



<http://www.diva-portal.org>

This is the published version of a paper published in *PLoS ONE*.

Citation for the original published paper (version of record):

Porth, I., Klapste, J., McKown, A D., La Mantia, J., Guy, R D. et al. (2015)

Evolutionary Quantitative Genomics of *Populus trichocarpa*.

PLoS ONE, 10(11): e0142864

<http://dx.doi.org/10.1371/journal.pone.0142864>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-113859>

RESEARCH ARTICLE

Evolutionary Quantitative Genomics of *Populus trichocarpa*

Ilga Porth^{1,2}, Jaroslav Klápště^{1,3}, Athena D. McKown¹, Jonathan La Mantia^{1,4}, Robert D. Guy¹, Pär K. Ingvarsson⁵, Richard Hamelin¹, Shawn D. Mansfield⁶, Jürgen Ehling⁷, Carl J. Douglas⁸, Yousry A. El-Kassaby^{1*}

1 Department of Forest and Conservation Sciences, University of British Columbia, Vancouver, BC V6T 1Z4, Canada, **2** Département des Sciences du Bois et de la Forêt, Faculté de Foresterie, de Géographie et de Géomatique, Université Laval, Québec, QC, G1V 0A6 Canada, **3** Department of Genetics and Physiology of Forest Trees, Czech University of Life Sciences, Prague, 165 21, Czech Republic, **4** Corn, Soybean and Wheat Quality Research Unit, United States Department of Agriculture, Wooster, Ohio, 44691 United States of America, **5** Department of Ecology and Environmental Science, Umeå University, Umeå, SE-901 87, Sweden, **6** Department of Wood Science, University of British Columbia, Vancouver, BC V6T 1Z4, Canada, **7** Department of Biology and Centre for Forest Biology, University of Victoria, Victoria, BC V8W 3N5, Canada, **8** Department of Botany, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

☞ These authors contributed equally to this work.

* y.el-kassaby@ubc.ca



OPEN ACCESS

Citation: Porth I, Klápště J, McKown AD, La Mantia J, Guy RD, Ingvarsson PK, et al. (2015) Evolutionary Quantitative Genomics of *Populus trichocarpa*. PLoS ONE 10(11): e0142864. doi:10.1371/journal.pone.0142864

Editor: Björn Hamberger, University of Copenhagen, DENMARK

Received: July 23, 2015

Accepted: October 27, 2015

Published: November 23, 2015

Copyright: © 2015 Porth et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are available within the paper and its Supporting Information files.

Funding: This work was supported by Genome British Columbia Applied Genomics Innovation Program (Project 103BIO) and Genome Canada Large-Scale Applied Research Project (Project 168BIO), funds to RDG, RCH, JE, SDM, CJD, and YE-K.

Competing Interests: The authors have declared that no competing interests exist.

Abstract

Forest trees generally show high levels of local adaptation and efforts focusing on understanding adaptation to climate will be crucial for species survival and management. Here, we address fundamental questions regarding the molecular basis of adaptation in undomesticated forest tree populations to past climatic environments by employing an integrative quantitative genetics and landscape genomics approach. Using this comprehensive approach, we studied the molecular basis of climate adaptation in 433 *Populus trichocarpa* (black cottonwood) genotypes originating across western North America. Variation in 74 field-assessed traits (growth, ecophysiology, phenology, leaf stomata, wood, and disease resistance) was investigated for signatures of selection (comparing Q_{ST} - F_{ST}) using clustering of individuals by climate of origin (temperature and precipitation). 29,354 SNPs were investigated employing three different outlier detection methods and marker-inferred relatedness was estimated to obtain the *narrow-sense* estimate of population differentiation in wild populations. In addition, we compared our results with previously assessed selection of candidate SNPs using the 25 topographical units (drainages) across the *P. trichocarpa* sampling range as population groupings. *Narrow-sense* Q_{ST} for 53% of distinct field traits was significantly divergent from expectations of neutrality (indicating *adaptive* trait variation); 2,855 SNPs showed signals of diversifying selection and of these, 118 SNPs (within 81 genes) were associated with adaptive traits (based on significant Q_{ST}). Many SNPs were putatively pleiotropic for functionally uncorrelated adaptive traits, such as autumn phenology, height, and disease resistance. Evolutionary quantitative genomics in *P. trichocarpa* provides an enhanced understanding regarding the molecular basis of climate-driven selection in forest trees and we highlight that important loci underlying *adaptive* trait variation also show relationship to climate of origin. We consider our approach the most comprehensive,

as it uncovers the molecular mechanisms of adaptation using multiple methods and tests. We also provide a detailed outline of the required analyses for studying adaptation to the environment in a population genomics context to better understand the species' potential adaptive capacity to future climatic scenarios.

Introduction

Knowledge about the genetic basis of *adaptive* quantitative traits in forest trees and genetic differentiation in response to selection facilitates the prediction of long-term responses to climate, but the genetic basis of adaptation is not comprehensively understood [1]. High levels of local adaptation due to consistent natural selection in a given environment resulted in local populations that have their highest fitness at their original provenance, and consequently, are differentiated from non-local populations. Within population diversity is fundamental to species survival in unpredictable environments, and therefore also relevant for conservation and forest management ([2,3]). Recent studies within forest trees have investigated the association of local climate and geography with either randomly identified loci (*Pinus taeda*: [4]; *Cryptomeria japonica*: [5], or candidate functional genes (*Picea abies*: bud set candidate genes, [6]; *Populus balsamifera*: flowering time candidate genes, [7]) to uncover genes underlying local adaptation. The genetic architecture underlying adaptive phenotypes of forest trees is generally highly complex (e.g. [8]). Therefore, untangling the relationships between adaptive loci and the role of climate in selection vs. neutral evolutionary processes is inherently difficult.

Evidence for potential adaptive significance of a genetic marker is often interpreted from ' F_{ST} outlier' analyses where genetic loci significantly differ in their allelic frequencies among populations. These 'outliers' can be efficiently detected using multilocus scans comparing patterns of nucleotide diversity and genetic differentiation to the simulated genome-wide neutral genetic background ([9,10]). For instance, this methodology has led to the detection of SNPs implicated in local climate adaptation in *Picea* ([11,12,13]). In order to obtain a detailed understanding of how populations have diverged in response to climate variation, such F_{ST} outliers can be tested for associations with an adaptive trait and an environmental variable to substantiate the evidence for their involvement in local adaptation ([14,15]). Integrating quantitative and population genomics is therefore essential to determine the degree to which genetic and phenotypic variation are driven by selection as opposed to neutral processes (e.g. genetic drift). Specifically, this allows for comprehensive information from genome-wide association studies (GWAS), Q_{ST} quantitative genetics analysis (i.e. 'top-down' approaches, [16]) and landscape population F_{ST} outlier analysis (i.e. 'bottom-up' approaches, [17]) be merged.

The existence of interaction effects among different loci within co-adapted gene complexes has long been recognized [18]. Yeaman (2013) suggested that ecological selection might even promote the physical clustering of locally adaptive loci through genomic rearrangements [19]. Landscape population genomics can identify genome regions significantly associated with spatial and temporal environmental gradients [3]. For instance, the study using natural *Arabidopsis* genotypes spanning the species' range revealed that local adaptation might be maintained by independent target loci enriched for molecular processes that exhibit their major genetic effects within distinct local environments but are neutral in others [20]. The geographic variation in the degree to which a genetic region under selection responds is termed "conditional neutrality" [21] and suggests a given species has not uniformly responded to an environmental pressure or that the pressure is not equally active across a species range. Importantly, the

(a) Temperature

(b) Precipitation

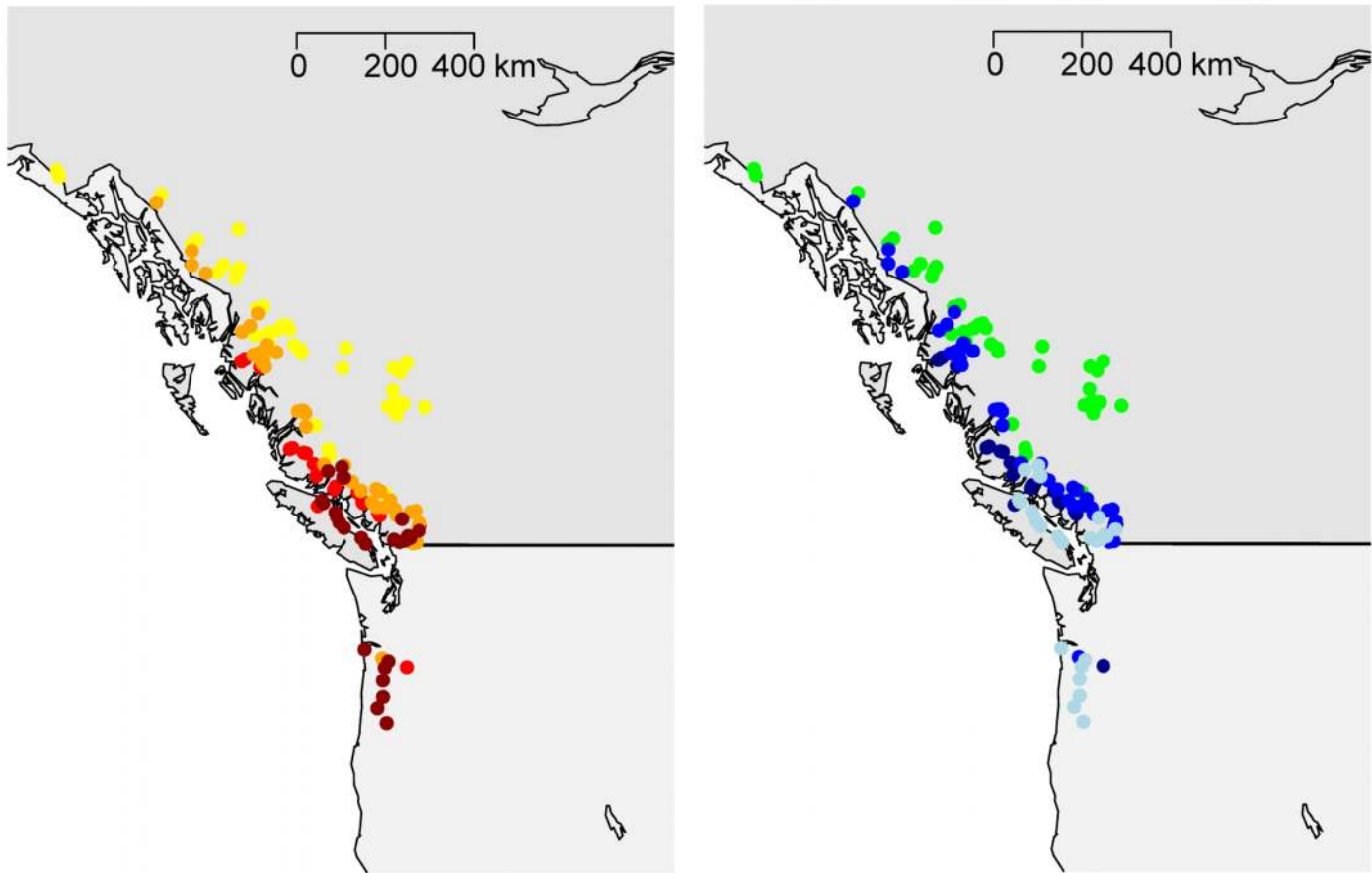


Fig 1. Geographical origins of 433 *P. trichocarpa* genotypes collected across 140 unique locations within the Pacific Northwest (British Columbia, Canada; Oregon, USA) and grouped into four distinct climate clusters using local temperature and precipitation records for location of origin. The climate regions were identified based on K-medoids clustering using the mean annual temperature (°C) between yrs 1971–2002 (MAT_1971–2002), the number of frost-free days (NFFD_1971–2002), and the mean annual precipitation (mm), observed between yrs 1971–2002 (MAP_1971–2002). Color coding is as follows: (a) population averages for MAT_1971–2002; NFFD_1971–2002: dark red (9.5°C; 287.1 d); red (8.1°C; 267.2d); orange (6.4°C; 215.2d); yellow (4.2°C; 175.4d); (b) population average for MAP_1971–2002: dark blue (2805.9mm); blue (1571.8mm); light blue (1517.0mm); green (744.2mm). We note here that canonical correlations between geography and ecology were high ($r = 0.9$ for the first canonical variable component).

doi:10.1371/journal.pone.0142864.g001

assessment of local adaptation in this work on *Arabidopsis* involves the study of fitness traits such as fecundity and survival (viability) ([20,22]). In addition, there also exist traits that increase fitness in one environment, but reduce it in another. Ecological genetics can more easily explore the genetic changes over time in annuals (due to their short generation times) involving multiple generations studied under a changing environment ([23,15]). This is less feasible for long-lived forest trees. However, the estimation of quantitative genetic parameters using SNP marker-inferred relatedness estimation to obtain *narrow-sense* estimates of heritability and Q_{ST} in wild populations [24] can allow monitoring adaptive genetic responses along an ecological time-scale [15].

