American Genetic Association

Journal of Heredity, 2018, 47–58
doi:10.1093/jhered/esx098
Symposium Article
Advance Access publication 8 November 2017

OXFORD

Symposium Article

Influence of Range Position on Locally Adaptive Gene–Environment Associations in Populus Flowering Time Genes

Stephen R. Keller, Vikram E. Chhatre, and Matthew C. Fitzpatrick

From the Department of Plant Biology, University of Vermont, Burlington, VT 05405 (Keller and Chhatre); and the Appalachian Laboratory, University of Maryland Center for Environmental Science, Frostburg, MD (Fitzpatrick). Vikram E. Chhatre is now at the Wyoming INBRE Bioinformatics Core, Department of Molecular Biology, University of Wyoming, Laramie, WY 82071.

Address correspondence to S. R. Keller at the address above, or e-mail: srkeller@uvm.edu

Received March 11, 2017; First decision May 18, 2017; Accepted November 4, 2017.

Corresponding Editor: Lynda Delph

Abstract

Local adaptation is pervasive in forest trees, which are characterized by large effective population sizes spanning broad climatic gradients. In addition to having relatively contiguous populations, many species also form isolated populations along the rear edge of their range. These rear-edge populations may contain unique adaptive diversity reflecting a history of selection in marginal environments. Thus, discovering genomic regions conferring local adaptation in rear edge populations is a key priority for landscape genomics to ensure conservation of genetic resources under climate change. Here, we report on adaptive gene–environment associations in single nucleotide polymorphisms (SNPs) from 27 genes in the *Populus* flowering time gene network, analyzed on a range-wide collection of >1000 balsam poplar trees, including dense sampling of the southern range edge. We use a combined approach of local adaptation scans to identify candidate SNPs, followed by modeling the compositional turnover of adaptive SNPs along multivariate climate gradients using gradient forests (GF). Flowering time candidate genes contained extensive evidence of climate adaptation, namely outlier population structure and gene–environment associations, along with allele frequency divergence between the core and edge of the range. GF showed strong allele frequency turnover along gradients of elevation and diurnal and temperature variability, as well as threshold responses to summer temperature and precipitation, with turnover especially strong in edge populations that occur at high elevation but southerly latitudes. We discuss these results in light of how climate may disrupt locally adaptive gene–environment relationships, and suggest that rear edge populations hold climate-adaptive variants that should be targeted for conservation.

Subject area: Population structure and phylogeography; Molecular adaptation and selection
Keywords: landscape genetics, local adaptation, range limits, rear edge, plant circadian clock

There is a growing awareness that the conservation of local adaptation in the face of climate change requires a comprehensive understanding of how evolutionary processes shape genetic diversity in different parts of a species' range (Hampe and Petit 2005; Sexton

et al. 2009; Woolbright et al. 2014). In northern hemisphere temperate or boreal biomes, populations along the southern or "rear edge" of species' range are an emerging priority for landscape studies identifying genes that contribute to local adaptation (Hampe and Petit 2005). For many species, the rear edge is closest to the estimated location of glacial refugia prior to postglacial range expansion (Petit et al. 2003), and is therefore predicted to contain genetic diversity not found elsewhere in the range (Lepais et al. 2013). Rear edge populations are also predicted to occupy unique positions along abiotic and biotic environmental gradients compared to the rest of the species range, such as higher elevations, warmer, longer, or more variable growing seasons, or novel biotic interactions with other species in their communities (Woolbright et al. 2014). Thus, range edge populations may harbor unique adaptive alleles or genotypic combinations of alleles in response to these environments not found elsewhere in the range (Hoffmann and Sgrò 2011). Lastly, populations near range edges may grow in marginal habitats near the limits of physiological tolerance; thus there is a need to identify climate-adapted diversity unique to these populations while they are still extant (Kawecki 2008; Hampe and Jump 2011).

Trees offer excellent study systems for using landscape genomic approaches to discover climate-adapted regions of the genome (Petit and Hampe 2006), and for identifying local populations or geographic regions with adaptive variants that could be utilized in conservation, restoration, and breeding efforts for biomass production (Savolainen et al. 2013; Evans et al. 2014; Bragg et al. 2015; Holliday et al. 2016). Previous work in *Populus* has demonstrated quantitative genetic clinal variation in phenology and growth traits along environmental gradients that determine growing season length (Dynesius and Jansson 2000; Keller et al. 2011b; Soolanayakanahally et al. 2013b; McKown et al. 2014). In particular, the timing of spring bud flush and late summer bud set—traits that together coordinate the timing of growth and dormancy with the permissive growing season—possess strong heritability in *Populus* (Olson et al. 2013) and exhibit clinal variation with growing degree days and daylength under common garden conditions (Keller et al. 2011b; Soolanayakanahally et al. 2013a; McKown et al. 2014).

A variety of functional studies pioneered in *Populus* have given us a basic model for the genetic control of vegetative phenology (Böhlenius et al. 2006; Hsu et al. 2011; Brunner et al. 2014). These studies strongly implicate genes in the plant circadian clock and associated flowering time pathway controlling tree phenology, principally genes responsible for photoperiod signaling (e.g., phytochromes and cryptochromes), circadian regulation of environmental signals (*GIGANTEA, EARLY FLOWERING-3*), and downstream signal integration via the *CONSTANS/FLOWERING LOCUS T* (CO/FT) regulon. Interactions among these genes and their responsiveness to external environmental cues control the timing of seasonal dormancy onset and release that governs local adaptation to growing season length. Appreciable allelic variability in these genes has been found to segregate in natural stands of *Populus*, with some genes exhibiting elevated population structure, clines in allele frequencies with climate, and genotype–phenotype associations (Ma et al. 2010; Rohde et al. 2011; Keller et al. 2012; Evans et al. 2014; Wang et al. 2014; Fitzpatrick and Keller 2015). These studies have provided some of the strongest support to date for identifying associations between environmental drivers of selection, the adaptive phenotypes responding to these environments, and the underlying genetic variants responsible for adaptation on the landscape (Sork et al. 2013; Rellstab et al. 2015).

Balsam poplar (*Populus balsamifera* L., Salicaceae) occurs in boreal forest ecosystems across North America. Balsam poplar's huge geographic range spans broad variation in climate, and populations show strong ecophysiological divergence along latitudinal and climatic gradients (Soolanayakanahally et al. 2009; Keller et al. 2011b). Extensive sampling in the core of the range has been conducted, and association mapping and local adaptation scans in single nucleotide polymorphisms (SNPs) from genes in the flowering time pathway revealed significant latitudinal divergence (Keller et al. 2011a), associations with bud phenology (Olson et al. 2013) and strong SNP–environment covariation with climate (Keller et al. 2012; Fitzpatrick and Keller 2015). However, rear edge populations are extensive in *P. balsamifera*, and their contributions to climate adaptation have not previously been investigated. Here, we report extensive new sampling of rear edge populations of *P. balsamifera* and analyze SNPs in the flowering time genetic pathway to address three key questions about the influence of range position on the genetics of local adaptation: 1) Are rear edge populations ecologically and/or genetically separated from the range core in terms of the climate inhabited and inter-population genetic connectivity? 2) Do rear edge populations show evidence of gene–environment associations that are distinct from the range core? 3) What parts of the range will experience the greatest loss of local adaptation due to disrupted gene–environment associations under future climates?

Methods

Sampling

We analyzed 1021 balsam poplar individuals collected from across its range. This included 443 individuals reported previously in Keller et al. (2012), as well as 578 new individuals genotyped for this study. This new sampling specifically targeted the southern rear edge of the species range, which we define here as locations south of 52°N latitude in regions west of 100°W longitude, and south of the Great Lakes in regions east of 100°W longitude (Figure 1). Three individuals from the IVI population previously reported by Keller et al. (2012) were combined with the COT population, which is located in close proximity. The total number of populations combined across the current study and Keller et al. (2012) was 90, consisting of 55 core populations and 35 rear edge populations. We also sampled 12 individuals from closely related *Populus* spp. to serve as outgroups for determining the ancestral versus derived state of each SNP allele. These outgroup samples were *Populus angustifolia* (1), *Populus deltoides* (1), *Populus tremula* (1), *Populus tremuloides* (1), and *Populus trichocarpa* (8).

DNA Extraction and SNP Genotyping

Germplasm collections from natural populations across the range were made during the winter months when trees were dormant (February–April). Vegetative cuttings of 6–10 cm were forced to flush under permissive temperatures in a greenhouse. Total genomic DNA was extracted from fresh leaf tissue using Qiagen 96 DNeasy kits, and DNA concentration estimated using the Qubit fluorometric BR assay.

