Available online: www.notulaebotanicae.ro
Print ISSN 0255-965X; Electronic 1842-4309
Not Bot Horti Agrobo, 2014, 42(2):299-309. DOI:10.15835/nbha4229739

Strategies to Identify Adaptive Genes in Hybridizing Trees like Oaks and Poplars

Oliver GAILING

Michigan Technological University, 1400 Townsend Drive, 49931 Houghton, MI, United States of America; ogailing@mtu.edu

Abstract

Ecologically divergent, hybridizing species such as oaks and poplars provide models to identify genomic regions under selection and adaptive alleles that are transferred between species in hybrid zones. Oaks show patterns of genomic divergence characteristic for early stages of speciation with gene flow, in which large genomic regions are homogenized by interspecific gene flow interspersed by smaller regions (outlier regions) with high interspecific differentiation as result of divergent selection. These outlier regions can be identified using genome scans in hybrid zones and anchored to the *Quercus robur* genome sequence which will become available in the near future. Combined outlier and association genetic approaches can assess the role of individual genes in outlier genomic regions in adaptive trait variation. In contrast, hybridizing poplar species show a pattern of genomic divergence with large genomic regions of high interspecific differentiation punctuated by smaller regions of low differentiation as the result of interspecific gene flow. Genome scans in multiple hybrid zones of interfertile poplar species and in populations outside the area of sympatry will allow for the identification of genes that are exchanged between species by interspecific gene flow using the *Populus trichocarpa* genome sequence as a reference. Again association genetic approaches can be used for the characterization of variation in these introgressed genes with adaptive trait variation. In the present paper, the application of genomic approaches to identify genes for adaptive species divergence and reproductive isolation, and introgressed genes between species is discussed.

Keywords: ecological speciation, genome scans, *Quercus*, *Populus*, reproductive isolation, Simple Sequence Repeats (SSRs)

Introduction

In early stages of speciation, genome scans in hybrid zones can discriminate genomic regions with high interspecific differentiation as result of divergent selection (outlier regions) from selectively neutral genomic regions that are homogenized by interspecific gene flow (e.g. Nosil *et al.*, 2009). The size of these outlier genomic regions increases with divergence time between species and the evolution of intrinsic and extrinsic barriers to interspecific gene flow (e.g. Via, 2009; Stolting *et al.*, 2013). In later stages of speciation, introgressed chromosomal blocks with low interspecific differentiation can be identified when populations from hybrid zones are compared with parental species populations outside the zones of sympatry.

Oaks and poplars represent model tree species to study early and later stages of speciation, and maintenance of species identity with gene flow. Thus, in both genera interspecific gene flow is common in contact zones between interfertile, but ecologically divergent species (e.g. Lepais *et al.*, 2009; Lexer *et al.*, 2005). However, oaks and poplars show different signatures of genomic divergence.

Hybridizing oaks show overall very low interspecific genetic differentiation punctuated by genomic regions of high differentiation (Goicoechea *et al.*, 2012; Lind-Riehl *et al.*, 2014; Scotti-Saintagne *et al.*, 2004b), a pattern that is predicted by the models of early stages of ecological speciation in the face of gene flow and strong divergent selection (Via, 2009, 2012; Via and West, 2008). This pattern of genomic divergence allows for screening of loci under divergent selection and linked genomic regions (outlier regions) in hybrid zones between ecologically divergent oak species (Scotti-Saintagne *et al.*, 2004b). Genome scans and the availability of high density linkage maps (Bodénès *et al.*, 2012) and a reference genome sequence in *Quercus robur* (Christophe Plomion, pers. comm.) will enable us to identify the genes that underlie these outlier regions.

In contrast, poplar species show patterns of genomic divergence indicative of later stages of speciation. In these later stages of speciation, genetic drift and selection within species has resulted in divergence of large genomic regions (Via, 2009) making it impossible to distinguish genomic regions involved in initial adaptive divergence between species from genomic regions that are differentiated as the result of genetic drift and within species selection responses. Thus, overall genome-wide genetic differentiation between interfertile hybridizing poplar species is magnitudes higher than in oaks, punctuated by smaller genomic blocks of low divergence (Stolting *et al.*, 2013; Fig. 1). Consequently, transfer of alleles (including adaptive alleles) between

ecologically divergent species can be analyzed when poplar hybrid zones are compared with populations outside the area of sympatry.

Comparison of hybrid classes in seedlings and adult trees of *Populus alba* and *Populus tremula* suggested strong viability selection against many but not all hybrid classes from seedling to adult life stage (Lindtke *et al.*, 2014). These results suggest that introgression of adaptive alleles between species is restricted to a few regions that are permeable to gene flow, but low interspecific differentiation could also be due to shared ancestral variation. The identification of genomic regions with low interspecific differentiation on the *P. trichocarpa* or *P. tremula* reference genomes (Tuskan *et al.*, 2006; *http://loblolly.ucdavis.edu/bipod/ftp/Genome_Data/genome/Pota/*) can help to distinguish between these two scenarios. Thus, blocks of introgressed regions in the other species' genomic background would indicate recent interspecific gene flow (Stolting *et al.*, 2013). Genes that underlie regions of low interspecific differentiation could be identified and tested for their potential adaptive importance using association genetic approaches (see below).

Oaks and poplars provide a model to compare the size and distribution of genomic blocks in early and later stages of speciation. The availability of genomic and experimental resources such as Whole Genome Sequences (WGS), high density linkage maps, Quantitative Trait Locus (QTL) and association populations in the model tree poplar (Evans *et al.*, 2014; Novaes *et al.*, 2009; Tuskan *et al.*, 2006) and in the related oaks (Alberto *et al.*, 2013; Bodénès *et al.*, 2012; Faivre Rampant *et al.*, 2011; Christophe Plomion, pers. comm.) will allow for the identification of genes that underlie outlier genomic regions in oaks and introgressed genomic regions in poplars. Outlier regions are expected to contain genes involved in adaptive divergence and reproductive isolation between species (Via, 2009). Introgressed genomic regions may be selectively neutral or contain genes of adaptive significance. Since both oaks and poplars show considerable within-species genetic variation for traits that discriminate

