ISSN 1346-7565

Acta Phytotax. Geobot. 76 (2): 85–109 (2025)
doi: 10.18942/apg.202508

The Genetic Structure and Demographic History of *Potentilla discolor* (Rosaceae), a Grassland Plant in East Asia, Based on Genome-Wide SNP Markers

Noriyuki Fujii[1*], Hyo Kobatake[2], Kazuharu Niki[2], Reona Taira[2], Takaya Iwasaki[3], Hajime Ikeda[4], Yoshihisa Suyama[5], Ayumi Matsuo[5], Kaori Takeshita (Murayama)[1], Andrey E. Kozhevnikov[6], Zoya V. Kozhevnikova[6], Hyoung-Tak Im[7], Jae-Hong Pak[8], Kyung Choi[9], Hengchang Wang[10], Tian-Gang Gao[11], Hongfeng Wang[12], Siqi Wang[12] and Akiko Soejima[1]

[1]Division of Natural Science, Faculty of Advanced Science and Technology, Kumamoto University, 2-39-1 Kurokami, Chuo, Kumamoto 860-8555, Japan. [*]nfujii@kumamoto-u.ac.jp (author for correspondence); [2]Course for Biological Sciences, Faculty of Science, Kumamoto University, 2-39-1 Kurokami, Chuo, Kumamoto 860-8555, Japan; [3]Natural Science Division, Faculty of Core Research, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan; [4]Department of Multidisciplinary Sciences, Graduate School of Arts and Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8902, Japan; [5]Kawatabi Field Science Center, Graduate School of Agricultural Science, Tohoku University, 232-3 Yomogida, Naruko-onsen, Osaki, Miyagi 989-6711, Japan; [6]Federal Scientific Center of the East Asia Terrestrial Biodiversity, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, Russia; [7]Department of Biology, Chonnam National University, 77 Yongbong-ro, Buk-gu, Gwangju 500-757, Republic of Korea; [8]Research Institute for Dok-do and Ulleung-do Island, Department of Biology, School of Life Sciences, Kyungpook National University, 80 Daehak-ro, Buk-gu, Daegu, 41566, Republic of Korea; [9]Division of Forest Biodiversity, Korea National Arboretum, 415 Gwangneungsumokwon-ro, Soheul-eup, Pocheon-si, Gyeonggi-do, 11186, Republic of Korea; [10]CAS Key Laboratory of Plant Germplasm Enhancement and Specialty Agriculture, Wuhan Botanical Garden, Chinese Academy of Sciences, Wuhan, 430074, Hubei, China; [11]State Key Laboratory of Systematic and Evolutionary Botany, Institute of Botany, Chinese Academy of Sciences, Beijing 100093, China; [12]School of Forestry, Northeast Forestry University, No. 26 Hexing Road, Xiangfang District, Harbin, China

The temperate grassland plant group known as the ‘Mansen elements’ is widely distributed across the grasslands of East Asia, including Japan. It has been hypothesized that these plants originated in continental regions of East Asia and migrated to Japan during the Quaternary period. However, their origins and migration history remain insufficiently elucidated. In this study, we attempted to infer the phylogeographic history of *Potentilla discolor*, a Mansen element member, using single nucleotide polymorphism (SNP) analysis via multiplexed inter-simple sequence repeat genotyping by sequencing (MIG-seq) and ecological niche modeling (ENM). The MIG-seq data revealed distinct genetic differentiation between the Japanese and continental populations. Demographic history analysis suggested that initial divergence likely occurred on the East Asian continent. This finding supports the hypothesis of a continental origin for *P. discolor*. In contrast, chloroplast DNA analysis, commonly used in traditional phylogeographic studies, failed to reveal genetic differentiation between the Japanese and continental populations observed in the SNP-based analyses and provided limited phylogeographic information. ENM analysis predicted that during the Last Glacial Maximum, suitable habitat for *P. discolor* extended continuously from central China to the Pacific coast of Japan via the exposed continental shelf of the East China Sea, suggesting the possibility that this species was introduced to Japan via that route.

Key words: Ecological niche modeling, Grassland, Japanese flora, Mansen elements, MIG-seq, Migration history, Population genetics, *Potentilla discolor*

Global climatic oscillations during the Quaternary significantly impacted the distribution of

numerous plant species on a global scale. Although East Asia was primarily free of ice sheets during the last glacial period (approximately 115,000–12,000 years ago) (Shi *et al.* 1986, Batchelor *et al.* 2019), climatic oscillations during the Quaternary influenced the distribution of vegetation in the region (Axelrod *et al.* 1996, Harrison *et al.* 2001, Qiu *et al.* 2011). The historical vegetation dynamics have been traditionally investigated through studies of pollen (Tsukada 1983, Yasuda & Miyoshi 1998, Ooi 2016) and plant macrofossils (Momohara 2016, 2018), both of which have been crucial in reconstructing regional vegetation histories. In recent decades, phylogeographic studies using molecular data have become increasingly common, allowing more detailed reconstructions of the origins of species or populations and migration histories through phylogenetic and population structure analyses (Avise 2000, 2004). The advent of next-generation sequencing (NGS) has further facilitated the acquisition of genome-wide single nucleotide polymorphism (SNP) data, which has greatly advanced demographic studies of populations (McCormack *et al.* 2013, Sakaguchi *et al.* 2018, Lu *et al.* 2020, Joyce *et al.* 2021, Magota *et al.* 2021, Sata *et al.* 2021, Sakaba *et al.* 2023, Kurata *et al.* 2024, Murakami *et al.* 2024). These developments have significantly enhanced our understanding of the distribution of plant species and their historical changes.

The Japanese Archipelago, located at the eastern edge of the Asian continent and stretching approximately 3,000 km from northeast to southwest, encompasses a wide range of climatic zones from subarctic to subtropical. Under these diverse climatic conditions, the archipelago is inhabited by approximately 5,600 species of vascular plants, demonstrating relatively high species diversity and endemism (Union of Japanese Societies for Systematic Biology 2003, Sabatini *et al.* 2022). The flora of Japan is closely related to that of the Asian continent and is classified as part of the Sino-Japanese Floristic Region, reflecting similarities in floral composition across East Asia (Takhtajan 1969, Good 1974). Recent phylogeographic studies have elucidated historical connections between the Japanese Archipelago and the Asian continent across various plant groups, including alpine species (Fujii *et al.* 1997, Ikeda *et al.* 2014), deciduous trees (Sakaguchi *et al.* 2012, Zeng *et al.* 2015), evergreen broad-leaved trees (Lee *et al.* 2013), and species exhibiting relict disjunct distributions (Li *et al.* 2008, Qiu *et al.* 2009, Qi *et al.* 2014, Sakaguchi *et al.* 2021). Many of those studies suggested that the Japanese flora is primarily related to the Asian continent, with evidence of vicariance between the Japanese and continental populations. However, those studies have not sufficiently addressed the question of whether the origin of certain plant species lies on the continent or within Japan itself, leaving this issue unresolved. Notably, some studies propose a Japanese origin for specific species, rather than a continental one (Ikeda *et al.* 2018, 2020, Kurata *et al.* 2022, Xia *et al.* 2022).

Many temperate grassland plants in Japan are widely distributed across northeastern China, Far East Russia, and the Korean peninsula. Within Japan, however, those species are primarily in the temperate southwestern regions and are largely absent from the northernmost island of Hokkaido (Koidzumi 1931, Hotta 1974, Murata 1988). Koidzumi (1931) referred to such native Japanese plants, which exhibit this distribution pattern, as "Mansen elements," named after the Japanese terms for Manchuria and Korea. Those plants are also sometimes described as "continental grassland relicts" (Ushimaru *et al.* 2018), based on their limited distribution in Japan. Early phytogeographic studies hypothesized that species of the Mansen element originated on the Asian continent and migrated to the Japanese Archipelago via the Korean peninsula during the cold climates of the Pleistocene (Kitamura 1957, Hotta 1974, Murata 1977, 1988, Tabata 1997). Furthermore, given the small morphological differences in the Mansen element species between the continental and Japanese populations, their migration presumably occurred during the recent cool, dry period (Murata 1977, 1988). However, phylogeographic studies specifically targeting the Mansen elements remain limited. Further research is needed to obtain a more detailed understanding

of their origin and migration history.

Recent phylogeographic analyses of the Mansen elements, using SNP data generated by next-generation sequencing, have been conducted on *Viola orientalis* (Maxim.) W. Becker (Violaceae) (Sata *et al.* 2021) and *Tephroseris kirilowii* (Turcz. ex DC.) Holub (Asteraceae) (Sakaba *et al.* 2023). Those studies offered new insights into the origins and diversification of those elements. The analysis of *V. orientalis* suggested that migration from the continent to Japan occurred during the Last Glacial Maximum (LGM), supporting the hypothesis of a continental origin for the Mansen elements. In contrast, the study of *T. kirilowii* revealed clear genetic differentiation between Japanese and continental populations, making it difficult to determine the phylogenetic origin of the Japanese populations. Those findings indicated that the Mansen elements may include not only species of continental origin but also those with distinct diversification histories. While the hypothesis of a continental origin is partially supported, the evolutionary and differentiation processes of individual Mansen element species likely reflect variability. Thus, refining the hypothesis regarding the origin and diversification of the member species will require further research.

In this study, *Potentilla discolor* Bunge (Rosaceae), a representative Mansen element species recognized by Hotta (1974) and Murata (1988), was selected as the target species. *Potentilla discolor* is a small perennial herb of sunny grasslands and dry meadows. Its distribution ranges from southwestern to northeastern China, the southern Primorsky region of Russia and the Korean Peninsula (Fig. 1). In Japan, its range is limited to Kyushu, Shikoku and the western part of Honshu (west of Aichi Prefecture) (Naruhashi 2001, Ikeda 2016). Due to the loss of grasslands, *P. discolor* is classified as Vulnerable (VU), with only small, isolated populations remaining (Ministry of the Environment Japan 2015). Species of *Potentilla* with a corolla resembling that of *P. discolor* are known to be entomophilous, employing honeybees, wasps, and bumblebees as pollinators (McIver & Erickson 2012). Although the achenes of *P. discolor* are ovoid and small (*ca*. 1 mm in length along the long axis, Ikeda 2016), they are considered gravity-dispersed because they lack appendages for wind dispersal. The chromosome number of *P. discolor*, $2n = 14$ (Naruhashi 1970), is considered to be diploid with a basic chromosome number of 7 (Delgado *et al.* 2000), making it advantageous for population genetic analysis. Given these characteristics, it is an appropriate model species for testing hypotheses related to the origins and migration history of the Mansen elements.

This study aimed to test hypotheses regarding the origins and migration history of the Mansen elements, using *Potentilla discolor* as a model species. We analyzed the phylogenetic relationships, genetic structure, and demographic history of populations of *P. discolor* based on SNP data obtained through multiplexed inter-simple sequence repeat genotyping by sequencing (MIG-seq). Additionally, we investigated geographic variation in chloroplast DNA (cpDNA), a marker commonly used in traditional phylogeographic studies, and compared the results with those derived from the SNP analyses. Ecological niche modeling (ENM) was used to estimate the potential distribution of *P. discolor* during the most recent glacial period (LGM, approximately 21,000 years before present (BP)). By integrating the datasets, we evaluated the origin and migration history of *P. discolor*. Our study has provided important insights into the formation processes of the grassland flora in East Asia.

Materials and Methods

Plant materials and DNA extraction

A total of 303 individuals from 22 populations (p1–p22) of *Potentilla discolor* were used in the present analyses (Fig. 1, Table 1). Among the samples, 264 individuals from 19 populations were selected for MIG-seq analysis; 237 individuals from 22 populations were utilized for cpDNA analysis. Populations p3 and p13 were excluded from the MIG-seq analysis due to insufficient sample size. Furthermore, population p9 was excluded due to its geographic proximity to p10,

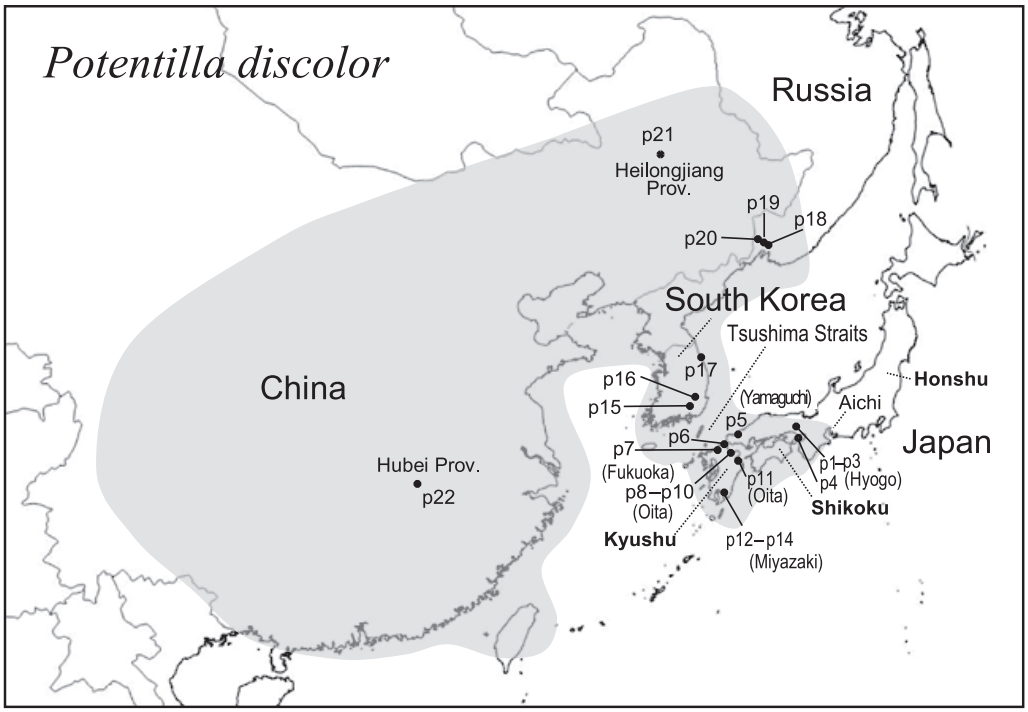

FIG. 1. Distribution map of *Potentilla discolor* and collection sites (p1–p22) of material used in this study. The approximate distribution range of *P. discolor* is depicted in gray.

