Genetic divergence of the sub-alpine shrubby variety, Quercus crispula var. horikawae, from the mountain oak species, Q. crispula, in Japan

Lerma San Jose-Maldia[1], Asako Matsumoto[2], Teruyoshi Nagamitsu[2], Saneyosshi Ueno[2], and Yoshihiko Tsumura[3]

[1]University of the Philippines Los Baños
[2]Forest Research and Management Organisation
[3]University of Tsukuba

September 4, 2023

Abstract

Ecotypic divergence in tree taxa often occurs in sub-alpine habitats, where environmental conditions are more stressful than those in lower elevations. In the mountain oak species in Japan, Quercus crispula (Qc), the sub-alpine shrubby variety, Q. crispula var. horikawae (Qch), has been recognized in central and northern Honshu. Although Qch has different phenotypes from Qc, genetic divergence between Qc and Qch has not been examined yet. Pairs of Qc and Qch populations in eight locations and additional Qc and Qch populations around these locations were investigated. Leaf size of Qch was smaller than that of Qc. Chloroplast DNA haplotypes were shared between the Qc and Qch populations. In genotypes at 29 nuclear microsatellite loci, genetic diversity did not differ between the Qc and Qch populations. Principal component analysis and a neighbor-joining tree of populations based on microsatellite genotypes demonstrated that 13 Qc populations and eight Qch populations were grouped separately, except for three Qch populations that were grouped to Qc. Climatic conditions in the eight Qch populations were characterized by lower temperature and heavier snowfall than those in the 16 populations of the genetic group of Qc. These results suggest the genetic divergence between Qc and Qch associated with sub-alpine climatic conditions, irrespective of leaf size. The origin of the sub-alpine Qch lineage and the history of ecotypic divergence should be investigated in future genomic studies.

Genetic divergence of the sub-alpine shrubby variety, *Quercus crispula* var. *horikawae* , from the mountain oak species, *Q. crispula* , in Japan

Lerma San Jose-Maldia[1,2], Asako Matsumoto[1], Teruyoshi Nagamitsu[1], Saneyoshi Ueno[1], Yoshihiko Tsumura[3]

[1]Forestry and Forest Products Research Institute, Forest Research and Management Organization, Tsukuba, Ibaraki 305-8687, Japan

[2]College of Forestry and Natural Resources, University of the Philippines Los Baños, College 4031, Laguna, Philippines

[3]Faculaty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-8572, Japan

Correspondence

nagamit@affrc.go.jp (T. Nagamitsu: TN) Orcid ID: 0000-0003-3366-7197

Tel.: +81-29-829-8262; Fax: +81-29-874-3720

Contacts

lsmaldia@up.edu.ph (L. San Jose-Maldia: LSJ-M) Orcid ID: 0000-0002-9237-635X

asakon@affrc.go.jp (A. Matsumoto: AM) Orcid ID: 0000-0002-3019-2985

saueno@affrc.go.jp (S. Ueno: SU) Orcid ID: 0000-0001-5571-0622

tsumura.yoshihiko.ke@u.tsukuba.ac.jp (Y. Tsumura: YT) Orcid ID: 0000-0002-9844-0914

Abstract

Ecotypic divergence in tree taxa often occurs in sub-alpine habitats, where environmental conditions are more stressful than those in lower elevations. In the mountain oak species in Japan, *Quercus crispula* (*Qc*), the sub-alpine shrubby variety, *Q. crispula* var. *horikawae* (*Qch*), has been recognized in central and northern Honshu. Although *Qch* has different phenotypes from *Qc* , genetic divergence between *Qc* and*Qch* has not been examined yet. Pairs of *Qc* and *Qch*populations in eight locations and additional *Qc* and *Qch*populations around these locations were investigated. Leaf size of*Qch* was smaller than that of *Qc* . Chloroplast DNA haplotypes were shared between the *Qc* and *Qch* populations. In genotypes at 29 nuclear microsatellite loci, genetic diversity did not differ between the *Qc* and *Qch* populations. Principal component analysis and a neighbor-joining tree of populations based on microsatellite genotypes demonstrated that 13 *Qc* populations and eight *Qch* populations were grouped separately, except for three*Qch* populations that were grouped to *Qc* . Climatic conditions in the eight *Qch* populations were characterized by lower temperature and heavier snowfall than those in the 16 populations of the genetic group of *Qc* . These results suggest the genetic divergence between *Qc* and *Qch* associated with sub-alpine climatic conditions, irrespective of leaf size. The origin of the sub-alpine *Qch* lineage and the history of ecotypic divergence should be investigated in future genomic studies.

Keywords

chloroplast DNA haplotype, climatic conditions, ecotypic divergence, expressed sequence tagged simple sequence repeat (EST-SSR), genetic structure, leaf characters

1 | Introduction

Ecotypic divergence in plants often occurs in alpine (including sub-alpine) habitats, where environmental conditions are more stressful than those in lower elevations (Konečná et al. 2019). Plants in alpine habitats suffer from stresses including freezing and snowfall, low and fluctuating temperature, strong wind, drought, low nutrient, and high UV radiation. Local adaptation to these stressful environments often leads to diversification in morphological and physiological traits, which results in divergence of alpine ecotypes (Wos et al. 2022). In cases that reproductive barriers arise between ecotypes in alpine and mountain zones due to not only temporal and spatial isolations in mating but also natural selection from local environments, genetic divergence occurs between ecotypes and proceeds toward speciation (Hirao et al. 2019). Because alpine habitats are geographically separated between different mountain ranges, two evolutionary processes can be applied to genetic divergence of an alpine ecotype across multiple mountain ranges (Holliday et al. 2016). One is parallel evolution of the alpine ecotype that occurred independently in different mountain ranges, where genetically-differentiated populations of the mountain ecotype are distributed (Trucchi et al. 2017; Szukala et al. 2023) (Figure S1a in Supporting Information). The other is stepping-stone colonization of different mountain summits by a single lineage of the alpine ecotype that had already diverged from the mountain ecotype (Bettin et al. 2007) (Figure S1b). Recent studies have demonstrated that the two evolutionary processes are responsible for the genetic divergence of alpine ecotypes in various plant taxa (Knotek et al. 2020; Bohutínská et al. 2021).

The Japanese Archipelago has mountain chains with alpine zones ranging north and south along the islands (Ohsawa and Ide 2011). Most of alpine plants in Japan are characterized by genetic divergence between central Honshu and northern Japan, the latter of which includes northern Honshu and/or Hokkaido (Fujii and Senni 2006). This genetic divergence are thought to result from multiple colonization of the Japanese

Archipelago by arctic plants during glacial cycles in the Pleistocene (Ikeda 2022). These plants migrated southward and colonized central Honshu in glacial periods. In subsequent post-glacial periods, the colonized populations were isolated in alpine zones of central Honshu and diverged from the populations that retreated northward. On the other hand, temperate plants in mountain zones in the Japanese Archipelago show various phylogeographic patterns (Ohsawa and Ide 2011). In temperate trees in mountain zones, populations in southern Japan tend to have higher genetic diversity, which reflects multiple refugia in glacial periods, than populations in northern Japan with relatively homogeneous genetic structure, which reflects expansion from the refugia in post-glacial periods (Tomaru et al. 2022). As a result, latitudinal genetic divergence is often found in mountain trees, while genetic boarders of the divergence are located in various places. In addition to the latitudinal divergence, genetic divergence between northeastern and southwestern coastal sides of the Japanese Archipelago sometimes occurs due to contrasting climate conditions on the opposite sides of the islands (Tsumura 2006; Tsumura 2022). Thus, plants in both alpine and mountain zones often show genetic differentiation within species among mountain ranges along the Japanese Archipelago.

A white oak (Section *Quercus*) species, *Quercus crispula*Blume (*Qc*), is common in cool-temperate forests in mountain zones in the Japanese Archipelago (Figure 1a–c). This species name is a synonym of *Q. mongolica* var. *crispula* (Blume) H. Ohashi, which is used in a taxonomic system widely accepted (Ohashi 1988; Aizawa et al. 2021). In this species, a sub-alpine variety, *Q. crispula*var. *horikawae* H. Ohba (*Qch*), is recognized (Ohba 1989) (Figure 1d–f). Because there is no combination of taxonomic naming under *Q. mongolica* Fisch. ex Ledeb., we follow nomenclature of Ohba (1989) in this study. *Qch* is usually found in sub-alpine zones or steep mountain slopes with heavy snowfall and is characterized by shrubby habit, bent trunk often decumbent near the ground, small leaf size, and dense hairs on the abaxial leaf surface (Ohba 2006). Because these phenotypes of *Qch* are discontinuous from those of *Qc*in a mountain range (Mt. Makihata, 37.0@N, 139.0@E, 1967 m) and are likely to be adaptive to sub-alpine environments (Noshiro 1984),*Qc* and *Qch* are regarded as different ecotypes. Genetic variation in nuclear-encoded allozymes in two *Qc* and two*Qch* populations did not indicate clear divergence between*Qc* and *Qch* (Tanimoto et al. 1992).

Genetic structure of *Qc* has been investigated using different genetic markers. In chloroplast DNA haplotypes, higher haplotype diversity was found in southern Japan, while a few haplotypes were dominated in northern Japan (Kanno et al. 2004; Okaura et al. 2007; Liu and Harada 2014). The southern boarders of the northern haplotypes were located in central Honshu (Liu and Harada 2014; Onosato et al. 2021). In genotypes of nuclear microsatellite (simple sequence repeat, SSR), higher genetic diversity in southern populations and gradual genetic divergence between northern and southern populations were found, supporting post-glacial northward colonization from southern refugia (Ohsawa et al. 2011). In single nucleotide polymorphism (SNP) in some nuclear genes, however, northern populations harbored high nucleotide diversity and fast decay of linkage equilibrium, which were comparable to southern populations (Quang et al. 2008). These findings suggest that genetic variation was maintained in some genes during northward colonization. Therefore, the ecotypic divergence of *Qch* should be considered in the demographic history of post-glacial northward colonization of *Qc* .

Here, we proposed three hypotheses: (1) parallel evolution of ecotypic divergence between *Qc* and *Qch* that occurred independently in genetically-differentiated populations in different mountain ranges (Figure S1a), (2) stepping-stone colonization of different mountain summits by *Qch* populations that belong to a single lineage divergent from *Qc* (Figure S1b), and (3) another case of genetic differentiation of populations of the two ecotypes in different mountain ranges (Figure S1c). To test the three hypotheses, we selected pairs of*Qc* and *Qch* populations in multiple mountain ranges along central to northern Honshu and measured climatic conditions, leaf characters, and genetic variation in chloroplast and nuclear genomes.

2 | Materials and Methods

2.1 | Sampling

We selected one *Qc* population (eight in total) and one or two*Qch* populations (10 in total) in each of eight locations as well as additional five *Qc* and one *Qch* populations around these locations (Table 1). The eight

locations were situated in mountain ranges with alpine zones in northern and central Honshu (2, 4, 6, 7, 8, 9, 10, and 11 in Figure 2). We identified the taxon, *Qc* or *Qch* , for each population based on growth habit, leaf size, and hair density on the abaxial leaf surface according to Ohba (2006). Three populations (05Hx, 06Hx, and 07Hx) were identified as *Qch* in the field observation but were grouped to *Qc* in the following genetic analysis (see Results 3.2). From each of the 24 populations, we sampled leaves of 18–32 trees (693 individuals in total) along mountain trails at intervals of at least 20 m in the summer in 2011–2012.

To measure leaf characters, we selected approximately three leaves from each of 4–8 individuals in each of 12 out of the 24 populations (155 leaves of 53 individuals in total; Table 1). For genetic analysis, we extracted DNA from leaves of all the 693 individuals in the 24 populations using DNeasy Plant Mini Kit (Qiagen, Hilden, Germany).