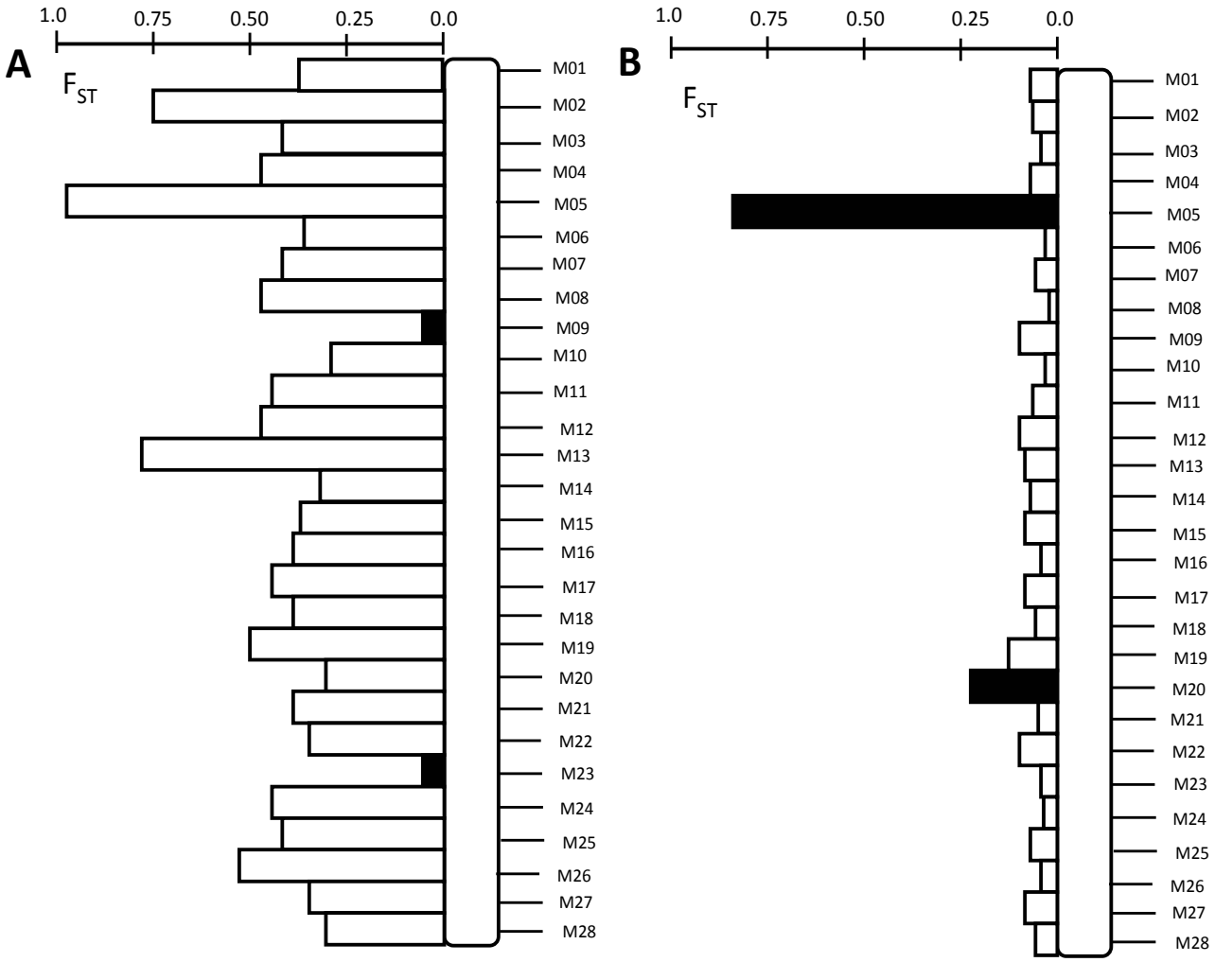

Fig. 1. Schematic representation of interspecific genetic differentiation (F_{ST}) at individual markers across a chromosomal segment in poplars (A) and oaks (B). Interspecific differentiation at most markers is magnitudes higher in hybridizing poplars than in hybridizing oaks (see text for details). Black bars indicate chromosomal regions marked by low interspecific differentiation in poplars and by high interspecific differentiation in oaks. Genome-scans using next generation sequencing could identify blocks of genes in these chromosomal segments that are exchanged by gene flow in poplars and that show signatures of divergent selection in oaks

between hybridizing species, some genes that are related to interspecific trait differences may also show significant associations with trait variation within species as shown for both poplars (Frewen *et al.*, 2000; Ingvarsson *et al.*, 2008) and oaks (Alberto *et al.*, 2013; Lind-Riehl *et al.*, 2014).

Current advances in interspecific genome scans and association genetics, and strategies to characterize outlier and introgressed genomic regions and to assess their effect on adaptive trait variation in oaks and poplars are presented in the present paper.

Hybridization in oaks

Oaks are known for their propensity to hybridize and are sometimes considered as a "worst case scenario for the biological species concept" due to recurrent interspecific gene flow (Coyne and Orr, 2004). On the other hand they provide a model to study ecological speciation and species coherence in the face of gene flow at the genome level (Goicoechea *et al.*, 2012; Lind-Riehl *et al.*, 2014; Scotti-Saintagne *et al.*, 2004b). Despite interspecific gene flow, oaks maintain genetic and morphological differences as well as specific ecological adaptations (Abrams, 1990; Curtu *et al.*, 2007; Kremer *et al.*, 2002) suggesting that strong divergent selection might contribute to species coherence (Curtu *et al.*, 2007; Scotti-Saintagne *et al.*, 2004b).

As an example, the ecologically divergent but interfertile oak species *Q. robur* and *Q. petraea* revealed a mosaic of genomic regions, in which regions of high interspecific differentiation (outlier regions) are surrounded by regions of low differentiation (Goicoechea *et al.*, 2012; Scotti-Saintagne *et al.*, 2004b). This genomic signature is predicted by models of early stages of ecological speciation in the face of gene flow and strong divergent selection (Via, 2009, 2012; Via and West, 2008). Linkage disequilibrium (LD) is expected to be higher in genomic regions under divergent selection than in neutral regions due a reduction in effective interspecific gene flow (Nosil *et al.*, 2009; Via, 2009, 2012; Via and West, 2008). This divergence hitchhiking could allow for the accumulation of alleles involved in reproductive isolation between species by reducing gene exchange around a gene under strong divergent selection (Via, 2009, 2012). Since the first whole genome sequence in *Q. robur* is nearing completion (Christophe Plomion, pers. comm.) and high density genetic linkage maps are currently developed in the European white oak section *Quercus* (Bodénès *et al.*, 2012; Gailing *et al.*, 2013) and in the North American species *Quercus rubra* (section Lobatae) (Jeanne Romero-Severson, pers. comm.), oaks provide an excellent model to identify genes that underlie these outlier genomic regions.

Hybridization in poplars

In contrast to oaks, hybridizing poplar species represent highly divergent species as reflected in the pattern of genomic divergence (Stolting *et al.*, 2013). Hybridization is frequent in contact zones between ecologically differentiated and genetically highly divergent poplar species, but gene exchange is limited by pre- and post-zygotic reproductive barriers, and distinct species boundaries are maintained (e.g. Keim *et al.*, 1989; Roe *et al.*, 2014; Whitham *et al.*, 1999). For example, genome scans at 11,976 Restriction Site Associated DNA (RAD)-sequencing derived loci in a *Populus alba* and *P. tremula* hybrid zone revealed pronounced selection against many, but not all hybrid classes, when adult and seedling generations were compared (Lindtke *et al.*, 2014). Likewise, effective interspecific gene flow in a hybrid zone between *Populus balsamifera* and *Populus deltoides* was much lower in the seedling generation than in the seeds suggesting selection in early life stages as a post-zygotic isolation mechanism (Roe *et al.*, 2014). Pre-zygotic barriers include phenological differences and incompatibilities resulting for example in asymmetric interspecific gene flow (Hamzeh *et al.*, 2007; Keim *et al.*, 1989; Roe *et al.*, 2014).

Mean genetic differentiation as measured by F_{ST} (Weir and Cockerham, 1984) at Simple Sequence Repeat (SSR) and Single Nucleotide Polymorphism (SNP) markers between hybridizing poplar species is magnitudes higher (mean $F_{ST} > 0.30$; Hersch-Green *et al.*, 2014; Stolting *et al.*, 2013) than between hybridizing oaks (mean $F_{ST} < 0.10$; Alberto *et al.*, 2013; Lind and Gailing, 2013; Mariette *et al.*, 2002) (see Fig. 1 for a schematic representation). For example, interspecific F_{ST} for 26 genomic SSR loci in a hybrid zone between *P. deltoides* and *P. angustifolia* ranged from $F_{ST} = 0.003$ to $F_{ST} = 0.96$ with a mean F_{ST} across all markers of 0.47 (Hersch-Green *et al.*, 2014). Likewise, comparatively high interspecific F_{ST} values were found in a contact zone between the hybridizing species *P. alba* and *P. tremula* ($F_{ST} = 0.37$ for SSRs, $F_{ST} = 0.634$ for SNPs; Stolting *et al.*, 2013).

