Disentangling the roles of history and local selection in shaping clinal variation of allele frequencies and gene expression in Norway spruce (*Picea abies*)

Jun Chen¹, Thomas Källman¹, Xiaofei Ma¹, Niclas Gyllenstrand², Giusi Zaina³, Michele Morgante³, Jean Bousquet⁴, Andrew Eckert⁶, Jill Wegrzyn⁵, David Neale⁵, Ulf Lagercrantz¹, and Martin Lascoux¹,⁷

¹Department of Ecology and Genetics, EBC, Uppsala University, 752 36 Uppsala, Sweden

²Department of Plant Biology and Forest Genetics, Swedish University of Agricultural Sciences, Uppsala, P.O. Box 7080, SE-750 07 Uppsala, Sweden

³Dipartimento di Scienze Agrarie e Ambientali, Universita di Udine, 33100 Udine, Italy

⁴Canada Research Chair in Forest and Environmental Genomics, Institute for Systems and Integrative Biology, Université Laval, Québec City, Québec Canada G1V 0A6

⁵Dept. of Plant Sciences, University of California Davis, Davis, CA 95616 USA

⁶Dept. of Biology, Virginia Commonwealth University, Richmond, VA 23284-2012, USA

⁷Laboratory of Evolutionary Genomics, CAS-MPG Partner Institute for Computational Biology, Chinese Academy of Sciences, Shanghai, China
Corresponding author: Martin Lascoux

E-mail: Martin.Lascoux@ebc.uu.se

Phone: 46 18 471 64 16

Fax: 46 18 471 64 25

Keywords: clinal variation, photoperiodic pathway, Norway spruce, FT genes
Abstract

Understanding the genetic basis of local adaptation is challenging due to the subtle balance among conflicting evolutionary forces that are involved in its establishment and maintenance. One system with which to tease apart these difficulties are clines in adaptive characters. Here we analyzed genetic and phenotypic variation in bud set, a highly heritable and adaptive trait, among 18 populations of Norway spruce (*Picea abies*), arrayed along a latitudinal gradient ranging from 47°N to 68°N. We confirmed that variation in bud set is strongly clinal using a subset of five populations. Genotypes for 137 single nucleotide polymorphisms (SNPs) chosen from 18 candidate genes putatively affecting bud set, and 308 control SNPs chosen from 264 random genes, were analyzed for patterns of genetic structure and correlation to environment. Population genetic structure was low ($F_{ST} = 0.05$), but latitudinal patterns were apparent among Scandinavian populations. Hence, part of the observed clinal variation should be attributable to population demography. Conditional on patterns of genetic structure, there was enrichment of SNPs within candidate genes for correlations with latitude. Twenty-nine SNPs were also outliers with respect to $F_{st}$. The enrichment for clinal variation at SNPs within candidate genes (i.e. SNPs in *PaGI, PaPhyP, PaPhyN, PaPRR7* and *PaFTL2*) indicated that local selection in the 18 populations, and/or selection in the ancestral populations from which they were recently derived, shaped the observed cline. Validation of these genes using expression studies also revealed that *PaFTL2* expression is significantly associated with latitude, thereby confirming the central role played by this gene in the control of phenology in plants.
Introduction

Local adaptation is a key process in the evolution of species. Understanding how local adaptation is established and maintained, however, is especially difficult as its establishment is contingent upon historical conditions and its maintenance depends on the balance among conflicting evolutionary forces (e.g. Yeaman and Otto 2011). It is a particularly challenging task in forest trees, because they have long generation times and therefore cannot be easily manipulated experimentally. For instance, transfer experiments are theoretically possible but practically difficult to implement. On the other hand, the analysis of the strong latitudinal clines displayed by forest trees for potentially adaptive traits such as bud set (Dormling 1973; Savolainen et al. 2007; Aitken et al. 2008) can provide crucial information on the forces involved in local adaptation and, in particular, on the relative parts played by demography and selection in the establishment of the cline. Furthermore, phenology in general, and flowering time and bud set in particular, have been extensively studied and strong candidate genes are available, many of which belong to the photoperiodic pathway including the circadian clock (Holliday et al. 2010; Gyllenstrand et al. 2007; Formara et al. 2010; Albani and Coupland 2010; Bergelson and Roux 2010; Hsu et al. 2011). Recent candidate gene studies on the genetic basis of clinal variation in phenology in European aspen (Ma et al. 2010) and Sitka spruce (Holliday et al. 2010; Lobo 2011) illustrate well the promises of this approach.

In the present study we focus on a conifer species, Norway spruce (*Picea abies* (L.) Kar.), and more specifically on clinal variation in the northwest part of its natural range. The clines in phenological traits observed in this species and in this part of its range are intriguing for several reasons. First, as in many other forest tree species (Jaramillo-Correa et al. 2001; Savolainen et al. 2007; Alberto et al. 2011; Kremer et al. 2011; Prunier et al. 2011), Norway spruce populations are strongly differentiated for bud set ($Q_{ST} = 0.74$, R. Liesch, unpublished data), a trait with high heritability (average $h^2$=0.63 (varying from 0.33 to 0.78 across 15 populations, R. Liesch, unpublished data)). This strong level of phenotypic differentiation for bud set contrasts with extensive gene flow and limited genetic differentiation among populations at neutral genetic markers. Genetic structure within and among Norway spruce populations is largely accounted for by two major geographically based groups ($F_{ST} \sim 0.10$ between groups; $F_{ST} \sim 0.05$ within groups), the Nordic-Baltic group and
the Alpine group (Heuertz et al. 2006). Second, tree populations in Northwestern Europe are young, as most of the region was covered by ice during the Last Glacial Maximum (LGM, around 18,000 ybp). The combined use of palynological data, macrofossils and maps of genetic diversity suggest that Norway spruce recolonized Scandinavia after the LGM from populations located in central Russia and possibly also from cryptic refugia in the northern part of European Russia (Binney et al. 2009; Väliranta et al. 2011). The spread westwards followed two main recolonization routes, a northern route over Finland and into northern Scandinavia and a southern route going through the Baltic states (Tollefsrud et al. 2009). Norway spruce reached eastern Finland about 6,500 years ago and east-central Sweden about 2,700 years ago, after which it took 100-500 years to reach the present population size and to replace the existing Pinus-Betula-Alnus forests (Giesecke and Bennett 2004, Seppä et al. 2009). This history is reflected in a complex population genetic structure across the extant range of this species despite extensive gene flow (Heuertz et al. 2006).

The recolonization of Scandinavia from different LGM refugia is not specific to Norway spruce and has been observed in various plant and animal species (e.g. Jaarola et al. 1999; De Carvalho et al. 2010). Because Scandinavia was recolonized very recently and from different refugia, the presence of steep clines in phenological traits raises important questions. Do the clines only reflect local adaptation on de novo mutations occurring after the recolonization, i.e. under strong local selection over a very short period of time? Or do they reflect the deployment in space of polymorphisms already segregating in the ancestral populations, i.e. through soft sweeps (Pritchard et al. 2010; Savolainen et al. 2011)? A correct interpretation of the clines, thus, requires that population history and the ensuing population genetic structure (e.g. pattern of isolation-by-distance) are taken into account. This is because differences in allele frequencies among populations (e.g. for between southern and northern Swedish spruce populations) could be due to expansion from different glacial refugia where selection, or simply random genetic drift, led to the fixation of different alleles. This, of course, would not necessarily imply that local selection favoring different alleles in different areas does not occur, but it would nonetheless imply that it might not be the only explanation to the observed clinal variation. Finally, clinal variation in phenology could also partly reflect phenotypic plasticity,
an aspect that was not investigated here, but that is important in responses by tree populations to climate (see Bresson et al. 2011).

In the present study, after validating the presence of clinal variation in bud set in a subsample of five populations, we examined SNP frequencies at candidate genes for timing of bud set in 18 Norway spruce populations for correlations with latitude (Figure S1). Various approaches, capturing different facets of the data, were used to detect and interpret the clinal variation at candidate gene SNPs. We then tested for clinal variation in gene expression in a subset of the genes for which clinal variation was detected for SNP allele frequencies. This study offers a clear illustration of the interplay between demographic history and selection in shaping genetic variation and of the intrinsic difficulty to tell the two apart (Li et al. 2012).