Targeted SNP genotyping was carried out using the Sequenom iPlex genotyping system, as described previously in Keller et al. (2012). Briefly, we used the published multiplexed SNP assays reported in Keller et al. (2012), which were derived from sequencing a range-wide panel of 24 *P. balsamifera* individuals for each of 27 genes in the flowering time pathway and associated pathways

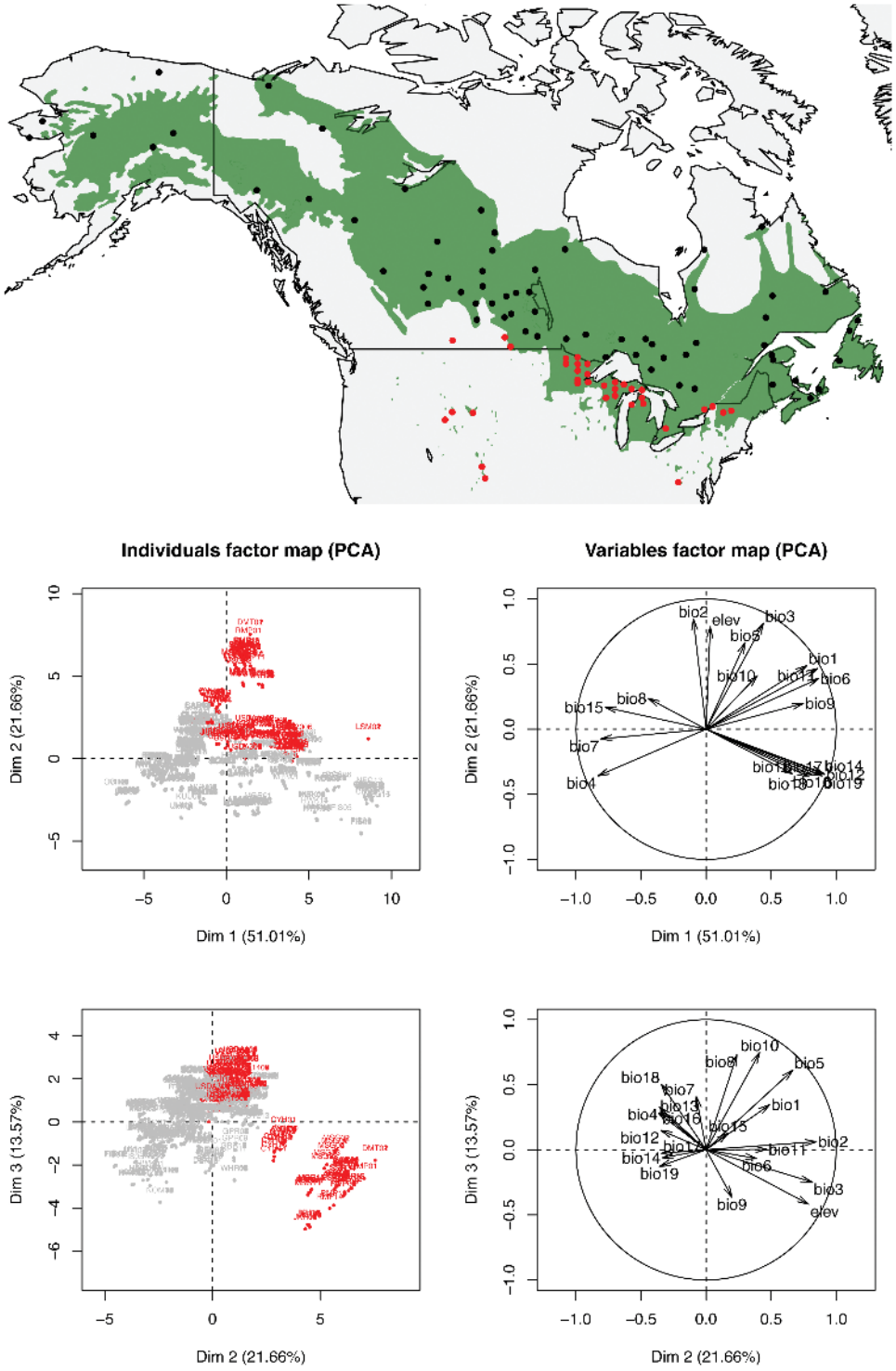

Figure 1. Range map showing sampling localities of 90 populations, with rear edge sites in red, and range core and northern edge sites in black. PCA shows range-wide differentiation in climate space based on 19 Bioclim variables plus elevation. PCA ordination of rear edge populations (red) from other range-wide populations (gray) along the first 3 principal components explained a cumulative 86% of the variance in climate among sites.

affecting vegetative and reproductive plant phenology (Keller et al. 2011a). These genes encompass functional roles involved in response to environment such as red and blue light photoreceptors, core and peripheral components of the plant circadian clock, downstream targets of the clock involved in phenology such as florigens, vernalization, and abscisic acid mediated response to stress and dormancy (Table 1, Supplementary Figure S1). In the current study, we collected SNP data from 14 multiplex genotyping assays targeting a total of 381 SNPs. We filtered SNPs for those showing low genotyping success rate across individuals, deviation from Hardy–Weinberg equilibrium due to excess heterozygosity (likely due to paralogy), inconsistent detection of copy number variants (i.e., 1 bp indels),

Table 1. Genes involved in flowering time and associated pathways used for SNP genotyping

Gene	Gene ID	Transcript ID	LG	Annotation	N SNPs
ABA insensitive-1B	*ABI1B*	Potri.006G224600	6	Abscisic acid	13
ABA insensitive-1D	*ABI1D*	Potri.018G060300	18	Abscisic acid	13
ABA insensitive-1D	*ABI3*	Potri.002G252000	2	Abscisic acid	15
Casein kinase-2 subunit β 3.4	CKB3.4	Potri.002G139900	2	Peripheral circadian clock	11
Constans-1	*CO1*	Potri.017G107500	17	Downstream target	10
Constans-2	*CO2*	Potri.004G108300	4	Downstream target	10
Cryptochrome-1.1	*CRY1.1*	Potri.005G164700	5	Photoreceptor	10
Cryptochrome-1.2	*CRY1.2*	Potri.002G096900	2	Photoreceptor	14
Early bolting in short days	*EBS*	Potri.002G226000	2	Downstream target	4
Early flowering-3	*ELF3*	Potri.003G045000	3	Peripheral circadian clock	33
Frigida	*FRI*	Potri.015G110100	15	Vernalization	7
Gigantea-2	*GI2*	Potri.002G064400	2	Peripheral circadian clock	19
Gigantea-5	*GI5*	Potri.005G196700	5	Peripheral circadian clock	22
Heme oxygenase-1.1	*HY1.1*	Potri.006G069700	6	Photoreceptor	9
Heme oxygenase-1.2	*HY1.2*	Potri.018G131700	18	Photoreceptor	17
Heme oxygenase-2.1	*HY2.1*	Potri.009G048000	9	Photoreceptor	11
Heme oxygenase-2.2	*HY2.2*	Potri.001G253700	1	Photoreceptor	6
Leafy	*LFY*	Potri.015G106900	15	Downstream target	16
Phytochome-A	*PHYA*	Potri.013G000300	13	Photoreceptor	5
Phytochome-B1	*PHYB1*	Potri.008G105200	8	Photoreceptor	6
Phytochome-B2	*PHYB2*	Potri.010G145900	10	Photoreceptor	8
Phytochome interacting factor-3	*PIF3*	Potri.013G001300	13	Photoreceptor	8
Flowering locus T-1	*PtFT1*	Potri.008G077700	8	Downstream target	4
Terminal flowering-1.1	*TFL1.1*	Potri.009G165100	9	Downstream target	2
Timing of cab-1	*TOC1*	Potri.015G061900	15	Central circadian clock	11
ZEITLUPE-1.1	*ZTL1.1*	Potri.018G090800	18	Photoreceptor	2
ZEITLUPE-2.9	*ZTL2.9*	Potri.006G166300	6	Photoreceptor	4

and loci that were only polymorphic with inclusion of outgroup taxa. The resulting filtered dataset consisted of 291 biallelic SNP loci genotyped in 1021 *P. balsamifera* individuals. Genotypes were converted to 012 format based on the number of derived alleles, using the outgroup taxa to determine the ancestral state at each locus by majority rule.

Statistical Analysis

Climate Associations with Range Position

We assessed climatic differences between core and edge populations using the Worldclim climate dataset at a resolution of 30 arc-seconds (~1 km) (Hijmans et al. 2005). We analyzed the 19 Bioclimatic (Bioclim) variables that use monthly means of temperature and precipitation to describe different biologically relevant climatic variables that take into account seasonality and extremes of climate. The 19 Bioclim variables plus the elevation (m) of each collection site were used in a principal component analysis (PCA) to summarize the major axes of climate variability, using the *FactomineR* package (Lê et al. 2008) in R (R Core Team 2014). Mean differences in climate principal components (PCs) between core and rear edge populations were analyzed with a Mann–Whitney *U* test.

