

Genomic and phenotypic architecture of a spruce hybrid zone (*Picea sitchensis* × *P. glauca*)

JILL A. HAMILTON,* CHRISTIAN LEXER[†] and SALLY N. AITKEN*

*Centre for Forest Conservation Genetics and Department of Forest Sciences, University of British Columbia, 3041-2424 Main Mall Vancouver, BC, Canada V6T 1Z4, [†]Unit of Ecology and Evolution, Department of Biology, University of Fribourg, Chemin du Musée 10, CH-1700 Fribourg, Switzerland

Abstract

Interspecific hybridization may enhance the capacity of populations to adapt to changing environments, and has practical implications for reforestation. We use genome-wide estimates of admixture and phenotypic traits for trees in a common garden to examine the extent and direction of gene flow across a *Picea* hybrid zone, testing assumptions of the bounded hybrid superiority and tension zone models of hybrid zone maintenance. Seeds were collected from the ecological transition zone spanning from maritime to continental climates across the *Picea sitchensis*–*P. glauca* contact zone, and 721 trees were planted in a common garden experiment within the hybrid zone. Individuals were genotyped using a panel of 384 candidate-gene single nucleotide polymorphisms (SNPs) putatively associated with adaptive traits in *Picea*, and phenotyped at age ten for height and autumn cold hardiness. Low interspecific heterozygosity in hybrids indicated that intrinsic reproductive barriers were too weak to prevent widespread recombination, although introgression appeared asymmetric with *P. sitchensis* dominating the zone. Whereas marker-based hybrid index was strongly correlated with climate and geography, phenotypic traits exhibited weak or no significant clines. Our results indicated that exogenous selection appeared to play a strong role in the distribution and structure of this hybrid zone, indicative of an environmentally determined bounded hybrid superiority model of hybrid zone maintenance, although endogenous mechanisms could not be ruled out. This study provides insight into the mechanisms underlying adaptation across ecologically transitional hybrid zones that will ultimately provide an additional tool in managing these economically important tree species.

Keywords: climate change, cold hardiness, hybrid zone, phenotypes, single nucleotide polymorphisms, spruce

Received 6 March 2012; revision received 26 July 2012; accepted 26 July 2012

Introduction

Natural hybrid zones are important biological systems for examining the roles of natural selection and gene flow in the maintenance of reproductive isolation and species differences. Interspecific hybridization is an important source of novel genetic variation in which new genetic recombinants may gain a further capacity for adaptive evolution (Rieseberg & Wendel 1993; Fritz *et al.* 2006). This is particularly important in a changing

climate, where there is growing evidence that some tree species may exhibit pronounced adaptational lag (Aitken *et al.* 2008). Introgression, where allelic variants are exchanged between species permanently and repeatedly, may increase a species' evolutionary potential and ability to adapt in the face of climate change and increasing environmental stochasticity. Although introgression between congeneric forest tree species is widely documented (Lexer *et al.* 2004), hybrid zones have traditionally remained a 'black box' in tree breeding programs, where the goal is to produce seed for reforestation optimally adapted to relatively local environments. Increased understanding of the mechanisms underlying local adaptation in hybrid zones may

Correspondence: Jill A. Hamilton, Fax: 604-822-9102; E-mail: jillahamilton@gmail.com

produce additional tools to mitigate the influence of climate change in managed forests.

Hybrid zone maintenance has a rich theoretical background that can be distilled into two types of models differing fundamentally in the mode of selection implicated: endogenous (non-environmental), or exogenous (environmental). The tension zone model developed in detail by Barton & Hewitt (1985) describes a hybrid zone shaped by endogenous genetic factors independent of the environment. In contrast, the bounded hybrid superiority model (Moore 1977; Arnold 1997) applies to hybrid zones in which exogenous selection shapes variation within the contact zone, and the fitness of hybrid and parental genotypes differ amongst environments (Wilhelm & Hilbish 1998). Furthermore, this model assumes that hybrids exhibit greater fitness relative to parental genotypes due to genotype-by-environment interaction within the ecologically transitional habitat (Moore 1977). Although these models have clearly differentiated assumptions, clinal variation resulting from geographic variation in environmental selection gradients (Endler 1977; Moore 1977) and intrinsic selection against hybrids are often indistinguishable (Barton & Hewitt 1985; Kruuk *et al.* 1999), and in some hybrid zones both extrinsic and intrinsic selection may be acting (e.g., *Artemisia tridentata*, Wang *et al.* 1997; Miglia *et al.* 2005). Teasing apart the mechanisms that contribute to hybrid zone maintenance may help us to understand the role of environment and genetic factors in hybrid zone persistence.

Evaluation of interspecific gene flow and the distribution of hybrid genotypes offers an understanding of the mechanisms, both extrinsic and intrinsic, that maintain distinct species (Field *et al.* 2011). Genome-wide scans of mapped loci provide increased resolution of potential locus-specific effects for maintaining species differences, whereas genomic clines for hybrid indices along ecological gradients may quantify relationships maintaining the hybrid zone (Nolte *et al.* 2009; Teeter *et al.* 2010; Gompert & Buerkle 2011). Finally, combining genetic and phenotypic data within a common garden allows for inference of both the extent and direction of gene flow, and ultimately aids in teasing apart genotype-by-environment interactions. This approach will have broad ecological, conservation, and economic implications in forest tree hybrid zones, where identification of seed source populations for reforestation within transitional habitat remains problematic. Moreover, where admixture results in an increased capacity to respond to a changing climate (Aitken *et al.* 2008), associating fitness proxies in long-lived tree species such as growth and cold hardiness, with climatic variation may aid in identifying those genotypes suitable for re-planting in areas where climatic conditions have changed or are predicted to change.

The introgression zone between Sitka spruce (*Picea sitchensis* (Bong) Carr.) and white spruce (*P. glauca* Moench Voss) spans maritime to continental climates along the Nass and Skeena river valleys in northern British Columbia. In British Columbia, Sitka spruce is a significant component of the Coastal Western Hemlock (CWH) ecosystem. This largest of spruce species thrives at a mean annual temperature of ~ 8 °C and mean annual precipitation exceeding two metres (Pojar *et al.* 1991). White spruce, a component of the Interior Cedar Hemlock (ICH) ecosystem characterized by cold, wet winters and warm, dry summers along with periodic extremes, comes into contact with Sitka spruce along river valleys where moist, cool climates offer transitional habitat between these coastal and inland ecosystems (Ketcheson *et al.* 1991). Early studies identified the zone of contact using morphology and phenology (Roche, 1969), whereas subsequent studies have characterized physiology (Grossnickle *et al.* 1996; Fan *et al.* 1997; Silim *et al.* 2001), insect resistance (O'Neill *et al.* 2002) and neutral genetic variation (Yeh & Arnott 1986; Sutton *et al.* 1994; Bennuah *et al.* 2004). The large amount of variation observed amongst populations across the hybrid zone has been attributed to steep climatic gradients.

Although there is strong evidence for the role of exogenous selection maintaining this hybrid zone, the fitness of parent species and hybrid individuals has not been directly assessed. Bennuah *et al.* (2004) described geographic clines in hybrid index based on a small number of expressed sequence tag markers along environmental gradients in the zone, suggesting environmental selection strongly influences genetic composition. However, clinal variation may also arise from endogenous selection, which was not examined within the Bennuah *et al.* (2004) study. By capitalizing on genomic resources developed for both Sitka and white spruce, and a 10-year-old common garden experiment previously established within the hybrid zone, we aim to better understand the mechanisms underlying the maintenance of this hybrid zone.

In the current study, we have used a genomic survey of single nucleotide polymorphisms (SNPs) to examine the composition of individuals collected from across the contact zone as well as reference pure parental species, all planted in a common garden experiment located within the hybrid zone. Using genomic composition estimated by hybrid index, and interspecific recombination evaluated using heterozygosity, we first asked whether the direction and extent of introgression across the contact zone fits predictions of a bounded hybrid superiority or a tension zone model of hybrid zone maintenance. We then compared associations between hybrid index and climate variables, indicative of exogenous selection, and distance variables, indicating a pattern of isolation-by-distance, to

infer potential mechanisms influencing genetic structure across this zone. Finally, we tested for differences in phenotypic traits among genotypic classes to gain further support for the presence or absence of hybrid superiority. This research will provide insight into the role of intrinsic and extrinsic factors in the maintenance of this hybrid zone, and inform seed transfer guidelines for current and future climates for forest trees in ecologically transitional areas.