**Material and Methods**

**Sampling and measurement of bud set:** Seeds of 303 Norway spruce trees were collected from 18 populations in four countries: Germany (1), Russia (1), Finland (7) and Sweden (9) at latitudes ranging from 47 °N to 68 °N (see Table 1 and Figure 1). In order to ascertain the presence of clinal variation in bud set, seeds were randomly sampled from five populations (SE-58.1 (Saleby), SE-61.6 (Istevallen), SE-62.7 (Högmansbod), SE-64.1 (Jämtland) and SE-66.4 (Jock) with latitudes ranging from 58°N to 66°N. Populations from northern Finland, that were used to analyze clinal variation in allele frequency, were not used here and were replaced by population SE-66.4 since their seeds had low germination rates. The number of seeds per population varied between 36 and 48 and the number of maternal trees, from 13 to 20. The seeds were germinated and the seedlings were then cultured in a growth chamber under continuous light (250 μmol m⁻² s⁻¹ light and 400-700 nm) and a temperature of 20°C for 8 weeks. Thereafter the seedlings were grown under increasing night length. Each photoperiod lasted 1 week starting with 1 week continuous light (LL), then 1 week with 22-h light /2-h dark, and so on. The dark period was extended by 1.5 hours each week until a photoperiod of 14.5 light / 9.5-h dark was reached. The latter lasted for 2 weeks (see Figure S2 for a detailed schedule).
Bud set was defined as two categories: 0, no bud at all and 1, when the bud started to appear until bud set was complete. Bud set was measured at the end of every photoperiodic treatment from week 5 (17.5-h light / 6.5-h dark) to week 8 (14.5 light / 9.5-h dark). The number of individuals that had set bud was counted in each population.

**DNA extraction and SNP genotyping:** Genomic DNA was extracted from megagametophytes using the DNeasy Plant Mini Kit (Qiagen, Germantown, MD, USA) then amplified using Illustra GenomiPhi V2 DNA Amplification Kit (GE Healthcare Life Sciences, Piscataway, NJ). A total of 445 single nucleotide polymorphisms (SNPs) representing 282 genes were genotyped using an Illumina GoldenGate 768 SNP array. The SNPs used in this study come from three different sources. The first set of 142 SNPs, including 137 candidate SNPs comes from our own resequencing efforts in a discovery panel of 48 Norway spruce alleles and comprises mostly SNPs from 18 putative candidate genes for bud set. The candidate genes were chosen among genes of the photoperiodic control pathway in *Arabidopsis thaliana* (Fornara et al. 2010) and based on information provided by experiments on bud set (e.g. Gyllenstrand et al. 2007) or on specific gene families (Karlgren et al. 2011) in Norway spruce. They were putatively involved in the photoperiodic pathway (PaCOL1, PaCOL2, PaCRY, PaPHYN, PaPHYO, PaPHYP, PaFTL2 and its promoter [formerly called FT4 in Gyllenstrand et al. 2007]), the circadian clock (PaCCA1 and its promoter, PaPRRI, PaPRR3, PaPRR7 and PaGI), or shoot apical development (PaMFTL1, PaHB3, PaKN1, PaKN2, PaKN4) of Norway spruce (see Table S1 for details). Multiple haplotypes were sequenced for each gene with an average number of 42 haplotypes/gene, except for three PaGI fragments for which only eight haplotypes were sequenced. The second source of 83 SNPs is the Comparative Re-Sequencing in Pinaceae (CRSP) project (http://dendrome.ucdavis.edu/crsp/). Sequences used for SNP discovery came from 800 primer pairs originally designed for loblolly pine (*Pinus taeda* L.). This list of primer sets was constructed without knowledge of putative gene function and was based solely on the set of 800 that maximized the production of high quality sequence in single validation samples from each of the seven species comprising the CRSP project. A single tree from Germany was used as the validation sample for Norway spruce. The SNP discovery panel consisted of 12 megagametophytes sampled in Poland, Latvia, Romania, Belarus,
Ukraine, Germany and Norway, which are fairly representative of the whole range. Finally, the remaining 220 SNPs come from the Arborea project (http://www.arborea.ulaval.ca/) as follows. A total of 1964 SNPs from 1485 genes initially identified and successfully genotyped in *P. glauca* were tested in 12 (diploid) individuals of Norway spruce from 12 populations in Central Europe including Latvia, Poland and Germany. These 1485 genes were transcription factors or candidate genes related to growth, wood formation, cell wall and lignin synthesis, which were deduced from expression studies in *P. glauca* or annotations in *Arabidopsis* homologs. Testing was done by submitting the panel directly to two Illumina GoldenGate SNP genotyping chips (Beaulieu *et al.* 2011; Pelgas *et al.* 2011), resulting in the discovery of 384 valid SNPs from 340 genes (transfer rate of 20%). Alleles with a frequency lower than 10% were removed from all SNP discovery panels. In summary, SNPs that were genotyped successfully were grouped into two categories: 137 candidate SNPs originating from 18 genes putatively involved in the photoperiodic pathway and 308 control SNPs originating from 264 genes not related to these functions. Importantly, the available information on the involvement of these candidate genes in the control of bud set is variable and consists of information in model species (*A. thaliana*, poplars) and physiological studies in Norway spruce. For example, while the level of expression of Pa*FTL2* correlates with the timing of bud set in Norway spruce (Gyllenstrand *et al.* 2007), no such direct evidence exists for other candidate genes. It is also worth noting that since many of the statistical tests used in this study are based on testing whether the number of candidate gene SNPs are overrepresented among the statistically significant SNPs, enlarging the candidate gene category with less well characterized genes will make these tests more conservative as enrichment will be less likely.

**Linkage disequilibrium:** Clinal variation of allele frequencies at SNPs unrelated to the targets of natural selection can be generated through linkage. Linkage disequilibrium (LD) among all pairs of SNPs, therefore, was estimated using the program Arlequin ver. 3.5 (Excoffier and Lischer 2010). Significant LD was identified at the 5% level using chi-square tests with a Bonferroni correction for multiple comparisons and was based on $r^2$ (Hill and Robertson 1966). Analysis was conducted for each population separately.
**Genetic structure**: To investigate population genetic structure, we used the Bayesian algorithm that is implemented in the program STRUCTURE ver. 2.3.3 (Pritchard et al. 2000; Hubisz et al. 2009) to delineate clusters of individuals based on their multilocus genotypes. All 18 populations and 264 unlinked control loci were used in this analysis. Linkage disequilibrium among SNPs within genes was limited as the number of SNPs per gene was small (see below). An admixture model with correlated allele frequencies was chosen and the analysis was performed for a number of clusters varying from $K = 1$ to $K = 22$. For each value of $K$ we performed 20 independent runs. Each run comprised 10,000 iterations of burn-in and 10,000 MCMC steps, as those values were found sufficient in pilot runs (longer burn-in (100,000) and MCMC (1,000,000) did not significantly change the results). To determine the optimal $K$ value, we used both the mode of posterior probabilities $\text{LnPr}(X|K)$ (Pritchard et al. 2000) and $\Delta K$ (Evanno et al. 2005).

Mantel tests were performed to test for isolation-by-distance among all sampled populations. A Mantel test of the correlation between a matrix of pairwise $F_{ST}$ between populations and the matrix of geographic distances was performed using the vegan library in R (Oksanen et al. 2011; R Core Development Team 2011). The ratio ($F_{ST}/(1-F_{ST})$) and the logarithm of geographic distance were used in this test as suggested by Rousset (1997). Finally, we calculated hierarchical $F$-statistics using the hierfstat library in R (Goudet 2005).

**Analysis of clinal variation in allele frequencies**: In order to detect candidate genes potentially involved with local adaptation, we treated each SNP independently and used different statistics to estimate the association between population allele frequencies and latitude. For all statistics, we examined the enrichment of SNPs located in candidate gene by calculating the ratio of candidate to control SNPs among those that were significant at 1% or 5% for the test statistics under consideration. We then used bootstrap resampling to assess the significance of the observed excess. Using this method should minimize the number of false positives, contingent upon the control SNPs largely reflecting neutral processes. By comparing the results obtained with the different methods, we obtained a set of candidate SNPs with a geographic pattern clearly distinct from the pattern at control SNPs.
Linear regression on latitude: If geographic variation is associated to an environmental gradient co-varying with latitude, we expect the allele frequencies to be clinal along latitudinal gradients. Ancestral and derived alleles at each SNP were determined using the southernmost population (RU-53.3, Bryansk) and nine individuals randomly sampled from four Brewer spruce (Picea breweriana S. Watson) populations as outgroups. In the few cases of conflict, we chose the SNP in Brewer spruce as the ancestral state. In order to measure the association of allele frequencies with population latitude, allele frequencies were first transformed using a square-root arcsine function (Berry and Kreitman, 1993) and then regressed on population latitude. Based on the results of the population structure analysis, three different datasets were considered when analyzing clines. We first considered the complete dataset. We then excluded the three southernmost populations. The GE-47.0 (Ruhpolding) and SE-58.3 (Saleby) populations belonged to a different cluster than other populations (see below). These two populations will create a pattern similar to the latitudinal cline, and we therefore excluded them from this analysis. The RU-53.3 (Bryansk) population is located much farther south from the rest of the populations and to avoid possible false positives, we excluded it. Finally, we excluded the four northernmost populations (FI-66.4, FI-67.0, FI-67.7 and FI-68.0) and considered a set including only populations from Central Fennoscandia. Since the second dataset seems the most appropriate for the analysis of clinal variation, it was analyzed in more detail than the two others.

