina.

Total RNA was extracted using a small-volume CTAB protocol developed by N. Street and J. Tucker, adapted from Chang et al. (1993). Each leaf segment was ground to a fine powder under liquid nitrogen in a small mortar using a pestle. The frozen tissue was placed in a pre-frozen 2.0 mL screw-cap microcentrifruge tube, and remained frozen until the extraction buffer was added. After adding 900 μL of CTAB extraction buffer (2% hexadecyltrimethylammoniumbromide, 2% polyvinylpyrrolidone 40 K 29-32, 100 mM Tris-HCl pH 8.0, 25 mM EDTA, 2 M NaCl) (warmed to 65°C) and 50 μL of β-mercaptoethanol, samples were vortexed vigorously and incubated for 5 mins. at 65°C. An equal volume of CHISAM (24:1 chloroform: isoamyl alcohol) was added to each sample, and the tubes were vortexed vigorously to emulsify. Samples were centrifuged at 12,000 g at 18°C for 10 minutes in a bench-top centrifuge (Eppendorf). The aqueous (upper) phase was transferred to a new 2.0 mL screw-cap tube, to which ¼ volume of 10 M lithium chloride was added. Samples were mixed thoroughly and placed at 4°C overnight to maximize precipitation of RNA. The following day the samples were centrifuged at 12,000 g at 4°C for 15 mins. The supernatant was removed, pellets resuspended in 700 μL SSTE (1 M NaCl, 0.5% SDS, 10 mM Tris-HCl pH 8.0, 1 mM EDTA) warmed to 60°C, and samples incubated at 60°C for 5 mins. An equal volume of CHISAM was added to each sample and tubes were vortexed to emulsify. Samples were centrifuged at 12,000 g at 18°C for 10 mins., at which time the aqueous phase was transferred to a 1.5 mL tube. An equal value of 100% ethanol chilled to -20°C was added to each sample, tubes were shaken to mix, samples were incubated at -20°C for 10 mins. to precipitate RNA. Prior to a final wash, samples were centrifuged at 12,000 g at 4°C for 10 mins., following which the supernatant was removed and the pellet washed in 1 mL of 70% ethanol chilled to -20°C. The solution was centrifuged at 12,000 g at 4°C for 1 mins., the supernatant removed, and the pellet dried completely before being redissolved in 20 μL of DEPC treated RNAse-free water. RNA concentration was quantified using a Nanodrop spectrophotometer (Nanodrop Technologies, Wilmington, Delaware, USA) and quality assessed with a Bioanalyzer RNA Nano kit (Agilent 2100 Bioanalyzer, Agilent Technologies, Waldbronn, Germany).

RNA processing was completed using Affymetrix 3’ IVT Express Kit following the manufacturer’s instructions. Gene expression was assessed through sample hybridization to the Poplar Genome Array using Affymetrix technology and the GeneChip® Hybridization, Wash, and Stain Kit following the manufacturer’s instructions. The Poplar Genome Array is composed of over 56,000 expressed and predicted genes, covering the majority of the *Populus* transcriptome, represented in over 61,000 probe sets on each array (Anonymous 2005). Each probe is a 25-mer oligonucleotide segment of an expressed sequence. A total of nine hybridizations were conducted: three leaf segments for each of three biological replicates of newly unfurled leaves (Figures 4.3 and 4.4). The Nottingham Arabidopsis Stock Centre (NASC) International Affymetrix Service (http://affymetrix.arabidopsis.info/) completed all aRNA preparations and hybridizations.

Microarray data quality assessment and statistical analyses were conducted using Robin v1.1.5 for Windows XP, a graphical user interface developed to run a variety of analyses from the Bioconductor project (Gentleman et al. 2004) that operates in the R language for statistical computing (Lohse et al. 2010). Initial data quality was assessed qualitatively from the boxplot of the distribution of unnormalized expression levels for each of the nine arrays. In addition, the range of expression levels was assessed using boxplots of the normalized unscaled standard error (NUSE). High-quality data should show a median standard error around 1.0, with values exceeding this benchmark indicating low data quality. Data were normalized using the robust multi-array average method of (Irizarry et al. 2003), as implemented in Robin, which normalizes expression for each probe across all chips in the experiment. Gene fragments differentially expressed among leaf sections as described in Figure 4.3 were identified using linear model-based approaches implemented using the limma package (Smyth 2004). For all analyses, the Benjamini-Hochberg false discovery rate (FDR) control procedure was applied for an FDR of 0.2 (Benjamini and Hochberg 1995; Lohse et al. 2010), an acceptable value when changes in gene expression are expected to be slight as is the case in this assessment lacking experimental conditions (Leakey et al. 2009).

Log fold-change data for each paired comparison [$log_2(\mathrm{Array}_2/\mathrm{Array}_1)$] were calculated to determine which genes displayed an absolute fold change greater than 2.0, indicating variation in expression among segments. These ratios were then assessed in MapMan v3.1.1 to test for significant differences in regulation of suites of genes involved in similar biological processes (Thimm et al. 2004; Usadel et al. 2009). The Wilcoxon rank sum test was applied to test the null hypothesis that a BIN (functional class) of genes displayed a pattern of expression no different than the other BINs using a log fold-change of 1.0. The Benjamini-Hochberg false discovery rate procedure was applied with an FDR of 0.05. Differential regulation of genes contributing to a similar biological pathway or function across the leaf lamina may provide insight into the developmental mechanisms regulating differential growth and overall leaf size and shape.

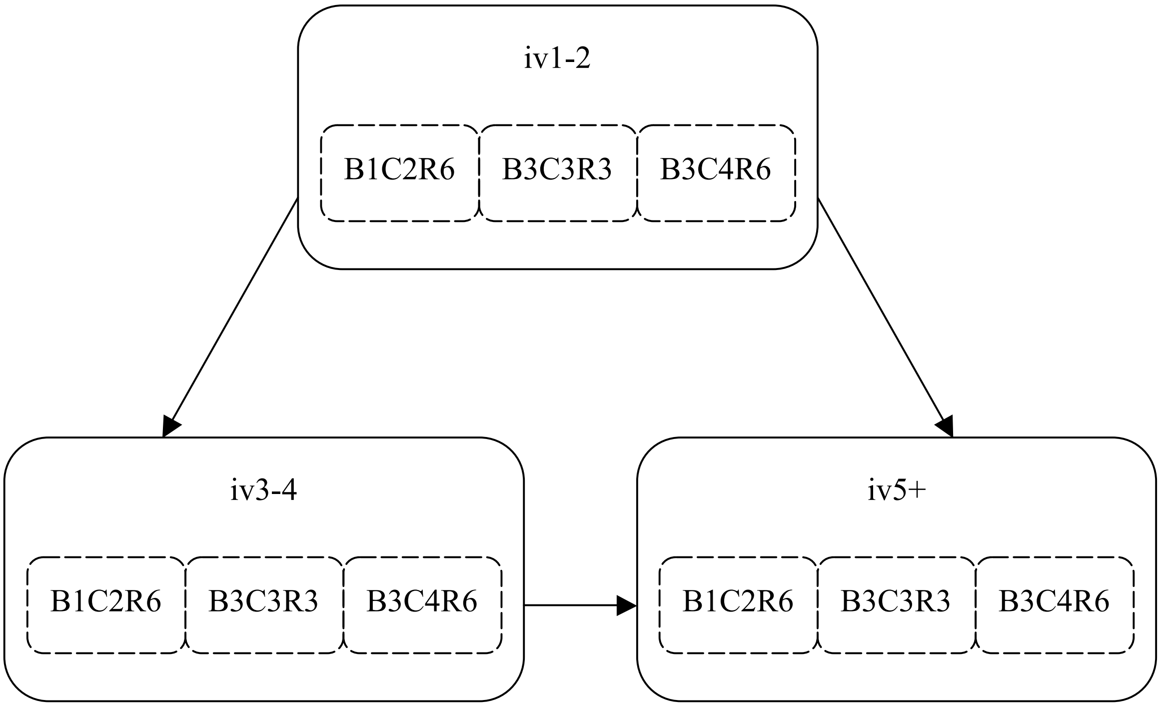


Figure 4.3 Schematic overview of microarray analyses conducted to identify gene fragments differentially regulated across the developing leaf lamina. Three biological replicates of genet N53 from Italy (B1C2,R6, B3C3R3, B3C4R6) were dissected into three leaf segments (iv1-2, iv3-4, iv5+). Gene expression was compared among leaf segments. In this model, the direction of the arrow indicates the subtraction of effects, e.g., ‘iv1-2 minus iv3-4’. Thus, genes downregulated in the base compared to the tip of the leaf will show a negative log_2 ratio. Genes upregulated in the base compared to the tip will show a positive log_2 ratio.

4.3 Results

4.3.1 Quantifying differences in leaf expansion among genets

Robust estimates of leaf size and interveinal area were obtained in the glasshouse experiment (Figure 4.4). The four genets displayed significant variation in growth patterns. Significant variation in overall leaf size was observed among genets ($F_{3,28} = 24.63$, $P<0.001$), over time (day since unfurling) ($F_{24,666} = 326.7$, $P<0.001$), and in the interaction between the two ($F_{72,666} = 22.40$, $P<0.001$). Differences in overall relative expansion rate, however, were non-significant among genets ($F_{3,28} = 1.49$, $P=0.239$), though significant differences were observed over time ($F_{23,632} = 1208$, $P<0.001$) and in the interaction Genet x Day ($F_{69,632} = 11.61$, $P<0.001$). These results indicate that the genets displayed significantly different leaf sizes and that they achieved the leaf sizes through different growth patterns as the leaf expanded. The lack of significant genet effect indicates newly unfurled leaves likely differ in size in proportion to the final leaf size, so that no difference in relative expansion was observed. Differences in leaf size varied among genets, across the leaf surface, and over the duration of leaf expansion, indicating final leaf shape to be determined by a complex genetically-determined developmental sequence (Table 4.2). The significant interaction between genet and day since unfurling indicates genets differ in the rate of leaf growth, which is apparent in the leaf expansion profiles (Figure 4.4). The Spanish genotype completed leaf expansion approximately four days earlier than the other genotypes, indicating the overall growth rate in this sample might be faster. Any effect of manipulating the developing leaves under transparencies for photography likely did not increase error significantly given the small variance in leaf size among replicates.

Relative expansion rate also displayed significant interaction between genet and day since unfurling, consistent with the differences in leaf expansion profiles, and indicated the growth rate varies among genotypes over the course of leaf expansion (Table 4.3). RER changed markedly over time, consistent with leaf growth slowing as the final size is achieved (Figure 4.4). Relative expansion rates varied across the leaf lamina, with the basal interveinal regions displaying greater expansion than the tip of the leaf (Figure 4.7). The significant interaction of genet x i.v. region indicated that the pattern of leaf expansion across the lamina varied among genets, consistent with the differences in leaf shape.

Table 4.2 Significant variation in the size of leaves and interveinal regions across the leaf lamina was detected among genets of *Populus nigra*.

Factor	**d.f.**	**MS**	***F***	***P***
Day since unfurling	24	5373373	288.5	<0.001
I.V. region	4	21830640	130.7	<0.001
Genet	3	30260127	20.7	<0.001
Day x i.v. region	96	268562	125.6	<0.001
Day x genet	72	357688	19.2	<0.001
Genet x i.v. region	12	2389905	14.3	<0.001
Genet x day x i.v.	288	30746	14.4	<0.001
Ramet´ (Genet)	28	1460852		
Day x Ramet´ (Genet)	672	18628		
I.V. x Ramet´ (Genet)	112	167025		
Error	2688	2139		
Total	3999			

Table 4.3 Significant variation in the relative expansion rate (RER) of leaves and interveinal regions was detected throughout the leaf development (days) and in the change in RER over time (genet x day), but not among genets.

Factor	**d.f.**	**MS**	***F***	***P***
Day since unfurling	23	3.29777	557.52	<0.001
I.V. region	4	0.18676	175.27	<0.001
Genet	3	0.01977	0.64	0.594
Day x i.v. region	92	0.01471	3.84	<0.001
Day x genet	69	0.03807	6.44	<0.001
Genet x i.v. region	12	0.00408	3.83	<0.001
Genet x day x i.v.	276	0.00501	1.31	0.001
Ramet´ (Genet)	28	0.03073		
Day x Ramet´ (Genet)	644	0.00592		
I.V. x Ramet´ (Genet)	112	0.00107		
Error	2576	0.00383		
Total	3839			

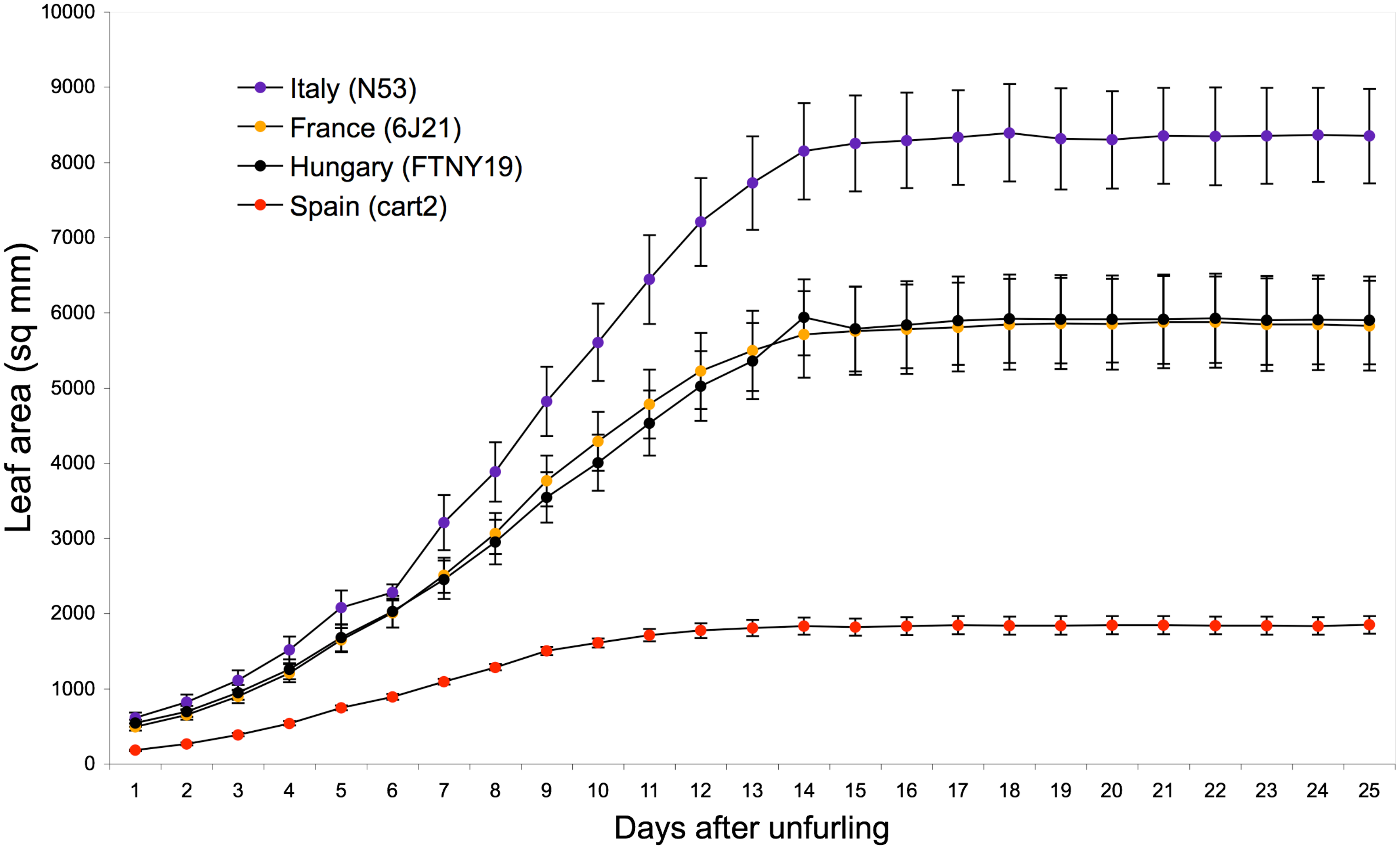


Figure 4.4 Leaf expansion profiles for four genets of *Populus nigra*. Points indicate the mean over eight replicates. Error bars represent the standard error of the mean.

4.3.2 Quantifying differences in epidermal cell size in expanding leaves

As cell growth and cell division interact in leaf development, cell size and cell number were assessed in young (newly unfurled) and fully expanded (mature) leaves from three replicates of each genet (Figure 4.5). The greatest differences in cell area were observed between leaf ages (Table 4.4). Significant variation was observed among interveinal regions, indicating that cell size may influence variation in leaf shape across the lamina. The significant interactions of genet x leaf age and of genet x leaf age x i.v. region indicated that cell growth may vary among genets (Table 4.4). Estimates of cell area from the newly unfurled leaves, and the statistical analyses of those data, may be subject to higher levels of experimental bias than those of mature leaves. The newly unfurled leaf displayed highly variable cell sizes within and between interveinal regions, evidently due to ongoing cell division in the leaf lamina along side expanding cells in developing stomata and veins (Figure 4.6). This variation likely increased the likelihood of experimental error in assessing biologically equivalent samples.

Differences in cell number were significant among genets and in the genets x leaf age interaction, and highly significant among all other factors (Table 4.4). These differences indicate genets displayed different patterns of cell division. Cell number varied across the leaf lamina differentially among genets and leaf ages (significant interaction effects), consistent with the increase in leaf area during leaf development and the unequal size of interveinal regions (Table 4.2).

The lack of significant difference in average cell size among genets and significant variation in cell number among genets and interveinal regions provides indirect evidence that differences in leaf area are driven by differences in cell number, i.e. cell division, and not cell growth.

Table 4.4 Cell size and number varies across the leaf lamina (interveinal regions, i.v.) and between newly unfurled and fully expanded leaves in four genets of *Populus nigra*.

	Trait:	**Cell area**			**Number Cells**		
Factor	**d.f.**	**MS**	***F***	***P***	**MS**	***F***	***P***
Leaf age	1	6258135	1482.8	<0.001	$2.73579\text{x}10^{14}$	75.2	<0.001
I.V. region	5	20326	5.7	<0.001	$2.93983\text{x}10^{13}$	26.1	<0.001
Genet	3	31156	2.1	0.177	$2.57654\text{x}10^{13}$	7.1	0.012
Leaf age x i.v.	5	5926	1.4	0.253	$2.87157\text{x}10^{13}$	25.5	<0.001
Genet x Leaf age	3	23491	5.6	0.023	$2.49816\text{x}10^{13}$	6.9	0.013
Genet x i.v.	15	6566	1.8	0.065	$5.47834\text{x}10^{12}$	4.9	<0.001
Genet x i.v. x leaf age	15	9615	2.2	0.022	$5.37963\text{x}10^{12}$	4.8	<0.001
Ramet' (Genet)	8	14728			$3.64597\text{x}10^{12}$		
Ramet' (Genet) x leaf age	8	4221			$3.63700\text{x}10^{12}$		
Ramet' (Genet) x i.v.	40	3594			$1.12639\text{x}10^{12}$		
Error	40	4300			$1.12642\text{x}10^{12}$		
Total	143						

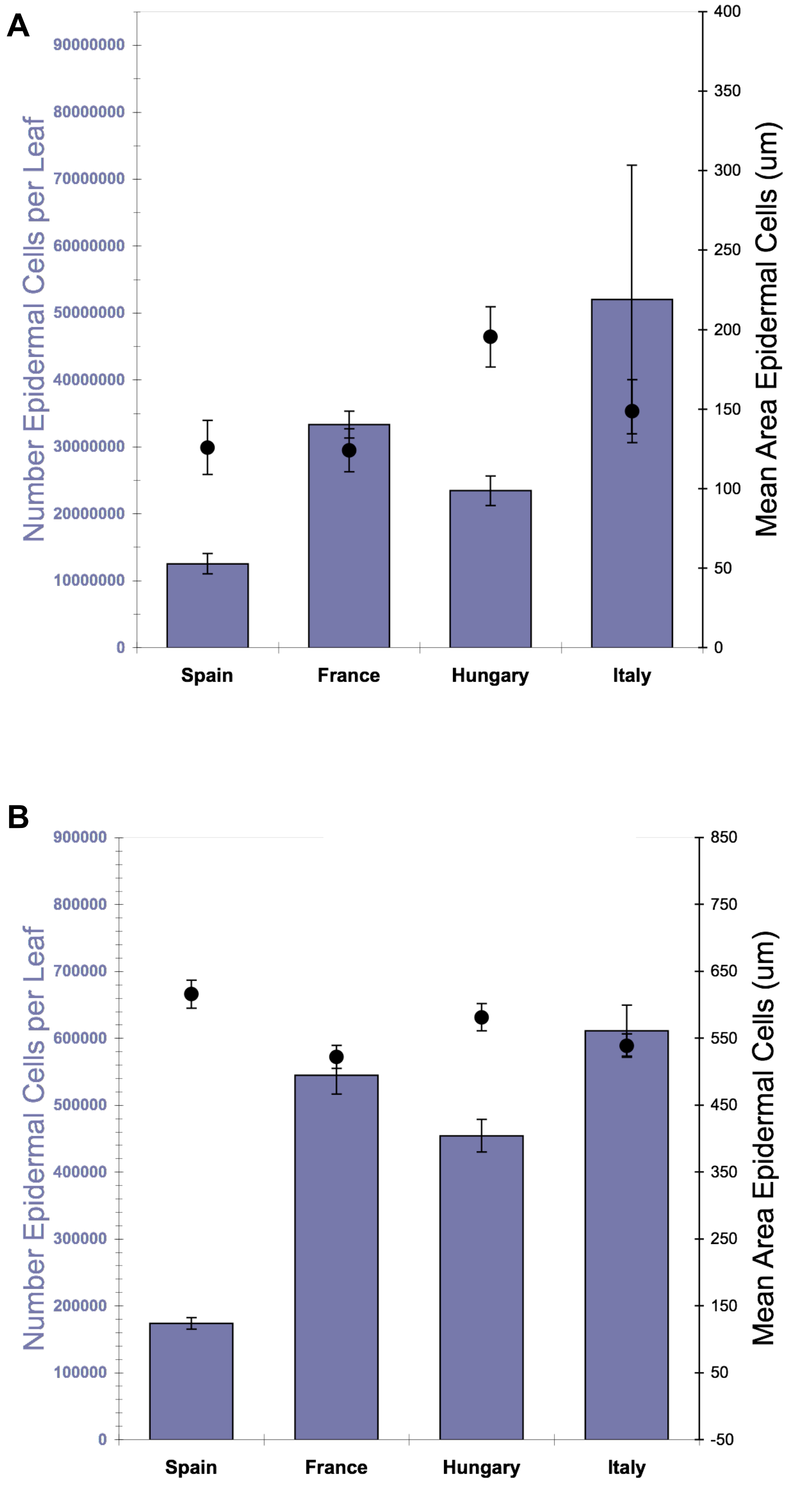


Figure 4.5 Cell number and size varies among genets of *Populus nigra*. The number of epidermal cells per leaf (leaf axis) was calculated from the individual interveinal region data. Average cell size was estimated from 10 epidermal cells. (A) Measurements from the youngest unfurled leaf. (B) Data for fully-expanded, mature leaves. Note different scales.

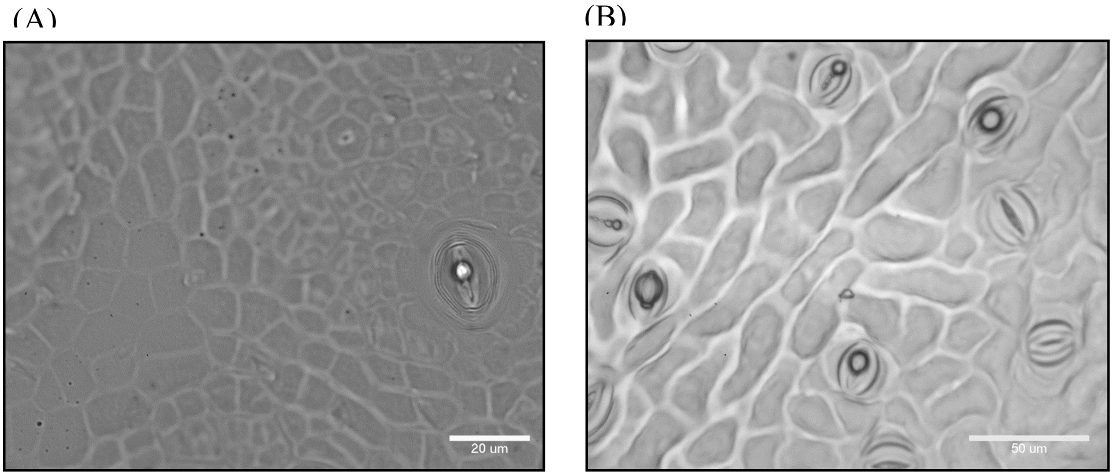


Figure 4.6 Cell imprints taken from newly unfurled (A) and fully expanded (B) leaves of a *Populus nigra* genet from Italy (N53). Cells in newly unfurled leaves are likely undergoing cell division, making estimates of mean cell size and cell number highly variable among samples. Note the difference in scale.

Figure 4.7 (following pages) Patterns of leaf and cell growth in four genets of *Populus nigra*. Each plate represents data for one genet, labelled at top. (A) Heat maps of relative expansion rates (RER) across the leaf lamina, at three timepoints and over the duration of the experiment, with colors representing RER for each interveinal region. (B) Interveinal region expansion profiles, with comparison to the overall leaf expansion profile (inset). (C) Variation in the average cell number per interveinal area (bars, left axis) and average epidermal cell area (lines, right axis) across the leaf lamina and among leaf ages. Note the difference in scales between the first unfurled leaf (left) and the fully expanded leaf (right) graphs. The number of cells per region is calculated from a sample in the center of the interveinal area, likely inflating estimates due to the small size of cells undergoing cell division in young leaves.

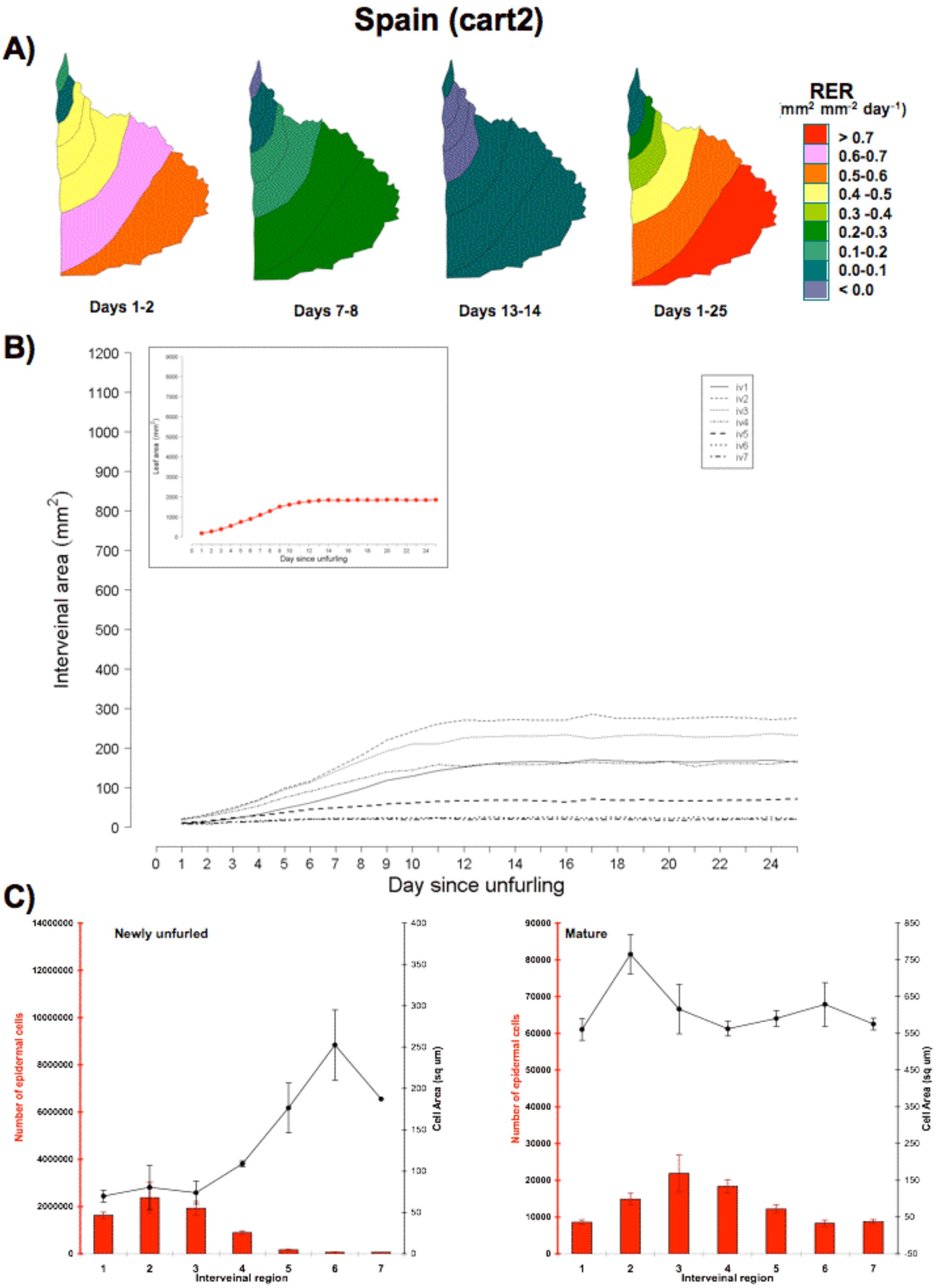

Spain (cart2)
A)
Days 1-2
Days 7-8
Days 13-14
Days 1-25
RER
(mm² mm⁻² day⁻¹)
🟥 > 0.7
0.6-0.7
0.5-0.6
0.4 -0.5
0.3 -0.4
0.2-0.3
0.1-0.2
0.0-0.1
< 0.0
B)
Interveinal area (mm²)
1200
1100
1000
900
800
700
600
500
400
300
200
100
0
Day since unfurling
0
1
2
3
4
5
6
7
8
9
10
12
14
16
18
20
22
24
Leaf area (mm²)
Day since unfurling
iv1
iv2
iv3
iv4
iv5
iv6
iv7
C)
Newly unfurled
Number of epidermal cells
14000000
12000000
10000000
8000000
6000000
4000000
2000000
0
Cell Area (sq um)
400
350
300
250
200
150
100
50
0
Interveinal region
1
2
3
4
5
6
7
Mature
Number of epidermal cells
90000
80000
70000
60000
50000
40000
30000
20000
10000
0
Cell Area (sq um)
850
750
650
550
450
350
250
150
50
-50
Interveinal region
1
2
3
4
5
6
7

Figure 4.7

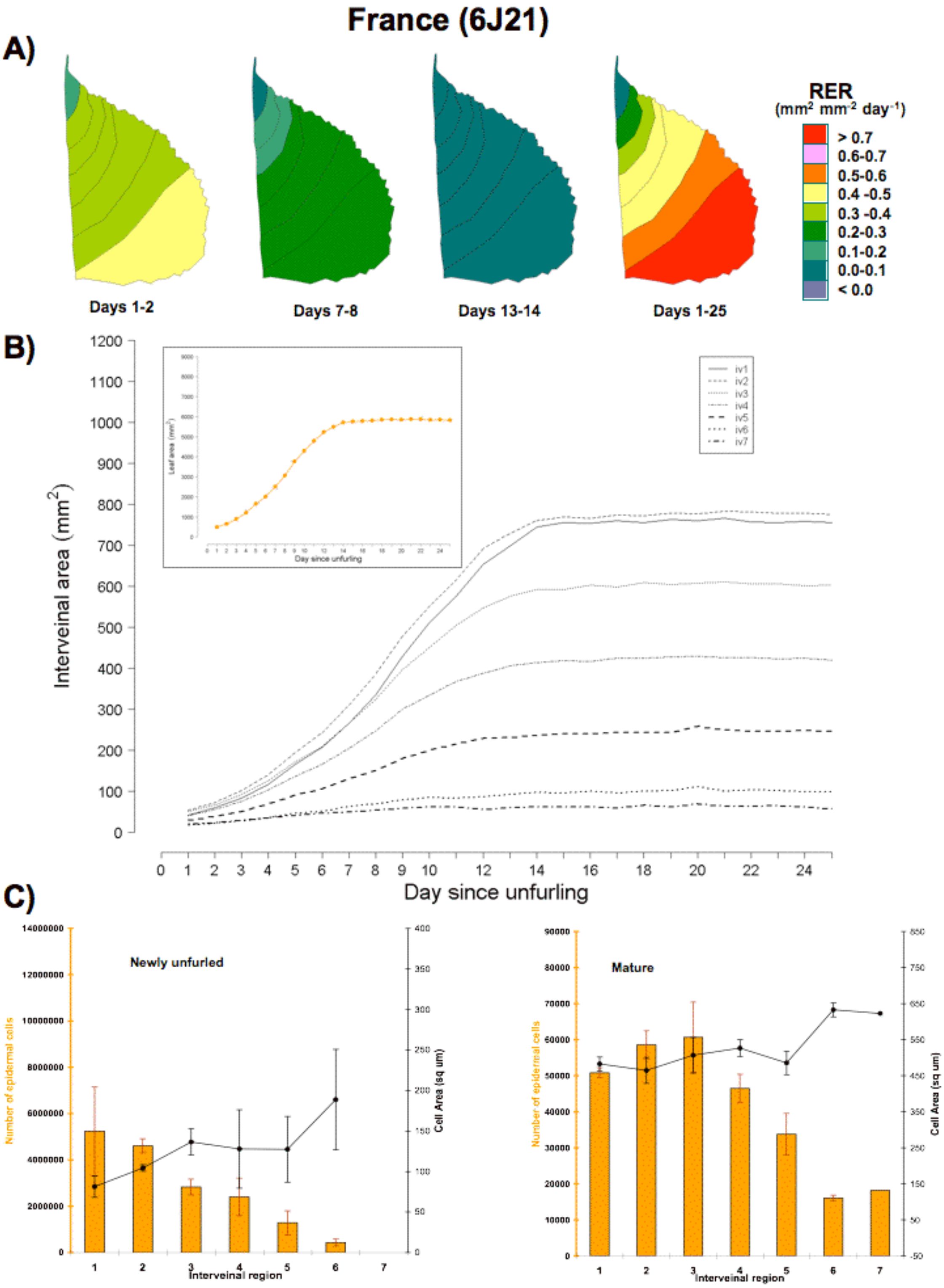


Figure 4.7, continued.

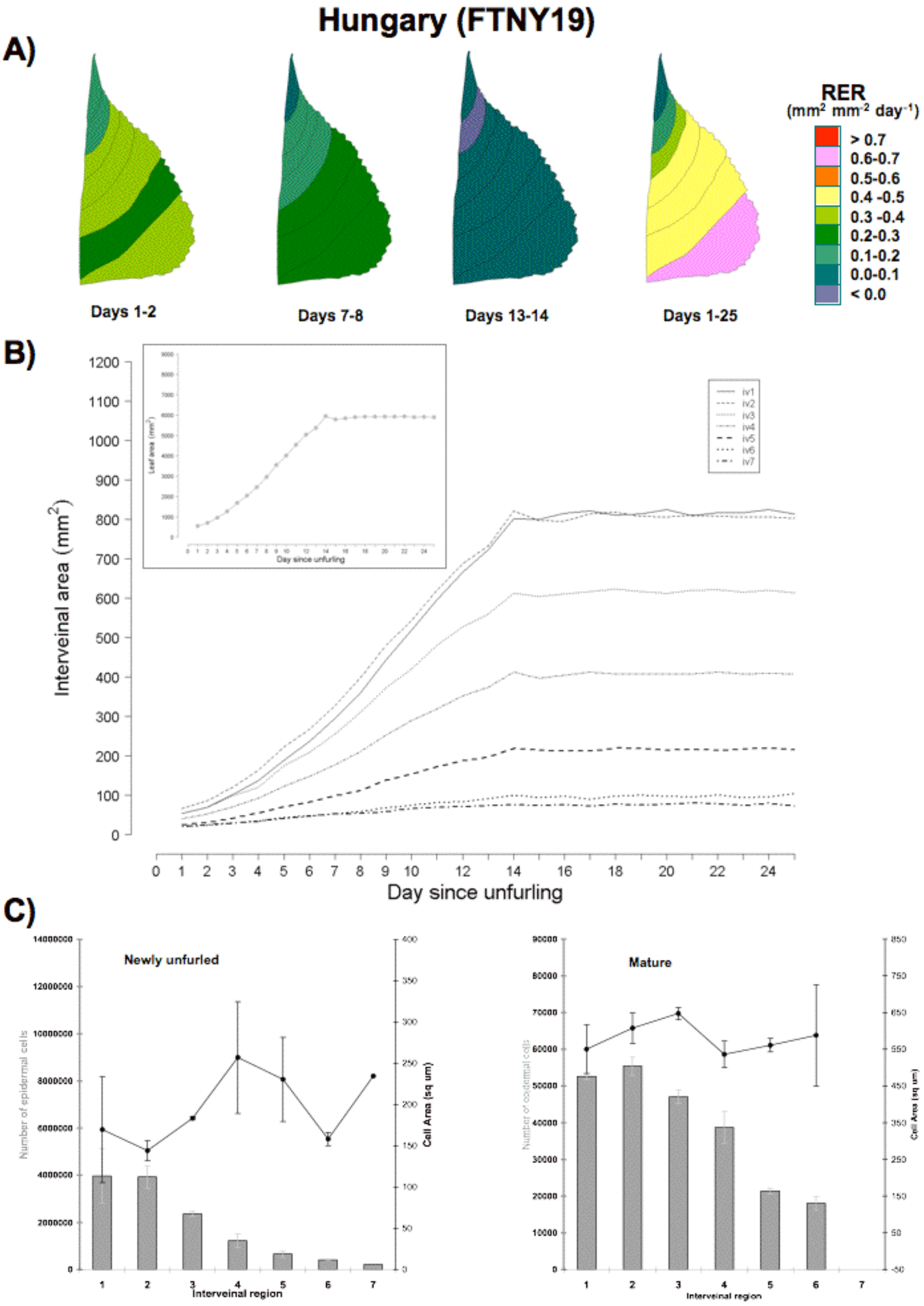


Figure 4.7, continued.

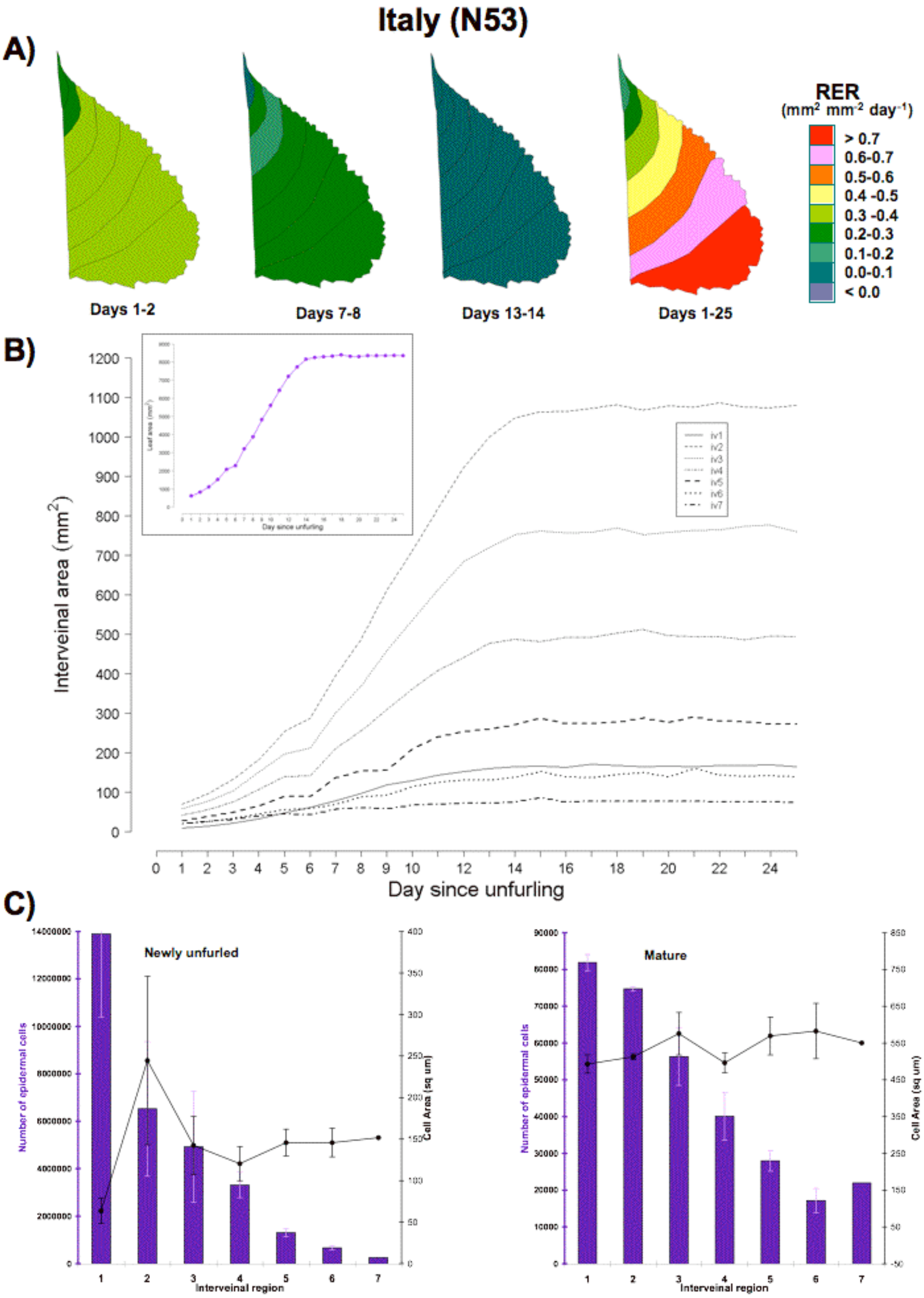


Figure 4.7, continued.

4.3.3 Relative contribution of cell size and cell division to leaf morphology

Differences in leaf area among genets and among interveinal regions within leaves are due to differences in cell number rather than differences in cell size (Figure 4.7). The base of leaves displayed greater relative expansion throughout the experiment than did the tips of the leaves (Figure 4.7 parts A). The area of the upper interveinal regions of the leaf (iv6 and iv7) are consistent among genets, indicating differences in regions iv1 through iv5 drive the significant variation in overall leaf area (Figure 4.7 parts B). However, though cell area varied significantly among interveinal regions, inspection of the data for each genet indicated the size of cells from the small-leaf genet from Spain (cart2) likely drove this trend (Figure 4.7 parts C). No evidence of differences in cell area across fully expanded mature leaves was found in the large-leaf trees from Italy (N53) or Hungary (FTNY19). The lack of difference in cell size and highly significant difference in cell number indicates that the high rates of relative expansion observed at the leaf base are driven by differences in cell division and not cell growth during leaf expansion.

4.3.4 Examination of transcriptome variation across the developing lamina

Given the variation in cell division observed across the leaf lamina, a microarray experiment was conducted to survey differences in gene expression across newly unfurled leaves. The nine arrays (three biological replicates, B1C2R6, B3C3R3, and B3C4R6, each having three leaf segments: i.v.1-2, i.v. 3-4, and i.v.5+) displayed consistent high-quality unnormalized data (Figure 4.8). Multivariate assessment using principal components analysis did not reveal similarities in expression patterns between genets or leaf segments, indicating either a low percentage of genes were differentially expressed, or that those under differential regulation were affected by unmeasured or unexplained variables (Figure 4.9).

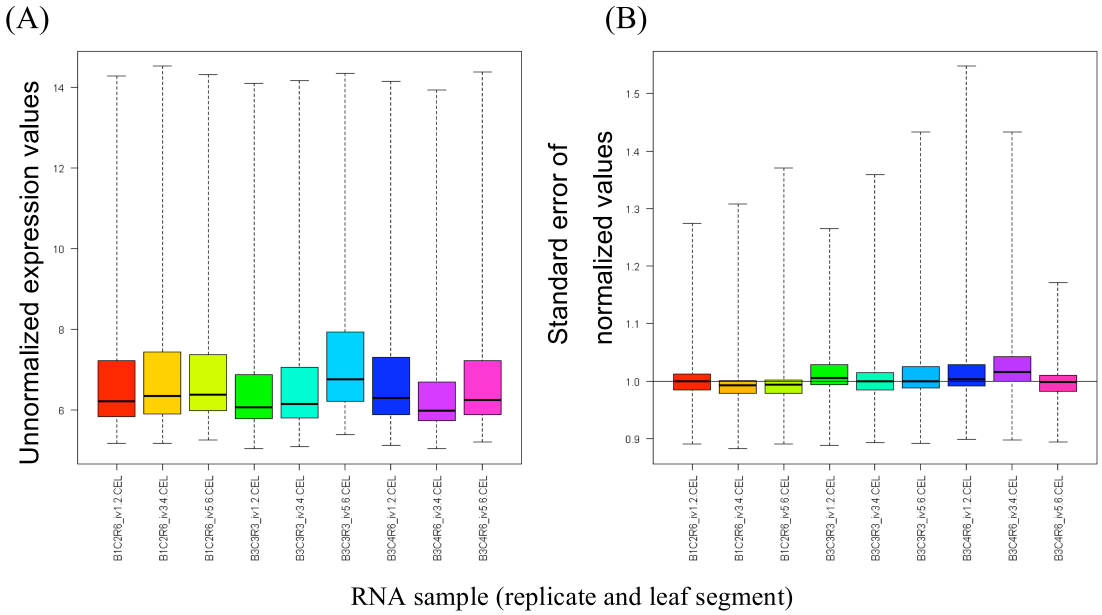


Figure 4.8 Microarray analysis of gene expression in nine samples of *Populus nigra* resulted in high-quality data from each chip. Each boxplot represents the array for one of three leaf segments (i.v.1-2, i.v.3-4 or i.v.5+) from one of three biological replicates of the Italian genet N53 (B1C2R6, B3C3R3, or B3C4R6). (A) Unnormalized expression levels detected for each array. (B) Normalized unscaled standard errors (NUSE) of expression levels for each array.