TABLE 1. The materials and their sources analyzed for MIG-seq and cpDNA variations.

Population code and taxon	Locality and voucher	No. of plants	
		MIG-seq	cpDNA
Potentilla discolor Bunge			
Japan			
p1	Miki-shi, Hyogo, Honshu, *N. Fujii F03471* (KUMA)	16	15
p2	Kato-shi, Hyogo, Honshu, *N. Fujii & K. Niki F03465* (KUMA)	15	13
p3	Himeji-shi, Hyogo, Honshu, *H. Ono F03464* (KUMA)	0	1
p4	Awaji-shi, Hyogo, Honshu, *M. Yokogawa July 23, 2018* (HYO)	8	6
p5	Shuhou-cho, Yamaguchi, Honshu, *Y. Ota & K. Niki F03475* (KUMA)	16	13
p6	Kitakyusyu-shi, Fukuoka, Kyushu, *N. Fujii et al. F03460* (KUMA)	15	16
p7	Fukuoka-shi, Fukuoka, Kyushu, *N. Fujii & K. Niki F03473* (KUMA)	13	10
p8	Nakatsu-shi, Oita, Kyushu, *N. Fujii et al. F03461* (KUMA)	13	15
p9	Usa-shi, Oita, Kyushu, *N. Fujii et al. F03552* (KUMA)	0	10
p10	Usa-shi, Oita, Kyushu, *N. Fujii et al. F03551* (KUMA)	14	2
p11	Oita-shi, Oita, Kyushu, *N. Fujii & K. Niki F03474* (KUMA)	15	10
p12	Kunitomi-cho, Miyazaki, Kyushu, *N. Fujii et al. F03462* (KUMA)	14	16
p13	Kunitomi-cho, Miyazaki, Kyushu, *N. Fujii et al. F03463* (KUMA)	0	1
p14	Kushima-shi, Miyazaki, Kyushu, *N. Fujii et al. F03547* (KUMA)	16	10
South Korea			
p15	Miryang, Gyeongsangbuk-do, *N. Fujii F03479* (KUMA)	11	11
p16	Yeongcheon, Gyeongsangbuk-do, *N. Fujii F03478* (KUMA)	15	14
p17	Gangneung, Gangwon-do, *N. Fujii F03480* (KUMA)	15	14
Russia			
p18	10km west of Ussuriysk, Primorsky Krai, *N. Fujii et al. F03468* (KUMA)	12	13
p19	20km northwest of Ussuriysk, Primorsky Krai, *N. Fujii et al. F03466* (KUMA)	16	15
p20	55km northwest of Ussuriysk, Primorsky Krai, *N. Fujii et al. F03467* (KUMA)	13	17
China			
p21	Daqing, Heilongjiang Prov., *H. Wang F03553* (KUMA)	15	5
P22	Wuhan, Hubei Prov., *T. Feng & H. Wang F03545* (KUMA)	12	10
	Total	264	237
Outgroup			
Potentilla chinensis Ser.			
	Yeongweol, Gangwon-do, South Korea, *N. Fujii F03482* (KUMA)	1	1

which may have resulted in redundancy. In each population, leaf samples were collected from individuals as far apart as possible and stored with silica gel beads. We also collected a sample of *Potentilla chinensis* Ser. from South Korea to serve as an outgroup. Outgroup selection was based on a molecular phylogenetic study that analyzed both chloroplast and nuclear DNA regions (Feng *et al.* 2017). That study demonstrated that *P. discolor* and *P. chinensis* were closely related species in the same Argentea clade. Voucher specimens collected during our study have been deposited in the herbarium of the Faculty of Science, Kumamoto University (KUMA), and the Museum of Nature and Human Activities, Hyogo Prefecture (HYO).

The sampled leaves were ground into a fine powder using a mixer mill (Multi-Beads Shocker MB455U, Yasui Kikai Corp., Osaka, Japan). Prior to DNA extraction, the samples were resuspended in the wash buffer described by Wagner *et al.* (1987) to remove polysaccharides; the washing procedure was performed twice. Total genomic DNA was extracted from the washed pellets using the cetyltrimethylammonium bromide (CTAB) method as described by Doyle and Doyle (1987).

MIG-seq analysis

A library for MIG-seq analysis was prepared using a slightly modified protocol (Suyama *et al.* 2022) from the original method by Suyama & Matsuki (2015). The first polymerase chain reaction (1st PCR) was performed with an annealing temperature of 38°C to enhance amplicon yield. Then, using a MiSeq Reagent Kit v.3 (150 cycles, Illumina), the library was sequenced on an Illumina MiSeq Sequencer, without 'DarkCycle' option (Illumina, San Diego, CA, USA). Next, using a FASTX Toolkit (http://hannonlab.cshl.edu/fastx_toolkit/), we removed low quality reads (q = 30, p = 40). To eliminate reads derived from extremely short library entries, we filtered out those containing adapter sequences using TagDust v.2.2 (Lassmann *et al.* 2009). The 'ustacks,' 'cstacks,' and 'sstacks' modules of Stacks v.1.48 software (Catchen *et al.* 2013) were utilized to detect SNPs in the quality-filtered reads with the following parameter settings: maximum distance (in nucleotides) allowed between stacks (M = 2), minimum depth of coverage required to create a stack (m = 3), and maximum distance allowed to align secondary reads to primary stacks (N = 4). The 'populations' module in Stacks was used to query SNPs with the following filtering settings (R = 0.1, --min-mac = 3, --max-obs-het = 0.6). The R parameter specifies the minimum percentage of individuals across populations required to process a locus. After the Stacks procedure, the dataset was filtered using TASSEL v.5 (Bradbury *et al.* 2007) as the following setting (Site Min Count = 80, Site Min Allele Freq = 0.01), and one individual in the Russian population (p20_6) with highly missing rate was excluded from the dataset. To avoid over-filtering of the SNP dataset for phylogenetic analysis, we followed the guidelines of Suissa *et al.* (2024). SNP retention was adjusted to a moderate level, retaining SNPs found in 45–75% of individuals. We converted the VCF file to a PHYLIP matrix using vcf2phylip.py (Ortiz 2019) and removed invariant sites using raxml_ascbias (https://github.com/btmartin721/raxml_ascbias), resulting in a dataset of 264 individuals, including one outgroup sample, with a total of 3,566 SNPs [genotyping rate (proportion not missing): 58.9%] (Dataset 1). This dataset was used for phylogenetic and network analyses.

For evaluation of genetic diversity, genetic structure and population demographic analyses, the 'ustakcs' and 'populations' modules in Stacks were used to query SNPs with the following filtering settings (ustacks: M = 2, m = 3, N = 4; populations: R = 0.5, --min-mac = 3, --max-obs-het = 0.6, --write-random-snp). After the Stacks procedure, the dataset was filtered using TASSEL v.5 as the following setting (Site min count = 159, Site Min Allele Freq = 0.01), and the outgroup sample (*P. chinensis*) and one individual in the Russian populations (p20_6) with highly missing rate were excluded, resulting in a dataset comprising 825 SNPs across 263 individuals. Outlier loci were removed using the R package pcadapt v.4.3.5 (Privé *et al.* 2020) in R v.4.3.2 (R Core Team 2023), with parameters set to $K = 5$ and

q-value < 0.01. Finally, the dataset consisted of 812 SNPs from 263 individuals, with a genotyping rate of 79.9% (Dataset 2). MIG-seq data are available in the International Nucleotide Sequence Database Collaboration (INSDC) Sequenced Read Archive under accession numbers DRR415804–DRR416068.

Data analysis based on MIG-seq data

The genetic diversity parameters of the mean number of different alleles per locus (*Na*), mean number of different alleles unique to a single population per locus (*PA*), mean number of locally common alleles per locus found in 25% or fewer populations (*LCA*), observed heterozygosity (H_O), expected heterozygosity (H_E), and unbiased expected heterozygosity (uH_E) were calculated for each population by GenAlEx v.6.501 (Peakall & Smouse 2012). Input files for GenAlEx were prepared by converting the post-pcadapt output to Genepop format via PGDspider v.2.1.1.5 (Lischer & Excoffier, 2012). The parameter of allelic richness (*AR*) (el Mousadik & Petit 1996) and fixation index (F_{IS}) for each population, along with their 95% confidence intervals, were estimated with 9,999 bootstraps using the R package diveRsity v.1.9.90 (Keenan *et al.* 2013). Additionally, Hardy-Weinberg Equilibrium (HWE) was tested for each population using an exact test implemented in the same R package. For the statistical test of differences between Japanese and continental populations, the *PA*, *LCA*, *AR*, H_E, and uH_E parameters were calculated for each locus in each population using the R package diveRsity v.1.9.90. To examine differences in each parameter between Japanese and continental populations, generalized linear mixed model (GLMM) analyses were conducted. Random effects were assigned to locus names, and distributions were applied as follows: a Poisson distribution for *PA* and *LCA* and a normal distribution for the remaining parameters. The analyses were performed using the R package lme4 v.1.1 (Bates *et al.* 2015).

We also conducted a molecular variation analysis (AMOVA) in Arlequin v.3.5.1.2 (Excoffier & Lischer 2010) to examine variation within and among populations. The hierarchy for that analysis was chosen based on the clustering from the STRUCTURE analysis. The significance of the covariance components associated with the different possible levels of genetic structure was tested using 1,023 permutations.

Phylogenetic trees were reconstructed using the maximum likelihood (ML) method implemented in RAxML v.8.2.12 (Stamatakis 2014) via the raxmlGUI v.2.0.10 interface (Edler *et al.* 2020). Since the dataset used in this study consisted of concatenated SNP data, a model selection analysis was not performed. Instead, the nucleotide substitution model was set to GTRGAMMA with correction for ascertainment bias (-model GTR+G+ASC_LEWIS), as this model is commonly used in phylogenetic studies and suitable for concatenated SNP datasets (Leaché *et al.* 2015). Branch support values were assessed with 1,000 bootstrap resamplings. The resulting phylogenetic tree was visualized using FigTree v.1.4.4 (Rambaut 2009) and rerooted with the outgroup sample *P. chinensis*.

We used STRUCTURE v.2.3.4 (Pritchard *et al.* 2000) to infer population genetic structure according to the admixture model with the allele frequency correlated model (Falush *et al.* 2003). The input file for STRUCTURE was prepared by converting the pcadapt-processed file into STRUCTURE format. Both the burn-in period and the number of MCMC repetitions after the burn-in period were set to 100,000. We assumed between 1 and 15 clusters (*K*), with 15 independent simulations for each *K*. To determine the optimal *K* value, we used STRUCTURE HARVESTER v.0.9.94 (Earl & Vonholdt 2012), which evaluates the log probability of the data [LnP(D)] and ΔK (delta *K*) based on the rate of change in [LnP(D)] between successive *K* values (Evanno *et al.* 2005). The online version of CLUMPAK v.1.1 (http://clumpak.tau.ac.il/index.html, Kopelman *et al.* 2015) was used to summarize the results from all runs and to generate representative bar plots for each *K* value. Additionally, we calculated the expected heterozygosity and F_{ST} values within each cluster using STRUCTURE. The F_{ST} values for each cluster represent the degree of genetic drift from a common ancestral population,

analogous to traditional F_{ST} values between clusters and their shared ancestral population (Pritchard *et al.* 2010).

For network analysis, pairwise genetic distances were calculated as binary Genetic Distances (GD) among individuals using TASSEL v.5. A phylogenetic network using the Neighbor-Net method (Bryant & Moulton 2004) was reconstructed using SplitsTree v.4.10 (Huson & Bryant 2006) on the basis of the calculated GD.

To infer the demographic history for *Potentilla discolor* using observed SNP variation, we employed explicit model comparisons in the program DIYABC-RF v.2.1.0 (Collin *et al.* 2021). DIYABC-RF integrates Approximate Bayesian Computation (ABC) with supervised machine learning based on Random Forests (RF) to evaluate alternative evolutionary scenarios. This approach effectively integrates the strengths of ABC and Random Forests, enabling precise model selection and enhanced handling of complex, high-dimensional datasets. As a first step, considering the genetic structure revealed by STRUCTURE and NeighborNet analyses, we classified the 19 studied populations into five regional population groups: Pop1, three Japanese populations (p1, p7, p14): Pop2, South Korean populations (p15–p17): Pop3, Far East Russian populations (p18–p20): Pop4, Chinese Heilongjiang population (p21): Pop5, Chinese Hubei population (p22). In Pop1, the three populations were selected based on the results of phylogenetic analyses to minimize sample size bias. Using Dataset 2, we excluded all missing SNPs across individuals within each group, resulting in a final dataset comprising 11 populations, 153 individuals and 732 SNPs (genotyping rate: 68.2%).