In this study, we integrated an extensive body of results on the genetics of wild *Populus trichocarpa* Torr. & A. Gray (black cottonwood) to understand adaptation to climate. All poplars, aspens, and cottonwoods (genus *Populus*) play important roles in natural ecosystems as pioneer species ([25,26]) and are economically important for various industrial products with an

increasing role as bioenergy crops ([27,28,29,30]). *Populus* species are still largely undomesticated with very low population differentiation indicative of extensive long-distance intraspecific gene flow [31]. In western North America, *P. trichocarpa* has an extensive cordilleran range (31–62°N), yet with no clear north-south differentiation in genetic diversity (and no decreasing genetic diversity with latitude), consistent with the species' colonization history from multiple potential glacial refugia [32]. Several studies have indicated subtle sub-structure in *P. trichocarpa* ([33,34,35]) relating to isolation-by-distance (IBD; *i.e.* the decrease of genetic similarity among populations with increasing geographical distance between these populations reflected in *continuous* patterns of genetic differentiation and allele frequency variation in the species [34] as opposed to natural barriers causing discrete local genetic clusters), introgression and adaptation [36]. We explored the extensive body of data on the genetics of *P. trichocarpa*, including genome-wide coverage of SNPs [35], and comprehensive GWAS results from wood characteristics [37], leaf rust fungus (*Melampsora xcolombiana*) resistance [38], biomass, eco-physiology, leaf stomata and phenology traits [39]. We studied the divergence patterns of phenotypic variation and SNPs among distinct climate clusters in 433 unrelated *P. trichocarpa* genotypes originally collected throughout the northern two-thirds of the species' latitudinal range (excluding the highly diverged Californian population Tahoe: [34], [40]). We tested whether phenotypic variation in traits was diverged among the climatic regions (based on non-neutral Q_{ST}), as would be expected of adaptive variation. We then predicted that SNPs that are most diverged among different climatic regions would be associated with mapped genes that underlie *adaptive* phenotypic variation [13].

In brief, we used an integrative analysis of quantitative traits and genetic markers to investigate climate adaptation in wild *P. trichocarpa* populations, we developed an integrative approach through merging genomic-based datasets and results. (1) The effects of individual loci were first separated from confounding population effects using spatial PCA (sPCA) to investigate the presence of local and global genetic structures. Following this assessment of population structure using genetic markers showing evidence of only one single genetic structure, distinct population clusters were generated based on climatic factors and this sub-population clustering was used in subsequent analyses (Fig 1). (2) The genetic differentiation in quantitative traits (narrow-sense Q_{ST}) among populations defined by climate clusters was calculated involving the estimation of relatedness based on genetic markers. (3) In parallel, the divergence of genetic markers (F_{ST} outlier analysis) among populations defined by climate clusters was assessed. (4) The significance of quantitative trait divergence among populations, as defined by climate clusters, was assessed by comparing the observed Q_{ST} values with the simulated distribution of Q_{ST} - F_{ST} for a neutral trait. If the null hypothesis was rejected, the trait was considered adaptive. (5) GWAS results identifying the SNP variants underlying adaptive traits were incorporated. If these SNP variants also corresponded to loci under selection (employing four different outlier detection methods), then, the SNP variants were considered adaptive. This comprehensive analysis of genomic and phenotypic information underscores the necessity of merging multiple datasets to more fully understand evolutionary genomics of *P. trichocarpa*.

Results

Population structure assessment

Negative eigenvalues from sPCA were negligible (Fig 2), suggesting no local genetic clusters. By comparison, the presence of IBD was verified by large positive eigenvalues (Fig 2). These results were further confirmed by the local and global tests within the “adegenet” program (see [Methods](#)). While, again, we did not detect local genetic structure in *P. trichocarpa* (local test $P = 0.937$), we did identify global genetic structure attributed to IBD (global test $P = 0.001$) that

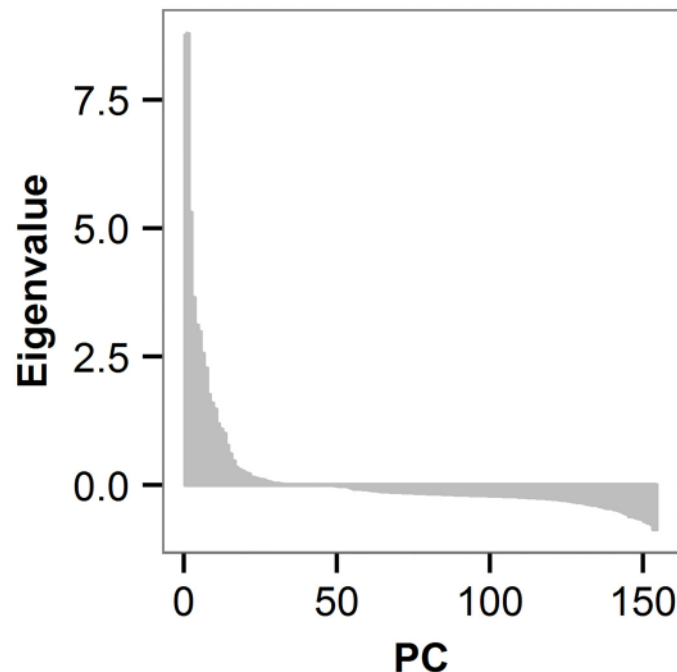


Fig 2. Identification of isolation-by-distance (IBD) among 433 *P. trichocarpa* genotypes based on spatial PCA. Large positive eigenvalues were indicative of IBD.

doi:10.1371/journal.pone.0142864.g002

was observed across the entire population involving the 140 unique geographical locations represented by one randomly chosen genotype.

Divergence of quantitative characters (Q_{ST}) among climate clusters

We calculated *narrow-sense* Q_{ST} values for 74 distinct field-assessed traits for the study population. Assessments included 16 wood, 12 biomass, 14 phenology, 18 ecophysiological, 13 leaf stomata, and one rust resistance phenotype (S1 Table). Observed Q_{ST} values for each trait were compared to the simulated distribution of $Q_{ST}-F_{ST}$ values for a neutral trait (simulating a range of possible demographic scenarios, see Methods). Among all traits, 53% (39/74 traits) had Q_{ST} values significantly different from zero and therefore were classified as *adaptive* (Table 1). The highest number of significant Q_{ST} values was observed among biomass traits (76%), phenology traits (70%), ecophysiology traits (64%) and leaf rust resistance (100%). By comparison, only 25% of wood-based traits had significant Q_{ST} values. Q_{ST} values for traits that significantly diverged among the four climate clusters ranged from 0.03 ($\delta^{15}N$, *i.e.* stable nitrogen isotope ratio) to 0.26 (bole biomass). Among all tested traits, the climatic clusters best explained the phenotypic variation in phenology based on the P_{ST} values, ranging from 17% (100% leaf yellowing) to 24% (bud set). Among wood characteristics, two cell wall sugar traits (% galactose and % arabinose in dry wood) and two wood ultrastructure attributes (fiber length and microfibril angle) showed significant Q_{ST} values. The climatic clusters explained 13 and 12% of the arabinose and galactose content, respectively.

Identification of SNPs under selection

Using both unsupervised and climate-based SPA, a total of 1,468 SNPs were identified being under selection at a 5% cutoff for each method (S2 Table). We also performed F_{ST} outlier

Table 1. h^2 , Q_{ST} , and h^2 corrected P_{ST} of adaptive traits ($P < 0.05$). Summary of 39 distinct adaptive traits of *P. trichocarpa* that diverged among different climate clusters (displayed are 59 tests for adaptation including tests for traits replicated in time, comprehensive results shown in S1 Table).

#	Trait	narrow-sense h^2	S.E.	narrow-sense Q_{ST}	S.E.	Variance explained by partitions	S.E.	P-value
1	Bole density_2012 ^a	0.4040	0.0402	0.0482	0.0522	0.0397	0.0429	0.0017
2	Bole mass_2012 ^a	0.1758	0.0430	0.2584	0.1788	0.1109	0.0877	0.0000
3	*Branches_2009 ^a	0.4898	0.0245	0.1567	0.1151	0.1541	0.1131	0.0000
4	H:D2+_2011 ^a	0.3753	0.0254	0.0321	0.0352	0.0243	0.0268	0.0178
5	*Height_2008 ^a	0.4540	0.0260	0.1133	0.0905	0.1040	0.0835	0.0000
6	*Height_2009 ^a	0.6543	0.0200	0.1132	0.0893	0.1432	0.1088	0.0000
7	*Height_2010 ^a	0.7378	0.0165	0.0900	0.0743	0.1274	0.1006	0.0000
8	*Height_2011 ^a	0.7092	0.0178	0.0792	0.0673	0.1087	0.0892	0.0000
9	*Height gain_2009 ^a	0.7504	0.0163	0.0952	0.0777	0.1364	0.1061	0.0000
10	*Height gain_2010 ^a	0.6217	0.0212	0.0477	0.0455	0.0586	0.0551	0.0019
11	*Height gain_2011 ^a	0.3372	0.0250	0.0490	0.0483	0.0337	0.0335	0.0016
12	Whole tree mass_2012 ^a	0.2279	0.0434	0.2323	0.1634	0.1225	0.0953	0.0000
13	*Volume_2009 ^a	0.3663	0.0256	0.1159	0.0925	0.0877	0.0718	0.0000
14	*Volume_2010 ^a	0.4519	0.0253	0.0945	0.0783	0.0862	0.0718	0.0000
15	*Volume_2011 ^a	0.5091	0.0243	0.0900	0.0751	0.0915	0.0760	0.0000
16	*Volume gain_2010 ^a	0.4441	0.0254	0.0913	0.0763	0.0820	0.0689	0.0000
17	*Volume gain_2011 ^a	0.4396	0.0253	0.0923	0.0771	0.0822	0.0691	0.0000
18	Amax/mass_2009 ^b	0.1349	0.0264	0.1822	0.1396	0.0579	0.0493	0.0000
19	Amax_2009 ^b	0.1916	0.0261	0.0596	0.0604	0.0240	0.0248	0.0007
20	Chlsummer_2009 ^b	0.2692	0.0292	0.1160	0.0968	0.0663	0.0577	0.0000
21	Chlsummer_2011 ^b	0.3078	0.0288	0.1438	0.1135	0.0939	0.0777	0.0000
22	C:N_2009 ^b	0.1631	0.0270	0.1423	0.1156	0.0518	0.0454	0.0000
23	d15N_2009 ^b	0.0882	0.0232	0.0257	0.0395	0.0047	0.0072	0.0446
24	Dleaf_2009 ^b	0.4872	0.0272	0.0269	0.0299	0.0263	0.0291	0.0371
25	gs_2009 ^b	0.4243	0.0279	0.0402	0.0401	0.0344	0.0343	0.0055
26	Leaves per bud_2011 ^b	0.3307	0.0310	0.0767	0.0695	0.0523	0.0482	0.0001
27	Leaves per bud_2012 ^b	0.4786	0.0297	0.0910	0.0765	0.0875	0.0735	0.0000
28	*LMAsummer_2010 ^b	0.2360	0.0281	0.0628	0.0644	0.0307	0.0322	0.0000
29	Narea_2009 ^b	0.1907	0.0278	0.0479	0.0525	0.0189	0.0211	0.0028
30	Nmass_2009 ^b	0.1592	0.0271	0.1409	0.1150	0.0500	0.0441	0.0000
31	WUE_2009 ^b	0.2457	0.0274	0.0731	0.0667	0.0373	0.0350	0.0000
32	AUDPC-2009 ^c	0.5322	0.0245	0.0490	0.0470	0.0521	0.0495	0.0017
33	AUDPC-2010 ^c	0.3937	0.0260	0.0723	0.0646	0.0579	0.0523	0.0002
34	AUDPC-2011 ^c	0.3132	0.0251	0.0848	0.0740	0.0551	0.0492	0.0001
35	*Active growth rate_2009 ^d	0.6094	0.0222	0.0390	0.0393	0.0471	0.0469	0.0083
36	*Bud set_2008 ^d	0.5970	0.0224	0.1390	0.1051	0.1617	0.1186	0.0000
37	*Bud set_2009 ^d	0.7390	0.0165	0.1790	0.1262	0.2438	0.1580	0.0000
38	*Bud set_2010 ^d	0.6483	0.0200	0.1708	0.1224	0.2108	0.1434	0.0000
39	Bud set186_2009 ^d	0.5247	0.0234	0.1988	0.1368	0.2067	0.1403	0.0000
40	Bud set186_2010 ^d	0.4041	0.0268	0.2125	0.1444	0.1792	0.1261	0.0000
41	*Height growth cessation_2009 ^d	0.7114	0.0178	0.1434	0.1072	0.1923	0.1354	0.0000
42	*Leaf drop_2008 ^d	0.5175	0.0244	0.1533	0.1137	0.1579	0.1160	0.0000
43	*Leaf drop_2009 ^d	0.5168	0.0237	0.2335	0.1525	0.2396	0.1547	0.0000
44	*Leaf drop_2010 ^d	0.5965	0.0214	0.1453	0.1088	0.1687	0.1225	0.0000
45	*Leaf lifespan_2010 ^d	0.6278	0.0208	0.0432	0.0419	0.0537	0.0514	0.0039
46	Canopy duration_2009 ^d	0.2409	0.0253	0.0944	0.0809	0.0480	0.0428	0.0000
47	*Canopy duration_2010 ^d	0.8119	0.0126	0.0462	0.0438	0.0729	0.0671	0.0024
48	Growth period_2009 ^d	0.3176	0.0255	0.1046	0.0862	0.0693	0.0589	0.0000
49	*Growth period_2010 ^d	0.7095	0.0176	0.1365	0.1032	0.1833	0.1308	0.0000
50	*Post-bud set period_2009 ^d	0.4222	0.0260	0.0332	0.0352	0.0282	0.0299	0.0187
51	*Post-bud set period_2010 ^d	0.5230	0.0237	0.1432	0.1075	0.1489	0.1106	0.0000
52	*100% Yellowing_2010 ^d	0.5886	0.0220	0.1498	0.1113	0.1718	0.1240	0.0000
53	*75% Yellowing_2010 ^d	0.5640	0.0227	0.0638	0.0571	0.0714	0.0632	0.0002
54	Arabinose ^e	0.8786	0.2227	0.0749	0.0707	0.1276	0.1079	0.0002
55	Fiber ^e	0.3027	0.2423	0.0825	0.1135	0.0446	0.0515	0.0000

(Continued)

Table 1. (Continued)

#	Trait	narrow-sense h^2	S.E.	narrow-sense Q_{ST}	S.E.	Variance explained by partitions	S.E.	P-value
56	Galactose ^e	0.9327	0.2089	0.0663	0.0621	0.1167	0.1002	0.0000
57	MFA1 ^e	0.4074	0.2383	0.0403	0.0539	0.0355	0.0419	0.0054
58	Ad_StomataNUM1 ^f	0.3165	0.0266	0.1229	0.0984	n.d.	n.d.	0.0129
59	Ad_STM_distribution ^f	0.1779	0.0351	0.1050	0.1041	n.d.	n.d.	0.0357

Note: P-value obtained by comparison of the observed $Q_{ST}-F_{ST}$ to the quantile of the simulated $Q_{ST}-F_{ST}$ distribution for a neutral trait [96].