Methods

Plant material and common garden experiment

Open-pollinated seeds were collected in 1997 from 29 locations throughout the introgression zone by the

British Columbia Forest Service. Populations from the Nass and Skeena River drainages were analysed together, where previous research by Benuah *et al.* (2004) indicated drainage distance up river valleys predicted genetic ancestry. Two additional allopatric populations, Sitka spruce on Haida Gwaii, British Columbia (HG) and white spruce from the Ottawa valley region, Ontario (ENA, Table 1 and Fig. 1) were also included in the study as reference parental populations. Cones were collected from the upper canopy of three to seven seed parents per location by helicopter. Open-pollinated progeny from each seed parent were grown for 1 year in a nursery, and ten seedlings per family were subsequently transplanted into a common garden near Kitwanga, British Columbia (55°17'N, 128°10'W, Fig. 1). A split-plot block design was used for the common

Table 1 Origins of Sitka × white spruce (*Picea sitchensis* × *P. glauca*) study populations, sample size (N) and associated climatic and geographic distance variables

Population	N	Latitude	Longitude	Elevation (m)	Drainage distance (km)	MAT (°C)	MAP (mm)	Continentality (°C)
Haida Gwaii (HG)*	24	53.65	-132.2	10	0	7.5	1482	13.6
Mt. Kaien	18	54.3	-130.33	600	0	6.2	3065	10.8
Seal Cove	18	54.32	-130.32	30	0	8	2901	10.2
Bella Bella	16	52.17	-128.16	60	0	8.6	2821	11.6
Ocean Falls	17	52.36	-127.69	60	24.825	7.8	4437	13.4
Ramsden Point	30	55.02	-130.13	20	42.878	7	3086	14.3
Kinkolith—Low	30	55	-129.93	30	51.011	6.6	2979	14.8
Exstew—Low	18	54.42	-129.07	30	85.175	5.6	1691	19.6
Eskers O/P—Low	24	54.41	-129.03	100	85.175	5.6	1639	19.3
Eskers O/P—Mid	18	54.41	-129.02	450	85.175	3.9	1581	19.8
Greenville—Low	24	55.09	-129.5	220	86.665	5.2	1700	18.8
Kitimat—Low	30	54.07	-128.54	100	108.438	6.3	2722	18.5
Hagensburg	18	52.39	-126.58	60	115.241	7	1787	16.6
Kitimat—High	30	54.2	-128.55	250	119.074	5.4	2044	19.1
Nass Camp—Low	30	55.3	-129.05	185	126.398	5.5	989	19.9
Nass	29	55.28	-129.02	100	127.246	5.5	1020	20.3
Douglas Creek—High	30	54.83	-128.73	800	150.426	2.5	1396	19.8
Douglas Creek—Mid	30	54.83	-128.69	450	152.721	4.6	1686	19.3
Kiteen—Low	29	55.44	-128.83	310	155.786	4.4	1083	20.3
Doreen—Low	24	54.83	-128.31	220	165.852	5	965	20
Talchako River—Low	18	52.15	-125.96	305	180.381	4.9	907	19.7
Cedarvale—Low	18	55.02	-128.35	110	187.612	5.4	832	20.3
Talchako River—High	18	52.06	-125.92	488	190.646	3.6	985	19.8
Cedarvale—Mid	18	55.03	-128.42	230	192.253	4.5	894	20.1
Donaker Creek	18	57.45	-131.08	300	217.39	1.8	692	26.5
Kitwanga—Low	24	55.37	-128.14	450	219.64	3.4	640	21.8
Kitwanga Lake—Mid	24	55.36	-128.13	680	221.451	2.8	673	21.8
Skeena crossing	18	55.1	-127.81	210	221.747	4.9	590	22.8
Meziadin	18	56.1	-129.3	265	233.807	4.7	908	21.1
Moricetown	18	55.03	-127.34	440	277.244	3.7	540	23.1
Eastern North America (ENA)*	42	45.38	-75.71	79.2	NA	6.3	914.2	NA

MAT, mean annual temperature; MAP, mean annual precipitation.

*Parental populations used to calculate hybrid index.

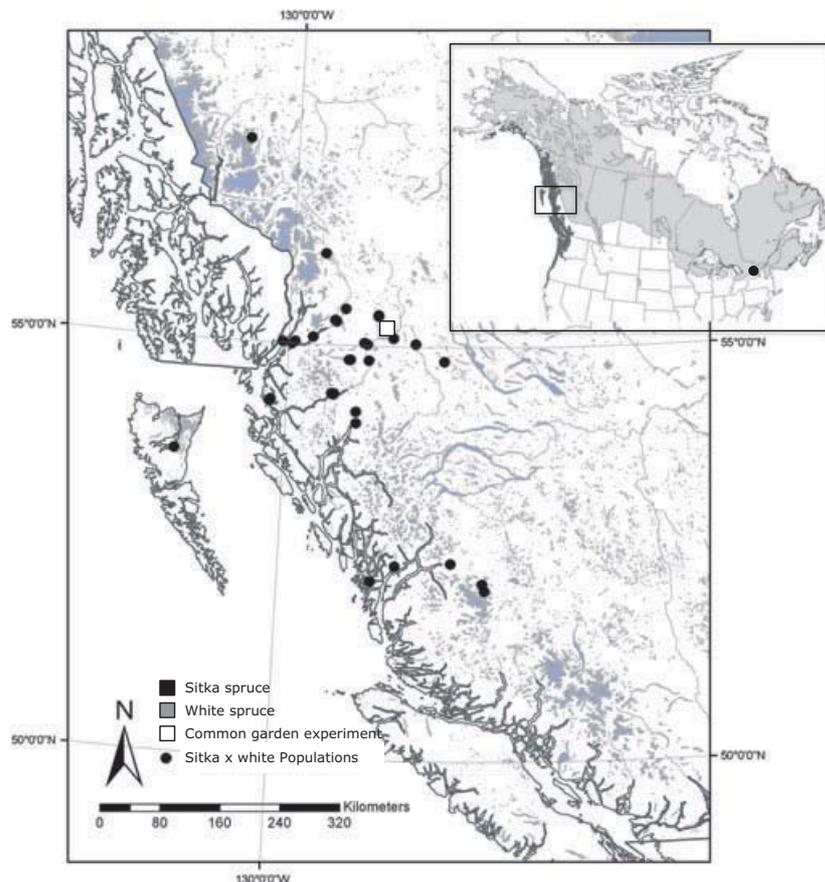


Fig. 1 Map of individuals sampled across the introgression zone between Sitka and white spruce in northwestern British Columbia, indicating provenance of origin of samples (circle) collected and planting location of the common garden experiment (square). Inset map provides North American range of *P. sitchensis* (dark grey) and *P. glauca* (light grey), as well as source of eastern North American population.

garden with 10 replications, where the main plots were regional sets and the subplots were sampling locations within regions to avoid shading and other competitive effects due to regional height differences negatively impacting growth or survival. Sampling location subplots were non-contiguous with individual populations randomized within main regional plots. On average, six progeny were sampled per open-pollinated family, for a total of 721 trees from sampling locations spanning the contact zone. Trees were measured for 10-year height, artificially tested to determine cold hardiness, and sampled for DNA extraction and genotyping.

Artificial freeze tests

Branches were harvested in late August 2010 from the common garden trial in Kitwanga, BC for artificial freeze testing using methods described in detail by Hannerz *et al.* (1999). Several Five millimetre needle segments per individual branch were cut and placed in a tube with 0.2 mL of distilled water and a trace of sil-

ver iodide to facilitate ice nucleation. Samples were frozen at three temperatures, -8 , -18 and -28 °C, in order to obtain intermediate injury levels for at least one temperature. Chamber temperature was lowered at a rate of 4 °C/h and then held for 1 h at each test temperature before removal. Control samples were kept at 4 °C for the duration. After freezing, 3.5 mL of distilled water was added to each sample and they were refrigerated for 24 h. Samples were then shaken for 60 min prior to initial measurement of electrolytic conductivity, heat-killed at 95 °C for 1.5 h, refrigerated for 24 h, re-shaken, and then re-measured for conductivity. Flint's Index of Injury I_t (Flint *et al.* 1967) was calculated as a percentage of injury as follows:

$$I_t = \frac{100(R_t - R_o)}{1 - R_o}$$

where $R_t = L_t/L_k$, $R_o = L_o/L_d$, estimated from a ratio of conductivity between frozen samples preceding (L_t) and following heat kill (L_k) with unfrozen control samples preceding (L_o) and following heat kill (L_d). Damage

levels at $-28\text{ }^{\circ}\text{C}$ were too high to be informative, therefore these data were not analysed further.