Outlier screening of candidate genes in red oak species (section Lobatae)

Species boundaries in the North American red oaks are often genetically and morphologically ambiguous (Aldrich and Cavender-Bares, 2011; Hipp and Weber, 2008) and the overall level of genetic differentiation at genomic SSRs and genic Expressed Sequence Tag (EST)-SSRs is comparatively low (e.g. < 5% between *Q. rubra* and *Q. ellipsoidalis*; Lind and Gailing, 2013; Sullivan *et al.*, 2013). Genetic assignment analyses at SSR and AFLP markers (Lind and Gailing, 2013; Owusu *et al.*, submitted; Sullivan *et al.*, submitted) and sharing of chloroplast haplotypes among species, but genetic differentiation among geographic regions independent of the species (Zhang *et al.*, in prep.) illustrated recurrent gene flow among species. Yet despite interspecific gene flow, species cluster genetically (Lind and Gailing, 2013) and maintain ecological distinctions such as different soil preferences and adaptations to drought (Abrams, 1990, 1992; Gailing, 2013).

Outlier screens using 36 genic EST-SSRs, that were originally developed in *Q. robur* and *Q. petraea* (Durand *et al.*, 2010), and eight genomic SSRs (Sullivan *et al.*, 2013) in the ecologically divergent species *Q. rubra* (drought averse) and *Q. ellipsoidalis* (drought tolerant) revealed four outlier loci, one of which was identified as under strong divergent selection in all four interspecific population pairs from three geographic regions (Lind-Riehl *et al.*, 2014). This microsatellite, FIR013, was nearly fixed for alternative alleles (138bp in *Q. ellipsoidalis*, 141bp in *Q. rubra*) in the two

species (interspecific F_{ST} = 0.38 – 0.79) in adult and seedling generations (Lind-Riehl *et al.*, 2014; Collins *et al.*, submitted). Heterozygous genotypes were rare in each region, but more frequent in the *Q. ellipsoidalis* populations which might indicate asymmetric gene flow between the more frequent *Q. rubra* and the disjunct *Q. ellipsoidalis* populations (Fig. 2). Likewise homozygotes for the 141bp-allele (the "*Q. rubra*"-allele) were more frequent in *Q. ellipsoidalis* populations than homozygotes for the 138bp-allele (the "*Q. ellipsoidalis*"-allele) in the *Q. rubra* populations (Fig. 2).

The EST from which FIR013 was derived was annotated as *CONSTANS-like 1* (*COL1*), a candidate gene for flowering time (Yano *et al.*, 2000). The tri-nucleotide microsatellite encodes a poly(E) repeat, and *Q. ellipsoidalis* is characterized by the deletion of one glutamine residue (138bp-allele) (Lind-Riehl *et al.*, 2014). Interestingly, a poly(E) repeat allele in *COL2B* was associated with growth

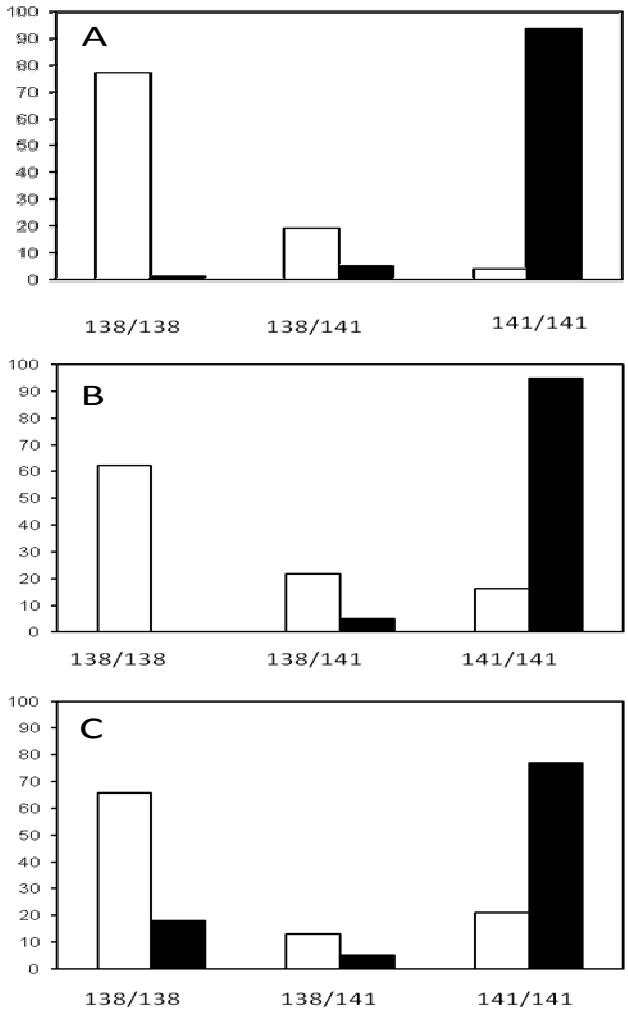

Fig. 2. Frequency (in percent) of genotypes at FIR013 (*COL1*) in neighboring *Quercus ellipsoidalis* (white bars) and *Q. rubra* (black bars) populations from three geographic regions. A: Baraga plains, B. Nicolet National Forest, C: Chequamegon National Forest. Detailed information on outlier screens and population locations is given in Lind-Riehl *et al.*, 2014)

cessation in *Populus tremula* (Ma *et al.*, 2010). Additionally, nucleotide variation in the *COL1* gene that differentiated between *Q. rubra* and *Q. ellipsoidalis* was also significantly associated with the timing of vegetative bud burst in *Q. petraea* (Alberto *et al.*, 2013). *CONSTANS-like* genes are zinc finger transcription factors involved in the photoperiod pathway of floral transition (Amasino, 2005) and are also involved in growth and development (Herrmann *et al.*, 2010; Hsu *et al.*, 2012). For example, a *CONSTANS-like* gene was associated with both height and flowering time in *Medicago sativa* (Herrmann *et al.*, 2010). Thus, these genes may play a role in adaptive divergence between species through both ecological and phenological divergence (Lind-Riehl *et al.*, 2014). Interestingly, *Q. ellipsoidalis* seedlings showed a significantly later, albeit overlapping, bud burst and higher mortality than *Q. rubra* seedlings from neighboring populations in a common garden trial (Gailing, 2013).

QTL and association mapping studies in red oaks could show whether variation in *COL1* is associated with phenology and growth traits. QTL mapping populations are available for *Q. rubra* and a high-density genetic linkage map is currently developed for this species (Jeanne Romero-Severson, pers. comm.). Likewise, range-wide provenance trials for *Q. rubra* with large adaptive trait variation are available for future association analyses (Kriebel *et al.*, 1976; Schlarbaum and Bagley, 1981) to test for an association of nucleotide variation in outlier genes with adaptive trait variation. Outlier loci between species can also be tested for allele frequency clines along environmental gradients within species and for associations with climatic / environmental parameters and adaptive traits (see Alberto *et al.*, 2013; Evans *et al.*, 2014). For example, whole genome selection scans and association analyses revealed a large number of genomic regions that showed both signatures of recent positive or divergent selection among populations and associations with adaptive trait variations in *Populus trichocarpa* (Evan *et al.* 2014). Finally, the allelic variants of *COL1* that differentiate between *Q. rubra* and *Q. ellipsoidalis* could be cloned and expressed in *Arabidopsis* wild type or late and early flowering mutants.

