

ORIGINAL ARTICLE

Importance of demographic history for phylogeographic inference on the arctic–alpine plant *Phyllodoce caerulea* in East Asia

H Ikeda¹, S Sakaguchi², V Yakubov³, V Barkalov³ and H Setoguchi⁴

Arctic–alpine plants have enormous ranges in the Northern Hemisphere. Phylogeographic studies have provided insights into their glacial survival as well as their postglacial colonization history. However, our understanding of the population dynamics of disjunct alpine populations in temperate regions remains limited. During Pleistocene cold periods, alpine populations of arctic–alpine species in East Asia were either connected to an ice-free Beringia refugium or they persisted with prolonged isolation after their establishment. To estimate which of these scenarios is more likely, we elucidated the genetic structure of *Phyllodoce caerulea* (Ericaceae) in Beringia and northern Japan, East Asia. Sequence variation in multiple nuclear loci revealed that *P. caerulea* can be distinguished into northern and southern groups. A demographic analysis demonstrated that the north–south divergence did not predate the last glacial period and detected introgression from *Phyllodoce aleutica*, relative widely distributed in East Asia, exclusively into the southern group. Therefore, although there has been genetic divergence between northern Japan and Beringia in *P. caerulea*, the divergence is unlikely to have resulted from their prolonged geographic separation throughout several cycles of glacial and interglacial periods. Instead, our study suggests that the introgression contributed to the genetic divergence of *P. caerulea* and that the range of *P. caerulea* was plausibly connected between northern Japan and Beringia during the last glacial period. Overall, our study not only provides a biogeographic insight into alpine populations of arctic–alpine plants in East Asia but also emphasizes the importance of careful interpretation of genetic structure for inferring phylogeographic history.

Heredity (2016) **116**, 232–238; doi:10.1038/hdy.2015.95; published online 4 November 2015

INTRODUCTION

Arctic–alpine plants have extensive ranges in the Northern Hemisphere encompassing the Arctic and high mountains in Europe, North America, central Asia and East Asia (see, for example, Hultén and Fries, 1986). As most of the current distribution ranges of arctic–alpine plants were covered with glaciers during the Pleistocene cold periods (Hultén, 1937), their present wide distributions have been formed through range shifts following the Pleistocene climate oscillations. Molecular investigations have provided insights into the biogeographic history of arctic–alpine plants with respect to postglacial colonization or glacial survival (see, for example, Abbott *et al.*, 2000; Alsos *et al.*, 2007). In contrast, our understanding of the Pleistocene population dynamics of alpine populations on mountains at temperate regions remains limited (Winkler *et al.*, 2012). Temperate alpine populations were not only the source of postglacial colonization to the Arctic (Schönswetter *et al.*, 2003) but also received migrants from the Arctic (Schönswetter *et al.*, 2006). These populations were also considered to have experienced prolonged persistence within temperate mountains (see, for example, Ikeda *et al.*, 2009; Winkler *et al.*, 2012).

East Asia is located south-west of Beringia, an important ice-free refugium of arctic–alpine species (Hultén, 1937), and was less glaciated during the Pleistocene (Frenzel *et al.*, 1992). These features of East

Asia have enabled arctic–alpine plants to either migrate there repeatedly or persist in isolation through several cycles of warm and cold periods. Therefore, current disjunct distributions of arctic–alpine plants in East Asia plausibly have a complicated history; that is, both repeated migrations and prolonged persistence through several cycles of warm and cold periods. High mountains in the Japanese Archipelago represent the southernmost range margins of arctic–alpine plants in this region. According to palaeobotanical studies, however, the tundra vegetation of Beringia did not spread into Hokkaido, the northern part of the Japanese Archipelago, during the last glacial maximum (Frenzel *et al.*, 1992). This implies that arctic–alpine plants may have colonized their alpine populations in Japan before the last glacial period. On the other hand, Japanese alpine communities are mostly dominated by species that also occur in the Arctic (for example, *Vaccinium vitis-idaea*, *Empetrum nigrum* and *Diapensia lapponica*). The similarity of Japanese alpine vegetation to that of the arctic tundra may represent recent colonization of arctic–alpine plants in this region. Phylogeographic studies of alpine plants in this region revealed genetic structures that are consistent with both of the hypotheses. Findings of distinct haplotypes between Hokkaido and Beringia (Fujii *et al.*, 1997, 1999) suggest the former hypothesis of prolonged isolation of Japanese populations from Beringia. In

¹Institute of Plant Science and Resources, Okayama University, Kurashiki, Japan; ²Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan; ³Institute of Biology and Soil Sciences, Russian Academy of Sciences, Far East Branch, Vladivostok, Russia and ⁴Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Japan

Correspondence: Dr H Ikeda, Institute of Plant Science and Resources, Okayama University, 2-20-1 Chuo, Kurashiki, Okayama 710-0046, Japan.
E-mail: ike@okayama-u.ac.jp

Received 16 March 2015; revised 18 September 2015; accepted 21 September 2015; published online 4 November 2015

particular, *Primula cuneifolia*, a species occupying habitats around snow-beds, harboured a unique haplotype in central Hokkaido (Fujii *et al.*, 1999). In contrast, other studies demonstrated genetic similarities between Hokkaido and Beringia (Shimizu-Inatsugi *et al.*, 2009; Ikeda *et al.*, 2014a), suggesting that alpine populations in Hokkaido and Beringia could have been connected during the last glacial period. Accordingly, further investigations are required to address whether alpine populations of arctic–alpine species in East Asia were either (1) persistent in isolation through several cycles of Pleistocene warm and cold periods after their establishment or (2) connected with Beringia during Pleistocene cold periods.

