MOLECULAR ECOLOGY

Molecular Ecology (2012) doi: 10.1111/mec.12067

The adaptive potential of *Populus balsamifera* L. to phenology requirements in a warmer global climate

MATTHEW S. OLSON,*† NICHOLAS LEVSEN,† RAJU Y. SOOLANAYAKANAHALLY,‡ ROBERT D. GUY,§ WILLIAM R. SCHROEDER,‡ STEPHEN R. KELLER¶ and PETER TIFFIN**

**Department of Biological Sciences, Texas Tech University, Lubbock, TX 79409, USA, †Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK 99709, USA, ‡Agroforestry Development Centre, Science and Technology Branch, Agriculture and Agri-Food Canada, Indian Head, SK, Canada S0G 2K0, §Department of Forest Sciences, University of British Columbia, Vancouver, BC, Canada V6T 1Z4, ¶Appalachian Laboratory, University of Maryland Center for Environmental Science, 301 Braddock Rd., Frostburg, MD 21532, USA, **Department of Plant Biology, University of Minnesota, 250 Biological Sciences Building, St. Paul, MN 55108, USA*

Abstract

The manner in which organisms adapt to climate change informs a broader understanding of the evolution of biodiversity as well as conservation and mitigation plans. We apply common garden and association mapping approaches to quantify genetic variance and identify loci affecting bud flush and bud set, traits that define a tree's season for height growth, in the boreal forest tree *Populus balsamifera* L. (balsam poplar). Using data from 478 genotypes grown in each of two common gardens, one near the southern edge and another near the northern edge of *P. balsamifera's* range, we found that broad-sense heritability for bud flush and bud set was generally high ($H^2 > 0.5$ in most cases), suggesting that abundant genetic variation exists for phenological response to changes in the length of the growing season. To identify the molecular genetic basis of this variation, we genotyped trees for 346 candidate single nucleotide polymorphisms (SNPs) from 27 candidate genes for the *CO*/*FT* pathway in poplar. Mixed-model analyses of variance identified SNPs in 10 genes to be associated with variation in either bud flush or bud set. Multiple SNPs within *FRIGIDA* were associated with bud flush, whereas multiple SNPs in *LEAFY* and *GIGANTEA 5* were associated with bud set. Although there was strong population structure in stem phenology, the geographic distribution of multilocus association SNP genotypes was widespread except at the most northern populations, indicating that geographic regions may harbour sufficient diversity in functional genes to facilitate adaption to future climatic conditions in many sites.

Keywords: association mapping, boreal forest tree, ecological genomics, heritability, population structure

Received 31 December 2011; revision received 13 August 2012; accepted 23 August 2012

Changing climate has repeatedly remoulded the range limits of temperate, boreal and arctic organisms (Umina *et al.* 2005; Saccheri *et al.* 2008), but the importance of the evolution of novel adaptive genotypes during these migrations remains controversial and perhaps underemphasized (Jump & Penuelas 2005). The role of adaptation is often overlooked because many species' fossil records appear to have tracked climate envelopes, indicating that species stay within preadapted climatic niche parameters as they migrate (Macdonald 1993; Williams *et al.* 2004; McLachlan *et al.* 2005). Nonetheless, although most species may not radically alter their niches during migration, novel genotypes or gene combinations may arise during adaptation to future climatic niches. Whether organisms respond to climate change by tracking the environment, by plastic adjustment to new environmental combinations or by adapting to

Correspondence: Matthew S. Olson, Fax: +1 806 742 2963; E-mail: matt.olson@ttu.edu

new environmental combinations influences the speed at which range edges may move and the ways that ecologists and evolutionary biologists view range shifts. For instance, if species track the environment or if only plastic adjustment is required, range shifts are primarily limited by the rate at which species can move. However, if novel adaptive genotypes are required, new combinations of standing genetic variation or new genetic mutations must be generated for successful range shifts, and the time required for these changes may be much longer than that required for species to simply move. From a management perspective, if novel genotypes are not required, conservationists and land managers can move genotypes to track their climatic niche; however, if genotypes adapted to future climate niches are uncommon or absent from contemporary populations, assisted migration that lacks breeding may be less successful.

Latitudinal clines in phenological variation in forest trees provide clear study systems to assess how current genotypes respond to changing climatic niches. These clines have been known to have a genetic basis since the mid-20th century (Pauley & Perry 1954; Farmer & Reinholt 1986) and are adaptive in several species (Hall *et al.* 2007; Holliday *et al.* 2010b; Keller *et al.* 2011b; Savolainen *et al.* 2011). The proper timing of spring growth and fall dormancy maximizes a tree's growing season, while minimizing frost damage in the spring and fall. For North American species at high latitudes, phenological clines must be regenerated after each ice age. Thus, present-day clines must have been established since the most recent North American glacial maxima (Breen *et al.* 2012; Levsen *et al.* 2012) and will most likely extend north in response to ongoing climate warming (Davis & Shaw 2001; IPCC 2007; Aitken *et al.* 2008).

We currently know little about either the genetic basis of clinal variation for phenological traits in trees, the patterns of growth and dormancy response to environmental cues across latitude, or the mechanisms limiting across-latitude migrations. Although responses to climate change are thought to result largely from tracking temperature (IPCC 2007; Soja *et al.* 2007), these predictions often fail to consider that temperature is only one aspect of the biotic and abiotic environment that determines a species range (Jackson & Overpeck 2000). For trees, perhaps the most important aspect of the abiotic environment that will stay constant, even under the most radical of climate change models, is photoperiod. Trees cue on critical photoperiod to predict the coming onset of fall, cease height growth and set terminal buds (Howe *et al.* 2000; Horvath *et al.* 2003); thus, photoperiodic cueing is an essential component of local adaptation. Although cold temperature and longer summer photoperiod are correlated across latitude, climate warming is likely to alter the correspondence between temperature and photoperiod at a particular latitude, decreasing the efficacy of preadaptation based on existing temperature–photoperiod regimes.

Studies of the genetic basis of variation in seasonal phenology (bud flush, bud set and leaf senescence), and the distribution of alleles influencing these phenotypes across latitude are necessary for understanding and predicting the responses of trees to climate change. Quantitative and association genetics analyses conducted in ecologically appropriate environments offer opportunities for identifying traits and genetic regions important for local adaptation across climate gradients (Eckert *et al.* 2009; Holliday *et al.* 2010a; Fournier-Level *et al.* 2011; Savolainen *et al.* 2011). In both annual (e.g. *Arabidopsis*, Koornneef *et al.* 1991) and perennial plants (e.g. *Populus*, Bohlenius *et al.* 2006), components of the *CONSTANS/FLOWERING LOCUS T* (*CO/FT*) regulatory module, which includes photoreceptors, clock genes and genes in the vernalization pathway, have been identified as candidates for controlling genetic variation in phenology in natural populations (Samach & Coupland 2000; Bohlenius *et al.* 2006; Ballerini & Kramer 2011). For instance, in *Arabidopsis thaliana* and *Arabidopsis lyrata*, natural variation in *FRIGIDA* (*FRI*) affects the vernalization requirement underlying clines in flowering time (Stinchcombe *et al.* 2004; Riihimaki *et al.* 2005); in natural populations of *Populus tremula*, association mapping identified a relationship between variation in *phytochrome B1* (*PHYB1*) and bud set (Ingvarsson *et al.* 2008); and an integrated analysis of QTL from natural and hybrid mapping populations of poplars identified QTL near *FT* and *GIGANTEA* (*GI*) (Rohde *et al.* 2011b).

Here, we use a collection of 478 genotypes sampled from throughout the current range of balsam poplar (*Populus balsamifera*), a foundation tree species of the North American boreal forest, to investigate the quantitative genetics underlying spring vegetative bud flush and fall bud set, the traits that define the season during which trees increase in height. We grew genotypes in each of two common garden experiments. One garden, located near the northern edge of the species range (Fairbanks, Alaska), allowed us to examine potential evolutionary responses at the leading edge of a potential range shift that may be caused by a warming climate. The second garden, located near the southern limit of the species range (Indian Head, Saskatchewan), allowed us to examine responses at the lagging edge of potential range shifts. In each garden, we planted replicated genotypes originating from throughout the range of *P. balsamifera*, including populations that currently grow north and south of both gardens (Fig. 1). We

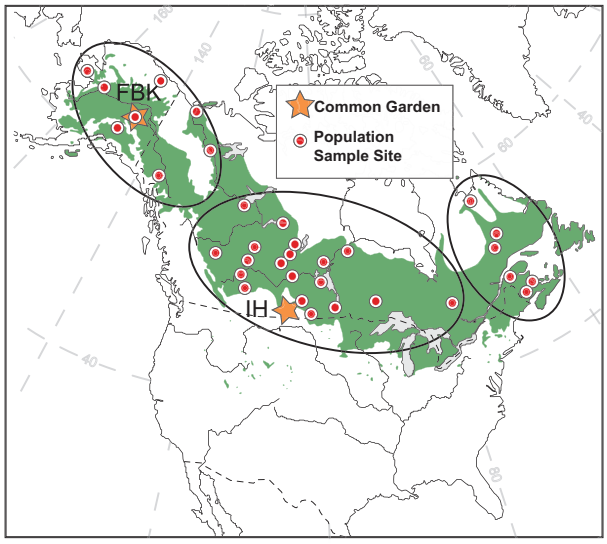

Fig. 1 Locations of genotype collection sites and common gardens of *Populus balsamifera* used for heritability estimates and association mapping. Red dots represent 33 populations from which tree cuttings were collected for establishment in two common gardens (orange stars) in Fairbanks, Alaska (FBK), and Indian Head Saskatchewan, Canada (IH). Green shading indicates the complete range of *P. balsamifera*. Ellipses surrounding populations indicate regional demes identified using a Structure analysis by Keller *et al.* (2010a).

address three general questions. First, because the response to selection is related to standing genetic variance (Fisher 1930), we estimated heritability for bud flush and bud set and asked whether it differed for trees grown at the southern and northern edges of the range. Second, we investigated whether allelic variation in genes from the *CO*/*FT* regulatory module underlies natural variation in bud flush and bud set in balsam poplar, and whether the same or different genes were identified depending on whether the traits were measured in the southern and northern garden. Finally, we asked how alleles associated with timing of bud flush and bud set are partitioned across populations and latitude in order to make predictions about the potential for adaptive response in phenology during climate change.