Population Structure

We examined population structure by calculating F_{ST} among populations using the *diveRsity* R package (Keenan et al. 2013), and by using discriminant analysis of principal components (DAPC) on the SNP genotype data using the *adegenet* R package (Jombart et al. 2010; Jombart and Ahmed 2011). We used K-means clustering to evaluate the number of groups consistent with the genotypes, and retained 150 PCs to be used as predictors in the discriminant analysis. We calculated the loadings of each SNP on the discriminant axes to assess which loci contributed most strongly to defining the population structure in the sample. We also evaluated genetic population structure with Dyer's Population Graph approach to analyze the network topology and conditional genetic distances among populations at a landscape scale (Dyer and Nason 2004; Dyer 2015). We estimated Population Graphs for all SNPs combined using the R package *popgraph*.

Tests for Local Adaptation

We tested for the action of local selection in structuring SNPs in the flowering time pathway based on identifying loci with 1) excess population structure relative to neutral expectations (F_{ST} outliers), and 2) testing for association between SNP genotypes and climate (gene–environment associations).

F_{ST} outliers were identified based on the *F*-model implemented in *Bayescan* (Foll and Gaggiotti 2008). We first analyzed divergence among the entire set of 90 populations to test for range-wide local adaptation. In addition, to investigate if different genes were targets of local selection in range core versus rear edge populations, we ran additional models separately for each range subset (55 core populations, 35 rear edge populations), for a total of 3 *Bayescan* runs. For each run, we set the prior odds of a locus being neutral versus under selection to 10 000 based on recommendations from simulated data under a variety of population structures, which suggest this stringent prior helps to minimize the false positive rate (Lotterhos and Whitlock 2014). In addition, our SNP data are derived from a priori candidate genes, and so would be expected to harbor a larger proportion of selected SNPs than a random sample from the genomic background, suggesting that our prior odds of 10 000 may be conservative. Additional run parameters consisted of 20 pilot MCMC chains (5000 iterations each) followed by a burn-in period of 50 000 iterations and sample collection of 100 000

iterations with a thinning interval of 10. Outlier loci for local adaptation were identified as those showing evidence of excess population differentiation (locus-specific $\alpha > 0$) at a false discovery rate (FDR) of 0.1.

We tested for SNP–climate associations indicative of local adaptation using the latent factor mixed model (*LFMM*) approach (Frichot et al. 2013). We used the first 3 axes from the PCA of the Bioclim data as predictors to test for associations between SNPs and climate. To control for population demography, we ran 10 replicates of *LFMM* for each of a range of latent factors from $K = 2$ to 10 (Frichot et al. 2013). Each replicate consisted of a burn-in of 10 000 iterations followed by data collection over 50 000 iterations. Corrected *P*-values were calculated by combining *z*-scores across the 10 replicates per *K* and dividing by the genomic inflation factor, as described by Frichot et al. (2013). The number of latent factors was chosen based on evaluating the genomic inflation factor that resulted in an unbiased *P*-value distribution (Frichot et al. 2013), and also by comparing choice of *K* to results from genotypic clustering using *fastStructure* (Raj et al. 2014). This resulted in $K = 5$ latent factors for the range-wide sample, $K = 5$ for the range core, and $K = 1$ for the rear edge. Outlier loci in *LFMM* were determined based on an FDR of 0.1.

We compared evidence for excess population structure and gene–environment association from *Bayescan* and *LFMM* to that obtained by *Bayenv2*, which uses a population matrix (Ω) proportional to pairwise F_{ST} to control for covariance in allele frequencies when testing SNP-wise local adaptation (Günther and Coop 2013). Because we don't have an a priori neutral set of SNPs, we estimated Ω using all SNPs. We used the first 3 climate PCs as predictors to test for gene–environment association, as in *LFMM*. We ran 10^5 iterations of the mcmc chain and compared Ω at intervals of 5000 steps. We found these to be highly congruent, and kept the final Ω matrix at the end of the 10^5 steps for further analysis. We then ran 3 replicate runs of *Bayenv2* for 10^5 iterations of the mcmc chain and averaged outputs over runs for each SNP. Because *Bayenv2* does not perform tests of significance, we ranked SNPs by their estimated population structure (XTX, similar to F_{ST}) and gene–environment association (Bayes factors) and retained SNPs in the top 5% quantile.

Landscape Genomics of Adaptive Gene–Environment Associations

To map the spatial turnover of adaptive variants on the landscape, we used the multivariate approach of gradient forest (GF) as described in Fitzpatrick and Keller (2015). Briefly, GF (Ellis et al. 2012) is an extension of the random forest machine-learning algorithm that models change in SNP allele frequencies across sites using nonlinear functions of environmental predictors (i.e., climate gradients). Being both nonlinear and multivariate, GF accommodates many of the biological complexities of clinal selection on allele frequencies that current methods in gene–environment association tests do not (Rellstab et al. 2015), including threshold responses to environmental gradients and simultaneous modeling of multiple, correlated predictor variables (Fitzpatrick and Keller 2015).

We used GF to analyze changes in derived allele frequency of flowering time SNPs along gradients of contemporary climate using the Bioclim variables. To minimize effects of strongly autocorrelated predictors, we reduced the climate dataset to 6 Bioclim variables that were minimally correlated (<0.75) with each other: bio2 (mean diurnal range), bio5 (maximum temperature of the warmest month), bio6 (minimum temperature of the coldest month), bio18 (summer precipitation), bio19 (winter precipitation), elevation (m), as well as latitude and longitude to account for geographic effects.

We restricted our GF analyses to SNPs that were polymorphic in at least 5 populations, using 500 regression trees per SNP. We used default values for the number of predictor variables randomly sampled as candidates at each split (2 in this case) and for the proportion of samples used for training (~0.63) and testing (~0.37) each tree. To evaluate where adaptive SNP–environment associations would become most disrupted under near-term climate change scenarios, we calculated an average "adaptive offset" measure by projecting the GF models to multiple scenarios of future climate for year 2050 based on 32 different Earth system models and one representative concentration pathway (RCP 8.5, a high-range emissions scenario consistent with current trajectories). This time period is short enough into the future that dispersal is unlikely to be a mitigating factor in predicting population response; therefore, a change in the position of a given locality along the SNP–environment turnover function provides a measure of the potential disruption of local adaptation (Fitzpatrick and Keller 2015). We interpret adaptive offset as a SNP-level prediction of future climate-stress due to loss of local adaptation. Future climate scenarios were downloaded from the CCAFS climate data center (http://www.ccafs-climate.org/).

Results

Climate Differences between Core and Rear Edge Populations

The 1021 individuals sampled covered a broad geographical and climatic range of environments. Collections ranged from 39.22°N to 69.10°N latitude (Figure 1), and spanned mean annual temperatures of –10.4 °C (COT and IVI) to 9.1 °C (LSM). PCA partitioned populations along PC1 (51% of variance) that described an axis of warm, wet sites with low seasonality (high PC1 scores) versus cooler, drier sites with greater seasonality (low PC1 scores). The PC2 axis explained 22% of the climate variance, and described an axis of high elevation warm sites with a large diurnal range (difference between daily maximum and minimum temperatures) versus low elevation cool sites with a low diurnal range. Populations along the southern range edge had significantly higher values of both PC1 and PC2 (Mann–Whitney *U* test, $P < 0.0001$) (Figure 1). Thus, rear edge populations of *P. balsamifera* differ from the range core by occupying climates characterized by greater temporal variability and high maximum temperatures (bio5, bio10), with associated high variance in diurnal temperatures (bio2, bio3). Interestingly, a third climate dimension (PC3; 14%) separated rear edge populations into 2 groups: a western mountain group in Colorado, Wyoming, and Alberta characterized by higher elevations and higher monthly and seasonal temperature variability (bio2, bio3), and a second group near the Great Lakes characterized by warmer growing season temperatures (bio5, bio8, bio10).