Considering the phylogenetic tree estimated by RAxML analysis, the population genetic structure detected by STRUCTURE and NeighborNet analyses and the results of multiple preliminary ABC-RF analyses under various demographic scenarios, we proposed eight scenarios (Scenarios 1–8) to estimate the demographic histories (Fig. S1). These scenarios were characterized by several demographic parameters, including divergence times in generation (t1–t4) and effective population sizes for Pop1–Pop5 (N1–N5) and ancestral populations (Na1–Na3, and Nanc). Scenarios 1–4 assume that the Hubei population (Pop5) diverged first from the others, with different orders among subsequent populations in each scenario. In contrast, Scenarios 5–8 assume that the Heilongjiang (Pop4), Far East Russian (Pop3), South Korean (Pop2), and Japanese (Pop1) populations diverged first from the others, respectively. Scenarios 1, 2, 4, 5, and 6 posit that the initial divergence occurred in the continental region before the populations reached Japan. Scenario 3 assumes an initial divergence between the central Chinese and Japanese populations, while Scenarios 7 and 8 propose that the initial divergence occurred between the Japanese and South Korean populations. Training sets included 100,000 datasets per scenario for the scenario choice and 100,000 datasets under a given scenario for parameter estimation steps, as recommended by Collin *et al.* (2021). Prior parameters for divergence times and effective population sizes were set as uniform distribution from 100 to 200,000. The posterior probability as well as the global and local error rates were used to evaluate scenario choice and the quality of prediction (Collin *et al.* 2021). We estimated the posterior distributions of demographic parameters under the best scenario and the precision of each parameter estimation by computing the posterior mean, median, and 90% credibility intervals (defined as the 5% and 95% quantiles) for each parameter (Raynal *et al.* 2019).

In the second-step analysis, we performed DIYABC-RF focused on the Japanese populations of *P. discolor* to estimate their demographic history. Considering the genetic structure in STRUCTURE and NeighborNet analyses, and the phylogenetic relationships of the ML tree, we classified the 11 studied populations into three regional population groups: Pop1, the Hyogo populations (p1, p2, p4); Pop2, the northern Kyushu to Yamaguchi population (p5, p6, p10, p11); Pop3, the Miyazaki populations (p12, p14). The northern Kyushu populations P7 and P8 were excluded from this analysis to avoid the complications of estimating complex scenarios due to their mixed genetic origins (Figs. 2–4). On the other hand,

samples from three South Korean populations (P15–17) were included as a reference group (Pop4). For this analysis, a dataset consisting of 573 SNPs from 170 individuals across 12 populations was constructed from Dataset 2. Considering the topology of the phylogenetic tree and the results of several preliminary ABC analyses using different demographic models, we estimated the demographic history under three scenarios (Scenarios A–C). (Fig. S2). Those scenarios were characterized by 10 demographic parameters, including divergence times in generations t1–t3, effective population sizes of Pop1–Pop4 (N1–N4) and the ancestral populations (Na1, Na2, Nanc). Scenario A assumes the initial divergence of the Hyogo populations located in the mid-western Honshu from the other populations, while Scenario B posits the earliest divergence of the Miyazaki populations in southern Kyushu. Scenario C, on the other hand, assumes a scenario in which the ancestor of Japanese population diverged almost simultaneously into three distinct groups. Prior parameters of divergence times and effective population sizes were set as uniform distribution from 100 to 100,000. Model choice and parameter estimation were performed under the same conditions as in the first-step analysis.

Direct sequencing analysis of cpDNA

The 5′-terminal region of *trnK* (including a portion of *matK*) in cpDNA was sequenced. In preliminary experiments involving three other cpDNA regions [*trnL–trnF* (Taberlet *et al.* 1991), *rpl16* intron (Jordan *et al.* 1996), and *trnS–trnG* (Shaw *et al.* 2005)], this region was found to be the most variable. The 5′-terminal region of *trnK* was amplified by PCR using the primers trnK11 and matK510R (Young *et al.* 1999). The 25 μL PCR reaction mixture contained the following reagents: 50–100 ng of template DNA, 2.5 μL of 10×PCR buffer, 0.2 mM of each deoxyribonucleotide, 2.0 mM of $MgCl_2$, 0.4 μM of each primer, and 0.5 U of ExTaq DNA polymerase (Takara Bio Corporation, Tokyo, Japan). Initial denaturation was performed for 3 min at 94 °C, followed by 30 cycles of denaturation at 94 °C for 1 min, primer annealing at 53 °C for 1 min, and extension at 72 °C for 2 min. The reactions were then extended by 7 min at 72 °C. After the PCR reaction, Illustra ExoProStar (GE Healthcare, Tokyo) was utilized to remove the excess primer and nucleotide. The same primers (final concentration 9.6 pM) were used for sequencing analyses to read in both directions. The sequencing of DNA was contracted out to Eurofins Genomics Corp. (Tokyo, Japan). The obtained nucleotide sequences were edited with ChromasPro v.1.7.6 (www.technelysium.com.au/ChromasPro.html) and aligned with MEGA v.10.1.8 (Kumar *et al.* 2018).

Data analyses using cpDNA sequences

DnaSP v.6.12.01 (Rozas *et al.* 2017) was used to determine cpDNA haplotypes based on the 5′-terminal region of *trnK*. In this analysis, only nucleotide substitutions were used to distinguish haplotypes and insertions/deletions (indels) were excluded. Haplotype diversity (*Hd*) and nucleotide diversity (π) were estimated among the Japanese populations and the continental populations using DnaSP. Populations with small sample sizes (p3, p10, and p13) were excluded from the analysis, and data from 233 individuals across 19 populations were used. To compare the genetic differentiation between Japanese and continental populations, F_{ST} values were calculated using Arlequin v.3.5.1.2. Additionally, a median-joining (MJ) network among the haplotypes was constructed using Network v.5.0.1 (Bandelt *et al.* 1999). The nucleotide sequence data were deposited in the DDBJ/GenBank/EMBL databases under accession numbers LC624803–LC624817.

ENM analysis

ENM was performed to evaluate the potential distribution of *P. discolor* under present climatic conditions and to project its range back to the LGM (*ca*. 21,000 years BP). Occurrence data for this species were obtained from Global Biodiversity Information Facility (GBIF) (GBIF.org. 2021), and were filtered to include samples collected after 1950, with voucher specimens, and latitude and longitude information (https://doi.org/10.15468/dl.vysprd). The geographic range for this analysis was defined as 90–150°E and

TABLE 2. Summary of genetic diversity of *Potentilla discolor* populations by regional and country groups using MIG-seq data[1].

	Na	*PA*	*LCA*	*AR*	H_O	H_E	uH_E
Japanese populations							
Hyogo (p1, p2, p4)	1.143	0.003	0.030	1.103	0.058	0.090	0.060
Northern Kyushu[2] (p5, p6, p10, p11)	1.283	0.003	0.031	1.225	0.079	0.089	0.091
Miyazaki (p12, p14)	1.274	0.005	0.039	1.217	0.084	0.090	0.091
All_Japan (p1-p14)	1.241	0.003	0.032	1.187	0.074	0.089	0.082
Continental populations							
Korea (p15-p17)	1.252	0.010	0.066	1.185	0.077	0.097	0.086
Russia (p18-p20)	1.378	0.005	0.117	1.278	0.092	0.108	0.112
China (p21, p22)	1.299	0.065	0.067	1.226	0.155	0.138	0.122
All_Continents	1.311	0.022	0.085	1.230	0.102	0.111	0.105

Na = Mean number of different alleles per locus; *PA* = Mean number of alleles unique to a single population per locus; *LCA* = Mean number of locally common alleles found in 25% or fewer populations per locus; *AR* = Allelic richness; H_O = Observed heterozygosity; H_E = Expected heterozygosity; uH_E = Unbiased expected heterozygosity.

[1]All values in the table represent means calculated from the data in Table S1, where a detailed description of the genetic diversity of each population is provided.

[2]Populations P7 and P8 were excluded due to the detection of mixed clusters. See the main text for further details.

15–60° N, based primarily on data from the GBIF database, supplemented by additional botanical references and online resources including Yakubov (1996), Oh & Pak (2001), Gu *et al.* (2003), Naruhashi (2001), the Science Museum Net (http://science-net.kahaku.go.jp/), and the China Virtual Herbarium (https://www.cvh.ac.cn/). The initial dataset, consisting of 949 occurrence records with voucher specimens, was rarefied using a 100 km distance filter with the R package spThin v.0.2.0 (Aiello-Lammens *et al.* 2015) to reduce spatial autocorrelation, resulting in 77 records for the present ENM simulations. A set of 19 bioclimatic variables at 2.5 arc-min resolution for the world under current conditions (1970–2000) was downloaded from WorldClim v.2.1 (www.worldclim.org; Fick & Hijmans 2017). Raster files of climate data within the study range were extracted using QGIS v.3.2.4 (QGIS Development Team 2022). Highly correlated variables (R > 0.8) were eliminated to prevent overfitting, using the R package ENMTools v.1.0.4 (Warren *et al.* 2021). Additionally, the variance inflation factor (VIF) was calculated using the R package usdm v.2.1-7 (Naimi *et al.* 2013, Naimi, 2017) to evaluate multicollinearity among climatic predictors, and variables with VIF > 3 were excluded. Finally, four of the 19 bioclimatic variables climate data were retained, including mean temperature of warmest quarter (Bio10), mean temperature of coldest quarter (Bio11), precipitation of warmest quarter (Bio18), and precipitation of coldest quarter (Bio19). Simulations of potential species habitats during the LGM were conducted using the Community Climate System Model v.4 (CCSM4; Gent *et al.* 2011) and the Model for Interdisciplinary Research on Climate (MIROC; Watanabe *et al.* 2011) at a 2.5 arc-minute resolution. MaxEnt v.3.4.4 (Phillips *et al.* 2017) was used with the four Bioclim layers and the occurrence data. A 20-fold cross-validation procedure was applied to test model performance. The area under the receiver operating characteristic curve (AUC) was used to evaluate model performance (Fielding & Bell 1997). Other parameters were set to default values in the MaxEnt analyses.

Results

Genetic diversity based on MIG-seq data

The genetic diversity of each *Potentilla discolor* population is detailed in Table S1, while the diversity among regional populations and between countries is summarized in Table 2. Indicative of the degree of genetic uniqueness, the mean number of private alleles (*PA*) and locally common alleles (*LCA*) in the continental populations (PA; 0.004–0.087, LCA; 0.046–0.133) was significantly higher than in Japanese populations (PA; 0–0.005, LCA; 0.023–0.041) ($p < 0.001$).

Values of *AR* and heterozygosities (H_E and uH_E), indicating genetic diversity, were significantly higher in continental populations (*AR*; 1.143–1.350, H_E; 0.091–0.165, uH_E; 0.074–0.133) compared to Japanese populations (*AR*; 1.051–1.247, H_E; 0.081–0.095, uH_E; 0.046–0.099) ($p < 0.001$). The fixation index (F_{IS}) of the population from Heilongjiang Province, China (p21) exhibited a relatively large deviation from zero (-0.693). However, no populations showed significant deviations in the HWE tests.

The global AMOVA showed that ~35 % of the variation was distributed among populations ($F_{ST} = 0.349$, $p < 0.001$), indicating substantial differentiation between populations (Table S2). The hierarchical AMOVA showed that ~16% of the variation was among groups (continental vs. Japanese populations), indicating moderate differentiation ($F_{CT} = 0.161$, $p < 0.001$).

Phylogenetic analysis based on MIG-seq data

In the ML tree of *Potentilla discolor* individuals and an outgroup using Dataset 1 (Fig. 2), the populations from northeastern Asia (p15–p21) and central China (p22) were paraphyletic to the clade of Japanese populations, although relationships among the continental populations remained ambiguous due to low bootstrap probabilities (BPs). Nevertheless, most populations were depicted as monophyletic. In the continental populations, three Far East Russian populations (p18–p20) formed a clade with high BP (91%), and this clade was sister to the Heilongjiang population (p21) in China with 65% BP. The three South Korean populations (p15–p17) made a single clade with 69% BP, and the clade was sister to the clade of Far East Russian populations and Heilongjiang populations with 29% BP. The Hubei population (p22) in China was positioned at the most basal part of the phylogenetic tree, albeit with limited statistical support.

On the other hand, the Japanese populations (p1, p2, p4–p8, p10–p12, and p14) were included in a single clade with high BP (98%). In the Japanese clades, two major clades (Clades I and II) were identified: Clade I, comprising the three Hyogo populations in mid-western Honshu with 98% BP, and Clade II, consisting of all Kyushu and Yamaguchi populations with 95% BP. In the latter clades, furthermore, two sub-clades (Clades III, and IV) were identified: Clade III, consisting of the northern Kyushu populations [Oita and Fukuoka, including most individuals from p7 and two from p8 (p8_12 and p8_13) populations] with 100% BP; and Clade IV, consisting of the other Kyushu and Yamaguchi populations (p5, one individual from p7 [p7_13], the majority of p8, and p10–p12, p14 populations) with 98% BP. The relationships within Clade IV were unclear, although two populations from Miyazaki (p12 and P14) in southern Kyushu formed a single clade (Clade V) with high BP (100%).

Genetic structure based on MIG-seq data

In the STRUCTURE analysis, the log-likelihood of the data, LnP(*K*), increased gradually as *K* increased from 2 to 10 (Fig. 3A). The ΔK values showed the highest peak at $K = 2$ (Fig. 3B). Five bar-plot diagrams are presented for $K = 2$, 3, 5, 7, and 10 (Fig. 3C). At $K = 2$, clear genetic differentiation was observed between the Japanese (p1, p2, p4, p5–p8, p10–p12, and p14) and continental (p15–p22) populations. At $K = 3$, the Hyogo populations (p1, p2, and p4) were distinguished from the other Japanese populations. At $K = 5$, the Miyazaki populations (p12 and p14) of Kyushu belonged to a different cluster. Furthermore, in the continental populations, genetic differentiation was observed between the South Korean populations (p15–p17), and the Far East Russian & Chinese populations (p18–p22). At $K = 7$, the two Chinese populations (p21 and p22) were assigned to separate clusters. At $K = 10$, in addition to the clusters observed at $K = 7$, the Fukuoka population in northern Kyushu (mainly p7) and the northern population in South Korea (p17) also formed distinct clusters.