Since the coefficient of determination, $R^2$, is a measure of the proportion of the total variance in frequencies that is explained by latitude, we used it as a statistic for “clinality” (Berry and Kreitman 1993). The significance of the $R^2$ value for each candidate SNP was obtained empirically by comparison to the distribution of $R^2$ values from the control SNPs and enrichment of significant SNPs among the candidate SNPs was tested for as described previously.

Regression Monte Carlo simulation: To investigate to what extent a significant clinal variation observed at one nucleotide site could be due to LD with another SNP, we used the Monte Carlo simulation approach developed by Berry and Kreitman (1993) (see also Verrelli and Eanes 2001). Each site is in turn considered as a “governing” locus and its effect on other loci assessed by Monte Carlo simulation. Details about this method are given in the Supplementary File.
Bayesian generalized linear mixed model (Bayenv) analysis: Another approach to detect clinal variation for SNP allele frequencies is to account for population history concomitantly with analysis of environmental correlations using a Bayesian generalized linear mixed model as implemented in the program Bayenv (Coop et al. 2010). This approach has recently been used successfully in humans (e.g. Hancock et al. 2011) and in loblolly pine (Eckert et al. 2010a) and more details on its implementation can be found in Coop et al. (2010), Hancock et al. (2008), Hancock et al. (2010) and in the Supplementary File. Briefly, environmental and geographic (e.g. latitude) factors are incorporated as fixed effects, while an estimated variance-covariance matrix of allele frequencies takes into account random effects due to shared population history. A null model of association of SNP frequency with population history is compared to an alternative model that also includes a linear effect of environmental and geographic variables. As a measure of the support for the correlation between SNP frequency and the environmental variables, a Bayes factor (BF) is calculated as the ratio of the posterior probabilities under the alternative and the null models. A BF of 3 is considered to reflect “substantial” evidence for selection and a BF larger than 10 indicates a “strong” support (Kass and Raftery 1995). However, because the null model is a pure drift model, i.e. an unlikely model in most cases, the value of the BF can be biased upward leading to an excess of false positives. To overcome this problem, we applied the method to all control SNPs to obtain an empirical distribution of BFs from these SNPs (Coop et al. 2010; Hancock et al. 2010). We scanned all 445 SNPs and picked the outliers according to this empirical distribution. The whole procedure was repeated eight times using eight different draws of the variance-covariance matrix describing population history. The results were averaged across the eight runs.

Fst outliers: Since the seminal study of Lewontin and Krakauer (1973), Wright’s fixation index, $F_{st}$, has been extensively used to detect recent selection from population genetics data (e.g. Beaumont and Nichols 1996; Beaumont and Balding 2004; Foll and Gaggiotti 2008). Loci under local selection are expected to show a larger $F_{st}$ value than neutral loci, while loci under balancing selection or purifying selection are expected to show a lower $F_{st}$ value. The available outlier tests using $F_{st}$ can be classified into two groups: tests based on an island model which assumes a common and unique migration pool (Beaumont and Nichols 1996; Beaumont and
Balding 2004; Foll and Gaggiotti 2008) and tests based on a hierarchical model that assumes that the population samples can be assigned to groups and that the migration rates are different among demes within groups and between groups (model implemented in Arlequin ver. 3.5, Excoffier et al. 2010). If the populations are hierarchically structured, ignoring it can lead to large number of false positives (see Excoffier et al. 2009; Narum and Hess 2011). Because an island model seems a reasonable assumption in the present case (see the results of STRUCTURE and hierarchical $F_{ST}$ analyses), we chose to analyze the data with BayeScan (Foll and Gaggiotti 2008; Fischer et al. 2010), which has proven to be one of the most reliable methods to detect outliers from the distribution of locus specific $F_{ST}$ (Pérez-Figueroa et al. 2010; Narum and Hess 2011). The program assumes that allele frequencies follow a multinomial Dirichlet distribution that arises under a wide range of demographic models. In this Bayesian approach, $F_{ST}$ has two components, a population specific component and a locus specific component. The alternative model for a given locus is retained when the locus specific component is necessary to explain the observed pattern of diversity. To account for the fact that we have different numbers of candidate and control SNPs, we used a prior odds ratio of 2, an approximate value of the ratio of control to candidate SNPs. We performed 20 pilot runs with 5000 length and 50,000 steps MCMC with additional 50,000 steps burn-in. Bayes factors as well as posterior probabilities were calculated to indicate how more likely the model with selection is compared to the neutral model. Outliers were identified at 5% and 1% significant levels of posterior probability corrected by the False Discovery Rate method (Benjamini and Hochberg 1995) as implemented in BayeScan.

**Candidate gene expression:** As a validation of outliers for environmental correlation with latitude and/or $F_{ST}$, the expression levels of PaGI, PaPRR7, PaFTL2 and PaCCA1 were assessed in the aforementioned sample of five Swedish populations. These genes were retained because they were either among the most significant in the Bayenv and BayeScan analyses (PaGI, PaPRR7, PaFTL2) or had strong physiological evidence as being involved with bud set (PaCCA1, U. Lagercrantz, unpublished data). Needles were sampled every fourth hour during the last 48 hours of each photoperiodic interval. Twelve seedlings were randomly chosen in each population and needles were collected and pooled. RNA was only extracted from needles sampled during the last 24-h and on photoperiodic treatments from week
1 to week 6. RNA was extracted according to manufacturer’s recommendations using STRN250 Spectrum™ Plant Total RNA kit (Sigma-Aldrich). cDNAs were synthesized from 0.5µg total RNA using Superscript III reverse transcriptase and random hexamer primers. Real-time PCR was performed according to Gyllenstrand et al. (2007) or on a MyiQ Real-Time PCR Detection System (Bio-Rad, Hercules, CA) with the following thermal conditions: 95°C for 7 minutes followed by 40 cycles of 95°C for 10 seconds and 60°C for 30 seconds (see Table S2 for a detail protocol and primer sequences). Each reaction was performed in duplicates containing 12µl DyNAmo™ Flash SYBR® Green (DyNAmo™ Flash SYBR® Green qPCR kit, Finnzymes, Espoo, Finland), 0.5mM of each primer and 4.75µl cDNA (diluted 1:100) in a total volume of 23.75µl. Samples were randomized in each step mentioned above, and α-tubulin was used as reference gene. Each sample was amplified twice with sample positions reversed in the plates. To compare expression levels, relative expression level of population x under photoperiod i at time-point t is defined as $R_{x|i}(t) = \Delta CT_{x|i}(t) - \Delta CT_{SN}$, where $\Delta CT$ is defined as $(CT_{REFERENCE} - CT_{TARGET})$ and $\Delta CT_{SN}$ is the average $\Delta CT$ value under continuous light treatment of population SE-58.3 (Saleby). To compare the amplitude between different photoperiod treatments at time-point t, we calculated the squared distance from the mean value within each treatment for every sample as:

$$\Delta D^2_{(xi)}(t) = (R_{(xi)}(t) - R_{mean})^2$$

where $R_{(xi)}(t)$ is the relative expression level and $R_{mean}$ is the mean relative expression level within that photoperiod treatment. For further analysis, we also repeated the whole procedure on the first 24-hour samples under 17.5-h light/6.5-h dark and 16-h light/8-h dark photoperiod treatments for PaFTL2 and PaPRR7. Nonlinear mixed models with harmonic terms were applied to represent the rhythmic expression patterns using the nlme library in R (Pinheiro et al. 2011). The models are described in details in Albert and Hunsberger (2005) and Yang and Su (2010).