Population Genetic Structure and Connectivity

Population F_{ST} averaged across all loci was 0.1070 (95% confidence interval (CI) = 0.0970–0.1149), and varied widely among population pairs (Supplementary File 1). The network topology of genetic structure across the landscape showed distinct patterns of connectivity between rear edge populations relative to the range core (Figure 2). Conditional genetic covariances among populations based on the SNP data showed that rear edge populations in the western mountains (RMP, JKH, SSR, MSG, CYH) formed an internally connected network, but with relatively few connections to populations in the core of the range. The latter consisted of geographically disparate edge–core connections, such

as between RMP (Colorado) and LON (Ontario), or between MSG (Wyoming) and USDA populations (Great Lakes region), probably a reflection of historical gene flow during postglacial expansion (Keller et al. 2010). Elsewhere in the rear edge, USDA populations around the Great Lakes showed a high degree of genetic connectivity with each other and populations in southern Ontario and New England, but relatively low connectivity with nearby core populations in northern Ontario, Manitoba, or Saskatchewan (Figure 2). Within the core of the range, populations showed high connectivity along a southeast to northwest trending axis from New England to NW Canada and Alaska. The exception to the broadly connected range core was the group of populations in far eastern Quebec and Atlantic Canada, which formed a separate relatively isolated network of populations, consistent with previous studies that identify this region as a distinct subpopulation (Keller et al. 2010).

The genetic isolation of rear edge populations compared to the range core and Atlantic Canada was also evident in the DAPC analysis (Figure 3). *K* means clustering of SNP genotypes identified 3 main groups, which DAPC separated into individuals from the western mountains (along DA-1), and between Atlantic Canada and the rest of the range core (along DA-2). The genes containing SNPs that contributed most strongly to the separation of these groups primarily involved 2 circadian genes upstream of the *CO/FT* regulon: namely, *EARLY FLOWERING-3 (ELF3)* which loaded strongly on DA-1 and distinguished rear edge populations in the western mountains, and 1 of the 2 *GIGANTEA* paralogs (*GI5*) which loaded strongly on DA-2, distinguishing the range core from Atlantic Canada populations.

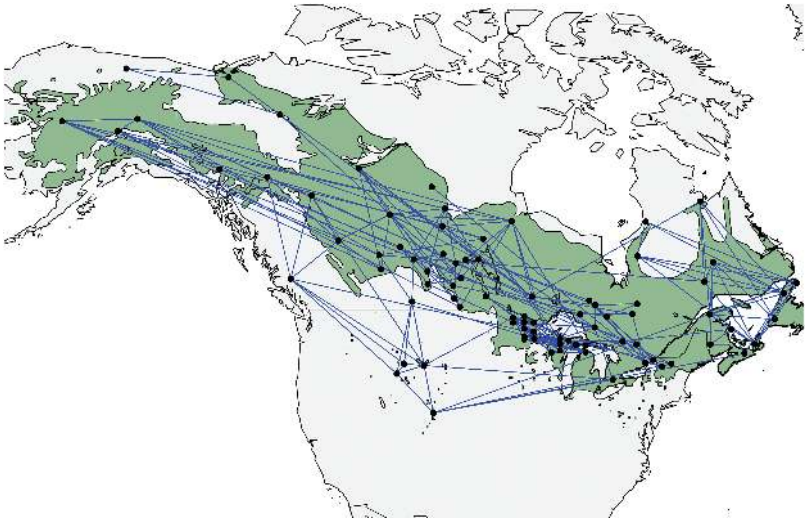

Figure 2. Population graph covariance network based on 291 SNPs in the flowering time pathway. Lines connecting populations reflect shared covariance in allele frequencies, conditional on other connections between populations in the network.

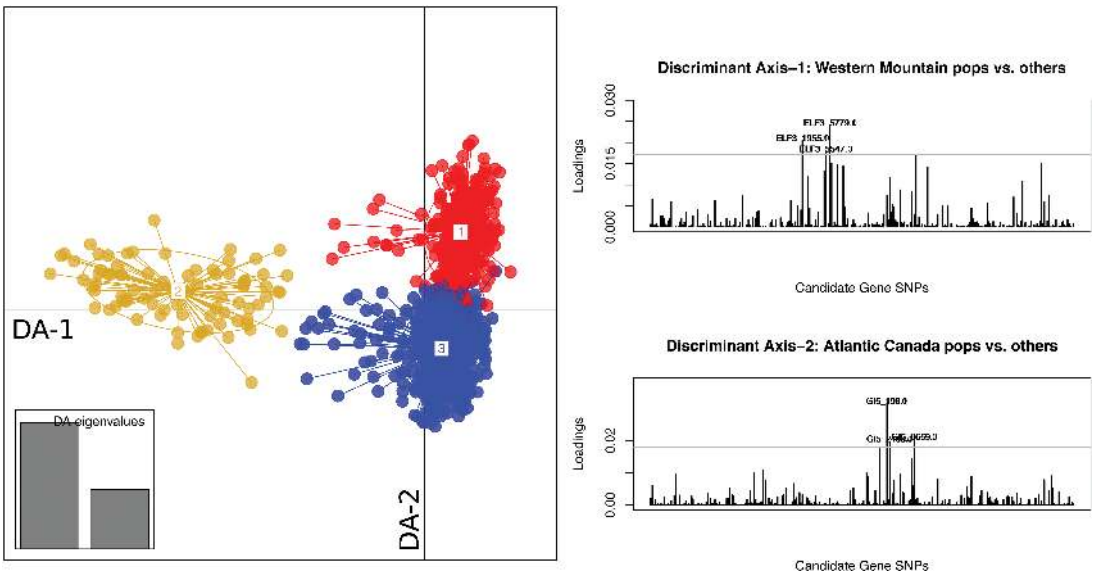

Figure 3. DAPC based on 291 flowering time SNPs across 1021 individuals. Shown are the biplot of individual coordinates along the first 2 discriminant axes (1: Atlantic Canada populations; 2: Western Mountain populations; and 3: Range Core populations), and loadings of candidate gene SNPs on the first 2 DA axes.

Local Adaptation in Rear Edge Versus Range Core Populations

Range-wide local adaptation was evident in the *Bayescan* analysis as elevated population structure (F_{ST}) for 13 SNPs, primarily in the *ELF3* (*N* = 7) and *GI5* genes (*N* = 5 SNPs; Supplementary Figure S2), both of which are part of the peripheral circadian clock (Supplementary Figure S1). The *GI5* SNP 5271 was also in the top 5% of XTX from *Bayenv2* (Supplementary File 2). One SNP in poplar *FLOWERING LOCUS T1* (*PtFT1*) also showed evidence for local adaptation in the range-wide sample.

Interestingly, the targets of local adaptation changed dramatically when analyzing range core and edge populations separately (Figure 4). Within the range core, there was still abundant evidence for local adaptation for SNPs in *GI5* but not in *ELF3*, while the rear edge showed no SNPs with elevated population structure (Figure 4). This suggests that allele frequency divergence in *ELF3* in the range-wide sample is primarily aligned between rear edge and core portions of the range, and not within local populations within each region, consistent with results from DAPC.

Tests of gene–environment associations in the range-wide sample using *LFMM* also strongly supported SNPs in *GI5* and *ELF3* as under climate-driven selection (Table 2). Significant associations in *ELF3* (*N* = 6 SNPs) were along a gradient of elevation and diurnal temperature variability (PC2), whereas significant associations in *GI5* (*N* = 7 SNPs) were associated with both PC2 and differences in overall temperature, precipitation, and their seasonality (PC1). Other genes showing significant range-wide association with PC1 included *ABi1B* (1 SNP), *CRY1.1* (3 SNPs), *EBS* (1 SNP), and *HY2.1* (2 SNPs); whereas significant association with PC2 was found for *ABi3* (2 SNPs), *CKB3.4* (2 SNPs), *CRY1.1* (1 SNP), *HY2.2* (1 SNP), *PHYB2* (1 SNP), *PtFT1* (1 SNP), and *ZTL2.9* (1 SNP). One SNP in *HY2.1* also showed associated with climate PC3. Five SNPs in the genes *GI5* and *CRY1.1* that were significant in LFMM were also in the top 5% of Bayes factors from *Bayenv2* (Table 2).

In contrast to the range-wide sample, gene–environment associations in the core-only were aligned exclusively with PC1 of the climate data (Table 2), in which many of the significant associations in the range core were also detected in the range-wide sample (9/14). Two SNPs in the genes *ABi1D* and *CRY1.1* were also in the top 5% of Bayes factors from *Bayenv2* (Table 2). The lack of gene–environment associations with PC2 among core-only samples points to adaptive divergence between core and edge populations in their response to diurnal climate variability, principally involving SNPs in *ELF3* (Table 2). Consistent with this idea, when edge populations were analyzed separately (edge-only), they showed relatively minimal local adaptation along PC1, but abundant local adaptation along PC2 and PC3. Many of the associations with PC2 in the rear edge sample involved *ELF3* SNPs (*N* = 11 SNPs), as well as heme oxygenase genes, *HY1.2* (2 SNPs), *HY2.1* (2 SNPs), and *HY2.2* (1 SNP). Additionally, edge-only populations contained outliers for local adaptation along PC3, unlike the range-wide and core-only samples. Four SNPs in *ELF3* were also in the top 5% of Bayes factors from *Bayenv2*, as was one SNP from the flowering time gene *PtFT1* (Table 2).