So far we performed outlier screens between *Q. rubra* and *Q. ellipsoidalis* population pairs from contrasting sites (mesic vs. xeric conditions) in the northern range of their sympatric distribution and identified *COL1* consistently as outlier under strong divergent selection across all interspecific comparisons. However, *Q. ellipsoidalis* populations are found to grow under more mesic conditions in some locations at the southern edge of its distribution range (A. Hipp, pers. comm.). Thus, interspecific population pairs with less pronounced differences in site preference and drought tolerance could be identified. The screening of outlier loci (e.g. *COL1*) in interspecific population pairs across different environments and soils could reveal distinct differentiation patterns depending on the steepness of environmental gradients between interspecific population pairs. Thus, species-discriminating genes might be nearly fixed on alternative alleles in highly contrasting environments as observed for *COL1* (Lind-Riehl *et al.*, 2014; Fig. 2), while introgression of adaptive alleles between species could be found in regions with less pronounced micro-environmental

differences between species pairs. First analyses within a narrow geographic range indicated that the frequency of introgressed alleles differed for the three geographic regions with slightly different climatic and edaphic conditions (Fig. 2; Lind-Riehl *et al.*, 2014).

Another outlier screen between the infertile and ecologically divergent species *Quercus velutina* and *Q. ellipsoidalis* identified two gene-based EST-SSRs as outliers, GOT040 and POR016 (Sullivan *et al.*, submitted). Both markers mapped on the same linkage group bin on linkage group 6 (≤ 15cM apart) of *Q. robur* (Durand *et al.*, 2010). Genetic variation and number of alleles at both markers were strongly reduced in all five *Q. ellipsoidalis* populations across the Great Lakes region indicative of a recent selective sweep (Sullivan *et al.*, submitted). GOT040 was annotated as a 40s ribosomal 16s-like protein and POR016 as serine/threonine-protein phosphatase 5-like protein. These outlier genes provide additional candidates for QTL and comparative mapping analyses as well as association genetic studies.

Genome scans in hybridizing poplars

Poplars can provide a model to identify genomic regions that are shared between hybridizing species as the result of recurrent interspecific gene flow (Stolting *et al.*, 2013). However, in addition to recurrent interspecific gene flow, allele sharing between species can be due to homoplasy and shared ancestral polymorphisms between species as the result of similar selection pressures in each species and/or insufficient time since species divergence (Stolting *et al.*, 2013). A genome scan in a *P. tremula* and *P. alba* hybrid zone revealed a genomic region of low interspecific differentiation of three megabases on chromosome 14, an incipient sex chromosome in poplar, possibly as result of interspecific gene flow (Stolting *et al.*, 2013). The genes underlying this region of low genetic divergence are not reported.

In silico mapping of loci with low interspecific differentiation to the *P. trichocarpa* reference genome (Tuskan *et al.*, 2006) or to the *P. tremula* draft genome assembly[1] can reveal blocks of low divergence on chromosomes indicative of recent interspecific gene flow (Scotti-Saintagne *et al.*, 2004b; Stolting *et al.*, 2013). Additionally, comparative genome scans in populations of parental species outside the area of sympatry and in hybrid zones could distinguish genomic regions that are shared between species as result of recurrent gene flow in hybrid zones from shared ancestral polymorphisms. However, it is unclear whether genomic regions that are exchanged among ecologically divergent poplar species contain adaptive alleles. The characterization of genomic regions with low interspecific differentiation in multiple hybrid zones and in control regions (non-hybrid zones outside the area of sympatry) from different environments could help to reveal whether introgression of alleles is dependent on environmental conditions. Expressional candidate genes of adaptive significance could be identified in these genomic regions based on the literature and publicly available microarray data for *Populus* (e.g. Poplar-Plant Expression data base[2], or Gene Expression Omnibus[3]), and comparative expression analyses could be performed in species and hybrids from species contact zones and from "pure" species reference populations.

A combined approach of genome scans for introgressed genomic regions and association analyses could reveal whether genes in these genomic regions are associated with adaptive trait variation. In a similar approach, whole genome outlier screens and association analyses for ecologically divergent populations of *P. trichocarpa* revealed outlier genes under positive or recent divergent selection that at the same time showed significant associations with adaptive traits (Evans *et al.*, 2014).

Comparative outlier regions across oak sections

Comparative outlier analyses across taxonomic groups can reveal genomic divergence unique to species pairs and parallel genomic divergence driven by natural selection as shown recently in an animal model (Soria-Carrasco *et al.*, 2014). The first outlier screens in oaks have been performed between the two European white oak species *Q. robur* and *Q. petraea*. Both species are interfertile and co-occur in most European forests, but have different soil preferences. Thus, *Q. petraea* prefers drier soils, while *Q. robur* is more frequent on nutrient rich soils which are temporally subjected to flooding (Breda *et al.*, 1993; Levy *et al.*, 1992; Zanetto *et al.*, 1994). A genome scan using AFLPs, SSRs and selected gene markers showed a mosaic of regions with low interspecific differentiation interspersed by regions with signatures of divergent selection and a hotspot of interspecific differentiation on linkage group 12 (Scotti-Saintagne *et al.*, 2004b). One SSR marker, QrZAG112, has been identified as an outlier between the two European white oak species *Q. robur* and *Q. petraea* (section *Quercus*) (Goicoechea *et al.*, 2012; Scotti-Saintagne *et al.*, 2004b) and between the species pair *Quercus alnifolia* and *Quercus coccifera* from section *Cerris* (Neophytou *et al.*, 2011). Likewise, the *COL1* gene as a strong outlier between *Q. rubra* and *Q. ellipsoidalis* (see above) showed elevated levels of interspecific differentiation between *Q. robur* and *Q. pedunculiflora* K. Koch (Curtu *et al.*, in prep.), two closely related species with different adaptations to drought (Enescu, 1993; Gailing and Curtu, 2014).

Comparative outlier screenings in different oak sections can further our understanding of how divergent selection drives species divergence and reproductive isolation between species. Next generation sequencing such as Restriction Site Associated DNA (RAD) sequencing (Baird *et al.*, 2008) can reveal genome-wide patterns of interspecific divergence (Keller *et al.*, 2013; Stolting *et al.*, 2013). For example, RAD libraries enriched for gene sequences can be sequenced in multiple population pairs to identify loci that are involved in adaptive species divergence. When RAD-seq data are anchored to the *Q.*

[1] *http://loblolly.ucdavis.edu/bipod/ftp/Genome_ Data/genome/Pota/*

[2] *http://www.plexdb.org/plex.php?database=Poplar*

[3] *http://www.ncbi.nlm.nih.gov/geo/*

robur whole genome sequence (WGS), genomic regions with signatures of divergent selection shared across taxonomic groups can be identified. Additionally, comparison of multiple interfertile taxa within sections (e.g. within section *Quercus* and *Lobatae*) with different divergence times could reveal the sequential accumulation of barrier loci (loci under divergent selection and/or involved in reproductive isolation between species) in the genome. For example, within section *Lobatae* the distribution range of the most drought tolerant red oak species *Q. ellipsoidalis* is largely overlapping with the range of *Q. velutina* and with that of *Q. rubra* as the most mesophytic species. Genetic assignment analysis at SSRs (Sullivan *et al.*, submitted) and AFLPs (Hipp and Weber, 2008; Sullivan *et al.*, submitted) suggested more frequent gene flow between *Q. ellipsoidalis* and the closely related species *Q. velutina* than between *Q. ellipsoidalis* and the more distantly related *Q. rubra*. Outlier screens at a limited number of genic and genomic SSRs between *Q. ellipsoidalis* on the one hand and *Q. velutina* and *Q. rubra* on the other hand revealed outliers unique to each species pair but also two genic outliers, GOT040 and POR016, common to both pairs (Sullivan *et al.*, submitted).