2.2 | Genotyping

To examine chloroplast (cp) DNA haplotypes, we selected 7–8 individuals from each of 17 out of the 24 populations (130 individuals in total; Table 1). We determined nucleotide sequences in four cpDNA regions: 3' to *rps* 2, *trn* T(UGU)–*trn* L(UAA) 5' exon, *rps* 16 intron, and *rp* L32–*trn* L(UAG) using a BigDye Terminator Sequencing Kit and a 3100 Genetic Analyzer (Thermo Fisher Scientific, Waltham, USA). Obtained sequences were assembled and manually edited using Sequencher 10.4.1 (Gene Codes Corporation, Ann Arbor, USA).

To examine nuclear (nc) microsatellite genotypes, we selected 32 expressed sequence tagged (EST) SSR loci (Ueno et al. 2008; Ueno and Tsumura 2008; Ueno et al. 2009a; Ueno et al. 2009b) (Data S1 in Supporting Information). We amplified these ncEST-SSR loci using a multiplex PCR Kit (Qiagen) and genotyped them using a 3100 Genetic Analyzer and Genotyper 3.2 software (Thermo Fisher Scientific).

Detailed methods of cpDNA sequencing and ncEST-SSR genotyping were described in our previous study (San Jose-Maldia et al. 2017).

2.3 | Genetic analyses

We classified cpDNA haplotypes solely based on nucleotide substitutions, because there were numerous gaps due to large indels in the*rp* L32–*trn* L(UAG) region. We described the frequency of haplotypes in each of the 17 populations.

To verify the quality of ncEST-SSR loci, we estimated the median frequency of null alleles at each locus using function popgenreport(mk.null.all) (Brookfield 1996) in package PopGenReport (Adamack and Gruber 2014) in R 4.2.2 (R_Core_Team 2019). We conducted chi-square tests for deviation of the Hardy-Weinberg equilibrium at each locus in each population using function popgenreport(mk.hwe). We removed loci with null alleles at substantially high (> 0.15) frequency and highly significant ($P < 0.001$) deviation in several (> 5) populations and used remaining loci in the following genetic analysis.

To evaluate genetic variation in ncEST-SSR genotypes, we conducted principal component analysis (PCA) of individual genotypes in the 24 populations and allele frequencies in individual populations using function dudi.pca in R package adegenet (Jombart 2008). We obtained the contributions of the first to fifth principal components (PCs) to total variation. We plotted the coordinates of the first and second PC values of each individual and each population to categorize them into genetic groups.

To examine genetic diversity in ncEST-SSR genotypes in each population, we calculated the allelic richness ($AR_{[32]}$, the number of alleles in the smallest sample size, 32 alleles) using function popgenreport(mk.allel.rich) in R package PopGenReport as well as the unbiased expected heterozygosity (H_E) and the inbreeding coefficient (F_{IS}) using functions Hs and Ho in R package hierfstat (Goudet 2005) in R 4.2.2. Significantly positive or negative F_{IS} in each population was verified by 2.5 and 97.5 percentiles estimated from bootstrapping over loci using function boot.ppfis. Differences inF_{IS} between the genetic groups of *Qc* and*Qch* (see Results 3.2) were verified by the Kruskal-Wallis rank-sum test using function kruskal.test in R 4.2.2. We examined latitudinal and elevational clines as well as differences between the genetic groups in

genetic diversity of populations. We applied a regression model: $y \sim \alpha + \beta_1 x_1 + \beta_2 x_2 + \gamma$, where y is AR [32] or H_E, α is an intercept, β_1 is a coefficient of x_1: latitude (@N), β_2 is a coefficient of x_2: elevation (m), and γ is a coefficient of the genetic groups, to data in the 24 populations using function lm in R 4.2.2.

To evaluate genetic structure in ncEST-SSR genotypes, we conducted model-based Bayesian clustering of individual genotypes in the 24 populations using STRUCTURE 2.3.3 (Pritchard et al. 2000). The admixture model of ancestries and correlated allele frequencies were used. Twenty replications were performed for each number of clusters (K) from 1 and 5, with 50000 sampling iterations after 10000 burn-in iterations. We evaluated increment in the log probability of data as K values increased and its variation among replications at each K value. We plotted the ancestry proportions of individuals at K values that were suitable for describing genetic structure.

To examine genetic distance between populations in ncEST-SSR genotypes, we calculated the Weir and Cockerham's pairwise F_{ST} using function pairwise.WCfst in R package hierfstat and the Nei's distance D_A using function dist.genpop in R package adegenet. We constructed a neighbor-joining (NJ) tree of the 24 populations based on the D_A matrix using function nj in R package ape (Paradis and Schliep 2019). To assess the robustness of the NJ tree, we generated multiple NJ trees based on D_A matrices from jackknifing of loci (removing one out of available loci) and calculated the frequency of generated trees at focal nodes that contained consistent populations with the NJ tree constructed from all the available loci.

2.4 | Climatic conditions and leaf characters

To evaluate climatic conditions that populations experienced, we obtained climatic value mesh data (Japan_Meteolorogical_Agency 1987) at the third meshes (about 1 km squares) of Japan Standardized Regional Mesh Codes (JIS X0410), where the 24 populations were located. Although the data have been updated (the latest version in 2020), we used the old version (1953–1983), because we evaluated climatic conditions before recent climate changes. From the data, we used the mean of daily-mean temperature (@C) in each month, the precipitation (mm) in each month, the maximum snow depth (cm) in each month from December to March (28 variables in total).

For the 155 leaves in the 12 populations, we obtained seven (a–g) measurements (mm) of leaf characters described in Noshiro (1984): a) blade length, b) blade width, c) inter-sinus blade width, d) distance to the widest point from base, e) length of the longest lateral vein, and f) intra-tooth vein length and g) tooth breadth at the end of the longest lateral vein.

To evaluate variations in climatic conditions among populations and in leaf characters among leaves, we conducted PCA using prcomp function in R 4.2.2. We obtained the contributions of the first to fifth PCs to total variation and the loadings of 28 variables in climatic conditions and seven variables in leaf characters to the first and second PCs. We plotted the coordinates of the first and second PC values of each population and these values of each leaf. To describe variation in leaf characters among populations, we calculated the median of PC values of leaves in each population and plotted these PC coordinates.

3 | Results

3.1 | Chloroplast DNA haplotypes

Among the determined cpDNA sequences, we discriminated three cpDNA haplotypes (A, B, and C) based on substitutions at four nucleotide sites in *rp* L32–*trn* L(UAG) region (Figure S2 in Supporting Information). *Qc* and *Qch* populations shared common haplotypes A and B (Table 2). Haplotype A ranged across northern (2: 40.7@N) and southern (11: 36.7@N) locations, whereas haplotype B occurred in restricted locations (4, 6, and 7: 38.8–40.1@N; Table 2, Figure 2). In location 4 (40.1@N), *Qc* population 04C had haplotype B, while *Qch* population 04H had haplotype A (Table 2). Haplotype C was found in the most southwestern location 11 (36.7@N; Table 2, Figure 2).

3.2 | Nuclear microsatellite genotypes

Among the 32 ncEST-SSR loci, we found three loci (CcC00610, FcC03095, and QmC01794), which had [?] 0.173 estimated frequency of null alleles and showed significant ($P < 0.001$) deviation from the Hardy-Weinberg equilibrium in [?] 8 populations (Data S1). Thus, we used remaining 29 out of the 32 loci in the following genetic analysis.

In PCA for allele frequencies of the 24 populations, the first and second PCs (PC1 and PC2) contributed to 22.5% and 17.0%, respectively, of variation among populations. The coordinates of PC1 and PC2 indicated that 13 *Qc* populations and eight *Qch* populations were grouped separately, except for three populations 05Hx, 06Hx, and 07Hx (Figure 3b). The three populations were identified as *Qch* in the field observation but were grouped to *Qc* in the ncEST-SSR variation. However, the separation of the genetic group of *Qc* (13*Qc* populations and three *Qch* populations 05Hx, 06Hx, and 07Hx) and the genetic group of *Qch* (eight *Qch* populations) was continuous, and three northern and southern marginal *Qch*populations 02H1, 02H2, and 11H, one *Qc* population 11C, and one*Qch* population 06Hx were located at intermediate positions in the PC coordinates (Figures 2, 3b). PCA of individual genotypes indicated that individuals of the genetic group of *Qch* tended to have higher values of PC1, which contributed to only 3.8% of variation among individuals, than individuals of the genetic group of *Qc* (Figure 3a). The PC1 values of individuals were overlapped between the genetic groups (Figure 3a).

The allelic richness (AR [32]) in the 24 populations was lower in higher latitudes ($P = 0.011$), was not related to elevation ($P = 0.172$), and was not different between the genetic groups of *Qc* and *Qch* ($P = 0.449$; Table 2, Figure S3a, b). The expected heterozygosity (H_E) was neither dependent on latitude nor elevation (P [?] 0.341) and was not different between the genetic groups of *Qc* and *Qch* ($P = 0.416$; Table 2, Figure S3c, d). The inbreeding coefficient (F_{IS}) was not significantly ($P < 0.05$) positive, except for three populations 06C, 08C, and 06H (0.054 [?] F_{IS} [?] 0.066), and were not different between the genetic groups of *Qc*and *Qch* ($P = 0.259$; Table 2).

The log probability of data (LnPD) in Bayesian clustering of individual genotypes showed a substantial increase from $K = 1$ to $K = 2$ and additional increases to $K = 3$ and to $K = 4$ (Figure S4). At $K = 5$, LnPD was unstable (Figure S4). At $K = 2$, two clusters were likely to represent the genetic groups of *Qc*and *Qch* (Figure 4a). Intermediate ancestry proportions of both clusters were frequently found, especially in some populations (for example, 11C and 11H; Figure 4a). At $K = 3$, the *Qc* cluster was divided into two clusters, one of which was dominant in population 05Hx (Figure 4b). At $K = 4$, the *Qch* cluster was divided into two clusters, one of which was frequent in northern populations 02H1, 02H2, and 04H of the genetic group of *Qch* (Figures 2, 4c).

The Weir and Cockerham's pairwise F_{ST} tended to be higher between the genetic groups (median F_{ST}= 0.046, 0.026 [?] F_{ST} [?] 0.107) than within the genetic group of *Qc* (median $F_{ST} = 0.019$, 0.005 [?] F_{ST} [?] 0.055) and the genetic group of*Qch* (median $F_{ST} = 0.039$, 0.015 [?]F_{ST} [?] 0.079; Data S1). A neighbor-joining (NJ) tree of populations based on the Nei's distanceD_A showed divergence between the genetic groups, although the length of a branch between the genetic groups was relatively short (Figure 5). Jackknifing over loci (removing one out of the 29 loci) generated 29 NJ trees, and seven populations of the genetic group of *Qch* , except for population 11H, were always grouped in the generated NJ trees (Figure 5).

3.3 | Climatic conditions

In PCA of climatic conditions in the 24 populations, PC1 and PC2 contributed to 52.5% and 20.4%, respectively, of variation among populations (Figure 6a). The mean of daily-mean temperature in every month increased as PC1 decreased, the precipitation from November to February and the maximum snow depth from December to March increased as both PC1 and PC2 increased, and the precipitation from April to September increased as PC2 decreased (Figure 6a).

In the coordinates of PC1 and PC2, four *Qch* populations of the genetic group of *Qch* were located in cold and snowy conditions with PC1 > 5 (Figure 6b). Other four *Qch* populations of the genetic group of *Qch* as well as four *Qc* and one*Qch* populations of the genetic group of *Qc* were located in intermediate conditions with –1 < PC1 < 3 (Figure 6b). Nine *Qc* and two *Qch* populations of the genetic group of *Qc* were located in

warm conditions with PC1 < –1 (Figure 6b). The southernmost *Qc* populations 13C and 14C were situated in rainy conditions in summer with PC2 < –6 (Figures 2, 6b).

3.4 | Leaf characters

In PCA of leaf characters, PC1 contributed to 81.4% of variation among leaves (Figure 7a). Because all the measurements of leaf parts increased as PC1 decreased, PC1 indicated leaf size (Figure 7a).

Leaves of *Qch* populations tended to have higher PC1 values (smaller leaves) than those of *Qc* populations, although the PC1 values were overlapped among the taxa (Figure 7b). The median of PC1 values in each population tended to be higher in eight *Qch*populations, including three populations 05Hx, 06Hx, and 07Hx of the genetic group of *Qc* , and tended to be lower in four *Qc*populations (Figure 7c).