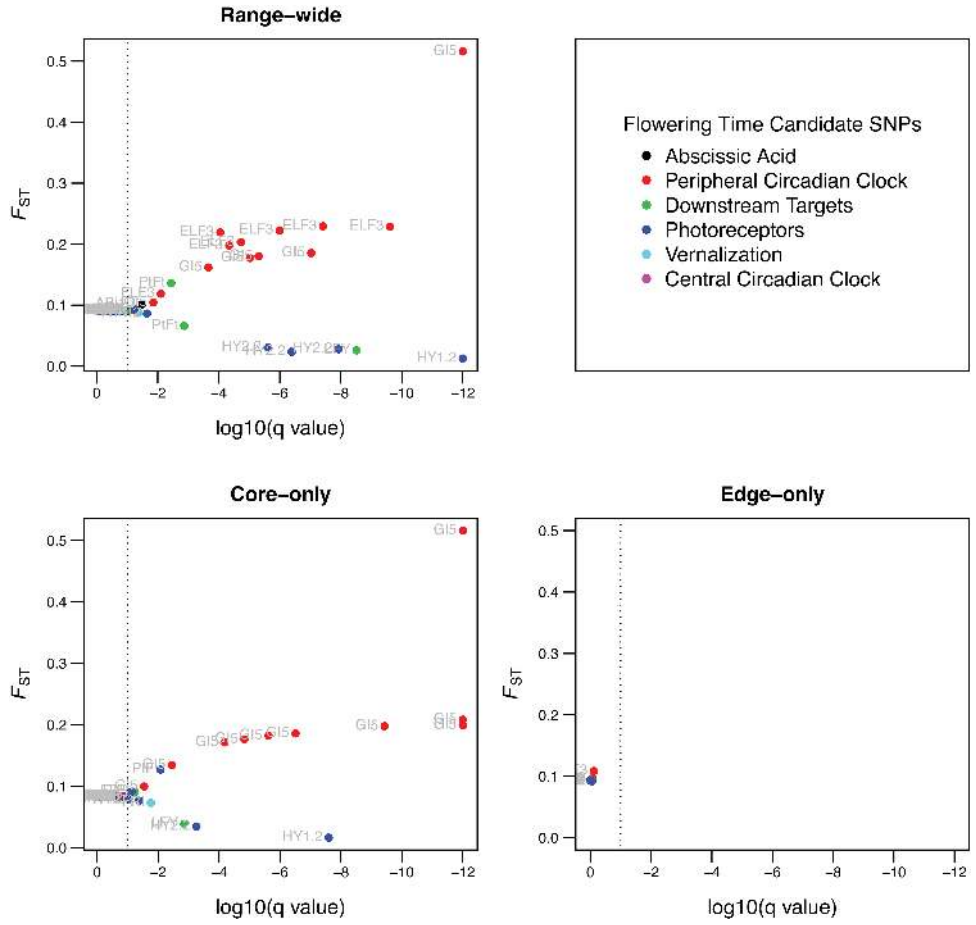

Figure 4. Bayescan plot identifying candidate outliers based on 291 flowering time SNPs. Threshold log10 *q* values for a FDR of 0.1 are shown as vertical lines. SNPs are color-coded according to their functional annotation within the flowering time network.

Table 2. Candidate SNPs showing association with climate PCs in different range positions

Candidate SNP	Range-wide (N = 1021)	Core-only (N = 734)	Edge-only (N = 290)
Climate PC1			
ABi1B_2376	0.0034		
ABi1D_1207		0.0022	
ABi1D_1354		*0.0018*	
CO1_576			0.0002
CRY1.1_974	0.0005	0.0001	
CRY1.1_3004	*0.0001*	<0.0001	
CRY1.1_3018	*0.0033*	*0.0004*	
EBS_609	0.0004	0.0010	
ELF3_5376		0.0022	
GI5_33	0.0008	**<0.0001**	
GI5_198		**0.0047**	
GI5_1950	**0.0005**	**0.0019**	
GI5_2405	**≤0.0001**	**<0.0001**	
GI5_8997		**0.0024**	
GI5_9585	0.0024		
GI5_9659	0.0001	**0.0003**	
HY2.1_236	0.0002		
HY2.1_1615	0.0002	<0.0001	
Climate PC2			
ABi1B_2055			0.0050
ABi1B_2274			<0.0001
ABi3_2670	<0.0001		
ABi3_3707	0.0013		0.0002
CKB34_1245	0.0041		
CKB34_1292			0.0008
CKB34_3188	0.0026		<0.0001
CKB34_3240	0.0023		0.0004
CO1_642			0.0106
CRY1.1_275	0.0001		<0.0001
CRY1.1_2742	*0.0047*		0.0012
CRY1.2_1187	0.0068		
EBS_609			0.0071
ELF3_75	0.0037		0.0002
ELF3_1099	0.0050		0.0016
ELF3_1955			*0.0018*
ELF3_2185	0.0019		
ELF3_2559	**0.0044**		<0.0001
ELF3_2620			0.0046
ELF3_4541			0.0105
ELF3_5376	**<0.0001**		*≤0.0001*
ELF3_5547			0.0008
ELF3_5779			0.0025
ELF3_5785			0.0050
ELF3_6167	0.0009		<0.0001
GI5_33	0.0010		0.0023
GI5_2612	0.0003		0.0002
GI5_9551	0.0002		0.0014
HY1.2_753			0.0043
HY1.2_1511			0.0010
HY2.1_173			0.0023
HY2.1_2491			<0.0001
HY2.2_4222			0.0092
HY2.2_4397	0.0062		
PHYB2_1417	<0.0001		<0.0001
PtFt_3044	**0.0002**		<0.0001
ZTL2.9_5936	0.0037		<0.0001
Climate PC3			
ABi1B_2274			0.0007
CRY1.1_275			<0.0001
ELF3_75			*0.0002*
ELF3_5376			*≤0.0001*
ELF3_6167			<0.0001
HY2.1_2491	0.0001		
PtFT1_3044			*≤0.0001*

Values are corrected P-values identified by *LFMM* (FDR = 0.1). Italicized SNPs are in the top 5% of Bayes factors from *Bayenv2*, and those in bold are F_{ST} outliers in *Bayescan* (FDR = 0.1).

Landscape Modeling of Local Adaptation under Current and Future Climates

GF analysis showed 191 of the 262 SNPs that were polymorphic in at least 5 populations exhibited allele frequency turnover along environmental gradients (R^2 > 0; Supplementary File 3). The 30 SNPs that LFMM identified as significant outliers in the range-wide analysis (Table 2) had a large portion of their allele frequency variation explained by the combined effects of climate, elevation, and geography (latitude and longitude) in multivariate GF models (R^2 per SNP: 0.07–0.71, mean = 0.41). Allele frequency turnover was especially strong along gradients of elevation, summer precipitation (bio18), maximum temperature (bio5), and mean diurnal temperature range (bio2) (Figure 5). Turnover along the portions of environmental gradients occupied by rear edge populations was prominent for higher elevations with a large diurnal range (Figure 5). Functional annotations of genes exhibiting the strongest turnover included the peripheral circadian clock (8/22 SNPs with R^2 > 0.5), followed by photoreceptors (7 SNPs), downstream targets (5 SNPs), and abscisic acid (2 SNPs).

We used GF to assess the potential for loss of local adaptation due to mismatches of climate-adaptive SNPs with future climate. We restricted this analysis to just those SNPs with strong consensus for climate-driven local adaptation—that is, those significant in both *Bayescan* and *LFMM* selection scans, and in the top 5% of Bayes factors from *Bayenv2* (Table 2). Four SNPs met these criteria: ELF3_5376, GI5_1950, GI5_2405, and PtFT1_3044. For these 4 SNPs, climate explained a large portion of the variance in allele frequencies in multivariate GF models (R^2 per SNP = 0.48–0.62). We compared the position of each population along allele frequency turnover functions for these 4 SNPs between current and future climate conditions to calculate the adaptive offset. Mapping this offset back onto the landscape revealed that most of the predicted loss of local adaptation between current and future climates is concentrated along the northern range margin in areas of the northwest such as Alaska and the Yukon, particularly for SNPs in *GI5* and *PtFT1* (Figure 6). *ELF3* showed some evidence of adaptive offset in the southern part of the range, but this was of lesser magnitude compared to offset of *GI5* and *PtFT1* SNPs in the north. Thus, despite the unique patterns of adaptive SNP variation found in rear edge populations, near-term climate change is not predicted to dramatically shift populations along the steepest portions of the environmental gradients in southern populations.