The expected heterozygosity within each cluster identified in the STRUCTURE analysis indicated relatively high genetic diversity in the continental populations compared to the Japanese populations. At $K = 2$, the expected heterozygosity of the cluster representing the Japanese populations was 0.1158, while that of the conti-

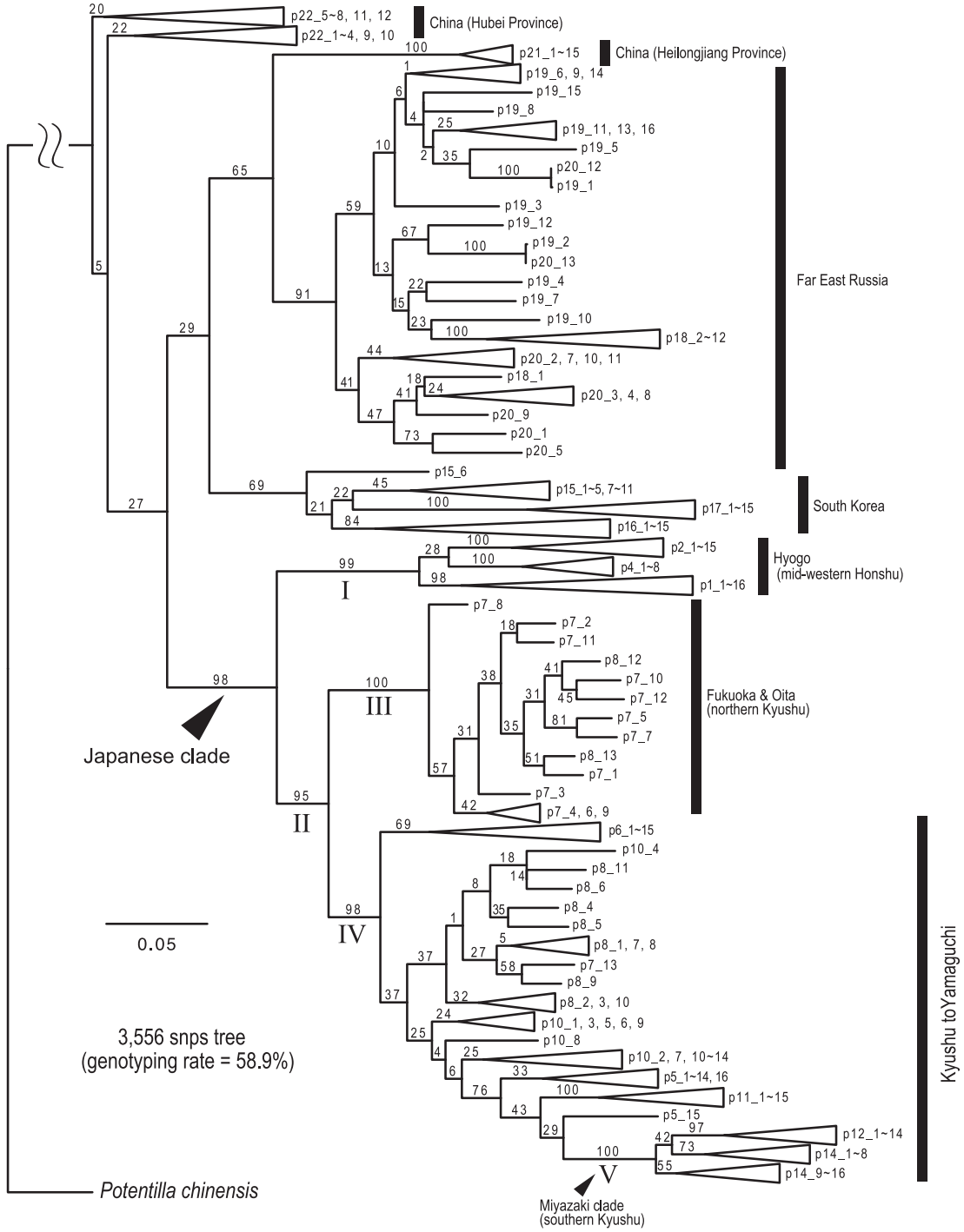

FIG. 2. Maximum likelihood (ML) tree based on 3,556 SNPs of 263 individuals from 22 populations of *Potentilla discolor*. The numbers along the branches represent bootstrap probabilities (%) derived from 1000 replicates. Population numbers correspond to those shown in Table 1. A clade consisting of individuals from a single population is collapsed into a triangle. The major lineages of the Japanese populations are represented using Roman numerals (Clades I–V).

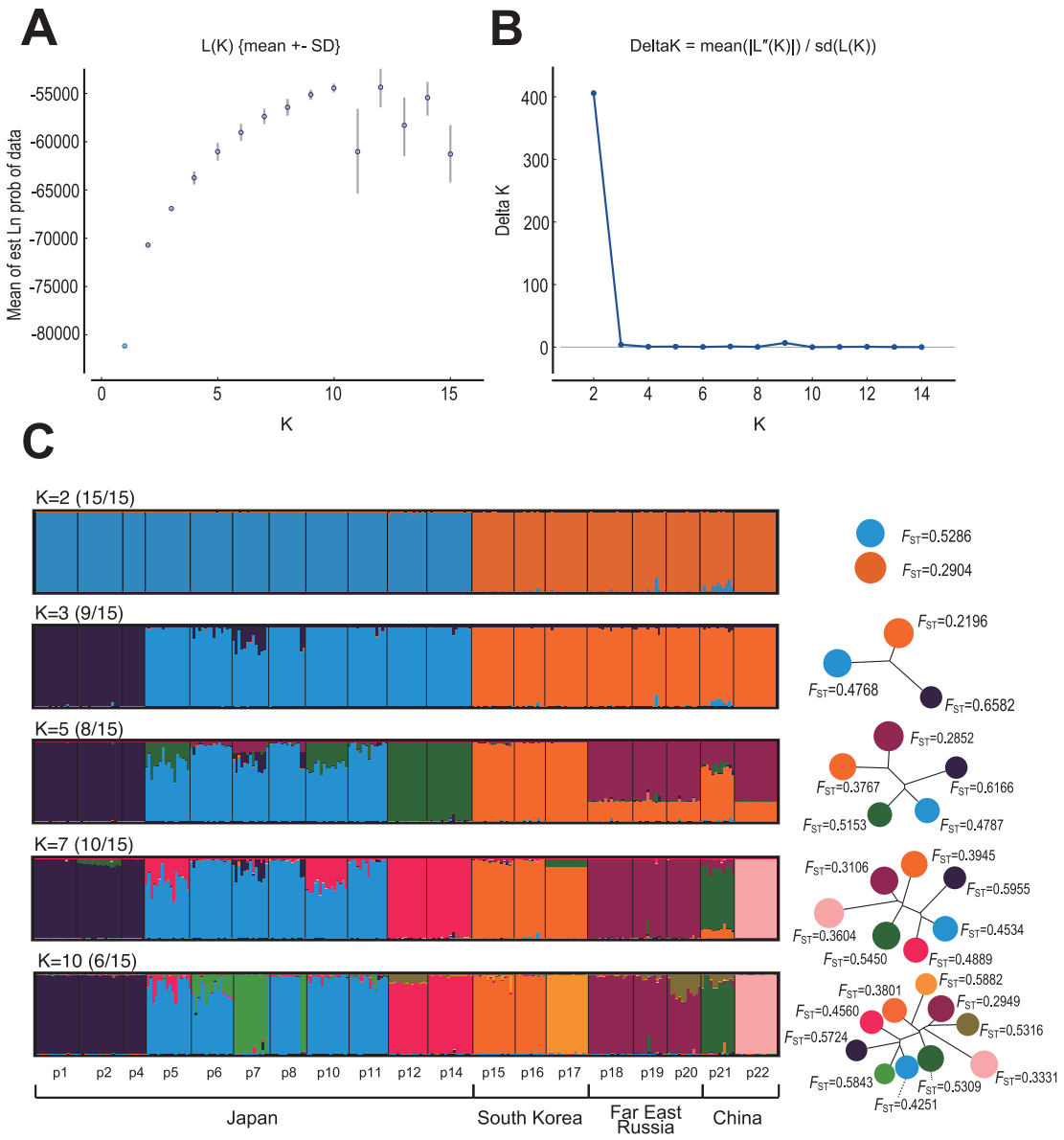

FIG. 3. Result of Bayesian clustering analyses in STRUCTURE of *Potentilla discolor*. **A**, Mean L(K) ± SD with 15 runs. **B**, Plots of delta-K values against $K = 2–14$. **C**, Bar plots show estimated probabilities of ancestral clusters of each sample. Numbers in parentheses indicate number of times the major pattern was observed in 15 iterations. Populations and their geographic areas are shown below bar plot. On right side of bar plots, the F_{ST} value of each cluster and amount of drift of each cluster from a common ancestral population is indicated. Dendrogram based on net nucleotide distance between clusters was drawn from the run with the highest likelihood. Circle size represents expected heterozygosity value within each cluster.

nental populations was 0.1455 (Fig. 3C). Among the Japanese populations, the clusters corresponding to the Hyogo populations (p1, p2, and p4) exhibited particularly low values ($K = 3$, 5, 7, and 10: 0.0804–0.0806). In contrast, within the continental populations, the clusters related to the Chinese and Russian populations (p18–p22) displayed higher expected heterozygosity compared to those corresponding to the South Korean populations (p15–p17) (K = 5, 7, and 10: 0.1216–

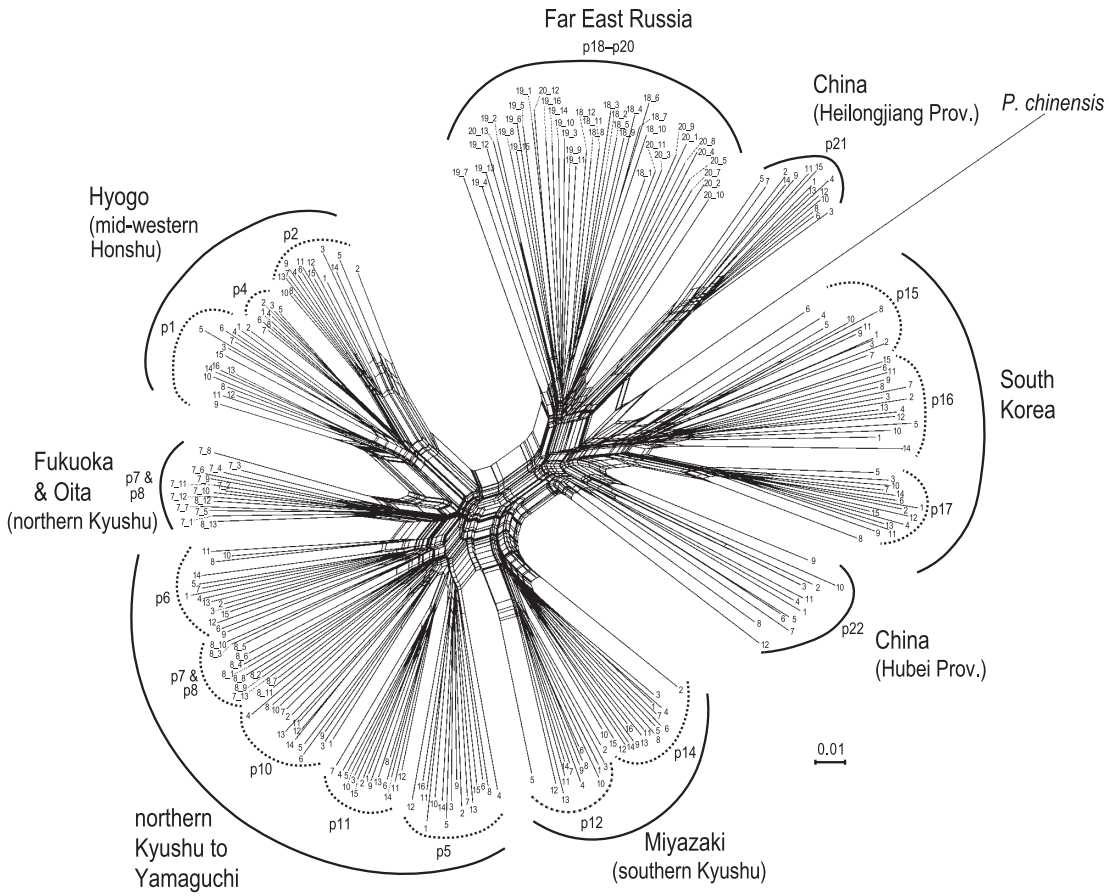

FIG. 4. NeighborNet network of 19 populations of *Potentilla discolor* constructed using SplitsTree. Individual and population numbers are omitted in this figure. When individuals within each population are clustered together, only individual numbers are indicated. Conversely, when interpopulation mixing is observed, both population numbers and individual numbers are shown.

0.1438 vs. 0.0840–0.1174).

Furthermore, the F_{ST} value for each cluster, representing the extent of genetic drift from a shared ancestral population, was relatively high in Japanese populations compared to the continental populations (K = 2: 0.5286 vs. 0.2904) (Fig. 3C). Among the Japanese populations, F_{ST} values were notably high in the Hyogo populations (K = 3, 5, 7, and 10: 0.5724–0.6582). Among the clusters of continental populations, the Heilongjiang population in China (p21) and the northern population in South Korea (p17) exhibited relatively higher F_{ST} values compared to those in Far East Russia (p18–p20), the southern population in South Korea (p15, p16), and the Hubei population in China (p22) (K = 7 and 10: 0.5309–0.5882 vs. 0.2949–0.3604).