4 | Discussion

Our results from ncEST-SSR genotypes do not support parallel evolution of ecotypic divergence between *Qc* and *Qch* in different mountain ranges (Figure S1a) but do suggest colonization of different mountain summits by *Qch* populations belonging to a single lineage that had already diverged from *Qc* (Figure S1b). First, the NJ tree indicated that the genetic group of *Qch* (eight *Qch*populations) and the genetic group of *Qc* (13 *Qc*populations and three *Qch* populations) and were separated in the tree topology as shown in Figure S1b. Second, the PCA demonstrated that the genetic groups of *Qc* and *Qch* were located at different positions in the coordinates of PC1 and PC2, irrespective of their geographic locations. Although the genetic groups of *Qc* and*Qch* are genetically divergent, their genetic differentiation ($F_{ST} = 0.046$) was lower than that between Japanese white oak species, *Qc* and *Q. dentata* Thunberg ($F_{ST} = 0.133$), *Qc* and *Q. serrata*Murray ($F_{ST} = 0.153$), and *Q. dentata* and*Q. serrata* ($F_{ST} = 0.211$) using ncEST-SSR loci (Nagamitsu et al. 2019). This low genetic differentiation between the genetic groups of *Qc* and *Qch* suggests recent divergence and/or gene flow between them, which is consistent with intermediate ancestry proportions of the *Qc* and *Qch*clusters frequently found in individuals of both genetic groups. Thus, the taxonomic treatment of *Qch* as the variety of species*Qc* seems appropriate (Ohba 1989), and the genetic groups of*Qc* and *Qch* can be recognized as ecotypes (Lowry 2012).

The genetic diversity in the genetic groups of *Qc* and *Qch*did not differ. This result suggests that the *Qch* ecotype has maintained its genetic variation in spite of restricted areas of its habitats in sub-alpine zones. The latitudinal cline in allelic richness implies northward colonization from southern refugia after the last glacial period. This cline is common in white oak species, *Q. aliena* Blume and *Q. serrata* (San Jose-Maldia et al. 2017) and*Qc* (Ohsawa et al. 2011) in Japan. The geographic distributions of cpDNA haplotypes in the *Qc* and *Qch* populations are consistent with the previous knowledge in the Japanese white oak species (Kanno et al. 2004; Okaura et al. 2007; Liu and Harada 2014; San Jose-Maldia et al. 2017; Onosato et al. 2021). These findings suggest that the *Qch* ecotype shares the post-glacial migration history through seed dispersal with the white oak species and has colonized sub-alpine zones toward northern mountain ranges. The genetic sub-structure within the genetic groups of *Qch* shown in the Bayesian clustering at $K = 4$ and the slightly higher genetic differentiation within *Qch* ($F_{ST} = 0.039$) than within *Qc* ($F_{ST} = 0.019$) may reflect stepping-stone colonization of sub-alpine zones during northward migration associated with founder effects and genetic drift.

Climatic conditions in the habitats of populations differed between the genetic groups of *Qc* and *Qch* . The climatic conditions of the genetic group of *Qch* were characterized by low temperature and heavy snow. This correspondence between climatic and genetic variations suggests that the *Qch* ecotype is isolated by temporal and spatial reproductive barriers and/or adapted to climatic environment in sub-alpine zones. In Moriyoshi (location 4), the *Qc* and*Qch* populations had different cpDNA haplotypes, suggesting seed dispersal barriers between mountain and sub-alpine zones in this mountain range. Although phenological shift of reproductive events and spatial separation in mountain topography potentially occur among elevations, pollen and seed dispersal is feasible, because wind dispersal of pollen ranges over long distances when tree populations are fragmented (Ortego et al. 2014), and seed dispersers move to search acorns among elevations (Gomez 2003; Bekku et al. 2019). Biotic and abiotic factors that differ between mountain and sub-alpine zones can affect

survival and growth of plants (Wos et al. 2022), although selective drivers responsible for local adaptation of the *Qch*ecotype are unclear. The balance between gene flow and natural selection may result in the weak genetic differentiation between the *Qc* and*Qch* ecotypes.

The three populations 05Hx, 06Hx, and 07Hx were identified as *Qch*based on phenotypes but were grouped to *Qc* based on ncEST-SSR genotypes. Leaf sizes observed in the three *Qch* populations were similar to those in five *Qch* populations of the genetic group of*Qch* and were smaller than those in four *Qc* populations of the genetic group of *Qc* . This unexpected result implies that environmental factors in the habitats of the three populations can induce the *Qch* phenotypes in spite of their genetic background of*Qc* . Because the climatic conditions in these habitats are not sub-alpine conditions as shown in the climatic PCA, specific factors, such as strong wind, drought, and low nutrient, may facilitate expression of the *Qch* phenotypes through morphological and physiological responses to these factors (Nagamitsu et al. 2019; Solé-Medina et al. 2022). In Oga (location 5) near the coast, for example, strong wind from the sea may lead to small leaves and shrubby habit. This phenotypic plasticity prevents us from defining diagnostic morphology to identify the taxon *Qch* as the sub-alpine ecotype recognized by ncEST-SSR genotypes. Common garden experiment in multiple environments is useful to clarify phenotypic plasticity and local adaptation mentioned above, which may help us to find the diagnostic morphology and to treat the taxon *Qch* properly.

We do not know the origin of the genetic group of *Qch* . The most plausible scenario is that the *Qch* ecotype derived from *Qc*in the Japanese Archipelago after the divergence between *Qc* and*Q. mongolica* , the latter of which is distributed in the continental northeastern Asia (Ohashi 1988). Reconstruction of geographic distributions in the last glacial period using species distribution modeling indicated that the past potential habitats of*Qc* existed in northern Honshu in addition to southwestern parts of the Japanese Archipelago (Onosato et al. 2021). Thus, marginal populations in this northern refugium could diverge from main populations in the southwestern refugia and colonize sub-alpine zones in mountain ranges in central and northern Honshu after the last glacial period.

In white oaks, most taxa are interfertile, and hybridization is involved in speciation and ecotypic divergence (Hipp et al. 2020). In northern Hokkaido, a coastal ecotype of *Qc* is derived from hybridization with *Q. dentata* and is treated as a hybrid taxon *Q*×*angustilepidota* Nakai (Nagamitsu et al. 2019; Nagamitsu et al. 2020). In central Honshu, *Q. mongolica* var. *mongolicoides*(H. Ohba) M. Aizawa is thought to originate from ancient hybridization between *Qc* and *Q. mongolica* , the latter of which had probably colonized the Japanese Archipelago during glacial cycles (Aizawa et al. 2018; Aizawa et al. 2021). In the continental northwestern Asia, *Q. mongolica* var. *liaotungensis*(Koidz.) Nakai (syn.: var. *undulatifolia* (H. Lev.) Kitam. & T. Hiroki), which is often treated as a separate species *Q. liaotungensis* Koidz. (syn.: *Q. wutaishanica* Mayr.), has shrubby habit (Aizawa et al. 2021). Thus, there is a possibility that *Q. liaotungensis* is involved in the origin of the *Qch* ecotype through ancient hybridization (Yang et al. 2016; Yang et al. 2018). The origin of the sub-alpine *Qch* lineage and the history of its ecotypic divergence should be investigated in future genomic studies including Asian white oak taxa.

Acknowledgements

The authors thank Ayako Kanazashi and Sei-ich Kanetani (Forestry and Forest Products Research Institute) and Megumi Kimura (Akira Prefectural University) for their help in the field work; Y. Taguchi for her assistance with the laboratory work.

References

Adamack AT, Gruber B (2014) PopGenReport: simplifying basic population genetic analyses in R. Methods in Ecology and Evolution 5:384–387. doi: 10.1111/2041-210X.12158

Aizawa M, Maekawa K, Mochizuki H, et al (2018) Unveiling the origin of Quercus serrata subsp. mongolicoides found in Honshu, Japan, by using genetic and morphological analyses. Plant Species Biology 33:174–190. doi: 10.1111/1442-1984.12207

Aizawa M, Maekawa K, Mochizuki H, Iizuka K (2021) Taxonomic revision of quercus serrata subsp. Mongolicoides. Acta Phytotaxonomica et Geobotanica 72:113–123. doi: 10.18942/apg.202017

Bekku YS, Kurokochi H, Matsuki Y, et al (2019) Genetic structure of Pinus parviflora on Mt. Fuji in relation to the hoarding behavior of the Japanese nutcracker. Ecosphere 10:3–8. doi: 10.1002/ecs2.2694

Bettin O, Cornejo C, Edwards PJ, Holderegger R (2007) Phylogeography of the high alpine plant Senecio halleri (Asteraceae) in the European Alps: In situ glacial survival with postglacial stepwise dispersal into peripheral areas. Molecular Ecology 16:2517–2524. doi: 10.1111/j.1365-294X.2007.03273.x

Bohutínská M, Vlček J, Yair S, et al (2021) Genomic basis of parallel adaptation varies with divergence in Arabidopsis and its relatives. Proceedings of the National Academy of Sciences of the United States of America. doi: 10.1073/pnas.2022713118

Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. Molecular Ecology 5:453–455. doi: 10.1046/j.1365-294X.1996.00098.x

Fujii N, Senni K (2006) Phylogeography of Japanese alpine plants: Biogeographic importance of alpine region of Central Honshu in Japan. Taxon 55:43–52. doi: 10.2307/25065527

Gomez JM (2003) Spatial patterns in long-distance dispersal of Quercus ilex acorns by jays in a heterogeneous landscape. Ecography 26:573–584.

Goudet J (2005) Hierfstat, a package for R to compute and test variance components and F-statistics. Molecular Ecology Notes 5:184–186.

Hipp AL, Manos PS, Hahn M, et al (2020) Genomic landscape of the global oak phylogeny. New Phytologist 226:1198–1212. doi: 10.1111/nph.16162

Hirao AS, Shimono Y, Narita K, et al (2019) Ecotypic divergences of the alpine herb Potentilla matsumurae adapted to fellfield–snowbed habitats across a series of mountain sky islands. American Journal of Botany 106:772–787. doi: 10.1002/ajb2.1290

Holliday JA, Zhou L, Bawa R, et al (2016) Evidence for extensive parallelism but divergent genomic architecture of adaptation along altitudinal and latitudinal gradients in Populus trichocarpa. New Phytologist 209:1240–1251. doi: 10.1111/nph.13643

Ikeda H (2022) Decades-long phylogeographic issues: complex historical processes and ecological factors on genetic structure of alpine plants in the Japanese Archipelago. Journal of Plant Research 135:191–201. doi: 10.1007/s10265-022-01377-w

Japan_Meteolorogical_Agency (1987) Explanation of climatic value mesh data. Document 14.

Jombart T (2008) Adegenet: A R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405. doi: 10.1093/bioinformatics/btn129

Kanno M, Yokoyama J, Suyama Y, et al (2004) Geographical distribution of two haplotypes of chloroplast DNA in four oak species (Quercus) in Japan. Journal of Plant Research 117:311–317. doi: 10.1007/s10265-004-0160-8

Knotek A, Konečná V, Wos G, et al (2020) Parallel alpine differentiation in Arabidopsis arenosa. Frontiers in Plant Science 11:1–12. doi: 10.3389/fpls.2020.561526

Konečná V, Nowak MD, Kolář F (2019) Parallel colonization of subalpine habitats in the central European mountains by Primula elatior. Scientific Reports 9:1–12. doi: 10.1038/s41598-019-39669-2

Liu HZ, Harada K (2014) Geographic distribution and origin of the chloroplast T/C-type in Quercus mongolica var. crispula in northeastern Japan. Plant Species Biology 29:207–211. doi: 10.1111/1442-1984.12008

Lowry DB (2012) Ecotypes and the controversy over stages in the formation of new species. Biological Journal of the Linnean Society 106:241–257. doi: 10.1111/j.1095-8312.2012.01867.x

Nagamitsu T, Shimizu H, Aizawa M, Nakanishi A (2019) An admixture of Quercus dentata in the coastal ecotype of Q. mongolica var. crispula in northern Hokkaido and genetic and environmental effects on their traits. Journal of Plant Research 132:211–222. doi: 10.1007/s10265-018-01079-2

Nagamitsu T, Uchiyama K, Izuno A, et al (2020) Environment-dependent introgression from Quercus dentata to a coastal ecotype of Quercus mongolica var. crispula in northern Japan. New Phytologist. doi: 10.1111/nph.16131

Noshiro S (1984) Variations of Quercus mongolica var. undulatifolia and var. grosseserrata on Mt. Makihata, Central Japan. Journal of Phytogeography and Taxonomy 32:116–126. doi: 10.24517/00056232

Ohashi H (1988) The new name instead of Quercus mongolica Fisch. var. grosseserrata (Bl.) Rehd. & Wilis. (Fagaceae). Journal of Japanese Botany 63:13–14.