Methods

Common gardens

Dormant stem cuttings were collected from trees in natural populations during the winter of 2005–2006 and rooted in a glasshouse at the Agroforestry Development Centre, Indian Head, Saskatchewan, Canada (Soolanayakanahally *et al.* 2009). Once the plants were 30 cm tall, they were moved outside and hardened off. In the spring of 2007, the Indian Head common garden was established (Fig. 1, 50.3°N 103.4°W, hereafter IH). In the winter of 2009, dormant stem cuttings taken from trees growing in IH and three additional populations from Alaska (Galena, Cottonwood Creek and Nome) were rooted in the Institute of Arctic Biology greenhouse in Fairbanks, AK, and that summer planted in a fallow field on the campus of the University of Alaska Fairbanks (Fig. 1, 64.8°N 147.7°W, hereafter FBK). Cuttings from two northern populations, Fairbanks and Inuvik, were added to FBK in June of 2010 using cuttings from IH.

Across both gardens, 478 genotypes from 33 populations were grown and sampled; however, because of differential mortality and planting in the two gardens, the FBK garden data set included 470 genotypes from 33 populations, whereas the IH data set included 433 genotypes from 30 populations (Fig. 1). Each garden was split into five blocks. Within each block, one clone from each of 11–15 individuals from each population was planted (for a total of up to five clonal replicates of each genotype). In both gardens, trees from the same population were planted adjacent to one another in a 3×5 grid, with individual trees separated by 3 m in IH and 2.5 m in FBK. Positions of genotypes were randomized within population grids, and positions of populations were randomized within blocks. Both IH and FBK gardens were tilled and treated with glyphosphate herbicide 3+ weeks prior to planting.

Traits

Data on the date of bud flush and bud set on all trees in both gardens were collected in 2010. Bud flush was scored every 2–3 days during the spring as the date at which leaves began to emerge from the bud. Because bud flush is controlled primarily by the accumulation of warm temperatures in the spring (Thornwaite 1948; Paus *et al.* 1986; Howe *et al.* 1999; Rohde *et al.* 2011a), the ordinal dates of bud flush were converted to cumulative growing degree days (cGDD) based on local temperature records in Fairbanks and Indian Head. cGDD was calculated as the accumulated sum of growing degree days (GDD) since January 1 to the ordinal date of bud flush. GDD for each day were calculated as $(T_{max} - T_{min})/2 - T_{base}$, where T_{max} and T_{min} were the daily maximum and minimum temperatures and T_{base} was 0 °C. Bud flush for three Alaska populations (Galena, Cottonwood Creek and Nome) was scored only at FBK and leaf flush in the Fairbanks and Inuvik populations was scored only at IH, because they were planted a year later than other populations at FBK.

Bud set was defined as the date when a large terminal bud with bud scales formed at the end of the terminal

shoot and was scored twice a week throughout the summer at IH and starting in July at FBK. Some trees exhibited lammasing (Kaya *et al.* 1994), a characteristic when buds form and break again at a later date. For these trees, we defined bud set as the last date that a terminal bud set for a particular tree. Because bud set is expected to be controlled primarily by critical photoperiod (Pauley & Perry 1954; Howe *et al.* 1999; but see Rohde *et al.* 2011a), the ordinal dates of bud set were converted to day length (DL) at the time of bud set based on the local latitudes of the two gardens using the CBM model (Forsythe *et al.* 1995) and assuming sunrise and sunset occurred when the centre of the sun was even with the horizon (p = 0 in eq. 3, Forsythe *et al.* 1995). Bud set for three Alaska populations (Galena, Cottonwood Creek and Nome) was scored only at FBK.

Genotyping

Two sets of single nucleotide polymorphisms (SNPs; defined herein as also including a few indel variants), a reference set and a candidate set previously used for other analyses (Keller *et al.* 2010a,b, 2012; Olson *et al.* 2010), were used for the association analysis. The 412 reference SNPs were developed from a 15-individual discovery panel of 590 different unlinked gene fragments (Olson *et al.* 2010) and assayed using a Sequenom iPLEX system (Keller *et al.* 2010a). The 346 candidate SNPs were developed from a 24-individual SNP discovery panel of 27 candidate genes in the *CO/FT* pathway (Keller *et al.* 2011a, 2012). From this panel, SNPs were screened for all individuals using SEQUENOM iPLEX or Sanger sequencing.

Analysis

Variance components for population and genotype within population, block and residual effects were calculated separately for each trait and each garden using PROC MIXED in SAS 9.2 (Table 1, SAS Institute 2009). Population and genotype(population) variance components were summed to calculate total genetic variance. Broad-sense heritability (H^2) was estimated as $\hat{H}^2 = \hat{\sigma}_g^2/[\hat{\sigma}_e^2 + \hat{\sigma}_g^2]$ where $\hat{\sigma}_g^2$ is the genotypic variance component and $\hat{\sigma}_e^2$ is the environmental variance (block + residual). Statistical significance of variance components was calculated using F tests constructed using the Method = type3 option in PROC MIXED.

Genotype by environment (G × E) effects were analysed using mixed-model analyses in PROC MIXED as

$$y_{ijk} = \mu + G_i + L_j + GL_{ij} + B_{k(j)} + \epsilon_{ijk}$$

where y_{ijk} is the observation of the ith genotype (G_i) in the jth garden location (L_j; Fairbanks or Indian Head); GL_{ij} is the ith genotype by jth garden interaction, in the kth block [$B_{k(j)}$] within gardens; μ is the overall mean; and ε_{ijk} is the error term. The genotype and block were treated as a random effects, and the garden location and GL_{ij} interaction were treated as fixed effects. Statistical significance of genotype by garden location effects was tested by comparing the −2 log-likelihood of full models with all effects and reduced models that did not include the genotype by garden interactions. Because heteroscedasticity across gardens can generate G × E, we also ran analyses values for each trait (bud flush or bud set) that were transformed so that the distribution had zero mean and unit variance.

Table 1 Hierarchical variance components and heritabilities for all trait and garden combinations

Trait	Garden	Variance component			*H*†
Growing season	IH	Vg	Total	599.9	0.85
			Pop	521.9**	0.74
			Geno(Pop)	78.1**	0.10
		Ve‡		105.3	
Growing season	FBK	Vg	Total	64.9	0.30
			Pop	45.4**	0.21
			Geno(Pop)	19.5**	0.09
		Ve‡			
Bud flush†	IH	Vg	Total	969.7	0.81
			Pop	656.2**	0.55
			Geno(Pop)	313.5**	0.26
		Ve		223.9	
Bud flush	FBK	Vg	Total	2540.9	0.47
			Pop	2026.8**	0.37
			Geno(Pop)	514.1**	0.09
		Ve		2905.7	
Bud set§	IH	Vg	Total	1.3	0.83
			Pop	1.1**	0.72
			Geno(Pop)	0.2**	0.10
		Ve			0.27
Bud Set	FBK	Vg	Total	1.5	0.50
			Pop	1.4**	0.44
			Geno(Pop)	0.2**	0.06
		Ve		1.5	

F tests from type 3 analysis of variance were used to determine the significance of variance components. Tests were constructed using the Method = option in PROC MIXED.
†Measured in cumulative growing degree days prior to date of bud flush
‡Ve includes block effects within each garden.
§Measured as the day length on the date of bud set.
*$P < 0.05$, **$P < 0.001$.

Association mapping

We employed a mixed linear model analysis implemented in TASSEL (Yu *et al.* 2006) to assess the significance of associations between candidate SNPs and

traits (bud flush and bud set) while accounting for previously identified population structure and kinship in our mapping population (Keller *et al.* 2010a). Block effects were not incorporated into these analyses because they were nonsignificant in both gardens ($P > 0.20$). Instead, a single value was used to represent the trait for each genotype in each garden and calculated as the mean of cGDD required at bud flush or the mean DL at bud set across replicates of each genotype. The population structure matrix (**Q**) was developed from a *structure* analysis (Pritchard *et al.* 2000) of the 412 SNP reference loci (Keller *et al.* 2010a,b) that identified three statistically supported partitions, which roughly defined an eastern deme, a central deme and a northern deme (Fig. 1 in Keller *et al.* 2010a); we refer to these demes as 'regions' herein. We compared the influence of two relatedness matrices for association mapping: (i) a phylogenetic kinship (**PK)** matrix calculated using EMMA (Kang *et al.* 2010) and (ii) a coancestry matrix (**A**) (also called additive genetic relatedness matrix), which was calculated as twice the kinship ($\mathbf{A} = 2\mathbf{K}$) (tassel users group http://groups.google.com/group/tassel, and Dr. Z. Zhang, personal communication). The kinship matrix (**K**) was calculated with SPAGeDi (Vekemans & Hardy 2004) using the method of Loiselle *et al.* (1995). Because poplar is dioecious and thus obligately outbred, we set the diagonal of the **K** matrix to 0.5 prior to calculation of the **A** matrix (Falconer & Mackay 1996). Individuals with negative kinship were assumed to be unrelated, and values were set to zero. The matrices differ in that the **A** matrix has a stronger theoretical relationship to kinship than the **PK** matrix, but the **PK** matrix is constrained to be positive semidefinite, a statistical property that allows calculation using SAS REML and EMMA algorithms (Kang *et al.* 2010). The results of the analyses using the two different matrices were qualitatively similar. We report only the results using the **A** matrix because they were slightly more conservative (Figs S1 and S2, Supporting information). Analyses that included both the **Q** and **A** matrices proved to correct for underlying structure better than analyses with only the **Q** matrix or only the **A** matrix (Fig. S1, Supporting information). Finally, we filtered loci for minor allele frequency (MAF) >1.5% prior to analysis. This liberal cut-off was chosen because under Hardy–Weinberg equilibrium 1% MAF generates >12 heterozygotes from a sample size of 430, a sufficiently large sample size for robust hypothesis tests across at least two genotypes.

In a further attempt to eliminate false positives, we ran all association analyses using the mixed model ($\mathbf{Q} + \mathbf{A}$) on our set of 412 unlinked reference SNPs compiled from a random suite of genes. The assumption is that because these are not candidate genes, few reference SNPs should show a significant association with our traits, and those that do are false positives. The proportion of reference SNPs with significantly greater *P*-values than our candidate genes was used as an empirical false discovery rate (FDR_{EMP}). A second false discovery rate (FDR_Q) was calculated using the q-value method (Storey 2002).