The NeighborNet network revealed population structures similar to those inferred by STRUCTURE and ML tree (Fig. 4). Two major clusters were identified: one comprising the Japanese populations and the other the continental populations. Within the Japanese cluster, four subclusters were identified: (1) the Hyogo populations (p1, p2, and p4) in mid-western Honshu; (2) the Fukuoka and Oita populations (all individuals in p7 except p7_13, and p8 except p8_12 and p8_13) in northern Kyushu; (3) populations spanning northern Kyushu to Yamaguchi in westernmost Honshu (p5, p6, p7_13, p8_12, p8_13, p10, and p11); and (4) the Miyazaki populations (p12 and p14) in southern Kyushu. In the continental

TABLE 3. Results of DIYABC-RF for scenario choice analyses.

First-step analysis using Japanese and continental populations										
Global error rate	Local error rate	Votes scenario[1]								Posterior probability
		1	2	3	4	5	6	7	8	
0.049	0.178	705	60	67	136	8	2	4	18	0.822 [Scen. 1]
Second-step analysis focusing on the Japanese populations										
Global error rate	Local error rate	Votes scenario[1]								Posterior probability
		A	B	C						
0.198	0.254	548	227	225						0.746 [Scen. A]

[1]For each scenario, see Figs. S1 & S2.

TABLE 4. Priors and parameter estimates of the selected scenarios in the DIYABC-RF analyses.

Priors for the parameters			Estimates for the parameters			
Type	Parameters	Range	Mean	Median	5% quantiles	95% quantiles
First-step analysis using Japanese and continental populations (Scen. 1; see Fig. 5 & Table 3)						
Population size	N1	100–100,000	43,892	41,655	24,915	71,086
	N2	100–100,000	37,436	35,059	20,811	60,823
	N3	100–100,000	50,237	47,399	25,790	81,821
	N4	100–100,000	44,136	41,072	19,666	78,475
	N5	100–100,000	49,076	47,913	24,083	78,096
	Na1	100–100,000	60,050	62,570	15,026	96,892
	Na2	100–100,000	56,248	55,919	13,531	96,536
	Nanc	100–200,000	169,501	176,598	119,897	197,901
Time of event	t1	100–50,000	12,582	11,091	1,109	30,582
	t2	100–50,000	23,221	22,101	11,718	35,844
	t3	100–100,000	25,243	24,698	13,457	38,452
	t4	100–150,000	40,787	40,027	24,670	61,176
Second-step analysis focusing on the Japanese populations (Scen. A; see Fig. 5 & Table 3)						
Population size	N1	100–50,000	9,624	8,712	3,363	20,237
	N2	100–50,000	13,271	11,896	4,264	24,500
	N3	100–50,000	7,381	6,396	1,420	16,948
	N4	100–50,000	27,610	28,259	12,641	42,389
	Na1	100–100,000	50,166	46,078	10,035	94,027
	Na2	100–100,000	43,628	39,515	9,895	92,805
	Nanc	100–100,000	88,418	90,972	67,630	99,865
Time of event	t1	100–50,000	4,506	4,194	1,227	9,408
	t2	100–50,000	9,870	9,194	3,315	18,869
	t3	100–100,000	21,364	20,681	10,389	34,345

cluster, four regional subclusters were identified: (1) the Far East Russian populations (p18–p20); (2) the South Korean populations (p15–p17); (3) the Heilongjiang population in China (p21); and (4) the Hubei population in China (p22). The outgroup accession, *Potentilla chinensis*, was linked to the cluster of continental populations.

DIYABC analysis based on MIG-seq data

Before conducting the main analyses in DI-

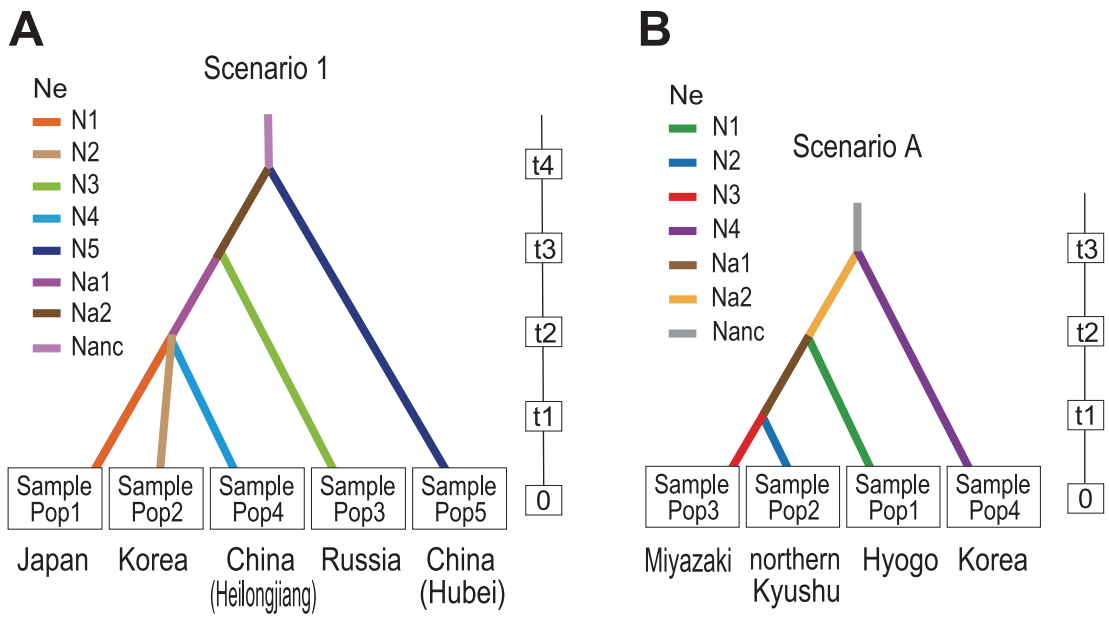

FIG. 5. Most supported demographic scenarios in the DIYABC-RF analyses. For a comparison with other scenarios, see Tables 3 & 4, and Figs. S1 & S2. **A**, First-step analysis using continental and Japanese populations classified into five groups: Pop1 (Japan), Pop2 (South Korea), Pop3 (Far East Russia), Pop4 (Heilongjiang Province, China), and Pop5 (Hubei Province, China). Parameters t# (t1–t4) represent timescale in generations; N# indicates effective population size for the following populations: Nanc (ancestral population of all groups), Na1 (ancestral population of Pop1–Pop3), Na2 (ancestral population of Pop1–Pop4), and N1–N4 (Pop1–Pop4 after divergence of each regional population). Parameter t1 in this scenario was set, for the sake of adjustment, to align with the number of divergences in other scenarios. **B**, Second-step analysis focusing on Japanese populations, with inclusion of South Korean populations as a comparative group. Populations were classified into four groups: Pop1 (Hyogo), Pop2 (northern Kyushu), Pop3 (Miyazaki), and Pop4 (South Korea). Parameters t# (t1–t3) represent timescale in generations; N# indicates effective population size for the following populations: Nanc (ancestral population of all groups), Na1 (ancestral population of Pop2 and Pop3), Na2 (ancestral population of Pop1–Pop3), and N1–N4 (Pop1–Pop4 after divergence of each regional population).

YABC-RF, we performed prior scenario checking using Principal Component Analysis (PCA). This analysis encompasses two datasets: one covering the entire distribution range and the other exclusively focusing on the Japanese populations. In both cases, the observed datasets were positioned within the clouds of simulated datasets (Figs. S1 & S2), indicating that the conditions defined for each analysis were suitable for random forest analysis.

In the first-step analysis, which included both the Japanese and continental populations, Scenario 1 was identified as the best-supported scenario, receiving 705 out of 1,000 votes with a posterior probability of 0.822 (Table 3, Fig. 5A, Fig. S1). This scenario assumes that the Hubei population in continental East Asia diverged first, followed by the divergence of the Far East Russian populations, and finally, the Heilongjiang, South Korean, and Japanese populations split into three lineages. In addition to the main analysis, preliminary analyses were conducted to test three alternative patterns in which the Heilongjiang, South Korean, and Japanese populations split in different orders. However, Scenario 1 received more votes than any of these alternative scenarios (data not shown). Subsequently, we conducted parameter estimation for Scenario 1, which received the highest number of votes (Table 4). The median effective population sizes for the five defined regional populations (N1–N5) ranged from 35,059 (90% confidence interval [CI]: 20,811–60,823) to 47,913 (CI: 24,083–78,096). The values for the populations in Japan, South Korea, and Heilongjiang Province were comparatively lower than those for Far East Russia and Hubei Prov-

ince (N1, N2, N4 < N3, N5). The median estimated divergence times (t1–t4) ranged from 11,091 generations ago (CI: 1,109–30,582) to 40,027 generations ago (CI: 24,670–61,176).

In the second-step analysis focusing on the Japanese populations, a comparison across the three scenarios indicated that Scenario A was the most favorable, receiving 548 out of 1,000 votes and achieving a posterior probability of 0.746 (Table 3, Fig. 5B, Fig. S2). However, the number of votes for Scenarios B and C were also relatively high (227, 225/1000 votes, respectively). We proceeded with the parameter estimation analysis for Scenario A with the highest number of votes (Table 4). The median values of effective population size of the four defined regional populations (N1–N4) ranged from 6,396 (CI: 3,363–20,237) to 28,259 (CI: 12,641–42,389).

Distribution and network of chloroplast DNA haplotypes

Using the *trnK* 5′-terminal region of cpDNA, 14 haplotypes (Types A–N) were detected (Table S3). There were four indels and 16 site changes among the 1,060 sites resulting from alignment. The two most common haplotypes were A (n = 82) and D (n = 78), followed by B (n = 25), F (n = 24), and K (n = 14). The remaining 10 haplotypes (C, E, G–N) were minor, with sample sizes ranging from 1 to 4. The four predominant haplotypes (A, B, D, and F) were on both the continent and in Japan (Fig. 6A). In contrast, the remaining haplotypes were unique to each population: Type C in Japan (p8); Types G, H, and K in South Korea (p15, p17); and Types I, J, L, M, and N in China (p21, p22). The MJ network illustrating relationships among these haplotypes is shown in Fig. 6B. The dominant haplotypes A and D were connected by four mutational steps via Types I and L.

In continental populations, twelve haplotypes (Types A, B, D, and F–N) were identified, with a haplotype diversity (*Hd*) of 0.832 and a nucleotide diversity (π) of 0.00283. In the Japanese populations, six haplotypes (Types A–F) were identified, with a haplotype diversity (*Hd*) of 0.617 and a nucleotide diversity (π) of 0.00212. The F_{ST} value, an index of genetic differentiation between Japanese and continental populations, was 0.02861 ($p = 0.260$), indicating no significant genetic differentiation between the groups.

Inference of distribution patterns by ENM

The MaxEnt models accurately predicted the current distribution pattern of *Potentilla discolor* (Fig. 7), achieving an AUC of 0.853 on average (range: 0.715–0.934), indicating the model's reliability. Climate variables contributed to the model as follows: mean temperature of the coldest quarter (Bio11, 57.2%), mean temperature of the warmest quarter (Bio10, 20.3%), precipitation of the warmest quarter (Bio18, 17.4%), and precipitation of the coldest quarter (Bio19, 5.1%).

To estimate the potential distribution during the LGM, the models were projected onto two paleoclimate scenarios (CCSM4 and MIROC) (Fig. 7). Both projections yielded similar predicted distribution patterns. High-probability suitable areas (> 0.75) were widely distributed across central and eastern China (*ca*. 25°N to 35°N). Additionally, suitable habitats were identified on the Pacific side of the Japanese Archipelago. In contrast, no suitable habitats were predicted in northeastern China, Far East Russia, or the Korean Peninsula.

Discussion

Genetic structure and migration history of P. discolor

The genetic structure analysis using MIG-seq data in this study clearly identified distinct clusters, separating Japanese populations from continental populations within *Potentilla discolor* (Figs. 3 & 4). The phylogenetic analysis revealed no evidence for monophyly of the continental populations as a whole, whereas the Japanese populations formed a monophyletic group with robust support (Fig. 2). The populations in the Japanese Archipelago, located at the eastern edge of the range of *P. discolor* and geographically isolated by the Tsushima Strait, likely experienced limited gene flow with continental populations, which may have contributed to a distinct genetic composition (Fig. 1).