^abiomass trait [45]

^becophysiology trait [45]

^cleaf rust resistance trait [38]

^dphenology trait [45]

^ewood trait [37]

^fleaf stomata traits [44]

*spatially adjusted trait [45]

^{||}the variance explained by climate clusters compared to the total variance was estimated as h^2 corrected P_{ST}

S.E. refers to standard errors

Active growth rate (cm day⁻¹)

Ad_StomataNUM1: Adaxial stomata numbers

Ad_STM_distribution: Adaxial stomata distribution

Amax/mass = photosynthetic rate per unit dry mass (μmol CO₂ mg⁻¹ s⁻¹)

Arabinose in dry wood (%)

AUDPC = (calculated) area under the disease curve, based on *M. xcolombiana* infection rating

Bole density (kg/m³)

Bole mass (kg)

Branch #

Bud set (day)

Bud set[†] (day): bud set dates considered only after summer solstice

C:N = carbon:nitrogen (mg mg⁻¹)

Canopy duration (days)

Chlsummer = chlorophyll content index (CCI)

D15N = stable nitrogen isotope ratio (‰)

Dleaf = net discrimination (‰)

Fiber: fiber length Lw (mm)

Galactose in dry wood (%)

Growth period (days)

gs = stomatal conductance (mol H₂O m⁻² s⁻¹)

H:D = height to diameter (cm:cm)

Height (cm)

Height gain (cm)

Height growth cessation (day)

Leaf drop (day)

Leaf lifespan (days)

Leaves per bud (#)

LMA = leaf mass per unit area (mg mm⁻²)

MFA1: microfibril angle at most recent growth ring (°)

Narea = nitrogen (mg mm⁻²)

Nmass = nitrogen (mg mg⁻¹)

Post-bud set period (days)

Volume (cm³)

Volume gain (cm³)

Whole tree mass (kg)

WUE = instantaneous water-use efficiency (μmol CO₂ mmol⁻¹ H₂O)

Yellowing, 100% (day)

Yellowing, 75% (day)

doi:10.1371/journal.pone.0142864.t001

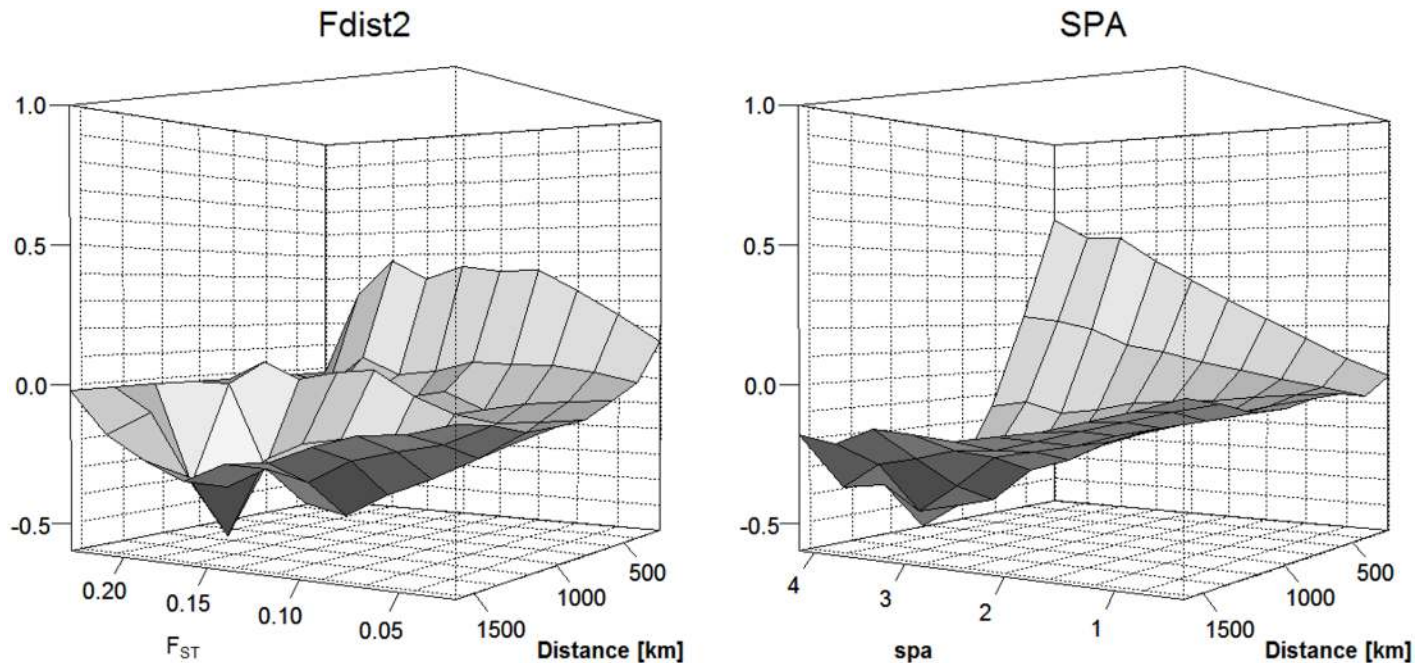


Fig 3. Comparison of two outlier detection methods (F_{ST} , SPA) for their efficiency to identify genetic selection signals under isolation-by-distance (IBD). Gene dispersal was tested employing Moran's test for spatial autocorrelation using 200km lags.

doi:10.1371/journal.pone.0142864.g003

analysis on climate clusters. While the mean F_{ST} value for the complete dataset (29,354 SNPs) was 0.0108, we obtained a mean neutral F_{ST} value (0.0078) after removing loci identified to be potentially under selection [41]. In the final analysis, all loci were tested against this neutral mean to identify a set of potential F_{ST} outliers relating to climate. Using 200k simulations in Fdist2, we identified 121 SNPs outside the 99% limits of the neutral distribution (S1 Fig) as potential candidates subjected to diversifying (positive) selection related to the four climate clusters. Among these, 88% of these climate-related 'outliers' were confirmed by allelic frequency correlation analysis with averages for climate variables within subpopulation (using multiple univariate logit regression models in SAM ($\alpha = 0.05$, S2 Table)), 77 of these loci persisted across different selection scan scenarios employed (unsupervised SPA, climate-based SPA, and F_{ST} analysis based on population subdivision [36]), and 48 SNPs were retrieved using association genetics (see below) (S2 Table). A comparison between Fdist and SPA testing gene dispersal and employing Moran's test for spatial autocorrelation (Fig 3) indicates, in general, the higher effectiveness of SPA to identify genetic selection signals under patterns of IBD.

A significant accumulation of F_{ST} outliers was identified on chromosome 15 (S1 Fig). The extent of linkage disequilibrium (LD) between all 121 outlier loci is presented in S2 Fig. In general, we found that LD was not substantial between SNPs from different genes. Incomplete LD can be caused by the possibility that SNPs are close to but not in complete LD with the causal variants (here probably due to 'tag SNP' design of the SNP chip array [35]) explaining why the observed LD between diverged loci is generally low [42]. One notable exception is two neighboring poplar genes (Potri.009G008600 and Potri.009G008500) initially annotated based on sequence homology to *Arabidopsis* genes as nitrate transporter types *ATNRT2:1* and *ATNRT2:4*, respectively. The allele frequencies of three SNPs and one SNP, respectively, in poplar orthologs of *ATNRT2:1* and *ATNRT2:4*, respectively, are strongly correlated to temperature ($R^2 > 0.9$; $P = 0.05$), while the remaining SNPs in both genes did not follow such a strong pattern (S2 Fig).

SNPs under diversifying selection and associated with quantitative traits

To corroborate findings of candidate loci putatively under diversifying selection based on climate, we compared these results with SNPs uncovered by associations with adaptive traits (showing non-neutral Q_{ST}). Among four GWAS studies in *P. trichocarpa*, a total of 619 SNPs had been identified with significant trait associations (at $\alpha = 0.05$): 410 with biomass, ecophysiology and phenology [39], 141 with wood property traits [43], 40 with *Melampsora xcolombiana* resistance [38], and 28 SNPs related to leaf stomata variation [44].

We compared four different outlier analyses to identify selection signals in 29,354 SNPs. Most trait-associated SNPs for which we detected selection signals were associated with adaptive traits (89%, S2 Table). The highest percentage of trait-associated SNPs in outlier analyses was found for climate-based F_{ST} outlier analysis (40% of the total number of outliers identified by the method; 48 SNPs), followed by geography-based F_{ST} outlier analysis (8%; 75 SNPs that were reported in [36]), unsupervised SPA (5%; 75 SNPs), and SPA with climate as a covariate (3%; 37 SNPs). In total, selection signals were detected for 151 trait-associated SNPs with 44% overlap among evaluation methods. Interestingly, there was a lack of genome-wide correlation between selection and association signal (Fig 4) and thus only dispersed association signals were detected among SPA selection signals (Fig 5, S2 Table). This result is probably a consequence of the structure correction methods employed in GWAS.

We retrieved a number of unique but also shared SNPs among the different analyses (Fig 6). Shared SNPs were highest for climate F_{ST} (75%) and geography-based F_{ST} (72%). Unsupervised SPA had the highest number of unique SNPs among the four methods (51%). We found 118 SNPs associated with adaptive traits (significant Q_{ST}) including 59 SNPs under diversifying selection shared among at least two outlier detection methods and 59 unique SNPs detected by climate F_{ST} , climate SPA and unsupervised SPA, respectively (S3 Table). A large number of SNPs (40%) that we identified as F_{ST} outliers using climate clustering were candidate SNPs

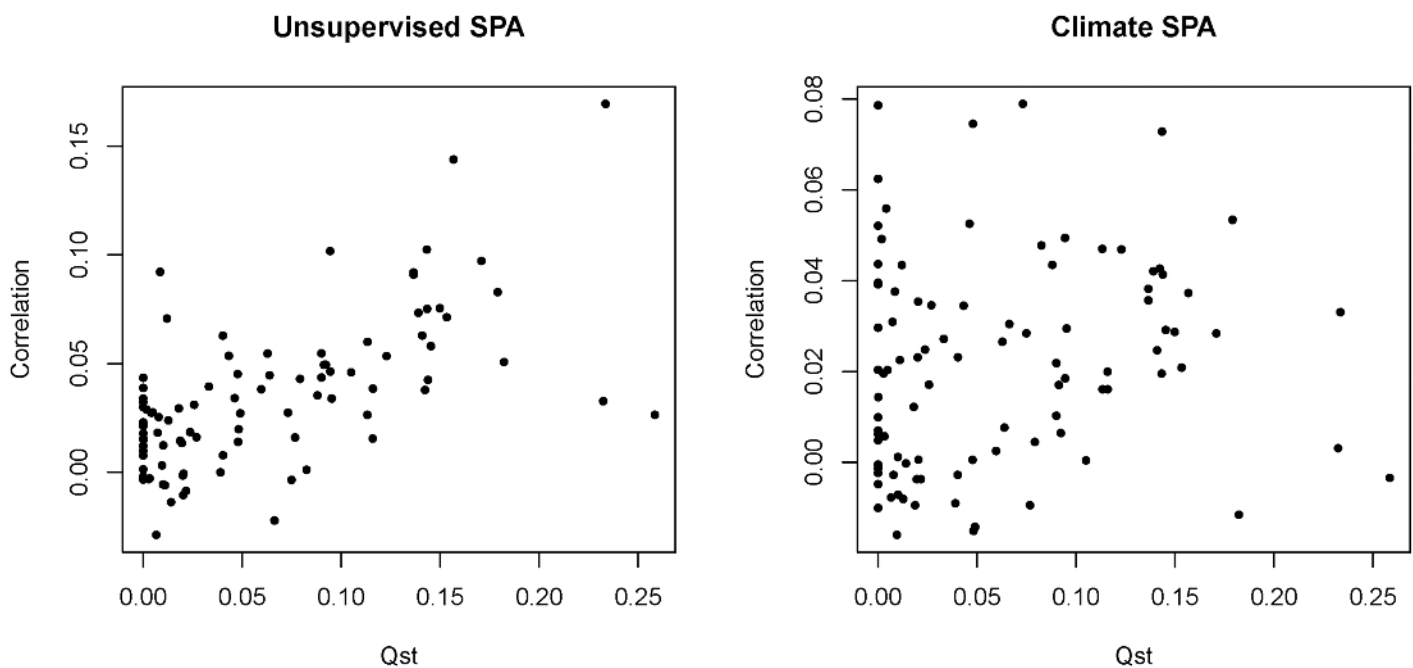


Fig 4. Genome-wide correlations between selection outliers and association signals based on 29k SNPs. Correlation of $-\log(P)$ versus spa was plotted against the trait's Q_{ST} .