Although comparative phylogeography, in which several co-occurring species are studied jointly, is the most efficient approach to addressing biogeographic history across communities (see, for example, Taberlet *et al.*, 1998), the current geographic patterns of genetic variations have been shaped not only by the historical range dynamics but also by other evolutionary forces (Avice, 2004). For example, introgression from divergent species into only one part of a species' geographic range results in a genetic differentiation between the region affected and the region not affected by introgression, as if these two regions had been isolated for prolonged periods. In particular, as co-occurrence of currently allopatric species may have occurred in the past because of the Pleistocene climate oscillations, introgression at historical sympatric areas may result in current genetic structures (see, for example, Pelsler *et al.*, 2012). Recent coalescent-based analyses of multiple loci enable us to examine introgression from related species (see, for example, Wegmann *et al.*, 2010; Hey, 2010). Such a statistical approach in phylogeography, statistical phylogeography (Knowles, 2009), is crucial for inferring a robust biogeographic history from genetic structure. In addition, although introgression influences genetic diversity throughout the genome, natural selection also influences geographic patterns of genetic diversity but only in some genes (Ikeda and Setoguchi, 2010). In particular, variation in regional environments within ranges of species and the associated distinct selection regimes may result in genetic differentiation that mimics prolonged geographic isolation following historical range shifts. Accumulation of database information regarding the global climate and development of relevant analyses techniques applying geographic information systems GIS (Hijmans *et al.*, 2005) allow us to explore climate heterogeneity throughout the ranges of species and assess the relevance of genetic structure with regional climate. Accordingly, careful interpretation of the genetic structure of each species is still important to learn their biogeographic histories.

Phyllodoce caerulea (L.) Bab. (Ericaceae) is an evergreen shrub with a distribution range encompassing the Arctic as well as high mountains in central and East Asia (Yamazaki, 1993). This species dominates alpine communities around snow-beds in high mountains in Hokkaido. Recent phylogenetic study based on sequences of 12 nuclear loci revealed that four of ca. seven *Phyllodoce* species including *P. caerulea* form a clade, and each of the four species are genetically distinguishable (Ikeda *et al.*, 2014b), although the species relationships remain unresolved because of gene flow between some pairs of species. In East Asia, *P. caerulea* frequently co-occurs with *Phyllodoce aleutica*, one of the four species in the same clade, and gene flow was detected between them (Ikeda *et al.*, 2014b). Although backcrosses rarely occur between *P. caerulea* and *P. aleutica* in their hybrid zones in Hokkaido (see, for example, Kameyama and Kudo, 2011), the genetic structure of *P. caerulea* might be influenced by admixture with *P. aleutica* that disturbs phylogeographic inference. Thus, the spatial pattern of genetic diversity of *P. caerulea* and the influence of introgression on the intraspecific genetic structure remains elusive.

In this study, we aimed to provide insights into historical population dynamics of alpine populations of arctic–alpine plants by investigating the biogeographic history of the snow-bed dominant arctic–alpine shrub *P. caerulea* in East Asia. We elucidate a geographic distribution of genetic diversity of *P. caerulea* in Hokkaido and Beringia using sequences of 13 nuclear loci from representative samples. The genetic structure is confirmed by sequence variation of 6 of the 13 loci using numerous samples in East Asia. The divergence history between populations in Beringia and Japan as well as the related species, *P. aleutica*, is estimated by applying an isolation with migration model (Hey, 2010). We did not include the other species of *Phyllodoce* in Japan, *Phyllodoce nipponica*, because this species represents a distinct lineage from the monophyletic group including *P. caerulea* and *P. aleutica*, and did not experience gene flow with these two species (Ikeda *et al.*, 2014b). Furthermore, we assessed the impact of climate heterogeneity on the genetic structure using GIS. In addition to the inference of biogeographic history of *P. caerulea*, we provide a cautionary example for conventional interpretations of genetic structure for phylogeographic histories by considering the influence of introgression from a related species and natural selection due to climate heterogeneity.

MATERIALS AND METHODS

Genetic analysis

The genetic structure of *P. caerulea* was explored using previously analysed sequences from 12 individuals of this species (Supplementary Table S1). These individuals represent populations in the Japanese Archipelago, the Kuril Islands, Sakhalin and Beringia including Kamchatka and are genetically distinguishable from relatives (Ikeda *et al.*, 2014b). In addition to the previously analysed 12 loci, we added C35 that was previously sequenced in *P. nipponica* (Ikeda and Setoguchi, 2013) but was not analysed in Ikeda *et al.* (2014b) because of the failure of PCR amplification in an outgroup species. Sequences of this locus were determined following the previous procedure (Ikeda and Setoguchi, 2013). Using unphased genotypic data based on haplotypes of each of 13 loci, the probability of assigning individuals into ancestral clusters ($K=1$ to 12) was estimated using STRUCTURE ver. 2.3 (Falush *et al.*, 2003) with 1.0×10^6 generations following a 5.0×10^5 generations burn-in period. The optimal number of clusters (K) was identified based on ΔK criterion (Evanno *et al.*, 2005) and the consistency of configurations among 20 replicate runs was evaluated by the symmetric similarity coefficient (H') calculated using CLUMPP (Jakobsson and Rosenberg, 2007).

Additional DNA samples, encompassing several individuals per population, were extracted from 11 populations in East Asia (Table 1 and Supplementary Table S1) with the same procedure as in the previous study (Ikeda *et al.*, 2014b). In total, 71 individuals from these populations were analysed to confirm the geographic structure detected across the above 12 representative samples. Six loci (C11, C12, C20, C35, *PHYB* and *PHYE*) were selected to estimate genetic variation, because these loci provide longer sequences among the 13 loci analysed above and should be informative to elucidate unbiased geographic patterns. PCR amplification, sequencing and subsequent sequence analyses, including inference of alleles by Bayesian statistical methods (Stephens and Donnelly, 2003), followed reported procedures (Ikeda and Setoguchi, 2013). The inferred sequences of two alleles per locus are applied to all of the following analyses except to construct the neighbour-net. As F1 hybrids of *P. caerulea* and *P. aleutica* frequently occur in high mountains of Hokkaido (for example, Kameyama and Kudo, 2011), the genetic structure was resolved including 12 individuals of *P. aleutica*. These individuals were analysed previously and lacked a signature of genetic admixture from other relatives (Ikeda *et al.*, 2014b). A relationship among individuals without missing data was constructed by a neighbour-net using SplitsTree (Bryant and Moulton, 2004). As linkages among alleles of each locus were not clarified in this study, unphased sequences of each locus were used after concatenation, where averaging over all possible resolutions was applied for all ambiguous codes (such as R, Y) under the Uncorrected P method. The genetic clusters across