Quantitative Trait Locus (QTL) mapping for adaptive traits in oaks

Co-location of outlier loci between species with Quantitative Trait Loci (QTL) for traits that are related to adaptive species differences provides additional evidence for the role of outlier loci in adaptive species divergence.

Quercus robur and *Q. petraea* co-occur in mixed stands across Europe, but show different ecological requirements. Morphologically, both species can be distinguished based on leaf morphological characters (Aas, 1993; Curtu *et al.*, 2007; Kremer *et al.*, 2002) and characters of the cupulae (Aas, 1993). While most samples can be assigned to the one or other morphological species using multivariate statistics, both species show a wide variation in traits that are used to distinguish between them (e.g. Curtu *et al.*, 2007; Kremer *et al.*, 2002). QTL for species-discriminating leaf traits have been characterized in intraspecific *Q. robur* crosses and, similar to outlier loci, were distributed across most of the 12 oak linkage groups (Gailing, 2008; Gailing *et al.*, 2013; Saintagne *et al.*, 2004; Scotti-Saintagne *et al.*, 2004a). Interestingly, one interspecific outlier locus, QrZAG96, co-located with a QTL for leaf morphological variation (petiole ratio and lobe-width ratio) on linkage group 10 in a *Q. robur* x *Q. robur* full-sib family (Saintagne *et al.*, 2004). One of the leaf traits, petiole ratio, showed very high interspecific differentiation (84%) in a mixed *Q. petraea* / *Q. robur* population (Saintagne *et al.*, 2004).

Likewise, *Q. robur* is less drought tolerant than *Q. petraea* and showed a lower water use efficiency (WUE) in adult trees that co-occurred in sympatric stands (Ponton *et al.*, 2001) as well as in seedling common garden experiments (Ponton *et al.*, 2002). A major QTL that explained more than 20% of the phenotypic variation in WUE was mapped on linkage group 11 in the *Q. robur* full-sib progeny in three consecutive years (Brendel *et al.*, 2008) suggesting considerable within-species genetic variation for this species-discriminating trait.

Quercus robur is more tolerant to waterlogging than *Q. petraea* (Parelle *et al.*, 2006; Parelle *et al.*, 2007a), and QTL for responses to waterlogging were detected on four linkage groups explaining each a moderate percentage of the total phenotypic variance in a *Q. robur* full-sib family (Parelle *et al.*, 2007b).

High density genetic linkage maps are now becoming available for F_1 interspecific *Q. robur* and *Q. petraea* and for *Q. petraea* full-sib families (Bodénès *et al.*, 2012) which will allow for a comparison of QTL positions between species. Backcrosses of F_1 interspecific hybrids with one or both of the parental species would be desirable to identify QTL that are associated with species differences in morphological and physiological traits.

A QTL mapping population is also available in *Q. rubra,* and the construction of a high-density genetic linkage map based on RAD tags and SSRs is underway (Jeanne Romero-Severson pers. comm.). Full-sib families are not available for the drought tolerant *Q. ellipsoidalis*, but single tree progeny have been obtained from *Q. ellipsoidalis* and from phenotypic intermediate individuals (putative *Q. rubra* x *Q. ellispoidalis* hybrids). Genotyping of open pollinated progeny is planned to identify full-sibs for genetic linkage and QTL mapping.

In order to validate the role of outlier loci in adaptive species differences, genetic and *in silico* mapping of outlier loci on genetic linkage maps can identify co-locations of outliers with QTL for adaptive and species differentiating traits. To narrow down the genomic region to individual genes, association mapping in unstructured populations of all genes in outlier regions in addition to reference loci should be performed. In the future, genome-wide outlier screens, and genetic linkage and association mapping using the same marker technique (e.g. single or double-digest RAD markers) can identify genes under divergent selection that are associated with adaptive species differences such as WUE or tolerance to waterlogging (root hypoxia).

Also, admixture mapping in natural hybrid zones was proposed as a method for the identification of QTL for traits that differ between hybridizing poplar species (Lindtke et al., 2013). The simultaneous screening for significant marker-phenotype associations (QTL) and outlier screens in oak hybrid zones may represent a way to identify candidate gene regions under divergent selection that at the same time are associated with adaptive trait differences. These genomic regions could be further validated in independent QTL and association populations.

QTL mapping for adaptive traits in poplars

QTL mapping studies for a variety of characters including growth and phenology traits have been conducted in intra- and interspecific poplar crosses (Dillen *et al.*, 2009; Fabbrini *et al.*, 2012). QTL for species differences can be mapped in interspecific pseudo-backcross pedigrees that segregate for traits that differentiate between parental species. For example, QTL mapping for 20 wood chemistry and growth traits under high and low nitrogen conditions in an interspecific pseudo-backcross pedigree of *P. trichocarpa* and *P. deltoides* (*P. trichocarpa* (clone 93-968) x *P. deltoides* (clone ILL-101) crossed with *P. deltoides* clone D124) identified a total of 63 QTL that explained from 3.58% to 11.36% of the

phenotypic variance. The *P. deltoides* allele was associated with the higher trait values in 29 QTL, the *P. trichocarpa* allele in 34 QTL (Novaes *et al.*, 2009). With the availability of the *P. trichocarpa* reference genome sequence, QTL regions can be anchored to the genome sequence and underlying genes can be identified (Muchero *et al.*, 2013; Novaes *et al.*, 2009).

In interspecific hybrid zones, full-sib pedigrees could be generated by crossing F_1 hybrids with both parental species to perform QTL analyses for adaptive species differences. However, the generation of full-sib families in these long-lived forest trees is time- and labor-intensive.

As a potential alternative, admixture mapping in hybrid zones holds promise for the identification of chromosomal regions that are associated with species differences (Lexer *et al.*, 2007; Lindtke *et al.*, 2013). Admixture mapping across multiple hybrid zones could identify stable marker-phenotype associations that could be further validated in segregating full-sib families and association populations. Thus, admixture mapping in an Italian hybrid zone between *P. tremula* and *P. alba* (n =219) for species discriminating leaf traits revealed polygenic inheritance with significant marker-phenotype associations on several chromosomes, each explaining between 2.3% to 18.2% of the phenotypic variance (mean phenotypic variance explained: 8.6%). Pooling of all samples from four hybrid zones including the Italian one identified mainly different QTL, but four markers were consistently associated with the same leaf traits in both data sets (Lindtke *et al.*, 2013).