SNP selection and genotyping

Newly flushed lateral shoots were collected in June 2009, flash frozen in the field, and transported in liquid nitrogen for genetic analysis. Fifty milligrams of tissue was ground to a fine powder using a Mixer Mill MM 400 (Retsch) and DNA was extracted using a modified CTAB protocol (Doyle & Doyle 1990). Pure Sitka ($N = 1088$), white ($N = 40$), and Engelmann spruce (*Picea engelmannii*, $N = 40$) were genotyped for 1536 single nucleotide polymorphisms (SNPs) in genes identified as putative candidates for roles in budset timing or cold-hardiness development (Holliday *et al.* 2008, 2010a), white pine shoot tip weevil (*Pissodes strobi*) resistance Porth *et al.* 2011 unpublished and growth. A subset of 384 SNPs (Pavy *et al.* 2008, Beaulieu *et al.* 2011, Pelgas *et al.* 2011) was selected from this study based on the degree of interspecific genetic differentiation ($F_{ST} \geq 0.20$) amongst the three spruce species. These spruce have a relatively recent common ancestor, therefore the likelihood of finding diagnostic alleles to develop hybrid indices given phylogeny (Ran *et al.* 2006) and history (Anderson *et al.* 2006; Mimura & Aitken 2007a) was unlikely. Consequently, genetic differentiation (F_{ST}) was used as a tool for selecting SNPs for development of hybrid indices where high allele frequency differentials between species contain more information about ancestry (Buerkle 2005). The selected SNPs are likely subject to some ascertainment bias in terms of under-representation of rare alleles because of the way in which SNPs were selected (Namroud *et al.* 2008).

Selected SNPs were genotyped using the Illumina bead array platform in conjunction with the GoldenGate allele-specific assay in a 96-well, 384-SNP format (Shen *et al.* 2005; Fan *et al.* 2006). This highly multiplexed assay hybridizes two oligos specific to the alleles at the SNP site along with a third oligo specific to genomic DNA downstream from the SNP. The primers used for PCR product generation contain fluorescent moieties, which differentiate homozygotes and heterozygotes based on fluorescent signals. Genotyping quality was examined using the GenomeStudio Genotyping Module (v1.0) for successfully genotyped SNPs ($N = 338$). Quality was assessed using a number of criteria, including call rates $>90\%$, a minimum 'GenTrain' score of 0.40, visual assessment of differentiation of homozygote and heterozygote clusters, and minimum intensity threshold for fluorescent signals from replicate beads (>1000). SNPs were ranked based on five categories (i) failed; (ii) not polymorphic; (iii) polymorphic, but low genotyping quality; (iv) clear genotyping, but extreme

heterozygote excess or heterozygote deficiency resulting in large deviations in Hardy-Weinberg expectations that may result from genotyping paralogues rather than homologues, or other genotyping inaccuracies and (v) high-quality genotyping. Of the initial 384 SNPs, 268 met the high quality standard of (v) for subsequent analysis, eight of which have exhibited significant associations with cold hardiness phenotypes within Sitka spruce (Holliday *et al.* 2010a, Table S1, Supporting Information).

Genomic admixture analysis and hybrid identification

To estimate admixture amongst individuals, we carried out analyses using two different approaches: A maximum likelihood estimate of hybrid index calculated using the program *hindex* (Buerkle 2005); and a Bayesian clustering approach implemented in *Structure* version 2 (Pritchard *et al.* 2000). *Hindex* requires a priori knowledge of pure parental individuals, here identified as allopatric populations—HG (Sitka spruce) and ENA (white spruce) to estimate the proportion of ancestry for either parent, where a hybrid index of zero corresponds to pure white spruce and one corresponds to pure Sitka spruce. *Structure*, however, estimates a posterior probability (q) of assignment to k genotypic cluster categories rather than a priori assignments of pure species. Multilocus genotypes describing the proportion of individual genotype assignment to k clusters, with the assumption of $k = 2$ representing two species, were used to estimate the proportion of individual parent species' contribution to the admixed gene pool. *Structure* runs were carried out using the admixture model assuming correlated allele frequencies with a burn-in step of 50 000, followed by 100 000 Markov chain Monte Carlo iterations. Ten replicate runs of models using genotypic clusters (k) from 1 to 10 confirmed $k = 2$ had the highest log-likelihood probability (Figs S1 and S2, Supporting information).

Using *hindex* and *Structure* provides additional confidence in assignment of genotypic classes. Pearson correlations between hybrid index and q -value were estimated to test whether a priori assignment of pure parental species prejudiced the calculation of ancestry. ENA individuals used to represent pure white spruce are from populations thousands of kilometres from the hybrid zone that have no possibility of introgression with *P. engelmannii*, which hybridizes extensively with white spruce in mountainous areas of southeastern British Columbia and southwestern Alberta. To confirm that ENA samples were representative of white spruce found in British Columbia, a subset of 120 loci genotyped both in ENA samples on the 384 SNP array and in white spruce individuals from Fort Nelson, British

Columbia on the 1536 SNP array (58° 49' 59", 123° 27' 0"), were further compared using *Structure*. Those results indicated a high posterior probability (q , where 0 is white spruce) that individual ENA and Fort Nelson, BC genotypes originated from the same genotypic cluster (average ENA = 0.01 ± 0.005 , average Fort Nelson = 0.006 ± 0.001 , Fig. S3, Supporting information).

Interspecific heterozygosity was estimated in the R-package *Introgress* (Gompert & Buerkle 2010) for each individual and compared with hybrid index to provide a means to categorize individuals into genotypic classes. Expectations followed Milne & Abbott (2008), where F1 individuals have both 100% interspecific heterozygosity and a hybrid index of 0.5 in which each species contributes equally to the genome-wide estimate of admixture (see Fig. 1 in Milne & Abbott 2008). Furthermore, due to the complex and overlapping nature of hybrid classes, e.g., an F2 may resemble an F3, classes are broadly defined as backcrosses to either parent (BCS, BCW) if they exhibited <100% proportion, but >75% ancestry from one parent, or as advanced-generation hybrids (FN) if ancestry proportions remained relatively equal between species, but interspecific heterozygosity decreased below 85%. In effect, classification was achieved based on genomic discontinuities visible in two-dimensional hybrid index-heterozygosity space for all SNPs (Fig. 2A). This analysis is similar to that implemented in *NewHybrids* (Anderson & Thompson 2002), distinguishing pure parentals from advanced-generation introgressants, and requires fewer assumptions than the Bayesian approach of *NewHybrids*. Although there is always still some uncertainty in class assignment with advanced generations, the use of broad categories for

each class allows for assignment of those more complex derivatives with reasonable confidence (Milne & Abbott 2008).

A lack of diagnostic (species-specific) markers results in higher estimates of interspecific heterozygosity for pure parent species than expected. We compared interspecific heterozygosity further, evaluating a subset of the 268 loci closest to being diagnostic for species, with an allele frequency differential >0.90 ($n = 31$, Fig. 2B). Comparing interspecific heterozygosity for these strongly differentiated loci with estimates for all SNPs provides insight into whether the apparent lack of early-generation hybrids and greater interspecific heterozygosity of parent species is due to ancestral polymorphisms.

Quantitative genetic data analysis

To better understand phenotypic variation in height and cold injury across the hybrid zone, an analysis of variance was conducted using the linear model:

$$y_{ijkl} = u + r_i + s_j + r_i * s_j + p_k(s_j) + r_i * p_i(s_j) + f(p(s_j))_{kl}$$

where r is the effect of rep i , s is the effect of set j , p_k is the effect of population k and $f(p)_{kl}$ is the effect of family l within population k . This full model was reduced to:

$$y_{ijk} = u + r_i + p_j + r_i * p_j + f(p)_{jk}$$

where r is the effect of rep i , p_j is the effect of population j and $f(p)_{jk}$ is the effect of family k within population j , as all terms involving regional set were not significant and had no effect on the model. Additive genetic variance and narrow sense heritability was

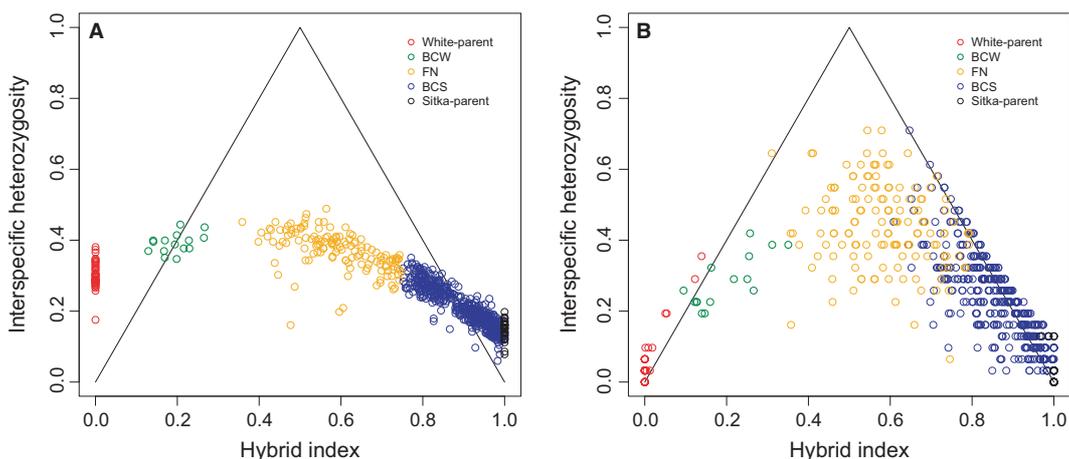


Fig. 2 Interspecific heterozygosity (vertical axis) vs. hybrid index (horizontal axis) using 268 candidate gene SNPs (A) and a subset of 31 SNPs that exhibited an allele frequency differential (δ) >0.90 (B). Plot based on all loci has been used to assign individuals to genotypic classes, where hybrid index 0 = white spruce and 1 = Sitka spruce. Individuals are classified as pure parental species (Sitka or white), backcrosses towards either Sitka spruce (BCS) or white spruce (BCW), or advanced generation hybrids (FN), based on discontinuities in two-dimensional space. See text for details.

estimated for each trait, assuming negligible influence of maternal effects, inbreeding, epistasis or correlated paternity (Streiff *et al.* 1999; Mimura & Aitken 2007b). Quantitative genetic differentiation amongst populations (Q_{ST}) was estimated using the relationship described in Spitze (1993) using PROC VARCOMP in SAS with the restricted maximum likelihood option (METHOD=REML).