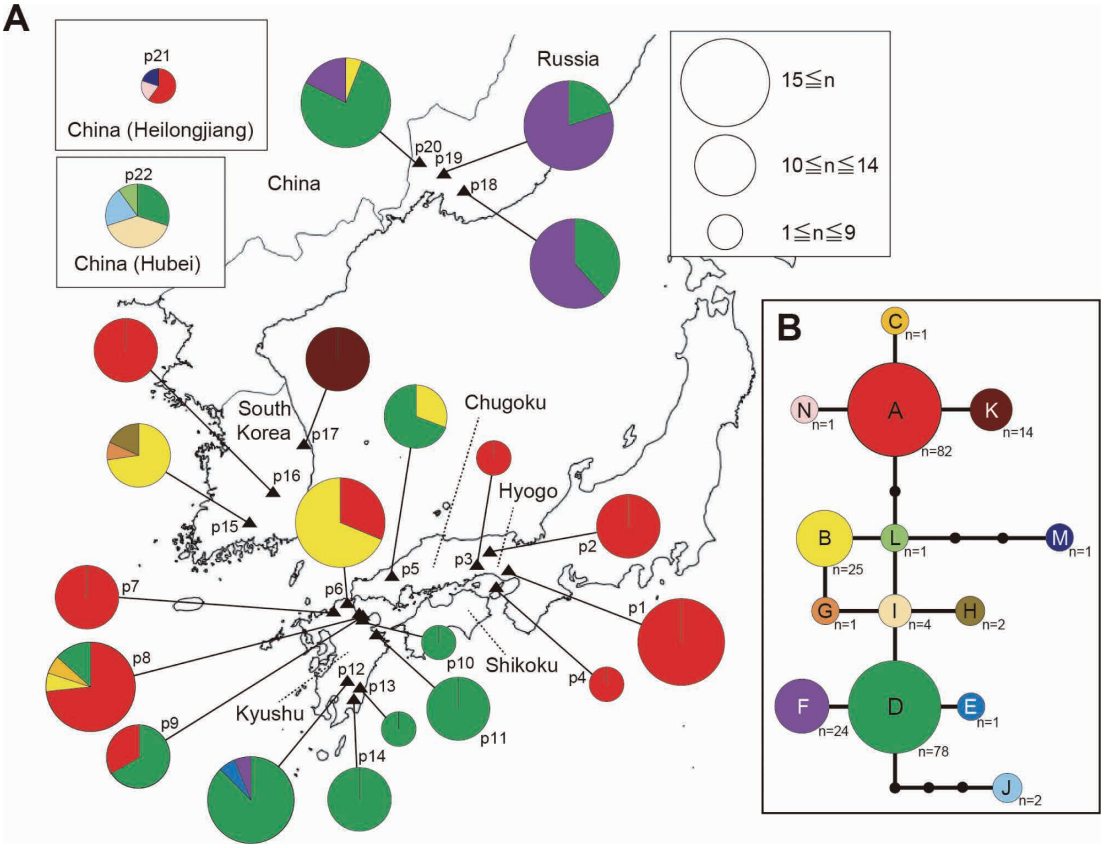

FIG. 6. Geographic distribution of chloroplast DNA haplotypes in *Potentilla discolor* (**A**) and median-joining network of haplotypes (**B**). Population numbers correspond to those listed in Table 1. In panel A, size of each circle represents number of samples per population. As indicated in upper right, the circles are categorized into three distinct groups. For locations of Chinese populations (p21 and p22), refer to Fig. 1.

The phylogenetic analysis in this study did not clearly resolve relationships among continental populations; however, regional clusters were clearly distinguishable, including those from South Korea, Far East Russia, Heilongjiang Province (China), and Hubei Province (China). Thus, to clarify the population demographic history of *Potentilla discolor*, we conducted DIYABC-RF analyses including the four regional populations and the Japanese populations. The results supported Scenario 1 (Table 3, Fig. 5A), which indicated that the population from Hubei Province, China, diverged first, followed by the Far East Russian populations. Therefore, the finding suggests that the initial diversification of *P. discolor* occurred within the eastern Asian continental populations, rather than between the Japanese and continental populations.

The ENM analysis during the LGM indicated that suitable distribution areas were widely present from central to eastern China, including within Hubei Province (Fig. 7). In the Late Pleistocene, four glacial-interglacial cycles are known to have occurred over the past 400,000 years (Lisiecki & Raymo 2005). Since we did not estimate the timing of the divergence in this study, it is not possible to specify which glacial period was involved in the continental differentiation of *P. discolor*. However, the results of the ENM analysis suggest that during a cold period such as the LGM, populations that had shifted their distribution to the southern continental region may have expanded their range northward during the subsequent post-glacial period.

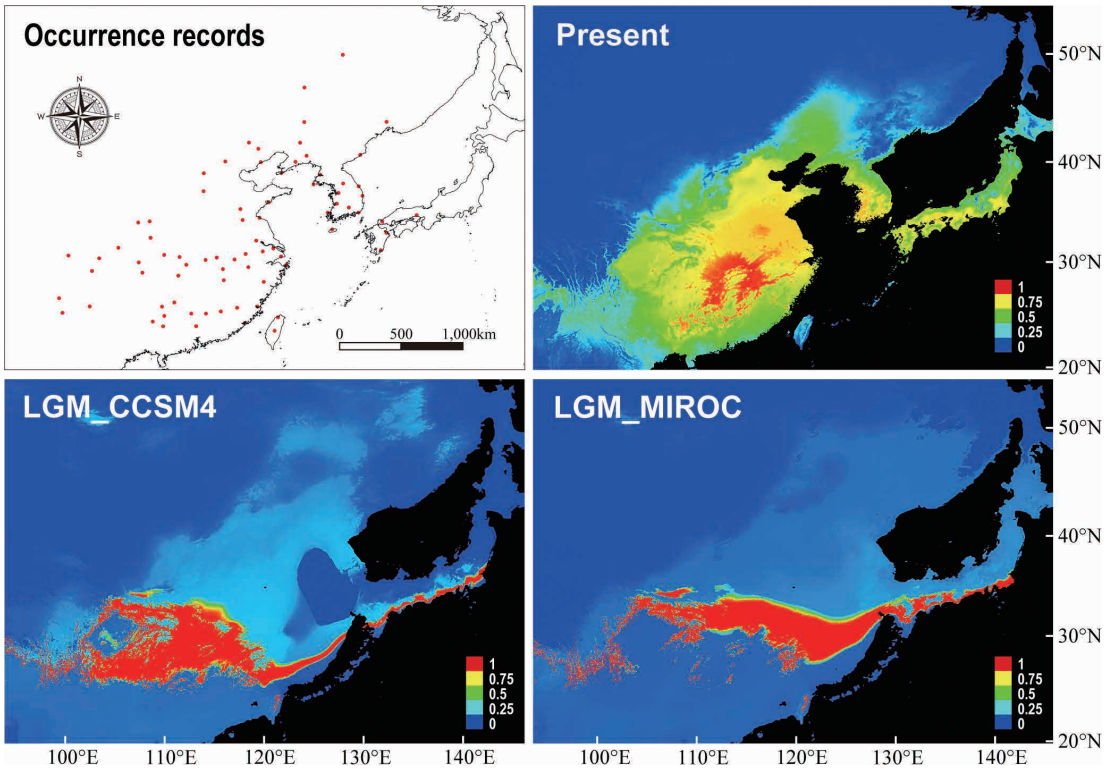

FIG. 7. Ecological niche models for *Potentilla discolor* at present and during Last Glacial Maximum (LGM) (CCSM and MIROC models, *ca*. 21,000 years BP). Color legend from blue to red indicates habitat suitability from low to high. In upper left figure, red dots indicate *P. discolor* occurrence records (77 points) used in this study.

For the populations from Heilongjiang Province (China), South Korea, and Japan, Scenario 1 assumes that they diverged into three distinct lineages from a common ancestor at approximately the same time after the Far Eastern Russian populations diverged (Table 3, Fig. 5A). However, the ENM analysis indicated that suitable habitat areas were largely absent on the Korean Peninsula and around Heilongjiang Province during the LGM (Fig. 7), implying that the current distribution in those regions was likely established after the LGM. Thus, the ancestral population may have expanded its range during the postglacial period and subsequently diverged into three lineages due to geographic isolation. But, given the complex population structure observed within the Japanese archipelago (as discussed in detail below), it is possible that their divergence occurred not after the LGM but rather during an earlier glacial period. Further investigation is required to clarify the timing and processes of population divergence and expansion in those regions.

Differentiation pattern within the Japanese Archipelago

Phylogenetic and genetic structure analyses revealed significant genetic differentiation between the Hyogo populations in mid-western Honshu and the Kyushu–Yamaguchi populations (Figs. 2–4). Furthermore, the southern Kyushu population from Miyazaki formed a distinct monophyletic group in the phylogenetic tree (Fig. 2) and a separate group in the cluster analyses (Figs. 3 & 4). The Fukuoka and Oita populations in northern Kyushu (p7, p8) will be discussed in detail in the following paragraphs. These findings regarding the Hyogo and Miyazaki populations suggest that, after the expansion of the distribution within the Japanese Archipelago, relict pop-

ulations became isolated at the edges of the range in mid-western Honshu and southern Kyushu. In the STRUCTURE analysis, the Hyogo and Miyazaki populations exhibited relatively high F_{ST} values (Fig. 3). The results suggest that strong genetic drift occurred in both populations, supporting the occurrence of isolation events likely associated with the bottleneck effect, which occurs when small isolated populations experience reduced genetic diversity. Additionally, demographic analysis supported a scenario (Scenario A) in which the Hyogo population (Pop1) diverged first, followed by the divergence of the northern Kyushu–Yamaguchi populations (Pop2) and the Miyazaki population (POP3) (Table 3, Fig. 5B). This result further corroborates the aforementioned scenario for genetic differentiation in the Japanese populations.

In the Fukuoka/Oita populations (p7, p8) in northern Kyushu, two distinct lineages (Clades III and IV) or genetic clusters were observed in each population (Figs. 2–4), thereby indicating that genetically divergent individuals occur sympatrically (*i.e.*, within the same geographic location) within a single population. The two populations were approximately 75 km apart, with little continuous habitat between them. Thus, it is unlikely that gene flow occurred frequently between them. The results may reflect the complex evolutionary history of the populations in northern Kyushu. However, the potential for experimental handling errors, such as sample contamination, cannot be excluded and further discussion on this matter is therefore limited. However, the presence of at least two distinct lineages or genetic clusters in northern Kyushu suggests that there may have been a broader range of suitable habitats for this species in the region in the past.

In our previous studies using MIG-seq and ddRAD-seq analyses with plants of the Mansen elements, genetic differentiation within the Japanese Archipelago has also been observed in *Viola orientalis* and *Tephroseris kirilowii* (Sata *et al.* 2021, Sakaba *et al.* 2023). However, genetic differentiation between the Kyushu and Honshu populations in *V. orientalis* was small. Although it was observed in the phylogenetic analysis, no differentiation was detected within Japan in the STRUCTURE analysis. Furthermore, in *T. kirilowii*, the STRUCTURE analysis revealed genetic differentiation in the southern Kyushu populations (Miyazaki & Kagoshima Pref.), while the genetic composition of populations from other regions was uniform. Considering that four clusters were recognized in *Potentilla discolor*, it is possible that the patterns of genetic differentiation within Japan differ significantly among species. Factors contributing to these differences may include variation in the phylogeographic history of each species. However, ecological factors, such as differences in habitat and seed dispersal patterns, are also likely to have played a significant role. For example, species of *Viola*, including *V. orientalis*, are myrmecochorous plants that produce seeds with elaiosomes, facilitating dispersal by ants (Beattie & Lyons, 1975), and *T. kirilowii* (Asteraceae) is considered to rely on wind dispersal through seeds equipped with a pappus. In contrast, the achenes of *P. discolor* are small and lack specialized dispersal appendages, indicating that they rely primarily on gravity for dispersal, which likely limits their ability to colonize distant or isolated habitats. The factors underlying the regional differentiation of *P. discolor* within Japan may therefore be closely linked to its limited dispersibility and the ecological constraints imposed by the seed dispersal mechanisms.

Origin of the Mansen elements

Previous phytogeographic studies hypothesized that species of the Mansen element originated on the Asian continent and migrated to Japan via the Korean peninsula during the cold climate of the Pleistocene (Kitamura 1957, Hotta 1974, Murata 1977, 1988, Tabata 1997). The continental origin of the Mansen elements was supported by a previous study on *Viola orientalis* (Sata *et al.* 2021). In the present study of *Potentilla discolor*, the ML phylogenetic tree did not provide direct evidence for a continental origin of the species (Fig. 2). However, demographic analysis suggested that populations from Hubei Province in China and Far East Russia contributed to the early diversification of the species (Scenario

1, Fig. 5A). In contrast, the scenario in which the Japanese population diverged first (Scenario 8), was not supported (Table 3, Fig. S1). Additionally, the network analysis showed that the cluster of *Potentilla chinensis* used as an outgroup was connected with the continental population clusters (Fig. 4). These findings support a continental origin for the Mansen elements.

The parameter estimates supported by Scenario 1 in the demographic analysis suggests that divergence between the continental and Japanese populations of *P. discolor* occurred approximately 22,000 (90% credibility interval = 11,718–35,844) generations ago (t2, Table 4). Although the precise generation time of *P. discolor*, this small perennial herb, remains undetermined, it is likely to be less than 10 years based on its life history. If past climate changes are associated with this divergence, it is possible that the event occurred during the last glacial period or possibly the preceding glacial periods. While the absolute divergence time remains unresolved, the relatively recent divergence in the late Quaternary supports the traditional hypothesis concerning the migration history of the Mansen elements.

Additionally, the migration routes from the continent to Japan remained unresolved in this study. If the ancestral population migrated to Japan via the Korean peninsula, it would be expected that the Japanese and South Korean populations would be genetically closely related. However, the ML tree did not support this hypothesis, as it did not provide evidence that the Japanese and South Korean populations form a sister group (Fig. 2). Furthermore, Scenario 1, supported by the demographic analysis, suggests that the populations from Japan, South Korea, and Heilongjiang Province diverged almost simultaneously, though the precise migration route to Japan remains unclear (Fig. 5A). The ENM analysis indicated that during the LGM, suitable habitats for *P. discolor* extended from central China to the Pacific side of the Japanese Archipelago (Fig. 7), suggesting that migration may have been *via* through the exposed continental shelf now occupied by the East China Sea. During the LGM, this region was exposed as land, and pollen analyses have revealed the expansion of grassland steppe vegetation (Kawahata & Ohshima 2004, Xu *et al.* 2010). Therefore, while the possibility of migration via this region cannot be excluded, the analysis based on the SNP data of *P. discolor* in this study did not provide results supporting this scenario. Regardless, given the limited sample size of continental populations in this study, increasing the number of continental populations in future analyses is necessary to better determine the migration route of *P. discolor* to Japan.