doi:10.1371/journal.pone.0142864.g004

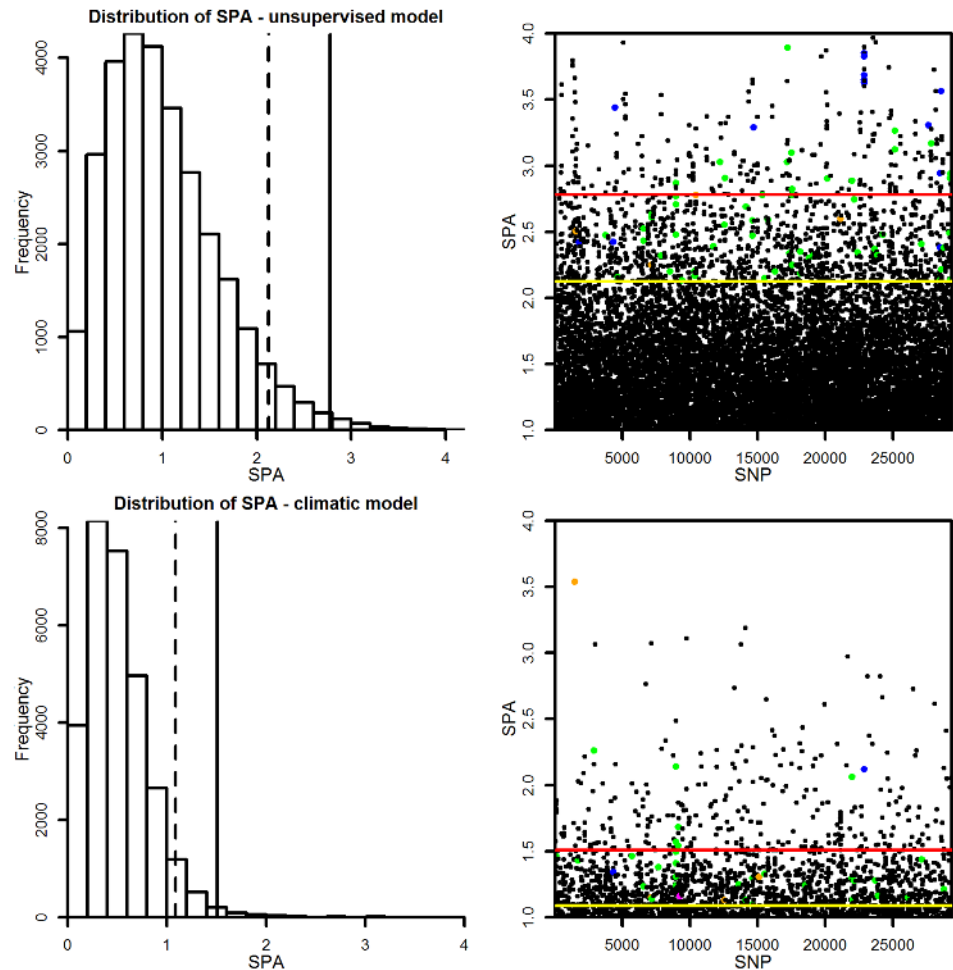


Fig 5. Individual SNPs under diversifying selection within genes mapping to quantitative trait variation. 5% cutoff: dashed and yellow lines; 1% cutoff: solid and red lines; ecology (biomass, ecophysiology, phenology, stomata)—green dots; wood properties (orange); rust resistance (blue).

doi:10.1371/journal.pone.0142864.g005

from association studies (S2 Table). The high number of trait-associated SNPs reflects both the polygenic nature of phenotypic traits (e.g., c.200 for bud set, [39]) and linkage disequilibrium (LD) to a lesser extent. The highest number of climate-based F_{ST} outliers associated with adaptive traits was found on chromosome 15 (12 SNPs), identifying a genomic region where SNPs putatively under selection to local climate generally may be clustered (S1 Fig).

We found that SNPs under potential climate selection matching putative causal variants from association studies consistently mapped to non-neutral Q_{ST} , adaptive traits (S1 and S2 Tables). Only one SNP associated with wood traits (within Potri.009G006500 annotated as *FRA8* associated with fiber length, [43]) was among the F_{ST} outlier loci. Comparatively, phenology traits were the most complex adaptive traits from the high match between the total number of associated SNPs and the proportion of SNPs with allele frequencies significantly diverged among climate clusters (S2 Table). In total, 118 SNPs were outliers under diversifying selection, associated with adaptive traits (significant Q_{ST}), and with many SNPs putatively pleiotropic for functionally uncorrelated adaptive traits, such as autumn phenology, height, and disease resistance (S3 Table). The 78 annotated poplar genes were largely derived from major gene functional group such as (1) transcription factors of several categories and (2) carbohydrate-related

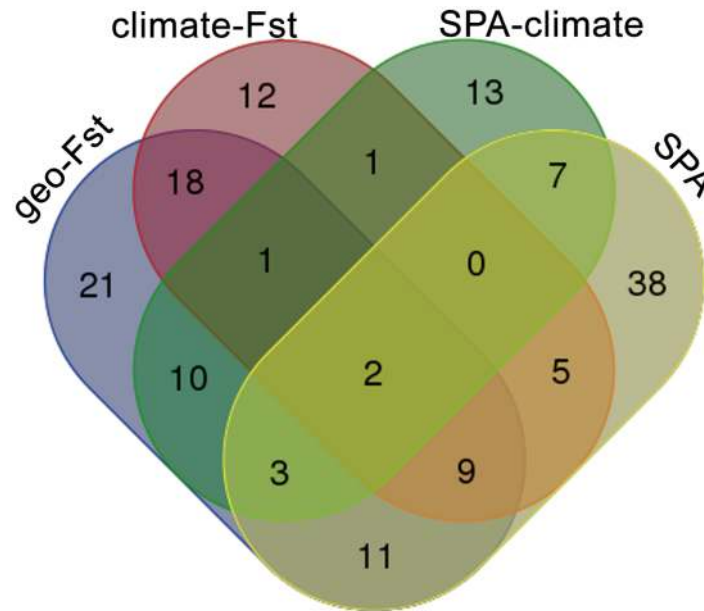


Fig 6. Venn diagram showing the numbers of unique and shared SNPs (totaling 151 trait-associated SNPs) among four different outlier detection approaches. F_{ST} using climate clusters, F_{ST} using geographical grouping, SPA analyses—with climate-based PCs incorporated as covariates and unsupervised, respectively. A subset of this information (118 SNPs) related to genetic polymorphisms associated solely with adaptive trait variation is provided in [S3 Table](#).

doi:10.1371/journal.pone.0142864.g006

genes, but also transporters. Among these transporters, two poplar genes (Potri.009G008600 and Potri.009G008500) annotated based on sequence homology to *Arabidopsis* genes as nitrate transporter types *ATNRT2:1* and *ATNRT2:4*, respectively, were highly pleiotropic for several adaptive traits ([S3 Table](#)).

Discussion

Evolutionary quantitative genomics

The main focus of our work involved identifying adaptive traits and their genetic basis in forest trees by employing both a quantitative genetics approach (Q_{ST} analysis) and population genomics [16] to uncover SNPs under strong selection (among c.29k tested genetic polymorphisms). Our analyses revealed that 53% of these traits produced significant narrow-sense Q_{ST} ([S1 Table](#)) underscoring that such quantitative traits are very likely related to adaption to local climatic conditions [45].

This study uses SNP marker-inferred relatedness estimation (*i.e.* the ‘animal model’) to obtain narrow-sense estimates of heritability and Q_{ST} in wild populations [24]. The quality of genetic estimates using the ‘animal model’ approach largely depends on the accuracy of relationship coefficient estimates and are affected by: 1) number and quality of markers [46], 2) variance in actual relatedness [47], and 3) how well the relationship estimates reflect the segregation of causal variants [48] Our present study is based on extensive, genome-wide SNPs [35] which can provide high accuracy for both the relationship coefficients and the estimated genetic parameters. However, samples from natural tree populations are subject to intensive gene flow (outcrossing) and generally show low levels of relatedness which can negatively affect heritability and Q_{ST} analyses.

Heritability is usually dependent on the population sampled (*i.e.* the observed allele frequency differences) and thus, can differ for smaller sampling sizes and/or specific sampling areas (*e.g.*, central vs. marginal regions of species distribution). Heritability estimates taken across a greater coverage of the species distribution are more likely to reflect evolutionary history of the traits (stabilizing vs. diversifying selection) rather than the effects of population subsampling. Sufficient variance in the actual relatedness is also required to reveal heritability in wild populations [47], although heritability, and indirectly, Q_{ST} estimates, can suffer from the inability to separate the pure additive genetics from environmental effects, specifically when relatedness is lacking. Thus, the presence of LD between markers and causal variants (QTLs) is crucial to recover the genetic parameters with sufficient precision. In the case of traits under diversifying selection, the additive genetic variance estimates (such as *narrow-sense* heritability) may also include a substantial QTL covariance component, in addition to the pure genetic variance. This is especially the case when many QTLs follow the same cline, and can further extend the additive genetic variance when the QTLs interact (*i.e.*, epistasis) [49] unless the epistasis is accounted for in the model [50]. Thus, heritability estimates for traits under diversifying selection (Table 1) may be upwardly biased (see below).

Heritability estimates are often interpreted as the capacity for adaptive evolution. In addition, epistatic interactions, specifically, the directional epistasis, have major effects through altering the genetic background (both, the additive genetic variances and the covariances, *i.e.* the allelic frequencies but also their effects) [51]. Hemani *et al.* (2013) outlined that for traits under selection, high levels of genetic variation are maintained and the traits evolve more slowly than expected, yet this could be attributed to high epistasis in traits under strong diversifying selection [42].

Selectively non-neutral genetic variants underlying traits adaptive to climate

Overall, the number of F_{ST} outlier SNPs underlying an adaptive trait correlated well with the total number of candidate SNPs associated with that trait ($r = 0.625$, $P = 0.0005$). Yet, the majority of trait associated SNPs were not F_{ST} outliers (S2 Table) and appeared to be unresponsive to selection for different climatic conditions, especially for phenology traits such as bud set, leaf drop or growth period. A previous simulation study suggested that differentiation in candidate loci is limited for complex traits in forest trees (*i.e.*, their F_{ST} values are similar to neutral values), despite their strong adaptive divergence among local populations (high Q_{ST}), due to large population sizes and high levels of gene flow [52]. Thus, highly polygenic adaptation (as observed in complex genetic traits) will not show sufficient allele frequency differentiation such that climatic clines in SNPs of candidate genes can be exhaustively detected.

We modelled the spatial structure of genetic variation using SPA (addressing gene flow under IBD), and SNPs identified via SPA were compared against GWAS-identified SNPs, climate-related F_{ST} outliers and geography-informed F_{ST} outliers. The majority of SNPs with steep allele frequency clines (based on unsupervised SPA) uncovered allele frequency correlations with the north-south cline (S2 Table). We noted that enrichment for particular genes, such as circadian rhythm/clock genes, was found in PC1 (a north-south population structure) [45] and that SNPs of these genes were among the highest ranked in SPA. Nonetheless, associations of circadian rhythm clock genes with strong correlations to environment were largely missing among the identified genetic associations for phenology traits (discussed in McKown *et al.* [39]). The interplay of IBD and natural selection was lost by the necessary structure correction in GWAS, however, evidence from gene expression or gene regulation that is also highly correlated with the trait under question might be possible to retrieve such SNPs of putative importance (Anonymous, [53]).

The presence of IBD in *P. trichocarpa* underscores the larger issue for investigating wild populations with quantitative genetics and population genomics approaches as IBD can confound population structure, association mapping, and outlier analyses. The power to detect local selection depends on several factors, including selection strength, the presence of distinct types of microenvironment heterogeneity, and the distance of gene dispersal compared to the overall spatial scale [54]. In our case, as the observed gene dispersal is ~500 km (Fig 3) and sampling is also discontinuous (Fig 1), this does not allow us to perform F_{ST} analysis on arbitrarily defined local populations because it will be more difficult to separate the stochastic noise (drift, migration) from the selection signal in smaller scale population subsampling leading to an excess of false positives [54]. Yet, selection pressures can differ along environmental clines. Thus, F_{ST} outliers should be investigated on the largest scale possible following the spatial distribution of the environment in order to identify spatial genetic structure. Nevertheless, IBD in wild populations will create some compromised statistical power in detecting local adaptation using specific pairs of populations that is unavoidable (Fig 3).