All data are archived on Dryad with provisional doi:10.5061/dryad.82201
Results

Bud set and clinal variation: Bud set was measured after each week of photoperiodic treatment and reflects induction some two weeks earlier. A clear clinal pattern was observed, even when the southernmost population (SE-58.3) is excluded (Figure 2). Under a 17.5-h light / 6.5-h dark photoperiod, no individual set bud in the two southernmost populations (SE-58.3 and SE-61.6), around 10-20% of individuals set bud in the two intermediate populations (SE-62.7 and SE-64.1), and about half in the northernmost population SE-66.4. If the critical night length for bud set is defined as the time when 50% of plants have set bud, we observe a clinal increase of the critical night length as the latitude decreases.

Linkage Disequilibrium: We computed a standard estimate of LD (i.e. $r^2$) across all individual alleles pooled within each population. Only 560 out of 94673 pairwise comparisons were significant using a chi-square test with Bonferroni correction ($P$-value= 5.28E-07), with a total of 157 comparisons with $r^2$ values $> 0.3$. The average level of $r^2$ detected in Norway spruce candidate genes is around 0.2 (Figure S3) and is similar to the value observed in other spruce species (Beaulieu et al. 2011). Most of these were intragenic and only a few were among genes. Most of the linkage disequilibrium was found in candidate genes, mainly because we had higher SNP density in these compared to control genes, but LD remained limited there too.

Genetic structure: Population genetic structure was not pronounced. The optimal number of clusters calculated with the $\Delta K$ method (Evanno et al. 2005) is $K = 2$ (Figure 3a and Figure S4a) and the estimated alpha is $> 1.5$ indicating that most individuals are admixed. In line with previous studies (Lagercrantz et al. 1990; Heuertz et al. 2006; Tollefsrud et al. 2008; Tollefsrud et al. 2009), the main two clusters consist of populations from the Baltic-Nordic domain (Russian, Finnish and Northern Swedish populations) and of populations completely or partly derived from the Alpine domain (GE-47.0 and SE-58.3). The Bryansk (RU-53.3) population that is located in Central Russia (i.e. in one of the putative LGM refugia from which recolonization of North-West Europe started) showed a surprisingly similar clustering pattern to the populations of central Fennoscandia. This was also reflected by a lower $F_{ST}$ value between RU-53.3 and the central Fennoscandia cluster than with others (Table S3). If we choose the mode of posterior probability distribution of the data
(Pritchard et al. 2000) as a criterion, the optimal number of clusters was $K = 3$ (Figure 3b and Figure S4b). The new cluster was dominated by the four populations from northern Finland that are located at latitudes higher than 66°N.

There was no pattern of isolation-by-distance among all populations (Mantel statistic $r = 0.17$, $P$-value = 0.143) or among populations when the GE-47.0 (Ruhpolding) and SE-58.3 (Saleby) populations were discarded (Mantel statistic $r = -0.25$, $P$-value = 0.982). Strong hierarchical structure was not apparent either. The lowest hierarchical level (level1) was between populations. A second level (level2) grouped populations into five clusters (GE-47.0, SE-58.3, RU-53.3, central Fennoscandia, Northern Finland) and the highest level (level3) considered the Alpine domain and Baltic-Nordic domain. The low hierarchical $F$-statistic values between these three levels ($F_{level1/total} = 0.017$, $F_{level2/level1} = 0.035$, $F_{level3/level2} = 0.05$) show that hierarchical genetic structure is negligible and thus applying an island model in the $F_{ST}$ outlier analysis should be reasonable.

**Analysis of clinal variation in allele frequencies:**

Using all of the data, including populations from the Alpine and Baltic-Nordic domains, a total of 54 candidate SNPs, out of 137 (39.4%), and 88 control SNPs, out of 308 (28.6%), have a regression slope significantly different from zero at the 5% significance level. To reduce the number of positive SNPs simply due to population structure, we removed the three southernmost populations (GE-47.0, RU-53.3, SE-58.3) and focused on populations located at latitude 61°N and above. A total of 92 out of 445 SNPs, 34 (24.8%) of which were candidates and 58 (18.8%) of which were controls, showed a significant linear regression with latitude at the 5% significance level (Table 2 and Figure 4). Finally, when we removed the four northernmost populations, and only kept populations from central Fennoscandia (latitudes ranging from 61°N to 65°N), there were only 28 significant SNPs including 11 from candidate genes.

The adjusted $R^2$ of regression of SNP frequency on latitude was estimated to represent how much of the variance in frequency could be explained by latitude and test for enrichment of the candidate SNPs among significant SNPs at the 5% significance
level. In all three datasets there was no enrichment in candidate SNPs among significant ones.

Observing that there are 92 SNPs in the dataset without the three most southern populations whose allele frequencies are significantly correlated with latitude and the presence of LD among candidate SNPs, we tested whether ‘clinality’ at one site affects that of flanking SNPs. A Monte Carlo simulation approach was used for this purpose and only the 34 significant candidate SNPs were investigated as they showed higher level of LD (Figure S5). Non-significant SNPs can be also affected by nearby variants under selection but they were ignored since they do not exhibit clinal variation and were unlikely to explain clinal variation at other segregating sites. Two broad patterns were apparent. One pattern involved clinality being driven by LD within and among genes, such as that observed at PaMFTL1 SNPs. In contrast, most of the effect of clinality for SNPs in the promoter of PaFTL2 came from within the promoter itself, as was already suggested by the pattern of LD. The analysis also showed a group of SNPs whose effect could be explained by a small number of other SNPs such as PaMFTL1_3658, PaFTL2pr_1951, PaPHYP_RIII274, PaGI_F2_9_987 or PaPRR7_F3_104. These SNPs had a higher level of ‘clinality’ than others and usually showed stronger signals in the two other tests (see results below and Table 2).

Bayesian generalized linear mixed model (Bayenv) analysis: The average BF estimated from the candidate SNPs was significantly higher than the one estimated from control SNPs (average BF = 3.26 and 1.48, respectively; Wilcoxon two sample test \( P \)-value = 0.0083). More importantly, there was a significant enrichment of candidate SNPs in the upper tail of the empirical distribution of BFs. The ratio at 10% tail (BF \( \geq \) 2.7) was 0.87 (27/31, candidate to control SNPs), and increased to 1.12 (18/16) in the 5% tail (BF \( \geq \) 4.6) and around 1.75 (7/4) in the 1% tail (BF \( \geq \) 10.4), (bootstrap < 1% for all tails, see Table 2). To remove the possible effect of control loci that were potentially under selection, we then excluded 14 control SNPs that were outliers in more than one of the tests (linear regression, Bayenv and \( F_{ST} \) outliers (see below)) and re-constructed the null model. This caused an increase of BF estimates in the upper 25% tail but the candidate outliers and the enrichment ratio remained the same (0.9, 1.06 and 2 at 10%, 5% and 1% tails, respectively).
**F_{st} outliers**: The average value of $F_{ST}$ across SNPs was 0.05. Candidate SNPs had significantly higher $F_{ST}$ values than control SNPs (mean $F_{st} = 0.056$ and 0.047, respectively; Wilcoxon rank sum test $P$-value = 0.00026). BayeScan identified 29 outliers at the 5% significance level (corrected FDR test), 16 of which were from candidate loci. At the 1% significance level, there were 11 candidate SNPs and 6 control SNPs identified as outliers (Figure 5). Fisher’s exact test indicated a significant enrichment of candidate SNPs in these tails ($P$-value = 0.0064 and 0.0054, respectively). At these significance levels, all outliers had high $F_{ST}$ values and therefore were putatively under positive selection. The exception was PaKN2b_1157 and two other control SNPs that had low $F_{st}$ values (~0.020).