Discrepancies between chloroplast DNA and MIG-seq analysis data

MIG-seq phylogenetic analysis showed that the Japanese populations were distinctly monophyletic (Fig. 2). However, cpDNA analysis did not reveal genetic differentiation between Japanese and continental populations. For example, haplotype D, one of the dominant haplotypes, was detected in all regions of Japan, Russia, and China (Hubei) (Fig. 6). Moreover, the genetic differentiation between Japanese and continental populations was significant based on F_{ST} values from the SNP data but not for cpDNA. Discrepancies between nuclear and chloroplast lineages are a relatively well-known phenomenon in plants, primarily due to incomplete lineage sorting based on ancestral polymorphisms, interspecific hybridization, or intraspecific hybridization among genetic groups (Wendel & Doyle 1998, Currat *et al.* 2008, Joly *et al.* 2009, Petit & Excoffier 2009, Pelser *et al.* 2010). One possibility is that populations of *P. discolor* that retained continental ancestral polymorphisms may have entered the Japanese Archipelago followed by genetic drift subsequently fixing specific haplotypes in each population. Regardless, the haplotype distribution pattern of the cpDNA showed limited geographic structure, indicating that phylogeographic information is sparse. Therefore, our study focused on determining the migration history of *P. discolor* based on MIG-seq data.

We would like to express our sincere gratitude to Tatsundo Fukuhara, Hiroshi Ikeda, Nobuki Kawano, Yoshiki Kobayashi, Tadashi Minamitani, Hajime Ono, Yoko Ota,

Yoshihiro Sawada, Atsuko Takano, Hirofumi Tsuji, Kouta Uki and Masashi Yokogawa (listed in alphabetical order, title omitted) for their assistance in collecting plant samples in Japan. We are also deeply grateful to Jiahao Shen and Tao Feng for their support in collecting samples in China, and to Woong Lee, Jin Suk Youn and Jung-Sim Lee for their help in South Korea. Our thanks go to Mitsuhiko Sato for his assistance with the MIG-seq experiments. We would like to extend our appreciation to all the members of our laboratory for their ongoing support of this research. Finally, we are grateful to an anonymous referee for valuable feedback during the review of our manuscript. This research was supported by JSPS KAKENHI Grant Number JP17H03721 and grants from the Institute of Plant Science and Resources, Okayama University (3037, 3140, & R236).

Supplementary material: The online version of this article (doi: 10.18942/apg.202508) contains supplementary material: Tables S1–S3, Figs. S1 & S2.

References

Aiello-Lammens, M. E., R. A. Boria, A. Radosavljevic, B. Vilela & R. P. Anderson. 2015. spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. Ecography 38: 541–545.

Avise, J. C. 2000. Phylogeography: the history and formation of species. Harvard University Press, Massachusetts.

Avise, J. C. 2004. Molecular markers, natural history, and evolution, Second edition. Sinauer Associates, Inc., Sunderland, Massachusetts.

Axelrod, D. I., I. Al-Shehbaz & P. H. Raven. 1996. History of the modern flora of China. *In*: Zhang, A. L. & S. G. Wu (eds.), Floristic characteristics and diversity of East Asian Plants, pp. 43–55. Springer, New York.

Bandelt, H.-J., P. Forster & A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16: 37–48.

Batchelor, C. L., M. Margold, M. Krapp, D. K. Murton, A. S. Dalton, P. L. Gibbard, C. R. Stokes, J. B. Murton & A. Manica. 2019. The configuration of Northern Hemisphere ice sheets through the Quaternary. Nat. Commun. 10: 3713.

Bates, D., M. Mächler, B. Bolker & S. Walker. 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67: 1–48.

Beattie, A. J. & N. Lyons. 1975. Seed dispersal in *Viola* (Violaceae): adaptations and strategies. Amer. J. Bot. 62: 714–722.

Bradbury, P. J., Z. Zhang, D. E. Kroon, T. M. Casstevens, Y. Ramdoss & E. S. Buckler. 2007. TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics 23: 2633–2635.

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

Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores & W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. Mol. Ecol. 22: 3124–3140.

Collin, F. D., G. Durif, L. Raynal, E. Lombaert, M. Gautier, R. Vitalis, J. M. Marin & A. Estoup. 2021. Extending approximate Bayesian computation with supervised machine learning to infer demographic history from genetic polymorphisms using DIYABC Random Forest. Mol. Ecol. Resour. 21: 2598–2613.

Currat, M., M. Ruedi, R. J. Petit & L. Excoffier. 2008. The hidden impact of invasions: significant introgression of local genes. Evolution 62: 1908–1920.

Delgado, L., F. Gallego & E. Rico. 2000. Karyosystematic study of *Potentilla* L. subgen. *Potentilla* (Rosaceae) in the Iberian Peninsula. Bot. J. Linn. Soc. 132: 263–280.

Doyle, J. J. & J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull. 19: 11–15.

Earl, D. A. & B. M. vonHoldt. 2012. STRUCTURE HARVESTER: a website and program for visualizing structure output and implementing the Evanno method. Conserv. Genet. Resour. 4: 359–361.

Edler, D., J. Klein, A. Antonelli, D. Silvestro & M. Matschiner. 2020. raxmlGUI 2.0: A graphical interface and toolkit for phylogenetic analyses using RAxML. Methods Ecol. Evol. 12: 373–377.

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

Excoffier, L. & H. E. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. Mol. Ecol. Resour. 10: 564–567.

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

Feng, T., M. J. Moore, M. H. Yan, Y. X. Sun, H. J. Zhang, A. P. Meng, X. D. Li, S. G. Jian, J. Q. Li & H. C. Wang. 2017. Phylogenetic study of the tribe Potentilleae (Rosaceae), with further insight into the disintegration of *Sibbaldia*. J. Syst. Evol. 55: 177–191.

Fick, S. E. & R. J. Hijmans. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. Int. J. Climatol. 37: 4302–4315.

Fielding, A. H. & J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. Environ. Conserv. 24: 38–49.

Fujii, N., K. Ueda, Y. Watano & T. Shimizu. 1997. Intra-

specific sequence variation of chloroplast DNA in *Pedicularis chamissonis* Steven (Scrophulariaceae) and geographic structuring of the Japanese "Alpine" plants. J. Plant Res. 110: 195–207.

GBIF.org. 2021. GBIF Home Page. <https://www.gbif.org> [accessed Jul. 11, 2021].

Gent, P. R., G. Danabasoglu, L. J. Donner, M. M. Holland, E. C. Hunke, S. R. Jayne, D. M. Lawrence, R. B. Neale, P. J. Rasch, M. Vertenstein, P. H. Worley, Z.-L. Yang & M. Zhang. 2011. The Community Climate System Model Version 4. J. Clim. 24: 4973–4991.

Good, R. 1974. The geography of the flowering plants, 4th edition. Longman, London.

Gu, C., C. Li, L. Lu, S. Jiang, C. Alexander, B. Bartholomew, A. R. Brach, D. E. Boufford, H. Ikeda, H. Ohba, K. R. Robertson & S. A. Spongberg. 2003. Rosaceae. *In*: Zhengyi, W., P. H. Raven & H. Deyuan (eds.), Flora of China. Vol. 9, Missouri Botanical Garden Press, St. Louis.

Harrison, S. P., G. Yu, H. Takahara & I. C. Prentice. 2001. Diversity of temperate plants in east Asia. Nature 413: 129–130.

Hotta, M. 1974. Evolutionary biology in plants III, History and geography of plants. Sanseido, Tokyo (in Japanese).

Huson, D. H. & D. Bryant. 2006. Application of phylogenetic networks in evolutionary studies. Molec. Biol. Evol. 23: 254–267.

Ikeda, H. 2016. *Potentilla* L. *In*: Ohashi, H., Y. Kadota, J. Murata, K. Yonekura, H. Kihara (eds.), Wild flowers of Japan, pp. 33–40. Heibonsha, Tokyo. (in Japanese).

Ikeda, H., H. Higashi, V. Yakubov, V. Barkalov & H. Setoguchi. 2014. 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., V. Yakubov, V. Barkalov, K. Sato & N. Fujii. 2020. East Asian origin of the widespread alpine snow-bed herb, *Primula cuneifolia* (Primulaceae), in the northern Pacific region. J. Biogeogr. 47: 2181–2193.

Ikeda, H., V. Yakubov, V. Barkalov & H. Setoguchi. 2018. Post-glacial East Asian origin of the alpine shrub *Phyllodoce aleutica* (Ericaceae) in Beringia. J. Biogeogr. 45: 1261–1274.

Jordan, W. C., M. W. Courtney & J. E. Neigel. 1996. Low levels of intraspecific genetic variation at a rapidly evolving chloroplast DNA locus in north American duckweeds (Lemnaceae). Amer. J. Bot. 83: 430–439.

Joly, S., P. A. McLenachan & P. J. Lockhart. 2009. A statistical approach for distinguishing hybridization and incomplete lineage sorting. Amer. Nat. 174: E54–E70.

Joyce, E. M., C. M. Pannell, M. Rossetto, J. Y. S. Yap, K. R. Thiele, P. D. Wilson & D. M. Crayn. 2021. Molecular phylogeography reveals two geographically and temporally separated floristic exchange tracks between Southeast Asia and northern Australia. J. Biogeogr. 48: 1213–1227.

Kato, M. 2011. Endemic plants of Japan. *In*: Kato, M. & A. Ebihara (eds.), A book series from the National Museum of nature and science No. 11, pp. 3–10. Tokai University Press, Kanagawa (in Japanese).

Kawahata, H. & H. Ohshima. 2004. Vegetation and environmental record in the northern East China Sea during the late Pleistocene. Glob. Planet. Change 41: 251–273.

Keenan, K., P. McGinnity, T. F. Cross, W. W. Crozier & P. A. Prodöhl. 2013. diveRsity: An R package for the estimation and exploration of population genetics parameters and their associated errors. Meth. Ecol. Evol. 4: 782–788.

Kitamura, S. 1957. Distribution of plants. *In*: Kitamura, S., G. Murata, M. Hori (eds.), Color. Illus. Herb. Plant Jap. (Sympetalae), pp. 246–264. Hoikusha, Osaka (in Japanese).

Koidzumi, G. 1931. Zengen (preface). *In*: Mayebara, K. (ed.), Florula Austro-Higoensis, Private publication (in Japanese).

Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg & I. Mayrose. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Mol. Ecol. Resour. 15: 1179–1191.

Kumar, S., G. Stecher, M. Li, C. Knyaz & K. Tamura. 2018. MEGA X: molecular evolutionary genetics analysis across computing platforms. Mol. Biol. Evol. 35: 1547–1549.

Kurata, S., S. Sakaguchi, S. K. Hirota, O. Kurashima, Y. Suyama & M. Ito. 2024. Phylogeographic incongruence between two related *Geranium* species with divergent habitat preferences in East Asia. Ecol. Res. 39: 273–288.

Kurata, S., S. Sakaguchi, H. Ikeda, S. K. Hirota, O. Kurashima, Y. Suyama & M. Ito. 2022. From East Asia to Beringia: reconstructed range dynamics of *Geranium erianthum* (Geraniaceae) during the last glacial period in the northern Pacific region. Plant Syst. Evol. 308: 28.

Lassmann, T., Y. Hayashizaki & C. O. Daub. 2009. TagDust–a program to eliminate artifacts from next-generation sequencing data. Bioinformatics 25: 2839–2840.

Leaché, A. D., B. L. Banbury, J. Felsenstein, A. N. de Oca & A. Stamatakis. 2015. Short tree, long tree, right Tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. Syst. Biol. 64: 1032–1047.

Lee, J. H., D. H. Lee & B. H. Choi. 2013. Phylogeography and genetic diversity of East Asian *Neolitsea sericea* (Lauraceae) based on variations in chloroplast DNA sequences. J. Plant Res. 126: 193–202.

Li, E. X., S. Yi, Y. X. Qiu, J. T. Guo, H. P. Comes & C. X. Fu. 2008. Phylogeography of two East Asian species in *Croomia* (Stemonaceae) inferred from chloroplast DNA and ISSR fingerprinting variation. Mol. Phylogenet. Evol. 49: 702–714.

Lischer, H. E. & L. Excoffier. 2012. PGDSpider: an automated data conversion tool for connecting population genetics and genomics programs. Bioinformatics 28: 298–299.

Lisiecki, L. E. & M. E. Raymo. 2005. A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. Paleoceanography 20: PA1003.

Lu, R. S., Y. Chen, I. Tamaki, S. Sakaguchi, Y. Q. Ding, D. Takahashi, P. Li, Y. Isaji, J. Chen & Y. X. Qiu. 2020. Pre-quaternary diversification and glacial demographic expansions of *Cardiocrinum* (Liliaceae) in temperate forest biomes of Sino-Japanese Floristic Region. Mol. Phylogenet. Evol. 143: 106693.

Magota, K., S. Sakaguchi, J. S. Lee, M. Yamamoto, D. Takahashi, A. J. Nagano & H. Setoguchi. 2021. Phylogeographic analysis of *Saxifraga fortunei* complex (Saxifragaceae) reveals multiple origins of morphological and ecological variations in the Japanese Archipelago. Mol. Phylogenet. Evol. 163: 107230.

McCormack, J. E., S. M. Hird, A. J. Zellmer, B. C. Carstens & R. T. Brumfield. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Mol. Phylogenet. Evol. 66: 526–538.

McIver, J. & K. Erickson. 2012. Pollination biology of *Potentilla recta* (sulfur cinquefoil) and its co-occurring native congener *Potentilla gracilis* in northeastern Oregon. Psyche 2012: 281732.

Ministry of the Environment Japan. 2015. Red data book 2014–Threatened wildlife of Japan, vol 8, vascular plants. Gyousei, Tokyo (in Japanese).

Momohara, A. 2016. Stages of major floral change in Japan based on macrofossil evidence and their connection to climate and geomorphological changes since the Pliocene. Quat. Int. 397: 93–105.