Polygenic and pleiotropic adaptation relating to climate

Our climate clustering partitioned the study population into four large, evenly-sized groups of individuals lending robustness to SNP detection even for lower frequency (recent) variants. In our study, the top two SNPs among climate related F_{ST} outliers showed strongest associations to climate partitions according to SAM analysis [Potri.010G250600 (*MSR2/ MANNAN SYNTHESIS RELATED 2* implicated in carbohydrate metabolism) and Potri.010G254400 (transporter *ATGCN4*) (S2 Table)]. In addition, six genes that harboured climate-related F_{ST} outlier SNPs have been identified as candidates for bud set in previous studies ([55,56]), yet these loci were not associated with bud set in our GWAS study ([39]; S2 Table), possibly through implementing the conservative population structure correction term in GWAS. Nevertheless, these genes may represent additional candidates for bud set, including Potri.003G218900 (*ACD1-LIKE*), Potri.009G015100 (senescence-associated family protein), Potri.014G170400 (*XER-ICO*), Potri.015G012500 (*IQ-domain 21*), Potri.018G015100 (chloroplast nucleoid DNA-binding protein), and Potri.019G078400 (leucine-rich repeat transmembrane protein kinase) (S2 Table).

Evidence is emerging that for perennial trees to effectively sense short day signals, *i.e.* critical day length in autumn phenology [57], a temperature optimum is required and genetically predetermined by the local climate of the individual's origin [58]. Allele frequencies for most of the SNPs that both associated with bud set and diverged among the climate clusters showed strong regression on the mean temperature variation of the climatic clusters (R^2 up to 0.94; S2 Table). A critical role for temperature, rather than precipitation, on bud set has also been found in *Picea* [12]. For autumn phenology, elevated temperatures can either accelerate or delay growth cessation depending on species or ecotype ([59,60]), but under climate warming, the overall effects on phenological timing in forest trees is unknown.

SNP allelic frequencies within both nitrate transporter genes *ATNRT2:4* and *ATNRT2:1* were strongly aligned with temperature variation ($R^2 \sim 90\%$) in *P. trichocarpa*. Moreover, these SNPs were pleiotropic for multiple autumn phenology traits, height, and leaf rust resistance (S3 Table). Nitrate transporters are generally important in plants, as nitrate is the main nitrogen source required for synthesis of nucleic and amino acids. Therefore, a regulation of nitrate distribution is crucial to modulate growth (biomass acquisition) in response to temperature or light conditions ([61,62]). Interestingly, there are only two poplar representatives within a phylogenetic sub-clade of NRT2 that is populated by as many as five Arabidopsis sequences (*ATNRT2.1/2.2/2.3/2.4/2.6*). This implies that a deletion event occurred in this clade whose

functional significance remains elusive to date [62]. Phylogenetic reconstruction coupled with gene expression analysis point at neo/subfunctionalisation of the two poplar nitrate transporters for long distance nitrate transport from roots, wood to leaves [62]. This acquisition of novel expression pattern and loss of the ancestral expression pattern demonstrates the signature of adaptive evolution in functional diversification in paralogous gene pairs [63].

In addition, our results revealed that adaptive genetic variants within both poplar nitrate transporters were also associated with leaf rust resistance ([38]; S3 Table). In *Arabidopsis*, loss of function of *ATNRT2.1* primes salicylic acid signaling and *PR1* up-regulation [64]. In poplar leaf rust inoculations, both *PTNRT2.4* and *PTNRT2.1* are strongly down-regulated in incompatible interactions, while no expression change is apparent in compatible interactions (J. La Mantia, personal observation). The identified nitrogen transporters might be important in nitrogen storage and nitrogen remobilization to recycle nutrients during the progression of leaf senescence [65]. They may also function in down-regulation of nitrogen assimilation during seasonal remodeling of tree phenology related to growth cessation induced by short photoperiods ([66,67]) and/or temperature [58]. The effect of temperature on rust aggressiveness is noted [68] and the climatic conditions which form a conducive environment for rust infection and disease duration likely provide a strong adaptive selection toward resistance.

Pectin esterase gene Potri.012G014500 (SNP scaffold_12_1811250) represents another example for which significant associations with climate (here: temperature) and several adaptive traits were found (S2 and S3 Tables). In fact, the allelic effects of this SNP related to hyposomaty also related to less rust infection ([45]). This is an illustrative example regarding the tradeoff between carbon gain and pest resistance under favourable climatic conditions relating to pathogen pressure ([45]).

Conclusions

The high adaptive potential of tree populations is considered the result of positive effects of long-distance gene flow based on its interactions with divergent selection across the contrasting environments [69], while local adaptation in forest trees with regards to climate-related traits is polygenic and recent [70]. For instance, interactions between temperature and photoperiodic cues were shown to influence bud set for short-term acclimation in poplar [58]. By combining quantitative genetics and population genomics analyses, our study contributes to an enhanced understanding of the molecular basis of adaptation to different local climate in an undomesticated perennial species (*P. trichocarpa*). The key findings provided SNPs whose allelic frequencies were most diverged among populations from different climate clusters and these SNPs tended to be associated with mapped genes underlying phenotypic variation. This phenotypic variation itself diverged among the different climate clusters. Our study dissected the influence of climate (specifically, temperature and precipitation), yet much of the variation in phenology is also attributed to photoperiod ([71,72,45]). The tight photoperiodic control of traits such as bud set, height growth cessation, and leaf senescence ([73,74,59]) is crucial both for resistance to cold temperatures and maximization of the growing season, particularly in trees originating from high-latitude and/or high elevation provenances ([75,56]). While we tested the influence of climate on the variation of other traits in *P. trichocarpa*, such as wood and biomass, we consider other local factors, such as soil condition (pH and minerals), soil/root microbial diversity, groundwater, and other ecological interactions also of potential importance. Reciprocal transplants will be necessary to elucidate the effects of gene \times environment plasticity on the expression of traits with spatially heterogeneous selection [76], but can focus on specific genes identified through a combined quantitative genomics analysis, such as the one proposed here. Forthcoming research can also scale trait-to-performance mapping in known pedigrees for the

assessment of SNP effects on fitness [77]. These findings will have important implications for the future management of natural forests, acting to guide efforts in facilitated adaptation to climate change via measure such as assisted gene flow [78].

Materials and Methods

Collection, genotyping, and phenotyping of *P. trichocarpa*

The germplasm used in this study was propagated under permission from the British Columbia Ministry of Forests, Lands & Natural Resource Operations. No endangered or protected species were used in this study. Plant material was collected from a population of 433 *P. trichocarpa* Torr. & A. Gray genotypes growing in a common garden. These genotypes came from 140 unique geographic locations spanning two thirds of the species' range (44–60°N, 121–138°W) ([79], Fig 1). Originally collected by the BC Ministry of Forests, Lands and Natural Resource Operations, individual genotypes were grown in two common gardens, Surrey, BC and Totem Field, University of British Columbia, BC. Genotypes were replicated across the two field gardens and the Totem Field individuals (established in 2008 [80]) were clonal propagations from Surrey site individuals (established in 2000 [79]).

Trees were genotyped using an Illumina iSelect array with 34,131 SNPs from 3,543 candidate genes designed for *P. trichocarpa* [35]. The characteristics of the poplar genome and array development are outlined in [35]. Briefly, the SNP array was designed to include genes of known importance (*i.e.* candidate genes) or genes based on expression analyses. Because of the rate of linkage disequilibrium (LD) decay in *P. trichocarpa*, between 67–134k SNPs would be required to include all common variants throughout the genome at LD = 0.2 (assuming a 403 Mb assembled genome length and an average of 3–6 kb for r^2 between common variants to drop to 0.2). Therefore, some SNPs were selected as representative SNPs to “tag” genes and genetic regions with high LD, and thus represent a group of SNPs (the haplotype). For this study, we further filtered array SNPs for: i) minor allele frequency (MAF) <0.05, ii) >10% missing data, and iii) Illumina's GenTrain score <0.5, thereby reducing SNP numbers to 29,354. This filtering is not biased towards higher frequency SNPs (*i.e.* older variants established at much higher frequencies within the population over time) as a wide distribution of allele frequencies (MAF>0.05) was considered for the analysis.

Phenotyping of genotype accessions within the common gardens and climate of origin data were obtained from previously published work (for full phenotyping details, see [38,37], [45]). In brief, phenology, ecophysiology, biomass [45], leaf stomatal anatomy [44] and leaf rust (*Melampsora xcolumbiana*) resistance traits [38] were repeatedly measured from accessions planted at the University of British Columbia's research field through replication in space (clonal ramets) and in time (measurements across years). Wood chemistry and ultrastructure traits were measured from wood cores of the nine-year-old ortets representing the same genotypes and growing in Surrey [37].

Assessment of population structure

Since forest tree species usually have extensive geographic ranges, exhibit extensive gene flow and have low levels of population stratification [81], we investigated whether the genetic variability due to non-random mating in our population was caused solely by isolation-by-distance (IBD), reflecting the large geographical distribution of our sample (cf. [36]), or also by natural barriers causing local genetic clusters. We performed spatial principal component analysis (sPCA) by using the “spca” function implemented in the R package “adegenet” [82] which is a spatially explicit multivariate analysis accounting for spatial autocorrelation processes and patterns of genetic variation. A K-nearest neighbours method with K = 10 was used as connection

network. Positional information for each genotype were transformed into Universal Transverse Mercator (UTM) coordinates using “convUL” in the R package “PBSmapping” [83]. Due to the occurrence of multiple genotypes with identical geographical coordinates (*i.e.* trees collected at the same latitude/longitude), we randomly selected a single genotype representing a geographical region (out of the total 140 locations). Eigenvalues for principal components from sPCA provided a cumulative picture about contributing factors, including the genetic variance and the spatial autocorrelation (through Moran’s I, see below). Large positive eigenvalues reflect the importance of the proportion of the genetic variance along with a strong positive autocorrelation in the global pattern (*i.e.* IBD), while large negative eigenvalues indicate the importance of the proportion of the genetic variance along with negative autocorrelation indicating the existence of discrete local genetic clusters.

We used the “global.test” and “local.test” functions in the “adegenet” package to infer the statistical significance of each type of genetic structure. These functions are based on a spectral decomposition of the connection matrix into Moran’s eigenvector map and test for association of those eigenvectors from Moran’s eigenvector map with Moran’s I [82]. To investigate gene dispersal, we employed a Moran I test for spatial autocorrelation ([84,54]). Moran’s I coefficients were investigated in 200 km spatial lags and the analysis was performed using “morantest” in the “spdep” R package [85]. Moran’s I coefficients were estimated as follows:

$$I = \frac{n}{\sum_{i=1}^n \sum_{j=1}^n w_{ij}} * \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \tag{1}$$

where *n* is the number of populations (*i.e.* unique geographical locations), *w_{ij}* is weight set at 0 or 1 depending on whether populations are considered neighbours in each 200 km lag test, *x_i* is the allele frequency in the *i*th population, and \bar{x} is the allele frequency across all populations.

Climatic zone clustering of *P. trichocarpa*

Since our initial investigation of population structure with sPCA indicated the presence of only one global structure consisting of IBD and lack of local discrete clusters, any marker-based inference about genetic clusters might be highly unreliable [86]. Therefore, we established population differentiation on the basis of climate envelopes ([12]). Clusters of individual genotypes were defined using climate of origin measures (*i.e.* independently of the genetic data). Climate variables were obtained using ClimateWNA [87] and included mean annual temperature (MAT; °C), number of frost-free days (NFFD), and mean annual precipitation (MAP; mm). Climate data were based on positional information (latitude, longitude, elevation) and 1971–2002 Canadian Climate Normals [45]. Using K-medoids clustering and the Calinski-Harabasz criterion [88], we split the study population into four groups with relatively balanced sample sizes of 87, 103, 142, and 101 representing climate classes #1–4, respectively. Clusters generally followed the western North American coastline inwards (Fig 1a & 1b).

Genetic differentiation in quantitative characters among populations defined by climate clustering

We tested phenotypic characteristics in *P. trichocarpa* for their adaptive potential (S1 Table). For *Q_{ST}*–*F_{ST}* comparisons, *Q_{ST}* values among the identified climate-related population groups were first estimated for each trait following [89] and [24], respectively.

The *narrow-sense Q_{ST}* was estimated by computing the variance components using the ‘animal model approach’ [90] following:

$$y = \mathbf{X}\beta + \mathbf{Z}p + \mathbf{Z}a + e \tag{2}$$

where β is a vector of fixed effects (intercept), p and a are vectors of random climate cluster and individual tree additive genetic effects, \mathbf{X} and \mathbf{Z} are incidence matrices assigning fixed and random effects to measurements in vector y , the cluster effects are following $p \sim N(0, \sigma_p^2)$ where σ_p^2 is the cluster variance, individual tree additive effects are following $a \sim N(0, \sigma_a^2 \mathbf{G})$ where σ_a^2 is the additive genetic variance and \mathbf{G} is the realized relationship matrix [91], using 29,354 SNPs estimated in R package “synbreed” [92] as follows:

$$\mathbf{G} = \frac{\mathbf{Z}\mathbf{Z}'}{2 \sum p(1-p)} \tag{3}$$

where \mathbf{Z} is $\mathbf{M}\mathbf{P}$, with \mathbf{M} the marker matrix with genotypes recoded into 0, 1 and 2 for the reference homozygote allele, the heterozygote and the alternative homozygote allele, respectively, and with \mathbf{P} the vector of doubled allele frequency; e is the vector of random residual effects following $e \sim N(0, \sigma_e^2 \mathbf{I})$ where σ_e^2 is the residual variance and \mathbf{I} is the identity matrix. The narrow sense Q_{ST} was estimated as follows:

$$Q_{ST} = \frac{\hat{\sigma}_p^2}{(\hat{\sigma}_p^2 + 2\hat{\sigma}_a^2)} \tag{4}$$

where $\hat{\sigma}_p^2$ and $\hat{\sigma}_a^2$ are the estimates of cluster and additive genetic variance representing among- and within-group trait variances attributable to additive effects.