Discussion

Populations at the rear edge of a species geographic range may differ from the range core in the selective environments experienced, in

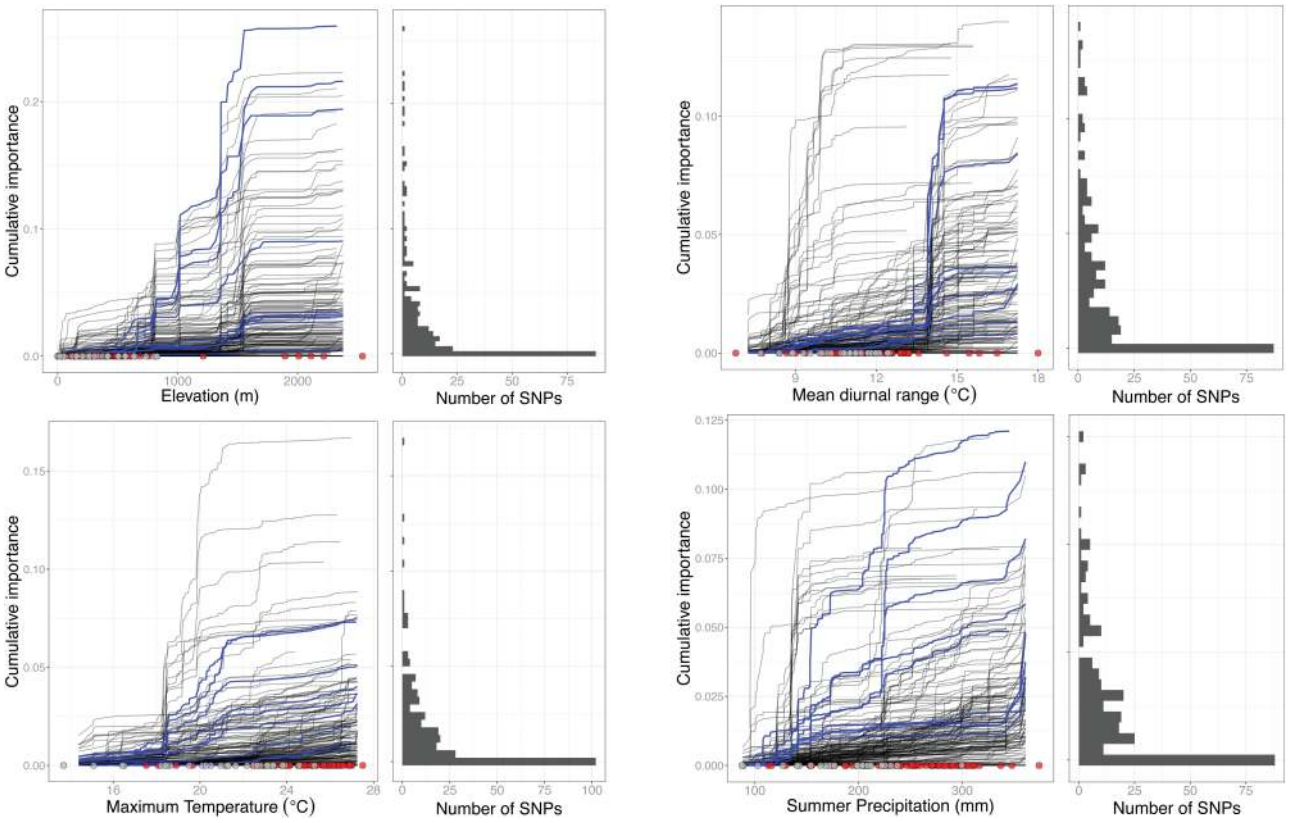

Figure 5. Compositional turnover in SNP frequencies modeled by GF. Shown are turnover plots and histograms of SNP-wise cumulative importance along the gradient for 4 of the 8 environmental predictors in the multivariate model GF model (Supplementary Figure S3). Each line in a turnover plot represents the cumulative variance in allele frequency for a single SNP explained by the environmental predictor, with SNPs in blue indicating significant outliers in both *Bayescan* and *LFMM*. Position of sites along the gradients are plotted as points along the *x*-axes for core (gray) or edge (red) populations.

their standing levels of genetic diversity, and in their genetic architecture of local adaptation (Hampe and Jump 2011). This makes rear edge populations particularly interesting focal regions to understand how range context affects the outcome of local adaptation, especially in the face of maladaptive gene flow (Bridle and Vines 2007; Kawecki 2008; Woolbright et al. 2014).

We found that rear edge populations differed from core populations in their climatic environment, and, as might be expected given their more southerly latitude, rear edge populations inhabited warmer, and generally wetter climates. What was less obvious was that rear edge populations also experienced more variable climates, both diurnally and seasonally across the year (Figure 1). Much of this appeared to be influenced by the higher elevations of western mountain populations, where growing season length and temporal variability can change dramatically over relatively short distances along steep local gradients such as elevation.

The separation of rear edge populations in climatic space was reinforced by the overall patterns of genetic connectivity, suggesting limited gene flow between the rear edge and the range core (Figure 2). Thus, despite the history of range expansion in *P. balsamifera* from ancestral diversity contained in southern refugial populations (Keller et al. 2010; Levsen et al. 2012), contemporary patterns of gene flow do not appear to suggest abundant shared diversity between the core and the rear edge. Rather, edge populations near the western mountains and southern Great Lakes show high internal connectivity, but limited connectivity with the range core. Rear edge populations of *P. balsamifera* in the western mountains also have a history of introgression from *P. angustifolia* and *P. trichocarpa*, and analysis of genome-wide SNP data indicate positive selection following introgression has contributed to adaptive evolution in these populations (Chhatre et al. Forthcoming). This scenario could facilitate the spread of adaptive variants among local subpopulations, while minimizing the influx of maladaptive alleles from divergent selective environments in the range core (Holt et al. 1997; Sexton et al. 2011). Modularity in the genetic network connectivity was also clear for populations from Atlantic Canada, while the remaining populations across the core showed broad connectivity along a NW-SE trending axis, suggesting abundant gene flow in the demographic center of the range.

These patterns of connectivity between the rear edge, Atlantic Canada, and the range core were mirrored in the DAPC ordination of SNP genotypes (Figure 3), which also revealed that the genes contributing most strongly to this regional divergence consisted of 2 peripheral circadian clock genes, *ELF3* and *GI5*. These genes also showed clear signatures of local adaptation as F_{ST} outliers and gene–environment associations, after accounting for population structure. While we do not have a putatively neutral set of SNPs from the genomic background to compare against, we were conservative in our testing approach to local adaptation, setting stringent prior odds for selection in *Bayescan* (1:10 000), controlling for background population structure by incorporating latent factors in *LFMM*, and using the population covariance of allele frequencies to control for demography in *Bayenv2*. Agreement among methods in identifying outliers was evident between *Bayescan* and *LFMM*, but few of