Significance of SNP-by-garden ($G \times E$) interactions could not be tested within the TASSEL framework, so SAS code for the mixed model (provided at http://www.maizegenetics.net/unified-mixed-model) was adjusted to accommodate testing the interactions. Because the **A** matrix generated by SPAGeDi was not positive semidefinite as required by SAS PROC MIXED, the **PK** matrix was substituted for these analyses. Because the random-effect **PK** matrix could not be implemented within a $G \times E$ framework, within each garden and for each trait we calculated the residuals from a model including both **PK** and **Q** matrices, but without fitting the SNP effect. Residuals were then combined across gardens for each trait and used as the dependent variable for testing the SNP, garden and SNP-by-garden effects. Bud flush and bud set traits were Studentized (zero mean, unit variance) prior to calculating the residuals to avoid false positives resulting from unequal variances across gardens.

To estimate the proportion of among-multilocus-genotype variation that could be explained with the candidate SNPs identified from the association analyses, we used least angle regression model (LAR in SAS PROC GLMSELECT, SAS Institute 2009) with predicted SNP genotype trait mean as the response variable. The best models were identified as those with minimum Schwarz Bayesian information criteria values, which were more conservative than Akaike information criteria for some analyses. The SNPs used as explanatory variables were chosen because the association analyses identified them as associated with phenotypic variance, so the expected proportion of variance explained is greater than zero—even in the absence of true association. Therefore, to estimate null expectations for the LAR analyses, we calculated linear regressions onto observed trait values using the most highly significant reference genes with genotypic means from each trait as the response variable. For each trait, we used the same number of reference SNPs as were identified in the LAR for candidate SNPs.

Pairwise gametic phase linkage disequilibrium (LD) between candidate SNPs was calculated as r^2 using dipdat (Hudson 2001), which uses Hill's (Hill 1974) method to estimate LD from diploid data, and was plotted using the LDHeatmap plugin for R (Graham *et al.* 2006).

Results

Patterns of phenotypic response

The season length for height growth (SLHG), calculated as the average number of days between bud flush and bud set, ranged among genotypes from 54 to 155 days in the Indian Head (IH) garden and from 61 to 124 days in the Fairbanks (FBK) garden (Fig. 2A, Fig. S3A, Supporting information). This is compared with an average of 123 frost-free days in Indian Head and 112 days in Fairbanks between 1960 and 1994, the period when records are available for both sites. Genetic differences among trees accounted for most of the variance in SLHG in IH (85%) and nearly one-third of the variance for SLHG in FBK (30%; Table 1). In both gardens, *c.* 10% of the genetic variation in SLHG was found within populations. Genotypes from the south exhibited longer height growing seasons than those from the north in both gardens, and a steeper relationship between latitude and SLHG in IH than in FBK generated a significant G × E effect (Fig. 2A, Table S1, Supporting information). Interestingly, northern genotypes exhibited longer height growing seasons in FBK than in IH (Fig. 2A), even though there are *c.* 11 more frost-free days at IH, an effect we can attribute to the inability of northern genotypes to grow in the short southern photoperiod.

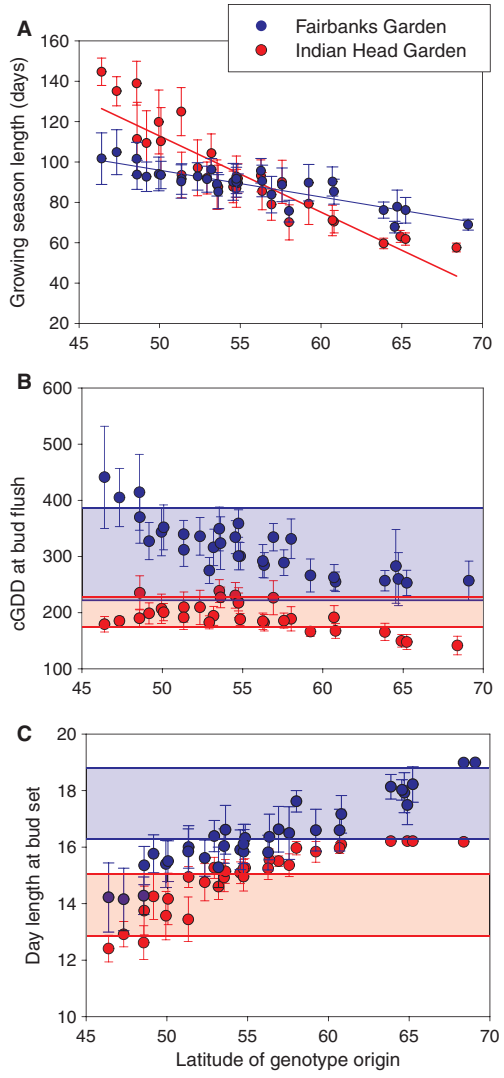

Fig. 2 Phenological patterns exhibited by genotypes in the Fairbanks (blue) and Indian Head gardens (red). Filled circles represent genotype means within populations (±1 SD). Blue shading represents the phenotype range of individuals from the nearest population to Fairbanks when growing in the Fairbanks garden. Red shading represents the phenotype range of individuals from the nearest population to Indian Head when growing in the Indian Head garden. (A) Number of days between bud flush and bud set (growing season length). (B) Cumulative growing degree days at each garden at the time of bud flush. (C) Day length at bud set.

The start of a tree's growing season is defined by bud flush, which is triggered primarily by the accumulation of warm temperatures in the spring (Thornwaite 1948; Paus *et al.* 1986; Howe *et al.* 1999; Menzel *et al.* 2006). Bud flush in *Populus balsamifera* started 29 calendar days and 138 cGDD earlier in the IH than in FBK (Fig. 2B, Figs S3A and S4B, Supporting information). In IH, genotypes from mid-latitude populations flushed buds the latest—an average of 10–15 days later than genotypes from above 65°N, which flushed earliest (Fig. 2B), whereas in FBK, trees from low latitudes flushed leaves the latest—an average of 10–12 days later than those from the far north (Fig. S4B, Supporting Information). The genetic correlation in cGDD at bud flush across gardens was 0.35 (Pearson's, 95% confidence limits 0.26 –0.44). Because the lowest-latitude trees originated primarily from the Atlantic provinces of Canada, which have climates that are more strongly affected by maritime influences than are interior Canadian provinces (Fig. 1), it is possible that these genotypes respond differently than interior and northern genotypes to environmental cues to flush buds. Genetic differences among trees accounted for most of the variation in the cGDD required for bud flush (Table 1). In the IH environment, 32% (=0.26/0.81) of the genetic variance was found within populations, but in FBK it was only 19%.

Assuming that local genotypes are locally adapted at each garden, their range of phenotypes can be used as a proxy for evaluating the latitudinal distance across which individuals could be moved and still exhibit seasonality similar to the local genotypes. In the IH, cGDD required for bud flush among trees from the nearest local population (by latitude) ranged from 186 to 225 (Fig. 2B); only populations above 60°N lat. expressed mean timing of bud flush outside that range.

In FBK, only the most southeastern populations (below 48°N lat.) expressed bud flush outside the range expressed by local genotypes (Fig. 2B).

Bud set of balsam poplars started approximately 1 month earlier and ended 10 days later at IH than it did at FBK (Fig. 2C, Figs S3B and S4C, Supporting information). The genetic correlation in DL at bud set across gardens was 0.75 (Pearson's, 95% confidence limits 0.71–0.79). Genetic differences among trees accounted for most of the variation in the DL at which buds set (Table 1), with *c*. 12% of genetic variance for timing of bud set present within populations in both gardens (Table 1). In both gardens, genotypes that originated from high latitudes set bud under longer photoperiods (earlier) than genotypes originating from low latitudes (Fig. 2C). At IH, northern genotypes set bud *c*. 90 days earlier than southern genotypes, whereas at FBK northern genotypes set bud 35–40 days earlier than southern genotypes (Fig. S4C, Supporting information). Interestingly, most genotypes collected from above 57°N set bud near the summer solstice in IH (16-hr daylight), which was likely as soon as they became competent to set bud. This extremely early bud set resulted in shorter growing seasons for northern genotypes in the southern (IH) than in the northern (FBK) garden (Fig. 2A).

Local genotypes set bud across a large range of critical photoperiods in both gardens (Fig. 2C). At IH, trees that originated from a nearby population set bud at DLs between 13 and 15 h, and average bud set in populations above 55°N fell outside the range of critical photoperiods expressed by local genotypes. At FBK, local genotypes set bud at DLs between 16.3 and 18.8 h, and average bud set in populations below 55°N fell outside the range of critical photoperiods expressed by local genotypes (Fig. 2C).

Association loci for bud flush and bud set

A discovery panel was used to identify 346 SNPs in 27 genes with homology to *CO/FT* pathway genes in *Arabidopsis* (Mouradov *et al*. 2002; Simpson & Dean 2002; Ehrenreich *et al*. 2009). Of these SNPs, 301 had MAF > 1.5% and were used for association mapping. One SNP from *GIGANTEA 5* (*GI5*) was associated with timing of bud flush in IH, and nine SNPs from five genes were associated with bud flush in FBK (Table 2). The *FRIGIDA* (*FRI*) gene accounted for almost half of the SNPs associated with bud flush in FBK (Table 2). In IH, the predicted average allelic effect of the *GI5_5271* minor allele on cGDD required for bud flush was −18.5 cGDD, and in FBK, all average effects of the minor alleles resulted in later bud flush, varying from 3.8 to 30.1 cGDD (Table 3). Three of the 12 SNPs, *ELF3_90*, *CRY1.2_2106* and *GI5_92*, also exhibited significant SNP × garden effects (Fig. 3), with larger phenotypic differences in FBK than in IH (Fig. 3).