**Summary of the analysis of SNP variation**: We combined the results from all of the analyses above (Figure S6) to obtain a short list of loci with evidence of clinal variation and/or selection (Table 2). Most of the SNPs displaying strong signals in all our analyses came from 5 genes: PaMFT1, PaFTL2, PaGI, PaPRR7 and PaPHYP. One particularly interesting SNP was PaFTL2pr_1951, which represents a recent mutation that does not exist in Brewer spruce or in the RU-53.3 (Bryansk) population, both of which served as outgroups, nor in populations from the Alpine domain (GE-47.0 (Ruhpolding) and SE-58.3 (Saleby)). The clinal pattern in allele frequency at PaFTL2pr_1951 cannot be explained by any other SNP, but PaFTL2pr_1951 affects the clinal variation at 14 other significant candidate SNPs. Previous studies on expression of FT-Like genes in Norway spruce have shown that the expression level of PaFTL2 correlates with bud set. Expression is induced by night lengths exceeding the population specific critical night length for bud set for trees from Romanian and Arctic populations, respectively (Gyllenstrand et al. 2007). The single nucleotide mutation in the PaFTL2 promoter might affect the divergent expression patterns in genotypes with varying critical night length. Other strong outlier SNP were located in putative circadian clock genes (PaPRR7, PaGI) that are likely regulators of PaFTL2 expression and bud set (Karlsgren A., N. Gyllenstrand and U. Lagercrantz unpublished). PaPRR7_F2_534 is a non-synonymous substitution causing an amino acid replacement from serine to threonine (TCT to ACT in nucleotide) and was found to be associated with two other SNPs, PaPRR7_F1_771 and PaPRR7_F3_104 (in intron and 3’ UTR region, respectively). PaGI_F2_9_987 and PaGI_F2_9_1470 are linked and both cause an amino acid replacement from histidine to tyrosine (CAT to
TAT in nucleotide). Using results from the NCBI BLASTP 2.2.25+ tool (Altschul et al. 1997), histidine is likely the conserved state, as all 64 hits from 47 species with high similarity (e-value < 6e-95) had a histidine in this position.

**Candidate gene expression:** For a subset of candidate genes, we compared the relative expression level among photoperiodic treatments using the pairwise Wilcoxon rank sum test implemented in R (R Development Core Team 2011). The mean expression level of PaFTL2 across all populations increased continuously as the night length increased from near 0 in constant light to 2.7 at night length of 3.5-h and 6.0 at night length of 8-h. In the two southernmost populations (SE-58.3 and SE-61.6), PaFTL2 was significantly upregulated at night length of 6.5-h compared to the expression level under constant light (P-value < 0.01, FDR corrected). In the two northernmost populations (SE-64.1 and SE-66.4), this change happened for a night length of 5-h (P-value < 0.01 and < 0.05, respectively). This change might happen even earlier in SE-66.4 as the P-values are 0.052 for a night length of 2-h and 3.5-h. In population SE-62.7, there was a significant increase in expression of PaFTL2 already at 2-h night length, which is earlier than expected based on previous results. Except for SE-58.3, all populations showed a second increase in expression level when the night length reached 8-h compared to the first increase level (P-value < 0.05 for SE-62.7 and < 0.001 for the other three). We therefore compared the relative expression level of PaFTL2 among populations under a night length of 6.5-h or 8-h. There is a significant difference (P-value < 0.001) in PaFTL2 expression between SE-58.3 (3.76) and other populations (5.14 on average) for a night length of 6.5-h (Figure 6). Furthermore, when the night length is 8-h, the mean expression levels of the five populations are significantly differentiated (SE-58.3 4.11, SE-61.6 6.22, SE-62.7 6.20, SE-64.1 6.70 and SE-66.4 6.73, P-value < 0.05) and linearly correlated with latitude (P-value = 0.048, adjusted R² = 0.75, Figure 7).

Unlike PaFTL2, the expression levels of PaCCA1, PaGI and PaPRR7 were similar across photoperiod treatments, and differences were not statistically significant. These genes are presumably located in the circadian clock (PaCCA1, PaGI and PaPRR7) or related to it (PaFTL2) and their expression level is expected to show a diurnal expression rhythm. To measure the amplitude of their rhythm, we calculated the squared distance from the mean level of relative expression in each photoperiod treatment. There was an increase of amplitude in 16-h light /8-h dark photoperiod in
all five populations in PaPRR7 (P-value < 0.05) compared to constant light. Furthermore, we represented their cycle trends using harmonic regression with one or two periods. The predicted curves of PaPRR7 and PaGI showed nice diurnal rhythms with a periodicity of around 24 hours (see Figure 6 and Table S4 for all estimated parameters), while neither PaFLT2 nor PaCCA1 fit the autoregressive model well and showed no clear diurnal rhythms.

**Discussion**

In the present study, we have used different approaches to identify SNPs in candidate genes that are potentially associated to the clinal variation observed for bud set in Norway spruce. The combined approaches identified a promising set of SNPs, in particular in the promoter of the PaFTL2 gene. The latter was furthermore accompanied by clinal variation in expression level. Below we discuss some caveats of the different approaches, how they relate to each other and how their joint use can shed light on the genetic control of bud set and local adaptation. In particular, we shall discuss the effect of population histories on the interpretation of clinal variation. Finally, we briefly outline the implications of the present results for the design of future association mapping studies for traits showing strong clinal variation and give some suggestions for future extensions of the current study.

**Ascertainment bias:** The SNPs analyzed here were discovered using three different discovery panels. All candidate SNPs originated from a large discovery panel (around 48 chromosomes), while the bulk of the control SNPs come from smaller discovery panels (12 chromosomes in the CRSP re-sequencing panel and 24 chromosomes in the Arborea genotyping panel). Albrechtsen et al. (2010) reported that the size and the geographic distribution of the discovery panel could lead to a bias in the site frequency spectrum. They also showed that as long as the discovery panel size is not too small (< 4 chromosomes), estimates of $F_{st}$ are rather robust to ascertainment bias since the bias has a similar effect on the total and among population variances. In our data, we observed a deficit of low frequency alleles in the control loci (mean allele frequency = 0.33 and 0.24 for control and candidate SNPs, respectively, P-value < 2.2e-16 in two-samples t-test). As expected, the control SNPs have significantly higher heterozygosity than candidate SNPs (mean heterozygosity = 0.29 and 0.25 for control and candidate SNPs respectively, P-value = 0.0238). $F_{ST}$ values among
populations are, however, of the same order of magnitude as estimates obtained through analysis of sequence data (mean $F_{st} = 0.056$ and 0.047, for candidate and control, respectively; 0.040 between Northern and Southern Sweden in Heuertz et al. 2006) although it should be pointed out that the two studies have many SNPs in common.

The control SNPs were also used to build a covariance matrix between populations for the null model used by Bayenv. The covariance matrix can be viewed as a model-based, population-specific estimate of $F_{st}$ (Nicholson et al. 2002; Coop et al. 2010). Hence, in both the Bayenv and BayeScan analysis, the results should not have been affected too strongly by ascertainment bias, and selection should indeed be the major cause of the substantial difference in $F_{st}$ estimates between control and candidate SNPs. This is supported by the analysis of the site frequency spectrum for control SNPs and candidate SNPs in the total data (Figure S7a) and for the significant SNPs detected in the three main analyses: clinal variation (Figure S7b), Bayenv analysis (Figure S7c) and the BayeScan analysis (Figure S7d). The site frequency spectrum for each of those sets of SNPs did not show any specific structure, and there was no clear over-representation of the candidate SNPs in the lowest allele frequency classes.

Population history, selection from standing variation and local selection: A major issue when trying to identify SNPs associated to local adaptation is to account for demographic processes that can leave the same signature (Väsemegi 2006; De Carvalho et al. 2010; Savolainen et al. 2011). Norway spruce, as well as many other plant and animal species, recolonized Northern Europe after the LGM along at least two main recolonization routes (Gieseke and Bennett 2004; Tollefsrud et al. 2009). Thus, the current northern European spruce forests were established during a rather short evolutionary period and from different sources. This is indeed what our data suggest. There are three main genetic clusters, but with the exception of the population from Germany and the four populations from Northern Finland, no populations can be assigned entirely to any of these three clusters. In the case of population SE-58.3, the presence of a group of individuals exhibiting a close similarity to individuals from the German population likely reflects the use of seeds imported from Germany and Belorussia in afforestation over the last century. Apart from Germany and the four populations from Northern Finland, the pattern in populations ranging from 60°N to 65°N, suggests a modest contribution from the
Alpine domain (the blue cluster in Fig. 3), as well as more substantial contributions from two Eastern LGM refugia (green and red clusters in Fig. 3). One of these two refugia, the one associated with populations from northern Finland, was probably located at a higher latitude than the other, a possibility suggested by recent paleogeographic studies (Väliranta et al. 2011). The green cluster could also reflect the survival of trees in Scandinavia or close to it during the LGM (Öberg and Kullman 2011), but if those scattered refugia had contributed to present-day populations, the similarity and the genetic closeness of the Russian population (RU-53.3) to populations from central Fennoscandia (Fig. 3b and Table S3) might be difficult to explain since it would run counter to the general direction of recolonization inferred from both genetic (Lagercrantz and Ryman 1990) and paleoecological studies (Gieseke and Bennett 2004).