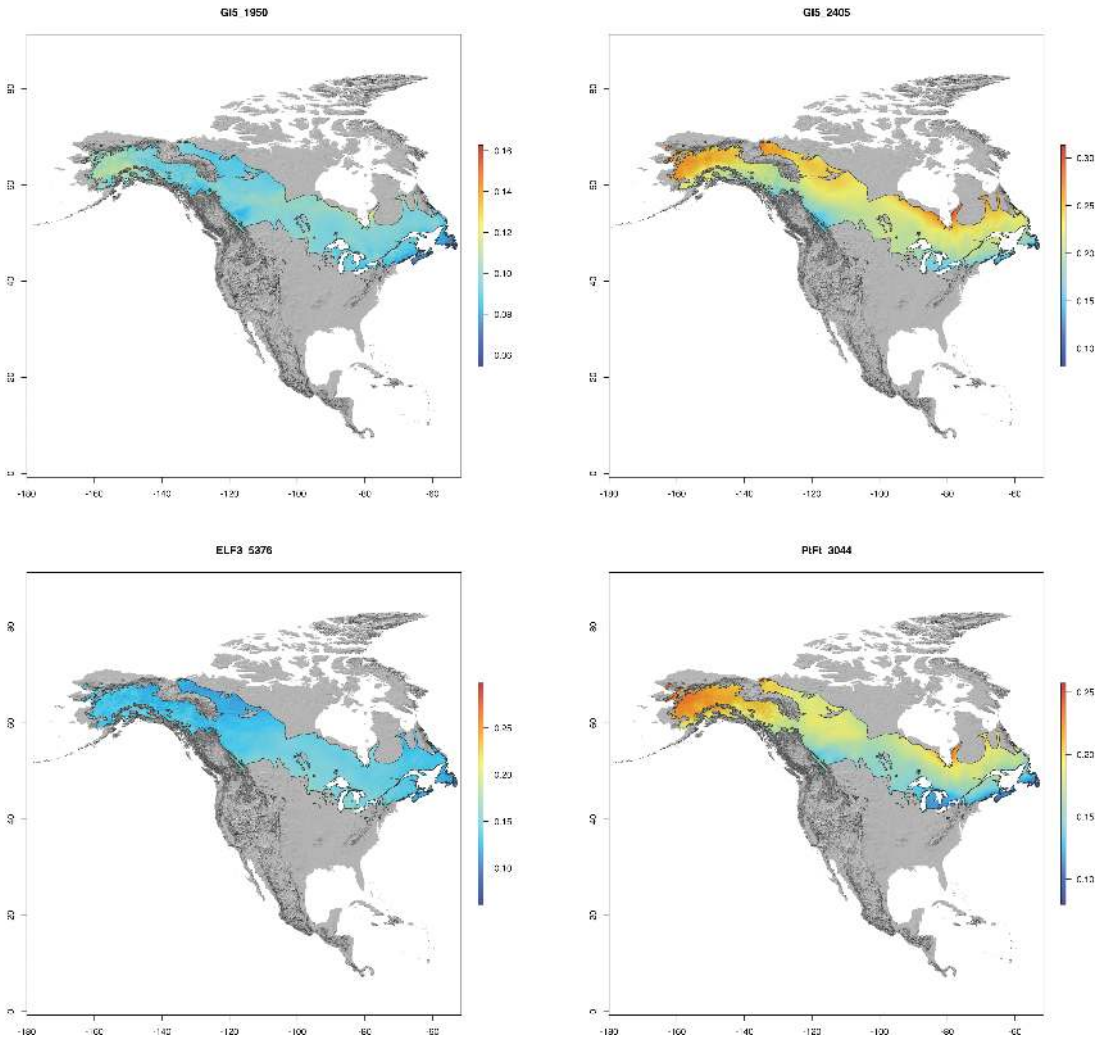

Figure 6. GF prediction of adaptive offset of local adaptation under future climate (2050) for 4 SNPs identified as significant outliers by all 3 local adaptation methods (Table 2). Offsets are based on modeled associations between climate and SNPs identified as being under local selection, with warmer colors indicating greater offset.

the significant loci identified by these 2 methods were corroborated by *Bayenv2* (Table 2). This probably reflects the influence of SNPs under local adaptation used to estimate the Ω matrix in *Bayenv2* in the absence of a neutral set of reference SNPs, thus leading to over-correction for population structure when testing for outliers. For example, no SNPs from *GI5* were in the top 5% of Bayes Factors from *Bayenv2* in the range core sample, yet we know from previous studies that included reference SNPs as a genomic control that *GI5* shows significant evidence of adaptive population structure and gene–environment association (Keller et al. 2012; Fitzpatrick and Keller 2015), as well as strong phenotypic associations with bud set and bud flush (Olson et al. 2013), 2 traits that are under strong local adaptation in *P. balsamifera* (Keller et al. 2011b). Nevertheless, it is probable that some SNPs that are unique to one analysis or range comparison are false positives, and thus interpretation of the significance of individual SNP loci must be made with caution. The general trend however is quite clear—the targets of local adaptation to climate in *P. balsamifera* have primarily been genes peripheral to the plant circadian clock (Supplementary Figures S1 and S2), with edge and core populations differing in the genetic architecture of this response (Figure 4, Table 2).

It is interesting to consider the molecular functions of *ELF3* and *GI5* as they relate to adaptive phenology and other circadian controlled traits. From work by our group and others in *Populus*, *GIGANTEA* is a strong candidate for locally adaptive phenology, based on elevated population structure, associations with temperature and precipitation, extended linkage disequilibrium, and phenotypic associations with bud flush and bud set (Rohde et al. 2011; Keller et al. 2012; Olson et al. 2013; Fitzpatrick and Keller 2015). *GI* is known to play diverse roles in a number of plant circadian responses in addition to phenology, including tolerance to heat and drought stress, germination, light signaling, and starch accumulation (Mishra and Panigrahi 2015). Thus, adaptive variants in *GI* may participate in any number of pleiotropic responses, but are especially likely in photoperiod-induced phenology (Sawa et al. 2007), consistent with the association of this genes with vegetative bud phenology (Rohde et al. 2011; Olson et al. 2013). Interestingly, *EARLY FLOWERING-3* also appears to regulate photoperiod-induced phenology as part of the evening complex of genes, and mutant *elf3* plants show photoperiod-independent flowering (Fowler 1999). Interestingly, *ELF3* is also known to interact directly with *GI*, and is known to regulate *GI* expression. This ability to modulate the expression of *GI* upstream

from the *CO/FT* regulon provides an additional target for selection to modify *GI* regulation and its pleiotropic consequences. Such regulation might be valuable in mountainous areas where the growing season may shrink rapidly with elevation, but the photoperiod does not. An intriguing follow-up study to this work would be to compare expression of *GI* and its downstream targets (*CO* and *FT*) using different 2-locus combinations of *ELF3* and *GI* to test if selection is maintaining an epistatic interaction between these genes that allows persistence in climatically variable habitats at southern latitudes.

We also showed that climate change is very likely to disrupt the SNP–environment association that exists for many genes, and is a major source of uncertainty when trying to predict responses of forest trees and other plants to climate change. Our estimation of adaptive offset based on projections of GF models to future climate revealed the most pronounced maladaptation may occur in the Pacific NW and interior Canada/Alaska (Figure 6), with the rear edge predicted to experience a less severe loss of location adaptation. Thus, while rear edge populations contain genetic diversity with putatively adaptive alleles not common elsewhere in the range, it does not appear that the climatic gradients to which they are adapted are the same gradients likely to shift under near-term climate warming. This may partially be a reflection of the influence of elevation on climate in the rear edge. Indeed, most rear edge populations show trends in *ELF3* frequency that scale with elevation and elevation-induced climate, which may provide some limited vertical refuge for trees as they attempt to respond to climate change.

Conclusions

Local adaptation in natural populations reflects the current and historical balance between gene flow and selection acting on alleles with location-dependent effects on fitness. Populations along the rear edge represent opportunities for local adaptation due to their occupancy of marginal environments, different levels of connectivity to other populations, and potential for unique variation not found elsewhere in the range. We confirm these predictions in balsam poplar, in which we found rear edge and core populations harbor divergent climate-adaptive alleles for genes in the flowering time pathway. Thus, rear edge populations should be targeted for germplasm conservation with the recognition that their unique evolutionary history may enable future conservation and breeding efforts aimed at mitigating climate change impacts on forest trees. Through spatially explicit analyses of turnover in the frequency of adaptive alleles across landscapes, studies such as ours can begin to identify the regions most at risk of loss of local adaptation in forests under changing environments.

Supplementary Material

Supplementary data are available at *Journal of Heredity* online.

Funding

This work was supported by National Science Foundation Plant Genome Research Program award (1461868 to S.R.K. and M.C.F.).

Acknowledgments

We are grateful to Lynda Delph and the American Genetics Association for organizing this symposium and special issue on genetics of local adaptation. We also appreciate the contributions of Raju Soolanayakanahally, Ron Zalesny, Andrew Elmore, Steven Guinn, and Matt Olson for tissue collections used for DNA extraction, and Peter Tiffin for contributions toward Sequenom iPlex genotyping development. Comments from 2 reviewers helped to improve the manuscript.

Data Availability

We have deposited the primary data underlying these analyses as follows: sampling locations and SNP genotypes: Dryad (doi: 10.5061/dryad.gp78p).

References

Böhlenius H, Huang T, Charbonnel-Campaa L, Brunner AM, Jansson S, Strauss SH, Nilsson O. 2006. CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science*. 312:1040–1043.

Bragg JG, Supple MA, Andrew RL, Borevitz JO. 2015. Genomic variation across landscapes: insights and applications. *New Phytol*. 207:953–967.

Bridle JR, Vines TH. 2007. Limits to evolution at range margins: when and why does adaptation fail? *Trends Ecol Evol*. 22:140–147.

Brunner AM, Evans LM, Hsu CY, Sheng X. 2014. Vernalization and the chilling requirement to exit bud dormancy: shared or separate regulation? *Front Plant Sci*. 5:732.

Chhatre VE, Evans LM, DiFazio SP, Keller SR. Forthcoming. Adaptive introgression and maintenance of a trispecies hybrid complex in range-edge populations of Populus. In press.