Table 1 Summary statistics of six nuclear loci of *Phyllodoce caerulea* in East Asia

Locus	Region ^a	n	bp	S	π	D	F_{CT}	P
C11	Total	124	412	6	0.0060	2.64	0.93	$P < 0.01$
	North	66	412	6	0.0020	-0.87		
	South	58	419	2	0.0002	-1.32		
C12	Total	138	440	13	0.0104	2.44	-0.13	$P = 0.80$
	North	74	440	10	0.0104	3.29		
	South	64	440	12	0.0105	2.33		
C20	Total	132	480	24	0.0083	-0.37	0.48	$P < 0.01$
	North	74	480	14	0.0044	-0.90		
	South	58	480	18	0.0072	-0.33		
C35	Total	132	521	14	0.0060	0.56	0.72	$P < 0.01$
	North	70	521	8	0.0028	-0.30		
	South	62	521	8	0.0020	-1.03		
PHYB	Total	138	510	6	0.0014	-0.73	0.43	$P < 0.01$
	North	74	510	5	0.0016	-0.40		
	South	64	510	2	0.0002	-1.32		
PHYE	Total	138	370	4	0.0004	-1.51	0.02	$P = 0.32$
	North	74	370	3	0.0006	-1.18		
	South	64	370	1	0.0001	-1.08		

Abbreviations: bp, length of sequences for analyses; D, Tajima's D; F_{CT} , genetic differentiation between North and South regions; n, number of sequences for analysis; P, P-values for F_{CT} evaluated by 10 000 permutation tests; S, number of segregating sites; π , average number of pairwise nucleotide differences per site calculated based on all sites.

The bold italic in D indicates a significance deviation from the standard neutral model ($P < 0.05$).

^aPopulations in Sakhalin and Beringia and those in Hokkaido and Kunashir Island are assigned in the North and South group, respectively.

these individuals were also identified by Bayesian clustering following the same procedure as the above-mentioned method.

Nucleotide diversity (π) (Nei, 1987) and Tajima's D (Tajima, 1989) were calculated for each locus using DnaSP ver. 5.10 (Librado and Rozas, 2009). Two individuals with genetic admixture with *P. aleutica* (see Results) were excluded. Neutral equilibrium for each locus was tested by 10 000 coalescent simulations based on Tajima's D. The geographic differentiation between populations in Hokkaido and Kunashir Island (South) and the more northern populations (North) was evaluated for each locus by fixation indexes (F_{CT}) in an analysis of molecular variance using Arlequin 3.5 (Excoffier *et al.*, 2005). The significance of the differentiation was tested by 10 000 permutations.

Estimating demographic history and examination of introgression

Introgression between *P. aleutica* and *P. caerulea* was examined by applying an isolation with migration model by incorporating the genetic structure of *P. caerulea* (Hey, 2010). As the number of loci is crucial for estimating demographic parameters with less variance (Arbogast *et al.*, 2002), we used sequences of 12 loci from 12 individuals that cover representative populations in East Asia and Alaska. C12 was excluded from the analysis because this locus exhibited a nonneutral pattern of polymorphisms (see Results). To incorporate the north-south differentiation of *P. caerulea* (see Results), we assumed two sister populations of *P. caerulea*, a northern population including Beringia and southern populations including Hokkaido as well as Kunashir Island. Thus, we assumed a topology of population where the northern and southern populations of *P. caerulea* were split following its divergence from *P. aleutica* (that is, ((North, South), *P. aleutica*)). We excluded the individual in Sakhalin from the analysis because its mixed genetic pattern cannot be unambiguously assigned into either northern or southern groups and because of its intermediate geographic location between the northern and southern groups.

Using the program IMA2 (Hey, 2010), all demographic parameters of the isolation with migration model (m : migration rates, θ : population sizes, t : divergence time) were estimated from 1.0×10^7 Markov chain Monte Carlo steps following a 100 000 generation burn-in period with 20 heated chains under the geometric increment model. Three independent runs were conducted, from which 100 000 genealogies were saved from each run.

Prior probability densities were optimized based on several preliminary runs using a wide range of densities ($m = 0.1$ (the mean of the exponential prior distribution); q (θ) = 0.5–2.5; $t = 0.5$ –1.5). Using all of the saved genealogies, the marginal posterior distribution, maximum likelihood estimates and 95% highest probability densities (HPDs) of each parameter were obtained. Following Ikeda *et al.* (2014b), the divergence time was scaled by geometric means of substitution rates per locus based on 5.3 – 7.8×10^{-9} substitutions/site/year. Likelihood ratio tests were conducted to examine whether migration rates greater than zero were statistically significant.

Climate difference

To characterize the climatic environments of each locality of *P. caerulea*, the values of 19 bioclimatic variables with 30 arc-sec resolution, which are provided by the Worldclim website (Hijmans *et al.*, 2005), were extracted using ArcGIS 10.2.2 (ESRI, Redlands, CA, USA). Latitude and longitude of the localities were extracted from collection records of *P. caerulea* downloaded from Arcots database (<http://arctos.database.museum/>) and geotags in photographs of *P. caerulea* taken during our fieldtrips in Hokkaido. In addition, the coordinates of populations used in the DNA analysis were included. After excluding the duplicated points within the same grid cells across all localities, principle component analysis was performed on scaled 19 bioclims using the `dudi.pca` function in R-package of 'ade4' (Chessel *et al.*, 2004).

RESULTS

Genetic structure in East Asia

Across 12 representative individuals of *P. caerulea*, the optimal number of clusters was two ($K = 2$, $\Delta K = 188.6$). The final configuration (Figure 1a) is consistent among 20 replicate runs ($H' = 1.00$) and revealed that individuals from Hokkaido and Beringia were assigned into distinct clusters. An individual from Sakhalin was assigned into each cluster with nearly equal probability (Supplementary Table S1).