Ohba H (1989) New names and notes of Japanese woody plants. Journal of Japanese Botany 64:321–329.

Ohba H (2006) Fagaceae. In: Iwatsuki K, Boufford DE, Ohba H (eds) Flora of Japan, IIa. Kodansha, Tokyo, pp 42–60

Ohsawa T, Ide Y (2011) Phylogeographic patterns of highland and lowland plant species in Japan. Alpine Botany 121:49–61. doi: 10.1007/s00035-010-0083-z

Ohsawa T, Tsuda Y, Saito Y, Ide Y (2011) The genetic structure of Quercus crispula in northeastern Japan as revealed by nuclear simple sequence repeat loci. Journal of Plant Research 124:645–654. doi: 10.1007/s10265-010-0402-x

Okaura T, Nguyen DQ, Ubukata M, Harada K (2007) Phylogeographic structure and late Quaternary population history of the Japanese oak Quercus mongolica var. crispula and related species revealed by chloroplast DNA variation. Genes and Genetic Systems 82:465–477. doi: 10.1266/ggs.82.465

Onosato K, Shitara T, Matsumoto A, et al (2021) Contact zone of two different chloroplast lineages and genetic guidelines for seed transfer in Quercus serrata and Quercus crispula. Plant Species Biology 36:72–83. doi: 10.1111/1442-1984.12296

Ortego J, Bonal R, Muñoz A, Aparicio JM (2014) Extensive pollen immigration and no evidence of disrupted mating patterns or reproduction in a highly fragmented holm oak stand. Journal of Plant Ecology 7:384–395. doi: 10.1093/jpe/rtt049

Paradis E, Schliep K (2019) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. Bioinformatics 35:526–528. doi: 10.1093/bioinformatics/bty633

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959. doi: 10.1093/genetics/155.2.945

Quang ND, Ikeda S, Harada K (2008) Nucleotide variation in Quercus crispula Blume. Heredity 101:166–174. doi: 10.1038/hdy.2008.42

R_Core_Team (2019) A language and environment for statistical computing. R Foundation for Statistical Computing

San Jose-Maldia L, Matsumoto A, Ueno S, et al (2017) Geographic patterns of genetic variation in nuclear and chloroplast genomes of two related oaks (Quercus aliena and Q. serrata) in Japan: implications for seed and seedling transfer. Tree Genetics and Genomes. doi: 10.1007/s11295-017-1202-4

Solé-Medina A, Robledo-Arnuncio JJ, Ramírez-Valiente JA (2022) Multi-trait genetic variation in resource-use strategies and phenotypic plasticity correlates with local climate across the range of a Mediterranean oak (Quercus faginea). New Phytologist. doi: 10.1111/nph.17968

Szukala A, Lovegrove-Walsh J, Luqman H, et al (2023) Polygenic routes lead to parallel altitudinal adaptation in Heliosperma pusillum (Caryophyllaceae). Molecular Ecology 32:1832–1847. doi: 10.1111/mec.16393

Tanimoto S, Inoue K, Shibata O (1992) Allozyme variation in Quercus crispula var. crispula and var. horikawae. Journal of Phytogeography and Taxonomy 40:1–4. doi: 10.24517/00055719

Tomaru N, Uchiyama K, Tamaki I, Sakaguchi S (2022) Genetic diversity and population genetic structure of forest tree species in Japan. Japanese Society of Forest Genetics and Tree Breeding, Hitachi

Trucchi E, Frajman B, Haverkamp THA, et al (2017) Genomic analyses suggest parallel ecological divergence in heliosperma pusillum (caryophyllaceae). New Phytologist 216:267–278. doi: 10.1111/nph.14722

Tsumura Y (2006) The phylogeographic structure of Japanese coniferous species as revealed by genetic markers. Taxon 55:53–66. doi: 10.2307/25065528

Tsumura Y (2022) Genetic guidelines for tree species and perspectives on the conservation and sustainable use of forests. Journal of Forest Research 27:83–95. doi: 10.1080/13416979.2022.2040096

Ueno S, Aoki K, Tsumura Y (2009a) Generation of expressed sequence tags and development of microsatellite markers for castanopsis sieboldii var. sieboldii (Fagaceae). Annals of Forest Science 66:509–509. doi: 10.1051/forest/2009037

Ueno S, Taguchi Y, Tomaru N, Tsumura Y (2009b) Development of EST-SSR markers from an inner bark cDNA library of Fagus crenata (fagaceae). Conservation Genetics 10:1477–1485. doi: 10.1007/s10592-008-9764-1

Ueno S, Taguchi Y, Tsumura Y (2008) Microsatellite markers derived from Quercus mongolica var. crispula (Fagaceae) inner bark expressed sequence tags. Genes and Genetic Systems 83:179–187. doi: 10.1266/ggs.83.179

Ueno S, Tsumura Y (2008) Development of ten microsatellite markers for Quercus mongolica var. crispula by database mining. Conservation Genetics 9:1083–1085. doi: 10.1007/s10592-007-9462-4

Wos G, Arc E, Hulber K, et al (2022) Parallel local adaptation to an alpine environment in Arabidopsis arenosa. Journal of Ecology 110:2448–2461. doi: 10.1111/1365-2745.13961

Yang J, Di X, Meng X, et al (2016) Phylogeography and evolution of two closely related oak species (Quercus) from north and northeast China. Tree Genetics and Genomes. doi: 10.1007/s11295-016-1044-5

Yang J, Vazquez L, Feng L, et al (2018) Climatic and soil factors shape the demographical history and genetic diversity of a deciduous oak (quercus liaotungensis) in Northern China. Frontiers in Plant Science 871:1–14. doi: 10.3389/fpls.2018.01534

Author contributions

Study design: LSJ-M and YT; Sampling: LSJ-M, AM, and SU; Phenotyping: LSJ-M; Genotyping: LSJ-M, AM, and SU; Data analysis: LSJ-M and TN; Writing manuscript: LSJ-M and TN

Conflicts of interest

The authors declare no conflicts of interest.

Funding

Genetic Guidelines for Restoration Programs using Genetic Diversity Information (2005–2009) from the Ministry of Environment

Data accessibility

Chloroplast DNA sequences are deposited in DNA Data Bank of Japan (DDBJ) with accessions: AB978567-AB979086. Other data are provided in Supporting Information (Data S1).

Table and figure legend

Table 1. Locations and sample sizes of *Quercus crispula*(*Qc*) and *Q. crispula* var. *horikawae* (*Qch*) populations. Three populations (05Hx, 06Hx, and 07Hx) were identified as*Qch* in the field observation but were grouped to *Qc* in genetic analysis (see Results 3.2).

Figure 1. Growth habits (a, d), leaf shape and size on adaxial surface (b, e), and leaf hairs on abaxial surface (c, f) of*Quercus crispula* (a–c) and *Q. crispula* var.*horikawae* (d–f). Photographs were taken by Lerma San Jose-Maldia (a–d) and Saneyoshi Ueno (e, f).

Figure 2. Locations of *Quercus crispula* (*Qc*) and*Q. crispula* var. *horikawae* (*Qch*) populations in central and northern Honshu. Numbers indicate locations shown in Table 1. Colors of circles indicate taxonomic and genetic categories (red:*Qc* , blue: *Qch* , and green: populations identified as*Qch* in the field observation but grouped to *Qc* in genetic analysis).

Table 2. Frequency of cpDNA haplotypes and genetic diversity of ncEST-SSR genotypes in *Quercus crispula* (*Qc*) and *Q. crispula* var. *horikawae* (*Qch*) populations. Population codes are shown in Table 1.

Figure 3. Principal component analysis of ncEST-SSR genotypes. (a) Ordinations of individuals. Contributions of first–fifth principal components (PCs) to total variation among individuals are 3.80%, 2.81%, 2.69%, 2.59%, and 2.38%, respectively. (b) Ordinations of populations. Contributions of first–fifth PCs to total variation among populations are 22.5%, 17.0%, 7.1%, 6.9%, and 5.8%, respectively. Population codes are shown in Table 1. Colors of circles indicate taxonomic and genetic categories (red: *Qc* , blue: *Qch* , and green: populations identified as *Qch* in the field observation but grouped to *Qc* in genetic analysis).

Figure 4. Bar plots of ancestry proportions of 2–4 clusters (*K*) in STRUCTURE analysis of ncEST-SSR genotypes. Colors of bars indicate inferred ancestries representing; (a) red: *Quercus crispula* (*Qc*) and blue: *Q. crispula* var. *horikawae*(*Qch*) at $K = 2$, (b) red: a part of *Qc* ancestry (*Qc* 1), orange: the other part of *Qc* ancestry (*Qc* 2), and blue: *Qch* at $K = 3$, and (c) red: *Qc* 1, orange:*Qc* 2, light blue: northern lineage of *Qch* ancestry, and blue: southern lineage of *Qch* ancestry at $K = 4$. Population codes are shown in Table 1. Colors of population codes indicate taxonomic and genetic categories (red: *Qc* , blue:*Qch* , and green: populations identified as *Qch* in the field observation but grouped to *Qch* in genetic analysis).

Figure 5. Neighbor-joining tree of populations based on Nei's distance in ncEST-SSR genotypes. Population codes are shown in Table 1. Colors of population codes indicate taxonomic and genetic categories (red: *Qc* , blue: *Qch* , and green: populations identified as*Qch* in the field observation but grouped to *Qc* in genetic analysis). Numbers ([?] 29) at nodes indicate the frequency of trees generated by jackknifing of loci (removing one out of available 29 loci) at focal nodes that contain consistent populations with the tree constructed from all the 29 loci.

Figure 6. Principal component analysis of populations by climatic conditions. (a) Loadings of climatic variables; T01–T12: monthly mean values of daily mean temperature (red arrows), P01–P12: monthly precipitation (blue arrows), and S12–03: monthly maximum snow depth (black arrows) to first and second principal components (PCs) and contributions (%) of first–fifth PCs to total variation. (b) Ordinations of populations in first and second PCs. Population codes are shown in Table 1. Colors of circles indicate taxonomic and genetic categories (red: *Qc* , blue: *Qch* , and green: populations identified as *Qch* in the field observation but grouped to*Qc* in genetic analysis).

Figure 7. Principal component analysis of leaves by leaf characters. (a) Loadings of seven measurements (arrows of a–g) to first and second principal components (PCs) and contributions (%) of first–fifth PCs to total variation among leaves. Ordinations of leaves (b) and populations (c) in first and second PCs. Population codes are shown in Table 1. Colors of circles indicate taxonomic and genetic categories (red: *Qc* , blue: *Qch* , and green: populations identified as *Qch* in the field observation but grouped to*Qc* in genetic analysis).

Supporting Information

Additional supporting information can be found online at https://doi.org/.