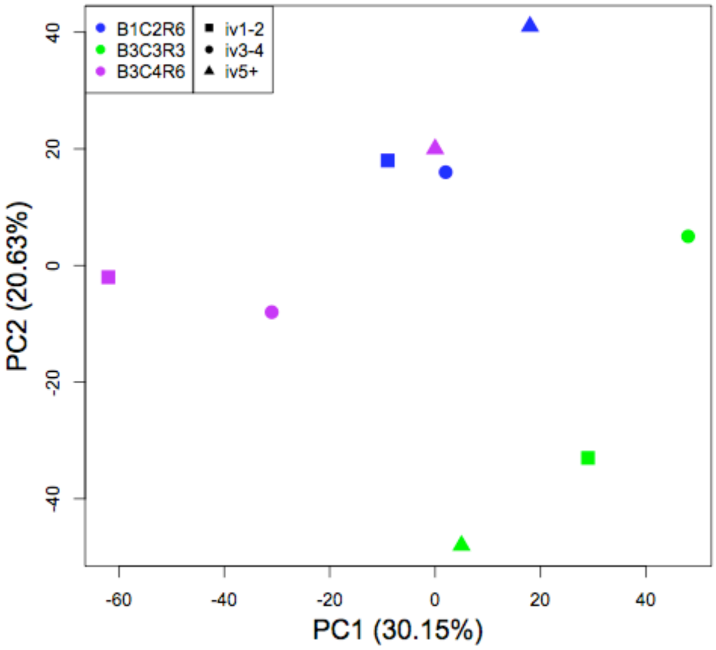


Figure 4.9 Principal components analysis of gene expression did not resolve similarities between genets or leaf segments. The first two components explained a similar percentage of variation in the observed gene expression (30.15 and 20.63). Points represent individual arrays, with samples designated by biological replicate by color (B1C2R6, B3C3R3, and B3C4R6) and by leaf segment by shape (iv1-2, iv3-4, iv5+).

Tests of differential gene expression using linear model-based methods failed to identify gene products significantly differentially regulated across the leaf lamina at a false discovery rate of 0.25. Six genes displayed $|log_2(Array_2/Array_1)| > 2$ in at least one comparison, indicating these genes likely vary in concentration among three leaf segments (Table 4.5). While most of the 13 gene models detected in the log_2FC assessment are of unknown function, four transcripts were annotated when interrogated on the *Populus trichocarpa* v2.0 genomic sequence (http://www.phytozome.net/poplar.php). Gene model POPTR_0002s00350.1 showed greater expression in the leaf base (i.v.1-2) compared to the apex (i.v.5+) and shows structure indicative of calcium ion binding. The increased expression of gene model POPTR_0006s14090.1 in the base compared to the apex was detected on two probes. This gene model contains an AP2/ERF domain characteristic of a transcription factor. Gene model POPTR_0013s12390.1 was down regulated in the base of the leaf compared to the tip, and encodes a β-ketoacyl-coenzymeA (coA) synthase protein. Finally, gene model POPTR_0017s02470.1, which encodes a hypothetical protein similar to a U-box domain-containing protein, was up regulated in the leaf base compared to the leaf apex.

Table 4.5 Thirteen probes indicated ten gene models were differentially expressed across the lamina of newly unfurled leaves of *Populus nigra*. Differential expression was identified from assessments of log_2 fold change ratios between leaf segments.

Identifier*	**Comparison**	***log_2* FC**[†]	**Gene Model**[‡]	**LG**[‡]	**Position (bp)**[‡]	**Details**
PtpAffx.158679.1.S1_at	iv1→iv5	2.20	POPTR_0002s00350.1	2	93058-94082	Hypothetical protein; similar to *Arabidopsis* AT5G42380.1: CML37 (CALMODULIN LIKE 37); calcium ion binding
PtpAffx.24353.1.A1_a_at	iv1→iv5	-2.39	POPTR_0005s21240.1(11)	5	20368536-20369153	Unknown function
	iv3→iv5	-2.64				
PtpAffx.24353.2.A1_x_at	iv1→iv5	-2.62	POPTR_0005s21240.1	5	20368536-20369153	Unknown function
	iv3→iv5	-3.11				
PtpAffx.249.75.A1_a_at	iv1→iv5	2.21	POPTR_0005s09450.1	5	6588192-6588772	Hypothetical protein
PtpAffx.249.75.A1_x_at	iv1→iv5	2.17	POPTR_0005s09450.1	5	6588192-6588772	Hypothetical protein
PtpAffx.155868.1.S1_s_at	iv1→iv5	2.14	POPTR_0006s14090.1	6	11398834 - 11399764	AP2/ERF domain-containing transcription factor

Identifier*	**Comparison**	***log*$_2$ FC**[†]	**Gene Model**[‡]	**LG**[‡]	**Position (bp)**[‡]	**Details**
PtpAffx.219707.1.S1_at	iv1→iv5	2.49	POPTR_0006s14090.1	6	11398834 - 11399764	AP2/ERF domain-containing transcription factor
PtpAffx.35351.2.S1_a_at	iv1→iv3	2.64	POPTR_0009s11760.1	9	9698962-9699749	Hypothetical protein
Ptp.1037.1.S1_s_at	iv1→iv5	-2.64	POPTR_0013s12390.1	13	13028915-13030756	Beta-ketoacyl-coa synthase family protein; acyltransferase/ catalytic/ transferase, transferring acyl groups other than amino-acyl groups
PtpAffx.67952.1.A1_at	iv1→iv5	2.08	POPTR_0017s02470.1	17	1781594-1782080	Hypothetical protein; similar to U-box domain-containing protein; similar to immediate-early fungal elicitor protein
Ptp.4450.1.S1_at	iv1→iv3 iv1→iv5	2.54 2.74	NA	NA	NA	Unknown function
PtpAffx.105254.1.S1_at	iv1→iv5	2.05	NA	NA	NA	Unknown function
PtpAffx.48329.2.S1_at	iv1→iv5	2.34	NA	NA	NA	Unknown function

*Affymetrix probe identifier

[†] *log*$_2$ fold change in expression, calculated as *log*$_2$(Array $_2$/Array$_1$). Genes down regulated in iv1 have a negative value; genes up regulated in iv1 have a positive value.

[‡]Gene model, linkage group (LG), position, and functional details from *P. trichocarpa* genome release 2.0 (http://www.phytozome.net/poplar.php). NA = not available.

In order to determine if groups of genes with similar functionality displayed similar patterns of expression across the segments of a leaf, even without significant fold-change differences in expression, the Wilcoxon rank sum test was applied to categorical BINs of gene models (Figure 4.10). Of the 233 secondary BINs identified in MapMan, 61 displayed significant trends in gene expression across the leaf lamina after corrections for multiple testing. Patterns of differential expression across the leaf lamina were grouped into four categories (Table 4.6). BINs in Category 1 displayed significantly more genes with increased expression in the leaf base (i.v. 1-2). Category 2 included BINs with a significant number of genes having decreased expression in the middle lamina (i.v. 3-4). BINs in Category 3 had significantly more genes up regulated in the apex of the leaf (i.v. 5+). BINs without a clear trend in expression were classified in Category 4 (Table 4.6).

The BINs containing a greater number of genes up regulated in the leaf base (i.v. 1-2, Category 1) are generally involved in energy production and substrate recovery (Table 4.6). Genes involved in the tricarboxylic acid cycle (TCA) organic acid transformation (BINs 8.1 and 8.2) and mitochondrial electron transport / ATP synthesis, including cytochrome c reductase (BIN 9.5), cytochrome c (BIN 9.6), and F1-ATPase (BIN 9.9), had increased expression in the leaf base, indicating genes related to energy metabolism to be differentially expressed in this leaf section. Genes involved in nitrogen degradation (BIN 12.3) and amino acid degradation (BIN 13.2) were also up regulated in the base of the leaf, indicating the cells are regenerating simple substrates. Genes related to three hormones were up regulated in the basal section: ethylene (BIN 17.5), salicylic acid (BIN 17.8), and, to a lesser extent, jasmonate (BIN 17.7), indicating hormone signals or stress response to be differentially controlled across the lamina.

The second category of differentially regulated genes displayed a significant number of elements down regulated in the middle section of the leaf (i.v. 3-4, Table 4.6). A variety of cellular functions displayed this pattern of expression, indicating the mid-lamina to have reduced transcription (BINS 27.1, 27.3, 27.4), amino acid activation (BIN 29.1), protein manipulation (BINs 29.3 and 29.6), cell wall precursor (BIN 10.1) and lipid (BINs 11.1 and 11.6) metabolism, and carbohydrate-active enzymes, including UDP glucosyl and glucoronyl transferases (BIN 26.2) and ß-1,3 glucan hydrolases (BIN 26.4). Together, this pattern indicates the mid-lamina to have reduced gene expression in a variety of functions, and may indicate an overall reduction in metabolism compared with the basal and apical sections.

(A) i.v. 1-2 → i.v. 3-4

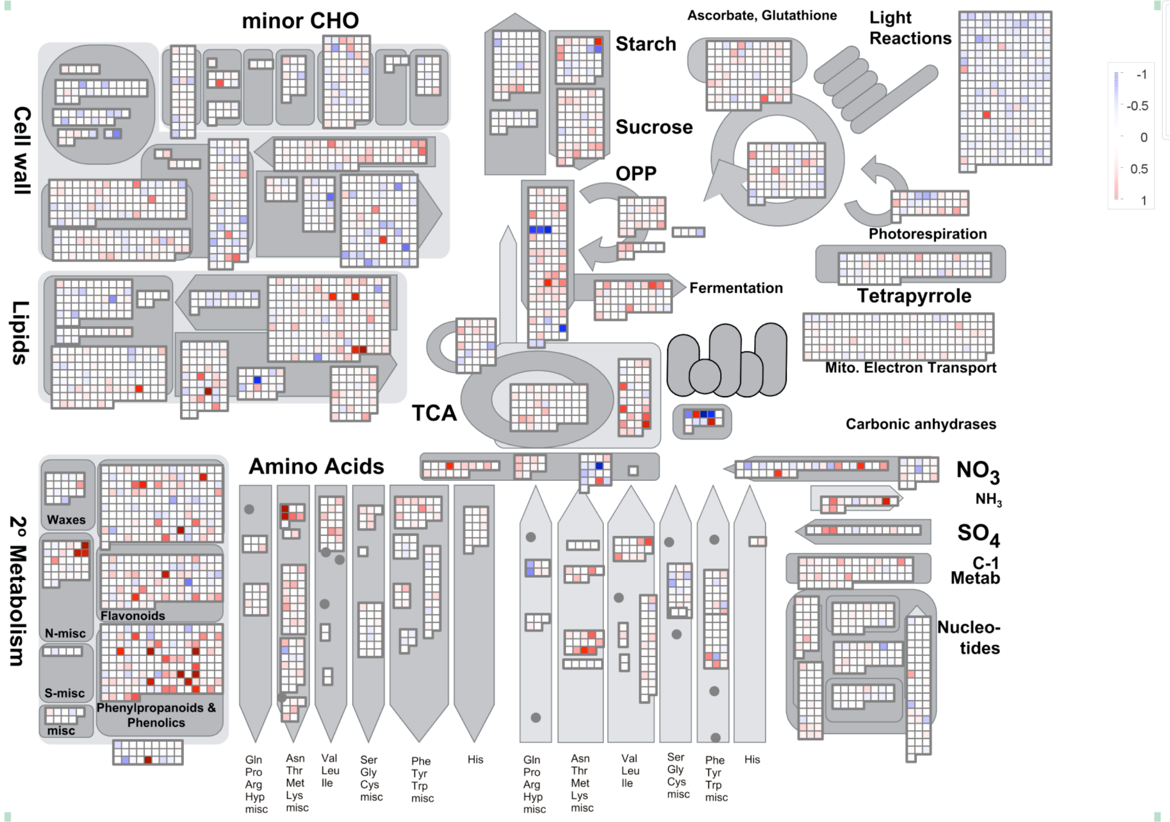


Figure 4.10 Overview of transcriptional differences between three segments of newly unfurled *P. nigra* leaves. Differences are presented as relative fold-changes in expression between (A) the base (i.v.1-2) and middle (i.v. 3-4) sections; (B) the middle (i.v.3-4) and apex (i.v.5+); and (C) the base (i.v. 1-2) and apex (i.v.5+) of the leaf. Each square represents a single gene model, with the color representing the change in expression. Negative (blue) values indicate lower expression in the more basal leaf segment; positive (red) values indicate greater expression in the more basal segment.

(B) i.v. 3-4 → i.v. 5+

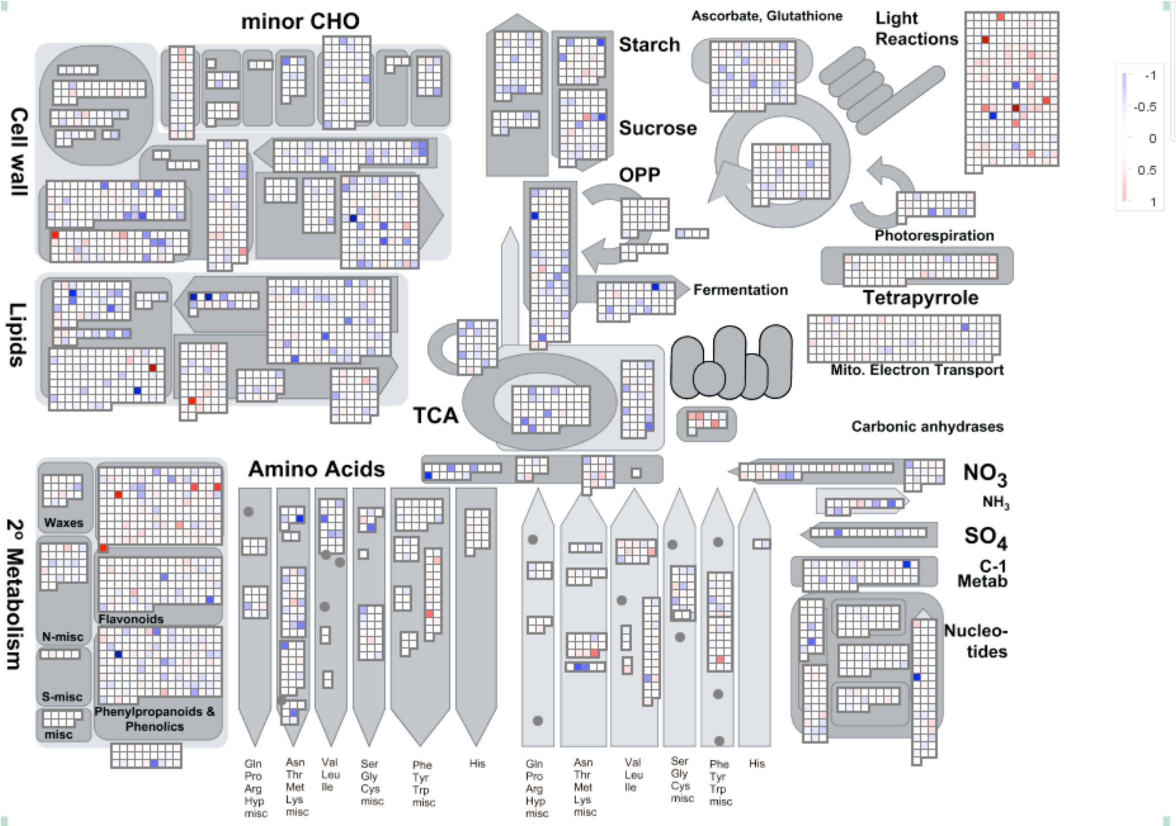


Figure 4.10 Continued.

(C) i.v. 1-2 → i.v. 5+

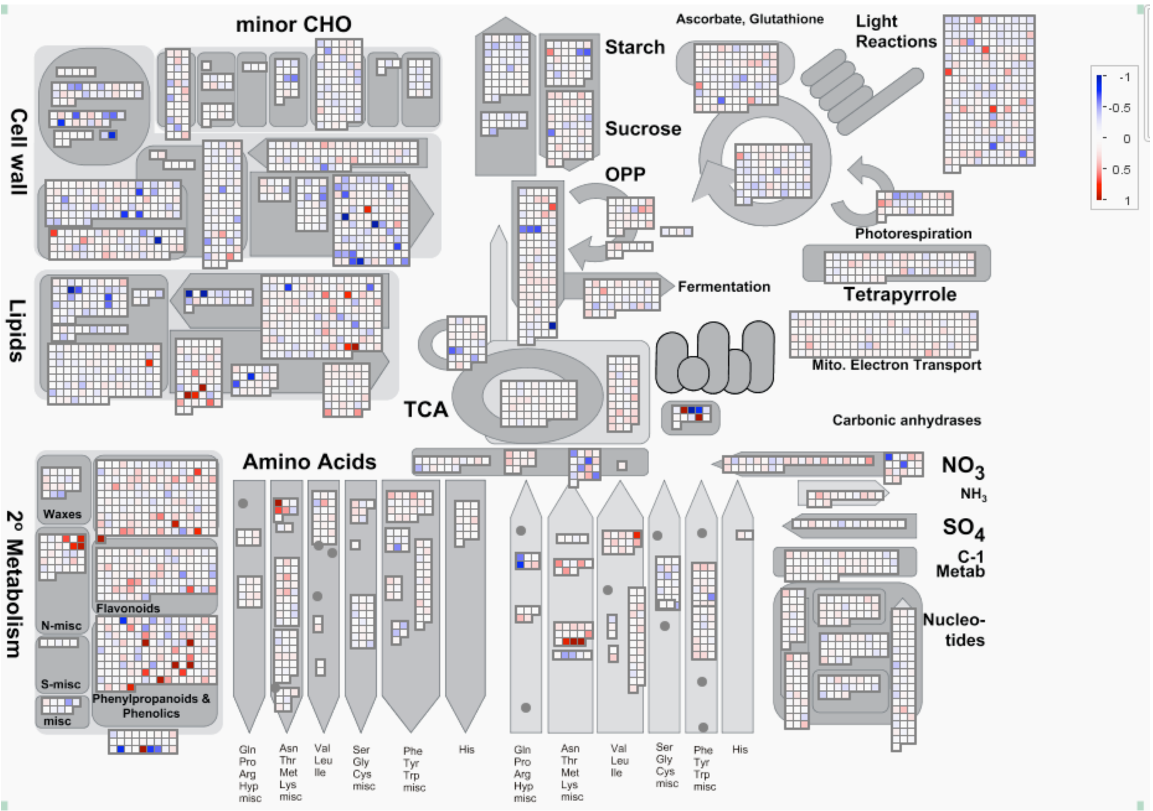


Figure 4.10 Continued.

Table 4.6 Differential regulation of functional categories (BINs, composed of a number of elements or probes) of genes across newly unfurled leaves of *P. nigra*. Wilcoxon rank sum tests identified non-random patterns of expression between leaf segments: i.v. 1-2 (base), i.v.3-4 (middle), and i.v.5+ (apex). Significant P-values are **bold**. The direction of gene expression is provided for significant comparisons: (+) = more elements up regulated in basal leaf segment; (-) = more elements down regulated in basal segment.

		Wilcoxon rank sum test P-value		
BIN	N elements	i.v.1 → i.v.3	i.v.3 → i.v.5	i.v.1 → i.v.5
Category 1: increased expression in leaf base				
8.1 TCA / org. transformation.TCA	55	**0.013 (+)**	0.352	0.463
8.2 TCA / org. transformation, other organic acid transformations	39	**0.005 (+)**	0.353	**0.047 (+)**
9.5 Mitochondrial electron transport/ATP synthesis, cytochrome c reductase	11	0.155	0.906	**0.027 (+)**
9.6 Mitochondrial electron transport/ATP synthesis, cytochrome c	15	**0.028 (+)**	0.311	0.556
9.9 Mitochondrial electron transport/ATP synthesis, F1-ATPase	26	0.067	0.494	**0.011 (+)**
12.3 N-metabolism, N-degradation	13	**0.023 (+)**	0.437	0.055
13.2 Amino acid metabolism degradation	153	**<0.001 (+)**	0.972	**0.001 (+)**
16.2 Secondary metabolism, phenylpropanoids	147	**0.004 (+)**	0.174	0.071
17.5 Hormone metabolism, ethylene	215	**0.016 (+)**	0.758	**0.008 (+)**
17.7 Hormone metabolism, jasmonate	40	0.268	0.522	**0.007 (+)**

		Wilcoxon rank sum test P-value		
BIN	N elements	i.v.1 → i.v.3	i.v.3 → i.v.5	i.v.1 → i.v.5
17.8 Hormone metabolism, salicylic acid	22	**0.028** (+)	0.443	**0.041** (+)
23.1 Nucleotide metabolism, synthesis	59	**0.002** (+)	0.231	0.269
26.21 Misc., protease inhibitor/seed storage/lipid transfer protein (LTP) family protein	28	**0.008** (+)	0.439	0.171
33.1 Development, storage proteins	57	0.194	0.606	**0.047** (+)
Category 2: decreased expression mid-leaf				
4.12 Glycolysis, enolase	11	**<0.001** (+)	**0.025** (-)	**0.006** (+)
10.1 Cell wall, precursor synthesis	71	**<0.001** (+)	**<0.001** (-)	0.656
11.1 Lipid metabolism, FA synthesis and FA elongation	172	**<0.001** (+)	**0.002** (-)	0.556
11.6 Lipid metabolism, lipid transfer proteins	10	**0.038** (+)	**0.025** (-)	0.804
13.1 Amino acid metabolism, synthesis	250	**<0.001** (+)	**0.002** (-)	**0.015** (+)
26.2 Misc., UDP glucosyl and glucoronyl transferases	307	**0.008** (+)	**<0.001** (-)	0.634
26.4 Misc., beta 1,3 glucan hydrolases	82	**0.015** (+)	**<0.001** (-)	0.155
27.1 RNA, processing	376	**<0.001** (+)	**<0.001** (-)	0.365
27.3 RNA, regulation of transcription	2622	**<0.001** (+)	**<0.001** (-)	0.249
27.4 RNA, RNA binding	182	**<0.001** (+)	**<0.001** (-)	0.880
29.1 Protein, aa activation	149	**0.017** (+)	**0.002** (-)	0.315
29.2 Protein, synthesis	730	**<0.001** (+)	**<0.001** (-)	**<0.001** (+)

		Wilcoxon rank sum test P-value		
BIN	N elements	i.v.1 → i.v.3	i.v.3 → i.v.5	i.v.1 → i.v.5
29.3 Protein, targeting	326	**0.008 (+)**	**<0.001 (-)**	0.082
29.6 Protein, folding	114	**0.014 (+)**	**0.002 (-)**	0.517
33.99 Development, unspecified	682	**0.002 (+)**	**<0.001 (-)**	0.864
34.9 Transport, metabolite transporters at the mitochondrial membrane	87	**<0.001 (+)**	**0.038 (-)**	0.902
Category 3: increased expression in leaf apex				
2.1 Major carbohydrate metabolism, synthesis	64	0.537	**0.001 (-)**	**<0.001 (-)**
10.6 Cell wall, degradation	160	0.104	0.096	**0.005 (-)**
11.2 Lipid metabolism, FA desaturation	19	0.280	0.552	**0.034 (-)**
11.3 Lipid metabolism, Phospholipid synthesis	52	0.776	**0.012 (-)**	**0.001 (-)**
11.8 Lipid metabolism, 'exotics' (steroids, squalene)	111	0.995	**0.022 (-)**	**0.047 (-)**
26.3 Misc. other, Ferredoxins and Rieske domain	7	0.200	0.141	**0.004 (-)**
26.13 Misc., acid and other phosphatases	59	0.180	**<0.001 (-)**	**0.029 (-)**
28.1 DNA, synthesis/chromatin structure	417	0.737	**<0.001 (-)**	**<0.001 (-)**
30.4 Signalling, phosphinositides	74	0.171	0.516	**0.015 (-)**
30.11 Signalling, light	120	0.864	0.094	**0.007 (-)**
31.2 Cell division	107	0.507	**<0.001 (-)**	**0.011 (-)**
34.2 Transport, sugars	110	0.511	0.112	**0.005 (-)**
34.3 Transport, amino acids	113	0.295	**<0.001 (-)**	**<0.001 (-)**

		Wilcoxon rank sum test P-value		
BIN	N elements	i.v.1 → i.v.3	i.v.3 → i.v.5	i.v.1 → i.v.5
34.13 Transport, peptides and oligopeptides	99	0.681	**0.002 (-)**	**<0.001 (-)**
Category 4: alternate pattern of expression				
1.1 Photosynthesis, light reaction	241	**<0.001 (-)**	**0.012 (+)**	**<0.001 (-)**
3.5 Minor CHO metabolism, others	70	0.159	**0.002 (-)**	0.424
4.9 Glycolysis, glyceraldehyde 3-phosphate dehydrogenase	11	0.333	**0.033 (-)**	0.700
5.1 Fermentation, aldehyde dehydrogenase	13	0.593	**0.018 (-)**	0.452
7.1 OPP, oxidative PP	26	**0.030 (+)**	0.832	0.100
11.9 Lipid metabolism, lipid degradation	119	**0.012 (+)**	0.701	0.193
15.1 Metal handling, acquisition	8	**0.001 (-)**	0.552	**0.013 (-)**
20.1 Stress, biotic	925	**0.001 (-)**	0.298	0.227
29.4 Protein, postranslational modification	1042	0.268	**<0.001 (+)**	**0.001 (-)**
29.5 Protein, degradation	1865	**<0.001 (+)**	**<0.001 (-)**	**0.001 (-)**
29.8 Protein assembly and cofactor ligation	47	0.919	**0.013 (+)**	0.052
30.2 Signalling, receptor kinases	1057	**<0.001 (-)**	**0.022 (+)**	0.316
30.5 Signalling, G-proteins	288	0.583	**0.008 (-)**	0.227
34.8 Transport, metabolite transporters at the envelope membrane	36	0.235	**0.004 (-)**	0.293
34.19 Transport, Major Intrinsic Proteins	63	0.834	**0.033 (-)**	0.122

BIN	N elements	Wilcoxon rank sum test P-value i.v.1 → i.v.3	 i.v.3 → i.v.5	 i.v.1 → i.v.5
34.21 Transport, calcium	25	**0.005 (-)**	0.163	**<0.001 (-)**
34.98 Transport, membrane system unknown	6	**0.042 (-)**	0.749	0.055

Several functional classes displayed a significant number of up regulated elements in the leaf apex (i.v. 5+, Category 3). Many of these functional classes are involved in synthesis and cell division, indicating the cell cycle may be increased in the leaf tip in newly unfurled leaves in *P. nigra* (Table 4.6). Gene classes involved major carbohydrate metabolism synthesis (BIN 2.1), lipid metabolism (BINs 11.3 and 11.8), signalling (BINS 30.4 and 30.11), and transport (BINs 34.2, 34.3, and 34.13) all displayed a significant number of elements up regulated in the leaf tip, indicating cells in the leaf apex are increasing transcription for synthesis of cellular components. In addition, functional classes related to cell wall degradation (BIN 10.6), DNA synthesis and chromatin structure (BIN 28.1) and cell division (BIN 31.2) displayed and increase in transcription in the leaf apex. Together, the pattern of gene expression indicates cell division to be up regulated in the tip compared to the basal sections of the leaf.

The fourth category of significantly regulated functional classes includes all other patterns of gene expression. Several functional classes displayed a small gradient across the leaf lamina, with a significant number of elements up or down regulated in one segment (Table 4.6). A subset of these BINs display a pattern similar to Category 2, with more genes down regulated in the middle segment (BINs 3.5, 4.9, 5.1, 29.5, 30.5, 34.8, and 34.19), consistent with the pattern or overall down regulation in the mid-lamina. Conversely, two functional groups displayed an increase in expression mid-leaf: the light reaction components of the photosystem II complex (BIN 1.1) and receptor kinase signalling molecules (BIN 30.2). The basal section (i.v. 1-2) displayed significantly fewer up-regulated genes related to light reactions than the middle and apical sections, indicating the distal regions of the leaf may provide the initial photosynthesis at dawn. Other functional groups displayed similar patterns to Category 1, with a slight increase in gene expression in the leaf base (BINs 7.1 and 11.9), while one bin displayed a pattern similar to an increase in regulation in the leaf apex (BIN 29.5).

Overall, the differences in gene expression between segments of newly unfurled leaves in *P. nigra*, while statistically slight, identify functional and biologically important differences in cellular activities across the leaf lamina (Figure 4.11). Genes related to energy production and substrate reclamation were more likely to be up regulated in the base of the leaf, while genes related to cell division were differentially expressed in the leaf apex. Many functional classes related to metabolism and basic cellular functions were down regulated in the middle segment, but genes related to photosynthesis and signalling were increased mid-leaf.

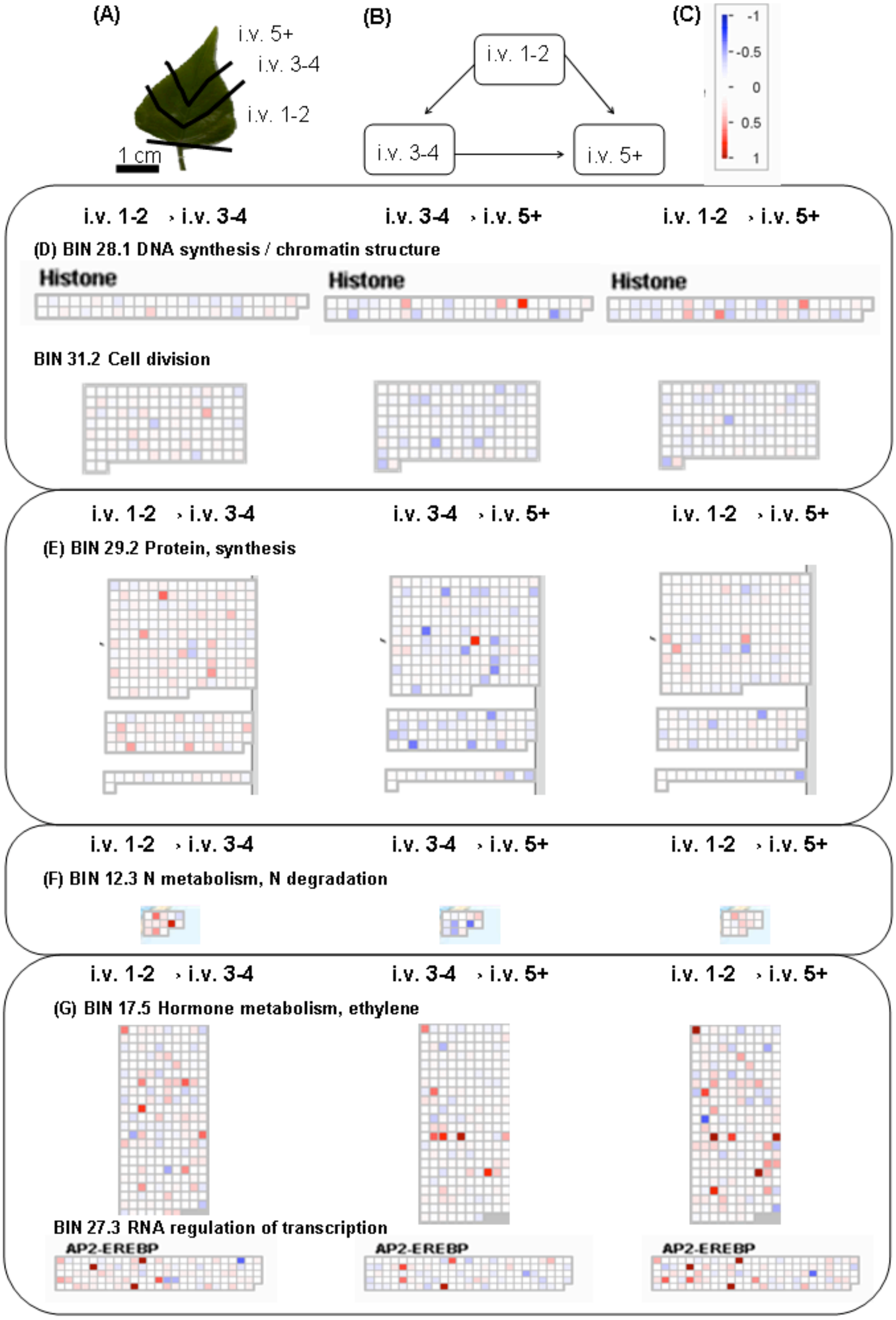


Figure 4.11 Variation in gene expression across the leaf lamina in *P. nigra*. (A) Gene expression was assessed in the basal (i.v.1-2), middle (i.v.3-4) and apex (i.v.5+) of newly unfurled leaves. (B) Expression log-fold change ($\log_2$FC) was determined for the more basal leaf section relative to the more apical section. Thus, categories of genes (BINs) with more elements up regulated in the basal section had a positive $\log_2$FC. (C) Positive $\log_2$FC values are depicted in red, negative $\log_2$FC values in blue. Each box represents a single element in the larger BIN (the grid). (D) Genes related to cell division were up regulated in the leaf tip. (E) Categories related to protein synthesis were down regulated in the middle section of the lamina. (F) Nitrogen degradation via GDH was up regulated in the leaf base. (G) Ethylene metabolism and transcription related to ethylene response were up regulated in the leaf base.

4.4 Discussion

4.4.1 Cell division drives variation in leaf morphology in *P. nigra*

Previous studies have described the significant genetic variance in leaf morphology in *P. nigra* across western Europe (Trewin et al. *in preparation*, Chapter 2, Chapter 3). Understanding the developmental basis of the morphological variation will help identify candidate genes that contribute to the variation and thus may be affected by natural selection. Identifying the genetic mechanisms is also of interest to breeding programs (e.g. for bioenergy) and the broader effort to unravel the complex control of leaf morphogenesis.

The relationship between cell size and cell number varies among species of *Populus*. It is well established that *Populus deltoides* displays a larger number of smaller cells, while leaves of *P. trichocarpa* have relatively fewer cells of larger size (Ridge et al. 1986; Ferris et al. 2002). In a previous study of single genets (genotypes) *P. nigra*, *P. alba*, and *P.* x *euramericana* (*P. deltoids* X *P. nigra*), *P. nigra* displayed a leaf size intermediate to the other species, with this size achieved via a smaller number of significantly larger abaxial epidermal cells (Ferris et al. 2001). Little work has been done investigating the cellular mechanisms controlling intraspecific variation in leaf morphology, however. *P. trichocarpa* from two sites differing in precipitation and temperature displayed significantly different leaf size and cell number, but not cell area, when grown in common gardens (Dunlap and Stettler 2001). In contrast to *P. trichocarpa*, three genets of *P. alba* displayed significant variation in epidermal cell area as well as cell number (Abbruzzese et al. 2009).

Here, the fine-scale assessment of leaf morphogenesis in four genets of *P. nigra* focused on the relationship between leaf area and epidermal cell traits. Leaf area differed among genets, though variation in cell size was nonsignificant among genets. Cell division influenced leaf morphogenesis and variation at the intraspecfic level. The observation that cell number differed among genets is more consistent with the cell theory than the organismal theory of growth (Fleming 2006b; Tisné et al. 2008).

The distinct morphology of the Spanish sample may reveal compensatory growth to affect morphology in this species, more consistent with the neo-cell theory of growth (Tsukaya 2005). Both in this glasshouse study and a field-based assessment indicated that the Spanish samples having significantly larger cell area, but smaller leaves composed of a smaller number of cells (Trewin et al, *in preparation*). This morphological distinctiveness is consistent with adaptation to the more arid environment of the Iberian peninsula (Chapter 2), and is similar to the pattern of slower growth and smaller leaves in xerophytic collections of *P. trichocarpa* (Dunlap and Stettler 2001). The genetic basis of the Spanish morphology (and partial compensation) may lie in a more general growth regulator (e.g. hormone) rather than genes specific to leaf morphogenesis (Ramirez-Parra et al. 2005).

Developmentally, the small-leaf Spanish phenotype appears to result from an early exit of the cell division phase of leaf morphogenesis. In sunflowers, cell division was shown to take place in the first two thirds of leaf development, with the final leaf expansion due to cell growth (Granier and Tardieu 1998). Exit from the cycle varies among species of *Populus*, with *P. trichocarpa* exiting the cell cycle early in leaf development, while *P. deltoides* maintains cell division late into morphogenesis (Van Volkenburgh and Taylor 1996). In the absence of increased cell growth, early termination of cell division is likely sufficient to result in smaller leaves in *P. nigra*. Statistical analysis of relative expansion rates in *P. nigra* detected a significant Genet X Day interaction (Table 4.3), indicating differences in growth rate among genets. The three larger-leaf samples cease leaf expansion around the fourteenth day after unfurling, while the Spanish genet completed leaf expansion around day eleven (Figure 4.4). This significant intraspecific variation complicates attempts to match developmental stages among samples from divergent genets. For example, the leaf plastochrone index (LPI) is commonly applied to classify leaves indirectly into developmental ages for comparison (Erickson and Michelini 1957). This method assumes an equal developmental rate among samples, however, such that a single reference leaf size may be selected where the leaves are growing exponentially. The significant Genet X Day interaction detected for four genets of *P. nigra*, along with the large range in mature leaf size and duration of exponential growth (Trewin et al. *in preparation*) indicate that a single LPI estimate may be inappropriate for this collection. As a result, direct estimates of growth rate and duration (e.g. flow cytometry) may be required to assess the cessation of cell division and timing of leaf maturation in *P. nigra*.

4.4.2 Variation in leaf shape resulted from differences in relative expansion rates across the leaf lamina

Analysis of the interveinal areas of each leaf revealed significant variation in relative growth rate across the leaf lamina, with relative expansion rate (RER) decreasing from base to tip. This pattern differs from the constant RER observed in *P.*

deltoides (Matsubara et al. 2006), but is consistent with the proximodistal gradient reported in *P.* x *euramericana* (Taylor et al. 2003) and *P. nigra* (Ferris et al. 2001). Tobacco (Trainotti et al. 2004) and sunflower (Granier and Tardieu 1998) leaves display a basipetal gradient of development, with the leaf apex ceasing cell division and maturing before the leaf base. In sunflower, this gradient was sufficient to explain differences in interveinal area, with overall expansion occurring at a constant rate across the lamina. Here, *P. nigra* displayed variation RER among interveinal areas both upon unfurling (during exponential growth, day 1-2) and over the entire leaf expansion (day 1-25), indicating expansion is not even across the lamina (Table 4.3, Figure 4.7).

Unlike the mean estimates over each leaf, both abaxial cell number and abaxial cell area varied among interveinal regions. Cell area differed between leaf ages (newly unfurled and mature) and interveinal regions, in contrast to evidence of uniform cell area in *P. deltoides* (Matsubara et al. 2006). Cell number varied significantly among leaf age, i.v. regions, genets, and each interaction. No clear pattern in the variation in cell size was detected across the lamina (Figure 4.7), yet variation in i.v. area was more correlated with cell number (R^2=0.76) greater than with cell area (R^2=0.064), consistent with cell division contributing to the proximodistal variation in RER and area.

Cell size is a phenotypically plastic trait sensitive to environmental conditions. Studies reporting variation in cell area across a leaf lamina are lacking in tree species. Actively growing wheat leaves display a sigmoidal distribution of cell sizes along the leaf blade, with cell area increasing from base to tip (Hu and Schmidhalter 2008). Cell size in *Arabidopsis* sepals, modified leaf structures on the outer whorl of flowers, is variable across the sepal structure and associated with increases in ploidy via endoreduplication (Roeder et al. 2010). High light can induce endoreduplication, with an associated increase of adaxial epidermal cell size in *Phaseolus vulgaris* (Kinoshita et al. 2008). *Trifolium repens* petioles displayed variation in both cell size and number in response to shade, but variation in petiole length in high light was attributable to differences in cell number but not cell size (Weijschede et al. 2008). Variation in cell size across the leaf lamina in newly unfurled leaves of *P. nigra* is most likely linked to the basipetal gradient of development. Alternatively, differences in the microclimate of ramets sampled for cellular assessment may have contributed to the variation in cell size observed across the leaf lamina. Additional studies comparing leaf morphology in controlled growth chambers and field conditions will be required to verify a genetic component to cell size variance within the leaf epidermis.

Together, the difference in cell number among genets and across the lamina confirms genetic control of leaf morphology in *P. nigra*. Investigation of regulatory differences such as transcriptional profiles may provide evidence of the genetic basis for this phenotypic variation.

4.4.3 Differences in gene expression revealed a proximodistal functional gradient across the developing leaf

Gene expression within a developing organ likely varies in a more muted fashion than differences due to stringent experimental manipulation. A twofold change in expression is typically considered biologically informative in manipulative tests (Groover et al. 2006; Leakey et al. 2009). Trends of consistent regulation among genes within functional groups may be more informative than individual log-fold change estimates (Leakey et al. 2009). Functional groups of genes were not evenly expressed across the lamina of newly unfurled leaves in *P. nigra*, indicating physiological functions may be compartmentalized across the organ or determined by a proximodistal developmental gradient. These patterns identified by Wilcoxon sum rank tests are biologically informative even in the absence of individual genes significantly regulated at a FDR of 0.25. Three distinct patterns of gene expression were revealed: functional classes that were up regulated in the basal section, decreased in the middle section, or up regulated in the apical segment of the leaf (Table 4.6). These patterns revealed differences in energy use and metabolism across the leaf.

The leaf base (interveinal regions 1 and 2) up regulated a significant number of genes related to amino acid and nitrogen degradation. The nitrogen degradation genes, more specifically, include glutamate dehydrogenase (GDH), a key enzyme involved in carbon and nitrogen metabolism in plants (Miyashita and Good 2008b). GDH catalyzes the reversible transfer of an amino group between glutamate and 2-oxoglutarate (α-ketoglutaric acid), and serves as a key link in nitrogen-carbon (C-N) balance (Miyashita and Good 2008b). GDH is induced by C starvation and dark conditions (Melo-Oliveira et al. 1996; Miyashita and Good 2008b) to oxidise glutamate to produce ammonium and oxoacids, the latter of which likely enter the TCA cycle as an alternate carbon source (Robinson et al. 1991; Miyashita and Good 2008b). Increased proteolysis has been reported in response to C starvation (Koch 1996; Brouquisse et al. 1998), consistent with the increase in transcripts related to amino acid degradation observed in the basal segment of *P. nigra* leaves here. GDH is also implicated as a key component of this broader amino acid metabolism, further supporting its role in the complex C starvation response (Miyashita and Good 2008a). Consistent with the induction of GDH as an alternative carbon source, the basal segment of newly unfurled leaves displayed an increase in transcription in genes related the TCA cycle (Table 4.6). Together, the induction of transcription related to amino acid and N degradation indicates the basal section of the leaf to be drawing on alternative C sources. These results may reflect a dark response as the leaves were harvested just after dawn, or they may reflect a putative energy gradient across the leaf lamina.

Genes involved in metabolism of the important plant phytohormones ethylene, salicylic acid and, to a lesser extent, jasmonic acid were significantly more likely to be induced in the leaf base than the more apical segments. These three hormones are involved in fine-tuning plant defensive response to herbivore attacks (Leon-Reyes et al. 2010). Salicylic acid is induced in response to biotrophic (e.g. fungal) pathogen attack (Baek et al. 2010), while jasmonic acid is induced in response to necrotrophic pathogens and insect herbivores (Leon-Reyes et al. 2010), though this distinction may be oversimplified (Thaler et al. 2010). Members of the ethylene response transcriptions factor family appear to mitigate the balance between the salicylic and jasmonic response pathways (Leon-Reyes et al. 2010). The reduced expression of biotic stress-related transcripts in the basal section indicates these patterns may not be strictly related to pathogen attack (Table 4.6). Alone, ethylene is primarily considered an inhibitor of plant growth, but recent evidence indicates that it may stimulate growth at low concentrations in developing organs (Pierik et al. 2006), meaning the elevated transcription ethylene metabolism and response factors in the leaf base may induce multiple developmental pathways. Genes related to the AP2/ERF (ethylene response factor) transcription factors were significantly up regulated in the basal section (Figure 4.10), including one gene with a logFC greater than 2 (Table 4.5). The AP2/ERF family of transcription factors is large and diverse, with members involved in developmental processes as well as response to biotic and abiotic stresses (Marsch-Martinez et al. 2006).

In contrast to those groups of genes up regulated in the leaf base, genes related to calcium transport were down regulated in the leaf base compared to the middle and apical sections (Table 4.6). Cytosolic calcium is involved in a variety of physiological functions within plant cells, including stomatal closure and the discrimination of osmotic changes (Sathyanarayanan and Poovaiah 2004). In *Arabidopsis*, cytosolic calcium concentration correspond to photoperiod, with calcium concentrations reaching a trough in the dark, rising in response to light, peaking several hours after dawn, and responding to changes in diurnal cycles (Love et al. 2004). The observed pattern of gene expression in *P. nigra* is consistent with a proximodistal gradient in the leaf response to light, with calcium transport occurring in the apical segments before the leaf base.

The majority of functional classes differentially regulated in the middle section of the leaf displayed reduced transcription compared with the basal or apical sections (Table 4.6, Category 2). Three classes displayed the opposite trend and had significantly more genes up regulated in the middle section: photosynthesystem II components involved in light reactions, signalling kinases, and genes related to posttranslational modification of proteins. The photosystem II complex houses the initial steps in oxidative photosynthesis, the process by which plants produce carbohydrates and yield atmospheric oxygen (Govindjee et al. 2010). The up regulation of genes related to photosystem II production in this study indicates the middle lamina respond to the light stimuli of dawn prior to, or at a greater magnitude than, the base and tip of the leaf. This pattern may be due to differential stimulus response across the leaf lamina or it may reflect compartmentalization of photosynthetic activity at dawn. The duration of this differential expression and physiological consequences of differential regulation of photosynthetic machinery across the lamina are intriguing topics for further study.