The genetic distinctiveness was also detected by analysing 71 individuals using 6 of the 13 loci. The neighbour-net exhibited two groups in *P. caerulea*, Hokkaido and Beringia, including Sakhalin (Figure 1b). *P. aleutica* formed another group and is distinguishable from *P. caerulea*. An individual in Daisetsu was located in the basal position of *P. caerulea* and close to *P. aleutica* (green star in Figure 1b). In Bayesian clustering, the optimal number of clusters across *P. caerulea* and *P. aleutica* was three ($K = 3$, $\Delta K = 3360.8$; Supplementary Figure S1). The final configuration (Figure 1c and Supplementary Table S1) was consistent among 20 replicate runs ($H' = 1.00$) and resolved the three genetic groups found in the neighbour-net, whereas several individuals in Sakhalin were admixed. Two individuals in Daisetsu admixed with *P. aleutica* with some probability; one was the basal individual in the neighbour-net and another was not included in the neighbour-net. These individuals may originate from contemporary hybridization between *P. caerulea* and *P. aleutica*.

The north-south genetic differentiation, with Sakhalin assigned to the northern group, was significantly supported at four of six loci (Table 1; analysis of molecular variance, $P < 0.05$). Two loci significantly deviated from neutral equilibrium when analysing all individuals (C11: $D = 2.64$, $P < 0.05$; C12: $D = 2.44$, $P < 0.05$). However, the results of the Bayesian clustering changed little after excluding these loci (Supplementary Figure S1).

Demographic history and introgression

Each demographic parameter was consistent among three replicated Markov chain Monte Carlo searches and had a single peak in their posterior probability distributions (Figure 2), indicating convergence of the present parameter estimations. All migration rates except for one ($m_{S>P}$) were the lowest estimates; that is, the highest peak of the

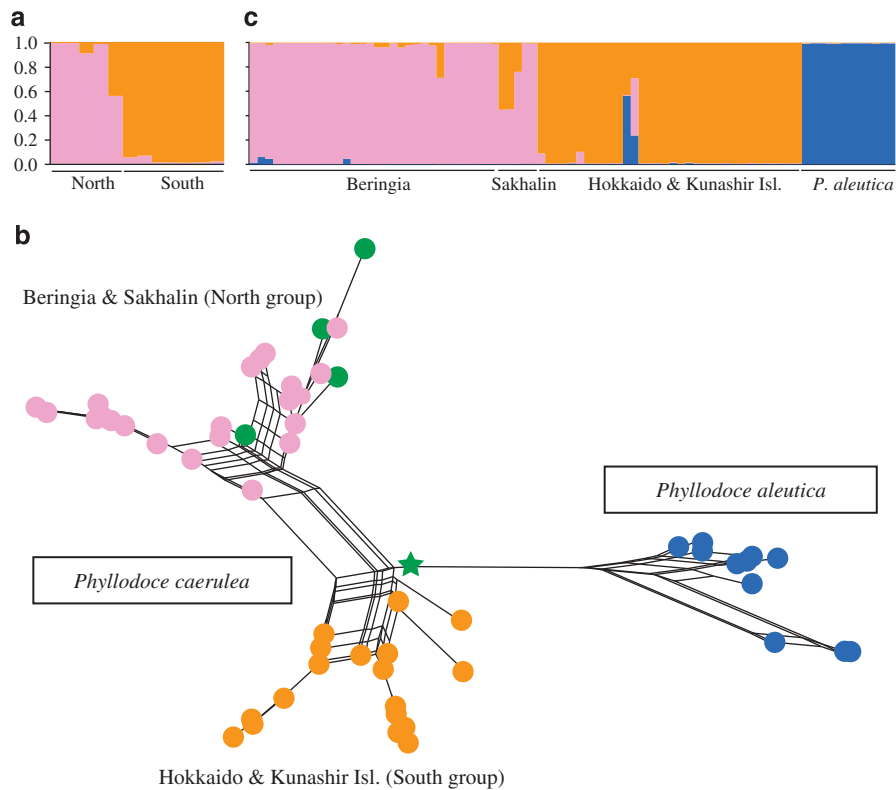


Figure 1 Geographic patterns of genetic variation. (a) The assignment probability into ancestral clusters inferred by Bayesian clustering in analysing representative individuals of *P. caerulea* based on 13 loci. (b) The relationships of individuals as inferred by a neighbour-net based on concatenated sequences from six loci. Green circles and star represent individuals of *P. caerulea* in Sakhalin and a putative hybrid individual in Daisetsu. (c) The assignment probability into ancestral clusters inferred by Bayesian clustering in analysing 71 individuals including *P. aleutica* based on 6 loci, respectively. The geographic regions are shown below the bars.

posterior probability distributions was at $m=0$, indicating that no gene flow occurred between the pair of populations. In contrast, the migration rate forward in time from *P. aleutica* to the southern group of *P. caerulea* was significantly larger than zero ($m_{S>Pa}=0.151$; $2LLR=12.35$, d.f. = 1, $P<0.001$).

The maximum likelihood estimates and their 95% HPDs of scaled divergence time between northern and southern populations ranged between 22 and 33 ka and between 6 and 87 ka (Table 2), respectively. The maximum likelihood estimates are a little earlier than the last glacial maximum (ca. 20 ka) (Gibbard and Van Kolfshoten, 2004) and the 95% HPD mostly corresponds to the last glacial period. The divergence time of *P. caerulea* and *P. aleutica* is ca. 210–310 ka with 95% HPD (82–600 ka). This estimation is older than a previous report (110–161 ka) (Ikeda *et al.*, 2014b), which may be caused by an inclusion of additional locus (C35) in the present study, whereas the present 95% HPD is in the range of that reported previously (35–630 ka).

Climate difference across populations of *P. caerulea* in East Asia and Beringia

The climate conditions for localities of *P. caerulea* differ between Hokkaido (hereafter including Kunashir Island) and Beringia plus Sakhalin along the first axial of principal component analysis (axis 1: 48.8%; Figure 3). All variables relevant to precipitation except for precipitation seasonality are highly correlated with this axis ($r>0.9$; Table 3). As represented by annual precipitation (Figure 3), *P. caerulea* grow in wetter conditions in Hokkaido (mean (s.d.) = 1608.0 (179.1) mm) than in Beringia (531.3 (179.1) mm) as well as Sakhalin.