Descriptions

Data S1. An excel file with data sheets: 1 (Climate): code of the third mesh (Japan standardized regional mesh code: JIS X0410), where populations are located, and climatic variables (climatic value mesh data, Japan Meteorological Agency 1987) in each mesh, 2 (LeafSize): measurements of three leaves of each individual, 3 (ESTSSR): nuclear EST-SSR genotypes of 693 individuals in 24 populations, 4 (HWEchisq): estimated frequency of null alleles in each locus and P value of χ^2 tests for deviation from the Hardy-Weinberg equilibrium in each locus and each population, 5 (Diversity): locations and genetic diversity indices in populations, 6 (Pairwise): PairwiseF_{ST} values (upper diagonal) and Nei's genetic distances (D_{A}; lower diagonal) between populations. 7 (StructureL): Log-likelihood values of runs with 1–5 clusters in STRUCTURE (Pritchard *et al* . 2000) analysis, 8 (StructureQ): ancestry proportions of 2–4 clusters in STRUCTURE analysis in each individual.

Figure S1. Three hypotheses (a–b) explaining genetic divergence among populations of sub-alpine (A) and mountain (M) ecotypes in northern (1) and southern (2) mountain ranges.

Figure S2. Chloroplast DNA haplotypes discriminated by substitutions at four nucleotide sites in *rp* L32–*trn* L(UAG) region. Nucleotide variation (a) and haplotype network (b) are shown.

Figure S3. Allelic richness of 32 alleles (a, b) and expected heterozygosity (c, d) of populations along latitudinal (a, c) and elevational (b, d) gradients. Colors of circles indicate taxonomic and genetic categories (red: *Qc* , blue: *Qch* , and green: populations identified as *Qch* in the field observation but grouped to *Qc* in genetic analysis).

Figure S4. Log-likelihood values of replications with 1–5 clusters in STRUCTURE analysis.

Table 1. Locations and sample sizes of *Quercus crispula* (*Qc*) and *Q. crispula* var. *horikawae* (*Qch*) populations. Three populations (05Hx, 06Hx, and 07Hx) were identified as *Qch* in the field observation but were grouped to *Qc* in the genetic analysis (see Results 3.2).

Population			Latitude	Longitude	Elevation	Sample size (number of individuals)		
Code	Location	Taxon	(°N)	(°E)	(m)	Leaf size	cpDNA	ncEST-SSR
01C	1: Imabetsu, Aomori	*Qc*	41.189	140.507	20	0	0	22
02C	2: Hakkoda, Aomori	*Qc*	40.663	140.816	690	4	8	32
02H1	2: Hakkoda, Aomori	*Qch*	40.634	140.883	1030	8	7	31
02H2	2: Hakkoda, Aomori	*Qch*	40.650	140.864	1060	5	0	25
03C	3: Hachimori, Akita	*Qc*	40.365	140.027	140	0	0	24
04C	4: Moriyoshi, Akita	*Qc*	40.052	140.617	430	0	8	32
04H	4: Moriyoshi, Akita	*Qch*	39.977	140.526	1260	0	7	32
05Hx	5: Oga, Akita	*Qch*	39.903	139.756	620	4	7	31
06C	6: Chokai, Akita	*Qc*	39.144	139.967	490	4	8	32
06Hx	6: Chokai, Akita	*Qch*	39.132	139.976	740	4	8	32
06H	6: Chokai, Akita	*Qch*	39.118	139.992	1220	4	7	32
07C	7: Hanadate, Yamagata	*Qc*	38.789	140.594	580	4	7	32
07Hx	7: Hanadate, Yamagata	*Qch*	38.791	140.600	810	4	8	32
08C	8: Gassan, Yamagata	*Qc*	38.531	139.956	750	4	8	32
08H	8: Gassan, Yamagata	*Qch*	38.539	139.999	1390	4	8	29
09C	9: Tadami, Fukushima	*Qc*	37.251	139.380	570	0	8	31
09H	9: Tadami, Fukushima	*Qch*	37.235	139.366	760	0	8	32
10C	10: Tanigawa, Gunma	*Qc*	36.838	138.964	690	0	8	18
10H	10: Tanigawa, Gunma	*Qch*	36.818	138.945	1460	0	8	32
11C	11: Hakuba. Nagano	*Qc*	36.665	137.825	1170	0	0	26
11H	11: Hakuba. Nagano	*Qch*	36.660	137.811	1660	4	7	32
12C	12: Shirakawa, Gifu	*Qc*	36.273	136.891	620	0	0	24
13C	13: Kiso, Nagano	*Qc*	35.839	137.508	1320	0	0	24
14C	14: Shitara, Aichi	*Qc*	35.149	137.506	1010	0	0	24
					Total	53	130	693

Table 2. Frequency of cpDNA haplotypes and genetic diversity of ncEST-SSR genotypes in *Quercus crispula* (*Qc*) and *Q. crispula* var. *horikawae* (*Qch*) populations. Population codes are shown in Table 1.

Population	cpDNA haplotype				ncEST-SSR genotype		
code	A	B	C		$AR_{[32]}$	H_E	F_{IS}
01C	–	–	–		5.50	0.672	0.038
02C	8	0	0		5.38	0.647	-0.001
03C	–	–	–		5.73	0.649	0.014
04C	0	8	0		5.95	0.653	0.012
06C	0	8	0		6.23	0.676	0.062
07C	0	7	0		5.70	0.653	0.018
08C	8	0	0		6.00	0.660	0.066
09C	8	0	0		6.05	0.673	0.046
10C	8	0	0		5.49	0.650	0.024
11C	–	–	–		6.04	0.669	0.032
12C	–	–	–		6.01	0.677	0.038
13C	–	–	–		5.81	0.662	0.013
14C	–	–	–		6.02	0.676	0.006
Qc total	32	23	0	Mean	5.84	0.663	0.028
02H1	7	0	0		5.43	0.666	0.014
02H2	–	–	–		5.57	0.665	-0.014
04H	7	0	0		5.71	0.675	0.042
05Hx	7	0	0		5.20	0.620	0.026
06Hx	1	7	0		5.63	0.685	0.027
06H	0	7	0		5.89	0.695	0.054
07Hx	0	8	0		5.53	0.649	0.013
08H	8	0	0		5.04	0.646	-0.029
09H	8	0	0		5.98	0.652	0.015
10H	8	0	0		6.07	0.671	0.005
11H	1	0	6		5.77	0.692	0.008
Qch total	47	22	6	Mean	5.62	0.665	0.015

Figure 1

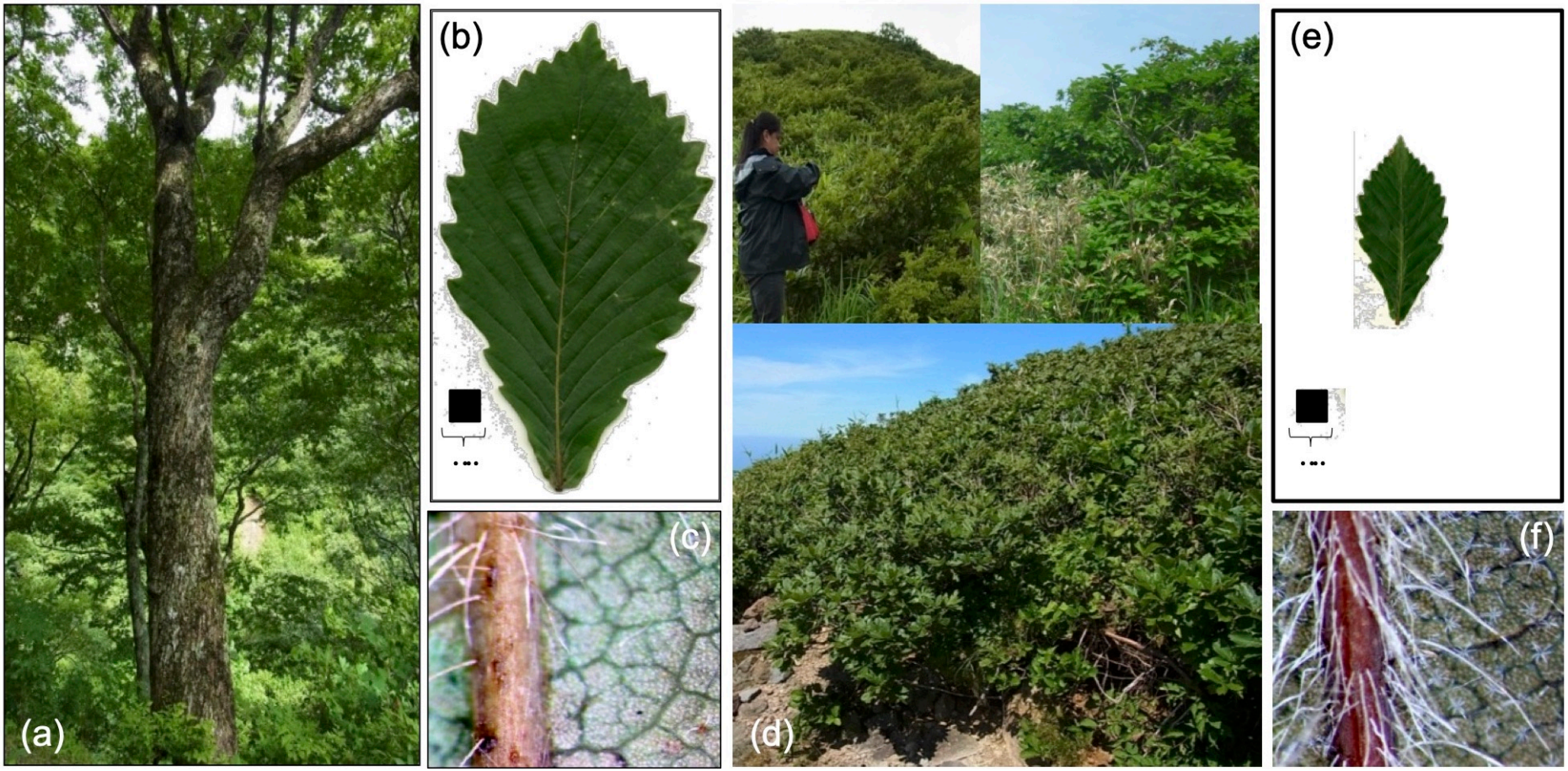


Figure 2

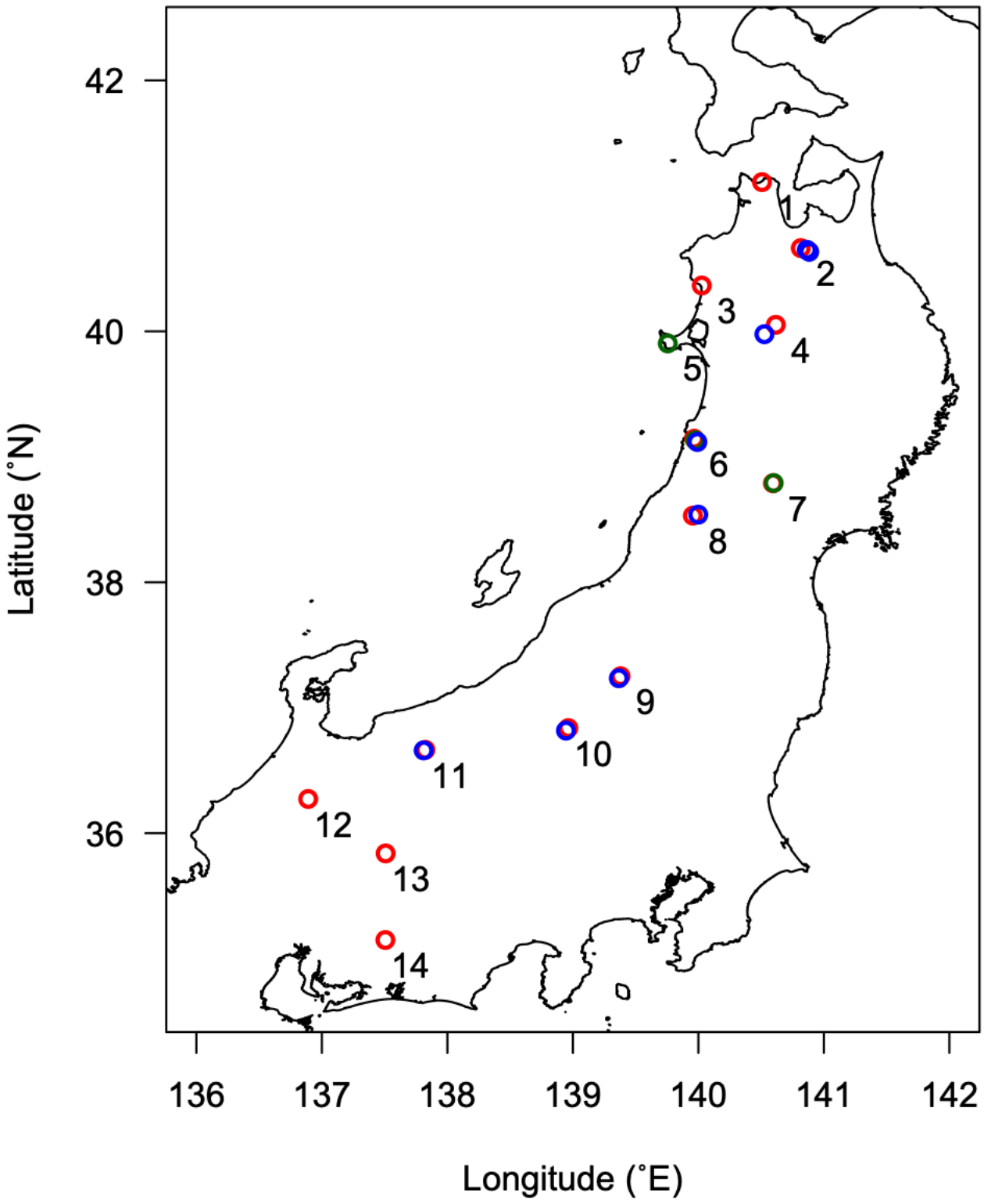

42
40
38
36
Latitude (˚N)
136
137
138
139
140
141
142
Longitude (˚E)
1
2
3
4
5
6
7
8
9
10
11
12
13
14