The measurements of all ecology and disease traits using clonal ramets (*i.e.* replication) enable estimating broad-sense Q_{ST} directly without the use of any relationship matrix, while narrow-sense Q_{ST} estimation was based on variance components estimated in the mixed linear model considering the realized relationship matrix [91] as in Eq 2. The model is identical to Eq 2 where the variance components for broad-sense Q_{ST} were estimated in the model considering a as the vector of clonal genotypic values following $a \sim N(0, \sigma_a^2 \mathbf{I})$ where σ_a^2 is the total genetic variance (including both additive and non-additive component) and e as the vector of ramet within clone effects following $e \sim N(0, \sigma_e^2 \mathbf{I})$. Then, the computed Q_{ST} values for each trait were compared to the average population differentiation estimate (F_{ST}) strictly based on neutral markers (see below) allowing inferences about trait evolution based on selection or genetic drift (neutral trait), [93].

Narrow-sense heritability (h^2) was based on variance components estimated in the mixed model as follows:

$$y = \mathbf{X}\beta + \mathbf{Z}a + e \tag{5}$$

where β is the vector of fixed effects (intercept and cluster) and a is the random vector of additive genetic effects following the description of Eq 2. The narrow-sense heritability was estimated as follows:

$$\hat{h}^2 = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}_e^2} \tag{6}$$

where $\hat{\sigma}_a^2$ and $\hat{\sigma}_e^2$ are estimates of additive genetic and residual variance, respectively. The phenotypic Q_{ST} (*i.e.* P_{ST}) ([89,24]) was estimated as follows:

$$P_{ST} = \frac{\hat{\sigma}_p^2}{(\hat{\sigma}_p^2 + 2\hat{h}^2\hat{\sigma}_e^2)} \tag{7}$$

where $\hat{\sigma}_p^2$ and $\hat{\sigma}_e^2$ are estimates of cluster and residual variance representing among- and

within-population variances, respectively, and \hat{h}^2 is the heritability estimated according to [37]. The variance components were estimated in ASReml software [94] using the mixed linear model following:

$$y = \mathbf{X}\beta + \mathbf{Z}p + e \quad (8)$$

where β is the vector of fixed effects (intercept) and p is the vector of random cluster effects, the effect of individuals within cluster is found within the error variance.

Identification of non-neutral SNPs and quantitative traits divergent among climate clusters

To identify SNPs putatively under selection and also associated with adaptive traits ([38,43,39]), we performed: 1) F_{ST} outlier analysis (using Fdist2) employing the same climate clusters as for Q_{ST} analysis, 2) unsupervised spatial ancestral analysis (SPA), and 3) SPA with climate as a covariate. Additionally, we compared our results with F_{ST} outlier analysis (using Fdist2 and BayeScan) that were reported in [36] using 25 topographic units separated by watershed barriers within the geographic area from Central Oregon, USA (44.3°N) to northern BC, Canada (59.6°N).

F_{ST} values for SNPs were calculated among the four climate clusters (for definition and calculation, see above). We implemented the Fdist2 program within the LOSITAN project [41] for SNP F_{ST} outlier detection. Fdist2 compares the distribution of F_{ST} values of sampled loci to the modeled neutral expectation of F_{ST} distribution using coalescent simulations [9]. We employed the infinite alleles mutation model (as we investigated SNPs), a subsample size of 50, and ran 200k simulations. F_{ST} values conditioned on heterozygosity and outside the 99% confidence interval were considered candidate outliers.

Since *P. trichocarpa* populations have known structure related to IBD ([36] and this study), we applied spatial ancestral analysis (SPA), a logistic regression-based approach [86], to detect SNPs with sharp allelic frequency changes across geographical space (implying candidates under selection). The unsupervised learning approach (using only genomic data) was employed to obtain SPA statistics. In addition, we tested SPA including the first two principal components (PCs) based on climate variables (explaining 91% of the variance) as covariates to determine individuals' location based on allele frequencies related to MAT, NFFD, and MAP climate components.

We investigated correlations between the outlier SNPs (based on climate clusters) and the environmental variables that defined the established climatic clusters (Fig 1). Subpopulation averages for MAT, NFFD, and MAP were tested for correlations with SNP allele frequencies employing multiple univariate logistic regression models with the spatial analysis method (SAM; [95]). The significance of correlations was assessed using three independent statistical tests (likelihood ratio and two Wald tests) implemented in SAM and applying an initial 95% confidence interval for the statistical tests. We used the Bonferroni correction method ($\alpha = 0.05$) for multiple testing resulting in $p < 6.887052 \times 10^{-5}$ for 726 tested models (242 alleles, three variables). Only those correlations that remained significant after Bonferroni correction for each of the three test statistics (*i.e.* the likelihood ratio and the two Wald tests) were retained.

Finally, we compared observed Q_{ST} values with the simulated distribution of Q_{ST} - F_{ST} values for a neutral trait using previously provided R scripts [96]. In brief, a range of possible demographic scenarios was tested simulating the distribution of Q_{ST} values based on mean F_{ST} for neutral markers and mean Q_{ST} for neutral traits ([97,98]). For a neutral trait, the expected Q_{ST} was estimated based on $\hat{\sigma}_p^2$ (*i.e.*, measured within-population variance; see above) and $\hat{\sigma}_a^2$ (*i.e.*, expected between-population variance) given in Eq 4. The distribution of σ_p^2 values was based

on σ_a^2 and the observed F_{ST} values of 29,233 SNPs present (total number reduced by removing outliers) within the simulated *neutral* envelope of F_{ST} values (F_{ST} outlier analysis) with Q_{ST} replaced by the F_{ST} in Eq 4. P -values were obtained by testing whether the null hypothesis that the estimated *narrow-sense* Q_{ST} for each tested trait is statistically equal to the expected Q_{ST} for a neutral trait [96].

Marker-trait association mapping

In previous analyses of marker-trait associations in *P. trichocarpa*, confounding effects of population stratification were adjusted using principal component analysis ([38,43,39] and a Q matrix population structure correction [39]. Phenological mismatch within the common garden can confound trait values [45], thus, association analyses included “area under the disease curve” resistance measures with adjustment for bud set [38] and all ecophysiological traits that were measured prior to bud set [39]. The Unified Mixed Model (a modification of the generalized linear model) was employed for marker-trait association mapping and is fully described ([38,43,39]). While necessary, the adjustment for confounding, cryptic genetic structure in the association analyses may have reduced the statistical power to detect associations. This is particularly problematic in species whose distribution is mainly along a one-dimensional cline or for which differentiation in ecological traits covaries with the species demographic history ([13,45]). Furthermore, the GWAS results may be biased towards common variants or variants with the greatest effects. This is related to the size of the SNP discovery panel (34k) [99] and the power to detect significant associations given the tested population sizes (334–448 individuals). As whole genome sequencing and phenotyping of thousands of genotypes would be required to comprehensively uncover the genetic architecture of complex traits, we consider the GWAS results informative but not exhaustive.

Supporting Information

S1 Fig. F_{ST} outlier loci detection in *P. trichocarpa* and distribution of outliers along the poplar chromosomes. (a) F_{ST} outlier loci detection and distribution of empirical F_{ST} estimates conditioned on expected heterozygosity (H_E). The envelope of values corresponding to neutral expectations at 99% CI level (with mean $F_{ST} = 0.0078$), solid line, was constructed with the infinite allele model according to [9]. (b) Distribution of the empirical F_{ST} estimates along the 19 poplar chromosomes and additional scaffolds (abbrev: scaff); the 121 identified outlier loci are indicated by red circles above their F_{ST} value bars. A goodness-of-fit test assuming a uniform distribution was performed to test whether the observed frequencies of ‘outlier loci’ along the 19 poplar chromosomes differed significantly from the expected value. Following the rejection of the null hypothesis (chi-square = 81.98 df = 18, P -value = 3.85e-10), we declared ‘outlier loci hotspots’ if the number of outliers at a given chromosome was equal or above the maximum value (*i.e.*, 20) for assessed outlier clusters from a randomly generated data set using the 118 outliers found across the 19 chromosomes, and running 1,000 replicates, which identified significant clustering of outliers on chromosome 15.

(PDF)

S2 Fig. Linkage disequilibrium between 121 identified F_{ST} outlier loci and relationship between F_{ST} outlier allele frequencies and climate variables in *P. trichocarpa*. Simple linear regression (R^2) of allelic frequencies (following arcsine transformation) on temperature and precipitation, respectively (mean annual temperature in °C: MAT_1971–2002; number of frost-free days: NFFD_1971–2002 and mean annual precipitation in mm: MAP_1971–2002, observed between yrs 1971–2002) calculated among the four distinct climate clusters (Fig 1);

Note: POPTR_0143s00200 was recently re-annotated to Potri.009G008500 and both genes are now assembled on chromosome 9 within 50kb of each other (new poplar genome assembly Phytozome v3). Both sequences are now described as tandem gene pair *PTNRT2.4A* (alias Potri.009G008600) and *PTNRT2.4B* (alias Potri.009G008500) with 97% DNA sequence similarity [62].

(PDF)

S1 Table. Comprehensive population differentiation estimates and h^2 corrected P_{ST} for *P. trichocarpa*: broad-sense and narrow-sense Q_{ST} for 58 distinct field traits; Q_{ST1} and narrow-sense Q_{ST} (Q_{ST2}) estimates for 16 wood traits.

(XLS)

S2 Table. Comprehensive summary table of all SNP detection results from GWAS [ecology [39]; rust [38]; stomata [44]; wood [43]] and outlier analysis (geographic F_{ST} [36], this study: climate F_{ST} , unsupervised SPA, climate SPA) for the black cottonwood population (presented in Fig 1) and using the 34k SNP chip [35]; adaptive traits (significant Q_{ST}) are in bold. In red and dark blue are 1% cutoffs ($spa = 2.78025$ and $spa = 1.50795$), in orange and light blue are 5% cutoffs ($spa = 2.12467$ and $spa = 1.08868$) in unsupervised SPA and climate SPA analyses, respectively.

(XLSX)

S3 Table. List of 118 SNPs associated with adaptive traits (significant Q_{ST} for at least one associated trait) including 59 SNPs under diversifying selection shared among at least two outlier detection methods and 59 unique SNPs detected by climate F_{ST} , climate SPA and unsupervised SPA, respectively. Comprehensive results are provided in S2 Table.

(XLS)

Acknowledgments

The authors thank Dr. Julien Prunier for help with 'Spatial analysis method' software.

Author Contributions

Conceived and designed the experiments: YE-K RDG RCH JE SDM CJD. Performed the experiments: IP ADM JL. Analyzed the data: JK IP. Contributed reagents/materials/analysis tools: PI. Wrote the paper: IP JK YE-K.

References

1. Savolainen O, Lascoux M, Merila J. Ecological genomics of local adaptation. *Nature Review Genetics*. 2013; 14(11):807–20.
2. Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S. Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*. 2008; 1(1):95–111. doi: [10.1111/j.1752-4571.2007.00013.x](https://doi.org/10.1111/j.1752-4571.2007.00013.x) PMID: [25567494](https://pubmed.ncbi.nlm.nih.gov/25567494/)
3. Allendorf FW, Hohenlohe PA, Luikart G. Genomics and the future of conservation genetics. *Nature Reviews Genetics*. 2010; 11(10):697–709. doi: [10.1038/nrg2844](https://doi.org/10.1038/nrg2844) PMID: [20847747](https://pubmed.ncbi.nlm.nih.gov/20847747/)
4. Eckert AJ, Bower AD, Gonzalez-Martinez SC, Wegrzyn JL, Coop G, Neale DB. Back to nature: ecological genomics of loblolly pine (*Pinus taeda*, Pinaceae). *Molecular Ecology*. 2010; 19(17):3789–805. doi: [10.1111/j.1365-294X.2010.04698.x](https://doi.org/10.1111/j.1365-294X.2010.04698.x) PMID: [20723060](https://pubmed.ncbi.nlm.nih.gov/20723060/)
5. Tsumura Y, Uchiyama K, Moriguchi Y, Ueno S, Ihara-Ujino T. Genome scanning for detecting adaptive genes along environmental gradients in the Japanese conifer, *Cryptomeria japonica*. *Heredity*. 2012; 109(6):349–60. doi: [10.1038/hdy.2012.50](https://doi.org/10.1038/hdy.2012.50) PMID: [22929151](https://pubmed.ncbi.nlm.nih.gov/22929151/)
6. Chen J, Kallman T, Ma X, Gyllenstrand N, Zaina G, Morgante M, et al. Disentangling the Roles of History and Local Selection in Shaping Clinal Variation of Allele Frequencies and Gene Expression in