A larger suite of associated SNPs was identified for bud set than for bud flush; 19 SNPs from eight genes were associated with bud set in IH, with over 1/3 of those from the *LEAFY* (*LFY*) gene, whereas 11 SNPs from four genes were significantly associated with bud set in FBK, with over half of those from *GI5* (Table 4). In IH, predicted minor allele average effects of significant SNPs on critical photoperiod for bud set varied from between −0.66 and 0.61 h, whereas in FBK, predicted average effects varied from −0.59 to 0.4 h

Table 2 SNPs associated with bud flush in the Indian Head and Fairbanks gardens

Garden	SNP symbol*	SNP†	MAF	*F*	*P*‡	R^2	R^2/V_G§	FDR_{EMP}¶	FDR_Q
Indian Head	***GI5_5271***	Indel+/− (I)	0.08	9.36	0.0001	0.04	0.09	0.008	0.0283
Fairbanks	***ELF3_90***	A/G (S)	0.07	15.92	2.3×10^{-7}	0.07	0.18	<0.002	6.1×10^{-5}
	CRY1.2_2106	A/G (S)	0.02	22.49	3.0×10^{-6}	0.05	0.01	0.0024	0.0004
	FRI_1780	G/T (I)	0.14	8.16	0.0003	0.04	0.09	0.0024	0.0123
	FRI_954	T/C (S)	0.17	8.10	0.0004	0.04	0.08	0.0024	0.0123
	GI5_92	T/C (I)	0.04	8.08	0.0004	0.04	0.10	0.0024	0.0123
	CKB34_452	C/T (S)	0.19	7.32	0.0008	0.04	0.08	0.0024	0.0217
	FRI_2845	A/G (I)	0.15	7.26	0.0008	0.03	0.08	0.0024	0.0217
	FRI_505	G/T (N) Ala:Ser	0.15	7.09	0.0009	0.03	0.07	0.0024	0.0232
	ELF3_1340	C/T (N) Pro:Ser	0.05	6.91	0.0011	0.03	0.08	0.0024	0.0254

MAF, minor allele frequency; SNP, single nucleotide polymorphisms.

*SNP named as Gene_position. SNP names in boldface were also significant at $FDR_Q < 0.05$ using the **A** matrix generated by EMMA;

†Major allele/minor allele. N, nonsynonymous; S, synonymous; I, intron. Amino acid changes provided for nonsynonymous SNPs.

‡Bonferroni critical value cut-off = 0.00017.

§R^2/V_G is an estimate of the proportion of genetic variation that is accounted for by the SNP.

¶An $FDR_{EMP} < 0.002$ means that none of the reference genes had *P*-values as small or smaller that the candidate–trait association.

Table 3 Predicted allelic effects on the cumulative growing degree days at initiation of bud flush from significant SNP–trait pairs calculated from adjusted means generated by the TASSEL model

Garden	SNP symbol*	2*a*†	*d*‡	*d*/*a*	α§
Indian Head	*GI5_5271*	31.9	−3.5	−0.2	−18.5
Fairbanks	*ELF3_90*	−176.1	−66.7	−0.8	28.3
	CRY1.2_2106	N/A¶	N/A¶	N/A¶	N/A¶
	FRI_1780	11.6	28.9	45.0	12.6
	FRI_954	−0.1	25.2	970.8	13.2
	GI5_92	51.8	56.1	2.2	23.8
	CKB34_452	−40.6	2.7	0.1	3.8
	FRI_2845	7.5	26.1	7.0	12.3
	FRI_505	10.2	26.7	5.2	11.5
	ELF3_1340	−50.7	7.2	0.3	30.1

SNP, single nucleotide polymorphisms.
*Number after gene indicates the position of SNP from the start of the gene.
†Calculated as the difference between the phenotype of major allele homozygote minus the minor allele homozygote.
‡Calculated as the difference between the phenotypes of the heterozygote and the average between the homozygotes.
§the average effect of the minor allele calculated as in Falconer & Mackay (1996).
¶Statistics undefined when one homozygote was not sampled.

(Table 5). No SNP × garden effects were significant for bud set (all $P < 0.05$).

LD of SNPs (≥10% MAF) within genes averaged $r^2 = 0.31$. Of the five genes (*CKB34*, *ELF*, *FRI*, *GI5* and *LFY*) with multiple SNPs associated with either bud flush or bud set, only *FRI* and *GI5* exhibited above average LD across all within-gene SNPs [$r^2 = 0.37$ and 0.69, respectively; Fig. S5 (Supporting information) shows LD heatmaps for each gene with significant association SNPs]. LD among SNPs may generate false-positive associations that are difficult to disentangle. Moreover, because each SNP is tested independently in the TASSEL model, the cumulative proportion of trait variance accounted for by SNPs cannot be estimated directly from these analyses. We addressed these factors by applying a LAR to identify the smallest set of significant association SNPs that best explains phenotypic variation for each trait. The best-fit sets of SNPs explained 11% and 30% of the observed variance in bud flush at IH and FBK, respectively, and 23% and 39% of the observed variance in bud set at IH and FBK, respectively (Fig. 4; Table 6). Because some of the variance from the LAR may be attributable to correlation with the trait due to coancestry or population structure, for comparison we calculated the proportion of variance from the genotypic means predicted by an equivalent

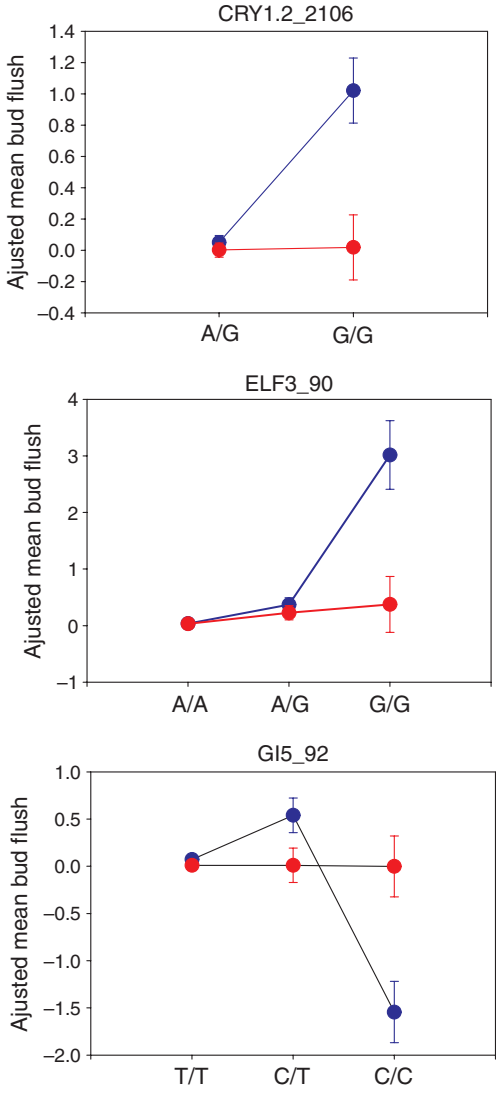

Fig. 3 Candidate loci that exhibited significant genotype by environment (SNP × Garden) interactions. Circles represent means for each genotype (±1 SD) growing in the Fairbanks (blue) or Indian Head (red) Garden. Bud flush was adjusted for population structure and coancestry prior to testing for SNP × garden effects. SNP, single nucleotide polymorphisms.

number of SNPs from our reference panel for each trait in each garden. The top reference SNPs explained 1% and 20% of the observed variance in bud flush at IH and FBK, respectively, and 36% and 12% of the observed variance in bud set at IH and FBK, respectively. This suggests that the proportion of variance explained by the candidate multilocus genotypes identified by LAR may be closer to the difference between these estimates, or *c*. 10% for bud flush in IH, *c*. 3% for bud flush in FBK, *c*. 27% for bud set in FBK and perhaps a very small proportion of variance of bud set in IH. Bud flush and bud set as predicted based on the

Table 4 Single nucleotide polymorphisms associated with bud set in the Indian Head and Fairbanks gardens

Garden	SNP symbol*	SNP†	MAF	*F*	*P*‡	R^2	$R^2 \times V_G$§	FDR_{EMP}¶	FDR_Q
Indian Head	***ELF3_1340***	C/T (N) Pro:Ser	0.05	13.27	2.7×10^{-6}	0.05	0.11	<0.002	0.0005
	ELF3_90	A/G (S)	0.07	11.66	1.2×10^{-5}	0.05	0.09	0.0024	0.0011
	LFY_1335	A/G (S)	0.38	8.38	0.0003	0.04	0.06	0.0024	0.0164
	LFY_2277	C/T (S)	0.12	7.76	0.0005	0.03	0.06	0.0024	0.017
	ELF3_407	G/A (I)	0.06	7.58	0.0006	0.03	0.06	0.0024	0.017
	LFY_1714	A/G (I)	0.49	7.44	0.0007	0.03	0.06	0.0024	0.017
	LFY_1326	C/G (S)	0.38	6.99	0.001	0.03	0.06	0.0048	0.0207
	GI5_268	A/G (I)	0.08	6.54	0.0016	0.03	0.05	0.0073	0.0261
	LFY_2996	C/T (3′U)	0.12	6.51	0.0017	0.03	0.06	0.0097	0.0261
	CKB34_3513	C/T (3′U)	0.04	6.45	0.0017	0.03	0.05	0.0097	0.0261
	LFY_1120	C/T (N) Pro:Ser	0.28	6.29	0.002	0.03	0.05	0.0097	0.0282
	LFY_1212	A/C (S)	0.41	6.15	0.0024	0.03	0.05	0.0097	0.0303
	GI5_9551	C/G (N) Gln:His	0.44	6.04	0.0026	0.03	0.05	0.0146	0.0312
	CRY1.2_2106	A/G (S)	0.02	9.06	0.0028	0.02	0.04	0.0146	0.0312
	GI2_8862	A/T (S)	0.07	5.91	0.0029	0.03	0.05	0.0146	0.0312
	GI5_198	C/T (I)	0.44	5.52	0.0043	0.02	0.05	0.0170	0.0433
	HY2.1_1961	C/A (I)	0.02	5.26	0.0055	0.02	0.04	0.0243	0.0485
	HY1.1_529	G/T (I)	0.41	5.22	0.0058	0.02	0.05	0.0243	0.0485
	CKB34_452	C/T (S)	0.19	5.20	0.0059	0.02	0.04	0.0243	0.0485
Fairbanks	***GI5_5271***	Indel +/− (I)	0.08	16.93	8.5×10^{-8}	0.06	0.15	<0.002	6.0×10^{-5}
	ABi1D_1595	G/T (I)	0.13	11.35	1.6×10^{-5}	0.04	0.10	<0.002	0.0004
	GI5_9585	T/C (N) Leu:Phe	0.45	7.48	0.0006	0.03	0.07	<0.002	0.0038
	ELF3_1340	C/T (N) Pro:Ser	0.05	7.48	0.0006	0.03	0.06	<0.002	0.0079
	GI5_198	T/C (I)	0.44	7.39	0.0007	0.03	0.06	<0.002	0.0095
	FRI_2927	T/G (S)	0.22	7.47	0.0008	0.06	0.11	<0.002	0.0012
	FRI_2732	A/G (S)	0.44	11.07	0.001	0.03	0.05	<0.002	0.0012
	ELF3_90	A/G (S)	0.07	7.01	0.001	0.02	0.06	<0.002	0.0012
	GI5_8997	T/A (I)	0.42	6.98	0.001	0.02	0.06	<0.002	0.0193
	GI5_9551	G/C (N) Gln:His	0.44	6.50	0.0017	0.02	0.06	<0.002	0.0193
	GI5_3966	C/G (I)	0.30	6.36	0.0019	0.02	0.06	<0.002	0.0208