Figure 3

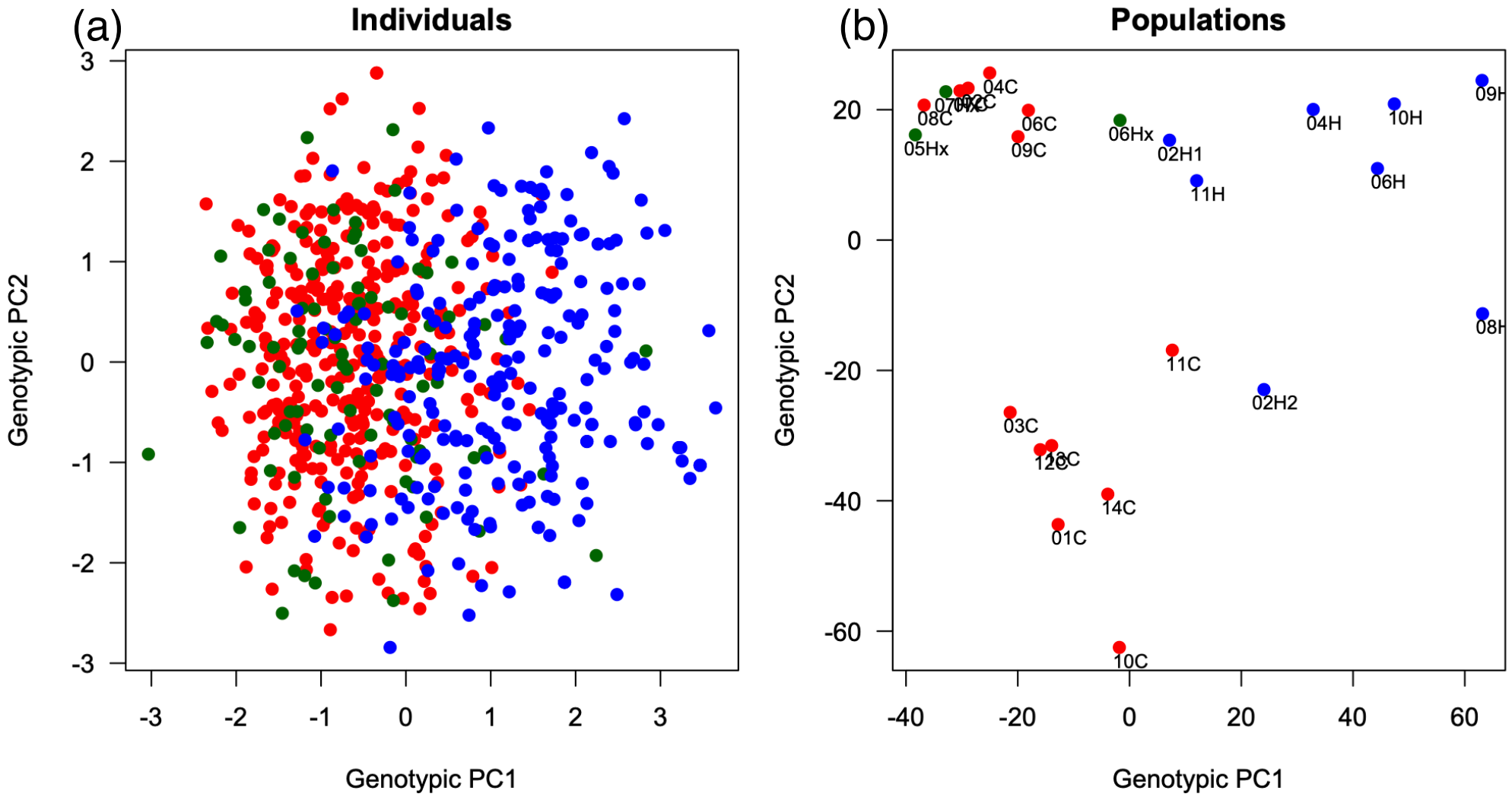

(a)
Individuals
Genotypic PC2
Genotypic PC1
3
2
1
0
-1
-2
-3
(b)
Populations
Genotypic PC2
Genotypic PC1
20
0
-20
-40
-60
-40
-20
0
20
40
60
04C
08C
06C
09C
05Hx
06Hx
02H1
04H
10H
09H
06H
11H
08H
11C
02H2
03C
14C
01C
10C

Figure 4

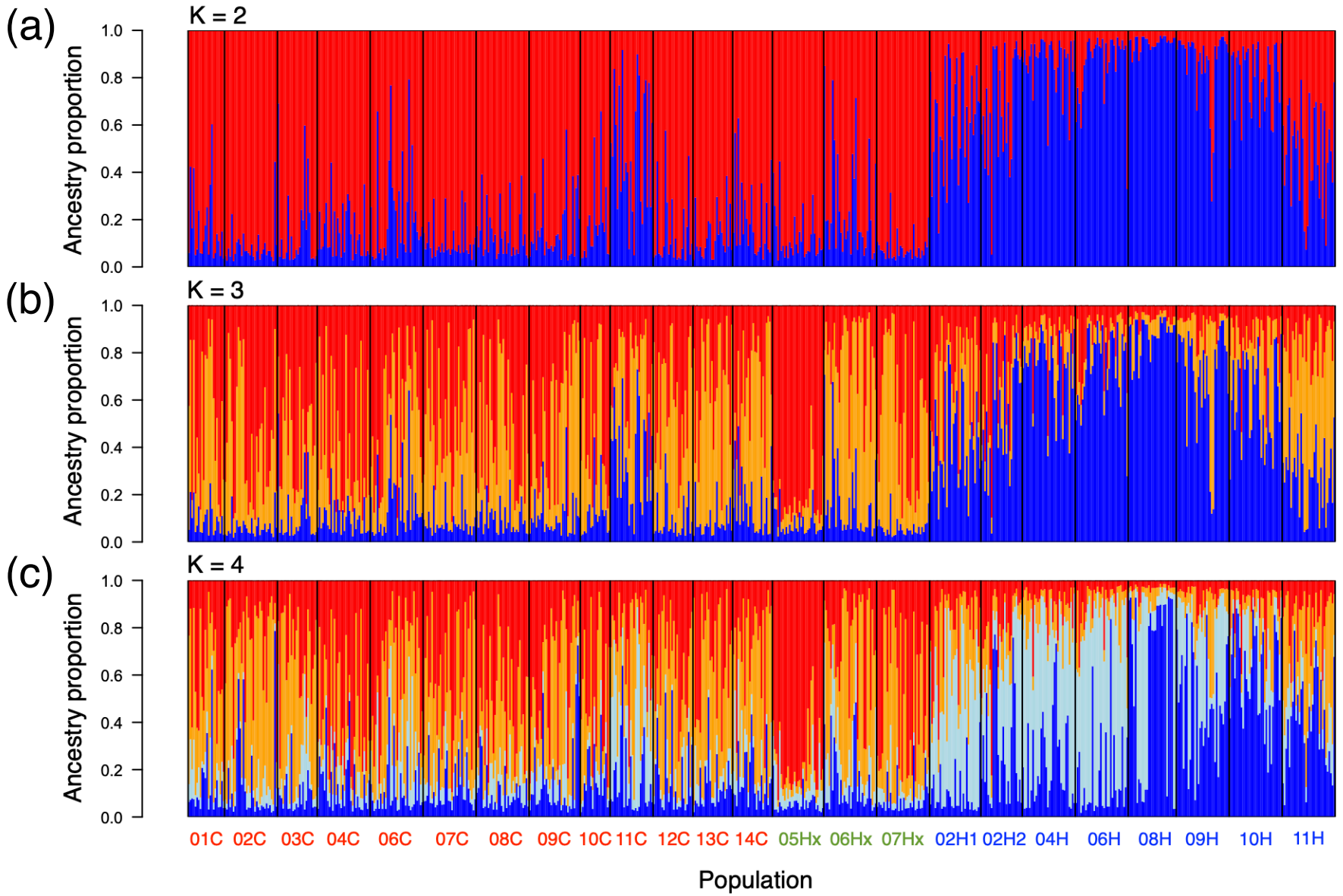

(a)
K = 2
(b)
K = 3
(c)
K = 4
Ancestry proportion
1.0
0.8
0.6
0.4
0.2
0.0
01C 02C 03C 04C 06C 07C 08C 09C 10C 11C 12C 13C 14C 05Hx 06Hx 07Hx 02H1 02H2 04H 06H 08H 09H 10H 11H
Population