Dyer RJ. 2015. Population graphs and landscape genetics. *Annu Rev Ecol Evol Syst*. 46:327–342.

Dyer RJ, Nason JD. 2004. Population graphs: the graph theoretic shape of genetic structure. *Mol Ecol*. 13:1713–1727.

Dynesius M, Jansson R. 2000. Evolutionary consequences of changes in species' geographical distributions driven by Milankovitch climate oscillations. *Proc Natl Acad Sci U S A*. 97:9115–9120.

Ellis N, Smith SJ, Pitcher CR. 2012. Gradient forests: calculating importance gradients on physical predictors. *Ecology*. 93:156–168.

Evans LM, Slavov GT, Rodgers-Melnick E, Martin J, Ranjan P, Muchero W, Brunner AM, Schackwitz W, Gunter L, Chen JG, *et al.* 2014. Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. *Nat Genet*. 46:1089–1096.

Fitzpatrick MC, Keller SR. 2015. Ecological genomics meets community-level modelling of biodiversity: mapping the genomic landscape of current and future environmental adaptation (M Vellend, Ed.). *Ecol Lett*. 18:1–16.

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

Fowler S. 1999. GIGANTEA: a circadian clock-controlled gene that regulates photoperiodic flowering in Arabidopsis and encodes a protein with several possible membrane-spanning domains. *EMBO J*. 18:4679–4688.

Frichot E, Schoville SD, Bouchard G, François O. 2013. Testing for associations between loci and environmental gradients using latent factor mixed models. *Mol Biol Evol*. 30:1687–1699.

Günther T, Coop G. 2013. Robust identification of local adaptation from allele frequencies. *Genetics*. 195:205–220.

Hampe A, Jump AS. 2011. Climate relicts: past, present, future. *Annu Rev Ecol Evol Syst*. 42:313–333.

Hampe A, Petit RJ. 2005. Conserving biodiversity under climate change: the rear edge matters. *Ecol Lett*. 8:461–467.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A. 2005. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol*. 25:1965–1978.

Hoffmann AA, Sgrò CM. 2011. Climate change and evolutionary adaptation. *Nature*. 470:479–485.

Holliday JA, Zhou L, Bawa R, Zhang M, Oubida RW. 2016. Evidence for extensive parallelism but divergent genomic architecture of adaptation

along altitudinal and latitudinal gradients in *Populus trichocarpa*. *New Phytol*. 209:1240–1251.

Holt RD, Gomulkiewicz R, Naturalist TA, Mar N. 1997. How does immigration influence local adaptation? A reexamination of a familiar paradigm. *Am Nat*. 149:563–572.

Hsu CY, Adams JP, Kim H, No K, Ma C, Strauss SH, Drnevich J, Vandervelde L, Ellis JD, Rice BM, *et al*. 2011. FLOWERING LOCUS T duplication coordinates reproductive and vegetative growth in perennial poplar. *Proc Natl Acad Sci U S A*. 108:10756–10761.

Jombart T, Ahmed I. 2011. adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics*. 27:3070–3071.

Jombart T, Devillard S, Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet*. 11:94.

Kawecki T. 2008. Adpatation to marginal habitats. *Annu Rev Ecol Evol Syst*. 39:321–342.

Keenan K, Mcginnity P, Cross TF, Crozier WW, Prodöhl PA. 2013. DiveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol Evol*. 4:782–788.

Keller SR, Levsen N, Ingvarsson PK, Olson MS, Tiffin P. 2011a. Local selection across a latitudinal gradient shapes nucleotide diversity in balsam poplar, *Populus balsamifera* L. *Genetics*. 188:941–952.

Keller SR, Levsen N, Olson MS, Tiffin P. 2012. Local adaptation in the flowering-time gene network of balsam poplar, *Populus balsamifera* L. *Mol Biol Evol*. 29:3143–3152.

Keller SR, Olson MS, Silim S, William SA, Peter T. 2010. Genomic diversity, population structure, and migration following rapid range expansion in the Balsam poplar, *Populus balsamifera*. *Mol Ecol*. 19:1212–1226.

Keller SR, Soolanayakanahally RY, Guy RD, Silim SN, Olson MS, Tiffin P. 2011b. Climate-driven local adaptation of ecophysiology and phenology in balsam poplar, *Populus balsamifera* L. (Salicaceae). *Am J Bot*. 98:99–108.

Lê S, Josse J, Husson F. 2008. FactoMineR: an R package for multivariate analysis. *J Stat Softw*. 25:1–18.

Lepais O, Muller SD, Saad-Limam SB, Benslama M, Rhazi L, Belouahem-Abed D, Daoud-Bouattour A, Gammar AM, Ghrabi-Gammar Z, Bacles CFE. 2013. High genetic diversity and distinctiveness of rear-edge climate relicts maintained by ancient tetraploidisation for *Alnus glutinosa*. *PLoS One*. 8:e75029.

Levsen ND, Tiffin P, Olson MS. 2012. Pleistocene speciation in the genus *Populus* (salicaceae). *Syst Biol*. 61:401–412.

Lotterhos KE, Whitlock MC. 2014. Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Mol Ecol*. 23:2178–2192.

Ma XF, Hall D, 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.

McKown AD, Guy RD, Klápště J, Geraldes A, Friedmann M, Cronk QC, El-Kassaby YA, Mansfield SD, Douglas CJ. 2014. Geographical and environmental gradients shape phenotypic trait variation and genetic structure in *Populus trichocarpa*. *New Phytol*. 201:1263–1276.

Mishra P, Panigrahi KC. 2015. GIGANTEA—an emerging story. *Front Plant Sci*. 6:8.

Olson MS, Levsen N, Soolanayakanahally RY, Guy RD, Schroeder WR, Keller SR, Tiffin P. 2013. The adaptive potential of *Populus balsamifera* L. to phenology requirements in a warmer global climate. *Mol Ecol*. 22:1214–1230.

Petit RJ, Aguinagalde I, de Beaulieu JL, Bittkau C, Brewer S, Cheddadi R, Ennos R, Fineschi S, Grivet D, Lascoux M, *et al*. 2003. Glacial refugia: hotspots but not melting pots of genetic diversity. *Science*. 300:1563–1565.

Petit RJ, Hampe A. 2006. Some evolutionary consequences of being a tree. *Annu Rev Ecol Evol Syst*. 37:187–214.

R Core Team. 2014. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raj A, Stephens M, Pritchard JK. 2014. fastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics*. 197:573–589.

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

Rohde A, Storme V, Jorge V, Gaudet M, Vitacolonna N, Fabbrini F, Ruttink T, Zaina G, Marron N, Dillen S, *et al*. 2011. Bud set in poplar—genetic dissection of a complex trait in natural and hybrid populations. *New Phytol*. 189:106–121.

Savolainen O, Lascoux M, Merilä J. 2013. Ecological genomics of local adaptation. *Nat Rev Genet*. 14:807–820.

Sawa M, Nusinow DA, Kay SA, Imaizumi T. 2007. FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. *Science*. 318:261–265.

Sexton JP, McIntyre PJ, Angert AL, Rice KJ. 2009. Evolution and ecology of species range limits. *Annu Rev Ecol Evol Syst*. 40:415–436.

Sexton JP, Strauss SY, Rice KJ. 2011. Gene flow increases fitness at the warm edge of a species' range. *Proc Natl Acad Sci U S A*. 108:11704–11709.

Soolanayakanahally RY, Guy RD, Silim SN, Drewes EC, Schroeder WR. 2009. Enhanced assimilation rate and water use efficiency with latitude through increased photosynthetic capacity and internal conductance in balsam poplar (*Populus balsamifera* L.). *Plant Cell Environ*. 32:1821–1832.

Soolanayakanahally RY, Guy RD, Silim SN, Song M. 2013a. Timing of photoperiodic competency causes phenological mismatch in balsam poplar (*Populus balsamifera* L.). *Plant Cell Environ*. 36:116–127.

Soolanayakanahally RY, Guy RD, Silim SN, Song M. 2013b. Timing of photoperiodic competency causes phenological mismatch in balsam poplar (*Populus balsamifera* L.). *Plant Cell Environ*. 36:116–127.

Sork VL, Aitken SN, Dyer RJ, Legendre P, Neale DB. 2013. Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genet Genomes*. 9:901–911.

Wang L, Tiffin P, Olson MS. 2014. Timing for success: expression phenotype and local adaptation related to latitude in the boreal forest tree, *Populus balsamifera*. *Tree Genet Genomes*. 10:911–922.

Woolbright SA, Whitham TG, Gehring CA, Allan GJ, Bailey JK. 2014. Climate relicts and their associated communities as natural ecology and evolution laboratories. *Trends Ecol Evol*. 29:406–416.