- Norway Spruce (*Picea abies*). *Genetics*. 2012; 191(3):865–81. doi: [10.1534/genetics.112.140749](https://doi.org/10.1534/genetics.112.140749) PMID: [22542968](https://pubmed.ncbi.nlm.nih.gov/22542968/)
7. Keller SR, Levsen N, Olson MS, Tiffin P. Local Adaptation in the Flowering-Time Gene Network of Balsam Poplar, *Populus balsamifera* L. *Molecular Biology and Evolution*. 2012; 29(10):3143–52. PMID: [22513286](https://pubmed.ncbi.nlm.nih.gov/22513286/)
 8. Holliday JA, Ralph SG, White R, Bohlmann J, Aitken SN. Global monitoring of autumn gene expression within and among phenotypically divergent populations of Sitka spruce (*Picea sitchensis*). *New Phytologist*. 2008; 178(1):103–22. doi: [10.1111/j.1469-8137.2007.02346.x](https://doi.org/10.1111/j.1469-8137.2007.02346.x) PMID: [18194148](https://pubmed.ncbi.nlm.nih.gov/18194148/)
 9. Beaumont MA, Nichols RA. Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B-Biological Sciences*. 1996; 263(1377):1619–26.
 10. Eveno E, Collada C, Guevara MA, Leger V, Soto A, Diaz L, et al. Contrasting patterns of selection at *Pinus pinaster* Ait. drought stress candidate genes as revealed by genetic differentiation analyses. *Molecular Biology and Evolution*. 2008; 25(2):417–37. PMID: [18065486](https://pubmed.ncbi.nlm.nih.gov/18065486/)
 11. Namroud M-C, Beaulieu J, Juge N, Laroche J, Bousquet J. Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. *Molecular Ecology*. 2008; 17(16):3599–613. doi: [10.1111/j.1365-294X.2008.03840.x](https://doi.org/10.1111/j.1365-294X.2008.03840.x) PMID: [18662225](https://pubmed.ncbi.nlm.nih.gov/18662225/)
 12. Prunier J, Laroche J, Beaulieu J, Bousquet J. Scanning the genome for gene SNPs related to climate adaptation and estimating selection at the molecular level in boreal black spruce. *Molecular Ecology*. 2011; 20(8):1702–16. doi: [10.1111/j.1365-294X.2011.05045.x](https://doi.org/10.1111/j.1365-294X.2011.05045.x) PMID: [21375634](https://pubmed.ncbi.nlm.nih.gov/21375634/)
 13. Holliday JA, Suren H, Aitken SN. Divergent selection and heterogeneous migration rates across the range of Sitka spruce (*Picea sitchensis*). *Proceedings of the Royal Society B-Biological Sciences*. 2012; 279(1734):1675–83.
 14. Luikart G, England PR, Tallmon D, Jordan S, Taberlet P. The power and promise of population genomics: From genotyping to genome typing. *Nature Reviews Genetics*. 2003; 4(12):981–94. PMID: [14631358](https://pubmed.ncbi.nlm.nih.gov/14631358/)
 15. Hansen MM, Olivieri I, Waller DM, Nielsen EE, Ge MWG. Monitoring adaptive genetic responses to environmental change. *Molecular Ecology*. 2012; 21(6):1311–29. doi: [10.1111/j.1365-294X.2011.05463.x](https://doi.org/10.1111/j.1365-294X.2011.05463.x) PMID: [22269082](https://pubmed.ncbi.nlm.nih.gov/22269082/)
 16. Sork VL, Aitken SN, Dyer RJ, Eckert AJ, Legendre P, Neale DB. Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes*. 2013:1–11.
 17. Stinchcombe JR, Hoekstra HE. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*. 2008; 100(2):158–70. PMID: [17314923](https://pubmed.ncbi.nlm.nih.gov/17314923/)
 18. Endler JA. Geographic variation, speciation, and clines. *Monographs in population biology*. 1977; 10:1–246. PMID: [409931](https://pubmed.ncbi.nlm.nih.gov/409931/)
 19. Yeaman S. Genomic rearrangements and the evolution of clusters of locally adaptive loci. *Proceedings of the National Academy of Sciences*. 2013; 110(19):E1743–51.
 20. Fournier-Level A, Korte A, Cooper MD, Nordborg M, Schmitt J, Wilczek AM. A Map of Local Adaptation in *Arabidopsis thaliana*. *Science*. 2011; 334(6052):86–9. doi: [10.1126/science.1209271](https://doi.org/10.1126/science.1209271) PMID: [21980109](https://pubmed.ncbi.nlm.nih.gov/21980109/)
 21. Schnee FB, Thompson JN. Conditional neutrality of polygene effects. *Evolution*. 1984; 38(1):42–6.
 22. Hancock AM, Brachi B, Faure N, Horton MW, Jarymowycz LB, Sperone FG, et al. Adaptation to Climate Across the *Arabidopsis thaliana* Genome. *Science*. 2011; 334(6052):83–6. doi: [10.1126/science.1209244](https://doi.org/10.1126/science.1209244) PMID: [21980108](https://pubmed.ncbi.nlm.nih.gov/21980108/)
 23. Anderson JT, Willis JH, Mitchell-Olds T. Evolutionary genetics of plant adaptation. *Trends in Genetics*. 2011; 27(7):258–66. doi: [10.1016/j.tig.2011.04.001](https://doi.org/10.1016/j.tig.2011.04.001) PMID: [21550682](https://pubmed.ncbi.nlm.nih.gov/21550682/)
 24. Pujol B, Wilson AJ, Ross RIC, Pannell JR. Are Q(ST)-F(ST) comparisons for natural populations meaningful? *Molecular Ecology*. 2008; 17(22):4782–5. doi: [10.1111/j.1365-294X.2008.03958.x](https://doi.org/10.1111/j.1365-294X.2008.03958.x) PMID: [19140971](https://pubmed.ncbi.nlm.nih.gov/19140971/)
 25. Eckenwalder JE. Systematics and evolution of *Populus*. Stettler RF B H, Heilman PE, Hinkley TM, editor. National Research Council of Canada Ottawa, ON, Canada: NRC Research Press; 1996.
 26. Cronk QCB. Plant eco-devo: the potential of poplar as a model organism. *New Phytologist*. 2005; 166(1):39–48. PMID: [15760349](https://pubmed.ncbi.nlm.nih.gov/15760349/)
 27. Carroll A, Somerville C. Cellulosic Biofuels. *Annual Review of Plant Biology*. 2009; 60:165–82. doi: [10.1146/annurev.arplant.043008.092125](https://doi.org/10.1146/annurev.arplant.043008.092125) PMID: [19014348](https://pubmed.ncbi.nlm.nih.gov/19014348/)
 28. Sannigrahi P, Ragauskas AJ, Tuskan GA. Poplar as a feedstock for biofuels: A review of compositional characteristics. *Biofuels Bioproducts & Biorefining-Biofr*. 2010; 4(2):209–26.

29. Stanton B, Neale D, Li S. *Populus* breeding: from the classical to the genomic approach. In: Jansson S, Groover AT, editor. *Genetics and Genomics of Populus*: Springer; 2010. p. 309–48.
30. Porth I, El-Kassaby YA. Using *Populus* as a lignocellulosic feedstock for bioethanol. *Biotechnology Journal*. 2015; 10(4):510–24. doi: [10.1002/biot.201400194](https://doi.org/10.1002/biot.201400194) PMID: [25676392](https://pubmed.ncbi.nlm.nih.gov/25676392/)
31. Slavov GT, Zhelev P. Salient Biological Features, Systematics, and Genetic Variation of *Populus*. *Genetics and Genomics of Populus*. 2010; 8:15–38.
32. Lexer C, Stoelting KN. Whole genome sequencing (WGS) meets biogeography and shows that genomic selection in forest trees is feasible. *New Phytologist*. 2012; 196(3):652–4. doi: [10.1111/j.1469-8137.2012.04362.x](https://doi.org/10.1111/j.1469-8137.2012.04362.x) PMID: [23043586](https://pubmed.ncbi.nlm.nih.gov/23043586/)
33. Slavov GT, Leonardi S, Adams WT, Strauss SH, DiFazio SP. Population substructure in continuous and fragmented stands of *Populus trichocarpa*. *Heredity*. 2010; 105(4):348–57. doi: [10.1038/hdy.2010.73](https://doi.org/10.1038/hdy.2010.73) PMID: [20531447](https://pubmed.ncbi.nlm.nih.gov/20531447/)
34. Slavov GT, DiFazio SP, Martin J, Schackwitz W, Muchero W, Rodgers-Melnick E, et al. Genome resequencing reveals multiscale geographic structure and extensive linkage disequilibrium in the forest tree *Populus trichocarpa*. *New Phytologist*. 2012; 196(3):713–25. doi: [10.1111/j.1469-8137.2012.04258.x](https://doi.org/10.1111/j.1469-8137.2012.04258.x) PMID: [22861491](https://pubmed.ncbi.nlm.nih.gov/22861491/)
35. Geraldes A, DiFazio SP, Slavov GT, Ranjan P, Muchero W, Hannemann J, et al. A 34K SNP genotyping array for *Populus trichocarpa*: Design, application to the study of natural populations and transferability to other *Populus* species. *Molecular Ecology Resources*. 2013; 13(2):306–23. doi: [10.1111/1755-0998.12056](https://doi.org/10.1111/1755-0998.12056) PMID: [23311503](https://pubmed.ncbi.nlm.nih.gov/23311503/)
36. Geraldes A, Farzaneh N, Grassa CJ, McKown AD, Guy RD, Mansfield SD, et al. Landscape genomics of *Populus trichocarpa*: the role of hybridization, limited gene flow, and natural selection in shaping patterns of population structure. *Evolution*. 2014; 68(11):3260–80. doi: [10.1111/evo.12497](https://doi.org/10.1111/evo.12497) PMID: [25065449](https://pubmed.ncbi.nlm.nih.gov/25065449/)
37. Porth I, Klápště J, Skyba O, Lai BS, Geraldes A, Muchero W, et al. *Populus trichocarpa* cell wall chemistry and ultrastructure trait variation, genetic control and genetic correlations. *New Phytologist*. 2013; 197(3):777–90. doi: [10.1111/nph.12014](https://doi.org/10.1111/nph.12014) PMID: [23278123](https://pubmed.ncbi.nlm.nih.gov/23278123/)
38. La Mantia J, Klapšte J, El-Kassaby YA, Azam S, Guy RD, Douglas CJ, et al. Association Analysis Identifies *Melampsora xcolombiana* Poplar Leaf Rust Resistance SNPs. *PloS One*. 2013; 8(11):e78423. doi: [10.1371/journal.pone.0078423](https://doi.org/10.1371/journal.pone.0078423) PMID: [24236018](https://pubmed.ncbi.nlm.nih.gov/24236018/)
39. McKown A, Klápště J, Guy R, Geraldes A, Porth I, Hannemann J, et al. Genome-wide association implicates numerous genes underlying ecological trait variation in natural populations of *Populus trichocarpa*. *New Phytologist*. 2014; 203(2):535–53. doi: [10.1111/nph.12815](https://doi.org/10.1111/nph.12815) PMID: [24750093](https://pubmed.ncbi.nlm.nih.gov/24750093/)
40. Evans LM, Slavov GT, Rodgers-Melnick E, Martin J, Ranjan P, Muchero W, et al. Population genomics of *Populus trichocarpa* identifies signatures of selection and adaptive trait associations. *Nature genetics*. 2014; 46(10):1089–96. doi: [10.1038/ng.3075](https://doi.org/10.1038/ng.3075) PMID: [25151358](https://pubmed.ncbi.nlm.nih.gov/25151358/)
41. Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G. LOSITAN: A workbench to detect molecular adaptation based on a F(st)-outlier method. *BMC Bioinformatics*. 2008; 9:323. doi: [10.1186/1471-2105-9-323](https://doi.org/10.1186/1471-2105-9-323) PMID: [18662398](https://pubmed.ncbi.nlm.nih.gov/18662398/)
42. Hemani G, Knott S, Haley C. An Evolutionary Perspective on Epistasis and the Missing Heritability. *PLoS Genetics* 2013; 9(2):e1003295. doi: [10.1371/journal.pgen.1003295](https://doi.org/10.1371/journal.pgen.1003295) PMID: [23509438](https://pubmed.ncbi.nlm.nih.gov/23509438/)
43. Porth I, Klapšte J, Skyba O, Hannemann J, McKown AD, Guy RD, et al. Genome-wide association mapping for wood characteristics in *Populus* identifies an array of candidate single nucleotide polymorphisms. *New Phytologist*. 2013; 200(3):710–26. doi: [10.1111/nph.12422](https://doi.org/10.1111/nph.12422) PMID: [23889164](https://pubmed.ncbi.nlm.nih.gov/23889164/)
44. McKown AD, Guy RD, Quamme L, Klápště J, La Mantia J, Constabel CP, et al. Association genetics, geography and ecophysiology link stomatal patterning in *Populus trichocarpa* with carbon gain and disease resistance trade-offs. *Molecular Ecology*. 2014; 23(23):5771–90. doi: [10.1111/mec.12969](https://doi.org/10.1111/mec.12969) PMID: [25319679](https://pubmed.ncbi.nlm.nih.gov/25319679/)
45. McKown AD, Guy RD, Klápště J, Geraldes A, Friedmann M, Cronk QCB, et al. Geographical and environmental gradients shape phenotypic trait variation and genetic structure in *Populus trichocarpa*. *New Phytologist*. 2014; 201(4):1263–76. doi: [10.1111/nph.12601](https://doi.org/10.1111/nph.12601) PMID: [24491114](https://pubmed.ncbi.nlm.nih.gov/24491114/)
46. Frenitiu FD, Clegg SM, Chittock J, Burke T, Blows MW, Owens IPF. Pedigree-free animal models: the relatedness matrix reloaded. *Proceedings of the Royal Society B-Biological Sciences*. 2008; 275(1635):639–47.
47. Ritland K, Ritland C. Inferences about quantitative inheritance based on natural population structure in the yellow monkeyflower, *Mimulus guttatus*. *Evolution*. 1996; 50(3):1074–82.
48. Lippert C, Quon G, Kang EY, Kadie CM, Listgarten J, Heckerman D. The benefits of selecting phenotype-specific variants for applications of mixed models in genomics. *Scientific Reports*. 2013; 3:1815. doi: [10.1038/srep01815](https://doi.org/10.1038/srep01815) PMID: [23657357](https://pubmed.ncbi.nlm.nih.gov/23657357/)