Receptor kinases are a superfamily of signal transduction enzymes involved in phytohormone perception, plant development, and defense response (Torii 2000). Leucine-rich repeat receptor kinases, such as those included in the functional class up regulated mid-leaf, form the basis for perception of extracellular stimuli and signal transduction into the cell cytoplasm (Torii 2000). One mechanism of calcium signalling within the cell is the phosphorylation of target proteins (Stone and Walker 1995; Sathyanarayanan and Poovaiah 2004), consistent with the increase in calcium transport transcripts in the middle segment compared to the basal section of the leaf. Similarly, the up regulation of 'posttranslational modifications of proteins' mid-leaf includes genes related to calcium binding and signalling (e.g. putative calcium-dependent protein kinases and calcium-binding EF hand proteins).

Functional classes of genes distinct from those up regulated in the basal section were significantly regulated in the apex of the leaf (Table 4.6). These categories included major carbohydrate and lipid metabolism, indicating that, unlike the leaf base that displayed transcription related to protein and nitrogen degradation, the leaf tip may be actively producing carbohydrates and complex compounds. The increased transcription of categories related to amino acid and protein transport, cell wall degradation, DNA chromatin production and cell division, are consistent with the cells in the leaf apex resuming cellular growth and division, distinct from the continuing dark response observed in the basal section of the leaf.

Together, these patterns of gene expression indicate this experiment revealed differences in photoresponse as well as developmental stage across the leaf lamina rather than transcriptional differences related to leaf morphogenesis. Tobacco leaves mature in a similar pattern, with the leaf apex switching from carbon sink to source prior to the leaf base (Meng et al. 2001). As *P. nigra* leaves display a gradient in cell division and RER, it is possible the variation in transciption reflects in part a basipetal gradient of development that influences the timing and mechanism of light response.

These results partially reflect the design of tissue harvesting that took place at dawn in order to minimize the effects of water or temperature stress in the greenhouse environment. If the stimulus of light at dawn results in a signalling cascade that converts cellular function from carbon and nitrogen catabolism to photosynthate-based cellular activities, then it appears these changes occur unequally across the leaf surface. More precisely, the apical region of the leaf began carbohydrate production and resumed other cellular functions (e.g. cell division) earlier or more quickly than did the basal section of the leaf. Given the critical nature of light perception and profound influence light has on plant morphology and development, it should not be surprising that transcriptional changes in response to daybreak should override developmental gene expression previously reported to dominate the transcriptome of leaf development in *Populus* (Sjödin et al. 2008; Street et al. 2008). These results pose additional questions related to stimulus perception, signal transduction and energy use compartmentalization in plants which should be addressed first by confirming differential transcription with quantitative real-time PCR, followed by physiological and metabolomic assessments in a controlled environment. Assessment of the spatial regulation of genes related to leaf growth should be undertaken during the daylight hours in a well watered and temperature controlled environment.

Chapter 5. Quantitative trait loci for insect damage and leaf development colocate in hybrid poplar

5.0 Overview

Leaves are not only the site of photosynthesis and gas exchange for the plant, but also the primary interaction with phytophagous insects. In order to identify the genetic basis of variation in insect association (damage) in hybrid poplar, a F_2 family *Populus trichocarpa* x *P. deltoides* was assessed for seven categories of insect leaf damage in a common garden experiment located in southern England. Data recorded in June and August identified significant variation in damage among the F_2 progeny, with some levels of damage exceeding those in the *P. trichocarpa* or F_1 parents, indicating transgressive phenotypes. Data were treated to quantitative trait loci analyses that identified 14 genomic regions correlating with damage in seven of the damage category-month combinations. Up to four QTL were identified for a single damage-month combination, with the genetic variation explaining up to 24% of the variance in the phenotype. When compared to QTL for leaf traits from previous studies, two insect association QTL were found to colocate to genetic hotspots, indicating leaf morphology and not strictly biochemical attributes may influence insect preference. Utilizing the *P. trichocarpa* v2.0 genome to identify all gene models within the QTL, analyses of the gene ontology (GO) classifications identified a number of GO categories significantly over-represented in the QTL for each damage-month combination. These functional classes help narrow lists of candidate genes for further identification of genes underlying variation in insect association in this economically and ecologically important genus. As the *Populus* species examined are native to North America, the genetic correlations with insect association likely include a genotype by environment effect such that insect association in another common garden may reveal different QTL. However, given the broad classifications of damage applied here and the application of network analyses, these results provide insight into the cellular components involved in insect association in a fast-growing forest tree.

5.1 Introduction

Photosynthetic plants display a remarkable diversity in leaf size and shape. This variation largely results from adaptation to local climates (Bailey and Sinnott 1916; Abrams 1994). For the most part, the adaptive nature of leaf traits has been investigated in relation to abiotic factors, primarily temperature and water availability (Dudley 1996b; Barboni et al. 2004; Meier and Leuschner 2008; Royer et al. 2008). In addition, the significant variation in leaf size observed in *Populus nigra* has been shown to correlate to both local climate and historic vicariance events (Chapters 2 and 3). A broad generalization of these studies indicates that drier or more xeric climates favour smaller leaves to minimize water loss due to transpiration, while moister and tropical climates are associated with larger leaves.

As leaves are the sites of photosynthesis and gas exchange, abiotic factors are likely the primary drivers of adaptive variation in leaf morphology. However, due to their prevalence and palatability, leaves are also the primary sites of interaction with herbaceous insects. The evolution of plant defences, and co-evolution between defences and insect herbivores, has been the subject of extensive study, especially in relation to agriculturally important systems (Kover and Caicedo 2001). Phytophagous insects may constitute a significant evolutionary pressure in natural populations of woody perennial species, despite the potential for these plants to vary biomass allocation in response to attack (Stevens et al. 2008). For example, the interaction between phytophagous insects and their avian prey may have influenced leaf morphology and architecture in *Quercus* systems (Park et al. 2008). In hybrid willow, the phenolic composition of leaves varied with genotype, and insect performance was correlated with plant genotype but was not influenced by drought treatment, meaning insect preference may be independent of some abiotic factors (Glynn et al. 2004). Variation in leaf palatability can influence the insect community associated with individual trees in natural populations (Bangert et al. 2006), thus leading to variation in multi-trophic level interactions (Bailey et al. 2006; Park et al. 2008; Broderick et al. 2010).

Populus spp., poplars, cottonwoods, and aspens, are fast growing, deciduous trees widely distributed across the Northern Hemisphere. *Populus* spp. are often considered keystone species in natural systems, and are associated with a large number of insect, vertebrate, and fungal species (Philippe and Bohlmann 2007). Many studies have investigated the genetic basis of resistance to fungal pathogens in *Populus*

(Newcombe et al. 1995; Newcombe and Bradshaw 1996; Newcombe et al. 1996; Woo and Newcombe 2003; Levée et al. 2009; Major et al. 2010). In addition, the biochemical basis of *Populus* resistance to several herbivores has been investigated, revealing the importance of such compounds as phenolic glycosides (Young et al. 2010), condensed tannins (Schweitzer et al. 2008), and endochitinases (Philippe and Bohlmann 2007). The biochemical and the associated genetic basis of the resistance of *Populus* to forest tent caterpillars (*Malacosoma disstria* Hubner), a herbivorous insect, have been studied in detail. Genetic and genomic studies have identified transcriptional products and gene families as candidates in the plant response to herbivory, including actual (Ralph et al. 2008) and simulated forest tent caterpillar feeding (Philippe et al. 2010), and Kunitz-type protease inhibitors in particular (Philippe et al. 2009). Following defoliation forest tent caterpillars appear to induce increases in phenolic glycosides in subsequent leaf flushes, potentially influencing the composition and feeding pattern of later-season insects (Donaldson and Lindroth 2008). These studies have confirmed that members of the genus *Populus* display genetic (i.e. adaptive) variation conferring resistance to major pests and pathogens in natural systems. The complex genetic control of this variation is still a topic of vigorous study.

In a broader approach, studies into the relationship between genetic variation in *Populus* host species and the diversity of the associated insect community have shed light on the complexities of community genetics (Bailey et al. 2006; Bangert et al. 2006; Wimp et al. 2007; Keith et al. 2010). In these studies of *Populus fremontii*, *P. angustifolia* and their hybrids, significant correlations have been identified between the genetic similarity of individual trees, the chemical properties of their leaves, and the structure of the insect communities associated with each individual (Bangert et al. 2006). In a fine-scale study, a significant correlation between genetic variation and insect community was detected while the correlation between chemical traits and insect community was non-significant, indicating that additional unidentified genetic factors influence community structure in natural stands (Wimp et al. 2007). Yet these studies have focused on anonymous genetic differences (AFLP markers) between unrelated genetic individuals from natural stands, a system that does not provide insight into which genomic regions underlie the genetic variance. While insect associations on hybrid *Salix* have revealed genomic regions (quantitative trait loci, QTL) associated with insect damage in willow (Rönnberg-Wästljung et al. 2006), to date no such study has been published for hybrid *Populus*.

This study aimed to determine whether there is a genetic basis for insect preference among hybrid poplar trees. First, variation in insect association, classified by damage type, was quantified through visual inspection of leaf damage in an F_2 pedigree of *P. trichocarpa* x *P. deltoides*. Second, a QTL analysis was conducted to identify genomic regions associated with variation in insect abundance. Third, collocation between QTL for insect abundance and QTL for leaf traits from previous studies was assessed to determine whether insect preference corresponds to leaf morphology as well as leaf chemistry. Fourth, using *Populus* genomic resources, a bioinformatics study was conducted to identify functional classes of genes over-represented in insect association QTL that may provide insight into the broader genetic control of these quantitative traits.

5.2 Methods

5.2.1 Study system and field assessment

In North America, *Populus trichocarpa* (black cottonwood) and *P. deltoides* (eastern cottonwood) are widely distributed and occur sympatrically (Eckenwalder 1984a). *P. trichocarpa* (sect. *Tacamahaca*) and *P. deltoides* (sect. *Aigeiros*) are most easily distinguished by leaf size, with the former displaying lanceolate leaves and the latter deltoid to cordate leaf bases (Eckenwalder 1984a). The species also differ in leaf color and margin, branch and bud color, and floral characteristics (Eckenwalder 1984a). Due to their interfertility and hybrid vigour when crossed, these species of *Populus* have been repeatedly chosen for hybrid analyses into the genetic basis of complex phenotypic traits (Ridge et al. 1986; Ceulemans et al. 1992; Rae et al. 2009).

POP1 (Family 331) is a F2 hybrid pedigree of *Populus trichocarpa* x *P. deltoides*, consisting of the maternal grandparent *P. trichocarpa*, 93-968, crossed with the paternal grandparent *P. deltoides*, ILL-129, to produce hybrid F1 progeny. Two of the F1 trees, the female 53-242 and male 53-246, were then crossed in two years to produce a full-sib hybrid pedigree (Family 331) (Bradshaw et al. 1994; Wu et al. 1998). This study analyzed 189 genets (genetic individuals) of this pedigree, including the maternal grandparent (93-968), the two F1 parents (53-242 and 53-246) and 184 of the F2 progeny. One pure *P. deltoides* genet was analyzed to assess the variation in the paternal species, but the paternal grandparent was not available for assessment. In the spring of 2000, three replicates (ramets) of each genet were planted in a fully replicated randomized block design as part of an ongoing short rotation coppice experiment located at the UK Forestry Commission field site at Headley, U.K. (51°07' N, 0°50' W). The trees were coppiced in the winter of 2009 prior to the start of the growing season, so this experiment assessed insect interactions on the first year of growth in the coppice cycle.

The genetic linkage map for the Family 331 pedigree was provided by G. Tuskan (pers. comm.) and includes microsatellite (simple sequence repeats, SSR) and fully-informative amplified fragment length polymorphism (AFLP) markers. For the linkage map, 350 of the F_2 progeny were genotyped for 91 microsatellite markers, and 165 F_2 progeny were genotyped for 92 AFLP markers. The linkage map resolved 21 Linkage Groups, which result in three of the 19 *Populus trichocarpa* chromosomes (5 and 8) aligning to two Family 331 linkage groups. In order to align the linkage map to the *P. trichocarpa* genome (v. 2.0), the forward and reverse primer sequence for each microsatellite in the linkage map were aligned to the genomic sequence using BLASTN with a word length of 4. Alignments were conducted using the tools provided on www.phytozome.net/poplar, accessed 3 May 2010.

Interactions between insects and trees were quantified as different categories of leaf damage. Plants were scored for damage in June and August 2009, with individual ramet scores derived as average scores for 30 leaves chosen at random per plant. Damage was scored in 8 categories. Percentage of leaf area lost was scored for chewer, skeletonizer, miner and gall damage, as well as rust infection. In addition, the number of leaf rollers was counted, and the presence or absence of sap suckers (on the leaf or stem) was scored. Measurements were made by three observers (J. DeWoody, M. Viger, and S. Travers) with the aid of printed guides depicting a scale of leaf damage (in percentages), photos of damage types for each class, and images of common phytophagous insect species. Prior to scoring damage in the field, the three observers practiced scoring damage on *Populus nigra* plants growing outside the Boldrewood glasshouse, and compared mean scores by each individual in order to standardize the interpretation of damage patterns. In addition, the observers intermittently scored the same tree and compared mean damage scores to maintain standardized measures while assessing the common garden. Finally, the observers changed starting locations for the August assessment in order to minimize the potential of influencing block effects by scoring the same plants twice.

5.2.2 Data analysis

Arcsine transformed data (for percentage measures) and square root transformed data (for count measures) were treated to a general linear model using the model:

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon, \tag{5.1}$$

where Y_{ij} is the phenotype of the ith genet in the jth block, α_i is the genet (within individual) effect, β_j is the block effect, and ε is the residual error. As there was no replication within blocks, this is a fully cross-factored model with no nesting. In addition, due to mortality of ramets in the field site, replication was reduced from the initial design so that approximately 30% of genets were represented by three replicates, 34% by two replicates, and 36% by a single replicate. Those traits found to have a significant block effect (June data for Chewer, Skeletonizer and Leaf Sucker; August data for Chewer, Skeletonizer, Galls, Leaf Roller, Leaf Sucker and Rust) were treated to a block correction by adding the difference of the block mean and the grand mean to each ramet score within each block. These block-corrected values were then treated to the same ANOVA model to gain corrected estimates of within and among genet variance. Broad-sense heritability (H^2) was estimated for each trait as:

$$H^2 = \sigma_B^2 / (\sigma_B^2 + \sigma_E^2), \tag{5.2}$$

where σ_B^2 was estimated from the among-genet mean squares (MS_B) and error mean squares (MS_E) from the GLM ($\sigma_B^2 = (MS_B - MS_E)/r$, where r is the number of replicates) and σ_E^2 was estimated from the residual variance (ε) (Rae et al. 2004).

In order to assess the relationship between different categories of data, Pearson's correlation coefficient was calculated for all pairs of damage categories and scoring month. Block-corrected data were used when appropriate. The correlation coefficients and 2-tailed measures of significance were calculated in SPSS v. 17.0 (IBM Corporation, Somers, New York).

The mean scores for each genet in each category of damage were treated to a quantitative trait loci (QTL) analysis. QTL were identified using Grid-QTL (Seaton et al. 2006), a Grid portal analysis system based on the algorithms used by QTLExpress (Seaton et al. 2002). QTL were identified using the ONE-QTL method, with significance determined from 1000 permutations using the chromosome-wide permutation option. For each significant QTL identified, its position (cM) on the linkage group, 95% confidence interval (CI), probability determined from the permutation tests, and percent variance in the trait explained were recorded. In order to describe differences in parental alleles, the paternal (*P. deltoides*) and maternal (*P. trichocarpa*) effects were calculated following Knott et al. (1997). In this method, each effect represents the difference in effect of the *P. trichocarpa* allele and *P. deltoides* allele inherited by either the maternal F1 or paternal F1 parent (Knott et al. 1997). As POP1 is an outcrossed pedigree, it is assumed that each grandparent (P0, pure parental species) was heterozygous at each locus so that the effects of four alleles are considered in each estimate. Thus positive estimates of maternal or paternal effects indicate the *P. trichocarpa* allele conveyed lower damage levels for that herbivore category, while negative values indicate the *P. deltoides* allele correlated with lower damage.

In addition to the well-described biochemical defences, herbivorous insects may select host trees based on developmental differences among genets (Wimp et al. 2007). As physical leaf traits may influence the insect community, a test for collocation between insect association and leaf trait QTL was conducted. First, a list of QTL for leaf traits identified in POP1 (Family 331) was collected from published (Rae et al. 2004; Rae et al. 2006; Street et al. 2006) and unpublished studies (Rodriguez-Acosta 2006) (Appendix 3). Eight categories of leaf traits were included: leaf area, leaf extension rate, leaf length, leaf width, leaf length:width ratio, leaf mass, absolute expansion rate, and specific leaf area. For studies testing the effects of abiotic treatments (Rae et al. 2006; Street et al. 2006), only QTL identified for control (ambient) traits were included. Second, the QTL identified for insect associations were added to the list. Third, the QTL were assessed for non-random alignment across the POP1 genetic map using a (5 cM) sliding-window approach as described in (Rae et al. 2009). This analysis tests for regions of the genome over-represented by QTL were the actual distribution random across linkage groups. Significance was determined from 2000 permutations of the QTL locations across all linkage groups. All collocation analyses and QTL plotting were conducted in qtlplots, a package developed for use in the R statistical environment (The R Project for Statistical Computing) by Nathaniel Street and is available from the author (nathaniel.street@plantphys.umu.se).

For each QTL occurring on a linkage group with sufficient microsatellite markers aligned to the *P. trichocarpa* genomic sequence (v. 2.0), the bp position of the 95% confidence interval for the QTL was estimated using homothetic projection. To do this, the bp positions of the two microsatellite loci closest to each QTL were used to estimate the local ratio of bp per cM for that chromosome region. This ratio was then used to calculate the approximate bp location of the ends of the 95% confidence interval for the QTL. All gene models within each QTL region were then identified using the bp locations of the confidence interval and the BioMart tool at www.phytozome.net/poplar (accessed October 2010).

In order to determine whether genes of similar function were over-represented in any of the damage categories, a gene ontology analysis was conducted for the group of gene models identified in each QTL. Gene ontologies represent a universal vocabulary to classify and describe the role of a gene product in an organism (The Gene Ontology Consortium 2000). Descriptions are provided in three independent classifications: the biological processes in which the gene product functions, the molecular function that the gene product serves, and the cellular component where that function takes place (The Gene Ontology Consortium 2000). Each gene or gene model may be assigned to many categories within each class of ontologies based on the sequence homology with other genes of known function. The gene ontology (GO) classification system provides a standardized method for comparing gene products across eukaryotic organisms and inferring gene function from previous studies (The Gene Ontology Consortium 2000). Here, the distribution of gene ontology categories observed in the combined list of gene models identified for each damage type (e.g. skeletonizer damage in June) was compared to the distribution of gene ontology categories for the *P. trichocarpa* genome as a whole. Tests were conducted using singular enrichment analysis (SEA), with multiple test adjustments made using Benjamini-Hochberg FDR for p=0.05. All analyses were conducted in agriGO v. 1.2, a web-based analysis service that provides gene model GO classification and background level analyses (Du et al. 2010).

5.3 Results

The hybrid pedigree of *Populus* displayed significant variation in levels of damage for the majority of categories scored, indicating that the insect community associated with different genets in a non-random manner (Table 5.1). For the June data and most of the categories of damage in August, the majority of trees showed little damage (Table 5.1). In June, three categories of damage showed significant differences in damage levels among genets: chewers, skeletonizers and sap suckers on leaves. In August, damage due to chewers, skeletonizers, sap suckers on leaves and rust infection all differed significantly among genets (Table 5.2). The distribution of damage levels in the F_2 progeny reveal transgressive segregation of susceptibility to insect damage, as the *P. trichocarpa* parent and F_1 progeny displayed lower damage levels than many F_2 genets (Figure 5.1).

The significance of correlation between damage levels varied among pairs of damage categories. Nine pairs of damage types were significantly correlated (Table 5.3). Three damage types were significantly positively correlated between early (June) and late (August) summer: chewer ($r = 0.213$, d.f. = 181, $P = 0.004$), sap suckers on leaves ($r = 0.996$, d.f. = 181, $P < 0.0001$), and skeletonizers ($r = 0.249$, d.f. = 181, $P = 0.001$). The presence of skeletonizers in June was also positively correlated with chewer ($r = 0.232$, d.f. = 181, $P = 0.002$) and leaf roller ($r = 0.194$, d.f. = 181, $P = 0.008$) damage in August. Similarly, the presence of leaf galls in June was positively correlated with the occurrence of leaf rollers in August ($r = 0.156$, d.f. = 181, $P = 0.035$). Skeletonizer damage in August was negatively correlated with the presence of sap suckers on leaves in June ($r = -0.219$, d.f. = 181, $P = 0.003$), indicating these species may avoid previously damaged leaves. Patterns of damage in August revealed a negative correlation between the presence of skeletonizers and leaf miners ($r = -0.157$, d.f. = 181, $P = 0.034$) and sap suckers on leaves ($r = -0.218$, d.f. = 181, $P = 0.003$), indicating that skeletonizers may deter other phytophagous insects, or that the insect community may be stratified across the tree canopy.

The lack of identification of individual insect species at the common garden site prevents interpretation of the data in terms of community richness or individual pest resistance. However, the two most common Coleoptera species were positively identified (Figure 5.2). The majority of the skeletonizer damage can likely be attributed to *Phratora vitellinae*. This species also feeds as a chewer in the adult form. The second beetle frequently observed was *Chrysomela populi*, and likely was a significant contributor to the chewer damage category.

Table 5.1. Average levels of leaf damage and broad-sense heritability (H^2) observed for eight categories of insect herbivory on an F_2 pedigree of hybrid poplar.

	P. trichocarpa	**F_1**	**F_1**		
	93-968	**53-242**	**53-246**	**$F_2 \pm s_{\bar{x}}$**	**H^2**
Category					
June data					
Chewer[*]	0.744	0.456	1.74	2.38 ± 0.176	0.164
Skeletonizer[*]	1.46	2.84	4.28	4.03 ± 0.298	0.284
Leaf miner[*]	0.022	0.0	0.033	0.009 ± 0.0007	0.000[§]
Gall[*]	0.0	0.0	0.0	0.005 ± 0.0004	0.000[§]
Leaf roller[†]	0.0	0.0	0.0	0.0006 ± 0.00004	0.027
Sap sucker, leaf[‡]	0.256	0.078	0.044	0.074 ± 0.005	0.229
Sap sucker, stem[‡]	0.0	0.0	0.0	0.002 ± 0.0001	0.000[§]
Rust[*]	0.0	0.0	0.0	0.014 ± 0.001	0.000[§]
August data					
Chewer[*]	1.2	2.48	2.17	5.37 ± 0.397	0.137
Skeletonizer[*]	1.56	1.69	1.72	3.09 ± 0.228	0.152
Leaf miner[*]	0.0	0.533	0.344	0.077 ± 0.006	0.098
Gall[*]	0.0	0.033	0.0	0.004 ± 0.0003	0.221
Leaf roller[†]	0.0	0.011	0.0	0.003 ± 0.0002	0.139
Sap sucker, leaf[‡]	0.0	0.033	0.156	0.117 ± 0.009	0.312
Sap sucker, stem[‡]	No damage				n/a
Rust[*]	1.39	22.8	24.8	19.4 ± 1.43	0.454

[*] Percent leaf area damaged

[†] Count of rollers present

[‡] Proportion leaves scored with damage present

[§] Negative calculation truncated to zero

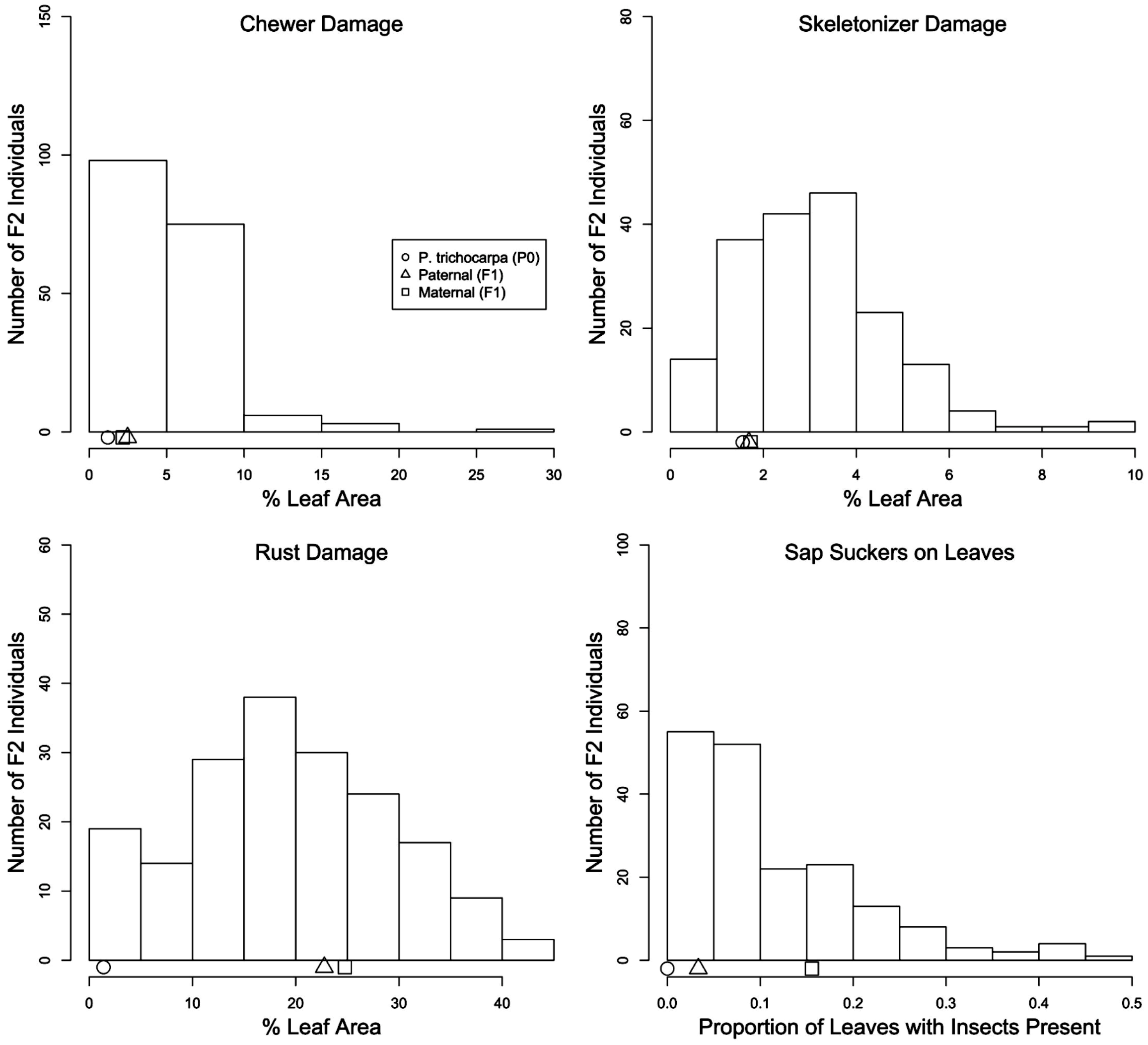


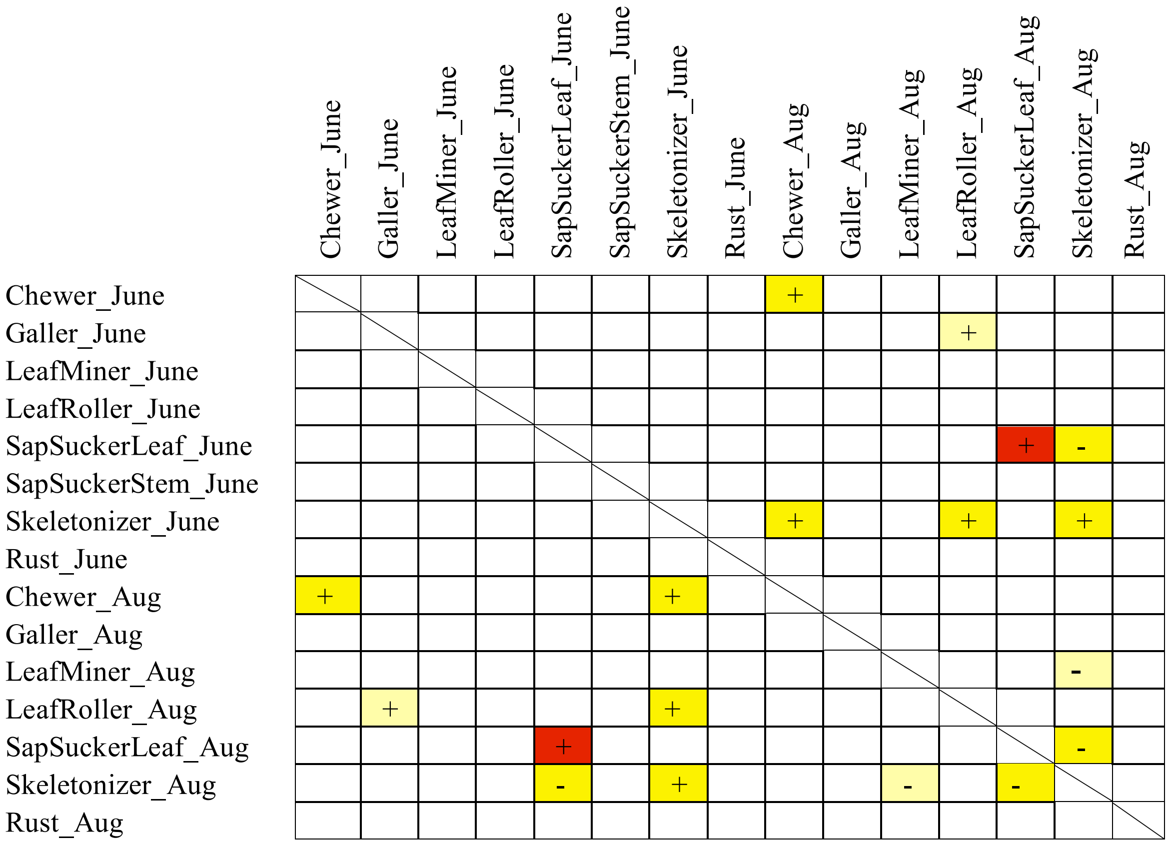
Figure 5.1. Distribution of levels of plant-insect interactions quantified as leaf damage in late summer (August). Histograms depict distribution of damage on a F2 hybrid poplar pedigree. Parental damage levels are depicted below the x-axis for the maternal (▲) and paternal (■) F1 parents, and the *P. trichocarpa* maternal grandparent (●).

Table 5.2. Analysis of variance identified significant genetic variation (factor Genet) in levels of insect damage among progeny of a hybrid pedigree.

	June			**August**		
Category	**Factor**	**MS**	***F***	**Factor**	**MS**	***F***
Chewer	Genet	34.18	$F_{188,187} = 1.44$**	Genet	23.22	$F_{186,175} = 1.40$*
	Block	102.2	$F_{2,187} = 4.29$*	Block	71.68	$F_{2,175} = 4.33$*
	Error	23.82		Error	16.56	
Skeletonizer	Genet	20.95	$F_{188,187} = 2.16$***	Genet	12.38	$F_{186,175} = 1.74$***
	Block	347.1	$F_{2,187} = 35.76$***	Block	492.9	$F_{2,175} = 69.3$***
	Error	9.71		Error	7.11	
Leaf miner	Genet	0.2801	$F_{188,187} = 0.70$	Genet	2.559	$F_{186,175} = 1.21$
	Block	0.0447	$F_{2,187} = 0.11$	Block	0.1801	$F_{2,175} = 0.09$
	Error	0.3981		Error	2.113	
Gall	Genet	0.1719	$F_{188,187} = 0.50$	Genet	0.1436	$F_{186,175} = 1.48$**
	Block	0.8466	$F_{2,187} = 2.46$	Block	0.4469	$F_{2,175} = 4.60$*
	Error	0.3438		Error	0.0971	
Leaf roller	Genet	0.0005	$F_{188,187} = 1.08$	Genet	0.002	$F_{186,175} = 1.23$
	Block	0.00002	$F_{2,187} = 0.05$	Block	0.007	$F_{2,175} = 3.43$*
	Error	0.0005		Error	0.002	
Sap sucker, leaf	Genet	131.4	$F_{188,187} = 1.48$**	Genet	189.5	$F_{186,175} = 1.95$***
	Block	856.9	$F_{2,187} = 9.65$***	Block	463.0	$F_{2,175} = 4.78$**
	Error	88.83		Error	96.92	
Sap sucker, stem	Genet	6.280	$F_{188,187} = 0.71$	Genet	No damage	n/a
	Block	21.38	$F_{2,187} = 2.42$	Block		
	Error	8.843		Error		
Rust	Genet	0.4259	$F_{188,187} = 0.54$	Genet	138.9	$F_{186,175} = 2.99$***
	Block	2.308	$F_{2,187} = 2.93$	Block	236.7	$F_{2,175} = 5.10$**
	Error	0.7866		Error	46.4	

*$P<0.05$; **$P<0.01$, ***$P<0.001$

Table 5.3 Correlation among leaf damage categories assessed in early (June) and late (August) summer in hybrid poplar.



	Chewer_June	Galler_June	LeafMiner_June	LeafRoller_June	SapSuckerLeaf_June	SapSuckerStem_June	Skeletonizer_June	Rust_June	Chewer_Aug	Galler_Aug	LeafMiner_Aug	LeafRoller_Aug	SapSuckerLeaf_Aug	Skeletonizer_Aug	Rust_Aug
Chewer_June									+						
Galler_June												+			
LeafMiner_June															
LeafRoller_June															
SapSuckerLeaf_June													+	-	
SapSuckerStem_June															
Skeletonizer_June									+			+		+	
Rust_June															
Chewer_Aug	+						+								
Galler_Aug															
LeafMiner_Aug														-	
LeafRoller_Aug		+					+								
SapSuckerLeaf_Aug					+									-	
Skeletonizer_Aug					-		+				-		-		
Rust_Aug															

* Colors represent 2-tail significance of Pearson’s correlation coefficients: $p>0.05$: white; $p<0.05$: light yellow; $p<0.01$, bright yellow; $p<0.0001$: red. The sign in each box represents the nature of the correlation: + = positive; - = negative.

Figure 5.2. Two Coleoptera species were frequently observed in the larval and adult form on hybrid poplar. (A) The majority of the skeletonizer damage can likely be attributed to *Phratora vitellinae*. (B) *Chrysomela populi* was likely a significant contributor to the chewer damage category. Photos by F. Lakatos and K. Tuba, University of West-Hungary.

A total of 14 QTL were identified for seven of the season-trait combinations assessed (Table 5.4). No significant QTL were associated with leaf rollers, gall damage, or sap suckers on stems. A smaller number of QTL were identified for the June data than August data, consistent with the higher levels of damage observed later in the season. For the June data, one QTL was identified for each of three damage categories: chewer, skeletonizer, and leaf miner. Each of the June QTL were located on a different linkage group (Table 5.4). These QTL explained between 3.0 and 7.7% of the phenotypic variance observed in these traits.

For the August data, 11 QTL were identified for four damage categories: chewer, skeletonizer, leaf miner, and sap suckers on leaves (Table 5.4). One QTL was identified for skeletonizer damage, four QTL identified for chewer damage, and three QTL for each of the other categories. The single QTL identified for August damage levels explained between 2.9 and 9.3% of the phenotypic variance observed in these traits. In total, the QTL observed for each trait explained 24% of the phenotypic variance in chewer damage, 5.8% in skeletonizer damage, 11% for leaf miner damage, and 14% for the presence of sap suckers on leaves.

Estimates of maternal and paternal effects for each QTL reveal the effects of the *P. trichocarpa* and *P. deltoides* alleles to vary among QTL (Table 5.4). Positive values indicate that the *P. trichocarpa* allele is associated with higher damage levels, while negative values indicate the *P. deltoides* allele from that parent correlated with higher damage. Two QTL for Chewer damage in August displayed large positive effects indicating the *P. trichocarpa* allele was associated with greater damage or insect association: one on linkage group (LG) Va and one on LG XVII. Three QTL displayed large negative parental effects: Chewer (LG Vb) and Skeletonizer (LG XIV) damage in June and Chewer damage in August on LG III. At these loci, the *P. deltoides* allele from each parent was associated with greater damage. QTL displaying parental effects of alternate signs indicate that the *P. trichocarpa* and *P. deltoides* parents were likely heterozygous for alleles of alternate effects (Knott et al. 1997).

The catalogue of QTL for eight leaf traits from three previous studies identified a total of 106 QTL. The number of QTL identified varied by trait, from a minimum of five (absolute expansion rate) to 33 (leaf area) (Table 5.5). When the insect association QTL were added to the list of leaf trait QTL, a total of 126 QTL were assessed for random distribution across the genome. The sliding window analysis identified four genetic ‘hotspots’ characterized by a higher number of QTL than predicted by random

(Figure 5.3). These hotspots occurred at the ‘start’ region of linkage groups VIIIa, IX, XII, and XIV. No insect association QTL were observed on LG IX, indicating that this hotspot solely reflects the genetic control of leaf traits. The hotspot on LG VIIIa was adjacent to an insect category QTL (Figure 5.3). The hotspot on LG XII included the QTL for sap suckers on leaves in August, and the hotspot on LG XIV included the QTL for skeletonizers in June, indicating that these regions may include both leaf development and insect interaction loci, or that the insect community may respond to variation in leaf development controlled by genes in this chromosomal region.

Table 5.4 Quantitative trait loci (QTL) identified for categories of insect damage assessed on an F_2 pedigree of hybrid poplar in June and August.

Category	**LG**	**Position (cM)**	**95% CI***	***P*-value†**	**% Variance‡**	**Paternal effect§**	**Maternal effect**
June							
Chewer	Vb	4	0-19	0.008	3.33	-0.985 (0.342)	-0.115 (0.348)
Miner	I	9	0-24	0.014	7.69	0.008 (0.003)	0.006 (0.003)
Skeletonizer	XIV	0	0-28	0.032	3.02	-0.709 (0.290)	-0.444 (0.290)
August							
Chewer	III	37	29-45	<0.001	9.27	-0.625 (0.275)	-1.085 (0.286)
Chewer	IV	59	45-84	0.012	4.55	0.829 (0.303)	-0.569 (0.341)
Chewer	Va	76	63-82	0.015	4.92	0.608 (0.296)	0.807 (0.299)
Chewer	XVII	54	35-69.5	0.007	5.41	0.469 (0.379)	1.313 (0.395)
Sap Sucker, Leaf	I	74	33-125	0.041	4.81	0.032 (0.015)	-0.035 (0.015)
Sap Sucker, Leaf	VI	144	134-144	0.006	6.73	-0.035 (0.009)	-0.012 (0.009)
Sap Sucker, Leaf	XII	17	0-24	0.037	2.88	0.014 (0.009)	-0.022 (0.010)
Leaf Miner	Va	19	0-41	0.020	4.14	0.045 (0.016)	0.031 (0.017)
Leaf Miner	VIIIa	27	13-28	0.009	3.65	-0.038 (0.016)	0.032 (0.016)
Leaf Miner	XVII	50	33-70	0.032	3.41	-0.054 (0.021)	-0.035 (0.022)
Skeletonizer	III	14	0-27	0.005	5.81	-0.065 (0.153)	-0.468 (0.131)

* 95% confidence interval in cM.
† Significance estimated from 1000 chromosome-wide permutations.
‡ The percent phenotypic variance accounted for by each QTL.
§ The paternal and maternal (F1) effects are provided with standard errors.

Table 5.5. Three previous studies identified 106 QTL for leaf traits in a hybrid pedigree of poplar (POP1).

Leaf Trait	**QTL Name**	**Reference**
Leaf area	leaf_area_SRC1	Rae et al., 2004
	leaf_area_SRC2	Rae et al., 2004
	leaf_area_mature_aCO2	Rae et al., 2006
	leaf_area_young_aCO2	Rae et al., 2006
	leaf_area_leaf_Italy	Rodriguez-Acosta, 2006
	leaf_area_leaf_UK	Rodriguez-Acosta, 2006
Leaf extension rate	leaf_extension_SRC	Rae et al., 2004
	leaf_extension_aCO2	Rae et al., 2006
	leaf_extension_droughtcontrol	Street et al., 2006
Leaf length	leaf_length_leaf_Italy	Rodriguez-Acosta, 2006
	leaf_length_leaf_UK	Rodriguez-Acosta, 2006
Leaf width	leaf_width_leaf_Italy	Rodriguez-Acosta, 2006
	leaf_width_leaf_UK	Rodriguez-Acosta, 2006
	leaf_width_droughtcontrol16d	Street et al., 2006
	leaf_width_droughtcontrol9d	Street et al., 2006
Length:width ratio	leaf_width:length_aCO2	Rae et al., 2006
	leaf_length:width_leaf_Italy	Rodriguez-Acosta, 2006
	leaf_length:width_leaf_UK	Rodriguez-Acosta, 2006
	leaf_length:width_droughtcontrol	Street et al., 2006
Leaf mass	dry_weight_leaf_Italy	Rodriguez-Acosta, 2006
	dry_weight_leaf_UK	Rodriguez-Acosta, 2006
Absolute expansion rate	absolute_expansion_aCO2	Rae et al., 2006
	absolute_expansion_droughtcontrol	Street et al., 2006
Specific leaf area	SLA_SRC	Rae et al., 2004
	SLA_aCO2	Rae et al., 2006
	SLA_leaf_Italy	Rodriguez-Acosta, 2006
	SLA_leaf_UK	Rodriguez-Acosta, 2006

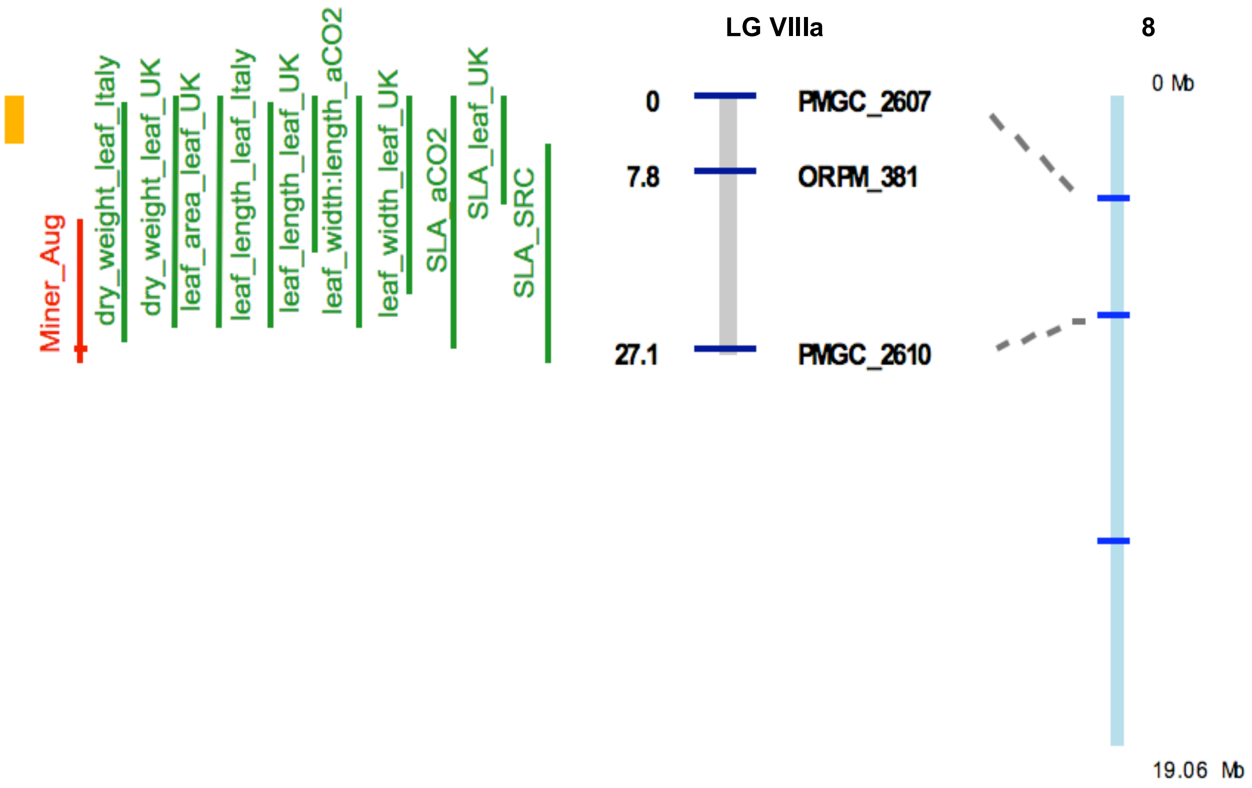


Figure 5.3 QTL for insect damage colocate to QTL hotspots identified for leaf traits. In each plot, the gray bar represents the POP1 linkage group, labelled in Roman numerals. The blue bar represents the *Populus trichocarpa* chromosome, with microsatellite loci (bold) designated by blue lines following the v2.0 genome. AFLP markers are indicated in gray. To the left of the linkage group, green bars represent leaf trait QTL, and red bars represent insect association QTL, each labelled by trait. QTL hotspots are indicated by the orange bar to the left of the QTL.