Genome-wide scans for marker-phenotype associations and for interspecific genetic differentiation in the same hybrid zones could provide an effective means to identify candidate genes that are involved in adaptive species differences, and to compare interspecific differentiation at these genes in hybrid zones and in populations outside the area of sympatry. Given the large within-species variation for most species-discriminating characters, association mapping in provenance trials of one species could also be used to test for marker-phenotype associations (Frewen *et al.*, 2000; Ingvarsson *et al.*, 2008).

Association mapping for adaptive traits in oaks

In outcrossing species like oaks, association mapping in provenance trials can narrow down QTL regions to individual genes (Neale and Savolainen, 2004). Association and outlier approaches can complement each other to identify genes under selection. For example, association and outlier analyses were performed at 106 candidate genes for bud burst in a *Q. petraea* common garden experiment consisting of 758 individuals from 32 natural populations sampled along latitudinal and altitudinal gradients (Alberto *et al.*, 2013). Outlier SNPs were identified in 15 genes, clinal variation was found for six SNPs in six genes, and associations with phenotypic or breeding values were observed in 14 SNPs from 12 genes. Out of the 14 SNPs identified under divergent selection, six were also correlated with geographic and climatic variables. Only one SNP in the Early light-induced protein1 (*ELIP1*) was identified as under divergent selection (for the latitudinal gradient) and was at the same time associated with the timing of bud burst (for the altitudinal gradient) (PVE = 6.6%). SNPs at three additional genes, Ribosomal protein (*L18a*), Aquaporin (*PIP1*) and Unknown protein 3 (*UNK3*) were identified as outliers under divergent selection, while closely linked SNPs at the same genes were associated with the timing of bud burst (3.3% - 8.8% PVE). Interestingly, nucleotide variation in the 3'-UTR of a major outlier locus between the red oak species *Q. rubra* and *Q. ellipsoidalis*, *COL1* (Lind-Riehl *et al.*, 2014), was also associated with bud burst in the *Q. petraea* provenance trial explaining 4.6% of the phenotypic variation (Alberto *et al.*, 2013).

Association mapping for adaptive traits in poplars

Whole genome selection scans and association analyses across a wide latitudinal range in *P. trichocarpa* identified 397 genomic regions with signatures of recent positive or divergent selection. These regions showed an overrepresentation of markers that were significantly associated with adaptive traits such as bud burst, bud set and height (Evans *et al.*, 2014). For example *FT1*, a gene involved in the maintenance of dormancy, was identified as an F_{ST} outlier under divergent selection and at the same time was significantly associated with bud burst at three different common garden sites.

Association studies of bud burst and bud set in a *P. tremula* common garden experiment identified phytochrome B2 (*phyB2*) as outlier with clinal variation across a latitudinal gradient (Ingvarsson *et al.*, 2006), and two non-synonymous SNPs in *phyB2* were associated with the timing of bud set in the same common garden trial (Ingvarsson *et al.*, 2008). *PhyB2* also co-located with QTL for bud burst and bud set in *P. deltoides* (southern latitude, 31°N) and *P. trichocarpa* (northern latitude, 48° N) interspecific crosses (Frewen *et al.*, 2000). Other association studies in *P. trichocarpa* identified significant marker-phenotype association for 16 wood chemistry traits. Individual markers explained a moderate level of the total phenotypic variation for each trait (3% -7%) (Porth *et al.*, 2013).

Selection against migrant alleles and allele combinations (reciprocal transplant experiments)

Selection against hybrids and migrants, and migrant alleles and allele combinations can be validated in reciprocal transplant experiments between both parental environments and between parental and intermediate (hybrid) environments. Especially fitness effects of outlier alleles (e.g. of *COL1*) could be evaluated in both parental and intermediate environments across multiple population pairs. However, migrant alleles will be rare in outliers that are nearly fixed on alternative alleles as it is the case for *COL1* in *Q. rubra* and *Q. ellipsoidalis* (i.e. the 141bp-allele is rare in *Q. ellipsoidalis*, the 138bp-allele is rare in *Q. rubra*) (Fig. 2; Lind-Riehl *et al.*, 2014). However, homozygotes and heterozygotes for migrant alleles might be more frequent in direct contact zones between species. Two of such sympatric *Q. rubra* / *Q. ellipsoidalis* stands have recently been identified in two geographic regions and genetic assignment analyses of adult trees and seeds are underway seed/seedling generations.

Comparisons of hybrid frequencies between adult and seedling generations provide indirect evidence for selection against hybrids in oaks (Curtu *et al.*, 2009) and poplar (Lindtke *et al.*, 2014), but direct evidence of selection against hybrids and migrant alleles from reciprocal transplant experiments is missing in both species.

Conclusions

Combined genome scans for outliers in hybrid zones, and association genetic approaches are necessary to identify genomic regions and genes that are involved in adaptive divergence and reproductive isolation between species (outlier regions). Oaks show a pattern of genomic divergence that makes them models for the identification of outlier regions and of individual genes that are associated with adaptive species differences.

Genome scans in genetically divergent and hybridizing poplar species can identify genomic regions that are exchanged between species. Underlying genes can be assessed for marker-phenotype associations in association populations. The availability of genomic resources and decreasing costs for sequencing will enable us to perform genome scans in different hybrids zones of several hybridizing species in both oaks and poplars. These comparative genome scan could identify patterns of genomic divergence which is unique to species pairs and patterns of parallel divergence driven by natural selection (Soria-Carrasco *et al.*, 2014) within both genera. Selection against migrant alleles can be tested in reciprocal transplant experiments between parental environments and between parental and hybrid environments.

Acknowledgements

Funding for the analyses in North American red oaks and the development of genomic resources was provided by the NSF Plant Genome Research program (NSF 1025974), the Hanes Trust, the USDA McIntire Stennis fund (1309051) and the Huron Mountain Wildlife foundation. Additional support was provided by the NSF program Sustainable Energy Pathways (1202012), the Biotech Research Center and by the Ecosystem Science Center of Michigan Technological University. I am also thankful to Jennifer Lind-Riehl who provided the published genotypic data for FIR013. This paper was presented at the International Symposium Forest and Sustainable Development held in Brasov between 24 and 25 October 2014.

References

Aas G (1993). Taxonomical impact of morphological variation in *Quercus robur* and *Q. petraea*: a contribution to the hybrid controversy. Ann For Sci 50(Suppl 1):107s-113s.

Abrams MD (1990). Adaptations and responses to drought in *Quercus* species of North America. Tree Physiol 7:227-238.

Abrams MD (1992). Fire and the development of oak forests in eastern North America, oak distribution reflects a variety of ecological paths and disturbance conditions. Bioscience 42:346-353.

Alberto FJ, Derory J, Boury C, Frigerio JM, Zimmermann NE, Kremer A (2013). Imprints of natural selection along environmental gradients in phenology-related genes of *Quercus petraea*. Genetics 195:495-512.

Aldrich P, Cavender-Bares J (2011). Quercus, p. 89-129. In: Kole C (ed). Wild crop relatives: Genomic and breeding resources, forest trees. Springer, Berlin.

Amasino RM (2005). Vernalization and flowering time. Curr Opin Biotech 16:154-158.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. Plos One 3(10), e3376.

Bodénès C, Chancerel E, Gailing O, Vendramin GG, Bagnoli F, Durand J, Goicoechea PG, Soliani C, Villani F, Mattioni C, Koelewijn HP, Murat F, Salse J, Roussel G, Boury C, Alberto F, Kremer A, Plomion C (2012). Comparative mapping in the Fagaceae and beyond with EST-SSRs. BMC Plant Biol 12(1):153.