MAF, minor allele frequency; SNP, single nucleotide polymorphisms.
*SNP named as Gene_position. SNP names in boldface were also significant at $FDR_Q < 0.05$ using the **A** matrix generated by EMMA.
†Major allele/minor allele. N, nonsynonymous; S, synonymous; I, intron. Amino acid changes provided for nonsynonymous SNPs.
‡Bonferroni critical value cut-off = 0.00017.
§$R^2 \times V_G$ is an estimate of the proportion of genetic variation that is accounted for by the SNP.
¶An $FDR_{EMP} < 0.002$ means that none of the reference genes had *P* values as small or smaller that the candidate–trait association.

multilocus SNP genotypes occurred earlier in the populations originating from the north (Fig. 4). Still, genotypes exhibiting a wide range of predicted effects were found across a wide range of latitudes.

Discussion

Latitudinal and climate-related patterns in bud flush and bud set exhibited strong genetic determination in *Populus balsamifera* that is consistent with local adaptation (Keller *et al.* 2011b). Trees from the south required a greater heat sum accumulation before bud flush and set bud at shorter critical photoperiods than trees from the north. Adaptation to future climates will require synchronization in the timing of both bud flush and bud set to track the season of suitable temperatures for growth (Savolainen *et al.* 2011). It is unclear, however, whether clines in bud flush and bud set both will be maintained and shift north or whether they will change relative to one another in the future, if *P. balsamifera* migration tracks climate warming.

The warmer temperatures that are predicted with ongoing climate change are expected to lead to bud flush occurring earlier because GDD will accumulate earlier in the spring (compare the responses of the same populations growing in warmer IH and cooler FBK). However, in both gardens, genotypes from the south, where it is warmer, consistently flushed buds later than genotypes from the north. Thus, if we imagine a future scenario with warmer spring temperatures, northern genotypes will still flush bud earlier than southern genotypes and be able to better take advantage of early season growing opportunities. Thus, in regard to early season growth, northern genotypes may have a fitness

Table 5 Effects on the day length at initiation of bud set from significant SNP–trait pairs calculated from adjusted means generated by the TASSEL model

Garden	SNP symbol*	2*a*†	*d*‡	*d*/*a*	α§
Indian Head	*ELF3_1340*	0.92	−0.27	−0.58	−0.66
	ELF3_90	1.14	0.01	0.01	−0.53
	LFY_1335	0.50	0.11	0.45	−0.14
	LFY_2277	−1.04	−0.29	−0.56	0.28
	ELF3_407	0.72	−0.18	−0.50	−0.49
	LFY_1714	0.42	−0.03	−0.13	−0.11
	LFY_1326	0.46	0.10	0.41	−0.13
	GI5_268	0.36	−0.22	−1.24	−0.33
	LFY_2996	−1.02	−0.29	−0.56	0.26
	CKB34_3513	0.93	0.84	1.82	0.28
	LFY_1120	0.51	0.26	1.01	−0.10
	LFY_1212	0.40	0.17	0.84	−0.10
	GI5_9551	0.19	−0.24	−2.52	−0.06
	CRY1.2_2106	N/A¶	N/A¶	N/A¶	N/A¶
	GI2_8862	−0.43	−0.63	−2.94	−0.31
	GI5_198	0.21	−0.23	−2.13	−0.06
	HY2.1_1961	0.81	0.23	0.57	0.61
	HY1.1_529	0.34	0.20	1.18	−0.08
	CKB34_452	0.14	0.35	5.08	0.11
Fairbanks	*GI5_5271*	1.28	0.10	0.16	0.30
	ABi1D_1595	−0.38	0.37	1.91	0.40
	GI5_9585	−0.29	−0.30	−2.13	0.08
	ELF3_1340	1.14	−0.05	−0.09	−0.59
	GI5_198	0.35	−0.26	−1.47	−0.07
	FRI_2927	−0.29	1.12	7.83	−0.11
	FRI_2732	N/A¶	N/A¶	N/A¶	N/A¶
	ELF3_90	2.06	0.68	0.66	−0.41
	GI5_8997	0.24	−0.30	−2.54	−0.03
	GI5_9551	0.29	−0.26	−1.85	−0.05
	GI5_3966	−0.52	−0.32	−1.24	0.09

SNP, single nucleotide polymorphisms.
*Number after gene indicates the position of SNP from the start of the gene.
†Calculated as the difference between the phenotype of major allele homozygote minus the minor allele homozygote.
‡Calculated as the difference between the phenotypes of the heterozygote and the average between the homozygotes.
§The average effect of the minor allele calculated as in Falconer & Mackay (1996).
¶statistics undefined when one homozygote was not sampled.

advantage over southern ones. Nonetheless, the fitness benefits and costs associated with smaller heat sum accumulation requirements for bud flush for northern compared to southern genotypes are unknown. One possibility is that northern genotypes may produce leaves with higher freeze tolerance (metabolic cost), allowing them to tolerate early season frost while taking advantage of early season growing opportunities (benefit), whereas southern genotypes may delay bud flush to avoid late spring frosts or extended mid-winter warm periods.

In both gardens, trees from the south set bud at shorter critical photoperiods than northern trees. In fact, photoperiods in the southern garden (IH) were never sufficiently long to maintain the growth of trees from the most northern populations, where they set bud soon after attaining photoperiodic competency (Soolanayakanahally *et al.* 2012). Conversely, when southern genotypes were placed in the north, the critical photoperiod signal for bud set was not encountered until relatively late in the fall; often weeks after local genotypes had already set bud. Thus, in current climates, southern genotypes may be susceptible to early fall frost damage when planted in the north, but if future climates are warmer, southern genotypes may be able to take advantage of later season growing opportunities that are unavailable to northern genotypes which are genetically constrained to set bud earlier.

Because photoperiod is the primary cue for timing of bud set and photoperiodic environments will not change with climate, fall bud set will not be delayed unless local populations evolve through migration or selection on local genotypes (Savolainen *et al.* 2011). In contrast, the temperature cue for bud flush will advance naturally as warming occurs earlier in the spring. Thus, we predict that the primary factor driving south to north migrations of genotype in forest trees will be selection on later bud set to take advantage of late season growing opportunities with warmer climates. Nonetheless, our data are suggestive that new genotypic combinations of bud flush and bud set may evolve as *P. balsamifera* moves north. If future northern trees benefit from both early bud flush and late bud set, new combinations of these traits that are not found in contemporary populations may arise and spread.

Loci associated with bud flush and bud set

Identifying the genes responsible for natural variation in seasonality allows deeper inquiry into the evolutionary process (Coop *et al.* 2009; Fournier-Level *et al.* 2011) and develops tangible markers for breeding and artificial selection. Three genes stood out as having multiple SNPs associated with one of the traits: four *FRIGIDA* (*FRI*) SNPs were associated with bud flush in FBK, seven *LEAFY* (*LFY*) SNPs were associated with bud set in IH, and six *GIGANTEA 5* (*GI5*) SNPs were associated with bud set in FBK. Because *FRI* plays a role in flowering after vernalization in *Arabidopsis thaliana* (Henderson *et al.* 2003), it is a strong candidate for traits associated with spring temperature responses, such as spring bud flush, which occurs simultaneously with flowering in mature poplars. In both *P. balsamifera* and *A. thaliana*, *FRI* homologs exhibit elevated diversity and are associated with natural variation in timing of early