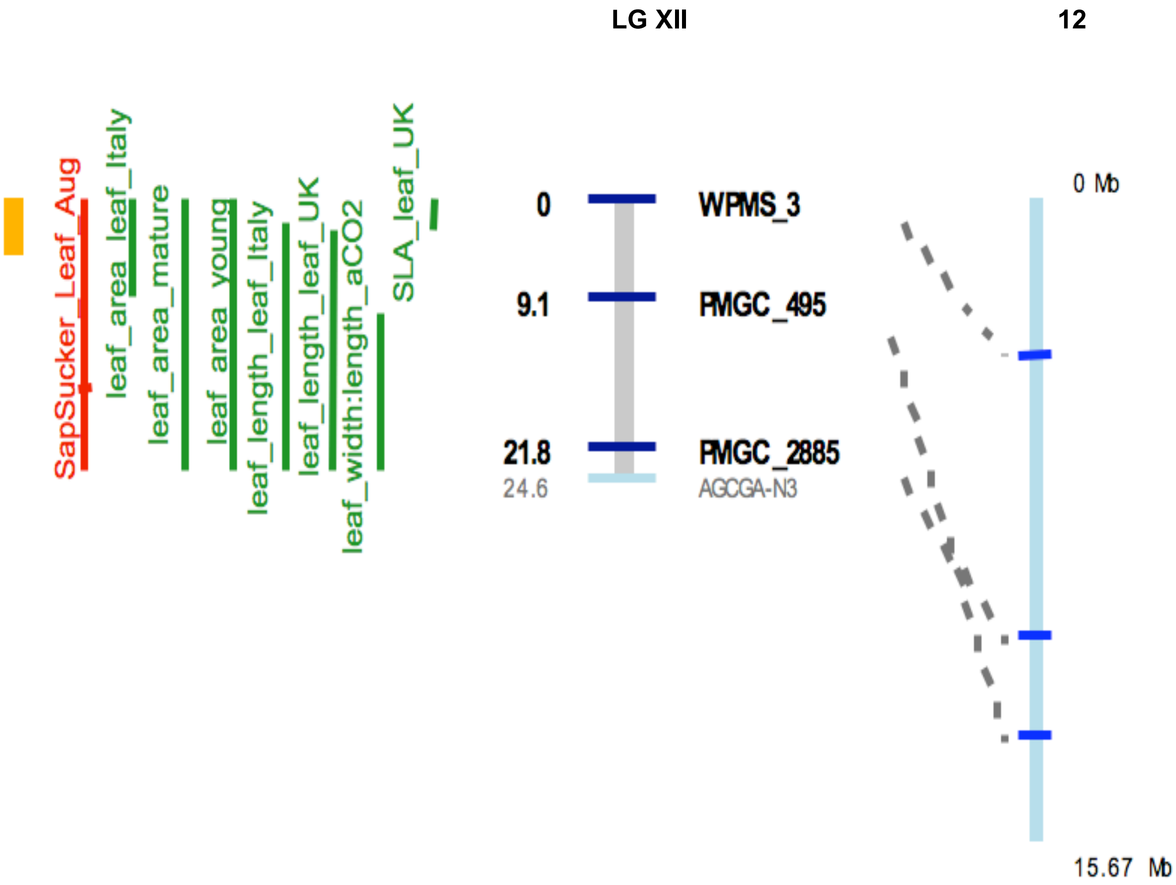


Figure 5.3. Continued.

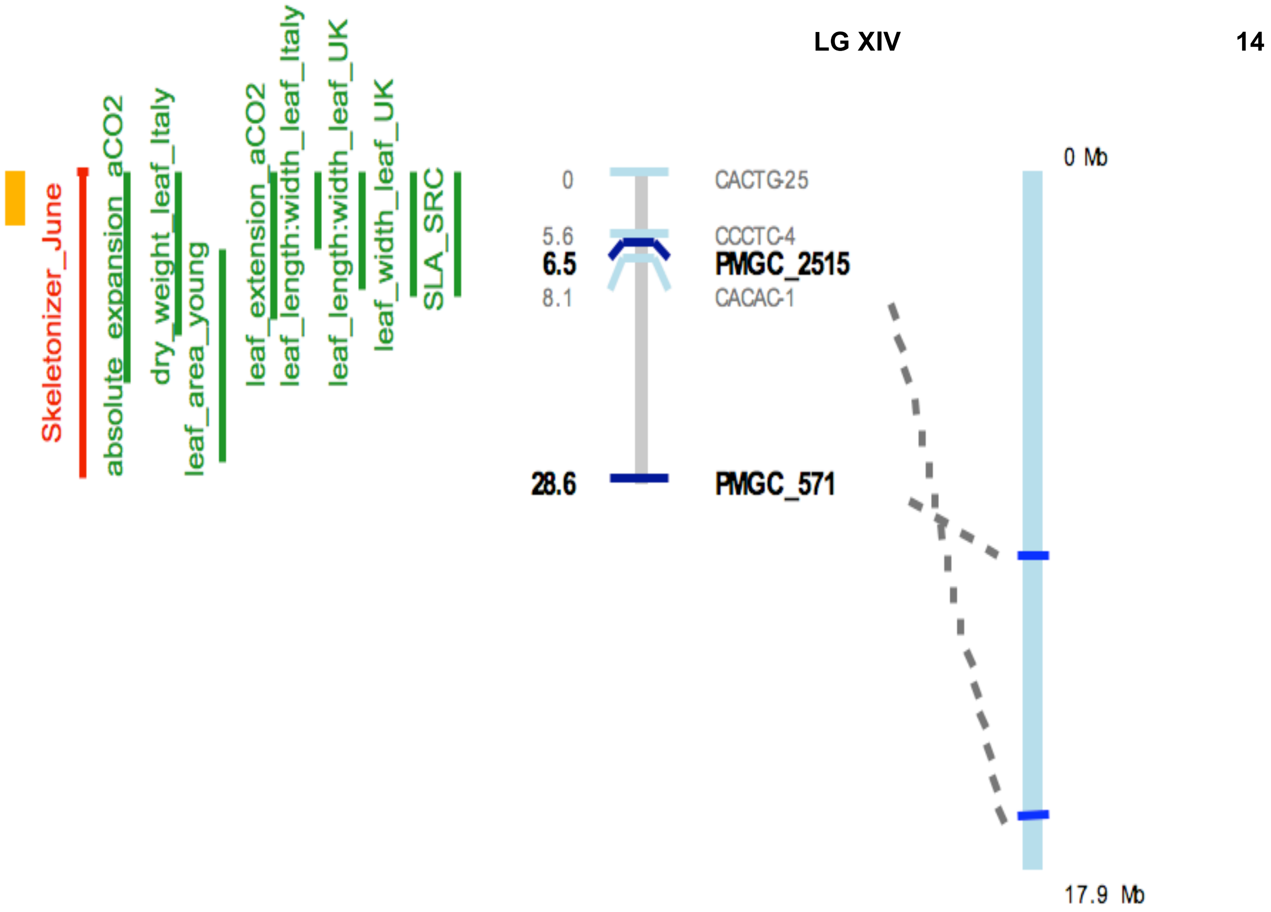


Figure 5.3. Continued.

Between 86 and 1910 gene models (open reading frames identified either from expressed sequence tags or predictive algorithms in the genome annotation) were identified within a single QTL when the *P. trichocarpa* genome was queried (Table 5.6). The high number of genes is partly due to the unsaturated nature of the genetic linkage map for the F_2 pedigree, which resulted in relatively large confidence intervals on some linkage groups. Analysis of the distribution of gene ontology (GO) classifications for each gene model within a category of QTL (e.g. chewer damage in August) revealed a non-random pattern of gene function. Two GO categories displayed non-significant results for one QTL each: the Cellular Components classes were no different than random in the skeletonizer QTL for the June data, and the Molecular Function classes were non-signficant for the skeletonizer QTL for August data (Table 5.5). All other damage category/GO class combinations displayed at least three significant GO categories, with a maximum of 114 GO categories being significantly over-represented in the chewer QTL for August damage (Table 5.6).

Similar GO classes were observed to be over-represented in QTL for the different damage types, indicating this analysis may have resolved functional components underlying the genetic basis of insect association in hybrid poplar. This pattern is apparent when examining the 10 most significantly over-represented categories for each QTL class (Tables 5.7, 5.8, and 5.9). For the Biological Processes classifications (Table 5.7), four GO categories were over-represented in more than half of the QTL categories assessed: GO:0008152 (metabolic processes), GO:0044238 (primary metabolic processes), GO:0006139 (nucleobase, nucleoside, nucleotide and nucleic acid metabolic process), and GO:0044237 (cellular metabolic process). These patterns indicate that genes involved with primary metabolism may be candidates for further study into the genetic basis of insect association. For the Cellular Components classification (Table 5.8), six categories were over-represented in more than half of the QTL categories: GO:0005623 (cell), GO:0044464 (cell part), GO:0005622 (intracellular), GO:0043226 (organelle), GO:0043229 (intracellular organelle), and GO:0044424 (intracellular part). These patterns provide evidence that gene involved in intracellular regulation or modification may be involved in insect host choice on hybrid poplar. For the Molecular Function classification (Table 5.9), two categories were over-represented in more than half of the QTL categories: GO:0003824 (catalytic activity) and GO:0016787 (hydrolase activity). These patterns highlight functional pathways over-represented in the QTL for each insect damage category (Figure 5.4).

Table 5.6. Analysis of gene models occurring within 95% confidence limits of QTL identified for categories of insect damage in hybrid poplar.

Category	**LG**	**QTL (cM)***	**N Gene Models**	**Number of GO Categories Over-represented**		
				Biological Processes	**Cellular Component**	**Molecular Function**
June						
Leaf miners	I	0 – 24	252	37	6	7
Skeletonizers	XIV	0 – 28	86	3	0	4
August						
Chewers	III	29 – 46	301	114	32	66
Chewers	IV	45 – 85	465			
Chewers	Va	62 – 86	396			
Chewers	XVII	35 – 69	455			
Miners	Va	0 – 41	352	60	23	62
Miners	VIIIa	12 – 27	292			
Miners	XVII	33 – 70	505			
Sap Suckers - leaves	I	32 – 125	1910	76	44	76
Sap Suckers – leaves	VI	134 – 144	118			
Sap Suckers - leaves	XII	0 – 24	833			
Skeletonizer	III	0 – 31	328	25	17	0

* The 95% confidence interval of the locus

Table 5.7. The ten most significantly over-represented Biological Processes GO categories in QTL by leaf damage type and month of assessment.

GO Category	**Description**	**N in QTL**	**In Genome**	**Adj. *P***
Leaf miner in June				
GO:0005975	carbohydrate metabolic process	14	1136	0.0017
GO:0044238	primary metabolic process	57	9502	0.002
GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	24	3290	0.0036
GO:0008152	metabolic process	66	12037	0.0036
GO:0044237	cellular metabolic process	49	8506	0.0036
GO:0060255	regulation of macromolecule metabolic process	16	1837	0.0036
GO:0080090	regulation of primary metabolic process	16	1821	0.0036
GO:0019222	regulation of metabolic process	16	1898	0.0044
GO:0050789	regulation of biological process	19	2490	0.0045
GO:0006350	transcription	15	1864	0.0051
Skeletonizer in June				
GO:0005975	carbohydrate metabolic process	13	1136	2.90E-07
GO:0008152	metabolic process	32	12037	0.01
GO:0044238	primary metabolic process	25	9502	0.017
Chewer in August				
GO:0044238	primary metabolic process	365	9502	6.20E-13
GO:0008152	metabolic process	432	12037	2.80E-11
GO:0043170	macromolecule metabolic process	289	7539	4.20E-11
GO:0044237	cellular metabolic process	315	8506	1.80E-10
GO:0009987	cellular process	376	10572	3.70E-10
GO:0044260	cellular macromolecule metabolic process	256	6824	1.50E-09
GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	133	3290	2.80E-07
GO:0006807	nitrogen compound metabolic process	147	3784	5.30E-07
GO:0019538	protein metabolic process	169	4499	5.30E-07
GO:0009058	biosynthetic process	159	4271	1.50E-06
Leaf miner in August				
GO:0044238	primary metabolic process	258	9502	3.00E-08
GO:0043170	macromolecule metabolic process	211	7539	3.10E-08
GO:0008152	metabolic process	310	12037	4.50E-08
GO:0044260	cellular macromolecule metabolic process	192	6824	4.50E-08
GO:0044237	cellular metabolic process	226	8506	1.90E-07
GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	101	3290	1.90E-06
GO:0006807	nitrogen compound metabolic process	109	3784	9.10E-06
GO:0009987	cellular process	258	10572	9.10E-06
GO:0009059	macromolecule biosynthetic process	91	3136	3.60E-05
GO:0034645	cellular macromolecule biosynthetic process	91	3130	3.60E-05
Sap suckers on leaves in August				
GO:0008152	metabolic process	710	12037	4.30E-15
GO:0044238	primary metabolic process	553	9502	6.60E-12
GO:0043170	macromolecule metabolic process	434	7539	1.50E-09
GO:0044237	cellular metabolic process	483	8506	1.50E-09
GO:0006807	nitrogen compound metabolic process	240	3784	3.60E-09
GO:0009987	cellular process	577	10572	6.30E-09
GO:0044260	cellular macromolecule metabolic process	387	6824	2.90E-08
GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	208	3290	3.60E-08
GO:0009058	biosynthetic process	249	4271	1.00E-06
GO:0010467	gene expression	169	2807	1.30E-05

Table 5.7 Continued.

GO Category	**Description**	**N in QTL**	**In Genome**	**Adj. *P***
Skeletonizers in August				
GO:0006350	transcription	17	1864	0.036
GO:0006629	lipid metabolic process	10	731	0.036
GO:0009058	biosynthetic process	33	4271	0.036
GO:0009889	regulation of biosynthetic process	16	1720	0.036
GO:0010468	regulation of gene expression	16	1734	0.036
GO:0010556	regulation of macromolecule biosynthetic process	16	1720	0.036
GO:0019219	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	16	1717	0.036
GO:0019222	regulation of metabolic process	17	1898	0.036
GO:0031323	regulation of cellular metabolic process	16	1743	0.036
GO:0031326	regulation of cellular biosynthetic process	16	1720	0.036

Table 5.8. The ten most significantly over-represented Cellular Components GO categories in QTL by leaf damage type and month of assessment.

GO Category	Description	N in QTL	N in Genome	Adj. *P*
Leaf miner in June				
GO:0016021	integral to membrane	12	1082	0.0044
GO:0031224	intrinsic to membrane	12	1270	0.0085
GO:0044425	membrane part	13	1551	0.01
GO:0016020	membrane	20	3219	0.021
GO:0005623	cell	36	7109	0.024
GO:0044464	cell part	36	7109	0.024
Chewer in August				
GO:0005623	cell	266	7109	6.50E-10
GO:0044464	cell part	266	7109	6.50E-10
GO:0005622	intracellular	170	4246	5.60E-09
GO:0043231	intracellular membrane-bounded organelle	93	2042	6.70E-08
GO:0043227	membrane-bounded organelle	93	2055	7.10E-08
GO:0044424	intracellular part	140	3536	1.10E-07
GO:0005634	nucleus	74	1626	7.90E-07
GO:0043226	organelle	113	2821	7.90E-07
GO:0043229	intracellular organelle	113	2821	7.90E-07
GO:0016020	membrane	109	3219	0.00069
Leaf miner in August				
GO:0005622	intracellular	125	4246	1.30E-06
GO:0005623	cell	187	7109	1.50E-06
GO:0044464	cell part	187	7109	1.50E-06
GO:0044424	intracellular part	105	3536	1.90E-06
GO:0043231	intracellular membrane-bounded organelle	67	2042	6.40E-06
GO:0043227	membrane-bounded organelle	67	2055	6.60E-06
GO:0005634	nucleus	56	1626	7.50E-06
GO:0043226	organelle	84	2821	8.90E-06
GO:0043229	intracellular organelle	84	2821	8.90E-06
GO:0043234	protein complex	33	965	0.00061
Sap suckers on leaves in August				
GO:0005623	cell	438	7109	1.20E-13
GO:0044464	cell part	438	7109	1.20E-13
GO:0005622	intracellular	267	4246	4.40E-10
GO:0043226	organelle	184	2821	1.10E-08
GO:0043229	intracellular organelle	184	2821	1.10E-08
GO:0044424	intracellular part	221	3536	1.10E-08
GO:0043231	intracellular membrane-bounded organelle	141	2042	1.70E-08
GO:0043227	membrane-bounded organelle	141	2055	2.20E-08
GO:0005634	nucleus	109	1626	2.50E-06
GO:0032991	macromolecular complex	101	1499	4.50E-06
Skeletonizers in August				
GO:0005623	cell	56	7109	0.0054
GO:0044464	cell part	56	7109	0.0054
GO:0005622	intracellular	35	4246	0.013
GO:0044424	intracellular part	30	3536	0.013
GO:0016021	integral to membrane	12	1082	0.016
GO:0032991	macromolecular complex	15	1499	0.016
GO:0043226	organelle	24	2821	0.016
GO:0043229	intracellular organelle	24	2821	0.016
GO:0031224	intrinsic to membrane	12	1270	0.034
GO:0043234	protein complex	10	965	0.034

Table 5.9. The ten most significantly over-represented Molecular Function GO categories in QTL by leaf damage type and month of assessment.

GO Category	Description	N in QTL	N in Genome	Adj. *P*
Leaf miner in June				
GO:0016798	hydrolase activity, acting on glycosyl bonds	9	623	0.009
GO:0003824	catalytic activity	64	12409	0.015
GO:0004553	hydrolase activity, hydrolyzing O-glycosyl compounds	8	593	0.015
GO:0003677	DNA binding	15	2000	0.027
GO:0048037	cofactor binding	8	774	0.03
GO:0016787	hydrolase activity	22	3729	0.044
GO:0003700	transcription factor activity	8	891	0.048
Skeletonizer in June				
GO:0004553	hydrolase activity, hydrolyzing O-glycosyl compounds	13	593	1.80E-10
GO:0016798	hydrolase activity, acting on glycosyl bonds	13	623	1.80E-10
GO:0016787	hydrolase activity	19	3729	3.50E-05
GO:0003824	catalytic activity	35	12409	0.0023
Chewer in August				
GO:0003824	catalytic activity	416	12409	1.30E-07
GO:0003676	nucleic acid binding	143	3583	5.30E-07
GO:0003677	DNA binding	86	2000	1.00E-05
GO:0005488	binding	386	12383	4.10E-05
GO:0016787	hydrolase activity	135	3729	6.80E-05
GO:0016740	transferase activity	160	4741	0.00034
GO:0000166	nucleotide binding	158	5094	0.0037
GO:0004175	endopeptidase activity	23	429	0.0037
GO:0004970	ionotropic glutamate receptor activity	7	65	0.0037
GO:0005083	small GTPase regulator activity	9	96	0.0037
Leaf miner in August				
GO:0003676	nucleic acid binding	120	3583	8.60E-09
GO:0003824	catalytic activity	315	12409	2.00E-07
GO:0005488	binding	293	12383	4.90E-05
GO:0016787	hydrolase activity	102	3729	0.00025
GO:0019904	protein domain specific binding	5	20	0.00039
GO:0003677	DNA binding	60	2000	0.00044
GO:0003700	transcription factor activity	33	891	0.00044
GO:0005216	ion channel activity	11	145	0.00044
GO:0015267	channel activity	11	147	0.00044
GO:0016740	transferase activity	120	4741	0.00044
Sap suckers on leaves in August				
GO:0003824	catalytic activity	739	12409	1.90E-16
GO:0016787	hydrolase activity	260	3729	1.10E-13
GO:0005488	binding	699	12383	2.70E-12
GO:0003676	nucleic acid binding	218	3583	4.50E-07
GO:0046872	metal ion binding	164	2557	6.30E-07
GO:0043167	ion binding	164	2576	7.10E-07
GO:0043169	cation binding	164	2576	7.10E-07
GO:0046914	transition metal ion binding	140	2131	9.00E-07
GO:0016491	oxidoreductase activity	155	2435	1.20E-06
GO:0016788	hydrolase activity, acting on ester bonds	64	830	1.60E-05

Figure 5.4. (following pages) GO classes over-represented in QTL for insect damage scored in hybrid poplar relative to the *P. trichocarpa* genome (v. 2.0). A sample of the damage categories and GO class combinations are presented. Each box represents a GO class, with the color indicating the level of significance after multiple testing adjustments, with white boxes representing non-significant categories. Dotted, dashed, and solid lines represent hierarchical relationships between classes with zero, one, or two significant terms connecting the line. The color of the line represents the type of relationship depicted (see key).

(A) Leaf miner damage in June

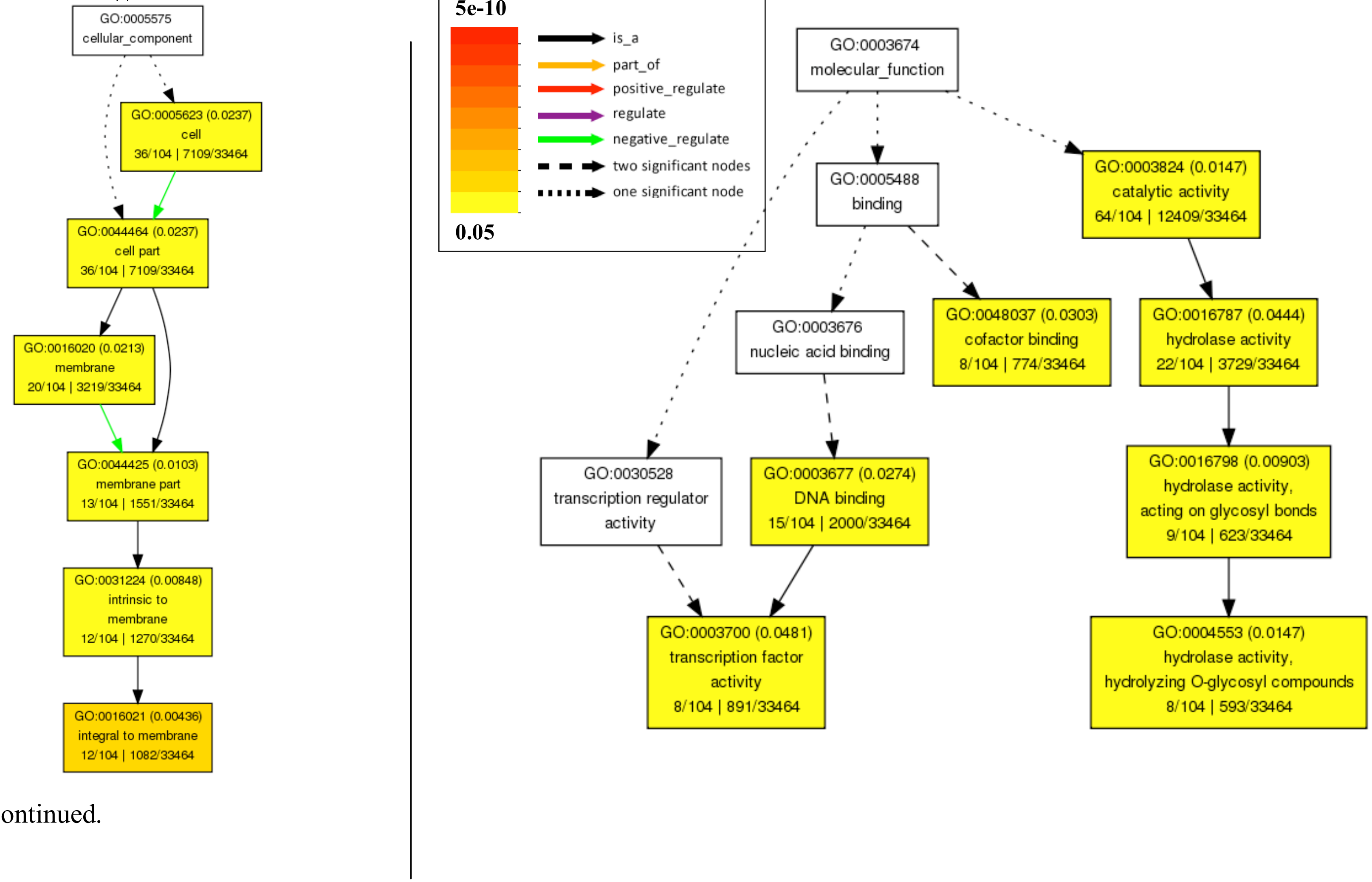


Figure 5.4. Continued.

(B) Skeletonizer damage in June

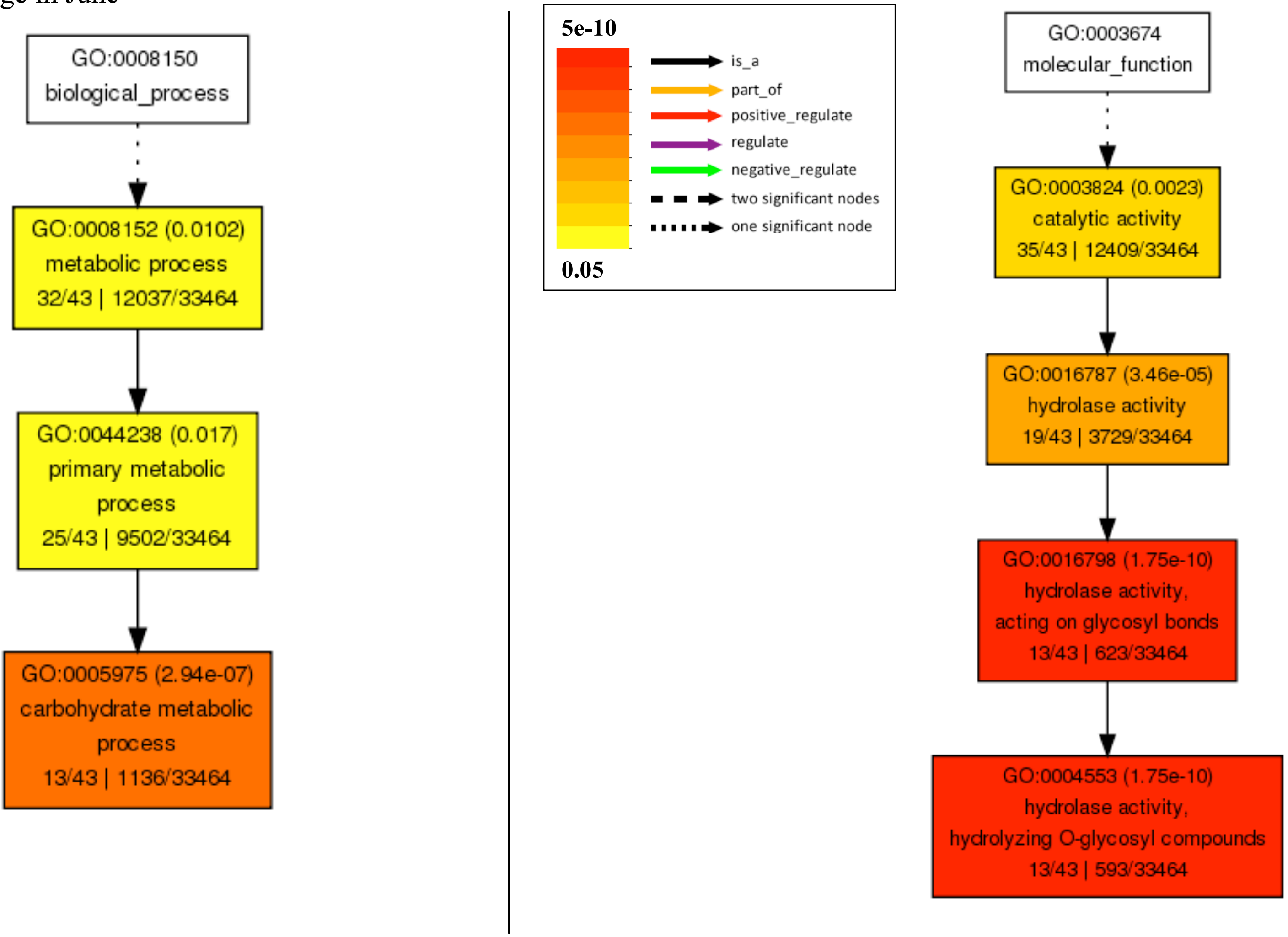


Figure 5.4. Continued.

(C) Leaf miner damage in August

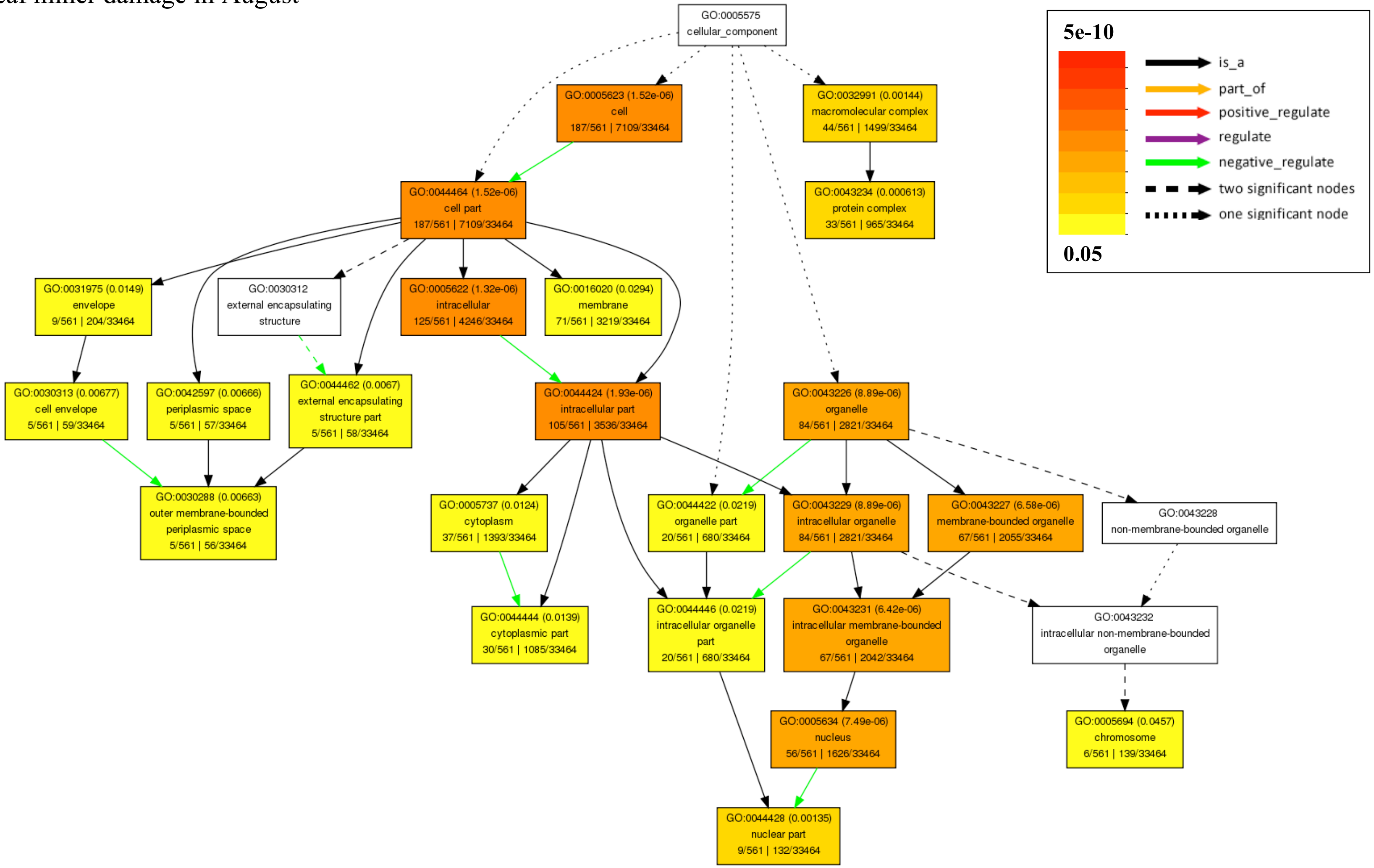


Figure 5.4. Continued.

(D) Skeletonizers in August

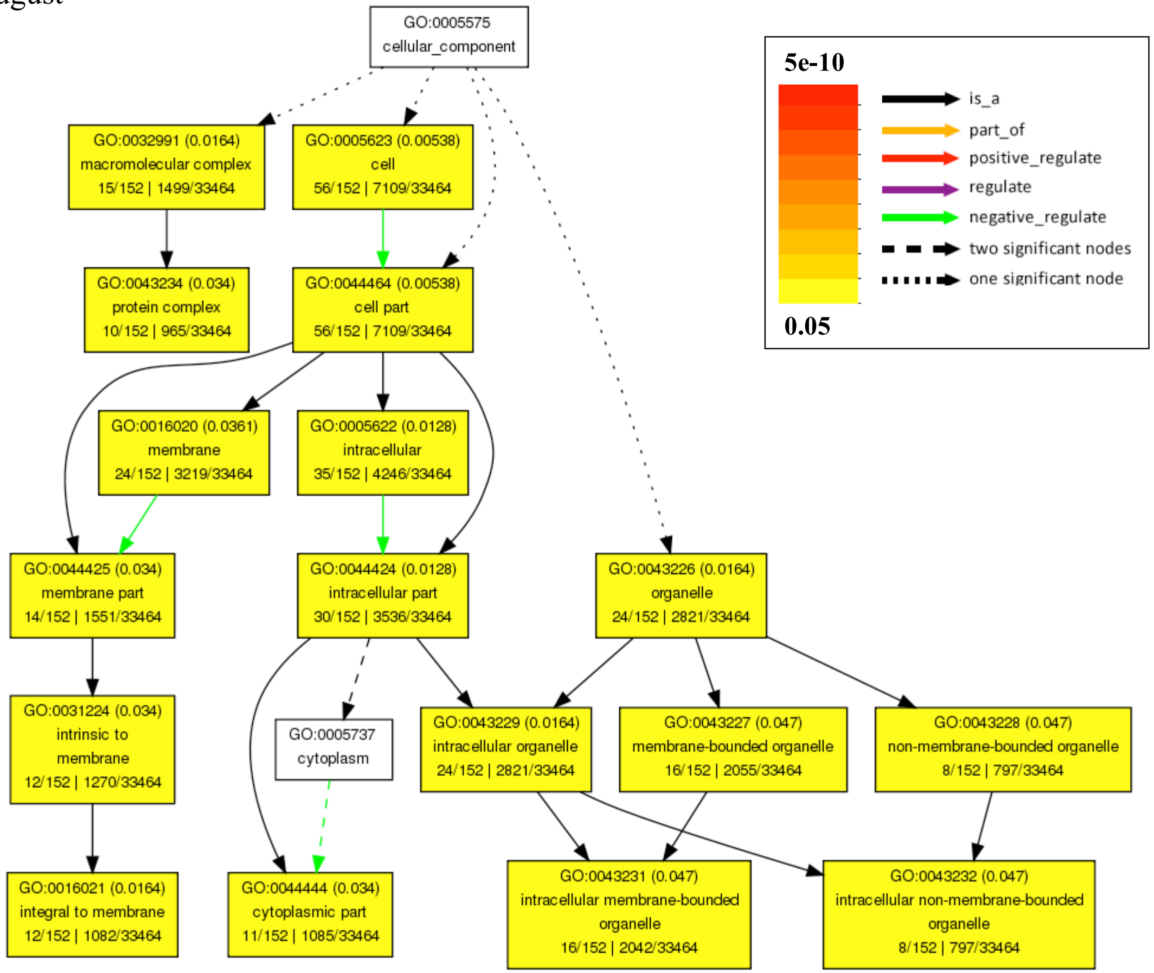


Figure 5.4. Continued.

5.4 Discussion

5.4.1 Phenotypic variation in insect association among F_2 hybrids

This common garden study revealed significant variation in the levels of insect association among F_2 progeny of a hybrid poplar family (POP1). Despite the broad categories used to quantify insect association (chewers, skeletonizers, etc.) and the use of visual inspection of leaf damage (rather than direct insect cataloguing), the significant differences indicate the insect community associates with different individuals (genets) in a non-random fashion. These differences have a genetic basis, with low to moderate heritability estimated for damage categories. This hybrid family of poplar continues to provide insight into the genetic basis of complex traits (Rodriguez-Acosta 2006; Street et al. 2006) (Rae et al. 2006; Rae et al. 2009).

For each damage category, some proportion of F_2 progeny displayed greater levels of insect association (damage) than either F_1 or the *P. trichocarpa* pedigree parents. Such transgressive phenotypes (more extreme than either parent rather than intermediate to the two) are possible outcomes of hybridization events (Rieseberg et al. 1999; Schwarzback et al. 2001). Though the actual *P. deltoides* parent of POP1 is not available in this common garden design, the vigorous growth and low insect damage of a related *P. deltoides* plant (data not presented) and the relatively low to moderate damage in the F_1 trees indicate that higher damage in the F_2 progeny are likely transgressive. The breakdown of co-adaptive gene complexes may provide insight into the adaptive nature of plant defences if specific biochemical differences can be attributed to the different damage levels. For example, coadaptation within the nuclear genome and among nuclear and cytoplasmic genomes has been investigated in the electron transport system in the intertidal copepod *Tigriopus califonicus* (summarized in Edmands et al. 2009). Examination of enzyme efficiency in multiple hybrid lines and backcrosses has revealed the occurrence and extent of the breakdown of coadapted gene complexes to vary. Some lines of cytonuclear hybrids displayed enzyme activity indicative of the breakdown of coadapted gene complexes, but others displayed greater activity with chimeric enzymes, indicating the most efficient protein combinations may not be represented in natural populations (Edmands et al. 2009). In hybrid *Salix*, concentrations of secondary chemicals (e.g. phenolic glycosides and 2’-cinnamoylsalicortin) varied widely in F_1 progeny, consistent with quantitative control of these traits (Orians 2000). The recombination events that occurred between the *P.*

trichocarpa and *P. deltoides* genomes in the establishment of the F_2 progeny likely resulted in novel combinations of biochemical and defensive genes which led to the large variance in leaf damage.

Levels of leaf damage were correlated among damage categories and among seasons. Genets damaged by chewers, skeletonizers or sap suckers on leaves in June were more likely to display higher levels of the same damage in August. This pattern is likely due in part to genetic basis of insect preference when choosing host plants, consistent with the heritable patterns of insect richness on *Populus fremontii*, *P. angustifolia* and naturally occurring hybrids (Wimp et al. 2007; Keith et al. 2010). Correlation within damage types may also reflect the potential sedentary nature of some insect species or life stages (e.g. aphids). Given the positive correlation between levels of skeletonizer damage in June and two significant damage types in August (chewer and leaf roller), assessing the level of skeletonizer association early in the season may serve as an indicator of greater insect community richness and overall susceptibility to attack. Other damage types were negatively correlated with skeletonizer damage occurring in August, however. Genets with skeletonizer damage in August were less likely to display sap suckers on leaves in June or August, or leaf miners in August. This pattern indicates a complex relationship among the phytophagous insect community on poplar, and may reveal avoidance of specific feeder types to the presence of other groups. In contrast to studies of *P. fremontii*, *P. angustifolia* and their hybrids, which identify the presence of an aphid species to be correlated with insect community composition (Keith et al. 2010), these data indicate the presence of skeletonizers to be correlated with insect abundance among the POP1 hybrid progeny. The negative association between insect communities may reflect the induction of facultative defences in response to one or the other herbivores (Young et al. 2010) or seasonal development (Klaper et al. 2001; Wimp et al. 2007).

Co-occurrence of aphids and mutualistic ant species were observed but not quantified on hybrid poplar. The mutualistic relationship between ant and aphid species involves ants tending the aphids, deterring other species of herbivorous insects, and feeding on the aphid honeydew (Styrsky and Eubanks 2007). In *Populus*, sugar levels were found to vary with the combination of aphids and ants present on the plant, indicating that in addition to changes in defensive compounds, changes in primary metabolism may result from aphid feeding (Fischer and Shingleton 2001). In this study, the presence of sap suckers on leaves (primarily aphids) in both assessments were negatively correlated with the occurrence of skeletonizers in August, indicating the ant attendants may affect the community composition in this hybrid pedigree. This negative correlation may confound the genetic associations identified in the QTL (see below). Additional assessment of insect preference in the absence of aphid-ant mutualists (e.g. in a controlled environment) will be required to determine the magnitude of the affect of ant presence on host plant choice of other insect species.

While this study focused on the genetic basis of insect interactions in hybrid poplar, the occurrence of a common fungal pathogen (rust, *Melampsora* spp.) was assessed in both June and August. The genetic basis of rust resistance is well studied in *Populus* (Woo and Newcombe 2003; Duplessis et al. 2009). The data collected for this hybrid pedigree may reveal genomic regions associated with rust resistance (data not presented), but these findings are complicated by the ecological system in which rust abundance was measured. Strict tests of genet susceptibility to *Melampsora* spp. infection in a controlled environment will be more appropriate and effective in identifying major-gene rust resistance in this hybrid pedigree.

5.4.2 Genomic regions correlated with insect association in hybrid poplar

The genetic variance in insect association correlated with genetic variation at markers assessed for the linkage map, identifying between one and four QTL for seven of the damage category/month combinations assessed. No previous study of QTL for insect association in *Populus* has been reported in the peer-reviewed literature, despite the economic and ecological importance of this genus. Each QTL explained a relatively small portion of the variance observed in each trait (between 2.9% and 9.3%), but cumulatively up to 24% of the variation was explained by the combination of identified QTL (e.g. for chewer damage in August). These levels are consistent with or slightly lower than other QTL studies in *Populus* hybrids. For example, a single QTL explained up to 19% of the variance observed in osmotic potential in a hybrid pedigree assessed in two contrasting environments (Tschaplinski et al. 2006), and up to 16.8% of the variance in bud flush (Frewen et al. 2000), while QTL related to biomass explained slightly less (5.0 to 9.8% variance explained) (Rae et al. 2009). QTL for insect damage in willow explained between 8 and 24% of the phenotypic variation (Rönnberg-Wästljung et al. 2006).

An investigation using similar methods in hybrid *Salix* reported similar variation among hybrid progeny and identified QTL for different damage types (Rönnberg-Wästljung et al. 2006). Rönnberg-Wästljung et al (2006) assessed insect damage in three replicate plantations and conducted a QTL analysis for two of the sites. The results indicate that the number and position of QTL vary by site, with a large proportion of QTL observed in one location not detected at another site. Differences in the composition of the local insect community were suggested as a mechanism for the lack of collocation (Rönnberg-Wästljung et al. 2006). In addition, QTL for different traits did not frequently overlap, and the few instances where QTL colocated were considered possible locations of defensive genes or gene complexes (Rönnberg-Wästljung et al. 2006).

Though the present study investigated damage at just one site, a similar pattern in the distribution of QTL among linkage groups was observed. QTL for different damage categories only colocated in one instance, with QTL for chewer and leaf miner damage in August occurring on LG XVII. This genomic region constitutes a candidate for the collocation of defense-related genes. However, unlike the *Salix* study (Rönnberg-Wästljung et al. 2006), this work assessed damage at two timepoints in the season, June and August. The levels of pest damage have been shown to vary throughout the growing season in *Populus* plantations (Coyle et al. 2006). QTL for three damage categories, chewer, leaf miner, and skeletonizer, were identified in both time points, although the loci for different timepoints were on different linkage groups for each category. This lack of colocation within damage types across the season may reflect the dynamic nature of leaf development and insect host choice during the season as well as the consequences of facultative and indirect defense mechanisms within plants. Given that data were collected for a single season, had low replication within the garden for several genets, and had limited power to detect QTL of small effect, these QTL likely constitute an underestimation of the genomic regions influencing these quantitative traits. Simultaneously considering all QTL, regardless of month of assessment, may provide the most complete understanding of the genetic basis of insect association in *Populus* (Rönnberg-Wästljung et al. 2006).

While collocation between insect-related QTL was infrequent, comparison of insect association QTL and leaf development QTL identified four genomic hotspots (i.e. genomic areas where a greater frequency of QTL colocate than predicted at random) where a greater number of QTL overlap than expected by a random distribution. Two of these hotspots did not include QTL for insect association, indicating genes in these regions likely contribute to leaf morphology alone. Two of the hotspots encompassed one insect association QTL: LG XII included a QTL for sap suckers on leaves in

August, and LG XIV included the QTL for skeltonizer damage in June. More precisely, the QTL for sap suckers on leaves colocates to QTL influencing leaf area and specific leaf area (leaf area per unit mass, a measure of thickness). These collocations indicate that sap suckers (here, primarily aphids) may prefer genets with leaves of a certain size or thickness. Additional data including the mean leaf size for each F_2 progeny will be required to quantify the nature and strength of any correlation between these traits. The QTL for skeletonizers in June colocated to QTL correlated with leaf expansion and extension rates, the leaf length to width ratios, and leaf mass and specific leaf area. In this case, skeletonizers appear to prefer leaves with certain growth characteristics, perhaps in a specific developmental stage or thickness. As with the sap suckers on leaves, additional phenotypic data are required to confirm any physical correlation between insect damage and leaf traits in the hybrid progeny. However, the identification of a single QTL for the skeletonizer damage in June, and the significant correlation between this damage and other damage levels in August together indicate the collocation with these leaf traits may reveal a link between leaf development or shape and insect preference. While leaf size is not conventionally thought to influence insect abundance, the effectiveness of physical defences such as trichomes has been established (Turnipseed 1977). In addition, tree phenology, in particular the timing of bud break, may significantly influence the insect community richness throughout the growing season. Stem and leaf modifying herbivores (e.g. gall makers, Wimp et al. 2007) in particular may require host leaves to be at a particular stage of development, such that variation in bud break and the initiation of leaf unfurling may affect host plant selection when the adults emerge. In a common garden study of hybrid poplars, stem and leaf modifying insects were correlated with plant genetics (AFLP diversity) but not plant chemistry, and differences in phenology were suggested as a mechanism for this pattern (Wimp et al. 2007). Correlation between insect association and phenological traits may provide additional insight into links between development and insect preference in this hybrid pedigree as well.

5.4.3 Genomic analyses identify functional categories of genes as candidates for further study

The significant over-representation of certain functional classes of genes (GO categories) may reveal genetic variation underlying insect preference not predicted a priori. These gene classes may constitute candidates for additional investigation, including association tests, functional assays, or sequence analyses. Functional differences in gene products corresponding to insect preference may provide insight into the physiological basis of insect susceptibility or resistance in hybrid poplar.