Statistical analysis

Spatial structure of relationships between introgression and the environment were examined using regressions of both climate and distance variables for provenances on genotypes (hybrid index) and phenotypes. The latitude, longitude and elevation of sampling locations were input into *ClimateBC v.3.2* (Wang *et al.* 2006) which estimates seasonal and annual climate variables based on geographic coordinates. Climate variables were selected from *ClimateBC* based on univariate linear regressions for hybrid index and phenotypic traits, identifying those variables that likely influence local adaptation producing the strongest clines and explaining the greatest proportion of variance. Non-linear physical drainage distance up river systems from the Pacific Ocean was estimated using *ArcGIS* (Version 10), replicating methods from Bennuah *et al.* (2004). Distance from the ocean in this region is strongly associated with climatic gradients from maritime to continental climates. Regressions were estimated from individual tree data using *R 2.5.1* (R Development Core Team 2007), excluding the ENA population due to its great distance from the hybrid zone. As non-linear clines are common across hybrid zones (Arnold 1997), both linear and non-linear regressions were tested. In addition, multiple linear stepwise regressions were performed to evaluate the relationship between geographic and climatic variables with genome-wide estimate of ancestry and putatively adaptive phenotypic traits (Table S2, Supporting Information).

Height and cold injury phenotypes were analysed as components of fitness, as conventional lifetime fitness measures are not possible in such long-lived organisms. Response variables were analysed using ANOVA to compare the means amongst genotypic classes, followed by a post-hoc analysis using Tukey's HSD (honest significant difference) test for multiple comparisons. These tests were accompanied by a non-parametric Kruskal–Wallis test, which does not assume a normal distribution.

Results

Analyses of admixture using both *hindex* and *Structure* indicate a high proportion of hybrid individuals and

asymmetry with more Sitka spruce ancestry than white spruce in the hybrid zone. Average hybrid index based on *hindex* was 0.78, and of the 721 individuals analysed, *Structure* identified 345 individuals with a q -value ≥ 0.9 , $35 \leq 0.1$, and 346 between 0.1 and 0.9, where a q -value equal to one indicates pure Sitka spruce. The high correlation between ancestry estimates from *Structure* and *hindex* ($r = 0.997$, $P < 0.001$), although *hindex* requires inclusion and a priori identification of pure parents and *Structure* does not, indicates that hybrid index estimates do not depend on pure parental samples being identified.

Low overall interspecific heterozygosity across all individuals (0.06–0.49) indicates that this hybrid zone is largely comprised of advanced generation recombinants, with weak or no barriers to reproduction between species. Individuals were assigned to genotypic classes by comparing hybrid index and interspecific heterozygosity jointly (Fig. 2A). The absence of an F1 generation (with 100% heterozygosity and intermediate hybrid index based on the criteria of Milne & Abbott (2008) is consistent with an advanced-generation hybrid zone and a lack of reproductive barriers between species. Although low interspecific heterozygosity suggests widespread recombination, the distribution of the majority of hybrid index estimates over 0.5 suggests that introgression is skewed towards Sitka spruce ancestry (Fig. 2A). This asymmetry is reflected in the distribution of genotypic classes based on *hindex*, where only 15 individuals were classified as backcrosses towards white spruce (BCW), compared to 139 individuals classified as advanced-generation hybrid class (FN), and 501 individuals classified as backcrosses to Sitka spruce (BCS). Evaluation of a subset of 31 near species-diagnostic loci with an allele frequency differential >0.90 produced lower values of interspecific heterozygosity for white spruce (zero to 0.35, mean = 0.05) and Sitka spruce (0 to 0.13, mean = 0.03, Fig. 2B) when compared to results for 268 loci. Patterns were consistent with results using the full set of markers, indicating that estimates of interspecific heterozygosity for pure species, although higher than expected, were consistent with widespread recombination and allele-sharing, rather than divergent ancestral polymorphisms (Stölting *et al.* 2012).

Narrow-sense heritabilities were estimated as 0.16 for cold injury at -8 °C, 0.17 at -18 °C and 0.17 for height, typical for these traits in conifers. Quantitative genetic differentiation (Q_{ST}) amongst all populations was moderate to low, at 0.35 for cold injury at -8 °C, 0.38 at -18 °C and 0.13 for height.

Regressions of both hybrid index and phenotypic traits on distance and climatic variables revealed strong geographic and climatic clines for hybrid index but

weak clines for phenotypic traits (Fig. 3, Table 2). Clines of hybrid index with independent variables other than mean annual temperature (MAT) were non-linear (Fig. 3). A quadratic regression for hybrid index on both drainage distance and continentality, and an inverse transformation for mean annual precipitation produced the best model fit. Drainage distance up rivers from the coast exhibited a strongly significant non-linear relationship with hybrid index with increasing distance from the coast to the interior (quadratic regression: $R^2 = 0.73$, $P < 0.001$, Fig. 3). This relationship was mirrored in the cline for MAP, which appeared to have a threshold pattern with a steep inflection observed at 1000 mm, above which hybrid index estimates remained close to one (Sitka spruce) and below which a steep decline in hybrid index was associated with decreased precipitation.

Significant clines were also observed for continentality (difference between mean warmest month and mean coldest month temperatures) and MAT. Whereas clinal variation in continentality explains a considerable proportion of the variance in hybrid index (quadratic regression: $R^2 = 0.41$, $P < 0.001$), MAT is more weakly associated with hybrid index ($R^2 = 0.28$, $P < 0.001$). Hybrid index exhibited significant clinal variation with respect to climate variables, however, drainage distance overall explained more variation than any one climate

variable. As drainage distance is highly correlated with several climatic variables (MAT $r = -0.80$, MAP $r = -0.80$, continentality $r = 0.89$, $P < 0.001$ all), both climate and drainage distance variables reflect the transition between maritime and continental climates characteristic of the introgression zone. Multiple linear stepwise regressions of hybrid index with drainage and climate variables (MAT, MAP, and continentality) indicated all variables were significantly related to hybrid index ($R^2 = 0.68$, $P < 0.001$), although drainage distance explained the most variance (F -value = 1344.06, $P < 0.001$).

Cold injury exhibited only shallow linear clines along climatic gradients and geographic distances, and somewhat surprisingly, there was no clinal variation detected for height (Table 2). Drainage distance was significantly associated with cold injury ($P < 0.001$), but explained just five to seven percent of the variation in cold injury amongst individuals. For all phenotypic traits, stepwise multiple linear regression with distance and climate variables did not improve the variance explained over univariate regressions.

Comparison of phenotypic characters across all genotypic classes using ANOVA with post-hoc tests for multiple comparisons, along with a non-parametric equivalent, Kruskal-Wallis, indicated means differed significantly amongst genotypic classes (Fig. 4, Table 3).

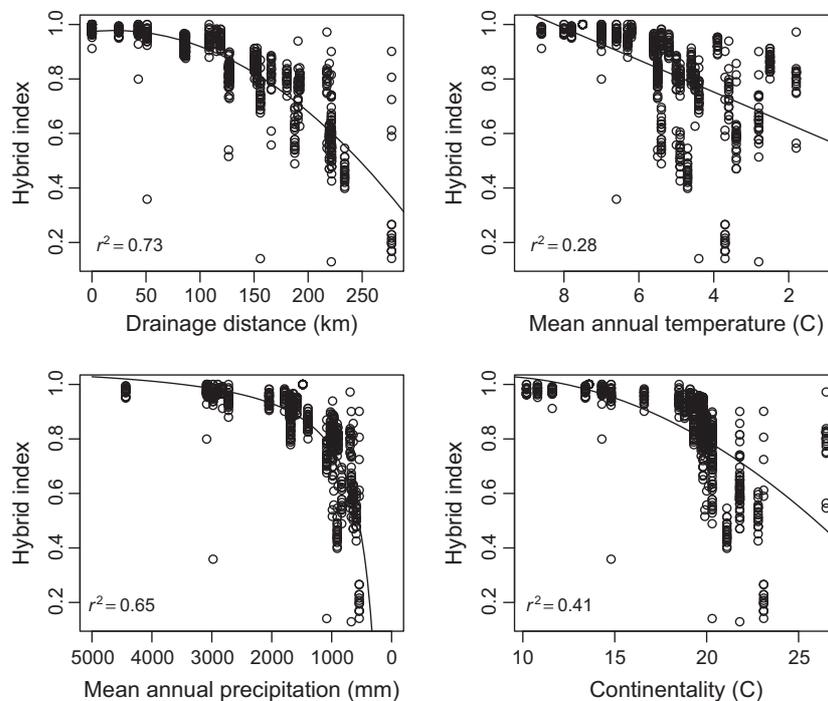


Fig. 3 Relationship between 268 SNP-based hybrid index (0 = white spruce, 1 = Sitka spruce) and geographic and climatic variables across 721 individuals spanning the Sitka-white introgression zone, including drainage distance (km), mean annual temperature ($^{\circ}\text{C}$), mean annual precipitation (mm), and continentality ($^{\circ}\text{C}$).