Independently of the exact geographical origin of the source populations, the spread of populations together with the existence of diverse origins could lead to serial founder effects and secondary contact, that in turn may result in neutral polymorphisms showing clines in allele frequencies (Novembre and Di Rienzo, 2009). Two recent studies have investigated the possibility that secondary contacts and admixture could play a role in local adaptation in Scandinavian tree species. In Scots pine (Pinus sylvestris), Savolainen et al. (2011) detected no admixture zone and speculated that adaptation occurred during the colonization process from standing variation. De Carvalho et al. (2011) did detect admixture in European aspen (Populus tremula) from central Sweden and argued that admixture facilitated adaptation from standing variation. Both studies differed from ours in major ways. The inference of population structure in Pinus sylvestris was based on only three populations and a limited number of markers. Therefore, it may not have had the power to detect a putative admixture zone. The study of De Carvalho et al. (2011) was based on seven populations, only one of them located in Scandinavia and the others being far from it. Admixture in European aspen was inferred in this single Swedish population located at 63°N, possibly reinforcing the impression of selection from standing variation.

In Norway spruce, while we did not observe any evidence of isolation-by-distance (Mantel test, P-value > 0.95), we did detect evidence of secondary contacts and admixture. In order to tell apart selection from demography we have used a three-step approach. First, we used simple regression analysis of allele frequencies on latitude
for differing subsets of the data. The number of SNPs showing clinal variation decreased sharply when populations from northern Finland were not considered, suggesting that part of the clinal variation reflected population structure. Since we did not correct for demography in these regression analyses, no enrichment in candidate genes was observed in the tails when the adjusted $R^2$ statistics was considered. This simply means that demography is, on the average, affecting all SNPs equally across the genome, irrespective of their status (candidate vs non-candidate) and that demography needs to be accounted for if one wants to identify SNPs also affected by selection. Second, since population structure is not accounted for in the regression analysis, we used a Bayesian generalized linear mixed model (Bayenv) on all populations (Coop et al. 2010; Hancock et al. 2008; Hancock et al. 2010; Hancock et al. 2011). This led us to discard many SNPs that had high $R^2$ in the standard cline analyses, suggesting that the strong clinal variation at those SNPs likely reflects population structure. However, since we also observed an enrichment of candidate SNPs when population structure was accounted for, it would therefore seem that the data also reflect selection, possibly from standing variation, rather than from new mutations. The latter is supported by the fact that most of the alleles that show significant SNPs are already present in the Russian population, which is close to the putative ancestral refugia. One of the most interesting SNPs, however, is PaFTL2pr_1951, which was absent in Brewer spruce, in the Russian population and in the German population, and therefore appears to be a new mutation. Third, we tested for $F_{ST}$ outliers in order to assess whether the enrichment in candidate SNPs that we detected with Bayenv was due to selection in past environments or since the populations colonized Scandinavia. In contrast to results in European aspen (Ma et al. 2010), we did detect significant SNPs when we used an $F_{ST}$-outlier approach, suggesting the presence of some amount of local selection. Furthermore, there was a good correspondence between the Bayenv and the $F_{ST}$-outlier approaches, with 12 SNPs picked by the two methods and 4 and 6 identified uniquely by Bayenv and the $F_{ST}$-outlier approach, respectively. Ma et al. (2010) argued that extensive long-distance gene flow in wind-pollinated forest trees (e.g. Robledo-Arnuncio 2011) will reduce population genetic structure and will also overwhelm selection and limit our ability to discover selection signal by erasing peaks of genetic differentiation (Ma et al. 2010). In our study, local selection seems to have been strong enough to generate
extreme $F_{st}$ values at certain loci so that they emerge as outliers in the analysis and furthermore, we also observe a significant enrichment of candidate SNPs among the significant ones. Prunier et al. (2011) obtained similar results in black spruce.

**Differentiation at quantitative trait loci:** Some authors (Latta 1998, Le Corre and Kremer 2003, Kremer and Le Corre 2011) have argued that low values of $F_{ST}$ will be observed even at loci underlying quantitative traits, and that this reflects a decoupling between the polymorphism of a quantitative trait and the polymorphism of its underlying quantitative trait loci. For instance, in the model of Kremer and Le Corre (2011), natural selection first acts on beneficial allelic associations at different loci in different populations, and only significantly affects allelic frequencies at individual loci at later time points. The high $F_{ST}$ values obtained for significant outlier SNPs (the mean $F_{ST}$ of outliers is 0.13 if the only value that was in the lower tail of the distribution is removed, Table 2) are also in contrast to results presented by Kremer and Le Corre (2011) where candidate SNPs and control SNPs exhibited similar values of $F_{ST}$. We believe that this is simply due to the fact that most of the candidate SNPs used in Kremer and Le Corre (2011) may actually not be associated to the studied quantitative traits. Nonetheless, as in Kremer and Le Corre (2011), Prunier et al. (2011) or Namroud et al. (2008), $F_{ST}$ values at candidate gene SNPs were still much smaller than $Q_{ST}$ values.

**Expression patterns of candidate genes:** Most of our putative candidate genes were chosen as representatives of the photoperiodic pathway and include putative photoreceptors, circadian clock genes and genes acting downstream of these genes. Even though the function of most of these candidate genes in Norway spruce is still poorly characterized, many of the significant SNPs were located in genes that are logical candidates based on current knowledge. One such example is the set of significant SNPs located in the PaPHYP gene. Physiological experiments have shown that genotypes from high latitudes do not respond to exposure to red light during an artificial long night and thus set buds, in contrast to genotypes from more southern latitudes (Clapham et al. 1998; Gyllenstrand et al. 2007). Based on sequence homology, PaPHYP corresponds to PhyB in angiosperms, which is the receptor that is predominantly responsible for responsiveness to red light (Schneider-Poetsch et al. 1998) and clinal variation at PhyP SNPs was observed in Scots pine (Garcia-Gil et al. 2003). Significant SNPs were also identified in putative circadian clock genes, in
particular in PaGI and PaPRR7. Our unpublished data support that these genes are indeed part of the Norway spruce circadian clock, and furthermore show divergent expression patterns in genotypes from extreme northern and southern latitudes, in particular divergent responses to night breaks (Karlgren A., N. Gyllenstrand and U. Lagercrantz unpublished). These results also agree well with studies in Sitka spruce and poplar where this category of genes was found to exhibit clinal variation and be associated with bud set (Holliday et al. 2010; Ma et al. 2010; Hall et al. 2011; Hsu et al. 2011).

If enhanced genetic structure in these candidate genes is truly caused by selection, we might expect a divergence in gene expression levels among alleles. As the circadian clock is composed of multiple integrated feedback loops, mutations in both regulatory and coding parts could affect the transcriptional patterns of multiple clock genes. When analyzing gene expression we used a bulk-sampling strategy with each population being represented by 12 randomly chosen individuals. Thus, the expression level should reflect the effect of each allele weighted by its frequency. This could undermine the estimation of the true expression divergence between the two alleles, as the effect of a slight allele frequency shift caused by weak selection may not be detected. In this study, the expression levels of PaPRR7, PaCCA1 and PaGI show very subtle differentiation along the latitudinal gradient. However, this marginal difference might still cause significant expression changes in downstream genes such as PaFTL2 and divergent bud set and growth cessation along the latitude in phenotype.

Our previous studies have shown that PaFTL2 expression is controlled by photoperiod and is correlated to bud set under various experimental conditions. Additionally, PaFTL2 expression differs between genotypes from extreme northern and southern latitudes (Gyllenstrand et al. 2007). In the present study, we found a weak but significant latitudinal cline in the expression of PaFTL2 correlated to bud set in a more limited part of the latitudinal range. In light of these data, the identification of several significant SNPs in the promoter of PaFTL2 is very encouraging, especially as some are located in or close to putatively important binding sites. Analyses using “Plant promoter analysis navigator” (PlantPAN, Chang et al., 2008) and “RNAhybrid” (Krüger and Rehmsmeier, 2006) indicate that the PaFTL2pr_1951 SNP is located in the 3’ end of a putative miRNA-target binding site of a member of the
miR169 or miR399 families, which are important transcriptional regulators in plants (reviewed by Bartel, 2004) and is located next to the binding site of an “E2F Consensus Element” transcript factor that might cause an affinity change.