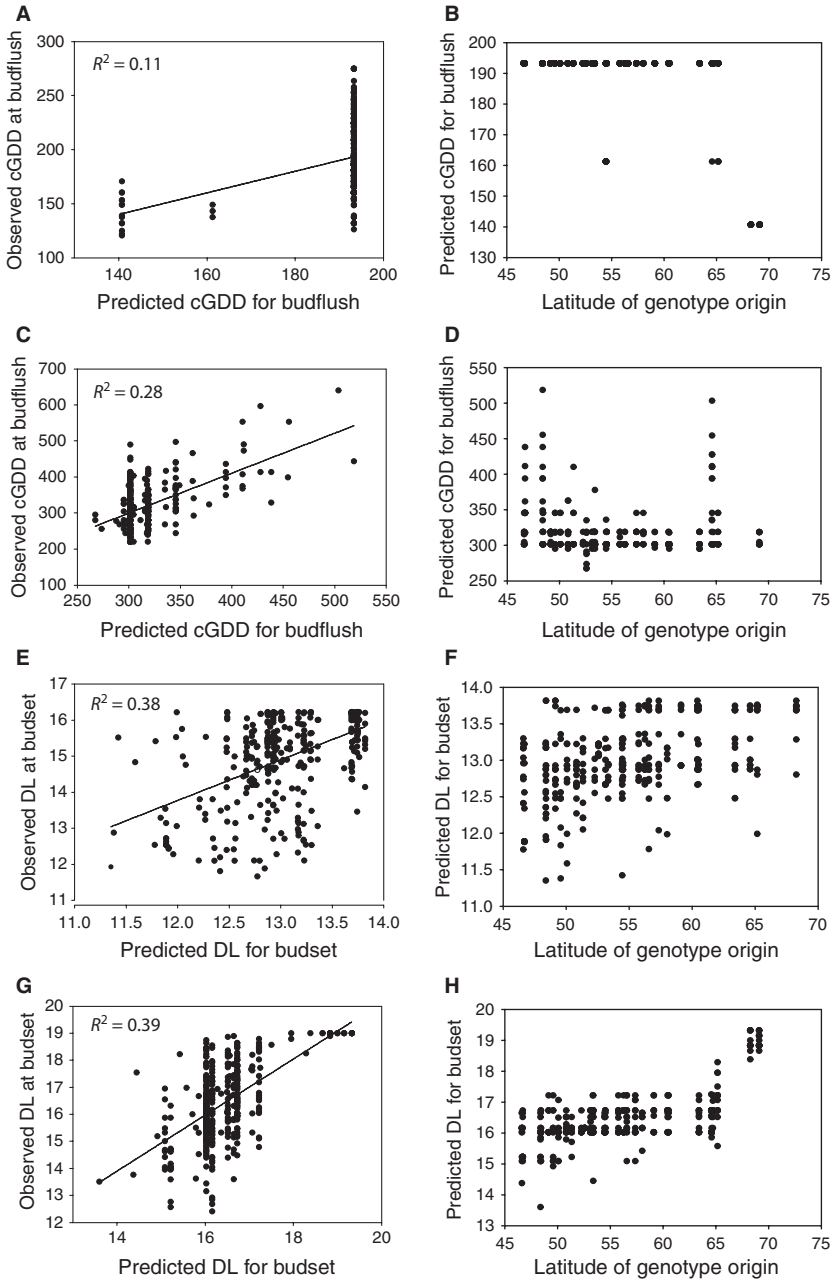

Fig. 4 Regressions of the predicted genotypic effects onto observed traits (panels A, C, E, G), and the relationships between latitude of genotype origin and predicted timing of bud flush and bud set from genotypes predicted by least angle regression (LAR) models (panels B, D, F, H). (A, B) Bud flush in the Indian Head garden. (C, D) Bud flush in the Fairbanks garden. (E, F) Bud set in the Indian Head garden. (G, H) Bud set in the Fairbanks garden. Single nucleotide polymorphisms identified by LAR and regression coefficients are shown in Table 6.

season developmental traits (Johanson *et al.* 2000; Le Corre *et al.* 2002; Keller *et al.* 2011a), indicating that allelic variants in *FR*I may provide a general source of developmental timing variation across angiosperms.

LFY is a master floral regulator in *Arabidopsis* that is expressed during floral initiation (Schultz & Haughn 1991; Moyroud *et al.* 2010). In *Arabidopsis*, *LFY* exhibits low diversity, consistent with a recent selective sweep (Olsen *et al.* 2002), but in *P. balsamifera*, *LFY* exhibits significantly elevated diversity and SNP variation that is associated with latitude (Keller *et al.* 2011a, 2012). The association of several *LFY* SNPs with bud set in the Indian Head garden suggests that it may play a role in affecting the timing of vegetative bud formation across latitude in environments with relatively short maximum DLs.

GI5 functions as a regulator of *CONSTANS* and *FLOWERING LOCUS T* (*FT*) as well as interacting with phytochromes to influence the internal circadian clock (Fowler *et al.* 1999; Mizoguchi *et al.* 2005; Bohlenius *et al.* 2006; Hsu *et al.* 2011). Several *GI5* SNPs exhibited significant associations with bud set in both gardens, and two SNPs *GI5_5271* and *GI5_92* exhibited associations with bud flush in the Indian Head and Fairbanks gardens, respectively. Some of the associations within *GI5* are likely false positives, generated by either high

Table 6 Regression parameter estimates and cumulative R^2

Model	Effect entered	Parameter estimate	Adjusted R^2
Indian Head bud flush	Intercept	161.27	
	GI5_5271	1.65	0.109
Fairbanks bud flush	Intercept	543.44	
	CRY1.2_2106	1.39	0.025
	ELF3_90	0.80	0.120
	ELF3_1340	0.83	0.212
	GI5_92	0.54	0.210
	FRI_2845	0.63	0.263
	FRI_505	0.15	0.283
Indian Head bud set	Intercept	11.73	
	CRY1.2_2106	2.54	0.036
	ELF3_1340	1.22	0.140
	GI5_198	1.40	0.237
	LFY_1335	0.20	0.236
	LFY_1212	0.90	0.288
	HY1.1_529	0.66	0.305
	GI5_268	0.51	0.358
	HY2.1_1961	−0.21	0.375
Fairbanks bud set	Intercept	15.29	
	GI5_5271	1.65	0.121
	GI5_198	1.61	0.229
	ELF3_1340	1.23	0.292
	ABi1D_1595	0.88	0.376
	ELF3_90	0.50	0.391

Because only one single nucleotide polymorphisms (SNP) was identified as significant in association analyses, a simple linear regression between predicted SNP effects and phenotypes was computed. For all other plots, the results of least angle regression models are plotted.

LD among SNPs (Fig. S4, Supporting information) or correlations between the bud flush and bud set traits (Platt *et al.* 2010). The geographic distribution of the *GI5_5271* deletion allele was unique among our sampled SNPs, being present only in northern populations (>64.5°N), and fixed in the northernmost *P. balsamifera* populations (INU and COT), with higher population structure than any other SNP we assayed. *GI5_5271* is found in individuals that exhibit the earliest bud set, under the longest photoperiods in Fairbanks. These patterns are consistent with expectations from a recent hard sweep in the far northern populations (Pritchard *et al.* 2010), but also may have been generated via allele surfing along the front of the expanding range (Klopfstein *et al.* 2006).

Across two common gardens, we identified candidate SNPs in nine genes in the *CO*/*FT* pathway associated with adaptive variation in bud flush and bud set in *P. balsamifera* (Tables 2 and 4). With the exception of *GI* (Rohde *et al.* 2011b), however, none of the genes we identified were found in previous QTL studies on *Populus*, which have implicated three phytochromes (*PHYA*, *PHYB1* and *PHYB2*; Frewen *et al.* 2000; Chen *et al.* 2002; Ingvarsson *et al.* 2008; Ma *et al.* 2010), *LATE ELONGATED HYPOCOTYL* (*LHY*; Ibanez *et al.* 2010; Ma *et al.* 2010), *ABSCISIC ACID INSENSITIVITY 1B* and 3 (Frewen *et al.* 2000; Chen *et al.* 2002; Rohde *et al.* 2002) and *FT* (Rohde *et al.* 2011b) as responsible for the variation in bud flush and bud set. Several of these previously identified QTL have been found using interspecific hybrid crosses, and genes identified may be responsible for interspecific difference, but not for variation within species. By contrast, in *Populus tremula*, *PHYB2* variation shows as a strong association with time of bud set (Hall *et al.* 2007; Ingvarsson *et al.* 2008; Ma *et al.* 2010), and this, together with the lack of evidence for any of the *PHY* genes being associated with the variation in bud flush or bud set in our analyses, suggests that different genes underlie clinal patterns of seasonality in different *Populus* species. Finally, we note that *GI* was associated with timing of bud set in Sitka and Norway spruce genotypes collected from a wide latitudinal extent (Holliday *et al.* 2010a; Chen *et al.* 2012), suggesting *GI* may contribute to natural variation in seasonality in a wide variety of trees.

Correspondence among studies

We found little overlap between the SNPs identified herein by association with phenotype and SNPs previously identified as targets of local adaptation, which used the same SNP data on the same *P. balsamifera* individuals (Keller *et al.* 2012). In fact, only seven SNPs identified as being associated with phenotypic variation in either FBK or IH were also identified by Fst-outlier scans or SNP–environment covariances as targets of local adaptation (Keller *et al.* 2012), and all of these were in *GI5* (all *GI5* SNPs in Tables 2 and 3 except *GI5_8862*). This may not be surprising, given that these analyses detect different signatures in the data. First, the association analyses presented here will detect only SNPs responsible for variation in the measured traits, whereas local adaptation scans identify SNPs but do not identify the phenotype upon which selection putatively acted. In other words, SNPs identified by local adaptation scans may be associated with local adaptation for traits other than bud flush or bud set. Second, SNPs identified here may contribute to variation in bud set and bud flush, but may not be responsible for local adaptation in these traits—the candidate SNPs we have identified account for only a small portion of total phenotypic variation. Third, local adaptation scans are searching for evidence of selection having shaped allele frequencies based on the location or environmental attributes of the site of origin where selection has been acting, whereas association analyses conducted with

phenotypes growing in a common garden detect loci responsible for variation in a novel environment. In fact, we found significant G × E effects across the IH and FBK gardens for three SNPs related to bud flush in FBK. Finally, although our gardens are located at or near the southern and northern range edges, some SNPs may not exhibit SNP–phenotype associations in either the FBK or IH garden environments, but are locally adapted to other environmental axes, such as those that differentiate the eastern and western parts of the species range. In sum, local adaptation screens may be more powerful tests for identifying locally adapted SNPs, whereas a common garden approach may be preferred for the identification of genes for breeding or making predictions about SNP influences in particular environments.

Contrary to the strong population-level differences in bud flush and bud set that we found, population structure was relatively weak for most of the SNPs identified as responsible for controlling timing of bud flush and bud set (Table S2, Supporting information reprints F statistics for our significant association SNPs from Keller *et al.* (2012)). Moreover, most of the multilocus genotypes identified via LAR were present across a wide latitudinal range (Fig. 4), indicating the lack of local genetic diversity may not be a near-term impediment to adaptation of the length of *P. balsamifera* growing seasons, with the notable exception of the most northern latitudes where unique multilocus genotypes are fixed. Regardless, given the widespread gene flow in *P. balsamifera* (Keller *et al.* 2010a), like many other forest trees (Sork & Smouse 2006), access to genetic variants should not impede adaptation.

External stimuli for bud flush and bud set

We used cGDD as the metric for assessing the signal for bud flush. Although cGDD is a reasonable approximation for the bud flush stimulus, the same genotypes planted in our two gardens did not flush buds after the same number of cGDD. This observation indicates that additional environmental stimuli are interacting with cGDD to stimulate bud flush, as have been identified in other studies (Paus *et al.* 1986). Determination of an accurate model of environmental factors that stimulate bud flush was beyond the scope of our study, but would aid in identifying the factors that generate G × E interactions across the environmental gradients inhabited by *P. balsamifera*.