49. Lynch M, Walsh B. Genetics and Analysis of Quantitative Traits. first ed. Sunderland, MA, USA: Sinauer Associates; 1998. 980 p.
50. Jannink J-L. Identifying quantitative trait locus by genetic background interactions in association studies. *Genetics*. 2007; 176(1):553–61. PMID: [17179077](#)
51. Carter AJR, Hermisson J, Hansen TF. The role of epistatic gene interactions in the response to selection and the evolution of evolvability. *Theoretical Population Biology*. 2005; 68(3):179–96. PMID: [16122771](#)
52. Kremer A, Le Corre V. Decoupling of differentiation between traits and their underlying genes in response to divergent selection. *Heredity*. 2012; 108(4):375–85. doi: [10.1038/hdy.2011.81](#) PMID: [21915150](#)
53. Anonymous. On beyond GWAS. *Nature Genetics*. 2010; 42(7):551. doi: [10.1038/ng0710-551](#) PMID: [20581872](#)
54. Epperson BK. *Geographical Genetics*. Princeton University Press, Princeton, New Jersey; 2003. 376 p.
55. Ruttink T, Arend M, Morreel K, Storme V, Rombauts S, Fromm J, et al. A molecular timetable for apical bud formation and dormancy induction in poplar. *Plant Cell*. 2007; 19(8):2370–90. PMID: [17693531](#)
56. Fabbrini F, Gaudet M, Bastien C, Zaina G, Harfouche A, Beritognolo I, et al. Phenotypic plasticity, QTL mapping and genomic characterization of bud set in black poplar. *BMC Plant Biology*. 2012; 12:47. doi: [10.1186/1471-2229-12-47](#) PMID: [22471289](#)
57. Petterle A, Karlberg A, Bhalerao RP. Daylength mediated control of seasonal growth patterns in perennial trees. *Current Opinion in Plant Biology*. 2013; 16(3):301–6. doi: [10.1016/j.pbi.2013.02.006](#) PMID: [23473967](#)
58. Rohde A, Bastien C, Boerjan W. Temperature signals contribute to the timing of photoperiodic growth cessation and bud set in poplar. *Tree Physiology*. 2011; 31(5):472–82. doi: [10.1093/treephys/tpr038](#) PMID: [21636689](#)
59. Kalcsits LA, Silim S, Tanino K. Warm temperature accelerates short photoperiod-induced growth cessation and dormancy induction in hybrid poplar (*Populus x spp.*). *Trees-Structure and Function*. 2009; 23(5):971–9.
60. Hanninen H, Tanino K. Tree seasonality in a warming climate. *Trends in Plant Science*. 2011; 16(8):412–6. doi: [10.1016/j.tplants.2011.05.001](#) PMID: [21640632](#)
61. Wang Y-Y, Tsay Y-F. Arabidopsis Nitrate Transporter NRT1.9 Is Important in Phloem Nitrate Transport. *Plant Cell*. 2011; 23(5):1945–57. doi: [10.1105/tpc.111.083618](#) PMID: [21571952](#)
62. Bai H, Euring D, Volmer K, Janz D, Polle A. The Nitrate Transporter (NRT) Gene Family in Poplar. *PLoS One*. 2013; 8(8):e72126. doi: [10.1371/journal.pone.0072126](#) PMID: [23977227](#)
63. Duarte JM, Cui LY, Wall PK, Zhang Q, Zhang XH, Leebens-Mack J, et al. Expression pattern shifts following duplication indicative of subfunctionalization and neofunctionalization in regulatory genes of Arabidopsis. *Molecular Biology and Evolution*. 2006; 23(2):469–78. PMID: [16280546](#)
64. Camanes G, Pastor V, Cerezo M, Garcia-Andrade J, Vicedo B, Garcia-Agustin P, et al. A Deletion in NRT2.1 Attenuates Pseudomonas syringae-Induced Hormonal Perturbation, Resulting in Primed Plant Defenses. *Plant Physiology*. 2012; 158(2):1054–66. doi: [10.1104/pp.111.184424](#) PMID: [22158760](#)
65. Himelblau E, Amasino RM. Nutrients mobilized from leaves of Arabidopsis thaliana during leaf senescence. *Journal of Plant Physiology*. 2001; 158(10):1317–23.
66. Black BL, Fuchigami LH, Coleman GD. Partitioning of nitrate assimilation among leaves, stems and roots of poplar. *Tree Physiology*. 2002; 22(10):717–24. PMID: [12091153](#)
67. Larisch C, Dittrich M, Wildhagen H, Lautner S, Fromm J, Polle A, et al. Poplar Wood Rays Are Involved in Seasonal Remodeling of Tree Physiology. *Plant Physiology*. 2012; 160(3):1515–29. doi: [10.1104/pp.112.202291](#) PMID: [22992511](#)
68. Chandrashekar M, Heather WA. Temperature sensitivity of reactions of populus spp to races of Melampsora-larici-populina. *Phytopathology*. 1981; 71(4):421–4.
69. Kremer A, Ronce O, Robledo-Arnuncio JJ, Guillaume F, Bohrer G, Nathan R, et al. Long-distance gene flow and adaptation of forest trees to rapid climate change. *Ecology Letters*. 2012; 15(4):378–92. doi: [10.1111/j.1461-0248.2012.01746.x](#) PMID: [22372546](#)
70. Le Corre V, Kremer A. The genetic differentiation at quantitative trait loci under local adaptation. *Molecular Ecology*. 2012; 21(7):1548–66. doi: [10.1111/j.1365-294X.2012.05479.x](#) PMID: [22332667](#)
71. Mimura M, Aitken SN. Adaptive gradients and isolation-by-distance with postglacial migration in Picea sitchensis. *Heredity*. 2007; 99(2):224–32. PMID: [17487214](#)

72. Soolanayakanahally RY, Guy RD, Silim SN, Song M. Timing of photoperiodic competency causes phenological mismatch in balsam poplar (*Populus balsamifera* L.). *Plant Cell and Environment*. 2013; 36(1):116–27.
73. Luquez V, Hall D, Albrechtsen BR, Karlsson J, Ingvarsson P, Jansson S. Natural phenological variation in aspen (*Populus tremula*): the SwAsp collection. *Tree Genetics & Genomes*. 2008; 4(2):279–92.
74. Fracheboud Y, Luquez V, Bjorken L, Sjodin A, Tuominen H, Jansson S. The Control of Autumn Senescence in European Aspen. *Plant Physiology*. 2009; 149(4):1982–91. doi: [10.1104/pp.108.133249](https://doi.org/10.1104/pp.108.133249) PMID: [19201914](https://pubmed.ncbi.nlm.nih.gov/19201914/)
75. Howe GT, Hackett WP, Furnier GR, Klevorn RE. Photoperiodic responses of a northern and southern ecotype of black cottonwood. *Physiologia Plantarum*. 1995; 93(4):695–708.
76. Whitlock MC. Evolutionary inference from Q(ST). *Molecular Ecology*. 2008; 17(8):1885–96. doi: [10.1111/j.1365-294X.2008.03712.x](https://doi.org/10.1111/j.1365-294X.2008.03712.x) PMID: [18363667](https://pubmed.ncbi.nlm.nih.gov/18363667/)
77. Lefèvre F, Boivin T, Bontemps A, Courbet F, Davi H, Durand-Gillmann M, et al. Considering evolutionary processes in adaptive forestry. *Annals of Forest Science*. 2013:1–17.
78. Aitken SN, Whitlock MC. Assisted Gene Flow to Facilitate Local Adaptation to Climate Change. *Annual Review of Ecology, Evolution, and Systematics*. 2013; 44:367
79. Xie C-Y, Ying CC, Yanchuk AD, Holowachuk DL. Ecotypic mode of regional differentiation caused by restricted gene migration: a case in black cottonwood (*Populus trichocarpa*) along the Pacific Northwest coast. *Canadian Journal of Forest Research*. 2009; 39(3):519–26.
80. McKown AD, Guy RD, Azam MS, Drewes EC, Quamme LK. Seasonality and phenology alter functional leaf traits. *Oecologia*. 2013; 172(3):653–65. doi: [10.1007/s00442-012-2531-5](https://doi.org/10.1007/s00442-012-2531-5) PMID: [23212534](https://pubmed.ncbi.nlm.nih.gov/23212534/)
81. Porth I, El-Kassaby Y. Assessment of the Genetic Diversity in Forest Tree Populations Using Molecular Markers. *Diversity*. 2014; 6(2):283.
82. Jombart T. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*. 2008; 24(11):1403–5. doi: [10.1093/bioinformatics/btn129](https://doi.org/10.1093/bioinformatics/btn129) PMID: [18397895](https://pubmed.ncbi.nlm.nih.gov/18397895/)
83. Schnute JT, Boers NM, Haigh R. PBS mapping 2: User's guide—Introduction. *Canadian Technical Report of Fisheries and Aquatic Sciences*. 2004; 2549:1–V.
84. Moran PAP. Notes on continuous stochastic phenomena. *Biometrika*. 1950; 37(1–2):17–23. PMID: [15420245](https://pubmed.ncbi.nlm.nih.gov/15420245/)
85. Bivand R. Spdep: spatial dependence: weighting schemes, statistics and models. R package version 0.5–77, Available online at <http://cran.r-project.org/src/contrib/Descriptions/spdep.html>. 2014.
86. Yang W-Y, Novembre J, Eskin E, Halperin E. A model-based approach for analysis of spatial structure in genetic data. *Nature Genetics*. 2012; 44(6):725–31. doi: [10.1038/ng.2285](https://doi.org/10.1038/ng.2285) PMID: [22610118](https://pubmed.ncbi.nlm.nih.gov/22610118/)
87. Wang T, Hamann A, Spittlehouse DL, Murdock TQ. ClimateWNA-High-Resolution Spatial Climate Data for Western North America. *Journal of Applied Meteorology and Climatology*. 2012; 51(1):16–29.
88. Di Giuseppe E, Jona Lasinio G, Esposito S, Pasqui M. Functional clustering for Italian climate zones identification. *Theoretical and Applied Climatology*. 2013; 114(1–2):39–54.
89. Saether SA, Fiske P, Kalas JA, Kuresoo A, Luigujoe L, Piernety SB, et al. Inferring local adaptation from Q(ST)-F-ST comparisons: neutral genetic and quantitative trait variation in European populations of great snipe. *Journal of Evolutionary Biology*. 2007; 20(4):1563–76. PMID: [17584249](https://pubmed.ncbi.nlm.nih.gov/17584249/)
90. Henderson CR. *Applications of Linear Models in Animal Breeding*. Guelph, ON: University of Guelph; 1984. 423 p.
91. VanRaden PM. Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science*. 2008; 91(11):4414–23. doi: [10.3168/jds.2007-0980](https://doi.org/10.3168/jds.2007-0980) PMID: [18946147](https://pubmed.ncbi.nlm.nih.gov/18946147/)
92. Wimmer V, Albrecht T, Auinger HJ, Schön CC. synbreed: a framework for the analysis of genomic prediction data using R. *Bioinformatics*. 2012; 28(15):2086–7. doi: [10.1093/bioinformatics/bts335](https://doi.org/10.1093/bioinformatics/bts335) PMID: [22689388](https://pubmed.ncbi.nlm.nih.gov/22689388/)
93. McKay JK, Latta RG. Adaptive population divergence: markers, QTL and traits. *Trends in Ecology & Evolution*. 2002; 17(6):285–91.
94. Gilmour AR, Gogel BJ, Cullis BR, Welham SJ, Thompson R. *ASReml User Guide Release 1.0*. Hemel Hempstead: VSN International Ltd; 2002.
95. Joost S, Bonin A, Bruford MW, Despres L, Conord C, Erhardt G, et al. A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology*. 2007; 16(18):3955–69. PMID: [17850556](https://pubmed.ncbi.nlm.nih.gov/17850556/)
96. Lind MI, Ingvarsson PK, Johansson H, Hall D, Johansson F. Gene flow and selection on phenotypic plasticity in an island system of rana temporaria. *Evolution*. 2011; 65(3):684–97. doi: [10.1111/j.1558-5646.2010.01122.x](https://doi.org/10.1111/j.1558-5646.2010.01122.x) PMID: [20825480](https://pubmed.ncbi.nlm.nih.gov/20825480/)

97. Lewontin RC, Krakauer J. Distribution of gene frequency as a test of theory of selective neutrality of polymorphisms. *Genetics*. 1973; 74(1):175–95. PMID: [4711903](#)
98. Whitlock MC, Guillaume F. Testing for Spatially Divergent Selection: Comparing Q(ST) to F-ST. *Genetics*. 2009; 183(3):1055–63. doi: [10.1534/genetics.108.099812](#) PMID: [19687138](#)
99. Geraldes A, Pang J, Thiessen N, Cezard T, Moore R, Zhao Y, et al. SNP discovery in black cottonwood (*Populus trichocarpa*) by population transcriptome resequencing. *Molecular Ecology Resources*. 2011; 11(Suppl 1):81–92. doi: [10.1111/j.1755-0998.2010.02960.x](#) PMID: [21429165](#)