**Implications**: One of the main aims of the statistical methods developed for association studies is to decrease the number of false positives due to population structure. This, however, could lead to false negatives if population history and differences in allele frequency at causal SNPs co-vary as may be the case in the present study (Atwell et al. 2010; Eckert et al. 2010b). As pointed out by Atwell et al. (2010), “plants from northern Sweden will tend to share cold-adaptive alleles at many causal loci as a result of selection, and marker alleles genome-wide as a result of demographic history”. Atwell et al. (2010) were referring to *A. thaliana* but our results and those of De Carvahlo et al. (2010) suggest that this may also fit well to Norway spruce and European aspen. In both species, the strong latitudinal clines observed at many SNPs may reflect the presence of different glacial refugia as well as different selection pressures.

In the present study, we have found evidence of clinal variation at SNPs in candidate genes using populations from a limited part of the Norway spruce natural range. We also observed that PaFTL2 expression varied along the latitudinal cline. Demonstrating that the identified SNPs are the truly causal for bud set, however, will remain a daunting task in a non-model species like Norway spruce. Ralph and Coop (2010) showed that parallel adaptation can occur within the same species, something that is of major interest in species like Norway spruce that have a fairly large global effective population size and a distribution that may allow for several clines to be established. Independent clines have been extensively used to study patterns of adaptation in Drosophila, and this seems to be a fruitful approach for detecting loci underlying local adaptation (e.g. Oakeshott et al. 1982; Turner et al. 2010). In order to validate the SNPs and loci identified in the present study, it would be highly interesting to test whether clinal variation at the loci identified here are also observed in other regions of the same species with a different history or in other species.

**Acknowledgements**: The research was supported by the European Community’s Sixth Framework Programme, under the Network of Excellence Evoltree and the
Seventh Framework Programme (FP7/ 2007-2013) under the grant agreement n° 211868 (Project Noveltree) and by the Swedish Research Council FORMAS grant dnr 217-2007-433 to ML. Part of the gene expression experiment was supported by the Nilsson-Ehle foundation. We thank Kerstin Jeppsson for help in the lab, David Clapham for help with plant management and VL Semerikov, YN Isakov, and the staff at the Swedish Forest Research Institute (Skogforsk) and at the Finnish Forest Research Institute (METLA) for seed samples. Thanks also to G Daoust (Canadian Forest Service) for supplying Arborea’s Norway spruce SNP discovery trees.

References


Atwell, S., Y. S. Huang, B. J. Vilhjálmsson, G. Willems, M. Horton et al., 2010
Genome-wide association study of 107 phenotypes in Arabidopsis thaliana
inbred lines. Nature 465: 627-631


Lobo, N.L., 2011 Clinal variation at putatively adaptive polymorphisms in mature populations of Sitka spruce (Picea sitchensis (Bong.) Carr.). MSc, University of British Columbia.


Whitlock M. C., and F. Guillaume, 2009 Testing for spatially divergent selection: comparing $Q_{ST}$ to $F_{ST}$. *Genetics* 183: 1055-1063


**Table 1.** Populations used in the present study

<table>
<thead>
<tr>
<th>Group</th>
<th>Population code</th>
<th>Population name</th>
<th>Country</th>
<th>#Individuals</th>
<th>Latitude (°N)</th>
<th>Longitude (°E)</th>
<th>Altitude (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpine domain</td>
<td>GE-47.0</td>
<td>Ruhpolding</td>
<td>Germany</td>
<td>30</td>
<td>47.00</td>
<td>12.23</td>
<td>1492</td>
</tr>
<tr>
<td></td>
<td>SE-58.3</td>
<td>Saleby</td>
<td>Sweden</td>
<td>15</td>
<td>58.37</td>
<td>13.15</td>
<td>70</td>
</tr>
<tr>
<td>Baltic-Nordic domain</td>
<td>RU-53.3</td>
<td>Bryansk</td>
<td>Russia</td>
<td>15</td>
<td>53.30</td>
<td>34.30</td>
<td>179</td>
</tr>
<tr>
<td>Central Fennoscandia</td>
<td>FI-61.5</td>
<td>StKu19</td>
<td>Finland</td>
<td>19</td>
<td>61.57</td>
<td>29.22</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>SE-61.6</td>
<td>Istevallen</td>
<td>Sweden</td>
<td>13</td>
<td>61.62</td>
<td>16.48</td>
<td>350</td>
</tr>
<tr>
<td></td>
<td>SE-61.8</td>
<td>Tjarnelund</td>
<td>Sweden</td>
<td>16</td>
<td>61.83</td>
<td>16.28</td>
<td>285</td>
</tr>
<tr>
<td></td>
<td>FI-62.0</td>
<td>StKu7</td>
<td>Finland</td>
<td>16</td>
<td>62.07</td>
<td>24.48</td>
<td>105</td>
</tr>
<tr>
<td></td>
<td>SE-62.6</td>
<td>Strangsund</td>
<td>Sweden</td>
<td>21</td>
<td>62.63</td>
<td>15.05</td>
<td>415</td>
</tr>
<tr>
<td></td>
<td>SE-62.7</td>
<td>Hogmansbod</td>
<td>Sweden</td>
<td>15</td>
<td>62.78</td>
<td>15.23</td>
<td>350</td>
</tr>
<tr>
<td></td>
<td>FI-63.0</td>
<td>StKu5</td>
<td>Finland</td>
<td>15</td>
<td>63.07</td>
<td>30.28</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td>SE-63.4</td>
<td>Mittberget</td>
<td>Sweden</td>
<td>14</td>
<td>63.48</td>
<td>17.67</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>SE-63.7</td>
<td>Hallen</td>
<td>Sweden</td>
<td>13</td>
<td>63.74</td>
<td>15.50</td>
<td>375</td>
</tr>
<tr>
<td></td>
<td>SE-64.1</td>
<td>Jämtland</td>
<td>Sweden</td>
<td>17</td>
<td>64.17</td>
<td>19.36</td>
<td>355</td>
</tr>
<tr>
<td></td>
<td>SE-65.3</td>
<td>Lillpite</td>
<td>Sweden</td>
<td>18</td>
<td>65.31</td>
<td>18.70</td>
<td>242</td>
</tr>
<tr>
<td></td>
<td>SE-66.4*</td>
<td>Jock</td>
<td>Sweden</td>
<td>48</td>
<td>66.41</td>
<td>22.44</td>
<td>145</td>
</tr>
<tr>
<td>Northern</td>
<td>FI-66.4</td>
<td>StKu2</td>
<td>Finland</td>
<td>12</td>
<td>66.40</td>
<td>26.88</td>
<td>175</td>
</tr>
<tr>
<td></td>
<td>FI-67.0</td>
<td>Kolari</td>
<td>Finland</td>
<td>27</td>
<td>67.04</td>
<td>26.36</td>
<td>264</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>----------</td>
<td>---------------</td>
<td>----------</td>
<td>-------</td>
<td>-----</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>Finland</td>
<td>FI-67.7</td>
<td>StKu19</td>
<td>Finland</td>
<td>17</td>
<td>67.72</td>
<td>26.05</td>
<td>280</td>
</tr>
<tr>
<td>FI-68.0</td>
<td>Muonio</td>
<td>Finland</td>
<td>Finland</td>
<td>10</td>
<td>68.00</td>
<td>24.15</td>
<td>342</td>
</tr>
</tbody>
</table>

*: SE-66.4 was used only in the expression study
Table 2: Summary of the SNP detected by the different approaches. See the main text for details on the different approaches.