Similarly, the same genotypes did not set bud under the same critical photoperiod when planted in our two gardens, as has been found in other common garden experiments in *Populus* (Rohde *et al.* 2011a). An obvious difference between our garden sites was that summer DLs were much longer in Fairbanks allowing variation in bud set among northern genotypes (above 57°N lat.) to be expressed (Table S1, Supporting information); however, the differences remained even when only genotypes with origins from lower latitudes were compared (Table S1, Supporting information). Another difference is that the rate of change in photoperiod is much faster in Fairbanks than in Indian Head; thus, the delay between when a critical photoperiod is perceived and manifestation of bud set (our measure of critical photoperiod) will be longer for plants grown in the north. Adjustment for this bias, however, will only serve to increase the differences between the gardens. Finally, temperature (Kalcsits *et al.* 2009; Rohde *et al.* 2011a) and soil moisture availability likely also influence timing of bud set; thus, more accurate assessments of the G × E interactions will require accurate measures of the entire suite of environmental signals affecting bud set.

Summary

We identified a suite of genes and loci as candidates for controlling bud flush and bud set in *Populus balsamifera*. Association mapping methods are expected to be most challenging when population structure is partitioned along the same geographic axes as variation in the trait, as is the case for *P. balsamifera* and many other species including humans (Coop *et al.* 2009). Nonetheless, we have shown that genetic variation controlling important traits is not always partitioned along axes defined by historical demography of *P. balsamifera* and detecting loci responsible for trait variation after adjustment for population structure and coancestry covariates is possible.

During the past century, the number of contiguous frost-free days during the growing season in Fairbanks Alaska has increased from 85 to over 125 days (Karl *et al.* 2009), so contemporary northern populations may no longer be the best adapted genotypes to northern environments (Beck *et al.* 2011). Although many environmental factors that we cannot anticipate are likely to change, the annual length of time with temperatures suitable for growth is likely to continue to increase. Timing of bud flush of northern genotypes will likely track earlier spring temperature increases because bud flush signals are temperature based, but bud set will not track later fall temperature increases because bud set is primarily cued by photoperiod. Therefore, southern genotypes, which set bud under shorter photoperiods, will likely be favoured when freezing temperatures occur later in the fall causing these genotypes to move north. It is unclear, however, whether the timing of bud flush in these southern genotypes, which required

longer accumulated GDD for bud flush than in northern genotypes, also will be favoured in future climate scenarios, or whether new combinations of traits conferring early bud flush under few cGDD and late bud set under shorter photoperiods will arise and spread. *Populus balsamifera* appears to have significant levels of genetic variation available within regions throughout most of its range to adaptively respond to these changes, with the notable exception of the most northern populations, which harbour unique variation associated with early bud set that is not found elsewhere. With sufficient lengthening of seasons, we might expect genotypes uniquely adapted to these very high latitudes to be lost, so studies of these variants have taken on some urgency.

Acknowledgements

We thank Mia Peterburs, Tanis Harvey, Pauline Ney, Amanda Robertson, Helena Storchova, Levi Wegner and Carol Lewis for help in planning and planting the Fairbanks common garden and Don Reynard for phenology monitoring at Indian Head garden. We thank the TASSEL team and Dr. Jianming Yu for assistance with association mapping methods. This study was supported by grants from Alaska EPSCoR (NSF award # EPS-0701898 and the state of Alaska) and the National Science Foundation Plant Genome Research Program (DBI-0701911 and 1137001).

References

Aitken SN, Yeaman S, Holliday JA, Wang TL, Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, **1**, 95–111.

Ballerini ES, Kramer EM (2011) In the light of evolution: a reevaluation of conservation in the CO–FT regulon and its role in photoperiodic regulation of flowering time. *Frontiers in Plant Science*, **2**, 81.

Beck PSA, Juday GP, Alix C *et al.* (2011) Changes in forest productivity across Alaska consistent with biome shift. *Ecology Letters*, **14**, 373–379.

Bohlenius H, Huang T, Charbonnel-Campaa L *et al.* (2006) CO/FT regulatory module controls timing of flowering and seasonal growth cessation in trees. *Science*, **312**, 1040–1043.

Breen AL, Murray DF, Olson MS (2012) Genetic consequences of glacial survival: the late Quaternary history of balsam poplar (*Populus balsamifera* L.) in North America. *Journal of Biogeography*, **5**, 918–928.

Chen THH, Howe GT, Bradshaw HD (2002) Molecular genetic analysis of dormancy-related traits in poplars. *Weed Science*, **50**, 232–240.

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

Coop G, Pickrell JK, Novembre J *et al.* (2009) The role of geography in human adaptation. *Plos Genetics*, **5**, e1000500.

Davis MB, Shaw RG (2001) Range shifts and adaptive responses to quaternary climate change. *Science*, **292**, 673–679.

Eckert AJ, Bower AD, Wegrzyn JL *et al.* (2009) Association genetics of coastal Douglas Fir (*Pseudotsuga menziesii* var. *menziesii*, Pinaceae). I. Cold-hardiness related traits. *Genetics*, **182**, 1289–1302.

Ehrenreich IM, Hanzawa Y, Chou L, Roe JL, Kover PX, Purugganan MD (2009) Candidate gene association mapping of *Arabidopsis* flowering time. *Genetics*, **183**, 325–335.

Falconer DS, Mackay TFC (1996) Introduction to Quantitative Genetics, 4th edn. Addison Wesley Longman Limited, Essex, UK.

Farmer REJ, Reinholt RW (1986) Genetic variation in dormancy relations of balsam poplar along a latitudinal transect in northwestern Ontario. *Silvae Genetica*, **35**, 38–42.

Fisher RA (1930) The Genetical Theory of Natural Selection. Clarendon Press, Oxford.

Forsythe WC, Rykiel EJ, Stahl RS, Wu HI, Schoolfield RM (1995) A model comparison for daylength as a function of latitude and day of year. *Ecological Modelling*, **80**, 87–95.

Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM (2011) A map of local adaptation in *Arabidopsis thaliana*. *Science*, **333**, 86–89.

Fowler S, Lee K, Onouchi H *et al.* (1999) GIGANTEA: a circadian clock-controlled gene that regulates photoperiodic flowering in *Arabidopsis* and encodes a protein with several possible membrane-spanning domains. *EMBO Journal*, **18**, 4679–4688.

Frewen BE, Chen THH, Howe GT *et al.* (2000) Quantitative trait loci and candidate gene mapping of bud set and bud flush in *Populus*. *Genetics*, **154**, 837–845.

Graham J, McNeney B, Blay S, Shin J (2006) LDheatmap: an R function for graphical display of pairwise linkage disequilibria between single nucleotide polymorphisms. *Journal of Statistical Software*, **16**, Code Snippet 3. http://www.jstatsoft.org/v16/c03.

Hall D, Luquez V, Garcia VM, St Onge KR, Jansson S, Ingvarsson PK (2007) Adaptive population differentiation in phenology across a latitudinal gradient in European Aspen (*Populus tremula*, L.): a comparison of neutral markers, candidate genes and phenotypic traits. *Evolution*, **61**, 2849–2860.

Hardy OJ, Vekemans X (2002) SPAGEDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Molecular Ecology Notes*, **2**, 618–620.

Henderson IR, Shindo C, Dean C (2003) The need for winter in the switch to flowering. *Annual Review of Genetics*, **37**, 371–392.

Hill WG (1974) Estimation of linkage disequilibrium in randomly mating populations. *Heredity*, **33**, 229–239.

Holliday JA, Ritland K, Aitken SN (2010a) Widespread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (*Picea sitchensis*). *New Phytologist*, **188**, 501–514.

Holliday JA, Yuen M, Ritland K, Aitken SN (2010b) Postglacial history of a widespread conifer produces inverse clines in selective neutrality tests. *Molecular Ecology*, **19**, 3857–3864.

Horvath DP, Anderson JV, Chao WS, Foley ME (2003) Knowing when to grow: signals regulating bud dormancy. *Trends In Plant Science*, **8**, 534–540.

Howe GT, Davis J, Jeknic Z *et al.* (1999) Physiological and genetic approaches to studying endodormancy-related traits in *Populus*. *HortScience*, **34**, 1174–1184.

Howe GT, Saruul P, Davis J, Chen THH (2000) Quantitative genetics of bud phenology, frost damage, and winter sur-

vival in an F-2 family of hybrid poplars. *Theoretical and Applied Genetics*, **101**, 632–642.

Hsu CY, Adams JP, Kim HJ *et al.* (2011) *FLOWERING LOCUS T* duplication coordinates reproductive and vegetative growth in perennial poplar. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, **108**, 10756–10761.

Hudson RR (2001) Two-locus sampling distributions and their application. *Genetics*, **159**, 1805–1817.

Ibanez C, Kozarewa I, Johansson M, Ogren E, Rohde A, Eriksson ME (2010) Circadian clock components regulate entry and affect exit of seasonal dormancy as well as winter hardiness in *Populus* trees. *Plant Physiology*, **153**, 1823–1833.

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

IPCC (2007) Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (eds Pachauri RK, Reisinger A). IPCC, Geneva, Switzerland.

Jackson ST, Overpeck JT (2000) Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology*, **26**, 194–220.

Johanson U, West J, Lister C, Michaels S, Amasino R, Dean C (2000) Molecular analysis of FRIGIDA, a major determinant of natural variation in *Arabidopsis* flowering time. *Science*, **290**, 344–347.

Jump AS, Penuelas J (2005) Running to stand still: adaptation and the response of plants to rapid climate change. *Ecology Letters*, **8**, 1010–1020.

Kalcsits LA, Silim S, Tanino K (2009) Warm temperature accelerates short photoperiod-induced growth cessation and dormancy induction in hybrid poplar (*Populus* × spp.). *Trees-Structure and Function*, **23**, 971–979.

Kang HM, Sul JH, Service SK *et al.* (2010) Variance component model to account for sample structure in genome-wide association studies. *Nature Genetics*, **42**, 348–U110.

Karl TR, Melillo JM, Peterson TC (2009) Global Climate Change Impacts in the United States. Cambridge University Press, Cambridge, MA.

Kaya Z, Adams WT, Campbell RK (1994) Adaptive significance of intermittent shoot growth in Douglas fir seedlings. *Tree Physiology*, **14**, 1277–1289.