Table 2 Results of linear regression of cold injury at two artificial freezing temperatures and tree height measured for 721 individuals sampled across the Sitka × white spruce hybrid zone and grown in a common garden, with geographic and climatic variables across all 721 individuals

	Cold injury -8 °C			Cold injury -18 °C			Height (cm)		
	R ² (%)	Slope	F-value (Pr > F)	R ² (%)	Slope	F-value (Pr > F)	R ² (%)	Slope	F-value (Pr > F)
Geographic variable									
Drainage distance (km)	7.0	-0.09	52.57 (<i>P</i> < 0.001)	7.0	-0.06	53.25 (<i>P</i> < 0.001)	0.0	0.00	0.00 (0.98)
Climatic variables									
MAT (°C)	6.0	3.59	40.15 (<i>P</i> < 0.001)	4.0	1.99	26.77 (<i>P</i> < 0.001)	0.0	-0.38	0.04 (0.56)
MAP (mm)	6.0	0.00	40.87 (<i>P</i> < 0.001)	3.0	0.00	20.29 (<i>P</i> < 0.001)	0.0	0.00	0.56 (0.77)
Continentality (°C)	6.0	-1.64	41.87 (<i>P</i> < 0.001)	4.0	-0.89	26.64 (<i>P</i> < 0.001)	0.0	0.90	0.99 (0.32)

MAP, mean annual precipitation; MAT, mean annual temperature.

However, pairwise comparisons between genotypic classes provided a clearer view of differentiation (Table 3). For cold injury at -8 °C, all individuals with membership in a hybrid class appeared to have greater cold hardiness than either pure parental species (Fig. 4). Multiple comparisons indicated that hybrid classes were not significantly different from each other, but were significantly different from both parental species classes, and Sitka and white spruce classes were not significantly different from each other (Table 3). However, this pattern was not observed at the colder temperature. Overall, at -18 °C, pure Sitka exhibited the greatest cold injury, and as the contribution of Sitka spruce decreased and white spruce increased, there was increased cold hardiness (Fig. 3). The BCW class appeared slightly more cold tolerant, on average, than white spruce at -18 °C, although not significantly different from white spruce (Table 3). Although there were significant differences amongst genotypic classes in height (Table 3),

there were no significant pairwise comparisons. This finding suggests that the differences observed are largely due to the shorter BCW class (Fig. 3).

Discussion

The genomic architecture of the Sitka-white spruce zone of introgression revealed some surprising paradoxes. Low overall interspecific heterozygosity based on all single nucleotide polymorphisms (SNPs), ranging 0.06–0.49, indicates an advanced-generation hybrid zone in which recombination is widespread. Variation in the proportion of ancestry across the zone appears to be strongly influenced by both geography and climate, indicating an important role for exogenous selection shaping the genomic architecture of the zone. However, asymmetric introgression with a predominance of Sitka spruce ancestry suggests pre- or post-zygotic barriers may influence structure within this zone. Surprisingly,

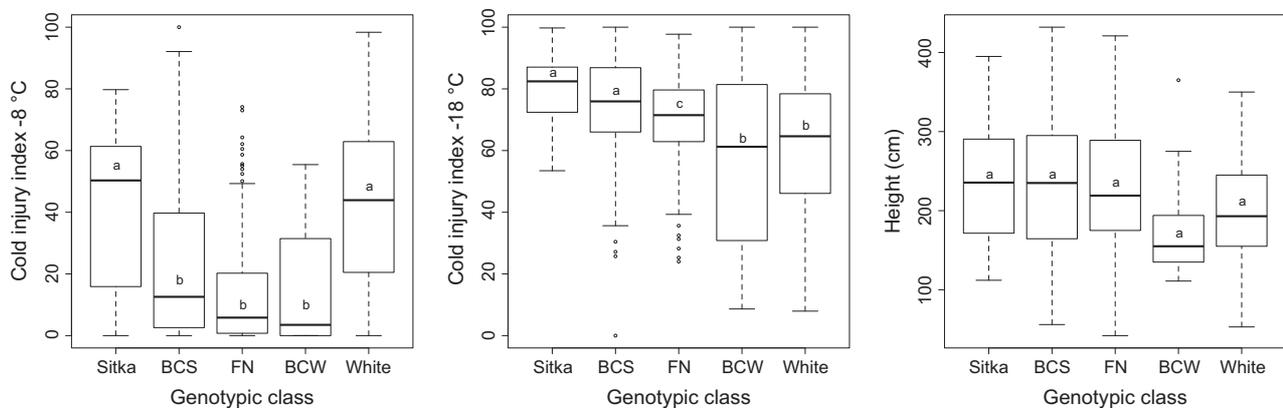


Fig. 4 Boxplots illustrating variation in phenotypic traits amongst genotypic classes including pure parentals (Sitka and white), backcrosses to either Sitka (BCS) or white spruce (BCW), and advanced-generation hybrids (FN). Phenotypic traits are cold injury at -8 and -18 °C estimated in artificial freeze tests, and height at age ten in the common garden experiment. Boxplots with same letter are not significantly different based on Tukey's comparison of means ($\alpha = 0.05$).

Table 3 ANOVA results for comparison of genotypic class across phenotypes putatively related to fitness in trees, along with non-parametric validation and significant pairwise comparisons based on Tukey's honest significant difference test for multiple comparisons

Source	SS	DF	MS	F (P-value)	X ² (P-value)
Cold injury -8 °C					
Genotype class	31 545	4	7886.3	14.874 (P < 0.001)	44.2922 (P < 0.001)
Error	367 430	693	530.2		
Cold injury -18 °C					
Genotype class	14 817	4	3704.2	15.046 (P < 0.001)	34.0644 (P < 0.001)
Error	170 114	691	246.2		
Height (cm)					
Genotype class	80 114	4	20 028.6	3.1126 (P = 0.01)	12.9968 (P = 0.01)
Error	4 562 176	709	6434.7		

a large proportion of phenotypic variation within the common garden remained unexplained by provenance geography or climate. Phenotypic variability suggests that hybrid classes are transgressive for cold hardiness at minimally damaging cold temperatures, but at lower temperatures, cold hardiness is lower in more maritime populations and greater in more continental populations, fitting expectations based on species' ecological differences.

Hybrid index structure

Most genotypes were classified as backcrosses toward Sitka spruce (69%) or as advanced-generation hybrids (FN—19%). There was no clear evidence of an F1 generation across the hybrid zone, consistent with an advanced-generation hybrid zone characterized by widespread interspecific recombination (Lexer *et al.* 2010; Teeter *et al.* 2010). The lack of early-generation hybrid derivatives is in line with selection for segregating hybrids within an ecologically transitional habitat, where hybrid genotypes may exhibit novel adaptations or combinations of traits (Rieseberg *et al.* 1999; Milne & Abbott 2008). Other hybrid zones studies have also observed a dearth of early-generation hybrid classes where F1-hybrid genotypes either make up a minority of hybrids present, or are absent (Milne & Abbott 2008; Lexer *et al.* 2010). Assignment of genotypic classes, however, comes with some expectations of error. Using our approach to class assignment, which was relatively flexible and broad, genotypes with very similar profiles are unlikely to be grouped inaccurately (Milne & Abbott 2008). The sampling scheme used for the common garden may have introduced a bias in the distribution of genotypic classes. This experiment was established to compare local adaptation and productivity amongst seed sources for reforestation within the hybrid zone, rather than at the periphery and beyond.

The lack of species-specific markers may introduce potential biases with respect to the genetic contribution of Sitka spruce or white spruce. However, these methods have been used to handle more variable markers (Nolte *et al.* 2009; Gompert *et al.* 2010), and largely congruent patterns are observed amongst all SNPs and a subset of those exhibiting allele frequency differentials over 0.90 (Fig. 2B). An additional caveat may be the use of SNPs pre-selected from putatively adaptive genes. Consistent with evidence from *Arabidopsis thaliana*, which suggests different genes may be targets of selection in different environments (Hancock *et al.* 2011), outlier analysis indicates only three of the 268 loci we studied here appear to be significant F_{ST} outliers. (J. A. Hamilton, C. Lexer, and S.N. Aitken unpublished). Exclusion of these loci from the present analysis did not change the results.