DISCUSSION

This study demonstrated that *P. caerulea* is genetically structured into a northern and southern group in East Asia (Figure 1). The present demographic history suggests that introgression from *P. aleutica* is involved in the north–south divergence (Figure 2 and Table 2). These results also suggest that prolonged isolation predating the last glacial period is not responsible for the divergence (Table 2; T_{0_low} and T_{0_high}). Notably, we revealed that without estimating the evolutionary history underlying a genetic structure, the genetic divergence could have suggested an alternative history where *P. caerulea* persisted in isolation in Hokkaido. Overall, our study not only provides an insight into the population dynamics of alpine populations of arctic–alpine plants in East Asia but also emphasizes the importance of careful interpretation of genetic structure in phylogeographic inferences.

Biogeographic history of *P. caerulea* in East Asia

The present north–south differentiation (Figure 1) is a genetic structure that is typical of those resulting from a prolonged isolation of populations, and this might suggest persistence of alpine populations of *P. caerulea* in Hokkaido with isolation from Beringia after their establishment. However, this scenario is in contradiction with the divergence time between northern and southern groups of *P. caerulea* (Table 2; T_{0_low} and T_{0_high}). Although we carefully applied the reported substitution rates (Ikeda *et al.*, 2014b), it should be noted that the rates are not specific to the focal species, and may fail to obtain an accurate time of the divergence. Under the assumption that the substitution rate we used is correct, prolonged range separation between Hokkaido and Beringia throughout several cycles of Pleistocene warm

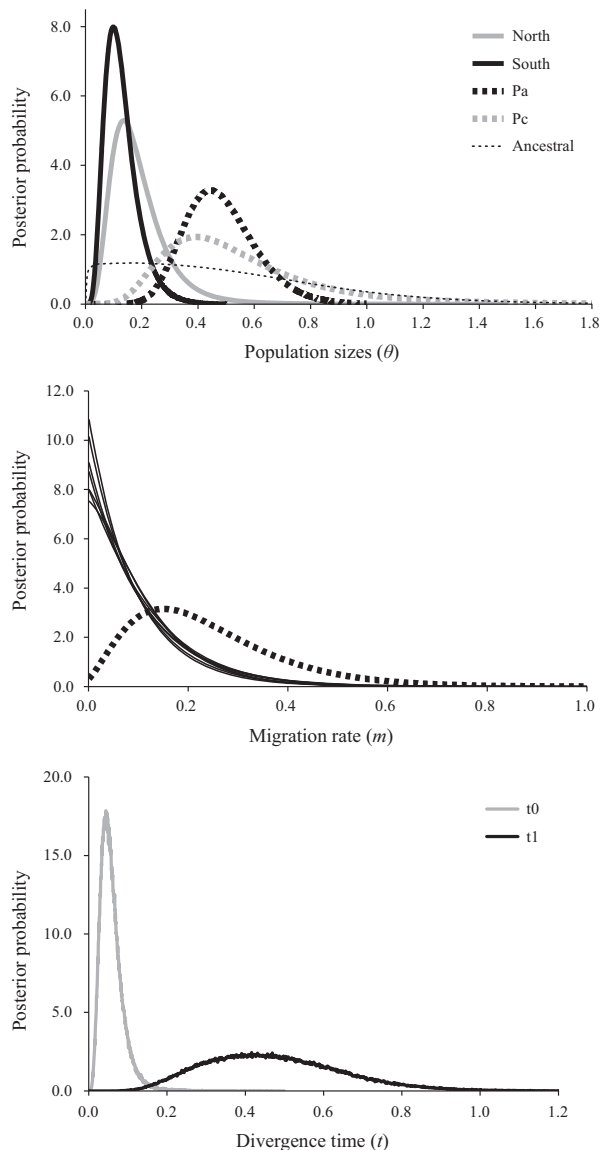


Figure 2 Posterior distributions of demographic parameters of an isolation with migration model regarding the divergence of north–south lineages of *P. caerulea* as well as *P. aleutica*. The dashed line indicates the migration rate forward in time from *P. aleutica* to the southern group of *P. caerulea*.

and cold periods is unlikely to explain the genetic divergence of *P. caerulea*.

In addition to the timing of divergence, the demographic analysis provided statistical evidence that introgression has occurred between southern populations of *P. caerulea* and *P. aleutica* (Figure 2 and Table 2). Given that *P. aleutica* is apparently distinguishable from *P. caerulea* (Figure 1), the genetic uniqueness of *P. caerulea* in Hokkaido could be attributable to the introgression. As hybrids of *P. aleutica* and *P. caerulea* are frequently observed in their sympatric populations, on-going hybridization and subsequent backcrosses might be responsible for the introgression. However, a previous study found an extremely low number of backcrossed individuals (0.0–2.5%) at three hybrid zones in Hokkaido (Kameyama and Kudo, 2011). In addition, although the clustering analysis sufficiently identified their putative hybrids (Figure 1c), all individuals of *P. caerulea* except for the putative hybrids exhibited no genetic

Table 2 Demographic parameters estimated by an isolation with migration (IM) model based on 12 nuclear loci excluding a putative nonneutral locus C12

Parameters	MLE	95% HPD
<i>Migration rate^a</i>		
$m_{N>S}$	0.001	(0.000–0.323)
$m_{S>N}$	0.001	(0.000–0.335)
$m_{N>Pa}$	0.001	(0.000–0.271)
$m_{Pa>N}$	0.001	(0.000–0.287)
$m_{S>Pa}$	0.151	(0.005–0.523)
$m_{Pa>S}$	0.001	(0.000–0.293)
$m_{AO>Pa}$	0.001	(0.000–0.327)
$m_{Pa>AO}$	0.001	(0.000–0.289)
<i>Population size^a</i>		
θ_N	0.136	(0.037–0.392)
θ_S	0.099	(0.032–0.249)
θ_{Pa}	0.444	(0.249–0.738)
θ_{AO}	0.393	(0.133–1.133)
θ_{A1}	0.174	(0.000–1.244)
<i>Divergence time</i>		
t_0	0.044	(0.012–0.116)
t_1	0.416	(0.163–0.806)
<i>Scaled divergence time; ka</i>		
T_{0_low}	22.2	(6.2–58.7)
T_{0_high}	32.6	(9.1–86.3)
T_{1_low}	210.9	(82.5–408.5)
T_{1_high}	310.3	(121.3–601.0)