The GO classes over-represented in more than half of the leaf damage categories provide insight into the potential physiological basis of insect choice. Only two GO classes based on molecular function were over-represented in over half of the damage categories assessed: catalytism (GO:003824) and hydrolase (GO:0016787) activity. Hydrolases are classified based on the type of bond the enzyme hydrolyzes. Though many enzyme families are poorly documented, hydrolases have been shown to perform diverse functions in the cell. Glycoside hydrolases hydrolyze sugars, either oligo- or polysaccharides (Henrissat and Davies 1997), and have been reported to be up-regulated following fungal infection in *Arabidopsis* (Mukherjee et al. 2010). The role of glycoside hydrolases in cellulose biosynthesis and cell wall modification is gaining increasing interest as it relates to biofuels research and plant breeding (Lopez-Casado et al. 2008). Expansins, or xyloglucan endotransglucosylase/hydrolases (XTH), are one category of glucoside hydrolases involved in cell wall disassembly or extension, two processes critical to leaf growth (Cho and Cosgrove 2000; Wagstaff et al. 2010). Serine hydrolases have been implicated in the production of secondary metabolites in plants, including the defense-related jasmonic acid and salicylic acid pathways (Kaschani et al. 2009). Ubiquitin C-terminal hydrolases are involved in auxin signalling by affecting protein stability (Busov et al. 2008). Identifying any differential expression of specific hydrolase genes in response to insect damage may further resolve which enzymes are involved in insect preference on poplar.

While a greater number of cellular component GO classes were over-represented in at least half of the QTL grouped by damage category, all six were involved in intracellular or organellar locations, indicating that variation in intracellular function may be critical to insect choice. The four biological process GO classes over-represented in more than half of the damage categories correspond to broad metabolic processes, and do not identify individual pathways as being critical to insect preference. However, the over-representation of cellular metabolic processes (GO:0044237) and nucleic acid metabolism processes (GO:0006139) correspond to the over-representation of intracellular components GO classes, indicating gene families related to intracellular metabolism may influence insect preference.

Ultimately, conclusions based on these GO analyses are limited by two factors. First, despite the use of the most current version of the *Populus* genome, the incomplete annotation and evidence of function of the majority of genes means that statistical significance may change as additional information is added to future releases. Second, these results investigate only the presence of genes within broad genomic regions, and make no assessment of gene function or genetic variation within coding regions.

5.4.4 Further investigation of the genetic control in insect preference in poplar

This assessment of insect association in hybrid poplar provided evidence of a genetic basis to the significant variation and identified areas of the *Populus* genome correlated with these traits. Several follow-up analyses will be required to relate genetic variation at candidate genes to the phenotypes in a causative manner. Focusing on a single damage type (e.g. chewer) or phytophagous species (e.g. aphids) may improve the power to identify causative alleles. Continuing to assess community-level variation may provide insight into the processes driving ecological relationships in this keystone genus.

First, repeating field measurements and, if possible, at a range of sites will identify robust QTL for insect associations and improve the possibility of identifying loci of small effect. Despite the potential for genotype by environment interactions to complicate QTL studies, multiple assessments can identify genetic hotspots for further analysis (Rodriguez-Acosta 2006; Rae et al. 2009). In addition, a survey of associated insect species with specimen collection and identification will allow estimates of species richness and community complexity which display robust genetic and inter-annual correlation in other *Populus* systems (Bailey et al. 2006; Keith et al. 2010).

Second, additional phenotypic information is required to identify the mechanism of insect preference when significant variation is identified. In addition to the measures of leaf traits discussed above (section 5.4.1), variation in secondary chemicals, physiology (e.g. photosynthesis), or development (e.g. bud break) may correlate with insect preference in some cases. Transcriptional variation assessed early and late in the growing season may identify genetic pathways potentially involved in facultative responses. Collocation of genes displaying significant variation in expression to QTL for that trait can prove a powerful method of narrowing candidate genes for further investigation (Rohde et al. 2011). Assessment of trees displaying extreme phenotypes (e.g. the top and bottom 5% when ranked by damage levels) may streamline such assessments.

Third, actual insect preference can be assessed using food choice tests in a controlled environment. These analyses can quantify preference between genets displaying different levels of damage in the common garden experiment, and confirm that insect choice, and not other factors, may be involved in the patterns. Finally, if allelic variation is identified in candidate genes collating to QTL identified for insect association, and if insect preference is determined to correlate with host plant genotype at that locus, transgenic analyses are required to verify the causative nature of genetic variation on phenotype.

Chapter 6. Modelling the effect of a changing climate on morphological and genetic variation

6.0 Overview

Testing hypotheses of evolutionary timescale is rarely possible using data from extant populations. Simulation studies provide an opportunity to examine the relative effect of different population processes on the evolution of population structure. Here, the Metapop simulation package was used to parameterize demographic, genetic (including a quantitative trait), and adaptive processes, to address questions regarding population divergence, colonization, and changing phenotypic selection. This individual-based simulation model was used to build three evolutionary events: first, population divergence under an isolation event; second, colonization of intermediate population by migrants from either refugia; and third, the effect of changing the intensity of phenotypic divergence among populations. For each event, the effect of varying the diversifying selection among populations, the pattern of phenotypic optima across the landscape, and stabilizing selection within populations was determined. Results indicate that even moderate levels of diversifying selection (variance in phenotypic optima) are sufficient to result in significant genetic and phenotypic differentiation among isolated refugia. Following colonization, the intensity of within-population stabilizing selection had a greater effect than the level of among-population variance in phenotype. Further, the pattern of phenotypic optima across the landscape, and not just the absolute level of variance in optima, significantly affected the phenotypic variation and genetic variance for the quantitative trait. When the intensity of phenotypic diversifying selection changed, little change in the genetic structure at neutral loci was observed, but the phenotypic differentiation and genetic variance varied. Decreases in phenotypic selection tended to result in lower phenotypic differentiation and genetic variance, while increases in phenotypic selection produced opposite trends. These results provide insight into potential adaptive responses of natural populations to changes in climate predicted over the next century.

6.1 Introduction

Previous chapters have shown that the hypothesis of strict local adaptation is insufficient to explain the pattern of morphological differentiation related to leaf shape observed in *Populus nigra*, and that historic vicariance and recolonization events likely influenced the phenotypic diversity observed today. These conclusions were based on the findings that adaptive traits that related leaf shape, cell size and tree biomass were not strictly correlated to the climatic conditions experienced by each natural population, but did vary with genetic differentiation at neutral molecular markers (Chapter 2). These conclusions assume that large-leaf trees from populations colonized from the eastern refugia will have a greater fitness than the local populations in the admixture zone in France, a hypothesis that must be tested by reciprocal transplant common garden experiments. The potential that gene flow since the most recent glacial maximum has been insufficient to introduce the most advantageous phenotypes (alleles) from the eastern refugia into the admixture zone provides an interesting hypothesis for analytical study. Will the current rate of gene flow and adaptation be sufficient for forest populations to adapt to a climate changing at an increasing rate?

Studying the morphology of extant populations restricts the scale of inference. Computer simulations provide the opportunity to investigate which processes significantly affect the evolution of complex traits. With the improvement of desktop computing power, computer simulations have addressed questions regarding postglacial recolonization, complex demographic structuring, and the possible effects of a changing climate on populations. Several studies of metapopulation dynamics in forest trees have shown that rare long distance dispersal events were likely critical to the recolonization following the most recent glacial maximum (Ibrahim et al. 1996), particularly for pine and oak species in Europe (LeCorre et al. 1997b; Austerlitz and Garnier-Gere 2003). The constant wave-front model of colonization from glacial refugia was insufficient to maintain variation in maternally inherited genomes, and did not produce the patchy distribution of neutral variation observed across European populations (Petit et al. 2004). These simulations provided insight into recolonization dynamics and the likely biological processes that took place following glaciations too rare to be documented in fossil records.

Alongside the simulations of recolonization, models incorporating overlapping generations revealed how demographic structuring affects the genetic differentiation in long-lived (tree) species compared with annual plants. In particular, the inclusion of age classes with different relative survival and fecundity improved the models and better represented forest species (Austerlitz et al. 2000). More precisely, the delay in reproduction produced by the juvenile stages allowed the arrival of a greater number of migrants before sexual reproduction took place in newly colonized populations, thereby reducing the initial founder effect and overall genetic differentiation across the landscape (Austerlitz et al. 2000). These patterns were observed without the rare long distance dispersal events found to be critical in other simulations.

The incorporation of selective parameters into models allows investigation of the effects of varying intensities of natural selection on the neutral and adaptive variation within populations. These simulations can model either adaptive differences due to environmental clines or due to changing climate conditions. A one-dimensional model incorporating outbreeding depression, or selection against hybrid individuals, during colonization between two divergent gene pools demonstrated that rare long distance dispersal and underdominance are sufficient to produce a narrow cline in phenotype observed in hybrid zones (Nichols and Hewitt 1994).

Investigations simulating changing climate conditions provide insight into evolutionary changes and possible future adaptation due to climate changes predicted over the next century. Models of very slow environmental changes comparable to those depicted in the fossil record revealed that selection might appear to anticipate environmental change (Sato and Waxman 2008). The categorical nature of genotypic changes in the simple model showed that the rare adaptive allele may increase in frequency faster than the environmental changes, resulting in novel phenotypes having a greater fitness than the current phenotypic optima. These changes were transient as the constant rate of environmental change reduced the fitness of the novel phenotypes (Sato and Waxman 2008). A more complex examination of selection acting on multiple loci in multiple populations over many generations investigated the effects of selection on phenotypic and genetic differentiation and genetic variance for the adaptive trait (LeCorre and Kremer 2003). By starting with nearly homogenous populations (demes), the relative effects of different levels of stabilizing selection within populations, diversifying selection among populations, and the outcrossing rate were identified. In these scenarios, the amount of gene flow among populations was the main determinant of genetic differentiation at low levels of selection, with high levels of gene flow homogenizing both genetic and phenotypic differentiation. When diversifying selection was strong, high levels of gene flow homogenized neutral markers but QTL differentiation remained elevated, creating heterogeneous levels of variance across the genome (LeCorre and Kremer 2003).

Ultimately, the ability of natural populations of forest trees to adapt to a changing climate may be compromised by the long-lived nature of these species. A recent detailed simulation parameterized to *Pinus sylvestris* (Scots pine) and *Betula pendula* (Silver birch) investigated the effects of demographic processes (age of maturation, dispersal, and mortality) on the potential of populations on a steep environmental cline to adapt to a changing climate (Kuparinen et al. 2010). In these scenarios the evolutionary adaptation of the populations lagged the environmental change (the mean phenotype did not reach the optima), but an increase in tree mortality lessened the lag of adaptation (Kuparinen et al. 2010). The authors conclude that the long lifespan and low mortality of extant populations minimizes the potential for adaptation, as natural selection is most effective on the juvenile stages and not mature trees (Kuparinen et al. 2010). This sample of previous simulation studies demonstrates the power of the technique to assess demographic and evolutionary processes of forest trees.

Given the power of simulation models to explore the relative effects of neutral and selective processes on genetic variation, the simulation program METAPOP will be used to explore the evolution of a quantitative trait in an idealized forest tree. This simulation package combines the models reported in LeCorre et al. (1997a), LeCorre and Kremer (2003) and Austerlitz and Garnier-Gere (2003) and uses information from demographic, quantitative genetic and selective parameters to simulate differentiation between populations of a forest tree species over time. By simulating a large vicariance event, colonization and migration between refugia, and various diversifying selection scenarios, this study addresses two objectives. First, how does the pattern of the distribution of phenotypic optima (advantageous morphologies) affect the genetic and phenotypic differentiation among populations? Second, what is the effect of changing the phenotypic optima and intensity of diversifying selection on the fitness, genetic differentiation, and phenotypic variation of a forest species?

6.2 The METAPOP Model

The METAPOP software used in this study combines functions from two simulation packages previously described in LeCorre and Kremer (2003) and Austerlitz and Garnier-Gere (2003). This deterministic model simulates populations over multiple generations to investigate the effects of varying migration, mutation, selection, and demographic parameters on population fitness and genetic structure. The power of the METAPOP simulation project lies in its ability to simulate demographic, genetic, and selective factors across multiple populations.

The METAPOP model is built on three categories of parameters: demographic, genetic, and selection factors (Table 6.1). The demographic parameters simulate a number *n* of populations composed of a maximum of *N* individuals. The program allows the definition of multiple age classes within each population, which simulates the overlap of generations, a trait that is critical to model the genetic structure of forest tree species accurately (Austerlitz and Garnier-Gere 2003). By defining the relative survival rate within an age class, transition rate between age classes, male and female fecundity, vegetative fecundity, and size ratio for each class, a complex life history model can be simulated for each population. In addition, gene flow is partitioned into seed and pollen dispersal components, with the probability of dispersal between each pair of populations defined in flow matrices.

The second category of parameters defined in the METAPOP software includes the genetic components. For the nuclear and cytoplasmic genomes, the number and type of loci assessed, the number of alleles per locus, the number of chromosomes in the nuclear genome and the recombination rate between nuclear loci are specified. In addition, the number and identity of loci under selection (linked to simulated QTL for the trait) is defined within the context of the nuclear linkage map. The relative paternal inheritance of cytoplasmic loci can be defined, as can the mutation rate for each nuclear and cytoplasmic locus. The initial allele frequencies for each population at each locus are specified as the prior values for the simulation.

The type and intensity of selection acting on the simulated quantitative trait are specified in the third category of parameters in METAPOP. Simulation is simulated as either genotypic or phenotypic, and phenotypic selection can be defined as either directional or stabilizing within each population. The intensity of phenotypic selection is also parameterized within each population. The strength and variance of diversifying selection among populations is defined by specifying a phenotypic optimum for each population.

The output of simulation includes the genetic differentiation at all loci, loci linked to QTL, and the quantitative differentiation among populations for the trait under selection. In addition, the relative fitness of each age class and population, estimates of inbreeding, additive, epistatic and genetic variance, and the variance of population mean estimates are logged. Finally, the software may produce a file of sample genotypes for each population in the simulation for the final generation of the scenario, which can then be used as the prior values in downstream simulations. By iterating the model using the posterior population values as the priors for the following generation, the evolution of genetic and phenotypic divergence can be simulated over multiple populations.

METAPOP was designed to simulate a canonical forest tree species, and as a result *Populus* species do not meet one of the key assumptions of the model. In particular, the model assumes trees to be monoecious, while *Populus* species are dioecious and have male and female flowers on separate genetic individuals. Dioecy is expected to increase outcrossing and gene flow among individuals, decrease fine-scale spatial autocorrelation within populations, and decrease differentiation among populations, particularly in long lived species like forest trees (Hamrick and Godt 1996). As a consequence, this study describes the potential evolutionary and adaptive consequences in a canonical forest tree species, not a dioecious species such as *Populus nigra*. The approach and results will be informative for the management and conservation of forest tree species.

Table 6.1 Parameters modeled by the METAPOP software to simulate the evolution of quantitative traits in multiple populations.

Demographic parameters	Genetic Parameters
Number of populations	Number of loci
Maximum number of individuals	Number of alleles per locus
Number of age classes	Recombination rates among loci
Relative size of each age class	Mutation rates
Survival and transition between classes	Number and linkage position of QTL
Fecundity	Paternal inheritance of cytoplasmic loci
Rate of outcrossing	Allele frequencies
Migration among populations	
Selection parameters	Output measurements
Type of selection within populations (genotypic or phenotypic)	Allele frequencies Genetic diversity
Intensity of selection within populations	Inbreeding
Intensity of selection among populations	Genetic differentiation
Additive, epistatic and dominance effects	Phenotypic differentiation
	Fitness

6.3 Methods

6.3.1 Model paramterization

In order to determine the effects of historic vicariance, population recolonization, and then changing phenotypic optima on the fitness, genetic structure and phenotypic differentiation, a number of model parameters were constant among the three scenarios (Table 6.2). For the demographic parameters, 100 populations were simulated on a 5 x 20 grid, with a maximum population of 1000 individuals within each. Not all populations were populated at the beginning of simulations (see section 6.3.2). Each populations was divided into four age classes, with two age classes representing seedling and sapling individuals, and thus incapable of reproduction. The two mature age classes differed in fecundity and size, but both classes were included in estimates of genetic differentiation. The relative size of each age class was proportioned that the maximum of 1000 individuals referred to age class 3 and corresponded to a maximum of 10000 individuals in age class 1 (seedlings), 2000 individuals in age class 2 (saplings), or 500 individuals of age class 4. In practice the maximum number of effective individuals was proportioned among age classes so that more than 1000 individuals may actually occur. No vegetative reproduction or self-fertilization was allowed. Total gene flow was partitioned into seed and pollen migration following (Petit et al. 1993), with the overall effective migration set to $N_e m$=10 per generation. Initiation tests indicated this to be the lowest migration rate sufficient for complete recolonization of the population space in a time span equivalent to that estimated for the post-Pleistocene migration (see section 6.3.2).

Several genetic parameters were constant throughout all scenarios (Table 6.2). The genetic parameters consisted of 100 nuclear AFLP (amplified fragment length polymorphism) loci simulated as dominant markers for estimates of genetic diversity and differentiation. A maximum of two alleles were allowed per locus. The 100 nuclear loci were evenly assigned to 10 linkage groups within recombination rates varying from 0.00025 to 0.5 (unlinked loci). Initial allele frequencies were determined from the Dirichlet distribution, with each subsequent scenario using the resulting allele frequencies as priors. The quantitative trait was linked to ten of the nuclear ALFP loci that served as QTL. Each QTL was located on a separate linkage group. Only additive genetic effects were assessed in these simulations; neither epistasis nor dominance was allowed.

The effects of the intensity of stabilizing selection within populations and diversifying selection among populations on phenotypic and genetic differentiation were explored through six different scenarios. Two levels of within-population stabilizing selection were examined: low selection (10) and moderate selection (1). Three levels of diversifying selection among populations were examined, quantified by the variance in phenotypic optima (V_{Zopt}): low (V_{Zopt} =1), moderate (V_{Zopt} =5) and high intensity (V_{Zopt} =10).

Table 6.2 Parameters assessed in all scenarios of population isolation, colonization, and changing selective optima examined through simulations.

Parameter	**Values examined**
Demographic parameters	
Number of populations	100
Maximum number of individuals	1000
Number of age classes	4
Relative size of each age class	0.1, 0.5, 1, 2
Survival probability within classes*	0, 0.4, 0.7, 0.9
Transition rate between classes	0.2, 0.5, 0.2
Relative fecundity of the two mature classes	1, 2
Rate of outcrossing	1.0
Genetic parameters	
Number of nuclear loci	100
Type of nuclear markers	AFLP dominant
Number of alleles per locus	2
Number of linkage groups	10
Mutation rate	1×10^{-6}
Number of QTL (nuclear markers under selection)	10
Distribution of QTL among linkage groups	Even, one per LG
Selection parameters	
Selection acted on	Phenotype
Hard or soft selection	Soft
Within population selection	Stabilizing
Among population selection	Diversifying
Additive allelic effects	Yes
Variance for dominance	0
Variance for epistasis	0

*The probability of an individual surviving in the same age class in the next generation

6.3.2 Simulating refugia isolation and population colonization

The initial parameters were used to simulate vicariance events where two refugia evolved under low ($V_{\mathrm{Zopt}} = 1$) or high ($V_{\mathrm{Zopt}} = 10$) diversifying selection. The two refugia were defined as ten populations each at either end of the 5 x 20 grid (Figure 6.1). Within each refuge, gene flow followed the island model, and no migration occurred beween refugia. The two levels of diversifying selection were crossed by low (10) or moderate (1) stabilizing selection within populations, for a total of four vicariance scenarios examined (Table 6.3). The vicariance event was simulated for 5000 generations, and each scenario was repeated 20 times.

The resulting population sizes and allele frequencies were used as the priors for simulations of colonization across the grid of populations. Colonization of the grid took place from the outer refugia in towards the center of the habitat over eight steps. At each step, migration was expanded to include the adjacent column of 5 populations following a two dimensional (2D) stepping stone model of migration (Figure 6.1). At the end of the eight rounds of colonization, the entire grid of populations were populated with the two refugia theoretically connect via the 2D stepping stone migration across the intermediate populations. Each round of colonization occurred over 500 generations, and following colonization the 2D stepping stone model was maintained for each scenario for 2000 generations, for a total of 6000 generations.

Twelve selection scenarios were examined during the recolonization (Table 6.3). Three levels of diversifying selection were examined: low ($V_{\mathrm{Zopt}} = 1$), moderate ($V_{\mathrm{Zopt}} = 5$), or high ($V_{\mathrm{Zopt}} = 10$). These three levels of variance were achieved through two distributions of phenotypic optima across the landscape of the population grid. First, the phenotypic optima were evenly distributed across the grid, with four quadrants of 25 populations assigned each optimum (Figure 6.2). Second, the phenotypic optima were unevenly distributed across the grid, with populations first colonized by the leftmost refuge being selected for a phenotype more similar to the distant refuge (Figure 6.3). This uneven distribution of phenotypic optima across the grid will provide insight into the possible evolutionary constraints of colonization into suboptimal environments. Finally, each model of diversifying selection (three levels of variance either evenly or unevenly distributed) was examined under low (10) and moderate (1) within-populations stabilizing selection. Each scenario was repeated 20 times.

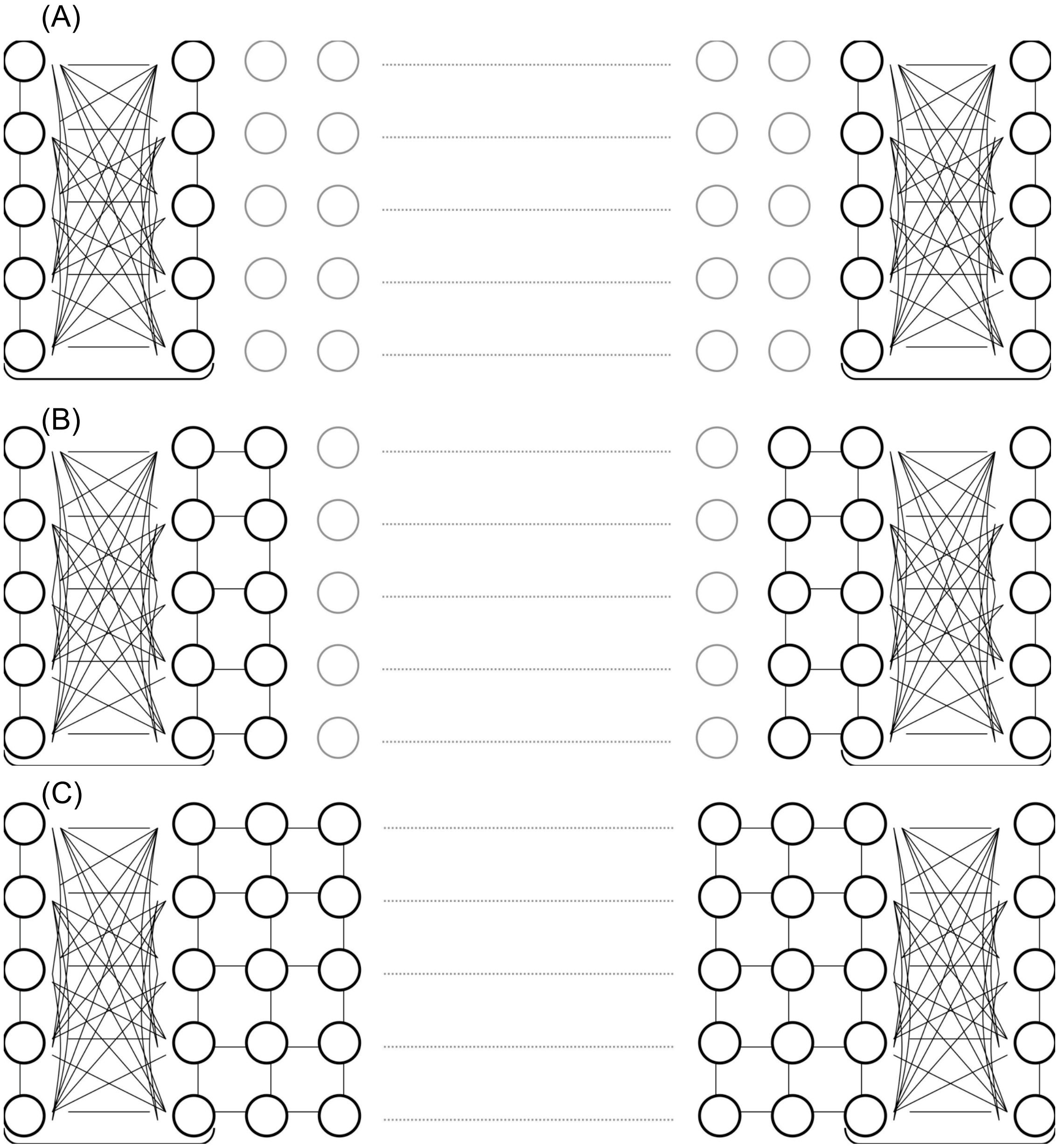


Figure 6.1 One hundred populations (circles) were simulated to investigate the effects of varying levels of diversifying selection on neutral and adaptive genetic structure. Lines represent gene flow between populations. (A) Two refugia composed of ten populations each (black circles marked by brackets) were simulated at the ends of a 5 by 20 grid of populations. (B) Following 5000 generations of isolation under the island model, the 80 populations between refugia (gray circles) were recolonized ten populations at a time. (C) After colonization to populate the 80 intermediate populations, gene flow between the two refugia followed the 2D stepping stone model for 2000 generations.

Table 6.3. Parameters used to explore the effects of vicariance, colonization and changing phenotypic optima on the phenotypic and genetic differentiation of populations. Each scenario was examined with 20 replicate simulations.

Parameter	**Values examined**
Vicariance scenarios (*N*=4)	
Variance in phenotypic optima among populations (V_{Zopt})	1, 10
Intensity of stabilizing selection within populations	10, 1
Colonization scenarios (*N*=12)	
Variance in phenotypic optima among populations (V_{Zopt})	1, 5, 10
Pattern of distribution of phenotypic optima	Even, uneven
Intensity of stabilizing selection within populations	10, 1
Changing phenotypic optima scenarios (*N*=24)	
Novel phenotypic optima, based on V_{Zopt} from prior scenario	{1→5
	1→10
	5→1
	5→10
	10→1
	10→5}
Pattern of distribution of phenotypic optima	Even, uneven
Intensity of stabilizing selection within populations	10, 1

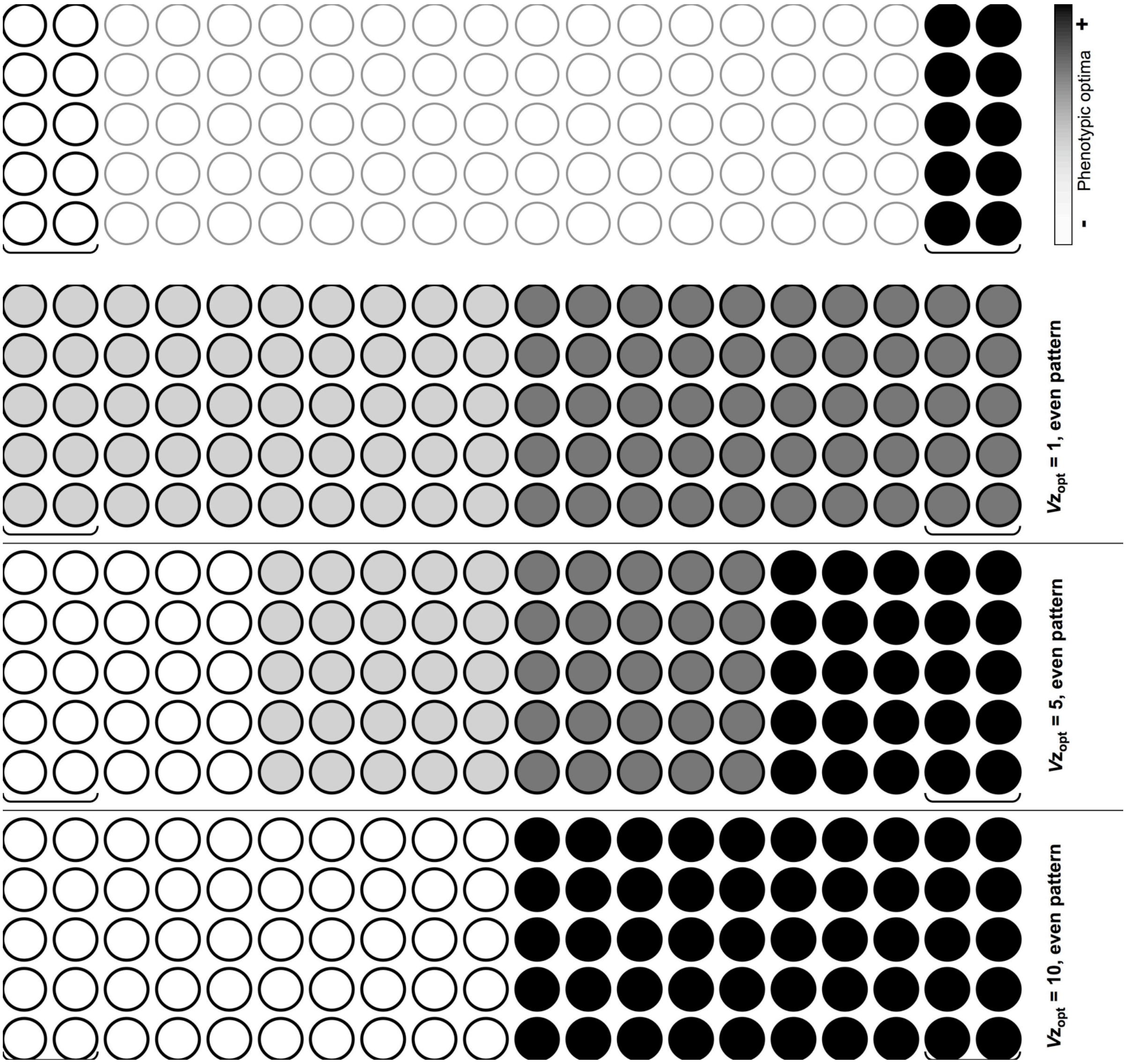


Figure 6.2 Three levels of variance in phenotypic optima (V_z), a measure of diversifying selection, were simulated with an even distribution of phenotypic optima across the grid. Each circle represents a population, with the initial refugia marked by brackets. The shading of the circle corresponds to the phenotypic optimum defined for that population.

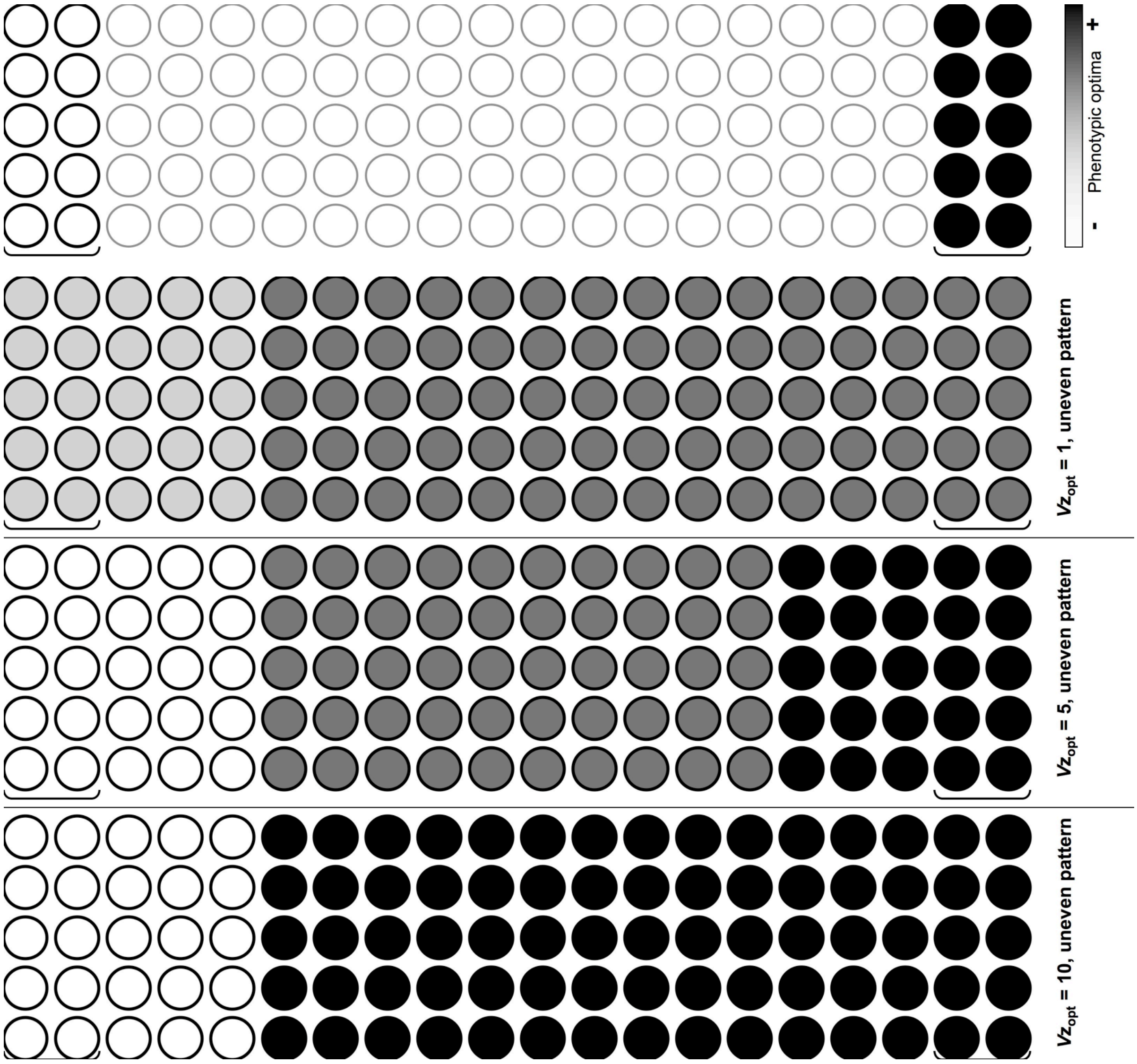


Figure 6.3 The same levels of differentiation (V_z) were modeled with phenotypic optima unevenly distributed along a grid in order to determine the effects of patterning on genetic and phenotypic differentiation. Symbols follow Figure 6.2.

6.3.3 Simulating the effects of changing phenotypic optima

The final population sizes and allele frequencies from each of the 12 colonization scenarios were used as priors to investigate the effects of a changing phenotypic optimum on populations fitness, genetic variance and phenotypic differentiation. A total of 24 scenarios were examined. For each colonization scenario, the level of population differentiation (quantified as variance in phenotypic optima, V_{Zopt}), was changed to the other two values to assess the effects of increasing or decreasing the intensity of diversifying selection (Table 6.3). For example, the final population sizes and allele frequencies from the low diversifying selection ($V_{\text{Zopt}}=1$) were used as priors for simulations of moderate ($V_{\text{Zopt}}=5$) and high ($V_{\text{Zopt}}=10$) variance in phenotypic optima. Six scenarios changing the phenotypic optima were conducted on the even patterns and uneven patterns of phenotypic optima, and each was tested for low (10) and moderate (1) within-population stabilizing selection. Each scenario was replicated 20 times.

6.3.4 Data analysis

Five measures of phenotypic and genetic structure among populations were assessed at the end of the vicariance, colonization, and changing optima scenarios. The overall genetic difference (gene differentiation) among populations was quantified as (Nei 1973) G_{ST}. In order to assess the effects of the various selection scenarios on gene differentiation, G_{ST} was estimated over all loci (G_{STall}) as well as only for loci linked to QTL (G_{STqtl}). Phenotypic differentiation was quantified as Q_{ST} (Spitze 1993). In order to assess the effects of different scenarios on the 'adaptability' of populations, the additive genetic variance (V_{A}) and heritability (h^2) were assessed. Differences among scenarios were assessed using general linear models, with the basic model:

$$Y_{ijk} = \mu + \upsilon_i + \omega_j + \phi_k + \varepsilon \tag{6.1}$$

where μ is the grand mean, υ_{I} is the i^{th} level of V_{Zopt}, ω_{j} is the j^{th} level of within-population selection intensity, ϕ_{k} is the k^{th} category of even or uneven distribution of phenotypic optima, and ε is the residual error. For scenarios of a change in phenotypic optima, categorical variables were used to distinguish υ_{I}. GLM analyses were implemented in SPSS (SPSS, Inc., Chicago, Illinois).

For the changing optima scenarios, in addition to the above measures, estimates of fitness for each population were calculated as the mean fitness over age classes.

Means of ten simulations were used to examine population fitness over generations under the changed phenotypic optima.

6.4 Results

6.4.1 The effects of divergent selection among isolated refugia on phenotypic differentiation among colonized populations

Two scenarios simulating isolation of refugia assessed the effects of varying the level of variance in phenotypic optima (V_{Zopt}), or the intensity of diversifying selection, on the genetic and phenotypic differentiation of populations. For these scenarios, the phenotypic optima of each population were equal within each refugium, but differed among refugia. The different levels of V_{Zopt} were sufficient to result in significant differences in genetic and phenotypic differentiation, as well as in the genetic variance and heritability in the adaptive trait (Table 6.4). The scenario having a lower V_{Zopt}, meaning the phenotypic optima were more similar across populations, resulted in lower levels of gene differentiation measured at all markers (G_{STall}) and those linked to QTL (G_{STqtl}) as well as lower levels of phenotypic differentiation (Q_{ST}) (Figure 6.4). Genetic variance and heritability for the adaptive trait also differed with V_{Zopt}, indicating that the levels of divergence between refugia will affect the genetic variance available for further adaptation. The intensity of stabilizing selection within each population also affected the measures of genetic and phenotypic differentiation. The effect was particularly pronounced for G_{STqtl} at low levels of differentiation (V_{Zopt}=1). However, the level of genetic variance (V_{G}) for an adaptive trait was determined by the levels of diversifying selection among populations and not by the stabilizing selection within populations (Table 6.4).

Scenarios modeling the colonization of populations between the two refugia investigated differences in V_{Zopt}, balancing selection within populations, and the distribution of phenotypic optima across the grid or landscape (the evenness of phenotypic optima). Each of these variables significantly affected levels of differentiation measured for genetic and phenotypic parameters (Table 6.5). Genetic and phenotypic differentiation generally increased with the variance in phenotypic optima, though the trend was more pronounced with lower stabilizing selection within populations (Figure 6.5). Greater levels of V_{Zopt} were particularly associated with higher levels of genetic variance (Figure 6.5). As with the scenarios of isolated refugia, the level of stabilizing selection within populations did not affect the genetic variance

(Table 6.5), indicating that V_{Zopt} was the determinant of genetic variance for an adaptive trait. The evenness of the distribution of phenotypic optima (Figures 6.2 and 6.3) also affected the levels of differentiation and variance. In general, the uneven pattern, where populations colonized from the leftmost refugium experience selection in favor of a phenotype more similar to the rightmost refugium, decreased genetic and phenotypic differentiation but increased genetic variance and heritability (Figure 6.5). As with the simulations of isolated refugia, the variance among the 20 replicates of each scenario was low, resulting in significant effects even when the absolute difference between measures (e.g. G_{STall}) were unlikely to be biologically meaningful.

Table 6.4. Varying levels of variance in phenotypic optima (V_{Zopt}) in scenarios of isolated refugia leads to significant variation in gene differentiation among all loci (G_{STall}) and among loci linked to QTL (G_{STqtl}), phenotypic differentiation (Q_{ST}), genetic variance (V_{G}), and heritablility (h^2) in the quantitative trait.

Source	**df**	**MS**	***F***	***P***
G_{STall}				
Variance in phenotypic optima	1	0.032	156.4	<0.001
Intensity within populations	1	0.020	97.32	<0.001
Error	77	0.000		
G_{STqtl}				
Variance in phenotypic optima	1	0.929	68.63	<0.001
Intensity within populations	1	1.202	88.77	<0.001
Error	77	0.014		
Q_{ST}				
Variance in phenotypic optima	1	0.008	83.90	<0.001
Intensity within populations	1	0.007	64.39	<0.001
Error	77	<0.000		
V_{G}				
Variance in phenotypic optima	1	0.186	370.6	<0.001
Intensity within populations	1	0.001	1.041	0.311
Error	77	0.001		
h^2				
Variance in phenotypic optima	1	0.002	68.92	<0.001
Intensity within populations	1	0.003	96.90	<0.001
Error	77	3.5E-5		

Table 6.5. Varying levels of variance in phenotypic optima (V_{Zopt}) and the distribution of the phenotypic optima across the population grid leads to significant variation in genetic and phenotypic structure in populations colonized from two divergent refugia. Abbreviations follow Table 6.5.

Source	**df**	**MS**	***F***	***P***
G_{STall}				
Even/uneven distribution	1	0.008	29.28	<0.001
Variance in phenotypic optima	2	0.048	182.7	<0.001
Intensity within populations	1	0.071	268.0	<0.001
Error	235	<0.001		
G_{STqtl}				
Even/uneven distribution	1	0.066	14.42	<0.001
Variance in phenotypic optima	2	0.387	84.59	<0.001
Intensity within populations	1	1.489	325.1	<0.001
Error	235	0.005		
Q_{ST}				
Even/uneven distribution	1	0.147	63.27	<0.001
Variance in phenotypic optima	2	0.279	120.0	<0.001
Intensity within populations	1	0.339	145.5	<0.001
Error	235	0.002		
V_{G}				
Even/uneven distribution	1	54.69	820.5	<0.001
Variance in phenotypic optima	2	18.59	278.9	<0.001
Intensity within populations	1	0.046	0.068	0.409
Error	235	0.067		
h^2				
Even/uneven distribution	1	0.101	93.86	<0.001
Variance in phenotypic optima	2	0.036	33.50	<0.001
Intensity within populations	1	0.364	337.6	<0.001
Error	235	0.001		

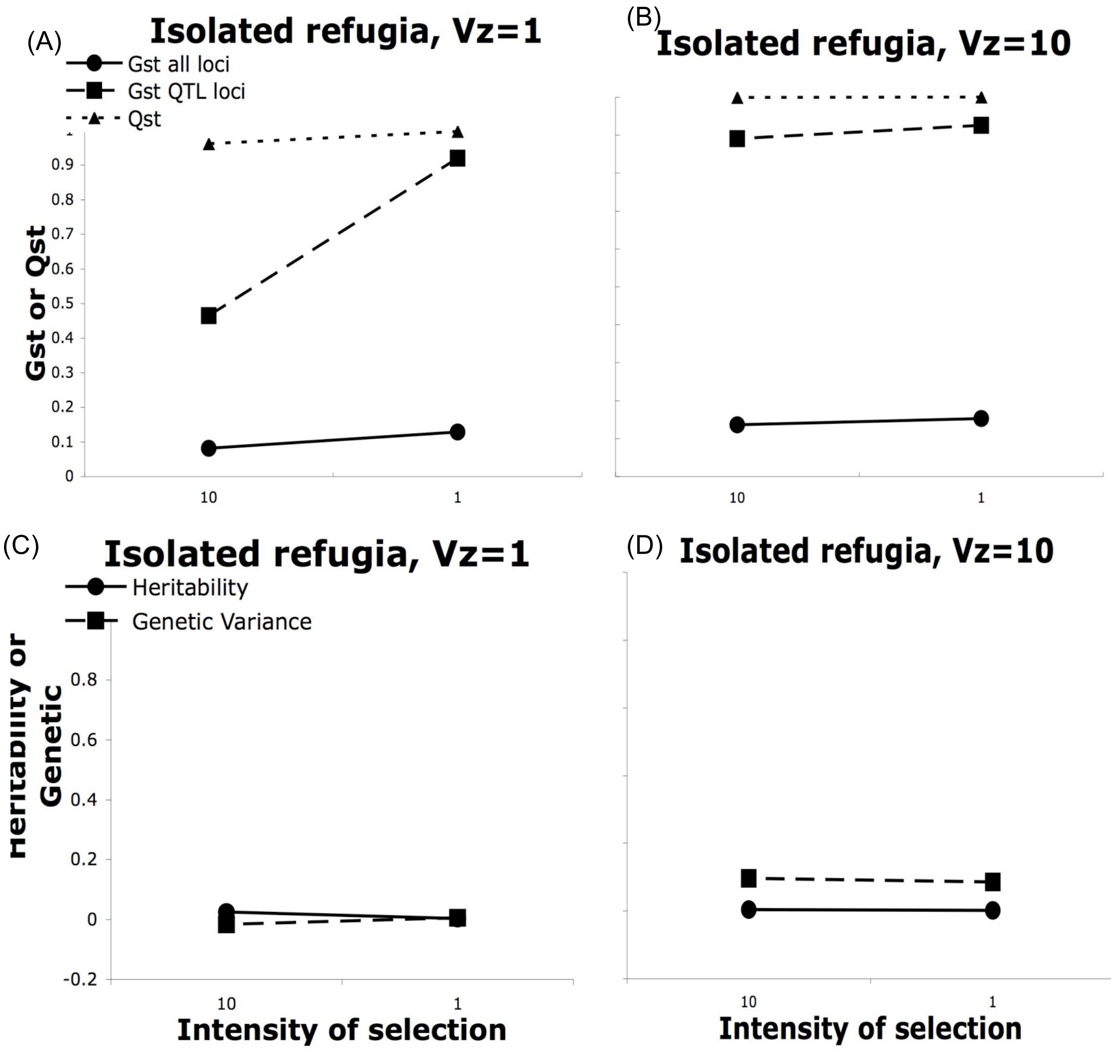


Figure 6.4 Simulation of isolated refugia under divergent selection resulted in (A, C) high levels of genetic and phenotypic differentiation but (B, D) low levels of heritability and genetic variance. Each scenario involved two refugia of 10 populations experiencing low (*V*zopt=1) or high (*V*zopt=10) levels of divergent selection and either moderate (10) or high (1) within-population stabilizing selection for 5000 generations. Plots represent mean values for 20 replicates of each scenario.