Fitness of genotypic classes

The ability to grow competitively and develop sufficient cold hardiness in advance of first frost is critical for the health and survival of boreal tree species (Morgenstern 1996; Howe *et al.* 2003). Increased cold hardiness in hybrids relative to parental genotypes within the hybrid zone provides support for the bounded hybrid superiority model. Although we have not directly tested the model using reciprocal transplants amongst parental species and hybrid environments, we have examined fitness proxies for both parental and hybrid genotypes within this zone. Interestingly, we saw that hybrid genotypes exhibited slightly greater cold hardiness than parental species on average when tested at -8 °C, but that with colder test temperatures this pattern changed (Fig. 4). Heterosis in early-generation hybrids usually results from dominant allelic effects, but is generally not retained in advanced-generation recombinants (Rieseberg & Carney 1998). As cold injury patterns at

–8 °C are common for all hybrid classes, none of which are early-generation hybrids, heterosis is unlikely. Rather, this pattern may reflect transgressive segregation due to complementary gene action of alleles differentiated in parental populations combining to produce phenotypic extremes (Rieseberg *et al.* 1999; Latta *et al.* 2006). Given the well-differentiated climatic niches exhibited between species, we predict that opposing allele combinations within the species may contribute to transgressive phenotypes for hybrid individuals. We are currently examining both genomic and geographic clines of mapped SNP loci that may allow us to identify regions of the genome that are well differentiated between parent species, along with locus-specific effects of candidate genes that may play a role in cold adaptation (J. A. Hamilton, C. Lexer and S. N. Aitken unpublished).

Surprisingly, there was little variation in height amongst genotypic classes, and no significant pairwise differences after accounting for multiple comparisons. The only class to exhibit slight differences was the backcross to white spruce class, which also exhibited the greatest cold tolerance across genotypic classes at both test temperatures. Although this class may exhibit the well-established trade-off between growth and development of cold hardiness, it also has the fewest members of any class ($N = 15$). Intrinsic incompatibilities may contribute to lower overall fitness and lower numbers; however, the lower numbers could be an artefact of sampling. In addition, the test site had a relatively benign, intermediate climate [mean annual temperature (MAT) = 2.9 °C and mean annual precipitation (MAP) = 802 mm]. Height data were analysed for two additional common garden sites on which this experiment was planted within the hybrid zone that exhibited more extreme temperatures (more continental site, MAT = 2.2 °C, MAP = 1057 mm; and more maritime site, MAT = 5 °C, MAP = 1578 mm), and a similar lack of significant provenance or climatic variance in height across genotypic classes was observed (unpublished).

Historic introgression

Low levels of interspecific heterozygosity combined with a high frequency of hybrids suggest this hybrid zone is the product of both long-term contact and widespread recombination between the two species (Fig. 2). Sitka and white spruce are closely related (Ran *et al.* 2006) and their current distribution and structure reflect events during the Pliocene epoch, and subsequent expansion and contraction of ice sheets during the Pleistocene (Morgenstern 1996). Recolonization history suggests these species came into contact most recently following the Pleistocene, but may have had repeated,

intermittent contact during glacial and interglacial cycles (Pisaric *et al.* 2003). Pollen records from eastern British Columbia suggest a dominance of *Picea* pollen, putatively attributed to white spruce, only 9000 YBP (Pisaric *et al.* 2003), whereas Sitka spruce rapidly expanded its range along the Pacific coast ~15 000 YBP following the Quaternary period (Pielou 1991; Mimura & Aitken 2007a; Shafer *et al.* 2010). Molecular evidence and recent climate envelope modelling support rapid expansion of Sitka spruce following a stepping stone model of serial recolonization along the Pacific coast, placing it within the Nass-Skeena region well in advance of white spruce (Mimura & Aitken 2007a; Holliday *et al.* 2010b). These findings suggest introgression may have first occurred between the putative local species (Sitka spruce) and later colonizing species (white spruce), contributing to asymmetrical patterns of introgression towards Sitka spruce (Currat *et al.* 2008; Petit & Excoffier 2009).

One consequence of the greater movement of Sitka spruce genome towards white spruce populations than the reverse is that new alleles or allelic combinations will expand the gene pool of Sitka spruce further into the hybrid zone over time, potentially constituting a selective advantage for Sitka genes within the hybrid zone (Keim *et al.* 1989). Introgression may provide novel allelic variation and allelic combinations, resulting in an increased ability to adapt to new conditions (Aitken *et al.* 2008).

Pre- and post-zygotic incompatibilities

Pre- and post-zygotic incompatibilities may also generate asymmetric introgression (Rieseberg & Soltis 1991; Rieseberg & Blackman 2010). Anecdotal evidence from artificial crosses between these species indicates no intrinsic post-zygotic barriers to reproduction exist in the F1 generation (J. King, personal communication) although genetic incompatibilities (Bombliès & Weigel 2007) that can cause hybrid breakdown in the F2 generation cannot be ruled out. The average wind direction from April to June within the hybrid zone (Environment Canada) suggests greater opportunity for coastal pollen to flow towards the interior than the reverse, which may contribute to asymmetry. The timing of reproduction, a potential pre-zygotic barrier, is determined jointly by highly heritable heat sum requirements (Hanninen & Tanino 2011) and the rate of heat sum accumulation in a given location. Heat sum accumulation above a threshold of 5 °C estimated from ClimateBC varies from 848 to 1571 degree days across the locations sampled, and higher heat sums on the coast may favour dispersal of Sitka spruce pollen to receptive white spruce over the reverse. Although genetic differences in heat sum requirements between

these species have not been studied, they may contribute to the patterns observed.

Genetic and phenotypic relationships with climate and distance

Strong clinal variation in SNP-based hybrid index indicates that the transition between maritime and continental climates along the Nass and Skeena rivers strongly impacts the genetic structure of populations throughout the contact zone. We observed similar climatic and geographic patterns for hybrid index using hundreds of potentially adaptive SNP markers to those seen by Bennuah *et al.* (2004) for a small number of putatively neutral markers. Strong relationships between hybrid index and drainage distance may be attributed to isolation by distance or to climate; however, concordant and strong correlations with climatic variables suggest environmental selection may be a more pervasive force influencing spatial structure. Although hybrid index exhibited significant clines with all climate variables tested, MAP was associated with particularly steep clines at ≤ 1000 mm of precipitation (Fig. 3). Introgressants may not tolerate high levels of precipitation, or may be out-competed by pure Sitka spruce under wetter conditions, but in drier environments there appears to be strong selection against Sitka spruce. This clinal variation in hybrid index along moisture, and to a lesser extent, temperature gradients suggests exogenous selection, and supports the bounded hybrid superiority model of hybrid zone persistence.

Clinal variation for phenotypic traits was surprisingly weak across the introgression zone. Cold injury phenotypes exhibited shallow clines, whereas height exhibited no clinal pattern. The large degree of variation within populations observed in phenotypes is typical of hybrid zones (Rieseberg & Carney 1998; Arnold *et al.* 2008). Different gene combinations can result in similar hybrid indices, but produce a range of phenotypic values. Our expectations were that increased variation due to introgression would be counter-balanced by spatially varying environmental selection, producing phenotypic clines associated with climatic variance. In Sitka spruce, phenotypic variation in cold hardiness and bud set phenology is associated with at least 35 SNPs, of which 15 show temperature and latitudinal clines in frequency across the species range (Holliday *et al.* 2010a). Whereas a large proportion of phenotypic variation is explained by geographic and climatic variation within Sitka spruce, patterns of phenotypic variation between Sitka and white spruce species are less clear. Based on the bounded hybrid superiority model of hybrid zone maintenance we do not necessarily expect strong clinal variation in phenotypes. Particularly where cold toler-

ance at moderate freezing temperatures confers a fitness advantage for hybrids, we would not expect a clinal gradient. Few individuals were sampled from the eastern end of the introgression zone, and this may also have contributed to the lack of significant clines amongst phenotypic traits. Additional common garden sites into allopatric environments, along with assessment of additional phenotypic traits such as drought tolerance and pest resistance, may have provided a more robust phenotypic test of the bounded hybrid superiority model.

Seed transfer and climate change

Our geographic and climatic clines for hybrid index provide guidance for how far seed can be transferred for reforestation and maintain local adaptation to current and future climates. Given the relatively flat clines across areas with high hybrid index values, it seems reasonable that seed can be moved amongst locations within this largely Sitka spruce area (hybrid index ≥ 0.80). Based on drainage distance, this corresponds to the area from the ocean to ~ 150 km up river valleys.