Abbreviations: AO, ancestral population of *P. caerulea*; A1, ancestral population of *P. caerulea* and *P. aleutica*; HPD, highest probability density; MLE, maximum likelihood estimate; N and S, northern and southern groups of *Phyllodoce caerulea*; Pa, *P. aleutica*; t_0 and t_1 , divergence time of northern and southern groups of *P. caerulea* and of *P. caerulea* and *P. aleutica*, respectively; T_{x_low} and T_{x_high} , scaled divergence times based on faster (7.8×10^{-9} substitutions/site/year) and slower (5.3×10^{-9} substitutions/site/year) substitutions rates, respectively (ka; thousand years). The bold value indicates significant larger than zero.

^a $m_{x>y}$ and θ_x , migration rate backward in time from population x to population y and population size of population x.

admixture with *P. aleutica*. Furthermore, *P. aleutica* is absent from two populations of *P. caerulea* in Hokkaido (6-Ris and 7-Yot), whereas individuals in these populations are genetically indistinguishable from those in populations with sympatric occurrence of *P. aleutica* (Figure 1). Therefore, on-going introgression does not sufficiently explain the genetic structure. Given that both *P. caerulea* and *P. aleutica* had the opportunity to spread more widely in East Asia during the last glacial period, their more extensive sympatric occurrence at that time plausibly allowed frequent genetic admixture. Taken together, introgression before the range expansion during the last glacial period rather than extant one could be involved in the genetic structure.

Although introgression with *P. aleutica* has contributed to the genetic divergence between northern and southern populations of *P. caerulea*, the genetic structure would also be shaped by genetic drift following the geographic separation between Hokkaido and Beringia after the last glacial maximum. Our results are not sufficient to distinguish which demographic scenario caused the genetic structure. In addition to the demographic processes, the climate difference between Hokkaido and Beringia (Figure 3) offers another possibility to shape the genetic structure; that is, the genetic divergence might be attributed to natural selection from environmental differences, where distinct alleles with adaptive advantage are fixed in each geographic

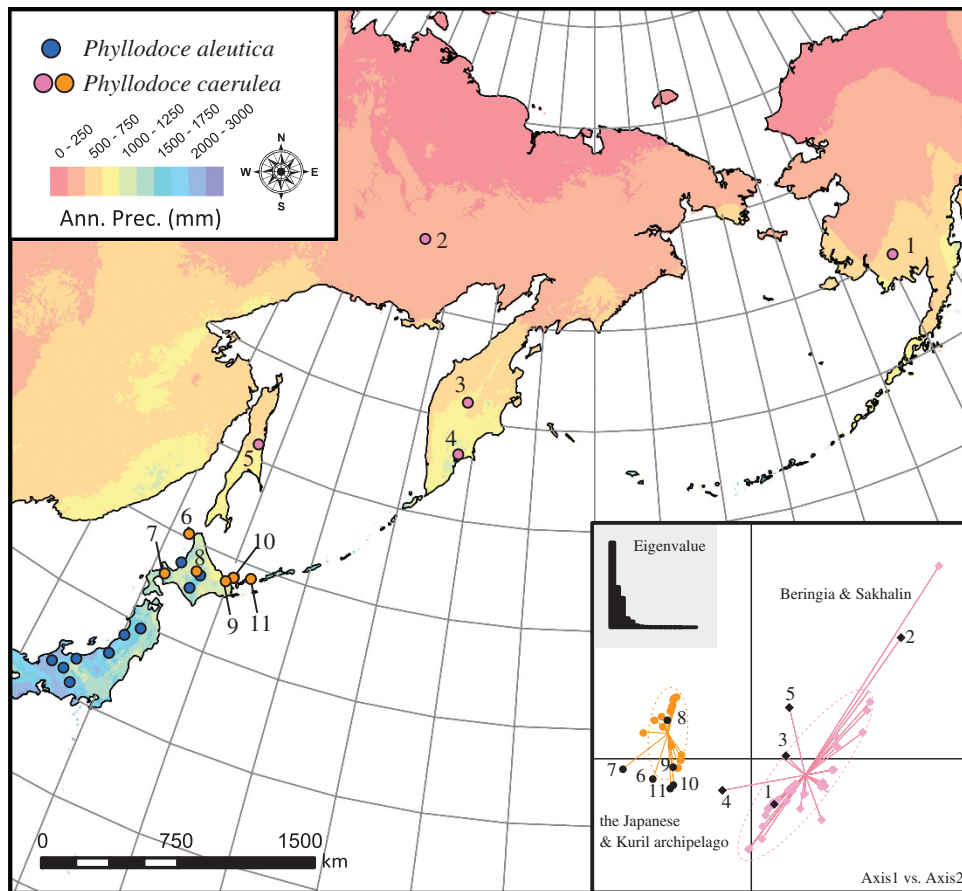


Figure 3 Climate differences within the range of *P. caerulea* in East Asia. The map shows populations used in the present study (coloured circles) together with annual precipitation (Ann. Prec. mm) in East Asia based on Bioclim. Pink, orange and blue colours in circles indicate the northern and southern groups of *P. caerulea* and *P. aleutica*, respectively. The number adjacent to the circles indicates the population numbers. The inserted box shows a result of principle component analysis based on 19 climate variables of Bioclim. The pink and orange plots represent localities of Beringia as well as Sakhalin and Hokkaido, respectively. The black symbols with numbers represent the populations used in the DNA analysis.

region. Consistent with the pattern that natural selection usually leaves a footprint on only a part of the genome, only two of six loci exhibited deviation from neutral evolution. However, excluding these loci had little influence on the genetic structure (Supplementary Figure S1), and therefore nonneutral evolution is not necessary to explain the genetic structure.