<table>
<thead>
<tr>
<th>SNP</th>
<th>Clinality</th>
<th>MC sampling: # SNPs</th>
<th>BayeScan</th>
<th>Bayenv</th>
<th>Annotation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
<td>affecting the SNP</td>
<td>affected by</td>
<td>$\log_{10}(P.O.)$</td>
<td>$F_{ST}$</td>
</tr>
<tr>
<td>PaPHYN_RIII185</td>
<td>0.256</td>
<td>25</td>
<td>7</td>
<td>-0.360</td>
<td>0.043</td>
</tr>
<tr>
<td>PaPHYN_RIII88</td>
<td>0.510</td>
<td>0</td>
<td>12</td>
<td>-0.427</td>
<td>0.051</td>
</tr>
<tr>
<td>PaPHYO_RIII510</td>
<td>0.213</td>
<td>33</td>
<td>5</td>
<td>-0.353</td>
<td>0.054</td>
</tr>
<tr>
<td>PaPHYP_RIII274</td>
<td>0.404</td>
<td>0</td>
<td>14</td>
<td>2.743</td>
<td>0.139**</td>
</tr>
<tr>
<td>PaPHYP_RIII76</td>
<td>0.408</td>
<td>1</td>
<td>10</td>
<td>-0.409</td>
<td>0.045</td>
</tr>
<tr>
<td>PaMFTL1_1050</td>
<td>0.316</td>
<td>23</td>
<td>9</td>
<td>2.467</td>
<td>0.129**</td>
</tr>
<tr>
<td>PaMFTL1_2091</td>
<td>0.383</td>
<td>3</td>
<td>11</td>
<td>2.013</td>
<td>0.127**</td>
</tr>
<tr>
<td>PaMFTL1_2136</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>0.126</td>
<td>0.072</td>
</tr>
<tr>
<td>PaMFTL1_2215</td>
<td>0.285</td>
<td>31</td>
<td>9</td>
<td>-0.492</td>
<td>0.048</td>
</tr>
<tr>
<td>PaMFTL1_3441</td>
<td>0.359</td>
<td>18</td>
<td>7</td>
<td>1.788</td>
<td>0.122**</td>
</tr>
<tr>
<td>PaMFTL1_3658</td>
<td>0.555</td>
<td>0</td>
<td>12</td>
<td>0.333</td>
<td>0.080</td>
</tr>
<tr>
<td>PaMFTL1_802</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>1.469</td>
<td>0.127*</td>
</tr>
<tr>
<td>PaFTL2pr_1560</td>
<td>0.419</td>
<td>6</td>
<td>9</td>
<td>-0.417</td>
<td>0.050</td>
</tr>
<tr>
<td>PaFTL2pr_1824</td>
<td>0.493</td>
<td>0</td>
<td>15</td>
<td>-0.347</td>
<td>0.053</td>
</tr>
<tr>
<td>PaFTL2pr_1951</td>
<td>0.641</td>
<td>0</td>
<td>14</td>
<td>1.540</td>
<td>0.121**</td>
</tr>
<tr>
<td>Gene</td>
<td>Value 1</td>
<td>Value 2</td>
<td>Value 3</td>
<td>Value 4</td>
<td>Value 5</td>
</tr>
<tr>
<td>--------------</td>
<td>----------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>PaFTL2pr_2173</td>
<td>0.459</td>
<td>6</td>
<td>13</td>
<td>0.850</td>
<td>0.101 *</td>
</tr>
<tr>
<td>PaFTL2pr_2509</td>
<td>0.441</td>
<td>2</td>
<td>13</td>
<td>1.915</td>
<td>0.133 **</td>
</tr>
<tr>
<td>PaFTL2pr_2694</td>
<td>0.387</td>
<td>7</td>
<td>13</td>
<td>1.159</td>
<td>0.108 *</td>
</tr>
<tr>
<td>PaFTL2pr_2790</td>
<td>0.396</td>
<td>11</td>
<td>12</td>
<td>1.284</td>
<td>0.112 *</td>
</tr>
<tr>
<td>PaGI_F2_9_1470</td>
<td>0.543</td>
<td>0</td>
<td>7</td>
<td>0.596</td>
<td>0.090</td>
</tr>
<tr>
<td>PaGI_F2_9_987</td>
<td>0.479</td>
<td>1</td>
<td>8</td>
<td>1000</td>
<td>0.170 **</td>
</tr>
<tr>
<td>PaGI_F6_8_1111</td>
<td>0.244</td>
<td>1</td>
<td>6</td>
<td>-0.061</td>
<td>0.036</td>
</tr>
<tr>
<td>PaHB3_385</td>
<td>0.349</td>
<td>0</td>
<td>8</td>
<td>-0.116</td>
<td>0.036</td>
</tr>
<tr>
<td>PaKN2a_253</td>
<td>0.240</td>
<td>33</td>
<td>6</td>
<td>-0.411</td>
<td>0.043</td>
</tr>
<tr>
<td>PaKN2b_1157</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>0.702</td>
<td>0.020 *</td>
</tr>
<tr>
<td>PaKN3a_211</td>
<td>0.377</td>
<td>0</td>
<td>12</td>
<td>-0.521</td>
<td>0.045</td>
</tr>
<tr>
<td>PaKN4b_242</td>
<td>0.251</td>
<td>33</td>
<td>8</td>
<td>0.415</td>
<td>0.088</td>
</tr>
<tr>
<td>PaCCA1_1302</td>
<td>0.288</td>
<td>22</td>
<td>7</td>
<td>-0.381</td>
<td>0.043</td>
</tr>
<tr>
<td>PaPRR1_1992</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>-0.529</td>
<td>0.047</td>
</tr>
<tr>
<td>PaPRR1_2990</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>-0.503</td>
<td>0.048</td>
</tr>
<tr>
<td>PaPRR1_3301</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>-0.511</td>
<td>0.048</td>
</tr>
<tr>
<td>PaPRR3_F1_2570</td>
<td>0.380</td>
<td>0</td>
<td>7</td>
<td>-0.319</td>
<td>0.055</td>
</tr>
<tr>
<td>PaPRR3_F1_2978</td>
<td>/</td>
<td>/</td>
<td>/</td>
<td>1.722</td>
<td>0.123 **</td>
</tr>
<tr>
<td>PaPRR3_F2_331</td>
<td>0.309</td>
<td>25</td>
<td>8</td>
<td>-0.504</td>
<td>0.048</td>
</tr>
<tr>
<td>Gene</td>
<td>P-value</td>
<td>X</td>
<td>Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>PaPRR3_F2_481</td>
<td>0.354</td>
<td>1</td>
<td>8</td>
<td>-0.369</td>
<td>0.055</td>
</tr>
<tr>
<td>PaPRR7_F1_1505</td>
<td>0.455</td>
<td>0</td>
<td>11</td>
<td>-0.237</td>
<td>0.060</td>
</tr>
<tr>
<td>PaPRR7_F1_771</td>
<td>0.683</td>
<td>2</td>
<td>11</td>
<td>2.467</td>
<td>0.126 **</td>
</tr>
<tr>
<td>PaPRR7_F2_417</td>
<td>0.230</td>
<td>33</td>
<td>7</td>
<td>-0.374</td>
<td>0.053</td>
</tr>
<tr>
<td>PaPRR7_F2_534</td>
<td>0.673</td>
<td>2</td>
<td>11</td>
<td>1000</td>
<td>0.138 **</td>
</tr>
<tr>
<td>PaPRR7_F3_104</td>
<td>0.689</td>
<td>2</td>
<td>11</td>
<td>1000</td>
<td>0.161 **</td>
</tr>
<tr>
<td>PaZTL_793</td>
<td>0.352</td>
<td>7</td>
<td>7</td>
<td>-0.343</td>
<td>0.041</td>
</tr>
</tbody>
</table>

Note: * P-value < 0.05, ** P-value < 0.01; Log10(posterior odds): 0.5-1: substantial evidence; 1-1.5: strong evidence, 1.5-2: very strong, >2: decisive evidence; Bayes factors: 3.2-10: substantial evidence; 10-100: strong evidence, > 100: decisive evidence
Figure captions

**Figure 1** Geographic range of Norway spruce (*Picea abies*) and location of the population samples used in the present study. The shaded area is the distribution of Norway spruce (modified from EUFORGEN 2009, www.euforgen.org) and the dots are the geographical locations of the populations used in the present study.

**Figure 2** Percentage of individuals setting bud in five populations under different photoperiodic treatments. “9.5 darkness 1” and “9.5 darkness 2” are week 7 and week 8 within the same photoperiodic treatment (see Material and Methods).

**Figure 3** Clustering analysis conducted in STRUCTURE. a. $K = 2$ b. $K = 3$.

**Figure 4** Examples of candidate SNPs that show a significant regression of the transformed allele frequency on latitude.

**Figure 5** Analysis of $F_{ST}$ outliers using Bayescan (Foll and Gaggiotti 2008). Estimates of $F_{ST}$ for candidate SNPs (solid circles) and background SNPs (open circles) are plotted against the logarithm of the posterior odds.

**Figure 6** Mean level per population of the relative RNA expression of four putative photoperiodic genes under different photoperiodic treatments. Autoregressive models were used to obtain the curves, except for PaCCA1 for which the fit was poor and different points were just joined by a line. Open boxes represent light periods and shaded boxes represent dark periods. a. PaPRR7 b. PaGI c. PaFTL2 and d. PaCCA1.

**Figure 7** Correlation of the mean level of expression of PaFTL2 under a night length of 8-h with latitude ($P$-value = 0.048, adjusted $R^2 = 0.75$).
Fig. 2
Fig. 3
Fig. 4
Fig. 5
Fig. 7