More care needs to be taken with seed transfer farther inland, where hybrid index is below 0.8 and climatic clines are steeper. Here we use the example of the Kitwanga population (drainage distance = 220 km, MAP 757 mm, MAT 5.4 °C, continentality 22.0 °C). The mean hybrid index of the 24 individuals genotyped from this population is 0.60, with a wide range amongst individuals from 0.47 to 0.86. Drainage distance explains most variation in hybrid index, and is the easiest measure for use in seed transfer decisions. If the goal is to plant seedlings with an average hybrid index within 10% of the value of local populations, seed could be collected from up to 25 km farther up- or downstream from the planting location. Given the high level of variation within populations, however, this may be overly restrictive.

Planning for future climates becomes more challenging. This region has already become warmer and wetter in the past 25 years compared with the base climate normal period of 1961–1990 (Mbogga *et al.* 2010). Using ClimateBC and the middle-of-the-road general circulation model CGCM A2 carbon dioxide scenario, the projected climate for Kitwanga, BC in the 2080s is for a slight increase in precipitation (to MAP = 803 mm), a large increase in temperature (to MAT = 9.4 °C) and a decrease in the difference between summer and winter mean temperatures (to Continentality = 20.1 °C). Assuming current populations are optimally adapted for our 1961–1990 climate normal period, the optimal hybrid index values derived from regressions in Fig. 3 for the 2080's are 0.68 for MAP, 0.79 for Continentality, and 1.0

for MAT. The lack of congruence amongst temperature and moisture variables makes it difficult to speculate on optimal hybrid index values for a warmer future, as the increase in precipitation does not match the increase in temperatures. Although this location may be warm enough for pure Sitka in the future, it may not be wet enough. However, the great variability in natural hybrid populations, like Kitwanga, suggests there is considerable standing variation in these populations for adaptation to future conditions.

Acknowledgements

We thank John King and the British Columbia Forest Service for establishing the common garden and providing phenotypic data for growth, Christine Chourmouzis, Lisa Erdle, Nina Lobo, and Jon Sweetman for field assistance, Pia Smets for freeze testing, Dorothea Lindtke for advice with analysis, and Loren Rieseberg and Kermit Ritland for helpful comments on this manuscript. The work was supported by Genome British Columbia, Genome Canada and the Province of British Columbia (grant to S.A.), by the Natural Science and Engineering Research Council of Canada (NSERC; grant to S.A.), and by a NSERC Canada Graduate Scholarship to J.H.

References

- Aitken SN, Yeaman S, Holliday JA, Wang T, Curtis-McLane S (2008) Adaptation, migration or extirpation: climate change outcomes for tree populations. *Evolutionary Applications*, **1**, 95–111.
- Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. *Genetics*, **160**, 1217–1229.
- Anderson LL, Hu FS, Nelson DM, Petit RJ, Paige KN (2006) Ice-age endurance: DNA evidence of a white spruce refugium in Alaska. *Proceedings of the National Academy of Sciences*, **103**, 12447–12450.
- Arnold ML (1997) *Natural Hybridization and Evolution*. Oxford University Press, New York, New York.
- Arnold ML, Cornman RS, Martin NH (2008) Hybridization, hybrid fitness and the evolution of adaptations. *Plant Biosystems – An International Journal Dealing with all Aspects of Plant Biology*, **142**, 166–171.
- Barton NH, Hewitt GM (1985) Analysis of hybrid zones. *Annual Review of Ecology and Systematics*, **16**, 113–148.
- Beaulieu J, Doerksen T, Boyle B *et al.* (2011) Association genetics of wood physical traits in the conifer white spruce and relationships with gene expression. *Genetics*, **188**, 197–214.
- Bennuah SY, Wang T, Aitken SN (2004) Genetic analysis of the *Picea sitchensis* × *glauca* introgression zone in British Columbia. *Forest Ecology and Management*, **197**, 65–77.
- Bomblies K, Weigel D (2007) Hybrid necrosis: autoimmunity as a potential gene-flow barrier in plant species. *Nature Reviews Genetics*, **8**, 382–393.
- Buerkle CA (2005) Maximum-likelihood estimation of a hybrid index based on molecular markers. *Molecular Ecology Notes*, **5**, 684–687.
- Currat M, Ruedi M, Petit RJ, Excoffier L (2008) The hidden side of invasions: massive introgression by local genes. *Evolution*, **62**, 1908–1920.
- Doyle JJ, Doyle JL (1990) Isolation of plant DNA from fresh tissue. *Focus*, **12**, 13–15.
- Endler JA (1977) *Geographic variation, speciation, and clines*. Princeton Univ. Press, Princeton, NJ.
- Fan S, Grossnickle SC, Sutton BCS (1997) Relationships between gas exchange adaptation of Sitka × interior spruce genotypes and ribosomal DNA markers. *Tree Physiology*, **17**, 115–123.
- Fan J, Gunderson K, Bibikova M *et al.* (2006) Illumina universal bead arrays. *Methods in Enzymology*, **410**, 57–73.
- Field DL, Ayre DJ, Whelan RJ, Young AG (2011) The importance of pre-mating barriers and the local demographic context for contemporary mating patterns in hybrid zones of *Eucalyptus aggregata* and *Eucalyptus rubida*. *Molecular Ecology*, **20**, 2367–2379.
- Flint HL, Boyce BR, Beattie DJ (1967) Index of injury – a useful expression of freezing injury to plant tissues as determined by the electrolytic method. *Canadian Journal of Plant Science*, **47**, 229–230.
- Fritz RS, Hochwender CG, Albrechtsen BR, Czesak ME (2006) Fitness and genetic architecture of parent and hybrid willows in common gardens. *Evolution*, **60**, 1215–1227.
- Gompert Z, Buerkle CA (2010) introgress: a software package for mapping components of isolation in hybrids. *Molecular Ecology Resources*, **10**, 378–384.
- Gompert Z, Buerkle CA (2011) Bayesian estimation of genomic clines. *Molecular Ecology*, **20**, 2111–2127.
- Gompert Z, Lucas LK, Fordyce JA, Forister ML, Nice CC (2010) Secondary contact between *Lycaeides idas* and *L. melissa* in the Rocky Mountains: extensive admixture and a patchy hybrid zone. *Molecular Ecology*, **19**, 3171–3192.
- Grossnickle SC, Sutton BCS, Folk RS, Gawley RJ (1996) Relationship between nuclear DNA markers and physiological parameters in Sitka × interior spruce populations. *Tree Physiology*, **16**, 547–555.
- Hancock AM, Brachi B, Faure N *et al.* (2011) Adaptation to climate across the *Arabidopsis thaliana* genome. *Science*, **334**, 83–86.
- Hannerz M, Aitken SN, King JN, Budge S (1999) Effects of genetic selection for growth on frost hardiness in western hemlock. *Canadian Journal for Forest Research*, **29**, 509–516.
- Hanninen H, Tanino K (2011) Tree seasonality in a warming climate. *Trends in Plant Science*, **16**, 412–416.
- Holliday JA, Ralph SG, White R, Bohlmann J, Aitken SN (2008) Global monitoring of autumn gene expression within and among phenotypically divergent populations of Sitka spruce (*Picea sitchensis*). *New Phytologist*, **178**, 103–122.
- Holliday JA, Ritland K, Aitken SN (2010a) Widespread, ecologically relevant genetic markers developed from association mapping of climate-related traits in Sitka spruce (*Picea sitchensis*). *New Phytologist*, **188**, 501–514.
- Holliday JA, Yuen M, Ritland K, Aitken SN (2010b) Postglacial history of a widespread conifer produces inverse clines in selective neutrality tests. *Molecular Ecology*, **19**, 3857–3864.
- Howe GT, Aitken SN, Neale DB *et al.* (2003) From genotype to phenotype: unraveling the complexities of cold adaptation in forest trees. *Canadian Journal of Botany*, **81**, 1247–1266.