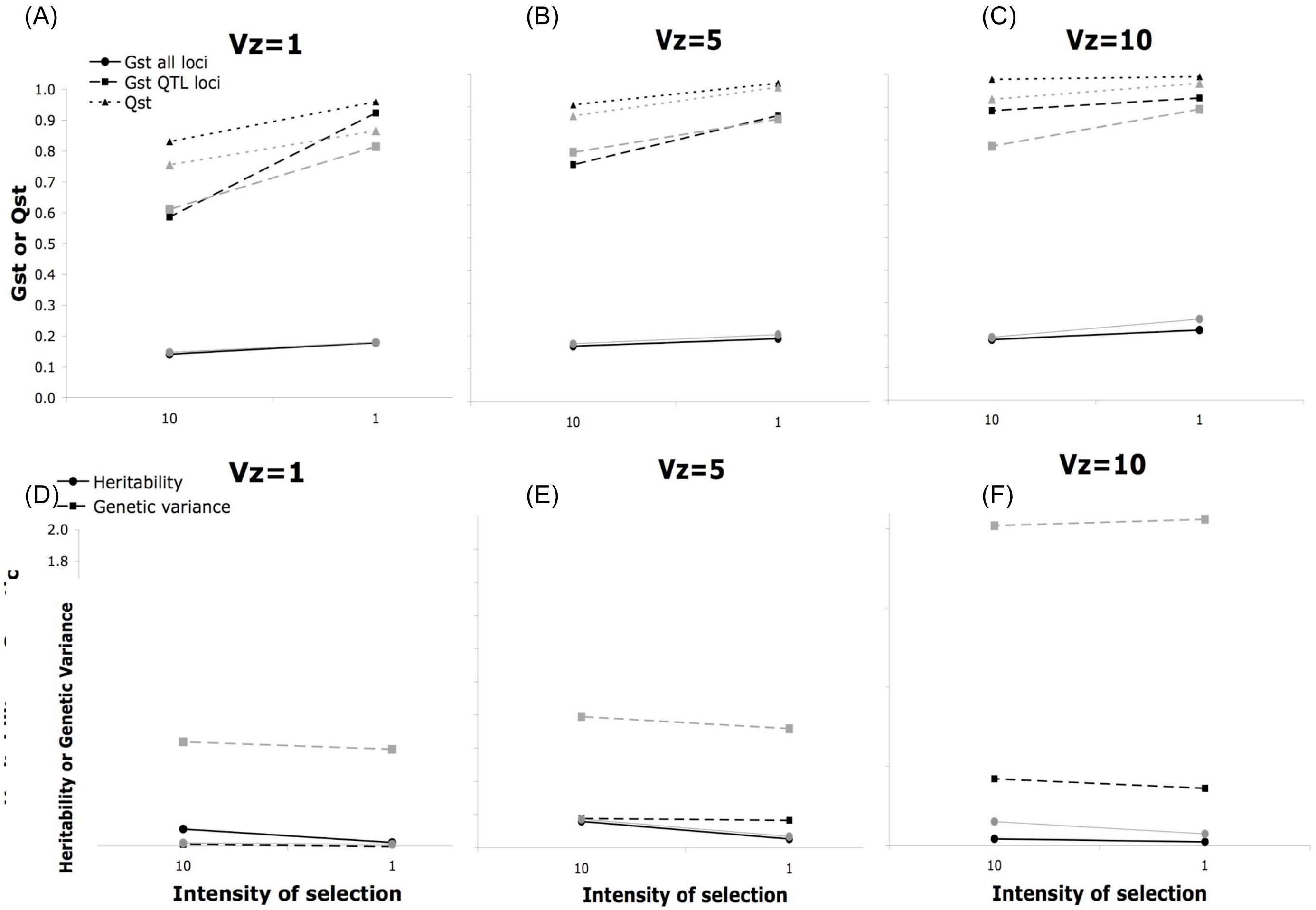


Figure 6.5 Heritability and genetic variance among 100 simulated populations following colonization from differentiated refugia. The variance in phenotypic optima among populations (V_z) and intensity of stabilizing selection within populations varied in each scenario. Black lines represent scenarios with an even distribution in phenotypic optima across the grid of populations. Gray lines represent scenarios with an uneven distribution of phenotypic optima.

6.4.2 The consequences of changing phenotypic optima on population structure and survival

The outputs from the colonization scenarios were then used as priors to assess the effect of changing the intensity of diversifying selection among populations. The immediate effect changing the phenotypic optima was assessed qualitatively for four populations under the two most extreme scenarios (V_{Zopt}=1 changed to V_{Zopt}=10, and V_{Zopt}=10 changed to V_{Zopt}=1). Each population was selected from the center of each quadrant of 25 populations on the 5 by 20 population grid, meaning all surrounding populations had equal phenotypic optima within that region. In particular, the four populations were chosen in order to determine if the difference in phenotypic optima across the grid (landscape) affected fitness differently. However, a similar response curve was found for all four populations and changing optima scenarios (Figure 6.6). Mean population fitness (the average of the mean fitness of each age class) decreased sharply in the first generations following the change in V_{Zopt}, but rebounded within 10 generations and fully recovered within 40 generations (Figure 6.6). The absolute level of population fitness was equivalent across the four scenarios examined, indicating the direction of change in V_{Zopt} may not affect the equilibrium fitness of populations. These patterns also indicate that when the variance in phenotypic optima among populations changes, adaptation, i.e. the achievement of a new phenotypic optima, occurs in a relatively short number of generations.

As with the scenarios of population isolation and colonization, the pattern of phenotypic optima (evenness), the level of change in V_{Zopt}, and the stabilizing selection within populations significantly affected measures of genetic and phenotypic differentiation and variance (Table 6.6). Unlike previous scenarios, the genetic variance varied with the direction and magnitude of the change in V_{Zopt}, indicating that genetic variance is affected when the phenotypic optima changes among populations. The greatest portion of the variance in the phenotypic measures of Q_{ST}, genetic variance and heritability was assigned to the difference in even versus uneven distribution of phenotypic optima, indicating that differences between the historic and current phenotypic optima may maintain phenotypic signatures through multiple adaptive events.

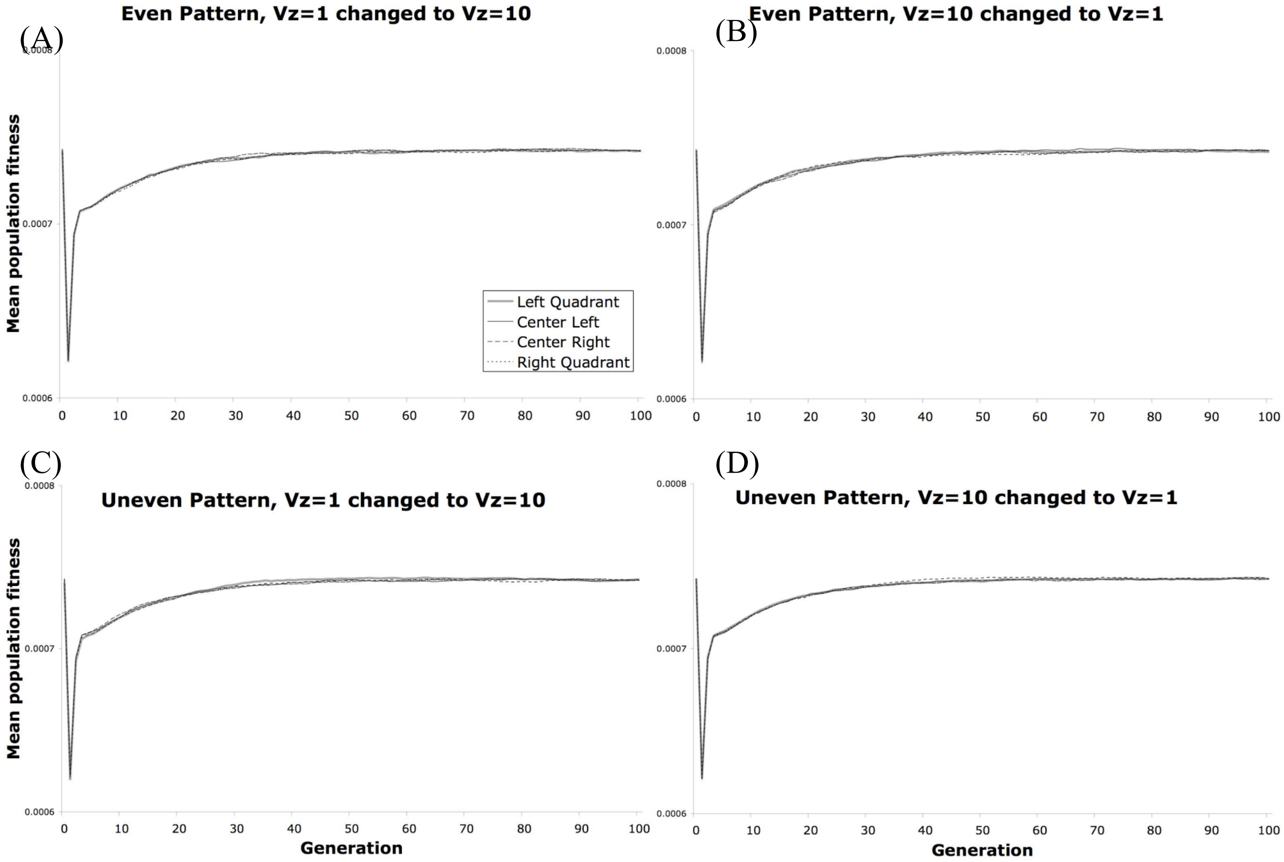


Figure 6.6 Mean population fitness recovered within 40 generations following a change in phenotypic optima. The centermost population from each quadrant across the grid (Figures 6.2 and 6.3) was assessed for the two most extreme scenarios under both the even pattern (A, B) and uneven pattern (C, D) of phenotypic optima.

Table 6.6. Changing the level of variance in phenotypic optima (V_{Zopt}) leads to significant variation in gene differentiation among all loci (G_{STall}) and among loci linked to QTL (G_{STqtl}), phenotypic differentiation (Q_{ST}), genetic variance (V_{G}), and heritablility (h^2) in the quantitative trait.

Source	**df**	**MS**	***F***	***P***
G_{STall}				
Even/uneven distribution	1	0.011	47.13	<0.001
Pattern of changing V_{Zopt}	5	0.022	96.82	<0.001
Intensity within populations	1	0.087	376.4	<0.001
Error	472	<0.001		
G_{STqtl}				
Even/uneven distribution	1	0.167	36.12	<0.001
Pattern of changing V_{Zopt}	5	0.114	24.71	<0.001
Intensity within populations	1	1.297	280.2	<0.001
Error	472	0.005		
Q_{ST}				
Even/uneven distribution	1	0.614	291.2	<0.001
Pattern of changing V_{Zopt}	5	0.053	25.24	<0.001
Intensity within populations	1	0.108	51.13	<0.001
Error	472	0.002		
V_{G}				
Even/uneven distribution	2	109.6	2963	<0.001
Pattern of changing V_{Zopt}	5	4.865	131.6	<0.001
Intensity within populations	1	0.742	20.06	<0.001
Error	472	0.037		
h^2				
Even/uneven distribution	1	0.473	278.5	<0.001
Pattern of changing V_{Zopt}	5	0.057	33.69	<0.001
Intensity within populations	1	0.103	60.48	<0.001
Error	472	0.002		

Rather than infer differences only from the absolute levels of genetic and phenotypic variation resulting from the change in V_{Zopt}, the relative change of each value was assessed by comparing the difference between final measures to those of the colonization scenario that served as the prior values. Only small changes were observed between initial and final values of G_{ST} over all loci (Figure 6.7). For loci linked to QTL and phenotypic differentiation, the level of within-population stabilizing selection affected the change in differentiation. For a decreasing V_{Zopt} (e.g. $V_{\mathrm{Zopt}}=10$ changed to $V_{\mathrm{Zopt}}=1$), scenarios of moderate stabilizing selection ($\omega=1$) resulted in smaller values G_{STqtl} and Q_{ST}, while scenarios of low stabilizing selection ($\omega=10$) resulted in differentiation closer to the initial values (Figure 6.7). When V_{Zopt} increased, however, measures of differentiation increased relative to the initial values for scenarios of low stabilizing selection, while scenarios of moderate stabilizing selection showed smaller changes. These effects were more pronounced in the more extreme scenarios than in those with a moderate prior variance in phenotypic optima ($V_{\mathrm{Zopt}}=5$).

The changes to measures of heritability were less pronounced than those for genetic and phenotypic differentiation (Figure 6.8). The genetic variance resulting from scenarios having an even distribution of phenotypic optima were generally lower and showed less change than those for the uneven pattern. Scenarios of changing V_{Zopt} having an uneven distribution of phenotypic optima displayed the greatest change in genetic variance. Decreases in V_{Zopt} resulted in lower values of genetic variance while increased V_{Zopt} resulted in higher genetic variance (Figure 6.8). These results indicate that changes in the intensity of diversifying selection may influence the potential of populations to adapt to further changes in phenotypic optima.

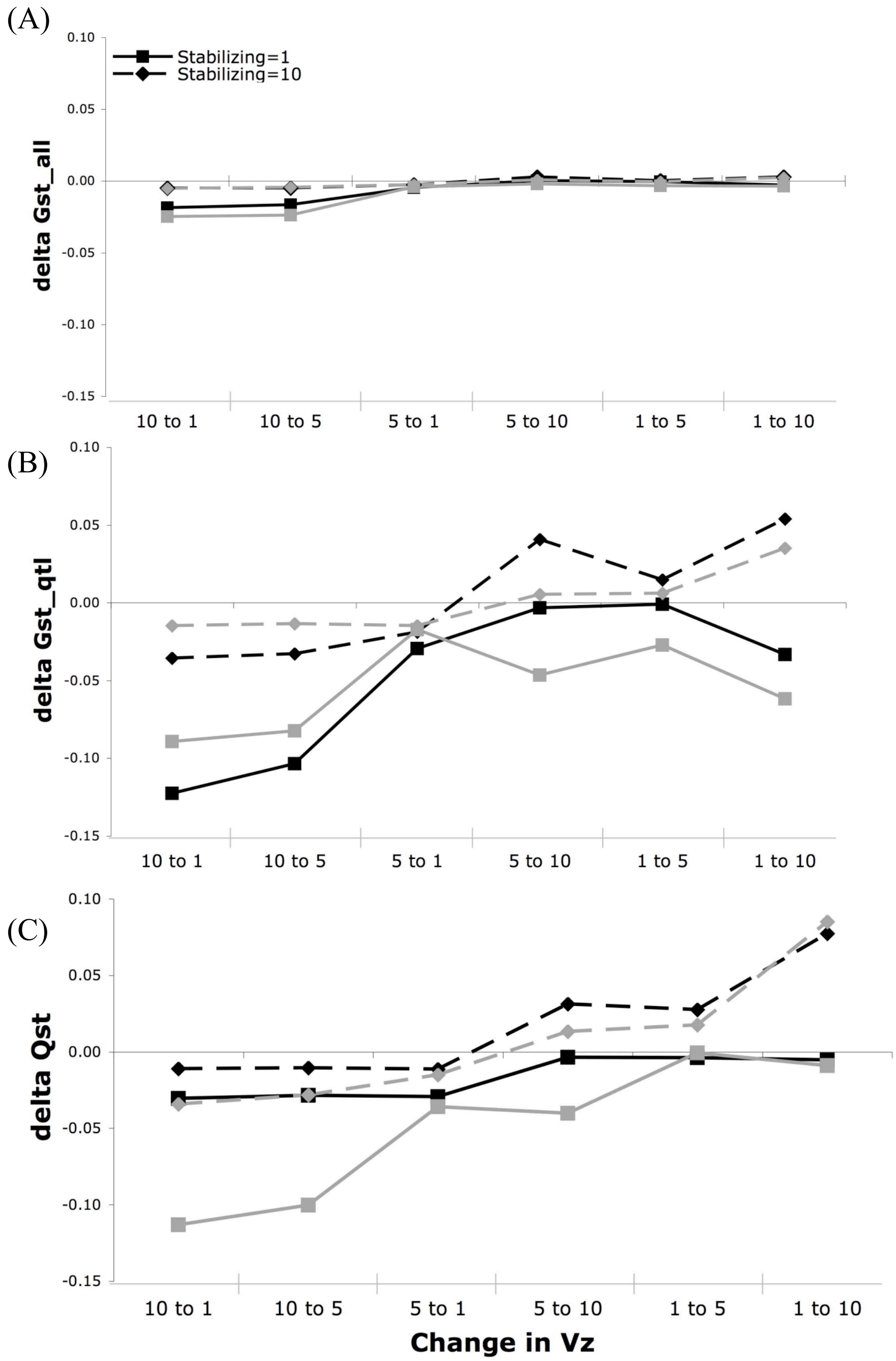


Figure 6.7 Effect of changing the variance in phenotypic optima (V_z) on the genetic and phenotypic differentiation among 100 simulated populations. (A) G_{ST} over all loci; (B) G_{ST} for loci linked to QTL; (C) Q_{ST}. The difference between initial and final values are plotted as delta values for each measure. Dashed lines represent scenarios with low within-population stabilizing selection (ω=10); solid lines represent moderate stabilizing selection (ω=1). Black symbols represent the pattern with an even distribution of phenotypic optima, and gray symbols represent the uneven pattern (Figures 6.2 and 6.3).

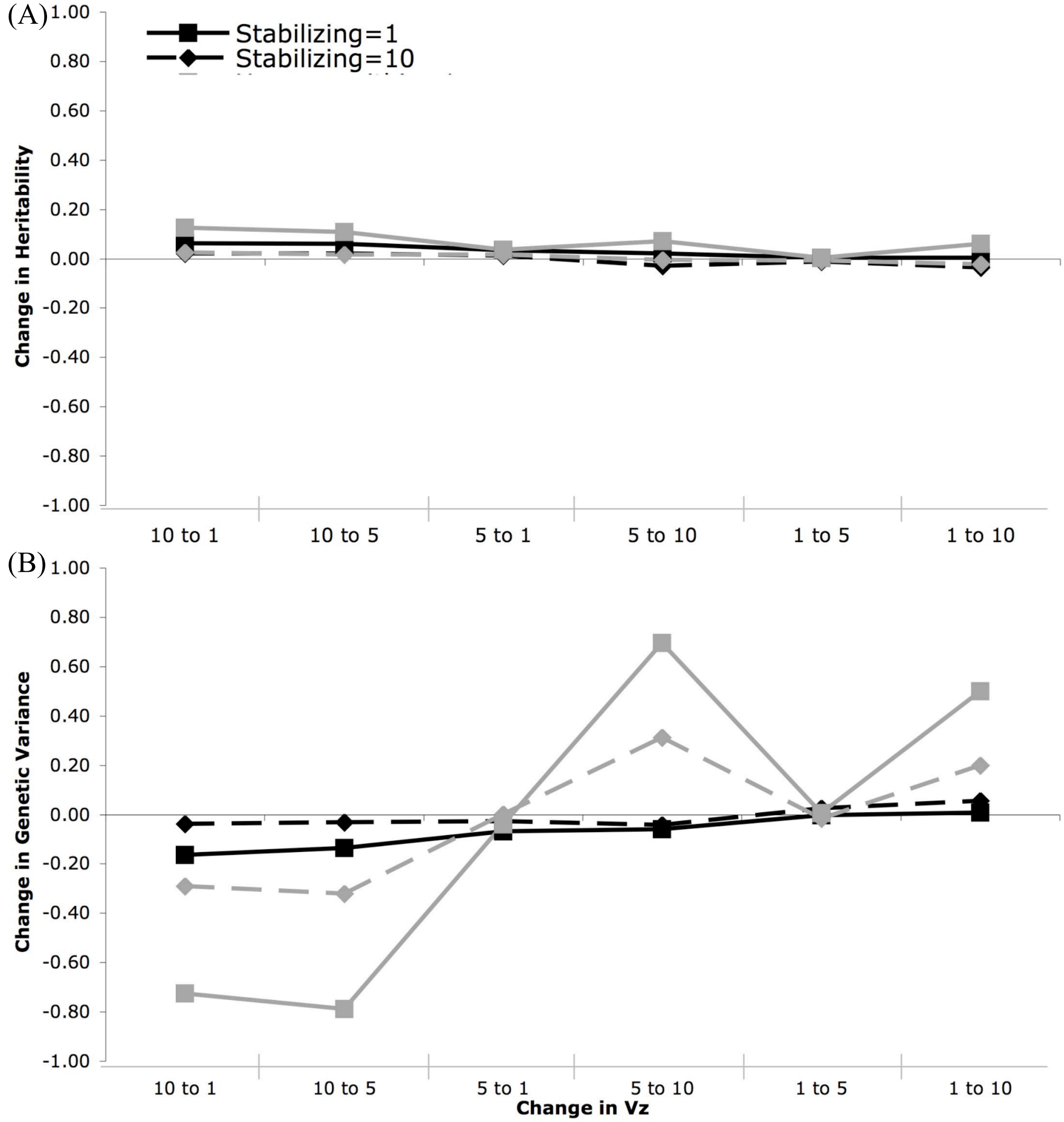


Figure 6.8 Effect of changing the variance in phenotypic optima (V_z) on (A) the heritability and (B) the genetic variance among 100 simulated populations. The difference between initial and final values are plotted as delta values for each measure. Dashed lines represent scenarios with low within-population stabilizing selection (ω=10); solid lines represent moderate stabilizing selection (ω=1). Black symbols represent the pattern with an even distribution of phenotypic optima, and gray symbols represent the uneven pattern.

6.5 Discussion

6.5.1 Genetic and phenotypic consequences of colonization from divergent refugia

Previous studies have simulated the genetic consequences of colonization without adaptation (LeCorre et al. 1997b; Austerlitz and Garnier-Gere 2003), of adaptation under different selection intensities that remained constant over time (LeCorre and Kremer 2003), and the speed of phenotypic adaptation in a population experiencing a moving optimum (Kopp and Hermisson 2007, 2009). This study investigated different patterns of selection across a hypothetic landscape (a grid of populations) and how changes to the intensity of selection affected the genetic and phenotypic characteristics of populations.

Population isolation coupled with divergent selection is one mechanism that may lead to speciation (Lande 1980). The significant genetic and phenotypic divergence observed between refugia following isolation (Table 6.5) is consistent with this theory of speciation, though these simulations assumed speciation did not take place, and that gene flow occurred among populations when propagules from the two refugia colonized the unoccupied space between. The differentiation observed in these simulations was greater than that described by LeCorre and Kremer (2003) using a previous version of METAPOP. These differences may be due to the particular parameterizations of the simulations. LeCorre and Kremer (2003) simulated 25 populations connected by the island model, whereas the scenarios here consisted of two isolated groups of 10 populations, a smaller effective population size, undergoing the same level of phenotypic differentiating selection.

Following colonization, the intensity of stabilizing selection within populations explained the greatest portion of variance in measures of genetic (G_{STall} and G_{STqtl}) and phenotypic (Q_{ST}) differentiation and heritability (Table 6.6). The importance of the intensity of stabilizing selection in defining a polygenic trait is well established (Lande 1980). For estimates of genetic (G_{STall} and G_{STqtl}) and phenotypic (Q_{ST}) differentiation, these results are consistent with a previous simulation of populations without colonization (LeCorre and Kremer 2003). However, the effects of within-population stabilizing selection on estimates of heritability were greater in the previous study (LeCorre and Kremer 2003). This difference may result from the differences in the pattern of phenotypic optima across the simulated landscape.

The distribution of phenotypic optima among populations significantly affected the genetic and phenotypic differentiation. That is, differentiation varied not just with the variance in the phenotypic optima (or the intensity of selection, V_{Zopt}) but also with the evenness of the specific phenotypic optima assigned to each population (Table 6.6). Differences between the phenotypic optima of the refugial source populations and the adaptive optima of the newly colonized populations significantly affect the resulting genetic and phenotypic structure of the species. The asymmetric distribution of optima modeled in this study is likely more realistic than the constant (e.g. LeCorre and Kremer 2003) or symmetric patterns (e.g. Kuparinen et al. 2010) typically applied in simulation studies. For instance, consider the European glacial refugia described in the Iberian, Italian and Balkan peninsulas (Hewitt 1999). If range expansion only occurred into sites with climate conditions matching the glacial refugium, then all genetic evidence of recolonization in Europe should follow a single pattern. As multiple colonization routes out of these refugia have been deduced from DNA marker and sequence data (Hewitt 1999; Petit et al. 2002; Palme et al. 2003), it is unlikely that each expanding population experienced equivalent environmental conditions. The asymmetric phenotypic selection, where the gene pool colonizing a site is less adapted than a propagule from a more distant refugium, results in a significant increase in the genetic variance and decrease in phenotypic and genetic differentiation relating to the adaptive trait (Figure 6.5). In particular, the greatest proportion of the variation in genetic variance (V_{G}) was attributed to the evenness of the pattern ($F_{1,235}$=820.5, P<0.001), indicating that heterogeneity in phenotypic optima across the landscape results in greater genetic heterogeneity within and among populations. As the level of additive genetic variance for an adaptive trait is one factor determining the rate of adaptation in a population (Lande and Shannon 1996), the differences in genetic variance here indicate populations may differ in their adaptive potential based on historic selective pressures.

While this examination of the genetic consequences of colonization from two divergent refugia provided novel insights, the model lacked rare long distance dispersal events shown to be a critical parameter in the postglacial recolonization of Europe by forest trees (Nichols and Hewitt 1994; LeCorre et al. 1997a; Austerlitz and Garnier-Gere 2003). Further, the pure diffusive model may result in a colonization rate slower than that observed in post-glacial evidence (Austerlitz and Garnier-Gere 2003). A mixed colonization model including local diffusion and rare long-distance parameters has been shown to increase the colonization rate and decrease the neutral genetic differentiation among resulting populations (LeCorre et al. 1997a). The magnitude and frequency of the long distance events influences the final genetic differentiation among populations due to the significant influence of founder effects when isolated populations are established at long distances from the wave front (Austerlitz and Garnier-Gere 2003). In these cases, even frequent local pollen dispersal cannot overcome the initial genetic differentiation established when the population was founded. Incorporation of a stochastic model with long distance dispersal may significantly influence the population structure of the asymmetric model in particular, providing a mechanism for propagules from the more distant refugia to establish and potentially out-complete colonists from the adjacent (maladapted) refugium.

6.5.2 Consequences of a change in intensity of phenotypic selection

These simulations provide insight into the potential adaptive responses of populations to a change in the variance of phenotypic optima among populations, which corresponds to directional selection within populations due to a change of phenotypic optima. This scenario differs from several studies in that it follows a discrete change in the most adaptive phenotypic value, rather than an optimum that changes at a constant rate over time (Lande and Shannon 1996; Kopp and Hermisson 2007, 2009; Kuparinen et al. 2010). While a constant rate of change may be more realistic to current climatic conditions than a single incremental step, the discrete model has been documented in natural populations and studied analytically (Boulding and Hay 2001), and allows examination of population responses during the change and after adaptation to the novel condition.

The mean fitness of each population quickly decreased in response to the change in phenotypic optima, but recovered within 40 generations (Figure 6.6). The rate of tree mortaility has been identified as a critical factor in determining the rate of adaptation to a constantly changing environment (Kuparinen et al. 2010), indicating that changes to the mortality parameters of these simulations may significantly alter this pattern. The translation of generation of a forest tree into generations of shorter-lived associated species may present significant conservation and management issues as forest populations decline or become more fragmented.

The changes in the variance in phenotypic optima may approximate the climatic changes predicted to occur in Europe over the next century. For instance, species with a southern range limit in the Mediterranean region may experience changes similar to an increase in variance among phenotypic optima as the climate becomes more extreme and less suitable for populations in this area. These simulations indicate that levels of differentiation reached following adaptation may differ from the initial values. Phenotypic differentiation may increase without a change in overall genetic structure (G_{STall}), while the genetic variance may increase markedly. Alternatively, populations at their northern range limit in central or northern Europe may experience changes approximating a decrease in variance among optimal phenotypes as conditions become more similar to the ideal (more southern) climate. Adaptation to the new climate may result in decreases in levels of genetic and phenotypic differentiation but slight increases in heritability, as indicated by these simulations.

As in the colonization scenarios, the magnitude of the changes depended on the level of stabilizing selection within populations. Thus, the intensity of the selection maintaining a homogeneous population significantly affected the resultant genetic and phenotypic structure of the populations. This trend was particularly apparent for changes in genetic variance for an uneven distribution of phenotypic optima. As previously discussed, the level of additive genetic variance in a population, together with the intensity of directional selection, determines the rate of adaptation of a quantitative trait (Lande and Shannon 1996). Consistent with this principle both the genetic variance and the level of phenotypic differentiation decreased with decreasing V_{Zopt}, indicating that populations became more phenotypically homogeneous.

6.5.3 Future scenarios to consider population responses to climate change

These results are based on models including only additive genetic effects for the trait of interest. However, epistatic and dominance effects will likely influence the response of populations to selective forces and the resulting phenotypic structure among populations. Investigations of the effect of dominance on Q_{ST}, the measure of phenotypic differentiation, indicate that the presence of recessive deleterious alleles at loci affecting a quantitative trait will tend to decrease Q_{ST} relative to F_{ST} in the absence of selection (Goudet 1995). In hypothesizing about the combined effects of dominance and directional selection, the authors suggest that the two processes may result in near equivalent levels of phenotypic and genetic differentiation, which may be misinterpreted as evidence of trait neutrality under the assumption of pure additive effects (Goudet 1995). Epistatic effects are also expected to decrease Q_{ST} relative to F_{ST} (Whitlock 1999; Lopez-Fanjul et al. 2003), though the interaction between epistasis and selection has not been investigated. In these investigations of additive selection, phenotypic differentiation was always much greater than the neutral genetic differentiation, consistent with the pure additive model applied. Revising the model parameters to include some level of epistatic or dominance interactions should be possible in the forthcoming version of the METAPOP software, and would provide insight into the interaction of these complex genetic processes.

Examination of allelic and genetic trends during the period of reduced population fitness (the period of adaptation) may provide additional insights. While several estimates of population structure are available as output (e.g. allele frequencies for each population at each generation), additional bioinformatics tools are required to fully analyze this information. For example, Kopp and Hermisson (2009) examined the rate of fixation of novel beneficial alleles in a population under stabilizing selection toward a moving optimum. Extracting per-generation changes in allele frequencies would allow examination of development of allelic clines and determination of which conditions produce the clines previously considered indicative of natural selection over a climatic gradient (e.g. Ingvarsson et al. 2006; Krutovsky et al. 2009).

Additional parameterizations of the model may prove informative. For instance, efforts to incorporate multiallelic loci (e.g. microsatellites) and cytoplasmic loci (e.g. chloroplast markers) may provide greater resolution of the effect of mutation and different dispersal patterns on the rate and pattern of adaptation. Adjusting demographic parameters to match empirical measures for a particular species (e.g. *Quercus petraea*) may provide insights directly applicable to management efforts. Attempts to parameterize the model to mimic the dioecious mating system of *Populus* species were not robust (data not presented), but additional work to simulate dioecious species (including animal models) will greatly improve the application and relevancy of this software.

Chapter 7. General Discussion

7.0 Insights from multiple scales: the landscape, the population, and the individual

The ability of natural populations to adapt or acclimate to climate change is likely to depend on the scale being considered (Svenning and Skov 2006). Thus, in order to build accurate management plans or predict population responses, multiple scales must be investigated. Here, the adaptive morphological variation in *Populus nigra* was considered at three scales. First, at the largest scale, the phenotypic and genetic structure of populations were considered across the landscape of western Europe (Chapter 2) and then into central Europe (morphological variation examined in Chapter 3). Simulation studies were also used to examine the genetic consequences of varying patterns of natural selection across a landscape, and the effects of changing the variance in phenotypic optima among those populations (Chapter 6).

Second, focusing on a single site (population), examination of insect preference on hybrid *Populus* revealed correlations between phenotype (leaf damage) and genotype, identifying QTL for insect associations. Further examination revealed two of these QTL to colocate to genetic ‘hot-spots’ for leaf development and morphology, linking developmental and biotic ecological factors in the genetic contribution to leaf morphology.

Third, focusing on individuals, significant variation in one adaptive trait, leaf size, was found to be driven by cell division and not cell growth, revealing a developmental difference between morphotypes (Chapter 4). The fundamental variation in cell number provides evidence to narrow future studies of candidate genes or physiological processes controlling adaptive morphology, and identifies a putative mechanism of natural selection driving phenotypic divergence.

7.1 Evolutionary history may affect future adaptive potential

At the largest scale, examination of patterns of genetic and morphological differentiation in *Populus nigra* across western Europe identified both isolation by distance and isolation by adaptation to affect genetic variation (Chapter 2). These patterns were consistent with differentiation reflecting not only current demographic and adaptive processes, but also historic differentiation between refugia and recolonization following the most recent glacial maximum (Cottrell et al. 2005). These patterns indicate that the restricted gene flow among populations of *P. nigra*, identified from estimates of F_{ST} (or analogous statistics) for microsatellite markers ranging from 0.08 (Smulders et al. 2008b) to 0.12 (Chapter 2), may be sufficient to minimize the homogenization of refugial genotypes and the migration of adaptively significant genes between historic gene pools. At a landscape scale, this may translate to different adaptive potentials among populations and may restrict population survival, particularly if migration is insufficient to allow populations to track the movement of optimal climates through time and across the changing climatic landscape (Ohlemuller et al. 2006; McLachlan et al. 2007).

The interaction between historic patterns of selection and future adaptation to changing conditions at the landscape scale was further investigated using simulation studies (Chapter 6). The effects of historic vicariance and differentiation were examined in two scenarios replicating the evolutionary trajectory hypothesized for *Populus nigra* based on chloroplast haplotypes (Cottrell et al. 2005) and current climatic similarities. Results indicate that dissimilar phenotypic optima between the source populations and newly established sites affected phenotypic differentiation and genetic variance among populations (Chapter 6). The adaptive potential of a population has been linked to the level of genetic variance related to adaptive traits, with greater variance correlated with a greater likelihood of adapting to changing conditions (Lande and Shannon 1996; Etterson 2004). Disparity between the phenotypic optima of migrants and the newly colonized site appears to increase the genetic variance for an adaptive trait, thus increasing the adaptive potential of the population.

In a theoretical exercise, extrapolating these results to the case of *P. nigra* allows predictions regarding the potential for populations to adapt to a changing climate. If populations in France were established by migrants originally more adapted to a different phenotypic optimum (Iberian climate) than that experienced currently (continental climate), these populations may contain greater genetic variance for the adaptive traits. This genetic variance may confer greater adaptive potential as the climate changes over the next century. In the broadest sense, the simulation studies are consistent with the evolutionary hypothesis that small-leaf tree populations in France maintain greater adaptive potential for the predicted change in climate.

Yet such extrapolation of the METAPOP model to *Populus* species requires caveats. *Populus* species vary from the canonical model of forest trees in having a dioecious mating system, whereby individual genets produce either male or female flowers. This mating system is predicted to decrease differentiation between populations (Moyle 2006), although *P. nigra* displays greater genetic differentiation than other forest trees (Chapter 2). Such significant departures from the demographic assumptions built in the METAPOP model will likely affect the simulation outcomes. Despite the genetic consequences of dioecy, however, the population genetic models built for forest species have not addressed this mating system.

Future simulation studies could begin by parameterizing the METAPOP scenarios for a dioecious model. Accurately simulating a dioecious mating system would broaden application of the METAPOP models not only to *Populus* species but also to animal systems. In addition, in order to simulate a changing environment more realistically, future scenarios should incorporate gradual and continual changes in phenotypic optima rather than a single discrete change. Both of these changes would allow more precise application of simulation studies to explore the evolutionary and adaptive processes affecting *Populus nigra.*

Additional genetic studies would help further resolve the relative roles of isolation by distance and isolation by adaptation in the morphologic and genetic structure of *Populus nigra*. The addition of chloroplast DNA marker data for the genets assessed in the common garden study (Chapter 2) would allow a direct test for the effects of historic isolation on current phenotypic differentiation. An alternative method to explore the effects of historic vicariance and ongoing admixture may be available upon the completion of an ongoing effort to resequence the *P. nigra* genome. This collaborative European effort will produce a complete genomic sequence for three divergent genets, and additional genotyping of 50 other genets from across western and central Europe. This genomic database may provide insight into the pattern of admixture across the *P. nigra* genome, as well as provide nucleotide data for whole genome scans (Storz 2005; Namroud et al. 2008).

7.2 Leaf traits as indicators of adaptation to the biotic environment

Leaf morphology varies broadly with environment and thus is considered an adaptive trait (Westoby and Wright 2006). The ecological factors thought to drive variation in leaf morphology are canonically considered abiotic variables such as water or light availability and temperature (Chapin et al. 1993; Meier and Leuschner 2008; Poorter and Rozendaal 2008). As the site of gas exchange and photosynthesis, the morphology of the leaf influences the water use efficiency of the plant on small and large scales. Within individual leaves, the stomatal patterning, leaf thickness and cuticle thickness influence the rate of water transpiration. Scaled up to include all the leaves on an individual plant, differences in stomatal conductance or water transpiration can have significant effects on the water relations and drought response of a tree (Jarvis 1976; Westoby and Wright 2006).

A growing body of literature indicates that interactions with the biotic environment may influence leaf morphology. Common garden studies of hybrid *Populus* species indicate that the biochemical properties of the leaves may be the mechanism influencing insect choice (Bangert et al. 2006), but examination of wild trees in a natural hybrid zone indicate that additional physiological factors, such as phenology, may be critical (Wimp et al. 2007). These findings indicate additional study of the interaction between the physical and developmental properties of leaves and insect preference is warranted.

Here, a study of quantitative genetic variation in insect preference on hybrid poplar revealed genetic markers correlated with insect interactions. Results indicated that both biochemical and developmental factors might be important (Chapter 5). The majority of QTL for insect association did not overlap, indicating the insect community responds to a large amount of genetic variation in poplar, some of which likely includes chemical and defense related genes. Additional bioinformatics, biochemical, and gene expression analyses will be necessary to identify the functional significance of the genetic variation within QTL. Yet two of the QTL colocated to genetic ‘hot-spots’ for leaf morphology. If the insects are responding to morphological variation apart from the biochemical properties, it is possible that small-scale variation in phenotype (individual leaf size or growth rate) can translate to ecological variation at the level of the community. Conversely, if the level of insect damage is sufficient to impact the fitness of the tree, then the insect community may affect the direction of selection on leaf traits in a population. Given the likelihood that natural selection is most effective in forest species at seedling establishment (Parry et al. 2004; Kuparinen et al. 2010), and that *Populus* trees have relatively fast regeneration times compared to canonical forest trees such as oak and pine, the effect of phytophagous insect damage concurrent with regeneration or seeding events may be significant.

While this study focused on a hybrid pedigree of *Populus*, a similar study could be conducted on an association population of *P. nigra* in order to assess the natural variation in insect abundance in this native species. Quantification of leaf damage in a common garden of *P. nigra* trees may reveal genetic and ecological processes contributing to the complex interaction between forest trees and the insect community. Correlations between leaf morphology or phenology and insect abundance may further resolve to what extent each physical property of the tree influences insect choice. Data from the common gardens could be compared to the natural ranges and abundance of different phytophagous species across the range of *P. nigra* in order to infer whether herbivory pressure may have influenced variation in leaf size on a regional or landscape scale.

7.3 Cell division as an adaptive developmental mechanism

The smallest scale of focus of this work compared leaf development among individuals to reveal that cell division and not cell growth drives variation in leaf size and shape in *Populus nigra* (Chapter 4). This finding identified a developmental mechanism (regulation of the cell cycle) controlling variation in a adaptive trait (leaf size). If natural selection drove differences in the rate or, more likely, duration of the cell division during leaf development among *P. nigra*, then variation in the sequence or regulation of genes related to the cell cycle should be considered for further study.

Variation in cell division may drive not just differences in leaf area but may underlie overall variation in growth rate and plant size observed in common garden (Chapter 2, Trewin 2008) and glasshouse studies (Chapter 3). Across plant species, leaf size is negatively correlated with the number of leaves produced (Yang et al. 2008), and often varies with the overall growth rate of the plant (Van Volkenburgh 1999). Cell cycle regulation may play a more important role in plant development than previously thought (Ramirez-Parra et al. 2005). While the mechanism controlling the duration of cell division in expanding leaves has not been identified, several genetic and physical factors are known to influence or initiate the cell cycle. Cell wall properties appear to control some aspect of cell division as the expression of expansin in a shoot apical meristem is sufficient to produce a leaf primordial (Pien et al. 2001). Osmotic pressure and carbohydrate availability also appear to affect the regulation of the cell cycle during leaf expansion (Van Volkenburgh 1999). Cell division is typically halted in response to stress such as water deficit (Zhu 2002). Given the developmental importance of the cell cycle, natural selection has likely driven the divergence of factors regulating the rate or duration of cell division in natural populations.

Differences in growth rate are also adaptive (Arendt 1997), so identifying a developmental mechanism linking cell division and overall growth in *P. nigra* may provide powerful tools for plant breeders and indicator traits for land managers. For example, two hormones that play dual roles in plant growth, auxin and cytokinins, control developmental processes as well as affect the regulation of the cell cycle (Ramirez-Parra et al. 2005). Thus, the adaptive mechanism of leaf size variation is as likely to be such genetic pathways as pluripotent growth hormones as are genes primarily involved in leaf initiation or development.

7.4 Conclusions

These studies provide novel insights into the complexities of the evolution of leaf morphology in an important floodplain tree, *Populus nigra*. Patterns of genetic and morphological differentiation indicated current phenotypic differences reflect historic vicariance events as well as local adaptation. Differences in leaf size among divergent morphotypes of *P. nigra* were due to differences in cell division and not cell size, a process important to multiple aspects of plant growth. Variation in insect preference among hybrid poplar identified quantitative trait loci (QTL) for different insect categories, with two loci collocating to QTL ‘hot-spots’ for leaf traits, indicating leaf morphology may influence insect choice. Simulation studies indicated asymmetric patterns of natural selection across a landscape might affect levels of phenotypic differentiation among and genetic variance within populations, which in turn affect a population's adaptive potential. Together, these results indicate that biotic and abiotic processes likely influenced the evolution of leaf size by affecting factors ranging from population-level variation to the control of cell division in this ecologically and economically important forest tree.

Appendix 1. Morphological variation in *P. nigra* previously assessed in a common garden experiment

	Population[‡]												
Trait*[†]	**1**	**2**	**3**	**5**	**6**	**7**	**8**	**9**	**11**	**12**	**13**	**14**	**15**
Leaf area (cm^2)	52.7 (2.7)	68.5 (3.9)	39.4 (12.7)	55.1 (9.0)	64.3 (7.7)	95.6 (6.0)	143.0 (5.4)	130.6 (7.3)	125.4 (9.2)	118.5 (10.3)	117.1 (7.3)	34.4 (3.6)	32.3 (3.0)
Leaf length (mm)	86.2 (2.4)	95.3 (2.6)	72.7 (13.8)	81.5 (6.0)	87.8 (5.2)	113.4 (3.6)	134.2 (2.3)	126.3 (4.6)	127.9 (3.1)	125.3 (4.8)	128.4 (5.5)	65.0 (3.4)	63.5 (2.7)
Leaf width (mm)	78.4 (2.4)	94.5 (2.7)	67.5 (12.5)	88.9 (9.4)	86.6 (5.6)	106.2 (3.2)	133.5 (2.6)	126.5 (4.2)	125.3 (5.3)	122.0 (3.0)	115.7 (5.7)	68.8 (3.7)	64.9 (3.0)
Length:width ratio	1.11 (0.02)	1.02 (0.02)	1.08 (0.01)	0.93 (0.04)	1.03 (0.03)	1.07 (0.03)	1.01 (0.01)	1.00 (0.02)	1.04 (0.03)	1.03 (0.05)	1.12 (0.05)	0.96 (0.03)	0.99 (0.03)
Specific leaf area (mm^2/mg)	1265.3 (113.2)	1399.1 (91.2)	442.0 (171.0)	1009.2 (203.8)	1472.9 (170.9)	2713.2 (167.4)	2823.4 (127.0)	2966.0 (133.7)	2861.7 (268.6)	2842.1 (110.0)	2885.9 (318.7)	1043.8 (240.2)	823.6 (159.4)
Mean cell area (mm^2)	355.4 (14.6)	296.6 (10.8)	298.3 (42.4)	464.4 (86.1)	360.0 (31.1)	337.7 (11.4)	312.0 (11.7)	323.9 (16.4)	359.3 (18.4)	295.4 (24.9)	278.9 (17.1)	330.3 (19.7)	360.8 (19.4)
Cell number per leaf (x10^6)	16.0 (1.2)	25.0 (1.8)	14.4 (5.6)	13.9 (4.0)	21.2 (3.7)	29.5 (2.3)	48.7 (2.7)	44.3 (3.4)	36.0 (2.9)	40.3 (2.4)	42.7 (3.4)	10.7 (1.3)	9.6 (1.2)

	Population‡												
Trait*†	**1**	**2**	**3**	**5**	**6**	**7**	**8**	**9**	**11**	**12**	**13**	**14**	**15**
Number stomata abaxial surface	24.5 (0.98)	26.4 (0.89)	27.0 (5.03)	23.6 (1.08)	26.81 (2.02)	27.2 (1.12)	29.4 (1.20)	28.7 (1.17)	25.1 (1.26)	32.6 (0.93)	28.83 (1.66)	25.5 (1.31)	24.5 (1.42)
Stomatal density abaxial surface	168.8 (11.3)	171.2 (8.5)	237.9 (7.0)	141.7 (16.8)	196.6 (16.2)	182.2 (11.8)	200.3 (11.3)	198.6 (14.4)	163.7 (12.2)	178.1 (17.9)	189.5 (13.9)	158.8 (14.9)	178.6 (12.7)
Stomatal index abaxial surface	8.1 (0.4)	8.6 (0.3)	8.7 (1.1)	7.6 (0.8)	8.4 (0.7)	9.4 (0.5)	8.9 (0.5)	8.4 (0.4)	8.8 (0.6)	8.5 (1.5)	8.8 (1.1)	8.6 (0.7)	9.9 (0.6)
Height (cm)	192.3 (6.0)	221.0 (6.5)	195.7 (11.7)	188.0 (15.1)	224.0 (14.8)	247.6 (6.0)	283.6 (6.2)	290.4 (8.4)	256.1 (10.7)	286.2 (15.4)	269.8 (4.6)	189.8 (10.2)	157.5 (7.4)
Diameter (mm)	9.94 (0.83)	13.26 (0.74)	9.33 (1.48)	8.40 (1.20)	13.75 (1.55)	16.32 (0.59)	19.50 (0.64)	19.95 (0.90)	17.00 (1.35)	21.40 (2.41)	20.75 (0.95)	9.39 (1.09)	6.98 (0.81)

*The mean measure for each population is provided with standard error in parenthesis.