Momohara, A. 2018. Influence of mountain formation on floral diversification in Japan, based on macrofossil evidence. *In*: Hoorn, C., A. Perrigo & A. Antonelli (eds.), Mountains, climate and biodiversity, pp. 459–474. Wiley Blackwell, Oxford.

el Mousadik, A. & R. J. Petit. 1996. Chloroplast DNA phylogeography of the argan tree of Morocco. Mol. Ecol. 5: 547–555.

Murakami, S., T. Ito, T. Uemachi, S. Fujii, A. Matsuo, Y. Suyama & M. Maki. 2024. Phylogenetic relationships and divergence time of *Hydrangea* sect. *Macrophyllae* (Hydrangeaceae) revealed by genome-wide SNPs. Plant Syst. Evol. 310: 27

Murata, G. 1977. Phytogeographical consideration on the flora and vegetation of Japan. Acta Phytotax. Geobot. 28: 65–83 (in Japanese).

Murata, G. 1988. The distribution on the continental elements in Japan. Flora and its characteristic in Japan, vol 17. Nihon no Seibutsu 2: 21–25 (in Japanese).

Naimi, B., N. A. S. Hamm, T. A. Groen, A. K. Skidmore & A. G. Toxopeus. 2013. Where is positional uncertainty a problem for species distribution modelling? Ecography 37: 191–203.

Naimi, B. 2017. Package 'usdm'. Uncertainty analysis for species distribution models. Wien <http//www. cran. r-project. org.>.

Naruhashi, N. 1970. A new natural hybrid between *Potentilla discolor* Bunge and *P. kleiniana* Wight subsp. *anemonefolia* (Lehm.) Murata. Acta Phytotax. Geobot. 24: 122–127.

Naruhashi, N. 2001. *Potentilla* L. *In*: Iwatsuki, K., D. E. Boufford, H. Ohba (eds.), Flora of Japan, vol. IIb, pp. 193–206. Kodansha, Tokyo.

Ortiz, E. M. 2019. vcf2phylip v2.0: convert a VCF matrix into several matrix formats for phylogenetic analysis. doi:10.5281/zenodo.2540861.

Oh, S. Y. & J. H. Pak. 2001. Distribution maps of vascular plants in Korea. Academybook Publishing Co., Seoul (in Korean).

Ooi, N. 2016. Vegetation history of Japan since the last glacial based on palynological data. Jap. J. Histor. Bot. 25: 1–101 (in Japanese).

Peakall, R. & P. E. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update. Bioinformatics 28: 2537–2539.

Pelser, P. B., A. H. Kennedy, E. J. Tepe, J. B. Shidler, B. Nordenstam, J. W. Kadereit & L. E. Watson. 2010. Patterns and causes of incongruence between plastid and nuclear Senecioneae (Asteraceae) phylogenies. Amer. J. Bot. 97: 856–873.

Petit, R. J. & L. Excoffier. 2009. Gene flow and species delimitation. Trends Ecol. Evol. 24: 386–393.

Phillips, S. J., R. P. Anderson, M. Dudík, R. E. Schapire & M. E. Blair. 2017. Opening the black box: an open-source release of Maxent. Ecography 40: 887–893.

Pritchard, J. K., M. Stephens & P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155: 945–959.

Pritchard, J. K., W. Wen & D. Falush. 2010. Documentation for STRUCTURE software: Version 2.3. University of Chicago, Chicago.

Privé, F., K. Luu, B. J. Vilhjalmsson & M. G. B. Blum. 2020. Performing highly efficient genome scans for local adaptation with R package pcadapt version 4. Mol. Biol. Evol. 37: 2153–2154.

QGIS Development Team. 2022. QGIS Geographic Information System. Open Source Geospatial Foundation <http://qgis.org>.

Qi, X.-S., N. Yuan, H. P. Comes, S. Sakaguchi & Y.-X. Qiu. 2014. A strong 'filter' effect of the East China Sea land bridge for East Asia's temperate plant species: inferences from molecular phylogeography and

ecological niche modelling of *Platycrater arguta* (Hydrangeaceae). BMC Evol. Biol. 14: 1–16.

Qiu, Y. X., Y. Sun, X. P. Zhang, J. Lee, C. X. Fu & H. P. Comes. 2009. Molecular phylogeography of East Asian *Kirengeshoma* (Hydrangeaceae) in relation to quaternary climate change and landbridge configurations. New phytol. 183: 480–495.

Qiu, Y. X., C. X. Fu & H. P. Comes. 2011. Plant molecular phylogeography in China and adjacent regions: Tracing the genetic imprints of Quaternary climate and environmental change in the world's most diverse temperate flora. Mol. Phylogenet. Evol. 59: 225–244.

Rambaut, A. 2009. FigTree ver. 1.4.4, <http://tree.Bio.Ed.Ac.Uk/software/figtree/> [accessed Apr. 27, 2022].

Raynal, L., J. M. Marin, P. Pudlo, M. Ribatet, C. P. Robert & A. Estoup. 2019. ABC random forests for Bayesian parameter inference. Bioinformatics 35: 1720–1728.

R Core Team. 2023. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/> [accessed Dec. 1, 2023].

Rozas, J., A. Ferrer-Mata, J. C. Sánchez-DelBarrio, S. Guirao-Rico, P. Librado, S. E. Ramos-Onsins & A. Sánchez-Gracia. 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol. Biol. Evol. 34: 3299–3302.

Sabatini, F. M., B. Jimenez-Alfaro, U. Jandt, M. Chytry, R. Field, M. Kessler, J. Lenoir, F. Schrodt, S. K. Wiser, M. A. S. Arfin Khan, F. Attorre, L. Cayuela, M. De Sanctis, J. Dengler, S. Haider, M. Z. Hatim, A. Indreica, F. Jansen, A. Pauchard, R. K. Peet, P. Petrik, V. D. Pillar, B. Sandel, M. Schmidt, Z. Tang, P. van Bodegom, K. Vassilev, C. Violle, E. Alvarez-Davila, P. Davidar, J. Dolezal, B. Herault, A. Galan-de-Mera, J. Jimenez, S. Kambach, S. Kepfer-Rojas, H. Kreft, F. Lezama, R. Linares-Palomino, A. Monteagudo Mendoza, J. K. N'Dja, O. L. Phillips, G. Rivas-Torres, P. Sklenar, K. Speziale, B. J. Strohbach, R. Vasquez Martinez, H. F. Wang, K. Wesche & H. Bruelheide. 2022. Global patterns of vascular plant alpha diversity. Nat. Commun. 13: 4683.

Sakaba, T., A. Soejima, S. Fujii, H. Ikeda, T. Iwasaki, H. Saito, Y. Suyama, A. Matsuo, A. E. Kozhevnikov, Z. V. Kozhevnikova, H. Wang, S. Wang, J. H. Pak & N. Fujii. 2023. Phylogeography of the temperate grassland plant *Tephroseris kirilowii* (Asteraceae) inferred from multiplexed inter-simple sequence repeat genotyping by sequencing (MIG-seq) data. J. Plant Res. 136: 437–452.

Sakaguchi, S., Y. Asaoka, D. Takahashi, Y. Isagi, R. Imai, A. J. Nagano, Y. X. Qiu, P. Li, R. Lu & H. Setoguchi. 2021. Inferring historical survivals of climate relicts: the effects of climate changes, geography, and population-specific factors on herbaceous hydrangeas. Heredity 126: 615–629.

Sakaguchi, S., T. Kimura, R. Kyan, M. Maki, T. Nishino, N. Ishikawa, A. J. Nagano, M. N. Honjo, M. Yasugi, H. Kudoh, P. Li, H. J. Choi, O. A. Chernyagina & M. Ito. 2018. Phylogeographic analysis of the East Asian goldenrod (*Solidago virgaurea* complex, Asteraceae) reveals hidden ecological diversification with recurrent formation of ecotypes. Ann. Bot. 121: 489–500.

Sakaguchi, S., Y. X. Qiu, Y. H. Liu, X. S. Qi, S. H. Kim, J. Han, Y. Takeuchi, J. R. Worth, M. Yamasaki, S. Sakurai & Y. Isagi. 2012. Climate oscillation during the Quaternary associated with landscape heterogeneity promoted allopatric lineage divergence of a temperate tree *Kalopanax septemlobus* (Araliaceae) in East Asia. Mol. Ecol. 21: 3823–3838.

Sata, H., M. Shimizu, T. Iwasaki, H. Ikeda, A. Soejima, A. E. Kozhevnikov, Z. V. Kozhevnikova, H. T. Im, S. K. Jang, T. Azuma, A. J. Nagano & N. Fujii. 2021. Phylogeography of the East Asian grassland plant, *Viola orientalis* (Violaceae), inferred from plastid and nuclear restriction site-associated DNA sequencing data. J. Plant Res. 134: 1181–1198.

Shaw, J., E. B. Lickey, J. T. Beck, S. B. Farmer, W. Liu, J. Miller, K. C. Siripun, C. T. Winder, E. E. Schilling & R. L. Small. 2005. The tortoise and the hare ii: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Amer. J. Bot. 92: 142–166.

Shi, Y., B. Ren, J. Wang & E. Derbyshire. 1986. Quaternary glaciation in China. Quat. Sci. Rev. 5: 503–507.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30: 1312–1313.

Suissa, J. S., G. Y. De La Cerda, L. C. Graber, C. Jelley, D. Wickell, H. R. Phillips, A. D. Grinage, C. S. Moreau, C. D. Specht, J. J. Doyle & J. B. Landis. 2024. Data-driven guidelines for phylogenomic analyses using SNP data. Appl. Plant Sci. 12: e11611.

Suyama, Y. & Y. Matsuki. 2015. MIG-seq: An effective PCR-based method for genome-wide single-nucleotide polymorphism genotyping using the next-generation sequencing platform. Sci. Rep. 5: 16963.

Suyama Y., S. K. Hirota, A. Matsuo, Y. Tsunamoto, C. Mitsuyuki, A. Shimura & K. Okano. 2022. Complementary combination of multiplex high-throughput DNA sequencing for molecular phylogeny. Ecol. Res. 37: 171–181.

Tabata, H. 1997. Meadow plants in Satoyama. *In*: Tabata, H. (ed.), Satoyama and its conservation, pp. 38–41. Hoikusha, Osaka (in Japanese).

Taberlet, P., L. Gielly, G. Pautou & J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17: 1105–1109.

Takhtajan, A. 1969. Flowering plants: origin and dispersal. Smithsonian Institution Press, Washington.

Tsukada, M. 1983. Vegetation and climate during the last glacial maximum in Japan. Quat. Res. 19: 212–235.

Union of Japanese Societies for Systematic Biology. 2003. Japanese biota species number survey, 1st edition <http://ujssb.org/biospnum/search.php> [accessed Nov. 30, 2023] (in Japanese).

Ushimaru, A., K. Uchida & T. Suka. 2018. Grassland biodiversity in Japan: threats, management and conservation. *In*: Squires, V. R., J. Dengler, H. Feng & L. Hua (eds.), Grasslands of the World, pp. 197–218. CRC Press, Boca Raton.

Wagner, D., G. Furnier, M. Saghai-Maroof, S. Williams, B. Dancik & R. Allard. 1987. Chloroplast DNA polymorphisms in lodgepole and jack pines and their hybrids. Proc. Natl. Acad. Sci. USA 84: 2097–2100.

Warren, D. L., N. J. Matzke, M. Cardillo, J. B. Baumgartner, L. J. Beaumont, M. Turelli, R. E. Glor, N. A. Huron, M. Simões, T. L. Iglesias, J. C. Piquet & R. Dinnage. 2021. ENMTools 1.0: an R package for comparative ecological biogeography. Ecography 44: 504–511.

Watanabe, S., T. Hajima, K. Sudo, T. Nagashima, T. Takemura, H. Okajima, T. Nozawa, H. Kawase, M. Abe, T. Yokohata, T. Ise, H. Sato, E. Kato, K. Takata, S. Emori & M. Kawamiya. 2011. MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments. Geosci. Model Dev. 4: 845–872.

Wendel, J. F. & J. J. Doyle. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. *In*: Soltis, D. E., P. S. Soltis & J. J. Doyle (eds.), Molecular systematics of plants II, pp. 265–296. Kluwer Academic Publishers, Massachusetts.

Xia, M. Q., R. Y. Liao, J. T. Zhou, H. Y. Lin, J. H. Li, P. Li, C. X. Fu & Y. X. Qiu. 2022. Phylogenomics and biogeography of *Wisteria*: implications on plastome evolution among inverted repeat-lacking clade (IRLC) legumes. J. Syst. Evol. 60: 253–265

Xu, D., H. Lu, N. Wu & Z. Liu. 2010. 30000-Year vegetation and climate change around the East China Sea shelf inferred from a high-resolution pollen record. Quat. Int. 227: 53–60.

Yakubov, V. V. 1996. *Potentilla* L. *In*: Charkevicz S. S. (ed.), Plantae Vasculares Orientalis Extremi Sovietici. vol. 8, pp. 168–206. Saint Petersburg, Nauka (in Russian).

Yasuda, Y. & N. Miyoshi. 1998. Illustrated vegetation history of Japanese Archipelago. Asakura-shoten, Tokyo. (in Japanese).

Young, N. D., K. E. Steiner & C. W. dePamphilis. 1999. The evolution of parasitism in Scrophulariaceae/Orobanchaceae: plastid gene sequences refute an evolutionary transition series. Ann. Missouri Bot. Gard. 86: 876–893.

Zeng, Y. F., W. T. Wang, W. J. Liao, H. F. Wang & D. Y. Zhang. 2015. Multiple glacial refugia for cool-temperate deciduous trees in northern East Asia: The Mongolian oak as a case study. Mol. Ecol. 24: 5676–5691.

Received December 2, 2024; accepted March 19, 2025