As a consequence, despite the genetic divergence between Beringia and Hokkaido, our study indicates that alpine populations of *P. caerulea* in Hokkaido plausibly connected with Beringia following the range expansion during the last glacial period. This scenario is congruent with previous inferences made on other tundra shrub species (*Cassiope lycopodioides*) (Ikeda *et al.*, 2014a) and on an arctic–alpine perennial herb (*Arabidopsis kamchatka*) (Shimizu-Inatsugi *et al.*, 2009). However, this scenario is not congruent with an inference on another snow-bed herb, *P. cuneifolia*, in which a uniquely diverged haplotype implies persistence of isolated populations in Hokkaido (Fujii *et al.*, 1999). Although further comparative phylogeography is necessary for a broad conclusion on the biogeographic relationship between Beringia and Hokkaido, this study provides insight into the population dynamics of alpine populations of arctic–alpine plants in East Asia. Notably, without estimating the evolutionary history underlying the genetic structure, the observed pattern of genetic divergence would support the alternative scenario of a history of prolonged isolation of the populations in Hokkaido. Thus, our study

emphasizes the importance of careful consideration of the influence of evolutionary history on the genetic structure for phylogeographic inference.

Impacts of introgression on the evolution of *P. caerulea* in East Asia

The lineage of *P. caerulea* in which introgression from *P. aleutica* has occurred grows in wetter climates than the lineage having no footprint of introgression (Figure 3). Although global climate data sets, as used in this study, imperfectly represent the habitats of alpine plants (Körner, 2003), the difference in precipitation may be relevant to the ecological feature of the snow-bed inhabiting *P. caerulea*, where the amount of snowfall has crucial impacts on its phenology (Kasagi and Kudo, 2003). Indeed, *P. caerulea* exclusively occurs in drier interior part of Alaska, whereas *P. aleutica* occurs in wetter coastal areas around the Pacific Ocean (<http://arctos.database.museum/>). According to this difference in geographic distributions, *P. aleutica* might favour a wetter climate than *P. caerulea*. As introgression of adaptive traits has long been suggested to be an important factor in plant evolution (Anderson, 1949), some traits with an adaptive advantage in a wetter climate might be transferred from *P. aleutica* to *P. caerulea*. Future investigations could provide insights into the evolutionary forces that have maintained the genetic structure resulting from introgression.

Table 3 Loadings of environmental variables on the first three principal components for *Phyllodoce caerulea* in East Asia and Beringia

Environmental variables	Axis 1	Axis 2	Axis 3
% Variation	48.8	23.2	17.0
Annual mean temp. (BIO1)	-0.79	-0.53	0.23
Mean temp. warmest quarter (BIO10)	-0.56	-0.12	0.72
Max temp. warmest month (BIO5)	-0.37	0.07	0.89
Mean temp. coldest quarter (BIO11)	-0.71	-0.70	0.02
Min temp. coldest month (BIO6)	-0.67	-0.74	0.00
Mean diurnal temp. range (BIO2)	0.32	0.23	0.76
Temp. annual range (BIO7)	0.49	0.76	0.40
Temp. seasonality (BIO4)	0.58	0.73	0.31
Mean temp. wettest quarter (BIO8)	-0.52	0.16	0.75
Mean temp. driest quarter (BIO9)	-0.06	-0.85	0.11
Isothermality (BIO3)	-0.10	-0.52	0.52
Annual prec. (BIO12)	-0.92	0.35	-0.10
Prec. wettest month (BIO13)	-0.92	0.26	-0.04
Prec. driest month (BIO14)	-0.91	0.36	-0.10
Prec. seasonality (BIO15)	0.63	-0.33	0.35
Prec. wettest quarter (BIO16)	-0.92	0.32	-0.07
Prec. driest quarter (BIO17)	-0.92	0.35	-0.09
Prec. warmest quarter (BIO18)	-0.90	0.37	-0.06
Prec. coldest quarter (BIO19)	-0.91	0.32	-0.12

Abbreviations: prec., precipitation; temp., temperature.

DATA ARCHIVING

All newly determined sequence data have been deposited in GeneBank and their accession numbers are cited in Supplementary Table S1.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank K Marr for collecting samples in Alaska and Canada and for editing the English of the draft manuscript, N Hashimoto and T Wakabayashi for supporting DNA experiments, anonymous reviewers for helpful comments, the October Hill Foundation for providing funds to collect samples from Alaska, the National Museum of Nature and Science for funding to HI and a Grant-in-Aid for Scientific Research for funding (KAKENHI) to HI (23657015) and HS (22405013), respectively.

Abbott RJ, Smith LC, Milne RI, Crawford RMM, Wolff K, Balfour J. (2000). Molecular analysis of plant migration and refugia in the Arctic. *Science* **289**: 1343–1346.

Alsos IG, Eidesen PB, Ehrich D, Skrede I, Westergaard K, Jacobsen GH *et al.* (2007). Frequent long-distance plant colonization in the changing Arctic. *Science* **316**: 1606–1609.

Anderson E (1949). *Introgressive Hybridization*. John Wiley & Sons: New York.

Arbogast BS, Edwards SV, Wakeley J, Beerli P, Slowinski JB (2002). Estimating divergence times from molecular data on phylogenetic and population genetic timescales. *Annu Rev Ecol Syst* **33**: 707–740.

Avise JC (2004). *Molecular Markers, Natural History and Evolution*. Sinauer Associates Inc.: Sunderland, MA, USA.

Bryant D, Moulton V (2004). Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Mol Biol Evol* **21**: 255–265.

Chessel D, Dufour A-B, Hioulose J (2004). The ade4 package-I-One-table methods. *R News* **4**: 5–10.

Evanno G, Regnaut S, Goudet J (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* **14**: 2611–2620.

Excoffier L, Laval G, Schneider S (2005). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online* **1**: 47–50.