Keller SR, Olson MS, Silim S, Schroeder WR, Tiffin P (2010a) Genomic diversity, population structure, and migration following range expansion in the balsam poplar, *Populus balsamifera*. *Molecular Ecology*, **19**, 1212–1226.

Keller SR, Olson MS, Silim S, Schroeder WR, Tiffin P (2010b) Data from: Genomic diversity, population structure, and migration following rapid range expansion in the balsam poplar, *Populus balsamifera*. Dryad Digital Repository. doi: 10.5061/dryad.1164.

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–U318.

Keller SR, Soolanayakanahally RY, Guy RD, Silim S, Olson MS, Tiffin P (2011b) Climate-driven local adaptation of ecophysiology and phenology in balsam poplar, *Populus balsamifera* L. (Salicaceae). *American Journal Of Botany*, **98**, 99–108.

Keller SR, Levsen N, Olson MS, Tiffin P (2012) Local adaptation in the flowering time gene network of balsam poplar, *Populus balsamifera* L. *Molecular Biology And Evolution*, **29**, 3143–3152.

Klopfstein S, Currat M, Excoffier L (2006) The fate of mutations surfing on the wave of a range expansion. *Molecular Biology And Evolution*, **23**, 482–490.

Koornneef M, Hanhart CJ, Vanderveen JH (1991) A genetic and physiological analysis of late flowering mutants in *Arabidopsis thaliana*. *Molecular & General Genetics*, **229**, 57–66.

Le Corre V, Roux F, Reboud X (2002) DNA polymorphism at the *FRIGIDA* gene in *Arabidopsis thaliana*: extensive nonsynonymous variation is consistent with local selection for flowering time. *Molecular Biology And Evolution*, **19**, 1261–1271.

Levsen ND, Tiffin P, Olson MS (2012) Pleistocene speciation in the genus *Populus* (Salicaceae). *Systematic Biology*, **61**, 401–412.

Loiselle BA, Sork VL, Nason J, Graham C (1995) Spatial genetic-structure of a tropical understory shrub, *Psychotria officinalis* (Rubiaceae). *American Journal Of Botany*, **82**, 1420–1425.

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.

Macdonald GM (1993) Fossil pollen analysis and the reconstruction of plant invasions. *Advances in Ecological Research*, **24**, 67–110.

McLachlan JS, Clark JS, Manos PS (2005) Molecular indicators of tree migration capacity under rapid climate change. *Ecology*, **86**, 2088–2098.

Menzel A, Sparks TH, Estrella N *et al.* (2006) European phenological response to climate change matches the warming pattern. *Global Change Biology*, **12**, 1969–1976.

Mizoguchi T, Wright L, Fujiwara S *et al.* (2005) Distinct roles of *GIGANTEA* in promoting flowering and regulating circadian rhythms in *Arabidopsis*. *Plant Cell*, **17**, 2255–2270.

Mouradov A, Cremer F, Coupland G (2002) Control of flowering time: interacting pathways as a basis for diversity. *Plant Cell*, **14**, S111–S130.

Moyroud E, Kusters E, Monniaux M, Koes R, Parcy F (2010) *LEAFY* blossoms. *Trends In Plant Science*, **15**, 346–352.

Olsen KM, Womack A, Garrett AR, Suddith JI, Purugganan MD (2002) Contrasting evolutionary forces in the *Arabidopsis thaliana* floral developmental pathway. *Genetics*, **160**, 1641–1650.

Olson M, Levsen N, Soolanayakanahally R *et al.* (2012) Data from: The adaptive potential of *Populus balsamifera* L. to phenology requirements in a warmer global climate. Dryad Digital Repository. doi: 10.5061/dryad.4p7t2.

Olson MS, Robertson AR, Takebayashi N, Silim S, Schroeder WR, Tiffin P (2010) Nucleotide diversity and linkage disequilibrium in Balsam Poplar (*Populus balsamifera*). *New Phytologist*, **186**, 526–536.

Pauley SS, Perry TO (1954) Ecotypic variation of the photoperiodic response in *Populus*. *Journal of the Arnold Arboretum*, **25**, 167–188.

Paus E, Nilsen J, Junttila O (1986) Bud dormancy and vegetative growth in Salix-polaris as affected by temperature and photoperiod. *Polar Biology*, **6**, 91–95.

Platt A, Vilhjalmsson BJ, Nordborg M (2010) Conditions under which genome-wide association studies will be positively misleading. *Genetics*, **186**, 1045–1052.

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

Pritchard JK, Pickrell JK, Coop G (2010) The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Current Biology*, **20**, R208–R215.

Riihimaki M, Podolsky R, Kuittinen H, Koelewijn H, Savolainen O (2005) Studying genetics of adaptive variation in model organisms: flowering time variation in *Arabidopsis lyrata*. *Genetica*, **123**, 63–74.

Rohde A, Prinsen E, De Rycke R, Engler G, Van Montagu M, Boerjan W (2002) PtABI3 impinges on the growth and differentiation of embryonic leaves during bud set in poplar. *Plant Cell*, **14**, 1885–1901.

Rohde A, Bastien C, Boerjan W (2011a) Temperature signals contribute to the timing of photoperiodic growth cessation and bud set in poplar. *Tree Physiology*, **31**, 472–482.

Rohde A, Storme V, Jorge V *et al.* (2011b) Bud set in poplar – genetic dissection of a complex trait in natural and hybrid populations. *New Phytologist*, **189**, 106–121.

Saccheri IJ, Rousset F, Watts PC, Brakefield PM, Cook LM (2008) Selection and gene flow on a diminishing cline of melanic peppered moths. *Proceedings Of The National Academy Of Sciences Of The United States Of America*, **105**, 16212–16217.

Samach A, Coupland G (2000) Time measurement and the control of flowering in plants. *BioEssays*, **22**, 38–47.

SAS Institute (2009) SAS/STAT 9.2 User's Guide, 2nd edn. SAS Institute Inc., Cary, North Carolina.

Savolainen O, Kujala ST, Sokol C *et al.* (2011) Adaptive potential of northernmost tree populations to climate change, with emphasis on scots pine (*Pinus sylvestris* L.). *Journal of Heredity*, **102**, 526–536.

Schultz EA, Haughn GW (1991) Leafy, a homeotic gene that regulates inflorescence development in *Arabidopsis*. *Plant Cell*, **3**, 771–781.

Simpson GG, Dean C (2002) Flowering – *Arabidopsis*, the rosetta stone of flowering time? *Science*, **296**, 285–289.

Soja AJ, Tchebakova NM, French NHF *et al.* (2007) Climate-induced boreal forest change: predictions versus current observations. *Global and Planetary Change*, **56**, 274–296.

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 and Environment*, **32**, 1821–1832.

Soolanayakanahally RY, Guy RD, Silim SN, Song M (2012) Timing of photoperiodic competency causes phenological mismatch in balsam poplar (*Populus balsamifera* L.). *Plant Cell and Environment*, doi: 10.1111/j.1365-3040.2012.02560.x.

Sork VL, Smouse PE (2006) Genetic analysis of landscape connectivity in tree populations. *Landscape Ecology*, **21**, 821–836.

Stinchcombe JR, Weinig C, Ungerer M *et al.* (2004) A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene *FRIGIDA*. *Proceedings of the National Academy Sciences of the United States of America*, **101**, 4712–4717.

Storey JD (2002) A direct approach to false discovery rates. *Journal of the Royal Statistical Society, Series B*, **64**, 479–498.

Thornwaite CW (1948) An approach toward a rational classification of climate. *Geographical Review*, **38**, 55–94.

Umina PA, Weeks AR, Kearney MR, McKechnie SW, Hoffmann AA (2005) A rapid shift in a classic clinal pattern in *Drosophila* reflecting climate change. *Science*, **308**, 691–693.

Vekemans X, Hardy OJ (2004) New insights from fine-scale spatial genetic structure analyses in plant populations. *Molecular Ecology*, **13**, 921–935.

Williams JW, Shuman BN, Webb T, Bartlein PJ, Leduc PL (2004) Late-quaternary vegetation dynamics in North America: scaling from taxa to biomes. *Ecological Monographs*, **74**, 309–334.

Yu JM, Pressoir G, Briggs WH *et al.* (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nature Genetics*, **38**, 203–208.

M.S.O., P.T., and W.R.S. conceived and designed the heritability experiments. M.S.O., and P.T. designed the association mapping experiments. M.S.O., N.L., R.Y.S., S.R.K., and P.T. contributed to data collection. M.S.O., N.L., and P.T. analyzed the data. M.S.O., and P.T. wrote the manuscript. All authors reviewed and edited the manuscript.

Data accessibility

Reference SNPs, annotations used for generating Q and A matrices, candidate SNPs used for association mapping and the sample phenotypes are available for download at www.popgen.uaf.edu, Keller et al. (2010b), and Olson et al. (2012). The sample locations can be found at www.popgen.uaf.edu.

Supporting information

Additional Supporting Information may be found in the online version of this article.

Table S1 Statistical tests of genotype by environment interactions (G × E) for timing of bud flush and bud set in the Fairbanks and Indian Head gardens.

Table S2 Hierarchical F statistics for each of the SNPs identified in association mapping studies reprinted from Keller *et al.* (2012).

Fig. S1 P-P plots for the association analyses of bud flush and bud set in *Populus balsamifera* measured in two different common gardens (Fairbanks, AK, and Indian Head, SK) and using SPAGeDi (Hardy & Vekemans 2002) to calculate the A matrix.

Fig. S2 P-P plots for the association analyses of bud flush and bud set in *Populus balsamifera* measured in two different common gardens (Fairbanks, AK, and Indian Head, SK) and using EMMA (Kang *et al.* 2010) to calculate the A matrix.

Fig. S3 Trait distributions in the Fairbanks (blue) and Indian Head (red) gardens.

Fig. S4 Phenological patterns exhibited by genotypes in the

Fairbanks (blue) and Indian Head gardens (red) plotted on a scale of Julian date and latitude of origin.

Fig. S5 Heat maps showing linkage disequilibrium (r^2) between loci within genes identified as associated with A) bud flush in the Fairbanks garden, (B) bud set in the Indian Head garden and (C) bud set in the Fairbanks garden.