†Data were collected by H. Trewin as part of her PhD thesis (Trewin, 2008).

‡Population designations follow Table 2.1.

Appendix 2. Microsatellite allele frequencies observed in 13 populations of *P. nigra*

Locus/						**Population**							
Allele	**1**	**2**	**3**	**5**	**6**	**7**	**8**	**9**	**11**	**12**	**13**	**14**	**15**
PMGC_14													
210									0.022				
219	0.017	0.023			0.042	0.188	0.014		0.087	0.100	0.056		0.054
222	0.100	0.057	0.083	0.083	0.125	0.263	0.108	0.190	0.022		0.056		
225	0.417	0.352	0.500	0.417	0.458	0.438	0.257	0.103	0.457	0.400	0.389	0.289	0.161
228	0.250	0.284	0.167	0.333	0.167	0.050	0.122	0.172	0.043	0.100		0.421	0.446
231	0.133	0.170	0.083	0.083	0.125	0.013	0.419	0.534	0.283		0.222	0.263	0.304
234	0.033	0.023	0.083			0.038	0.068						
237													0.018
240							0.014						
243	0.050	0.091	0.083	0.083	0.083				0.087	0.400	0.278	0.026	0.018
255						0.013							
PMGC_486													
153		0.011											0.019
159							0.039					0.059	0.037
161		0.011				0.013	0.158	0.200				0.029	0.093
163	0.483	0.216	0.167	0.167	0.231	0.163	0.105	0.100	0.065	0.300	0.222	0.265	0.204
165	0.017	0.045	0.056	0.083		0.038	0.066	0.133					
167	0.133	0.159	0.500	0.083	0.154	0.100	0.026	0.050				0.088	0.074
169	0.083	0.034				0.013	0.145	0.100	0.022		0.056	0.088	0.111
171		0.011		0.083		0.125	0.092	0.083	0.283	0.100	0.333		0.019
173	0.167	0.250	0.111		0.385	0.463	0.132	0.117	0.391	0.300	0.111	0.176	0.056
175	0.100	0.216	0.056	0.333	0.192	0.025	0.158	0.083	0.087	0.200		0.059	0.111
177	0.017	0.023			0.038		0.026	0.050	0.043			0.147	0.111

Locus/	Population												
Allele	**1**	**2**	**3**	**5**	**6**	**7**	**8**	**9**	**11**	**12**	**13**	**14**	**15**
179			0.111			0.063			0.043		0.056		
181		0.011		0.250				0.017				0.088	0.130
187		0.011					0.053	0.067	0.043	0.100	0.222		
189									0.022				
193													0.037
PMGC_2088													
175							0.012						
181	0.083	0.100			0.154	0.037	0.012	0.067	0.059	0.100			0.036
183		0.011											
185	0.267	0.156	0.250	0.083		0.049	0.012	0.033				0.400	0.357
187												0.067	0.089
189	0.067	0.122	0.500	0.167	0.154	0.122	0.468	0.417	0.137	0.100	0.082	0.300	0.107
191		0.022		0.250	0.115	0.061	0.023	0.033	0.020			0.167	0.196
193	0.200	0.211	0.250	0.167	0.115	0.305	0.023	0.050	0.274	0.200	0.244	0.067	0.089
195	0.017	0.044			0.077		0.152	0.083					
197	0.150	0.122			0.077	0.110				0.100	0.082		0.089
199	0.067	0.133			0.038	0.280	0.129	0.167	0.117	0.200			0.018
201	0.017	0.011		0.083	0.038		0.035	0.017	0.059				0.018
203		0.022		0.083	0.038					0.100			
205					0.038								
207	0.033	0.011											
209	0.100	0.022		0.167		0.037	0.023	0.133	0.234	0.200	0.326		
211		0.011			0.154		0.012						
213							0.012						
null							0.087		0.101		0.267		
PMGC_2163													
243		0.012											
245	0.067	0.098	0.333	0.083	0.077	0.250	0.256	0.234	0.136				
247						0.045	0.064	0.047	0.091	0.100	0.056		0.034
249	0.067	0.098	0.083	0.083	0.077	0.011		0.047		0.200		0.182	0.086
251	0.017	0.024			0.038	0.034	0.026						

Locus/						**Population**							
Allele	**1**	**2**	**3**	**5**	**6**	**7**	**8**	**9**	**11**	**12**	**13**	**14**	**15**
253							0.013						
255		0.012				0.239	0.090	0.016	0.159		0.111		
257	0.033	0.024											0.069
259	0.117	0.098		0.083			0.038	0.047					
261	0.017												
263	0.033	0.159		0.083	0.192	0.011		0.016			0.056	0.318	0.207
265	0.300	0.171	0.333	0.583	0.115	0.307	0.064	0.094	0.227	0.300	0.444	0.455	0.379
267	0.050	0.049	0.083		0.154	0.034	0.115	0.078	0.182	0.200	0.278		0.017
269	0.017	0.061			0.038		0.013	0.063					0.103
271		0.012	0.083					0.031		0.100			
273	0.017	0.049					0.051	0.063	0.091		0.056		0.017
275	0.250	0.110	0.083	0.083	0.231	0.034	0.167	0.109	0.091	0.100		0.045	0.086
277		0.012			0.038			0.016					
279		0.012					0.026		0.023				
281					0.038	0.034	0.077	0.109					
283	0.017												
285								0.031					
PMGC_2818													
148	0.323	0.202	0.500	0.167	0.167	0.523	0.200	0.172	0.250	0.300	0.222	0.237	0.362
150							0.013						
152	0.129	0.107		0.167		0.114	0.275	0.094	0.023				0.017
154	0.403	0.440	0.357	0.417	0.708	0.011	0.225	0.266	0.182	0.300	0.111	0.421	0.483
156		0.024		0.083	0.042	0.023	0.038	0.047					
158	0.065	0.119	0.071			0.170	0.163	0.266	0.318	0.100	0.333	0.316	0.103
160						0.011	0.013	0.016		0.100		0.026	
162	0.048	0.048	0.071	0.167	0.083	0.023	0.075	0.141	0.068		0.056		0.017
164		0.012				0.011			0.114	0.200	0.222		
166	0.032	0.048				0.102			0.045		0.056		
168						0.011							
174													0.017

Locus/	**Population**												
Allele	**1**	**2**	**3**	**5**	**6**	**7**	**8**	**9**	**11**	**12**	**13**	**14**	**15**
PMGC_2879													
155	0.194	0.114	0.063				0.012					0.028	0.052
157		0.011	0.063										0.017
161	0.016					0.083	0.037	0.032					
163	0.016	0.023				0.012		0.016	0.023			0.083	
165	0.016	0.011		0.500	0.136							0.167	0.172
167	0.161	0.114	0.250	0.167	0.045	0.167	0.012		0.136		0.167	0.056	0.052
169	0.016	0.068				0.036	0.024					0.083	0.052
171	0.032		0.063					0.048	0.023				
173		0.011					0.012						
175							0.024	0.016				0.056	0.034
177	0.048	0.091	0.125		0.045		0.012	0.097	0.045			0.056	0.069
179	0.032		0.063				0.024						0.034
181	0.097	0.239	0.125	0.083	0.091	0.071	0.110	0.145	0.023	0.200		0.222	0.207
183	0.065	0.091		0.083			0.073	0.097	0.091	0.200	0.278	0.028	0.069
185	0.016	0.023				0.107	0.098	0.113	0.250	0.100	0.056	0.028	0.034
187	0.113	0.045	0.125		0.273	0.012	0.244	0.145	0.205	0.400	0.333		0.017
189		0.011		0.083	0.045		0.037	0.048	0.023			0.028	0.121
191	0.048	0.011	0.063		0.182	0.238	0.085	0.016	0.068				0.017
193		0.045	0.063	0.083	0.091	0.048	0.061	0.048					
195	0.016	0.011			0.045	0.143	0.073	0.097	0.091		0.111	0.167	0.052
197						0.024	0.012		0.023				
199	0.032	0.011			0.045	0.024	0.012	0.016		0.100			
201		0.011					0.024	0.032					
209								0.016			0.056		
211						0.012							
213	0.081	0.057				0.024	0.012	0.016					
WPMS_14													
249	0.032	0.012										0.324	0.208
255	0.129	0.143		0.083	0.188	0.081	0.108	0.086	0.043			0.088	0.021
258	0.016	0.036		0.083		0.047	0.257	0.155	0.022			0.029	

Locus/	**Population**												
Allele	**1**	**2**	**3**	**5**	**6**	**7**	**8**	**9**	**11**	**12**	**13**	**14**	**15**
264	0.081	0.083	0.167	0.167	0.094	0.081	0.027	0.034	0.022			0.088	0.042
267	0.306	0.131	0.167		0.141	0.058		0.017	0.087	0.100			0.021
270			0.167		0.094	0.023	0.014		0.152		0.083		
273	0.177	0.202		0.167		0.384	0.135	0.259	0.261	0.200	0.417	0.118	0.104
276	0.048	0.071	0.167	0.083	0.188	0.093	0.311	0.207	0.022			0.029	0.333
279	0.016	0.024			0.047	0.174	0.014	0.017	0.217	0.400	0.250		0.042
282		0.012		0.083		0.012	0.014				0.167		
285					0.047			0.052	0.022		0.083	0.059	
291		0.024		0.083	0.047				0.043				
294									0.022				
297				0.083			0.014						
300							0.014						
303							0.014					0.029	
306	0.081	0.214	0.167	0.167		0.023	0.068	0.103	0.065	0.200		0.176	0.104
309	0.016	0.012				0.012	0.014	0.017	0.022	0.100		0.029	0.125
312		0.012				0.012		0.052					
315	0.097	0.024	0.167									0.029	
null					0.154								
WPMS_18													
245								0.016					
248	0.121	0.075	0.111	0.083	0.231	0.012	0.125	0.129	0.045	0.100			
251	0.534	0.350	0.333	0.500	0.115	0.256	0.263	0.355	0.409	0.300	0.667	0.375	0.444
254		0.025											
257	0.172	0.250	0.167		0.231	0.593	0.313	0.210	0.295	0.300	0.222	0.175	0.074
260				0.083									0.037
263	0.017	0.050	0.167				0.150	0.177					0.019
266	0.069	0.238	0.111	0.167	0.346	0.047	0.063	0.065	0.159	0.200		0.225	0.278
269	0.017		0.111	0.167	0.077	0.035	0.075	0.032	0.091	0.100	0.111	0.225	0.148
272	0.017												
275	0.052					0.058	0.013	0.016					
278		0.013											

Locus/						**Population**							
Allele	**1**	**2**	**3**	**5**	**6**	**7**	**8**	**9**	**11**	**12**	**13**	**14**	**15**
WPMS_20													
246	0.317	0.313		0.250	0.231	0.122	0.1698	0.167	0.227	0.200	0.167		0.259
252	0.250	0.263	0.500	0.333	0.154	0.195	0.249	0.400	0.068	0.100		0.500	0.056
258	0.233	0.338	0.500	0.333	0.500	0.280	0.1471	0.250	0.409	0.500	0.556	0.500	0.667
264	0.200	0.088		0.083	0.115	0.146	0.3169	0.183	0.045				0.019
270						0.256	0.0226		0.045	0.100	0.111		
276									0.205	0.100	0.167		
null							0.0947						

Appendix 3. Previously described QTL used to identify genomic 'hot-spots' for leaf morphology

QTL	**LG**	**Position***	**Lower**[†]	**Upper**[‡]	**Citation**
absolute_expansion_aCO2	I	0	0	18	Rae et al., 2006
absolute_expansion_aCO2	XIV	0	0	20	Rae et al., 2006
absolute_expansion_droughtcontrol	X	12	12	58	Street et al., 2006
absolute_expansion_droughtcontrol	IX	0	0	4	Street et al., 2006
dry_weight_leaf_Italy	VIIIa	1	1	26	Rodriguez-Acosta, 2006
dry_weight_leaf_Italy	XIV	0	0	15	Rodriguez-Acosta, 2006
dry_weight_leaf_UK	I	9	9	24	Rodriguez-Acosta, 2006
dry_weight_leaf_UK	IV	7	7	31	Rodriguez-Acosta, 2006
dry_weight_leaf_UK	IX	0	0	10	Rodriguez-Acosta, 2006
dry_weight_leaf_UK	Va	65	65	82	Rodriguez-Acosta, 2006
dry_weight_leaf_UK	VIIIa	0	0	25	Rodriguez-Acosta, 2006
dry_weight_leaf_UK	XVII	20	20	78	Rodriguez-Acosta, 2006
leaf_area_leaf_Italy	IX	0	0	9	Rodriguez-Acosta, 2006
leaf_area_leaf_Italy	XII	0	0	9	Rodriguez-Acosta, 2006
leaf_area_leaf_UK	IV	7	7	34	Rodriguez-Acosta, 2006
leaf_area_leaf_UK	IX	0	0	9	Rodriguez-Acosta, 2006
leaf_area_leaf_UK	Va	67	67	82	Rodriguez-Acosta, 2006
leaf_area_leaf_UK	VIIIa	0	0	25	Rodriguez-Acosta, 2006
leaf_area_leaf_UK	XI	31	31	55	Rodriguez-Acosta, 2006
leaf_area_leaf_UK	XVII	23	23	73	Rodriguez-Acosta, 2006
leaf_area_mature_aCO2	I	141	141	9	Rae et al., 2006
leaf_area_mature_aCO2	IV	48	48	66	Rae et al., 2006
leaf_area_mature_aCO2	IX	0	0	9	Rae et al., 2006
leaf_area_mature_aCO2	VI	97	97	125	Rae et al., 2006
leaf_area_mature_aCO2	X	30	30	62	Rae et al., 2006
leaf_area_mature_aCO2	XI	51	51	55	Rae et al., 2006
leaf_area_mature_aCO2	XII	0	0	24	Rae et al., 2006
leaf_area_mature_aCO2	XIX	16	16	32	Rae et al., 2006

QTL	LG	Position*	Lower[†]	Upper[‡]	Citation
leaf_area_mature_aCO2	XV	12	12	35	Rae et al., 2006
leaf_area_mature_aCO2	XVII	49	49	78	Rae et al., 2006
leaf_area_SRC1	IV	77	77	84	Rae et al., 2004
leaf_area_SRC1	VI	60	60	78	Rae et al., 2004
leaf_area_SRC1	XIX	6	6	15	Rae et al., 2004
leaf_area_SRC1	XVII	0	0	34	Rae et al., 2004
leaf_area_SRC2	IX	0	0	8	Rae et al., 2004
leaf_area_SRC2	VI	59	59	78	Rae et al., 2004
leaf_area_SRC2	XIX	0	0	17	Rae et al., 2004
leaf_area_SRC2	XVII	64	64	78	Rae et al., 2004
leaf_area_young_aCO2	I	0	0	22	Rae et al., 2006
leaf_area_young_aCO2	VI	66	66	90	Rae et al., 2006
leaf_area_young_aCO2	VII	26	26	62	Rae et al., 2006
leaf_area_young_aCO2	XI	20	20	55	Rae et al., 2006
leaf_area_young_aCO2	XII	0	0	24	Rae et al., 2006
leaf_area_young_aCO2	XIV	7	7	27	Rae et al., 2006
leaf_area_young_aCO2	XVII	7	7	61	Rae et al., 2006
leaf_extension_aCO2	I	0	0	22	Rae et al., 2006
leaf_extension_aCO2	III	84	84	95	Rae et al., 2006
leaf_extension_aCO2	XI	45	45	54	Rae et al., 2006
leaf_extension_aCO2	XIV	0	0	14	Rae et al., 2006
leaf_extension_droughtcontrol	III	0	0	65	Street et al., 2006
leaf_extension_SRC	II	28	28	59	Rae et al., 2004
leaf_extension_SRC	VI	56	56	69	Rae et al., 2004
leaf_extension_SRC	VII	24	24	51	Rae et al., 2004
leaf_extension_SRC	XVIII	54	54	84	Rae et al., 2004
leaf_length:width_droughtcontrol	III	26	26	68	Street et al., 2006
leaf_length:width_droughtcontrol	Va	24	24	63	Street et al., 2006
leaf_length:width_leaf_Italy	III	66	66	95	Rodriguez-Acosta, 2006
leaf_length:width_leaf_Italy	IX	0	0	3	Rodriguez-Acosta, 2006
leaf_length:width_leaf_Italy	XIV	0	0	7	Rodriguez-Acosta, 2006
leaf_length:width_leaf_Italy	XVIII	22	22	47	Rodriguez-Acosta, 2006
leaf_length:width_leaf_UK	Va	28	28	64	Rodriguez-Acosta, 2006
leaf_length:width_leaf_UK	VI	73	73	88	Rodriguez-Acosta, 2006

QTL	LG	Position*	Lower[†]	Upper[‡]	Citation
leaf_length:width_leaf_UK	XI	31	31	54	Rodriguez-Acosta, 2006
leaf_length:width_leaf_UK	XIV	0	0	11	Rodriguez-Acosta, 2006
leaf_length:width_leaf_UK	XVIII	30	30	49	Rodriguez-Acosta, 2006
leaf_length_leaf_Italy	VIIIa	1	1	25	Rodriguez-Acosta, 2006
leaf_length_leaf_Italy	XII	2	2	24	Rodriguez-Acosta, 2006
leaf_length_leaf_UK	I	140	140	165	Rodriguez-Acosta, 2006
leaf_length_leaf_UK	IV	6	6	31	Rodriguez-Acosta, 2006
leaf_length_leaf_UK	VIIIa	0	0	17	Rodriguez-Acosta, 2006
leaf_length_leaf_UK	XII	3	3	24	Rodriguez-Acosta, 2006
leaf_length_leaf_UK	XV	32	32	61	Rodriguez-Acosta, 2006
leaf_length_leaf_UK	XVII	22	22	65	Rodriguez-Acosta, 2006
leaf_width:length_aCO2	I	149	149	174	Rae et al., 2006
leaf_width:length_aCO2	II	23	23	49	Rae et al., 2006
leaf_width:length_aCO2	IV	72	72	84	Rae et al., 2006
leaf_width:length_aCO2	Va	54	54	77	Rae et al., 2006
leaf_width:length_aCO2	VIIIa	0	0	25	Rae et al., 2006
leaf_width:length_aCO2	XII	10	10	24	Rae et al., 2006
leaf_width_droughtcontrol16d	II	19	19	60	Street et al., 2006
leaf_width_droughtcontrol16d	IX	0	0	2	Street et al., 2006
leaf_width_droughtcontrol9d	XVIII	0	0	30	Street et al., 2006
leaf_width_leaf_Italy	IX	0	0	9	Rodriguez-Acosta, 2006
leaf_width_leaf_UK	IV	7	7	36	Rodriguez-Acosta, 2006
leaf_width_leaf_UK	Va	58	58	82	Rodriguez-Acosta, 2006
leaf_width_leaf_UK	VI	115	115	144	Rodriguez-Acosta, 2006
leaf_width_leaf_UK	VIIIa	0	0	21	Rodriguez-Acosta, 2006
leaf_width_leaf_UK	XI	27	27	55	Rodriguez-Acosta, 2006
leaf_width_leaf_UK	XIV	0	0	12	Rodriguez-Acosta, 2006
leaf_width_leaf_UK	XVII	19	19	70	Rodriguez-Acosta, 2006
SLA_aCO2	VIIIa	0	0	27	Rae et al., 2006
SLA_aCO2	X	33	33	56	Rae et al., 2006
SLA_leaf_Italy	Va	19	19	73	Rodriguez-Acosta, 2006
SLA_leaf_Italy	X	0	0	25	Rodriguez-Acosta, 2006
SLA_leaf_UK	IV	0	0	20	Rodriguez-Acosta, 2006
SLA_leaf_UK	VI	11	11	49	Rodriguez-Acosta, 2006

QTL	**LG**	**Position***	**Lower**†	**Upper**‡	**Citation**
SLA_leaf_UK	VII	0	0	13	Rodriguez-Acosta, 2006
SLA_leaf_UK	VIIIa	0	0	12	Rodriguez-Acosta, 2006
SLA_leaf_UK	XII	0	0	3	Rodriguez-Acosta, 2006
SLA_SRC	II	98	98	133	Rae et al., 2004
SLA_SRC	III	7	7	34	Rae et al., 2004
SLA_SRC	VIIIa	5	5	28	Rae et al., 2004
SLA_SRC	VIIIb	18	18	25	Rae et al., 2004
SLA_SRC	X	37	37	62	Rae et al., 2004
SLA_SRC	XIV	0	0	12	Rae et al., 2004

*Position of QTL in cM

†Lower position of 95% confidence interval in cM

‡Upper position of 95% confidence interval in cM

List of References

Abbruzzese, G., I. Beritognolo, R. Muleo, M. Piazzai, M. Sabatti, G. Scarascia Mugnozza, and E. Kuzminsky. 2009. Leaf morphological plasticity and stomatal conductance in three *Populus alba* L. genotypes subjected to salt stress. Environmental and Experimental Botany 66:381-388.

Abrams, M. D. 1994. Genotypic and phenotypic variation as stress adaptations in temperate tree species: a review of several case studies. Tree Physiology 14:833-842.

Ackerly, D. D., and P. B. Reich. 1999. Convergence and correlations among leaf size and function in seed plants: a comparative test using independent contrasts. American Journal of Botany 86:1272-1281.

Al Afas, N., N. Marron, and R. Ceulemans. 2007. Variability in *Populus* leaf anatomy and morphology in relation to canopy position, biomass production, and varietal taxon. Annals of Forest Science 64:521-532.

Anonymous. 2005. GeneChip(R) Poplar Genome Array Data Sheet *in* I. Affymetrix, ed. Downloaded from company webpage 17 September 2010, High Wycombe, United Kingdom.

Antonovics, J. 1968. Evolution in closely adjacent plant populations VI. Manifold effects of gene flow. Heredity 23:508-524.

Arendt, J. D. 1997. Adaptive intrinsic growth rates: An integration across taxa. The Quarterly Review of Biology 72:149-177.

Arens, P., H. Coops, J. Jansen, and B. Vosman. 1998. Molecular genetic analysis of black poplar (*Populus nigra* L.) along Dutch rivers. Molecular Ecology 7:11-18.

Austerlitz, F., and P. Garnier-Gere. 2003. Modelling the impact of colonisation on genetic diversity and differentiation of forest trees: interaction of life cycle, pollen flow and seed long-distance dispersal. Heredity 90:282-290.

Austerlitz, F., S. Mariette, N. Machon, P.-H. Gouyon, and B. Godelle. 2000. Effects of colonization processes on genetic diversity: differences between annual plants and tree species. Genetics 154:1309-1321.

Baek, D., P. Pathange, J.-S. Chung, J. Jiang, L. Gao, A. Oikawa, M. Y. Hirai, K. Saito, P. W. Pare, and H. Shi. 2010. A stress-inducible sulphotransferase sulphonates salicylic acid and confers pathogen resistance in *Arabidopsis*. Plant, Cell & Environment 33:1383-1392.

Bailey, I. W., and E. W. Sinnott. 1916. The climatic distribution of certain types of angiosperm leaves. American Journal of Botany 3:24-39.

Bailey, J. K., S. Wooley, R. L. Lindroth, and T. G. Whitham. 2006. Importance of species interactions to community heritability: a genetic basis to trophic-level interactions. Ecology Letters 9:78-85.

Bangert, R. K., R. J. Turek, B. Rehill, G. M. Wimp, J. A. Schweitzer, G. J. Allan, J. K. Bailey, G. D. Martinsen, P. Keim, R. L. Lindroth, and T. G. Whitham. 2006. A genetic similarity rule determines arthropod community structure. Molecular Ecology 15:1379-1391.

Barboni, D., S. P. Harrison, P. J. Bartlein, G. Jalut, M. New, I. C. Prentice, M.-F. Sanchez-Goni, A. Spessa, B. Davis, and A. C. Stevenson. 2004. Relationships between plant traits and climate in the Mediterranean region: A pollen data analysis. Journal of Vegetation Science 15:635-646.

Benetka, V., J. Mottl, K. Vackova, M. Pospiskova, and M. Dubsky. 1999. Estimation of the introgression level in *Populus nigra* L. populations by means of isozyme gene markers. Silvae Genetica 48:218-223.

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B 57:289-300.

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

Bonnin, I., J.-M. Prosperi, and I. Olivieri. 1996. Genetic markers and quantitative genetic variation in *Medicago truncatula* (Leguminosae): A comparative analysis of population structure. Genetics 143:1795-1805.

Boulding, E. G., and T. Hay. 2001. Genetic and demographic parameters determining population persistence after a discrete change in the environment. Heredity 86:313-324.

Braatne, J. H., S. J. Brunsfeld, V. D. Hipkins, and B. L. Wilson. 2006. Naturalization of plains cottonwood (*Populus deltoides* subsp *monilifera*) along river drainages west of the Rocky Mountains. Western North American Naturalist 66:310-320.

Bradshaw, H. D. 1996. Molecular genetics of *Populus*. Pp. 183-199 *in* R. F. Stettler, H. D. Bradshaw, P. E. Heilman, and T. M. Hinckley, eds. Biology of *Populus* and its Implications for Management and Conservation. NRC Research Press, Ottawa.

Bradshaw, H. D., Jr., R. Ceulemans, J. Davis, and R. Stettler. 2000. Emerging model systems in plant biology: Poplar (*Populus*) as a model forest tree. Journal of Plant Growth Regulation 19:306-313.

Bradshaw, H. D., Jr., and R. F. Stettler. 1995. Molecular genetics of growth and development in *Populus*. IV. Mapping QTLs with large effects on growth, and phenology traits in a forest tree. Genetics 139:963-973.

Bradshaw, H. D., Jr., M. Villar, B. D. Watson, K. G. Otto, S. Stewart, and R. F. Stettler. 1994. Molecular genetics of growth and development in *Populus*. III. A genetic linkage map of a hybrid poplar composed of RFLP, STS, and RAPD markers. Theoretical and Applied Genetics 89:167-178.

Bright, K. L., and M. D. Rausher. 2008. Natural selection on a leaf-shape polymorphism in the ivyleaf morning glory (*Ipomoea hederacea*). Evolution 62:1978-1990.

Broderick, N. A., E. Vasquez, J. Handelsman, and K. F. Raffa. 2010. Effect of clonal variation among hybrid poplars on susceptibility of gypsy moth (Lepidoptera: Lymantriidae) to *Bacillus thuringiensis* subsp. *kurstaki*. Journal of Economic Entomology 103:718-725.

Brouquisse, R., J.-P. Gaudillere, and P. Raymond. 1998. Induction of a Carbon-Starvation-Related Proteolysis in Whole Maize Plants Submitted to Light/Dark Cycles and to Extended Darkness. Plant Physiol. 117:1281-1291.

Burg, K., A. Helmersson, P. Bozhkov, and S. von Arnold. 2007. Developmental and genetic variation in nuclear microsatellite stability during somatic embryogenesis in pine. Journal of Experimental Botany 58:687-698.

Busov, V., A. Brunner, and S. Strauss. 2008. Genes for control of plant stature and form. New Phytologist 177:589-607.

Byrne, M. E., R. Barley, M. Curtis, J. M. Arroyo, M. Dunham, A. Hudson, and R. A. Martienssen. 2000. Asymmetric leaves1 mediates leaf patterning and stem cell function in *Arabidopsis*. Nature 408:967-971.

Cervera, M. T., V. Storme, B. Ivens, J. Gusmao, B. H. Liu, V. Hostyn, J. Van Slycken, M. Van Montagu, and W. Boerjan. 2001. Dense genetic linkage maps of three *Populus* species (*Populus deltoides, P. nigra* and *P. trichocarpa*) based on AFLP and microsatellite markers. Genetics 158:787-809.

Cervera, M. T., V. Storme, A. Soto, B. Ivens, M. Van Montagu, O. P. Rajora, and W. Boerjan. 2005. Intraspecific and interspecific genetic and phylogenetic relationships in the genus *Populus* based on AFLP markers. Theoretical and Applied Genetics 111:1440-1456.

Ceulemans, R., G. Scarascia-Mugnozza, B. M. Wiard, J. H. Braatne, T. M. Hinckley, R. F. Stettler, J. G. Isebrands, and P. E. Heilman. 1992. Production physiology and morphology of *Populus* species and their hybrids grown under short rotation. I. Clonal comparisons of 4-year growth and phenology. Canadian Journal of Forest Research 22:1937-1948.

Chang, S., J. Puryear, and J. Cairney. 1993. A simple and efficient method for isolating RNA from pine trees. Plant Molecular Biology Reporter 11:113-113.

Chapin, F. S., K. Autumn, and F. Pugnaire. 1993. Evolution of suites of traits in response to environmental stress. American Naturalist 142:S78-S92.

Cheliak, W. M., and B. Dancik. 1982. Genetic diversity of natural populations of a clone-forming tree *Populus tremuloides*. Canadian Journal of Genetics and Cytology 24:611-616.

Chen, K., Y. H. Peng, Y. H. Wang, H. Korpelainen, and C. Y. Li. 2007. Genetic relationships among poplar species in section Tacamahaca (*Populus* L.) from western Sichuan, China. Plant Science 172:196-203.

Chen, T. H. H., G. T. Howe, and H. D. Bradshaw. 2002. Molecular genetic analysis of dormancy-related traits in poplars. Weed Science 50:232-240.

Cho, H. T., and D. J. Cosgrove. 2000. Altered expression of expansin modulates leaf growth and pedicel abscission in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences of the United States of America 97:9783-9788.

Cnops, G., S. Jover-Gil, J. L. Peters, P. Neyt, S. De Block, P. Robles, M. R. Ponce, T. Gerats, J. L. Micol, and M. Van Lijsebettens. 2004. The *rotunda2* mutants identify a role for the *LEUNIG* gene in vegetative leaf morphogenesis. Journal of Experimental Botany 55:1529-1539.

Cole, C. T. 2005. Allelic and population variation of microsatellite loci in aspen (*Populus tremuloides*). New Phytologist 167:155-164.

Colmer, T. D., and L. A. C. J. Voesenek. 2009. Flooding tolerance: suites of plant traits in variable environments. Functional Plant Biology 36:665-681.

Cookson, S. J., A. Radziejwoski, and C. Granier. 2006. Cell and leaf size plasticity in *Arabidopsis*: what is the role of endoreduplication? Plant Cell and Environment 29:1273-1283.

Cottrell, J. E., V. Krystufek, H. E. Tabbener, A. D. Milner, T. Connolly, L. Sing, S. Fluch, K. Burg, F. LeFevre, R. Achard, S. Bordacs, K. Gebhardt, B. Vornam, M. J. M. Smulders, A. H. Vanden Broeck, J. Van Slycken, V. Storme, W. Boerjan, S. Castiglione, T. Fossati, N. Alba, D. Agundez, C. Maestro, E. Notivol, J. Bovenschen, and B. C. Van Dam. 2005. Postglacial migration of *Populus nigra* L.: lessons learnt from chloroplast DNA. Forest Ecology and Management 206:71-90.

Coyle, D. R., M. D. Coleman, J. A. Durant, and L. A. Newman. 2006. Multiple factors affect pest and pathogen damage on 31 *Populus* clones in South Carolina. Biomass and Bioenergy 30:759-768.

Dennison, M. D., and A. J. Baker. 1991. Morphometric Variability in Continental and Atlantic Island Populations of Chaffinches (Fringilla coelebs). Evolution 45:29-39.

DeWoody, J., J. D. Nason, and V. D. Hipkins. 2006. Mitigating scoring errors in microsatellite data from wild populations. Molecular Ecology Notes 6:951-957.

Diemer, M. 1998. Leaf lifespans of high-elevation, aseasonal Andean shrub species in relation to leaf traits and leaf habit. Global Ecology and Biogeography Letters 7:457-465.

Donaldson, J. R., and R. L. Lindroth. 2008. Effects of variable phytochemistry and budbreak phenology on defoliation of aspen during a forest tent caterpillar outbreak. Agricultural and Forest Entomology 10:339-410.

Doncaster, C. P., and A. J. H. Davey. 2007. Analysis of Variance and Covariance: How to choose and construct models for the life sciences. Cambridge University Press, Cambridge, England.

Du, J., S. D. Mansfield, and A. Groover. 2009. The *Populus* homeobox gene *ARBORKNOX2* regulates cell differentiation during secondary growth. The Plant Journal 60:1000-1014.

Du, Z., X. Zhou, Y. Ling, Z. Zhang, and Z. Su. 2010. agriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Research 38:W64-W70.

Dubreuil, M., M. Riba, and M. Mayol. 2008. Genetic structure and diversity in *Ramonda myconi* (Gesneriaceae): effects of historical climate change on a preglacial relict species. Am. J. Bot. 95:577-587.

Dudley, S. A. 1996a. Differing selection on plant physiological traits in response to environmental water availability: A test of adaptive hypotheses. Evolution 50:92-102.

Dudley, S. A. 1996b. The response to differing selection on plant physiological traits: evidence for local adaptation. Evolution 50:103-110.

Duminil, J., S. Fineschi, A. Hampe, P. Jordano, D. Salvini, G. G. Vendramin, and R. J. Petit. 2007. Can population genetic structure be predicted from life-history traits? American Naturalist 169:662-672.

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

Duplessis, S., I. T. Major, F. Martin, and A. Seguin. 2009. Poplar and pathogen interactions: Insights from *Populus* genome-wide analyses of resistance and defence gene families and gene expression profiling. Critical Reviews in Plant Sciences 28:309-334.

Eckenwalder, J. E. 1984a. Natural intersectional hybridization between North American species of *Populus* (Salicaceae) in sections *Aigeiros* and *Tacamahaca*. II. Taxonomy. Canadian Journal of Botany 62:325-335.

Eckenwalder, J. E. 1984b. Natural intersectional hybridization between North American species of *Populus* (Salicaceae) in sections *Aigerios* and *Tacamahaca*. III. Paleobotany and evolution. Canadian Journal of Botany 62:336-342.

Eckenwalder, J. E. 1996. Systematics and evolution of *Populus*. Pp. 7-32 *in* R. F. Stettler, H. D. Bradshaw, P. E. Heilman, and T. M. Hinckley, eds. Biology of *Populus* and its Implications for Management and Conservation. NRC Research Press, Ottawa, Canada.

Eckert, C. G., D. Manicacci, and S. C. H. Barrett. 1996. Genetic drift and founder effect in native versus introduced populations of an invading plant, *Lythrum salicaria* (Lythraceae). Evolution 50:1512-1519.

Edmands, S., S. L. Northrup, and A. S. Hwang. 2009. Maladapted gene complexes within populations of the intertidal copepod *Tigriopus californicus*? Evolution 63:2184-2192.

Ehrenreich, I. M., and M. D. Purugganan. 2006. The molecular genetic basis of plant adaptation. American Journal of Botany 93:953-962.

Ehrenreich, I. M., P. A. Stafford, and M. D. Purugganan. 2007. The genetic architecture of shoot branching in *Arabidopsis thaliana*: a comparative assessment of candidate gene associations *vs*. quantitative trait locus mapping. Genetics 176:1223-1236.

Erickson, D. L., C. B. Fenster, H. K. Stenoien, and D. Price. 2004. Quantitative trait locus analyses and the study of evolutionary process. Molecular Ecology 13:2505-2522.

Erickson, R. O., and F. J. Michelini. 1957. The Plastochron Index. American Journal of Botany 44:297-305.

Etterson, J. R. 2004. Evolutionary potenital of *Chamaecrista fasciculata* in relation to climate change. II. Genetic architecture of three populations reciprocally planted along an environmental gradiant in the Great Plains. Evolution 58:1459-1471.

Felsenstein, J. 2005. PHYLIP (Phylogeny Inference Package). Distributed by the author. Department of Genome Sciences, University of Washington, Seattle.

Ferris, R., L. Long, S. M. Bunn, K. M. Robinson, H. D. Bradshaw, A. M. Rae, and G. Taylor. 2002. Leaf stomatal and epidermal cell development: identification of putative quantitative trait loci in relation to elevated carbon dioxide concentration in poplar. Tree Physiology 22:633-640.

Ferris, R., M. Sabatti, F. Miglietta, R. F. Mills, and G. Taylor. 2001. Leaf area is stimulated in *Populus* by free air CO_2 enrichment (POPFACE), through increased cell expansion and production. Plant, Cell and Environment 24:305-315.

Fischer, M. K., and A. W. Shingleton. 2001. Host plant and ants influence the honeydew sugar composition of aphids. Functional Ecology 15:544-550.

Fleming, A. 2006a. The integration of cell proliferation and growth in leaf morphogenesis. Journal of Plant Research 119:31-36.

Fleming, A. 2006b. Leaf initiation: The integration of growth and cell division. Plant Molecular Biology 60:905-914.

Flint, J., and R. Mott. 2001. Finding the molecular basis of quantitative traits: successes and pitfalls. Nature Reviews Genetics 2:437-445.

Fossati, T., F. Grassi, F. Sala, and S. Castiglione. 2003. Molecular analysis of natural populations of *Populus nigra* L. intermingled with cultivated hybrids. Molecular Ecology 12:2033-2043.

Fossati, T., G. Patrignani, I. Zapelli, M. Sabatti, F. Sala, and S. Castiglione. 2004. Development of molecular markers to assess the level of introgression of *Populus tremula* into *P. alba* natural populations. Plant Breeding 123:382-385.

Francis, D. 2007. The plant cell cycle - 15 years on. New Phytologist 174:261-278.

Freimer, N. B., S. K. Service, R. A. Ophoff, A. J. Jasinska, K. McKee, A. Villeneuve, A. Belisle, J. N. Bailey, S. E. Breidenthal, M. J. Jorgensen, J. J. Mann, R. M. Cantor, K. Dewar, and L. A. Fairbanks. 2007. A quantitative trait locus for variation in dopamine metabolism mapped in a primate model using reference sequences from related species. Proceedings of the National Academy of Sciences 104:15811-15816.

Frewen, B. E., T. H. H. Chen, G. T. Howe, J. Davis, A. Rohde, W. Boerjan, and H. D. Bradshaw, Jr. 2000. Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. Genetics 154:837-845.

Gallo, L. A., and T. Geburek. 1991. Genetics of isozyme variants in *Populus tremula*, *Populus tremuloides* and their hybrids. Euphytica 53:225-234.

Gaudet, M., V. Jorge, I. Paolucci, I. Beritognolo, G. Scarascia-Mugnozza, and M. Sabatti. 2008. Genetic linkage maps of *Populus nigra* L. including AFLPs, SSRs, SNPs, and sex trait. Tree Genetics & Genomes 4:25-36.

Gentleman, R. C., V. J. Carey, D. M. Bates, B. Bolstad, M. Dettling, S. Dudoit, B. Ellis, L. Gautier, Y. Ge, J. Gentry, K. Hornik, T. Hothorn, W. Huber, S. Iacus, R. Irizarry, F. Leisch, C. Li, M. Maechler, A. J. Rossini, G. Sawitzki, C. Smith, G. Smyth, L. Tierney, J. Y. H. Yang, and J. Zhang. 2004. Bioconductor: open software development for computational biology and bioinformatics. Genome Biology 3:R80.

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

Gibbs, J. R., and A. Singleton. 2006. Application of genome-wide single nucleotide polymorphism typing: Simple association and beyond. Plos Genetics 2:1511-1517.

Glynn, C., A.-C. Rönnberg-Wästljung, R. Julkunen-Tiitto, and M. Weih. 2004. Willow genotype, but not drought treatment, affects foliar phenolic concentrations and leaf-beetle resistance. Entomologia Experimentalis et Applicata 113:1-14.

Gonzalez-Martinez, S. C., R. Alia, and L. Gil. 2002. Population genetic structure in a Mediterranean pine (*Pinus pinaster* Ait.): a comparison of allozyme markers and quantitative traits. Heredity 89:199-206.

Gonzalez-Martinez, S. C., K. V. Krutovsky, and D. B. Neale. 2006. Forest-tree population genomics and adaptive evolution. New Phytologist 170:227-238.

González-Martínez, S. C., S. Mariette, M. M. Ribeiro, C. Burban, A. Raffin, M. R. Chambel, C. A. M. Ribeiro, A. Aguiar, C. Plomion, R. Alía, L. Gil, G. G. Vendramin, and A. Kremer. 2004. Genetic resources in maritime pine (*Pinus pinaster* Aiton): molecular and quantitative measures of genetic variation and differentiation among maternal lineages. Forest Ecology and Management 197:103-115.

Goudet, J. 1995. Fstat version 1.2: a computer program to calculate F-statistics. Journal of Heredity 86:485-486.

Govindjee, J. Kern, J. Messinger, and J. Whitmarsh. 2010. Photosystem II. Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd., Chichester.

Granier, C., and F. Tardieu. 1998. Spatial and temporal analyses of expansion and cell cycle in sunflower leaves. Plant Physiology 116:991-1001.

Grivet, D., V. L. Sork, R. D. Westfall, and F. W. Davis. 2008. Conserving the evolutionary potential of California valley oak (*Quercus lobata* Née): a multivariate genetic approach to conservation planning. Molecular Ecology 17:139-156.

Groover, A., S. Mansfield, S. DiFazio, G. Dupper, J. Fontana, R. Millar, and Y. Wang. 2006a. The *Populus* homeobox gene *ARBORKNOX1* reveals overlapping mechanisms regulating the shoot apical meristem and the vascular cambium. Plant Molecular Biology 61:917-932.

Groover, A. T., S. D. Mansfield, S. P. DiFazio, G. Dupper, J. R. Fontana, R. Millar, and Y. Wang. 2006b. The *Populus* homeobox gene *ARBORKNOX1* reveals overlapping mechanisms regulating the shoot apical meristem and the vascular cambium. Plant Molecular Biology 61:917-932.

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

Hamrick, J. L., and M. J. W. Godt. 1996. Effects of life history traits on genetic diversity in plant species. Philosophical Transactions of the Royal Society B-Biological Sciences 351:1291-1298.

Hamzeh, M., and S. Dayanandan. 2004. Phylogeny of *Populus* (Salicaceae) based on nucleotide sequences of chloroplast TrnT-TrnF region and nuclear rDNA. American Journal of Botany. 91:1398-1408.

Han, K.-H., M. P. Gordon, and S. Strauss. 1996. Cellular and molecular biology of *Agrobacterium*-mediated transformation of plants and its application to genetic transformation of *Populus*. Pp. 201-222 *in* R. F. Stettler, H. D. Bradshaw, P. E. Heilman, and T. M. Hinckley, eds. Biology of *Populus* and its Implications for Management and Conservation. NRC Research Press, Ottawa.

Hardy, G. H. 1908. Mendelian Proportions in a Mixed Population. Science 28:49-50.

Hay, A., M. Barkoulas, and M. Tsiantis. 2006. ASYMMETRIC LEAVES1 and auxin activities converge to repress *BREVIPEDICELLUS* expression and promote leaf development in *Arabidopsis*. Development 133:3955-3961.

He, J.-S., X. Wang, D. F. B. Flynn, L. Wang, B. Schmid, and J. Fang. 2009. Taxonomic, phylogenetic, and environmental trade-offs between leaf productivity and persistence. Ecology 90:2779-2791.

Heinze, B. 2008. Genetic traces of cultivated hybrid poplars in the offspring of native *Populus nigra* in Austria. Preslia 80:365-374.

Henrissat, B., and G. Davies. 1997. Structural and sequence-based classification of glycoside hydrolases. Current Opinion in Structural Biology 7:637-644.

Hewitt, G. M. 1999. Post-glacial re-colonization of European biota. Biological Journal of the Linnean Society 68:87-112.

Hijmans, R. J., S. E. Cameron, J. L. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology 25:1965-1978.

Hipkins, V. D., and J. H. Kitzmiller. 2004. Genetic variation and clonal distribution of quaking aspen in the central Sierra Nevada. Transactions of the Western Section of the Wildlife Society 40:32-44.

Holderegger, R., S. Angelone, S. Brodbeck, D. Csencsics, F. Gugerli, S. E. Hoebee, and R. Finkeldey. 2005. Application of genetic markers to the discrimination of European Black Poplar (*Populus nigra*) from American Black Poplar (*P. deltoides*) and Hybrid Poplars (*P. x canadensis*) in Switzerland. Trees (Berlin) 19:742-747.

Horiguchi, G., A. Ferjani, U. Fujikura, and H. Tsukaya. 2006. Coordination of cell proliferation and cell expansion in the control of leaf size in *Arabidopsis thaliana*. Journal of Plant Research 119:37-42.

Hu, Y., and U. Schmidhalter. 2008. Spatial and temporal quantitative analysis of cell division and elongation rate in growing wheat leaves under saline conditions. Journal of Integrative Plant Biology 50:76-83.

Ibrahim, K. M., R. A. Nichols, and G. M. Hewitt. 1996. Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. Heredity 77:282-291.

Imbert, E., and F. LeFevre. 2003. Dispersal and gene flow of *Populus nigra* (Salicaceae) along a dynamic river system. Journal of Ecology 91:447-456.

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

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

International Rice Genome Sequencing Project. 2005. The map-based sequence of the rice genome. Nature 436:793-800.

Inze, D. 2005. Green light for the cell cycle. Embo Journal 24:657-662.

Irizarry, R. A., B. Hobbs, F. Collins, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, and T. P. Speed. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics (Oxford, England) 4:249-264.

Jarvis, P. G. 1976. The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 273:593-610.

Jermstad, K. D., D. L. Bassoni, N. C. Wheeler, and D. B. Neale. 1998. A sex-averaged genetic linkage map in coastal Douglas-fir (*Pseudotsuga menziesii*[Mirb.] Franco var '*menziesii'*) based on RFLP and RAPD markers. TAG Theoretical and Applied Genetics 97:762-770.

Johannsen, W. 1911. The genotype conception of heredity. American Naturalist 45:129-159.

Jombart, T., D. Pontier, and A.-B. Dufour. 2009. Genetic markers in the playground of multivariate analysis. Heredity 102:330-341.

Kajba, D., D. Ballian, M. Idžojtić, and S. Bogdan. 2004. The differences among hairy and typical European black poplars and the possibly role of the hairy type in relation to climatic changes. Forest Ecology and Management 197:279-284.