Figure 5

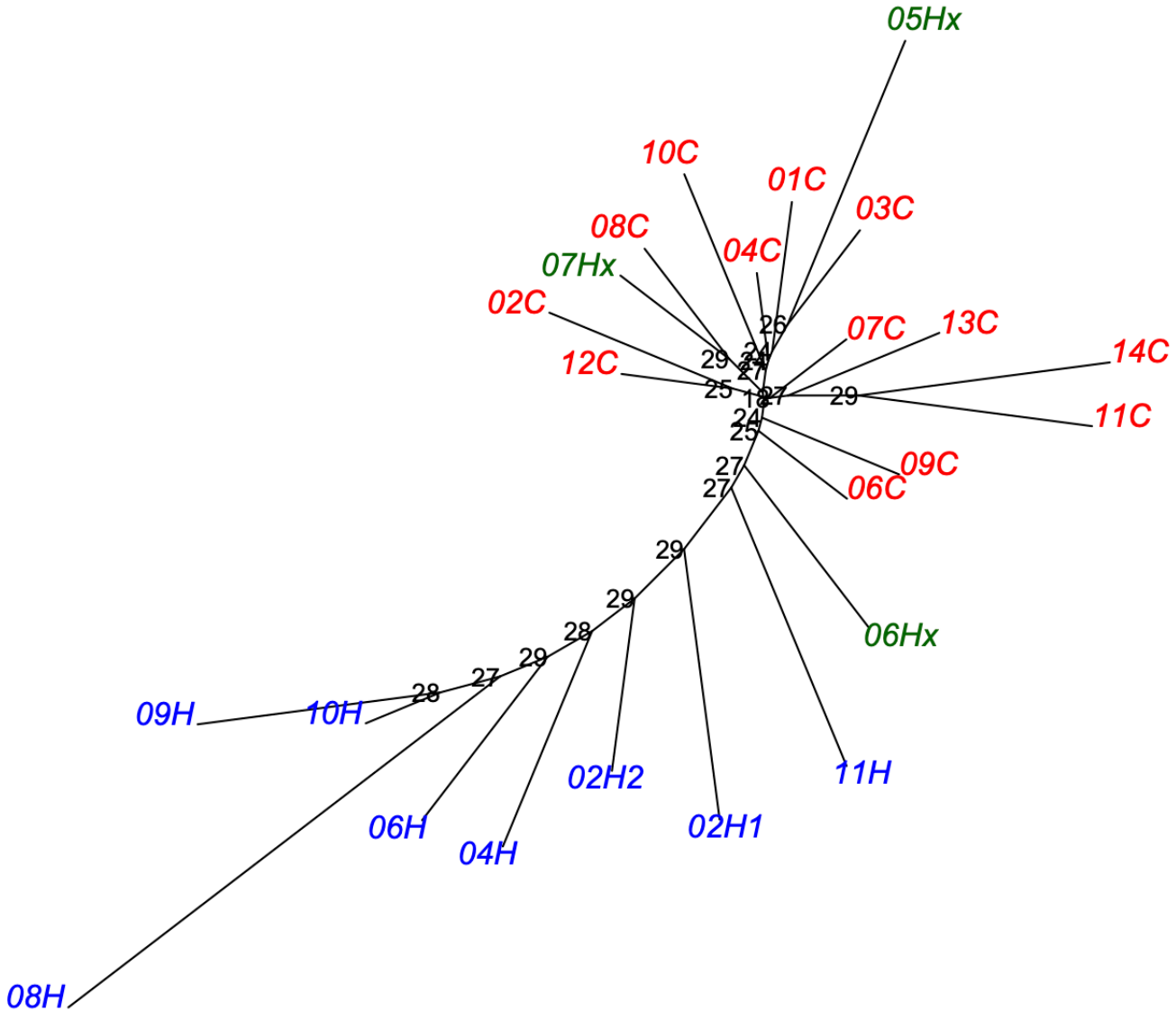

05Hx
10C
01C
03C
08C
04C
07Hx
02C
07C
13C
14C
12C
11C
09C
06C
06Hx
11H
02H2
02H1
06H
04H
09H
10H
08H

Figure 6

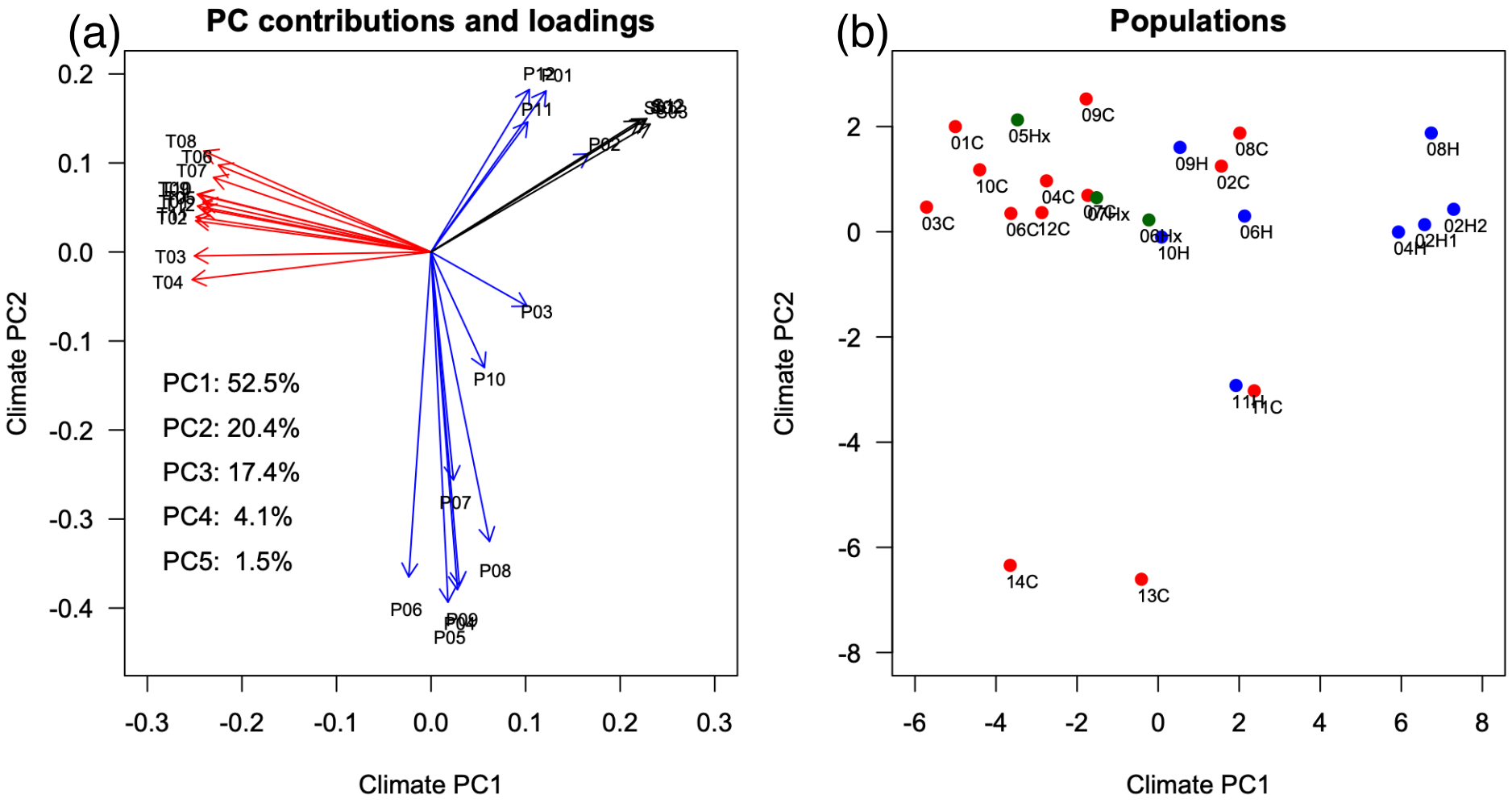
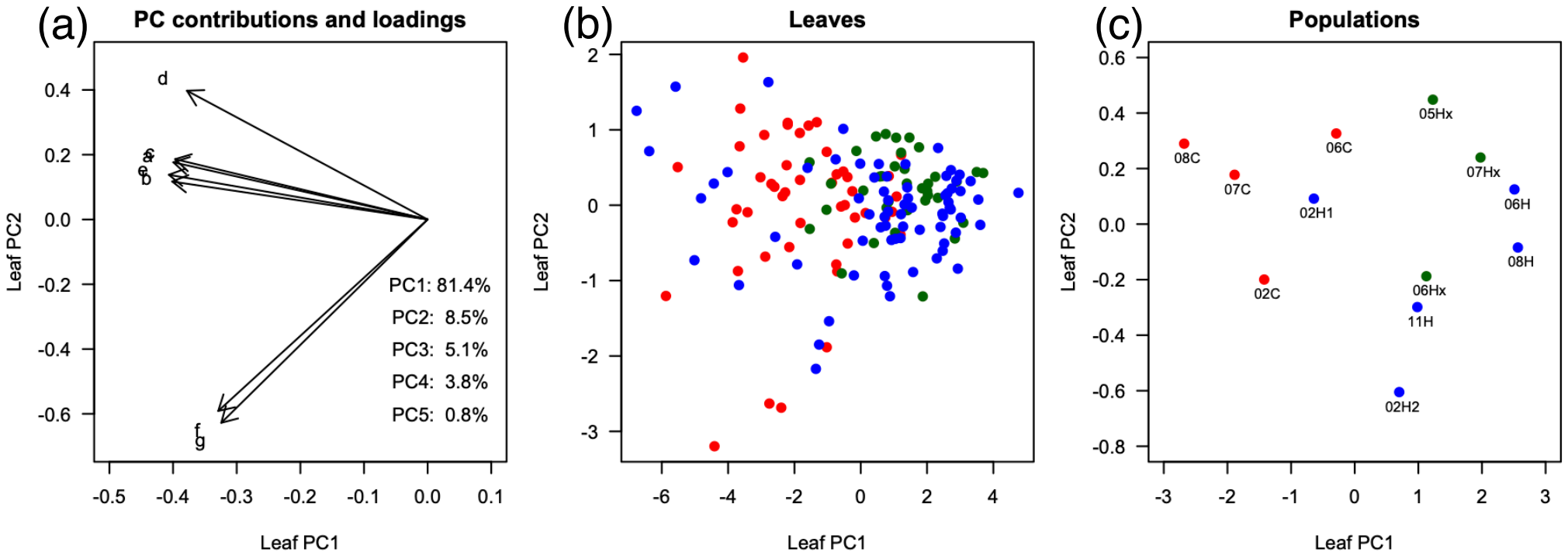

(a)
PC contributions and loadings
Climate PC2
Climate PC1
PC1: 52.5%
PC2: 20.4%
PC3: 17.4%
PC4: 4.1%
PC5: 1.5%
T08
T06
T07
T03
T04
P03
P10
P07
P08
P06
P05
P02
P11
(b)
Populations
09C
01C
05Hx
08C
08H
09H
02C
10C
04C
03C
06C
12C
07Hx
06H
02H2
02H1
04H
10H
11H
11C
14C
13C

Figure 7


(a)
PC contributions and loadings
d
PC1: 81.4%
PC2: 8.5%
PC3: 5.1%
PC4: 3.8%
PC5: 0.8%
Leaf PC2
Leaf PC1
(b)
Leaves
Leaf PC2
Leaf PC1
(c)
Populations
08C
07C
06C
05Hx
07Hx
06H
02H1
08H
02C
06Hx
11H
02H2
Leaf PC2
Leaf PC1

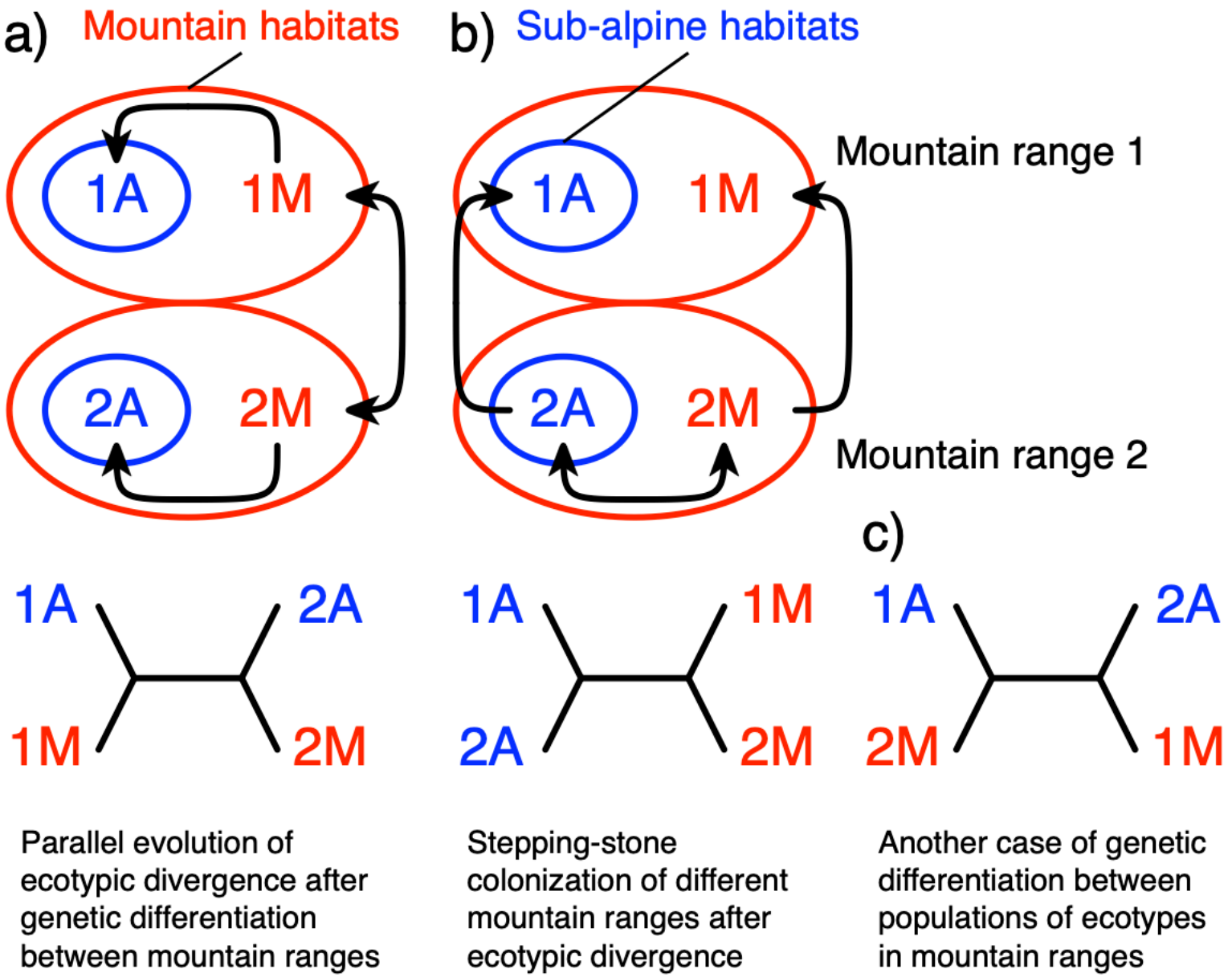


Figure S1. Three hypotheses (a–b) explaining genetic divergence among populations of sub-alpine (A) and mountain (M) ecotypes in northern (1) and southern (2) mountain ranges.