Breda N, Cochard H, Dreyer E, Granier A (1993). Field comparison of transpiration, stomatal conductance and vulnerability to cavitation of *Quercus petraea* and *Q. robur* under water-stress. Ann Sci Forest 50:571-582.

Brendel O, Le Thiec D, Scotti-Saintagne C, Bodenes C, Kremer A, Guehl J-M (2008). Quantitative Trait loci controlling water use efficiency and related traits in *Quercus robur* L. Tree Genet Genomes 4:263-278.

Coyne JA, Orr HA (2004). Speciation. Sinauer Associates, Sunderland.

Curtu AL, Gailing O, Finkeldey R (2007). Evidence for hybridization and introgression within a species-rich oak (*Quercus* spp.) community. BMC Evol Biol 7:218.

Curtu AL, Gailing O, Finkeldey R (2009). Patterns of contemporary hybridization inferred from paternity analysis in a four-oak-species forest. BMC Evol Biol 9:284.

Dillen SY, Storme V, Marron N, Bastien C, Neyrinck S, Steenackers M, Ceulemans R, Boerjan W (2009). Genomic regions involved in productivity of two interspecific poplar families in Europe. 1. Stem height, circumference and volume. Tree Genet Genomes 5:147-164.

Durand J, Bodénès C, Chancerel E, Frigero J-M, Vendramin GG, Sebastiani F, Buonamici A, Gailing O, Koelewijn H-P, Villani F, Mattioni C, Cherubini M, Goicoechea PG, Herran A, Ikaran Z, Cabane C, Ueno S, de Daruvar A, Kremer A, Plomion C (2010). A fast and cost-effective approach to develop and map EST-SSR markers: oak as a case study. BMC Genomics 11:570.

Enescu V (1993). A test of half-sib progenies of greyish oak, *Quercus pedunculiflora* K Koch. Ann For Sci 50(Suppl. 1):439s -443s.

Evans LM, Slavov GT, Rodgers-Melnick E, Martin J, Ranjan P, Muchero W, Brunner AM, Schackwitz W, Gunter L, Chen J-

G, Tuskan G, Difazio S (2014). Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. Nat Genet 46:1089-1096.

Fabbrini F, Gaudet M, Bastien C, Zaina G, Harfouche A, Beritognolo I, Marron N, Morgante M, Scarascia-Mugnozza G, Sabatti M (2012). Phenotypic plasticity, QTL mapping and genomic characterization of bud set in black poplar. BMC Plant Biol 12(1):47.

Faivre Rampant P, Lesur I, Boussardon C, Bitton F, Martin-Magniette M, Bodenes C, Le Provost G, Berges H, Fluch S, Kremer A, Plomion C (2011). Analysis of BAC end sequences in oak, a keystone forest tree species, providing insight into the composition of its genome. BMC Genomics 12:292.

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

Gailing O (2008). QTL analysis of leaf morphological characters in a *Quercus robur* full-sib family (*Q. robur* x *Q. robur* ssp. *slavonica*). Plant Biol 10:624-634.

Gailing O (2013). Differences in growth, survival and phenology in *Quercus rubra* and *Q. ellipsoidalis* seedlings. Dendrobiology 70:71-79.

Gailing O, Bodenes C, Finkeldey R, Kremer A, Plomion C (2013). Genetic mapping of EST-derived Simple Sequence Repeats (EST-SSRs) to identify QTL for leaf morphological characters in a *Quercus robur* full-sib family. Tree Genet Genomes 9:1361-1367.

Gailing O, Curtu AL (2014). Interspecific gene flow and maintenance of species integrity in oaks. Ann Forest Res 57:5-18.

Goicoechea PG, Petit RJ, Kremer A (2012). Detecting the footprints of divergent selection in oaks with linked markers. Heredity 109:361-371.

Hamzeh M, Sawchyn C, Perinet P, Dayanandan S (2007). Asymmetrical natural hybridization between *Populus deltoides* and *P. balsamifera* (Salicaceae). Can J Botany 85:1227-1232.

Herrmann D, Barre P, Santoni S, Julier B (2010). Association of a CONSTANS-LIKE gene to flowering and height in autotetraploid alfalfa. Theor Appl Gen 121:865-876.

Hersch-Green EI, Allan GJ, Whitham TG (2014). Genetic analysis of admixture and patterns of introgression in foundation cottonwood trees (Salicaceae) in southwestern Colorado, USA. Tree Genet Genomes 10:527-539.

Hipp AL, Weber JA (2008). Taxonomy of Hill's oak (*Quercus ellipsoidalis*: Fagaceae): Evidence from AFLP data. Syst Bot 33:148-158.

Hsu CY, Adams JP, No K, Liang HY, Meilan R, Pechanova O, Barakat A, Carlson JE, Page GP, Yuceer C (2012). Overexpression of Constans homologs CO1 and CO2 fails to alter normal reproductive onset and fall bud set in woody perennial poplar. Plos One 7(9), e45448.

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

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

Keim P, Paige KN, Whitham TG, Lark KG (1989). Genetic analysis of an interspecific hybrid swarm of *Populus* - occurence of unidirectional introgression. Genetics 123:557-565.

Keller I, Wagner CE, Greuter L, Mwaiko S, Selz OM, Sivasundar A, Wittwer S, Seehausen O (2013). Population genomic signatures of divergent adaptation, gene flow and hybrid speciation in the rapid radiation of Lake Victoria cichlid fishes. Mol Ecol 22:2848-2863.

Kremer A, Dupouey JL, Deans JD, Cottrell J, Csaikl U, Finkeldey R, Espinel S, Jensen J, Kleinschmit J, Van Dam B, Ducousso A, Forrest I, Lopez de Heredia U, Lowe AJ, Tutkova M, Munro RC, Steinhoff S, Badeau V (2002). Leaf morphological differentiation between *Quercus robur* and *Quercus petraea* is stable across western European mixed oak stands. Ann For Sci 59:777-787.

Kriebel HB, Bagley WT, Deneke FJ, Funsch RW, Roth P, Jokela JJ, Merritt C, Wright JW, Williams RD (1976). Geographic variation in *Quercus rubra* in North Central United States plantations. Silvae Genet 25:118-122.

Lepais O, Petit RJ, Guichoux E, Lavabre JE, Alberto F, Kremer A, Gerber S (2009). Species relative abundance and direction of introgression in oaks. Mol Ecol 18:2228-2242.

Levy G, Becker M, Duhamel D (1992). A comparison of the ecology of pedunculate and sessile oaks - radial growth in the center and northwest of France. Forest Ecol Manag 55:51-63.

Lexer C, Buerkle CA, Joseph JA, Heinze B, Fay MF (2007). Admixture in European *Populus* hybrid zones makes feasible the mapping of loci that contribute to reproductive isolation and trait differences. Heredity 98:74-84.

Lexer C, Fay MF, Joseph JA, Nica MS, Heinze B (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. Mol Ecol 14:1045-1057.

Lind-Riehl J, Sullivan A, Gailing O (2014). Evidence for selection on a CONSTANS-like gene between two red oak species. Annals of Botany 113:967-975.

Lind J, Gailing O (2013). Genetic structure of *Quercus rubra* L. and *Q. ellipsoidalis* E. J. Hill populations at gene-based

EST-SSR and nuclear SSR markers. Tree Genet Genomes 31:231-239.