Kaschani, F., C. Gu, S. Niessen, H. Hoover, B. F. Cravatt, and R. A. L. van der Hoorn. 2009. Diversity of serine hydrolase activities of unchallenged and *Botrytis*-infected *Arabidopsis thaliana*. Molecular & Cellular Proteomics 8:1082-1093.

Kearsey, M. J., and A. G. L. Farquhar. 1998. QTL analysis in plants; where are we now? Heredity 80:137-142.

Keith, A. R., J. K. Bailey, and T. G. Whitham. 2010. A genetic basis to community repeatability and stability. Ecology 91:3398-3406.

Kelleher, C. T., R. Chiu, H. Shin, I. E. Bosdet, M. I. Krzywinski, C. D. Fjell, J. Wilkin, T. Yin, S. P. DiFazio, J. Ali, J. K. Asano, S. Chan, A. Cloutier, N. Girn, S. Leach, D. Lee, C. A. Mathewson, T. Olson, K. O'Connor, A.-L. Prabhu, D. E. Smailus, J. M. Stott, M. Tsai, N. H. Wye, G. S. Yang, J. Zhuang, R. A. Holt, N. H. Putnam, J. Vrebalov, J. J. Giovannoni, J. Grimwood, J. Schmutz, D. Rokhsar, S. J. M. Jones, M. A. Marra, G. A. Tuskan, J. Bohlmann, B. E. Ellis, K. Ritland, C. J. Douglas, and J. E. Schein. 2007. A physical map of the highly heterozygous *Populus* genome: integration with the genome sequence and genetic map and analysis of haplotype variation. Plant Journal 50:1063-1078.

Kelly, C. K., M. W. Chase, A. De Bruijn, M. F. Fay, and F. I. Woodward. 2003. Temperature-based population segregation in birch. Ecology Letters 6:87-89.

Kinoshita, I., A. Sanbe, and E. i. Yokomura. 2008. Difference in light-induced increase in ploidy level and cell size between adaxial and abaxial epidermal pavement cells of Phaseolus vulgaris primary leaves. J. Exp. Bot.:ern055.

Klaper, R., K. Ritland, T. A. Mousseau, and M. D. Hunter. 2001. Heritability of phenolics in *Quercus laevis* inferred using molecular markers. Journal of Heredity 92:421-426.

Knott, S. A., D. B. Neale, M. M. Sewell, and C. S. Haley. 1997. Multiple marker mapping of quantitative trait loci in an outbred pedigree of loblolly pine. Theoretical and Applied Genetics 94: 810-820.

Koch, K. E. 1996. Carbohydrate-modulated gene expression in plants. Annual Review of Plant Physiology and Plant Molecular Biology 47:509-540.

Kohn, M. H., W. J. Murphy, E. A. Ostrander, and R. K. Wayne. 2006. Genomics and conservation genetics. Trends in Ecology & Evolution 21:629-637.

Kopp, M., and J. Hermisson. 2007. Adaptation of a quantitative trait to a moving optimum. Genetics 176:715-719.

Kopp, M., and J. Hermisson. 2009. The genetic basis of phenotypic adaptation I: Fixation of beneficial mutations in the moving optimum model. Genetics 182:233-249.

Koskela, J., S. G. M. de Vries, D. Kajba, and G. von Wühlisch. 2004. *Populus nigra* Network, report of the seventh (25-27 October 2001, Osijek, Croatia) and eighth meetings (22-24 May 2003, Treppeln, Germany). International Plant Genetic Resources Institute, Rome, Italy.

Kover, P. X., and A. L. Caicedo. 2001. The genetic architecture of disease resistance in plants and the maintenance of recombination by parasites. Molecular Ecology 10:1-16.

Kremer, A., J. Kleinschmit, J. Cottrell, E. P. Cundall, J. D. Deans, A. Ducousso, A. O. König, A. J. Lowe, R. C. Munro, R. J. Petit, and B. R. Stephan. 2002. Is there a correlation between chloroplastic and nuclear divergence, or what are the roles of history and selection on genetic diversity in European oaks? Forest Ecology and Management 156:75-87.

Krutovsky, K., J. B. St. Clair, R. Saich, V. D. Hipkins, and D. B. Neale. 2009. Estimation of population structure in coastal Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco var. *menziesii*] using allozyme and microsatellite markers. Tree Genetics & Genomes 5:641-658.

Kuittinen, H., A. Mattila, and O. Savolainen. 1997. Genetic variation at marker loci and in quantiative traits in natural populations of *Arabidopsis thaliana*. Hereidy 79:144-152.

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

Lande, R. 1980. Genetic variation and phenotypic evolution during allopatric speciation. The American Naturalist 116:463-479.

Lande, R., and S. Shannon. 1996. The role of genetic variation in adaptation and population persistence in a changing environment. Evolution 50:434-437.

Latta, R. G. 1998. Differentiation of allelic frequencies at quantitative trait loci affecting locally adaptive traits. American Naturalist 151:283-292.

Latta, R. G., and J. K. McKay. 2002. Genetic population divergence: markers and traits. Trends in Ecology & Evolution 17:501-502.

Leakey, A. D. B., E. A. Ainsworth, S. M. Bernard, R. J. C. Markelz, D. R. Ort, S. A. Placella, A. Rogers, M. D. Smith, E. A. Sudderth, D. J. Weston, S. D. Wullschleger, and S. Yuan. 2009. Gene expression profiling: opening the black box of plant ecosystem responses to global change. Global Change Biology 15:1201-1213.

LeCorre, V., S. Dumolin-Lapegue, and A. Kremer. 1997a. Genetic variation at allozyme and RAPD loci in sessile oak *Quercus petraea* (Matt.) Liebl.: the role of history and geography. Molecular Ecology 6:519-529.

LeCorre, V., and A. Kremer. 2003. Genetic variability at neutral markers, quantitative trait loci and trait in a subdivided population under selection. Genetics 164:1205-1219.

LeCorre, V., N. Machon, R. J. Petit, and A. Kremer. 1997b. Colonization with long-distance seed dispersal and genetic structure of maternally inherited genes in forest trees: a simulation study. Genetical Research 69:117-125.

Legionnet, A., and F. Lefevre. 1996. Genetic variation of the riparian pioneer tree species *Populus nigra* L. I. Study of population structure based on isozymes. Heredity 77:629-637.

Leon-Reyes, A., Y. Du, A. Koornneef, S. Proietti, A. P. Korbes, J. Memelink, C. M. J. Pieterse, and T. Ritsema. 2010. Ethylene signaling renders the jasmonate response of *Arabidopsis* insensitive to future suppression by salicylic acid. Molecular Plant-Microbe Interactions 23:187-197.

LeRoy, C. J., T. G. Whitham, P. Keim, and J. C. Marks. 2006. Plant genes link forests and streams. Ecology (Tempe) 87:255-261.

Levée, V., I. T. Major, C. Levasseur, L. Tremblay, J. MacKay, and A. Séguin. 2009. Expression profiling and functional analysis of *Populus* WRKY23 reveals a regulatory role in defense. New Phytologist 184:48-70.

Lexer, C., C. A. Buerkle, J. A. Joseph, B. Heinze, and M. F. Fay. 2007. Admixture in European Populus hybrid zones makes feasible the mapping of loci that contribute to reproductive isolation and trait differences (September 20, 2006). Heredity 98:183-184.

Lexer, C., M. F. Fay, J. A. Joseph, M. S. Nica, and B. Heinze. 2005. Barrier to gene flow between two ecologically divergent *Populus* species, *P. alba* (white poplar) and *P. tremula* (European aspen): the role of ecology and life history in gene introgression. Molecular Ecology 14:1045-1057.

Liu, Z., and G. R. Furnier. 1993a. Comparison of allozyme, RFLP, and RAPD markers for revealing genetic variation within and between trembling aspen and bigtooth aspen. Theoretical and Applied Genetics 87:97-105.

Liu, Z., and G. R. Furnier. 1993b. Inheritance and linkage of allozymes and restriction fragment length polymorphisms in trembling aspen. Journal of Heredity 84:419-424.

Lohse, M., A. Nunes-Nesi, P. Kruger, A. Nagel, J. Hannemann, F. M. Giorgi, L. Childs, S. Osorio, D. Walther, J. Selbig, N. Sreenivasulu, M. Stitt, A. R. Fernie, and B. Usadel. 2010. Robin: An intuitive wizard application for R-based expression microarray quality assessment and analysis. Plant Physiology 153:642-651.

Lopez-Casado, G., B. R. Urbanowicz, C. M. B. Damasceno, and J. K. C. Rose. 2008. Plant glycosyl hydrolases and biofuels: a natural marriage. Current Opinion in Plant Biology 11:329-337.

Lopez-Fanjul, C., A. Fernandez, and M. Toro. 2003. The effect of neutral nonadditive gene action on the quantitative index of population divergence. Genetics 164:1527-1633.

Love, J., A. N. Dodd, and A. A. R. Webb. 2004. Circadian and diurnal calcium oscillations encode photoperiodic information in *Arabidopsis*. Plant Cell 16:956-966.

Loveless, M. D., and J. L. Hamrick. 1984. Ecological determinants of genetic structure in plant populations. Annual Review of Ecology and Systematics 15:65-95.

Luikart, G., and J.-M. Cornuet. 1998. Empirical evaluation of a test for identifying recently bottlenecked populations from allele frequency data. Conservation Biology 12:228-237.

Luikart, G., W. B. Sherwin, B. M. Steele, and F. W. Allendorf. 1998. Usefulness of molecular markers for detecting population bottlenecks via monitoring genetic change. Molecular Ecology 7:963-974.

Lynch, M., M. Pfrender, K. Spitze, N. Lehman, J. Hicks, D. Allen, L. Latta, M. Ottene, F. Bogue, and J. Colbourne. 1999. The quantitative and molecular genetic architecture of a subdivided species. Evolution 53:100-110.

Lynch, M., and B. Walsh. 1998. Genetics and analysis of quantitative traits. Sinauer Associates, Inc., Sunderland, Massachusettes.

Major, I. T., M.-C. Nicole, S. Duplessis, and S. Armand. 2010. Photosynthetic and respiratory changes in leaves of poplar elicited by rust infection. Photosynthetic Research 104:41-48.

Manly, B. F. J. 1997. Randomization, Bootstrap and Monte Carlo Methods in Biology. Chapman & Hall, London.

Marcotrigiano, M. 2010. A role for leaf epidermis in the control of leaf size and the rate and extent of mesophyll cell division. American Journal of Botany 97:224-233.

Marsch-Martinez, N., R. Greco, J. D. Becker, S. Dixit, J. H. W. Bergervoet, A. Karaba, S. de Folter, and A. Pereira. 2006. BOLITA, an *Arabidopsis* AP2/ERF-like transcription factor that affects cell expansion and proliferation/differentiation pathways. Plant Molecular Biology 62:825-843.

Matsubara, S., V. Hurry, N. Druart, C. Benedict, I. Janzik, A. Chavarria-Krauser, W. Achim, and U. Schurr. 2006. Nocturnal changes in leaf growth of *Populus deltoides* are controlled by cytoplasmic growth. Planta 223:1315-1328.

McKay, J. K., and R. G. Latta. 2002. Adaptive population divergence: markers, QTL and traits. Trends in Ecology & Evolution 17:285-291.

McLachlan, J. S., J. J. Hellmann, and M. W. Schwartz. 2007. A framework for debate of assisted migration in an era of climate change. Conservation Biology 21:297-302.

Meier, I., and C. Leuschner. 2008. Leaf size and leaf area index in *Fagus sylvatica* forests: competing effects of precipitation, temperature, and nitrogen availability. Ecosystems 11:655-669.

Melo-Oliveira, R., I. C. Oliveira, and G. M. Coruzzi. 1996. *Arabidopsis* mutant analysis and gene regulation define a nonredundant role for glutamate dehydrogenase in nitrogen assimilation. Proceedings of the National Acadamy of Sciences USA 93:4718-4723.

Meng, Q., K. Siebke, P. Lippert, B. Baur, U. Mukherjee, and E. Weis. 2001. Sink-source transition in tobacco leaves visualized using chlorophyll fluorescence imaging. New Phytologist 151:585-595.

Merila, J., and P. Crnokrak. 2001. Comparison of genetic differentiation at marker loci and quantitative traits. Journal of Evolutionary Biology 14:892-903.

Mitton, J. B., and K. L. Duran. 2004. Genetic variation in piñon pine, *Pinus edulis*, associated with summer precipitation. Molecular Ecology 13:1259-1264.

Miyashita, Y., and A. G. Good. 2008a. Glutamate deamination by glutamate dehydrogenase plays a central role in amino acid catabolism in plants. Plant Signaling & Behavior 3:842-843.

Miyashita, Y., and A. G. Good. 2008b. NAD(H)-dependent glutamate dehydrogenase is essential for the survival of *Arabidopsis thaliana* during dark-induced carbon starvation. Journal of Experimental Botany 59:667-680.

Miyazawa, S.-I., N. J. Livingston, and D. H. Turpin. 2006. Stomatal development in new leaves is related to the stomatal conductance of mature leaves in poplar (*Populus trichocarpa* x *P. deltoides*). Journal of Experimental Botany 57:373-380.

Morgante, M., and F. Salamini. 2003. From plant genomics to breeding practice. Current Opinion in Biotechnology 14:214-219.

Morreel, K., G. Goeminne, V. Storme, L. Sterck, J. Ralph, W. Coppieters, P. Breyne, M. Steenackers, M. Georges, E. Messens, and W. Boerjan. 2006. Genetical metabolomics of flavonoid biosynthesis in *Populus*: a case study. Plant Journal 47:224-237.

Moya-Larano, J., and G. Corcobado. 2008. Plotting partial correlation and regression in ecological studies. Web Ecology 8:35-46.

Moyle, L. C. 2006. Correlates of genetic differentiation and isolation by distance in 17 congeneric *Silene* species. Molecular Ecology 15:1067-1081.

Mukherjee, A. K., M.-J. Carp, R. Zuchman, T. Ziv, B. A. Horwitz, and S. Gepstein. 2010. Proteomics of the response of *Arabidopsis thaliana* to infection with *Alternaria brassicicola*. Journal of Proteomics 73:709-720.

Namroud, M.-C., J. Beaulieu, N. Juge, J. Laroche, and J. Bousquet. 2008. Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. Molecular Ecology 17:3599-3613.

Nason, J. D., J. L. Hamrick, and T. H. Fleming. 2002. Historical vicariance and postglacial colonization effects on the evolution of genetic structure in *Lophocereus*, a Sonoran desert columnar cactus. Evolution 56:2214-2226.

Neale, D. B. 2007. Genomics to tree breeding and forest health. Current Opinion in Genetics & Development 17:539-544.

Nei, M. 1972. Genetic distance between populations. The American Naturalist 106:283-292.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proceedings of the National Acadamy of Sciences USA 70:3321-3323.

Newcombe, G., and H. D. Bradshaw, Jr. 1996. Quantitative trait loci conferring resistance in hybrid poplar to *Septoria populicola*, the cause of leaf spot. Canadian Journal of Forest Research 26:1943-1950.

Newcombe, G., H. D. Bradshaw, Jr., G. A. Chastagner, and R. F. Stettler. 1996. A major gene for resistance to *Melampsora medusae* f. sp. *deltoidae* in a hybrid poplar pedigree. Phytopathology 86:87-94.

Newcombe, G., T. Bradshaw, G. Chastagner, and R. Stettler. 1995. A major gene for resistance in a native tree species, *Populus trichocarpa*, and its hybrids. Phytopathology 85:1140.

Newton, A. C., T. R. Allnutt, A. C. M. Gillies, A. J. Lowe, and R. A. Ennos. 1999. Molecular phylogeography, intraspecific variation and the conservation of tree species. Trends in Ecology & Evolution 14:140-145.

Nichols, R. A., and G. M. Hewitt. 1994. The genetic consequences of long distance dispersal during colonization. Heredity 72:213-217.

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

O'Reilly, P. T., m. F. Canino, K. M. Bailey, and P. Bentzen. 2004. Inverse relationship between FST and microsatellite polymorphism in the marine fish, walleye pollock (*Theragra chalcogramma*): implications for resolving weak population structure. Molecular Ecology 13:1799-1814.

Ochieng, J. W., A. W. T. Muigai, and G. N. Ude. 2007. Localizing genes using linkage disequilibrium in plants: Integrating lessons from the medical genetics. African Journal of Biotechnology 6:650-657.

Ohlemuller, R., E. S. Gritti, M. T. Sykes, and C. D. Thomas. 2006. Towards European climate risk surfaces: the extent and distribution of analogous and non-analogous climates 1931-2100. Global Ecology and Biogeography 15:395-405.

Orians, C. M. 2000. The effects of hybridization in plants on secondary chemistry: implications for the ecology and evolution of plant-herbivore interactions. American Journal of Botany. 87:1749-1756.

Palme, A. E., Q. Su, A. Rautenberg, F. Manni, and M. Lascoux. 2003. Postglacial recolonization and cpDNA variation of silver birch, *Betula pendula*. Molecular Ecology 12:201-212.

Park, C.-R., T. Hino, and H. Itô. 2008. Prey distribution, foliage structure, and foraging behavior of insectivorous birds in two oak species (*Quercus serrata* and *Q. variabilis*). Ecological Research 23:1015-1023.

Parkhurst, D. F., and O. L. Loucks. 1972. Optimal leaf size in relation to environment. Journal of Ecology 60:505-537.

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

Parry, M. A. J., D. Habash, and J. L. Araus. 2004. Optimisation of water use by plants. Annals of Applied Biology 144:125-126.

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

Peakall, R., P. E. Smouse, and D. R. Huff. 1995. Evolutionary implications of allozyme and RAPD variation in diploid populations of dioecious buffalograss *Buchloë dactyloides*. Molecular Ecology 4:135-147.

Penuelas, J., P. Prieto, C. Beier, C. Cesaraccio, P. de Angelis, G. de Dato, B. A. Emmett, M. Estiarte, J. Garadnai, A. Gorissen, E. K. Lang, G. Kroel-Dulay, L. Llorens, G. Pellizzaro, T. Riis-Nielsen, I. K. Schmidt, C. Sirca, A. Sowerby, D. Spano, and A. Tietema. 2007. Response of plant species richness and primary productivity in shrublands along a north-south gradient in Europe to seven years of experimental warming and drought: reductions in primary productivity in the heat and drought year of 2003. Global Change Biology 13:2563-2581.

Petit, R. J., R. Bialozyt, P. Garnier-Gere, and A. Hampe. 2004. Ecology and genetics of tree invasions: from recent introductions to Quaternary migrations. Forest Ecology and Management 197:117-137.

Petit, R. J., S. Brewer, S. Bordacs, K. Burg, R. Cheddadi, E. Coart, J. Cottrell, U. M. Csaikl, B. van Dam, J. D. Deans, S. Espinel, S. Fineschi, R. Finkeldey, I. Glaz, P. G. Goicoechea, J. S. Jensen, A. O. Konig, A. J. Lowe, S. F. Madsen, G. Matyas, R. C. Munro, F. Popescu, D. Slade, H. Tabbener, S. G. M. de Vries, B. Ziegenhagen, J.

L. de Beaulieu, and A. Kremer. 2002. Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. Forest Ecology and Management 156:49-74.

Petit, R. J., A. Kremer, and D. B. Wagner. 1993. Finite island model for organelle and nuclear genes in plants. Heredity 71:630-641.

Philippe, R., and J. Bohlmann. 2007. Poplar defense against insect herbivores. Canadian Journal of Botany 85:1111-1126.

Philippe, R. N., S. G. Ralph, C. Külheim, S. I. Jancsik, and J. Bohlmann. 2009. Poplar defense against insects: genome analysis, full-length cDNA cloning, and transcriptome and protein analysis of the poplar Kunitz-type protease inhibitor family. New Phytologist 184:865-884.

Philippe, R. N., S. G. Ralph, S. D. Mansfield, and J. Bohlmann. 2010. Transcriptome profiles of hybrid poplar (*Populus trichocarpa* × *deltoides*) reveal rapid changes in undamaged, systemic sink leaves after simulated feeding by forest tent caterpillar (*Malacosoma disstria*). New Phytologist 188:787-802.

Picotte, J., D. Rosenthal, J. Rhode, and M. Cruzan. 2007. Plastic responses to temporal variation in moisture availability: consequences for water use efficiency and plant performance. Oecologia 153:821-832.

Pien, S. p., J. Wyrzykowska, S. McQueen-Mason, C. Smart, and A. Fleming. 2001. Local expression of expansin induces the entire process of leaf development and modifies leaf shape. Proceedings of the National Academy of Sciences of the United States of America 98:11812-11817.

Pierik, R., D. Tholen, H. Poorter, E. J. W. Visser, and L. A. C. J. Voesenek. 2006. The Janus face of ethylene: growth inhibition and stimulation. Trends in Plant Science 11:176-183.

Podolsky, R. H., and T. P. Holtsford. 1995. Population structure of morphological traits in *Clarkia dudleyana* I. Comparison of Fst between allozymes and morphological traits. Genetics 140:733-744.

Poorter, H. 1989. Interspecific variation in relative growth rate: on ecological causes and physiological consequences. Pp. 45-68 *in* H. Lambers, M. L. Cambridge, H. Konings, and T. L. Pons, eds. Causes and consequences of variation in growth rate and productivity of higher plants. SPB Academic, The Hague, The Netherlands.

Poorter, H., Ü. Niinemets, L. Poorter, I. J. Wright, and R. Villar. 2009. Causes and consequences of variation in leaf mass per area (LMA): a meta-analysis. New Phytologist 182:565-588.

Poorter, L., and D. Rozendaal. 2008. Leaf size and leaf display of thirty-eight tropical tree species. Oecologia 158:35-46.

Pritchard, J. K., M. Stephens, and P. Donnelley. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Rae, A. M., R. Ferris, M. J. Tallis, and G. Taylor. 2006. Elucidating genomic regions determining enhanced leaf growth and delayed senescence in elevated CO_2. Plant Cell and Environment 29:1730-1741.

Rae, A. M., M. P. C. Pinel, C. Bastien, M. Sabatti, N. R. Street, J. Tucker, C. Dixon, N. Marron, S. Y. Dillen, and G. Taylor. 2008. QTL for yield in bioenergy *Populus*: identifying GxE interactions from growth at three contrasting sites. Tree Genetics & Genomes 4:97-112.

Rae, A. M., K. M. Robinson, N. R. Street, and G. Taylor. 2004. Morphological and physiological traits influencing biomass productivity in short-rotation coppice poplar. Canadian Journal of Forest Research 34:1488-1498.

Rae, A. M., N. Street, K. Robinson, N. Harris, and G. Taylor. 2009. Five QTL hotspots for yield in short rotation coppice bioenergy poplar: The Poplar Biomass Loci. BMC Plant Biology 9:23.

Rae, A. M., P. J. Tricker, S. M. Bunn, and G. Taylor. 2007. Adaptation of tree growth to elevated CO2: quantitative trait loci for biomass in Populus. New Phytologist 175:59-69.

Raible, C. C., C. Casty, J. Luterbacher, A. Pauling, J. Esper, D. C. Frank, U. Buntgen, A. C. Roesch, P. Tschuck, M. Wild, P. L. Vidale, C. Schar, and H. Wanner. 2006. Climate variability-observations, reconstructions, and model simulations for the Atlantic-European and Alpine region from 1500-2100 AD. Climatic Change 79:9-29.

Rajora, O. P. 1990. Genetics of allozymes in *Populus deltoides* Marsh., *Populus nigra* L., and *Populus maximowiczii* Henry. Journal of Heredity 81:301-308.

Rajora, O. P., and B. P. Dancik. 1995. Chloroplast DNA variation in *Populus*. II. Interspecific restriction fragment polymorphisms and genetic relationships among *Populus deltoides, P. nigra, P. maximowiczii,* and *P.* x *canadensis*. Theoretical and Applied Genetics 90:324-330.

Ralph, S. G., H. J. E. Chun, D. Cooper, R. Kirkpatrick, N. Kolosova, L. Gunter, G. Tuskan, C. Douglas, R. Holt, S. J. M. Jones, M. Marra, and J. Bohlmann. 2008. Analysis of 4,664 high-quality sequence-finished poplar full-length cDNA clones and their utility for the discovery of genes responding to insect feeding. BMC Genomics 9:57.

Ramirez-Parra, E., B. Desvoyes, and C. Gutierrez. 2005. Balance between cell division and differentiation during plant development. International Journal of Developmental Biology 49:467-477.

Regier, N., S. Streb, C. Cocozza, M. Schaub, P. Cherubini, S. C. Zeeman, and B. Frey. 2009. Drought tolerance of two black poplar (*Populus nigra* L.) clones: contribution of carbohydrates and oxidative stress defence. Plant, Cell & Environment 32:1724-1736.

Reinhardt, B., E. Hänggi, S. Müller, M. Bauch, J. Wyrzykowska, R. Kerstetter, S. Poethig, and A. J. Fleming. 2007. Restoration of *DWF4* expression to the leaf margin of a *dwf4* mutant is sufficient to restore leaf shape but not size: the role of the margin in leaf development. The Plant Journal 52:1094-1104.

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution 43:223-225.

Ridge, C. R., T. M. Hinckley, R. F. Stettler, and E. Van Volkenburgh. 1986. Leaf growth characteristics of fast-growing poplar hybrids *Populus trichocarpa* x *P. deltoides*. Tree Physiology 1:209-216.

Rieseberg, L., M. A. Archer, and R. K. Wayne. 1999. Transgressive segregation, adaptation and speciation. Heredity 83:363-372.

Robinson, S. A., A. P. Slade, G. G. Fox, R. Phillips, R. G. Ratcliffe, and G. R. Stewart. 1991. The role of glutamate dehydrogenase in plant nitrogen metabolism. Plant Physiology. 95:509-516.

Rodriguez-Acosta, M. 2006. The genetic basis of drought tolerance in *Populus*. Pp. 283. PhD Thesis, School of Biological Sciences. University of Southampton, Southampton.

Roeder, A. H. K., V. Chickarmane, A. Cunha, B. Obara, B. S. Manjunath, and E. M. Meyerowitz. 2010. Variability in the control of cell division underlies sepal epidermal patterning in *Arabidopsis thaliana*. PLoS Biol 8:e1000367.

Roff, D. A. 2007. A centennial celebration for quantitative genetics. Evolution 61:1017-1032.

Rogers, A. R., and H. C. Harpending. 1983. Population structure and quantitative characters. Genetics 105:985-1002.

Rohde, A., V. Storme, V. Jorge, M. Gaudet, N. Vitacolonna, F. Fabbrini, T. Ruttink, G. Zaina, N. Marron, S. Dillen, M. Steenackers, M. Sabatti, M. Morgante, W. Boerjan, and C. Bastien. 2011. Bud set in poplar – genetic dissection of a complex trait in natural and hybrid populations. New Phytologist 189:106-121.

Rönnberg-Wästljung, A.-C., I. Åhman, C. Glynn, and O. Widenfalk. 2006. Quantitative trait loci for resistance to herbivores in willow: field experiments with varying soils and climates. Entomologia Experimentalis et Applicata 118:163-174.

Rousset, F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. Genetics 145:1219-1228.

Royer, D. L., J. C. McElwain, J. M. Adams, and P. Wilf. 2008. Sensitivity of leaf size and shape to climate within *Acer rubrum* and *Quercus kelloggii*. New Phytologist 179:808-817.

Sathyanarayanan, P. V., and B. W. Poovaiah. 2004. Decoding Ca^{2+} signals in plants. Critical Reviews in Plant Sciences 23:1 - 11.

Sato, M., and D. Waxman. 2008. Adaptation to slow environmental change, with apparent anticipation of selection. Journal of Theoretical Biology 252:166-172.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nature Biochemistry 18:233-234.

Schwarzback, A. E., L. A. Donovan, and L. Rieseberg. 2001. Transgressive character expression in a hybrd sunflower species. American Journal of Botany 88:270-277.

Schweitzer, J. A., M. D. Madritch, J. K. Bailey, C. J. LeRoy, D. G. Fischer, B. Rehill, R. L. Lindroth, A. E. Hagerman, S. Wooley, S. C. Hart, and T. G. Whitham. 2008. From genes to ecosystems: the genetic basis of condensed tannins and their role in nutrient regulation in a *Populus* model system. Ecosystems 11:1005-1020.

Scotti, I., G. G. Vendramin, L. S. Matteotti, C. Scarponi, and M. Sari-Gorla. 2000. Postglacial recolonization routes for *Picea abies* K. in Italy as suggested by the analysis of sequence-characterized amplified region (SCAR) markers. Molecular Ecology 9.

Seaton, G., C. S. Haley, S. A. Knott, M. Kearsey, and P. M. Visscher. 2002. QTL Express: mapping quantitative trait loci in simple and complex pedigrees. Bioinformatics 18:339-340.

Seaton, G., J. Hernandez, J. A. Grunchec, I. White, J. Allen, D. J. De Koning, W. Wei, D. Berry, C. Haley, and S. Knott. 2006. GridQTL: A Grid portal for QTL mapping of compute intensive datasets. Proceedings of the 8th World Congress on Genetics Applied to Livestock Production, Belo Horizonte, Brazil.

Shen, W. H. 2001. The plant cell cycle: G1/S regulation. Euphytica 118:223-232.

Sjödin, A., K. Wissel, M. Bylesjö, J. Trygg, and S. Jansson. 2008. Global expression profiling in leaves of free-growing aspen. BMC Plant Biology 8:61.

Slatkin, M. 1993. Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47:264-279.

Smith, R. L., and K. J. Sytsma. 1990. Evolution of *Populus nigra* (Sect. *Aigeiros*): introgressive hybridization and the chloroplast contribution of *Populus alba* (Sect. Populus). American Journal of Botany 77:1176-1187.

Smouse, P. E., J. C. Long, and R. R. Sokal. 1986. Multiple regression and correlation extensions of the Mantel test of matrix correspondence. Systematic Zoology 35:627-632.

Smouse, P. E., and R. Peakall. 1999. Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. Heredity 82:561-573.

Smulders, M., R. Beringen, R. Volosyanchuk, A. Vanden Broeck, J. van der Schoot, P. Arens, and B. Vosman. 2008a. Natural hybridisation between *Populus nigra* L. and *P* . x *canadensis* Moench. hybrid offspring competes for niches along the Rhine river in the Netherlands. Tree Genetics & Genomes 4:663-675.

Smulders, M. J. M., J. E. Cottrell, F. Lefevre, J. van der Schoot, P. Arens, B. Vosman, H. E. Tabbener, F. Grassi, T. Fossati, S. Castiglione, V. Krystufek, S. Fluch, K. Burg, B. Vornam, A. Pohl, K. Gebhardt, N. Alba, D. Agundez, C. Maestro, E. Notivol, R. Volosyanchuk, M. Pospiskova, S. Bordacs, J. Bovenschen, B. C. van Dam, H. P. Koelewijn, D. Halfmaerten, B. Ivens, J. van Slycken, A. Vanden Broeck, V. Storme, and W. Boerjan. 2008b. Structure of the genetic diversity in black poplar (*Populus nigra* L.) populations across European river systems: Consequences for conservation and restoration. Forest Ecology and Management 255:1388-1399.

Smulders, M. J. M., J. van der Schoot, P. Arens, and B. Vosman. 2001. Trinucleotide repeat microsatellite markers for black poplar (*Populus nigra* L.). Molecular Ecology Notes 1:188-190.

Smyth, G. K. 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Statistical Applications in Genetics and Molecular Biology 3:Article 3.

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

Sterky, F., R. R. Bhalerao, P. Unneberg, B. Segerman, P. Nilsson, A. M. Brunner, L. Charbonnel-Campaa, J. J. Lindvall, K. Tandre, S. H. Strauss, B. Sundberg, P. Gustafsson, M. Uhlen, R. P. Bhalerao, O. Nilsson, G. Sandberg, J. Karlsson, J. Lundeberg, and S. Jansson. 2004. A Populus EST resource for plant functional genomics. Proceedings of the National Academy of Sciences of the United States of America 101:13951-13956.

Stevens, M. T., E. L. Kruger, and R. L. Lindroth. 2008. Variation in tolerance to herbivory is mediated by differences in biomass allocation in aspen. Functional Ecology 22:40-47.

Stevens, M. T., M. G. Turner, G. A. Tuskan, W. H. Romme, L. E. Gunter, and D. M. Waller. 1999. Genetic variation in postfire aspen seedlings in Yellowstone National Park. Molecular Ecology 8:1769-1780.

Stone, J. M., and J. C. Walker. 1995. Plant protein kinase families and signal transduction. Plant Physiology. 108:451-457.

Storme, V., A. Vanden Broeck, B. Ivens, D. Halfmaerten, J. Van Slycken, S. Castiglione, F. Grassi, T. Fossati, J. E. Cottrell, H. E. Tabbener, F. LeFevre, C. Saintagne, S. Fluch, V. Krystufek, K. Burg, S. Bordacs, A. Borovics, K. Gebhardt, B. Vornam, A. Pohl, N. Alba, D. Agundez, C. Maestro, E. Notivol, J. Bovenschen, B. C. Van Dam, J. van der Schoot, B. Vosman, W. Boerjan, and M. J. M. Smulders. 2004. *Ex-situ* conservation of black poplar in Europe: genetic diversity in nine gene bank collections and their value for nature development. Theoretical and Applied Genetics 108:969-981.

Storz, J. F. 2005. Using genome scans of DNA polymorphism to infer adaptive population divergence. Molecular Ecology 14:671-688.

Street, N., A. Sjodin, M. Bylesjo, P. Gustafsson, J. Trygg, and S. Jansson. 2008. A cross-species transcriptomics approach to identify genes involved in leaf development. BMC Genomics 9:589.

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

Styrsky, J. D., and M. D. Eubanks. 2007. Ecological consequences of interactions between ants and honeydew-producing insects. Proceedings of the Royal Society B, Biological Sciences 274:151-164.

Svenning, J. C., and F. Skov. 2006. Potential impact of climate change on the northern nemoral forest herb flora of Europe. Biodiversity and Conservation 15:3341-3356.

Taylor, G. 2002. *Populus*: *Arabidopsis* for forestry. Do we need a model tree? Annals of Botany (London) 90:681-689.

Taylor, G., P. J. Tricker, F. Z. Zhang, V. J. Alston, F. Miglietta, and E. Kuzminsky. 2003. Spatial and temporal effects of free-air CO_2 enrichment (POPFACE) on leaf growth, cell expansion, and cell production in a closed canopy of poplar. Plant Physiology 131:177-185.

Thaler, J. S., A. A. Agrawal, and R. Halitschke. 2010. Salicylate-mediated interactions between pathogens and herbivores. Ecology 91:1075-1082.

Theodoris, G., N. Inada, and M. Freeling. 2003. Conservation and molecular dissection of ROUGH SHEATH2 and ASYMMETRIC LEAVES1 function in leaf development. Proceedings of the National Academy of Sciences of the United States of America 100:6837-6842.

Thimm, O., O. Blasing, Y. Gibon, A. Nagel, S. Meyer, P. Kruger, J. Selbig, L. A. Muller, S. Y. Rhee, and M. Stitt. 2004. MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. The Plant Journal 37:914-939.

Tisné, S., M. Reymond, D. Vile, J. Fabre, M. Dauzat, M. Koornneef, and C. Granier. 2008. Combined genetic and modeling approaches reveal that epidermal cell area and number in leaves are controlled by leaf and plant developmental processes in Arabidopsis. Plant Physiology 148:1117-1127.

Torii, K. U. 2000. Receptor kinase activation and signal transduction in plants: an emerging picture. Current Opinion in Plant Biology 3:361-367.

Trainotti, L., A. Pavanello, and G. Casadoro. 2004. Differential expression of genes in apical and basal tissues of expanding tobacco leaves. Plant Science 167:679-686.

Trenberth, K. E., P. D. Jones, P. Ambenje, R. Bojariu, D. Easterling, A. Klein Tank, D. Parker, F. Rahimzadeh, J. A. Renwick, M. Rusticucci, B. Soden, and P. Zhai. 2007. Observations: Surface and atmospheric climate change. *in* S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, eds. Climate Change 2007: The physical science basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change., Cambridge, United Kingdom.

Trewin, H. 2008. Linking transcript, QTL and association mapping to understand the genetic control of leaf size and shape in *Populus*. PhD Thesis. School of Biological Sciences. University of Southampton, Southampton.

Tschaplinski, T. J., G. A. Tuskan, M. M. Sewell, G. M. Gebre, D. E. Todd, and C. D. Pendley. 2006. Phenotypic variation and quantitative trait locus identification for osmotic potential in an interspecific hybrid inbred F-2 poplar pedigree grown in contrasting environments. Tree Physiology 26:595-604.

Tsukaya, H. 2005. Leaf shape: genetic controls and environmental factors. International Journal of Developmental Biology 49:547-555.

Turesson, G. 1922. The genotypical response of the plant species to the habitat. Hereditas 3:211-350.

Turnipseed, S. G. 1977. Influence of trichome variations on populations of small phytophagous insects in soybean. Environmental Entomology 6:815-817.

Tuskan, G. A., S. DiFazio, S. Jansson, J. Bohlmann, I. Grigoriev, U. Hellsten, N. Putnam, S. Ralph, S. Rombauts, A. Salamov, J. Schein, L. Sterck, A. Aerts, R. R. Bhalerao, R. P. Bhalerao, D. Blaudez, W. Boerjan, A. Brun, A. Brunner, V. Busov, M. Campbell, J. Carlson, M. Chalot, J. Chapman, G. L. Chen, D. Cooper, P. M. Coutinho, J. Couturier, S. Covert, Q. Cronk, R. Cunningham, J. Davis, S. Degroeve, A. Dejardin, C. dePamphilis, J. Detter, B. Dirks, I. Dubchak, S. Duplessis, J. Ehlting, B. Ellis, K. Gendler, D. Goodstein, M. Gribskov, J. Grimwood, A. Groover, L. Gunter, B. Hamberger, B. Heinze, Y. Helariutta, B. Henrissat, D. Holligan, R. Holt, W. Huang, N. Islam-Faridi, S. Jones, M. Jones-Rhoades, R. Jorgensen, C. Joshi, J. Kangasjarvi, J. Karlsson, C. Kelleher, R. Kirkpatrick, M. Kirst, A. Kohler, U. Kalluri, F. Larimer, J. Leebens-Mack, J. C. Leple, P. Locascio, Y. Lou, S. Lucas, F. Martin, B. Montanini, C. Napoli, D. R. Nelson, C. Nelson, K. Nieminen, O. Nilsson, V. Pereda, G. Peter, R. Philippe, G. Pilate, A. Poliakov, J. Razumovskaya, P. Richardson, C. Rinaldi, K. Ritland, P. Rouze, D. Ryaboy, J. Schmutz, J. Schrader, B. Segerman, H. Shin, A. Siddiqui, F. Sterky, A. Terry, C. J. Tsai, E. Uberbacher, P. Unneberg, J. Vahala, K. Wall, S. Wessler, G. Yang, T. Yin, C. Douglas, M. Marra, G. Sandberg, Y. V. de Peer, and D. Rokhsar. 2006. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). Science (Washington D C) 313:1596-1604.

Tuskan, G. A., K. E. Francis, S. L. Russ, W. H. Romme, and M. G. Turner. 1996. RAPD markers reveal diversity within and among clonal and seedling stands of aspen in Yellowstone National Park, U.S.A. Canadian Journal of Forest Research 26:2088-2098.

Tuskan, G. A., L. E. Gunter, Z. K. Yang, T. Yin, M. M. Sewell, and S. P. DiFazio. 2004. Characterization of microsatellites revealed by genomic sequencing of *Populus trichocarpa*. Canadian Journal of Forest Research 34:85-93.

Usadel, B., F. Poree, A. Nagel, M. Lohse, A. Czedik-Eysenberg, and M. Stitt. 2009. A guide to using MapMan to visualize and compare omics data in plants: a case study in the crop species, maize. Plant, Cell & Environment 32:1211-1229.

Vanden Broeck, A., J. Cottrell, P. Quataert, P. Breyne, V. Storme, W. Boerjan, and J. Van Slycken. 2006. Paternity analysis of *Populus nigra* L. offspring in a Belgian plantation of native and exotic poplars. Annals of Forest Science 63:783-790.

van der Schoot, J., M. Pospiskova, B. Vosman, and M. J. M. Smulders. 2000. Development and characterization of microsatellite markers in black poplar (*Populus nigra* L.). Theoretical and Applied Genetics 101:317-322.

van Oosterhout, C., W. F. Hutchinson, D. P. M. Willis, and P. Shipley. 2004. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4:535-538.

Van Volkenburgh, E. 1999. Leaf expansion - an integrating plant behaviour. Plant, Cell & Environment 22:1463-1473.

Van Volkenburgh, E., and G. Taylor. 1996. Leaf growth physiology. Pp. 283-299 *in* R. F. Stettler, H. D. Bradshaw, P. E. Heilman, and T. M. Hinckley, eds. Biology of *Populus* and its Implications for Management and Conservation. NRC Research Press, Ottawa, Canada.

Vasemägi, A., and C. R. Primmer. 2005. Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. Molecular Ecology 14:3623-3642.

Wagstaff, C., G. J. J. Clarkson, F. Zhang, S. D. Rothwell, S. C. Fry, G. Taylor, and M. S. Dixon. 2010. Modification of cell wall properties in lettuce improves shelf life. Journal of Experimental Botany 61:1239-1248.

Waldmann, P., and S. Andersson. 1998. Comparison of quantitative genetic variation and allozyme diversity within and between populations of *Scabiosa canescens* and *S. columbaria*. Heredity 81:79-86.

Wang, H. L., S. D. Yang, and C. L. Zhang. 1997. The photosynthetic characteristics of differently shaped leaves in *Populus euphratica* Olivier. Photosynthetica (Prague) 34:545-553.

Wang, J., and A. Caballero. 1999. Developments in predicting the effective size of subdivided populations. Heredity 82:212-226.

Wei, F., E. Coe, W. Nelson, A. K. Bharti, F. Engler, E. Butler, H. Kim, J. L. Goicoechea, M. Chen, S. Lee, G. Fuks, H. Sanchez-Villeda, S. Schroeder, Z. Fang, M. McMullen, G. Davis, J. E. Bowers, A. H. Paterson, M. Schaeffer, J. Gardiner, K. Cone, J. Messing, C. Soderlund, and R. A. Wing. 2007. Physical and genetic structure of the maize genome reflects its complex evolutionary history. PLoS Genetics 3:e123.

Weijschede, J., K. Antonise, H. de Caluwe, H. de Kroon, and H. Huber. 2008. Effects of cell number and cell size on petiole length variation in a stoloniferous herb. American Journal of Botany 95:41-49.

Weinberg, W. 1908. Über den Nachweis der Vererbung veim Menschen. Jahreshefte des Vereins für vaterländische Naturkunde in Wüttemberg 64:368-382.

Westoby, M., and I. J. Wright. 2006. Land-plant ecology on the basis of functional traits. Trends in Ecology & Evolution 21:261-268.

Whitlock, M. C. 1999. Neutral additive variance in a metapopulation. Genetical Research 74:215-221.

Williams, B. K. 1983. Observations of the use of discriminant analysis in ecology. Ecology 84:1283-1291.

Wimp, G. M., S. Wooley, R. K. Bangert, W. P. Young, G. D. Martinsen, P. Keim, B. Rehill, R. L. Lindroth, and T. G. Whitham. 2007. Plant genetics predicts intra-annual variation in phytochemistry and arthropod community structure. Molecular Ecology 16:5057-5069.

Winfield, M., and F. M. R. Hughes. 2002. Variation in *Populus nigra* clones: Implications for river restoration projects in the United Kingdom. Wetlands 22:33-48.

Woo, K.-S., and G. Newcombe. 2003. Absence of residual effects of a defeated resistance gene in poplar. Forest Pathology 33:81-89.

Wright, S. 1931. Evolution in Mendelian populations. Genetics 16:97-159.

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

Wu, R., H. D. Bradshaw, Jr., and R. F. Stettler. 1997. Molecular genetics of growth and development in *Populus* (Salicaceae). V. Mapping quantitative trait loci affecting leaf variation. American Journal of Botany 84:143-153.

Wu, R., H. D. Bradshaw, Jr., and R. F. Stettler. 1998. Developmental quantitative genetics of growth in Populus. Theoretical and Applied Genetics 97:1110-1119.

Wu, R. L., Y. F. Han, J. J. Hu, J. J. Fang, L. Li, M. L. Li, and Z. B. Zeng. 2000. An integrated genetic map of *Populus deltoides* based on amplified fragment length polymorphisms. Theoretical and Applied Genetics 100:1249-1256.

Yang, D., G. Li, and S. Sun. 2008. The generality of leaf size versus number trade-off in temperate woody species. Annals of Botany 102:623-629.

Yin, T. M., S. P. DiFazio, L. E. Gunter, D. Riemenschneider, and G. A. Tuskan. 2004. Large-scale heterospecific segregation distortion in *Populus* revealed by a dense genetic map. Theoretical and Applied Genetics 109:451-463.

Yin, T. M., X. Y. Zhang, M. R. Huang, M. X. Wang, Q. Zhuge, S. M. Tu, L. H. Zhu, and R. L. Wu. 2002. Molecular linkage maps of the *Populus* genome. Genome 45:541-555.

Young, B., D. Wagner, P. Doak, and T. Clausen. 2010. Induction of phenolic glycosides by quaking aspen (*Populus tremuloides*) leaves in relation to extrafloral nectaries and epidermal leaf mining. Journal of Chemical Ecology 36:369-377.

Zhu, J.-K. 2002. Salt and drought stress signal transduction in plants. Annual Review of Plant Biology 53:247-273.

Ziegenhagen, B., S. Gneuss, G. Rathmacher, I. Leyer, R. Bialozyt, B. Heinze, and S. Liepelt. 2008. A fast and simple genetic survey reveals the spread of poplar hybrids at a natural Elbe river site. Conservation Genetics 9:373-379.

Zsuffa, L., E. Giordano, L. D. Pryor, and R. F. Stettler. 1996. Trends in poplar culture: some global and regional perspectives. Pp. 515-539 *in* R. F. Stettler, H. D. Bradshaw, P. E. Heilman, and T. M. Hinckley, eds. Biology of *Populus* and its Implications for Management and Conservation. NRC Research Press, Ottawa, Canada.