- Keim P, Paige KN, Whitham TG, Lark KG (1989) Genetic analysis of an interspecific hybrid swarm of *Populus*: occurrence of unidirectional introgression. *Genetics*, **123**, 557–565.
- Ketcheson MV, Braudmandl TF, Meidinger D *et al.* (1991) Interior Cedar-Hemlock Zone. *Ecosystems of British Columbia*.
- Kruuk LEB, Baird SJE, Gale KS, Barton NH (1999) A comparison of multilocus clines maintained by environmental adaptation or by selection against hybrids. *Genetics*, **153**, 1959–1971.
- Latta RG, Gardner KM, Johansen-Morris AD (2006) Hybridization, recombination, and the genetic basis of fitness variation across environments in *Avena barbata*. *Genetica*, **129**, 167–177.
- Lexer C, Henize B, Alia R, Rieseberg LH (2004) Hybrid zones as a tool for identifying adaptive genetic variation in outbreeding forest trees: lessons from wild annual sunflowers (*Helianthus* spp.). *Forest Ecology and Management*, **197**, 49–64.
- Lexer C, Joseph JA, van Loo M *et al.* (2010) Genomic admixture analysis in European *Populus* spp. reveals unexpected patterns of reproductive isolation and mating. *Genetics*, **186**, 699–712.
- Mbogga MS, Wang X, Hamann A (2010) Bioclimate envelope model predictions for natural resource management: dealing with uncertainty. *Journal of Applied Ecology*, **47**, 731–740.
- Miglia KJ, McArthur ED, Moore WS *et al.* (2005) Nine-year reciprocal transplant experiment in the gardens of the basin and mountain big sagebrush (*Artemisia tridentata*: Asteraceae) hybrid zone of Salt Creek Canyon: the importance of multiple-year tracking of fitness. *Biological Journal of the Linnean Society*, **86**, 213–225.
- Milne RI, Abbott RJ (2008) Reproductive isolation among two interfertile *Rhododendron* species: low frequency of post-F1 hybrid genotypes in alpine hybrid zones. *Molecular Ecology*, **17**, 1108–1121.
- Mimura M, Aitken SN (2007a) Adaptive gradients and isolation-by-distance with postglacial migration in *Picea sitchensis*. *Heredity*, **99**, 224–232.
- Mimura M, Aitken SN (2007b) Increased selfing and decreased effective pollen donor number in peripheral relative to central populations in *Picea sitchensis* (Pinaceae). *American Journal of Botany*, **94**, 991–998.
- Moore WS (1977) An evaluation of narrow hybrid zones in vertebrates. *The Quarterly Review of Biology*, **52**, 263–277.
- Morgenstern EK (1996) *Geographic Variation in Forest Trees*. UBC Press, Vancouver, British Columbia.
- Namroud M-C, Beaulieu J, Juge N, Laroche J, Bousquet J (2008) Scanning the genome for gene single nucleotide polymorphisms involved in adaptive population differentiation in white spruce. *Molecular Ecology*, **17**, 3599–3613.
- Nolte AW, Gompert Z, Buerkle CA (2009) Variable patterns of introgression in two sculpin hybrid zones suggest that genomic isolation differs among populations. *Molecular Ecology*, **18**, 2615–2627.
- O'Neill GA, Aitken SN, King JN, Alfaro RI (2002) Geographic variation in resin canal defenses in seedlings from the Sitka spruce × white spruce introgression zone. *Canadian Journal of Forest Research*, **32**, 390–400.
- Pavy N, Pelgas B, Beauseigle S *et al.* (2008) Enhancing genetic mapping of complex genomes through the design of highly-multiplexed SNP arrays: application to the large and unsequenced genomes of white spruce and white spruce. *BMC Genomics*, **9**, 21–38.
- Pelgas B, Bousquet J, Meirmans PG, Ritland K, Isabel N (2011) QTL mapping in white spruce: gene maps and genomic regions underlying adaptive traits across pedigrees, years and environments. *BMC Genomics*, **12**, 145–168.
- Petit RJ, Excoffier L (2009) Gene flow and species delimitation. *Trends in Ecology & Evolution*, **24**, 386–393.
- Pielou EC (1991) *After the Ice Age: the Return of Life to Glaciated North America*. University of Chicago Press, Chicago, Illinois.
- Pisaric MFJ, Holt C, Szeicz JM, Karst T, Smol JP (2003) Holocene treeline dynamics in the mountains northeastern British Columbia, Canada, inferred from fossil pollen and stomata. *The Holocene*, **13**, 161–173.
- Pojar J, Klinka K, Demarchi DA (1991) *Coastal Western Hemlock Zone*.
- Porth I, Hamberger B, White R, Ritland K (2011) Defense mechanisms against herbivory in *Picea*: sequence evolution and expression regulation of gene family members in the phenylpropanoid pathway. *BMC Genomics*, **12**, 608–634.
- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.
- R Development Core Team (2007) *R: A Language and Environment for Statistical Computing* (ed. Computing RFFS), Vienna, Austria.
- Ran J-H, Wei X-X, Wang X-Q (2006) Molecular phylogeny and biogeography of *Picea* (Pinaceae): implications for phylogeographical studies using cytoplasmic haplotypes. *Molecular Phylogenetics and Evolution*, **41**, 405–419.
- Rieseberg LH, Blackman BK (2010) Speciation genes in plants. *Annals of Botany*, **106**, 439–455.
- Rieseberg LH, Carney SE (1998) Plant hybridization. *New Phytologist*, **140**, 599–624.
- Rieseberg LH, Soltis DE (1991) Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants*, **5**, 65–84.
- Rieseberg LH, Wendel JF (1993) *Introgression and Its Consequences in Plants. Hybrid Zones and the Evolutionary Process*. Oxford University Press, New York, New York.
- Rieseberg LH, Archer MA, Wayne RK (1999) Transgressive segregation, adaptation and speciation. *Heredity*, **83**, 363–372.
- Roche L (1969) A genealogical study of the genus *Picea* in British Columbia. *New Phytologist*, **68**, 505–554.
- Shafer ABA, Cullingham CI, côté SD, Coltman DW (2010) Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, **19**, 4589–4621.
- Shen R, Fan J, Campbell D *et al.* (2005) High-throughput SNP genotyping on universal bead arrays. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, **573**, 70–82.
- Silim SN, Guy RD, Patterson TB, Livingston NJ (2001) Plasticity in water-use efficiency of *Picea sitchensis*, *P. glauca* and their natural hybrids. *Oecologia*, **128**, 317–325.
- Spitze K (1993) Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. *Genetics*, **135**, 367–374.
- Stölting KN, Nipper R, Lindtke D *et al.* (2012) Genomic scan for single nucleotide polymorphisms reveals patterns of divergence and gene flow between ecologically divergent species. *Molecular Ecology*, doi: 10.1111/mec.12011.
- Streiff R, Ducouso A, Lexer C *et al.* (1999) Pollen dispersal inferred from paternity analysis in a mixed oak stand of *Quercus robur* L. and *Q. petraea* (Matt.) Liebl. *Molecular Ecology*, **8**, 831–841.
- Sutton BCS, Pritchard SC, Gawley JR, Newton CH (1994) Analysis of Sitka spruce – interior spruce introgression in British

Columbia using cytoplasmic and nuclear DNA probes. *Canadian Journal for Forest Research*, **24**, p. 278.

- Teeter KC, Thibodeau LM, Gompert Z *et al.* (2010) The variable genomic architecture of Isolation between hybridizing species of house mice. *Evolution*, **64**, 472–485.
- Wang H, McArthur ED, Sanderson SC, Graham JH, Freeman DC (1997) Narrow hybrid zone between two subspecies of Big Sagebrush (*Artemisia tridentata*: Asteraceae). IV. Reciprocal transplant experiments. *Evolution*, **51**, 95–102.
- Wang T, Hamann A, Spittlehouse D, Aitken SN (2006) Development of scale-free climate data for western Canada for use in resource management. *International Journal of Climatology*, **26**, 383–397.
- Wilhelm R, Hilbish TJ (1998) Assessment of natural selection in a hybrid population of mussels: evaluation of exogenous vs endogenous selection models. *Marine Biology*, **131**, 505–514.
- Yeh FC, Arnott JT (1986) Electrophoretic and morphological differentiation of *Picea sitchensis*, *Picea glauca*, and their hybrids. *Canadian Journal of Forest Research*, **16**, 791–798.

This study is part of the PhD thesis of J.A.H. on the genetic mechanisms underlying local adaptation in the *Picea sitchensis* × *P. glauca* hybrid zone. J.A.H. is broadly interested in the genetic consequences of introgression within natural tree populations. C.L. is interested in evolutionary genomics of adaptation and speciation in plants and provided analytical support. S.N.A. is a forest geneticist investigating climate change outcomes for trees using population genomic approaches, and is interested in ecological and conservation genetics.

Data accessibility

Population origin data, phenotypic traits, and SNP genotype data: DRYAD entry doi:10.5061/dryad.s11b6

Supporting information

Table S1 Summary of 268 candidate gene SNPs used for genetic analysis in hindex, including species-specific alleles and their frequency within reference Sitka spruce (QCI) and white spruce (ENA) populations, respectively and the allele frequency differentials (δ) between the two species.

Table S2 Summary of multiple stepwise regressions performed in *R* of distance (drainage distance), and climate variables (mean annual temperature, mean annual precipitation, and continentality) with dependent variables; hybrid index, cold injury, and height.

Fig. S1 Population clustering analyses based on ten replicate *Structure* runs for Sitka spruce, white spruce, and hybrid populations.

Fig. S2 Genetic relationship amongst 31 populations estimated using *Structure* and genotypic data based on 268 SNP loci for $K = 2$ for Sitka spruce (red) and white spruce (green) individuals and their hybrids.

Fig. S3 Genetic relationship between three ‘reference’ populations (Haida Gwaii, eastern North America and Fort Nelson) estimated using *Structure* and genotypic data based on 120 SNP loci for $K = 2$ and 3 for Sitka spruce (green) and white spruce (red).