Falush D, Stephens M, Pritchard JK (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**: 1567–1587.

Frenzel B, Pécsi M, Velichko AA, Starkel L (1992). *Atlas of Paleoclimates and Paleoenvironments of the Northern Hemisphere: Late Pleistocene-Holocene*. Hungarian Academy of Sciences: Budapest.

Fujii N, Ueda K, Watano Y, Shimizu T (1997). Intraspecific sequence variation of chloroplast DNA in *Pedicularis chamissonis* Steven (Scrophulariaceae) and geographic structuring of the Japanese 'Alpine' plants. *J Plant Res* **110**: 195–207.

Fujii N, Ueda K, Watano Y, Shimizu T (1999). Further analysis of intraspecific sequence variation of chloroplast DNA in *Primula cuneifolia* Ledeb. (Primulaceae): implications for biogeography of the Japanese alpine flora. *J Plant Res* **112**: 87–95.

Gibbard P, Van Kolfschoten T (2004). The Pleistocene and Holocene epochs. In: Gradstein FM, Ogg JG, Smith AG (eds). *A Geologic Time Scale*. Cambridge University Press: Cambridge. pp 441–452.

Hey J (2010). Isolation with migration models for more than two populations. *Mol Biol Evol* **27**: 905–920.

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

Hultén E (1937). *Outline of the History of Arctic and Boreal Biota during the Quaternary Period: Their Evolution during and after the Glacial Period as indicated by the Equiformal Progressive Areas of Present Plant Species*. Thule: Stockholm, Sweden.

Hultén E, Fries M (1986). *Atlas of North European Vascular Plants*. Koeltz Scientific Books: Königstein.

Ikeda H, Higashi H, Yakubov V, Barkalov V, Setoguchi H (2014a). Phylogeographical study of the alpine plant *Cassiope lycopodioides* (Ericaceae) suggests a range connection between the Japanese archipelago and Beringia during the Pleistocene. *Biol J Linn Soc* **113**: 497–509.

Ikeda H, Senni K, Fujii N, Setoguchi H (2009). High mountains of the Japanese archipelago as refugia for arctic-alpine plants: phylogeography of *Loiseleuria procumbens* (L.) Desvoux (Ericaceae). *Biol J Linn Soc* **97**: 403–412.

Ikeda H, Setoguchi H (2010). Natural selection on *PHYE* by latitude in the Japanese archipelago: insight from locus specific phylogeographic structure in *Arctostaphylos nana* (Ericaceae). *Mol Ecol* **19**: 2779–2791.

Ikeda H, Setoguchi H (2013). A multilocus sequencing approach reveals the cryptic phylogeographical history of *Phyllodoce nipponica* Makino (Ericaceae). *Biol J Linn Soc* **110**: 214–226.

Ikeda H, Yakubov V, Barkalov V, Setoguchi H (2014b). Molecular evidence for ancient relicts of arctic-alpine plants in East Asia. *New Phytol* **203**: 980–988.

Jakobsson M, Rosenberg NA (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**: 1801–1806.

Kameyama Y, Kudo G (2011). Clarification of the genetic component of hybrids between *Phyllodoce caerulea* and *Phyllodoce aleutica* (Ericaceae) in Hokkaido, northern Japan. *Plant Species Biol* **26**: 93–98.

Kasagi T, Kudo G (2003). Variations in bumble bee preference and pollen limitation among neighboring populations: comparisons between *Phyllodoce caerulea* and *Phyllodoce aleutica* (Ericaceae) along snowmelt gradients. *Am J Bot* **90**: 1321–1327.

Knowles LL (2009). Statistical phylogeography. *Annu Rev Ecol Syst* **40**: 593–612.

Körner C (2003). *Alpine Plant Life*, 2nd edn. Springer: Heidelberg, Germany.

Librado P, Rozas J (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.

Nei M (1987). *Molecular Evolutionary Genetics*. Columbia University Press: New York.

Pelster PB, Abbott RJ, Comes HP, Milton JJ, Möller M, Looseley ME *et al.* (2012). The genetic ghost of an invasion past: colonization and extinction revealed by historical hybridization in *Senecio*. *Mol Ecol* **21**: 369–387.

Schönswetter P, Paun O, Tribsch A, Niklfeld H (2003). Out of the Alps: colonization of Northern Europe by East Alpine populations of the Glacier Buttercup *Ranunculus glacialis* L. (Ranunculaceae). *Mol Ecol* **12**: 3373–3381.

Schönswetter P, Popp M, Brochmann C (2006). Rare arctic-alpine plants of the European Alps have different immigration histories: the snow bed species *Minuartia biflora* and *Ranunculus pygmaeus*. *Mol Ecol* **15**: 709–720.

Shimizu-Inatsugi R, Lihová J, Iwanaga H, Kudoh H, Marhold K, Savolainen O *et al.* (2009). The allopolyploid *Arabidopsis kamchatica* originated from multiple individuals of *Arabidopsis lyrata* and *Arabidopsis halleri*. *Mol Ecol* **18**: 4024–4048.

Stephens M, Donnelly P (2003). A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet* **73**: 1162–1169.

Taberlet P, Fumagalli L, Wust-Saucy AG, Cosson JF (1998). Comparative phylogeography and postglacial colonization routes in Europe. *Mol Ecol* **7**: 453–464.

Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.

Wegmann D, Leuenberger C, Neuenschwander S, Excoffier L (2010). ABCtoolbox: a versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics* **11**: 116.

Winkler M, Tribsch A, Schneeweiss GM, Brodbeck S, Gugerli F, Holderegger R *et al.* (2012). Tales of the unexpected: phylogeography of the arctic-alpine model plant *Saxifraga oppositifolia* (Saxifragaceae) revisited. *Mol Ecol* **21**: 4618–4630.

Yamazaki T (1993). Ericaceae 5. *Phyllodoce*. In: Iwatsuki K, Yamazaki T, Boufford DE, Ohba H (eds). *Flora of Japan IIIa*. Kodansha: Tokyo. p 9.

Supplementary Information accompanies this paper on Heredity website (<http://www.nature.com/hdy>)