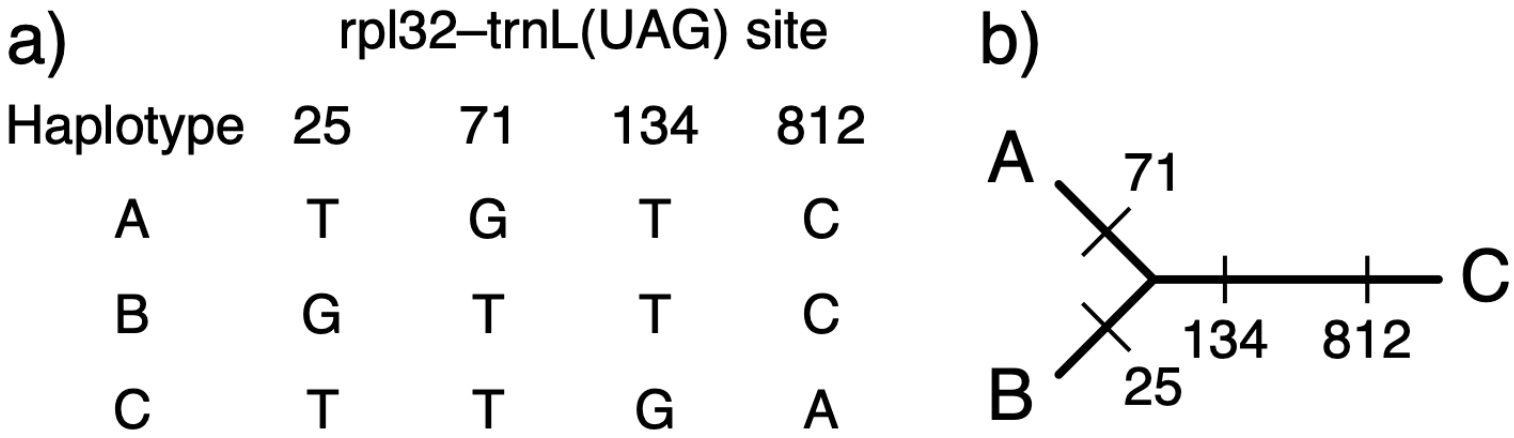


Figure S2. Chloroplast DNA haplotypes discriminated by substitutions at four nucleotide sites in rpL32–trnL(UAG) region. Nucleotide variation (a) and haplotype network (b) are shown.

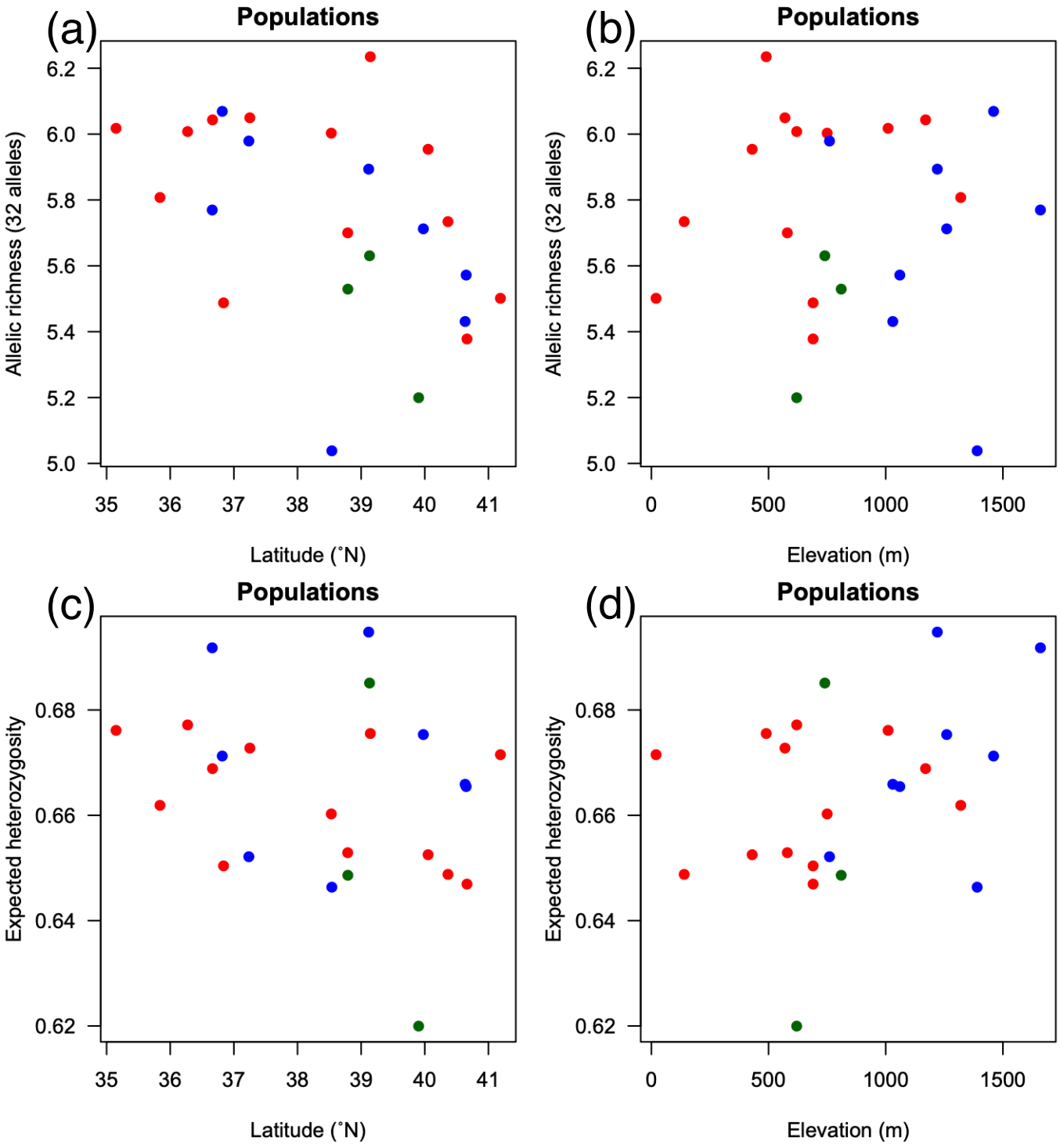


Figure S3. Allelic richness of 32 alleles (a, b) and expected heterozygosity (c, d) of populations along latitudinal (a, c) and elevational (b, d) gradients. Colors of circles indicate taxonomic and genetic categories (red: Qc, blue: Qch, and green: populations identified as Qch in the field observation but grouped to Qc in genetic analysis).

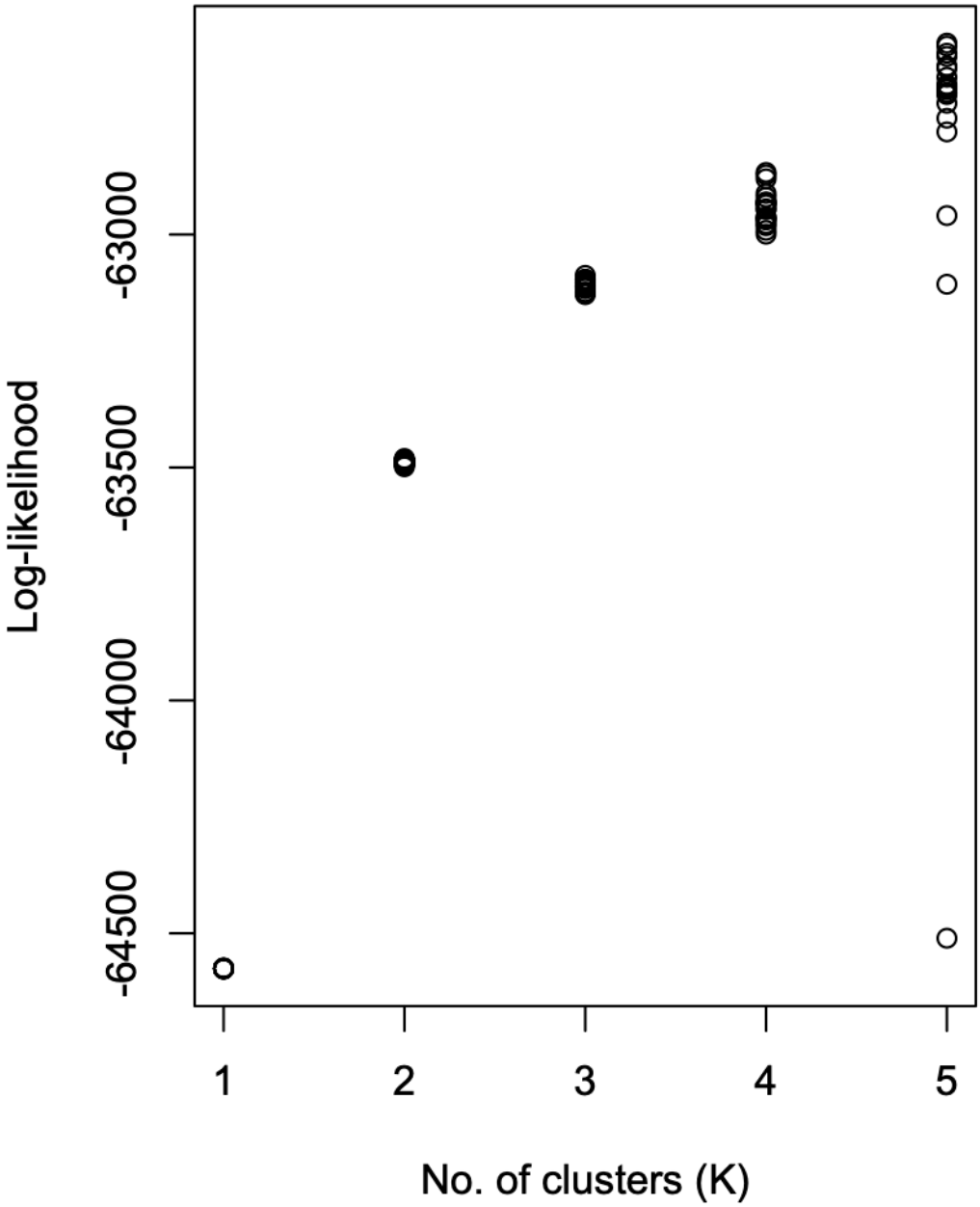


Figure S4. Log-likelihood of replications with 1–5 clusters in STRUCTURE analysis.