Lindtke D, Gompert Z, Lexer C, Buerkle CA (2014). Unexpected ancestry of *Populus* seedlings from a hybrid zone implies a large role for postzygotic selection in the maintenance of species. Mol Ecol 23:4316-4330.

Lindtke D, Gonzalez-Martinez SC, Macaya-Sanz D, Lexer C (2013). Admixture mapping of quantitative traits in *Populus* hybrid zones: power and limitations. Heredity 111:474-485.

Ma XF, Hall D, St Onge KR, Jansson S, Ingvarsson PK (2010). Genetic differentiation, clinal variation and phenotypic associations with growth cessation across the *Populus tremula* photoperiodic pathway. Genetics 186:1033-1044.

Mariette S, Cottrell J, Csaikl UM, Goikoechea P, König AO, Lowe AJ, Van Dam BC, Barreneche T, Bodenes C, Streiff R, Burg K, Groppe K, Munro RC, Tabbener H, Kremer A (2002). Comparison of levels of genetic diversity detected with AFLP and microsatellite markers within and among mixed *Q. petraea* (Matt.) Liebl. and *Q. robur* L. stands. Silvae Genet 51:72-79.

Muchero W, Sewell MM, Ranjan P, Gunter LE, Tschaplinski TJ, Yin TM, Tuskan GA (2013). Genome anchored QTLs for biomass productivity in hybrid *Populus* grown under contrasting environments. Plos One 8(1), e54468.

Neale DB, Savolainen O (2004). Association genetics of complex traits in conifers. Trends Plant Sci 9:325-330.

Neophytou C, Dounavi A, Fink S, Aravanopoulos FA (2011). Interfertile oaks in an island environment: I. High nuclear genetic differentiation and high degree of chloroplast DNA sharing between *Q. alnifolia* and *Q. coccifera* in Cyprus. A multipopulation study. Eur J For Res 130:543-555.

Nosil P, Funk DJ, Ortiz-Barrientos D (2009). Divergent selection and heterogeneous genomic divergence. Mol Ecol 18:375-402.

Novaes E, Osorio L, Drost DR, Miles BL, Boaventura-Novaes CRD, Benedict C, Dervinis C, Yu Q, Sykes R, Davis M, Martin TA, Peter GF, Kirst M (2009). Quantitative genetic analysis of biomass and wood chemistry of *Populus* under different nitrogen levels. New Phyt 182:878-890.

Parelle J, Brendel O, Bodenes C, Berveiller D, Dizengremel P, Jolivet Y, Dreyer E (2006). Differences in morphological and physiological responses to water-logging between two sympatric oak species (*Quercus petraea* Matt. Liebl., *Quercus robur* L.). Ann For Sci 63:849-859.

Parelle J, Brendel O, Jolivet Y, Dreyer E (2007a). Intra- and interspecific diversity in the response to waterlogging of two co-occurring white oak species (*Quercus robur* and *Q. petraea*). Tree Physiol 27:1027-1034.

Parelle J, Zapater M, Scotti-Saintagne C, Kremer A, Jolivet Y, Dreyer E, Brendel O (2007b). Quantitative trait loci of tolerance to waterlogging in a European oak (*Quercus robur* L.): physiological relevance and temporal effect patterns. Plant Cell Environ 30:422-434.

Ponton S, Dupouey JL, Breda N, Dreyer E (2002). Comparison of water-use efficiency of seedlings from two sympatric oak species: genotype x environment interactions. Tree Physiol 22:413-422.

Ponton S, Dupouey JL, Breda N, Feuillat F, Bodenes C, Dreyer E (2001). Carbon isotope discrimination and wood anatomy variations in mixed stands of *Quercus robur* and *Quercus petraea*. Plant Cell Environ 24:861-868.

Porth I, Klapste J, Skyba O, Hannemann J, McKown AD, Guy RD, DiFazio SP, Muchero W, Ranjan P, Tuskan GA, Friedmann MC, Ehlting J, Cronk QCB, El-Kassaby YA, Douglas CJ, Mansfield SD (2013). Genome-wide association mapping for wood characteristics in *Populus* identifies an array of candidate single nucleotide polymorphisms. New Phyt 200:710-726.

Roe AD, MacQuarrie CJK, Gros-Louis MC, Simpson JD, Lamarche J, Beardmore T, Thompson SL, Tanguay P, Isabel N (2014). Fitness dynamics within a poplar hybrid zone: I. Prezygotic and postzygotic barriers impacting a native poplar hybrid stand. Ecology and Evolution 4:1629-1647.

Saintagne C, Bodenes C, Barreneche T, Pot D, Plomion C, Kremer A (2004). Distribution of genomic regions differentiating oak species assessed by QTL detection. Heredity 92:20-30.

Schlarbaum SE, Bagley WT (1981). Intraspecific genetic variation of *Quercus rubra* L., Northern red oak. Silvae Genet 30:50-56.

Scotti-Saintagne C, Bodenes C, Barreneche T, Bertocchi E, Plomion C, Kremer A (2004a). Detection of Quantitative Trait Loci controlling bud burst and height growth in *Quercus robur* L. Theor Appl Gen 109:1648-1659.

Scotti-Saintagne C, Mariette S, Porth I, Goicoechea PG, Barreneche T, Bodenes C, Burg K, Kremer A (2004b). Genome scanning of interspecific differentiation between two closely related oak species (*Quercus robur* L. and *Q. petraea* (Matt.) Liebl.). Genetics 168:1615-1626.

Soria-Carrasco V, Gompert Z, Comeault AA, Farkas TE, Parchman TL, Johnston JS, Buerkle CA, Feder JL, Bast J, Schwander T (2014). Stick insect genomes reveal natural selection's role in parallel speciation. Science 344:738-742.

Stolting KN, Nipper R, Lindtke D, Caseys C, Waeber S, Castiglione S, Lexer C (2013). Genomic scan for single nucleotide polymorphisms reveals patterns of divergence and gene flow between ecologically divergent species. Mol Ecol 22:842-855.

Sullivan A, Lind J, McCleary TS, Romero-Severson J, Gailing O (2013). Development and characterization of genomic and gene-based microsatellite markers in North American red oak species. Plant Mol Biol Rep 31:231-239.

Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, et al. B. (2006). The genome of black

cottonwood, *Populus trichocarpa* (Torr. & Gray). Science, 313(5793):1596-1604.

Via S (2009). Natural selection in action during speciation. P Natl Acad Sci USA 106:9939-9946.

Via S (2012). Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. Philos T Roy Soc B 367:451-460.

Via S, West J (2008). The genetic mosaic suggests a new role for hitchhiking in ecological speciation. Mol Ecol 17:4334-4345.

Weir BS, Cockerham CC (1984). Estimating F-statistics for the analysis of population structure. Evolution 38:1358-1370.

Whitham TG, Martinsen GD, Floate KD, Dungey HS, Potts BM, Keim P (1999). Plant hybrid zones affect biodiversity: Tools for a genetic-based understanding of community structure. Ecology 80:416-428.

Yano M, Katayose Y, Ashikari M, Yamanouchi U, Monna L, Fuse T, Baba T, Yamamoto K, Umehara Y, Nagamura Y, Sasaki T (2000). Hd1, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene CONSTANS. Plant Cell 12:2473-2483.

Zanetto A, Roussel G, Kremer A (1994). Geographic variation of inter-specific differentiation between *Quercus robur* L. and *Quercus petraea* (Matt.) Liebl. Forest Genetics 1:111-123.