Journal of Plant Research (2021) 134:1181–1198
https://doi.org/10.1007/s10265-021-01339-8

REGULAR PAPER – TAXONOMY/PHYLOGENETICS/EVOLUTIONARY BIOLOGY

Phylogeography of the East Asian grassland plant, *Viola orientalis* (Violaceae), inferred from plastid and nuclear restriction site-associated DNA sequencing data

Haruna Sata[1] · Midori Shimizu[1] · Takaya Iwasaki[2] · Hajime Ikeda[3] · Akiko Soejima[4] · Andrey E. Kozhevnikov[5] · Zoya V. Kozhevnikova[5] · Hyoung-Tak Im[6] · Su-Kil Jang[7] · Takayuki Azuma[8] · Atsushi J. Nagano[9] · Noriyuki Fujii[4,10]

Received: 30 March 2021 / Accepted: 18 August 2021 / Published online: 30 September 2021

Abstract
To elucidate the origin and migration history of the "Mansen elements," a group of temperate grassland plants mainly distributed in northeastern Asia, phylogeographic analyses based on chloroplast DNA markers and double-digest restriction site-associated DNA sequencing (ddRAD-seq) data were performed on *Viola orientalis*, one of the representative species of the group. Phylogenetic analyses using ddRAD-seq data revealed that the populations of *V. orientalis* were clustered into five clades, among which the continental clades made of populations from Russia and Korea diverged more than 100,000 years earlier than the Japanese clades. The Japanese clade likely diverged during the last glacial period, followed by a further post-glacial divergence into the Kyushu and the Honshu subclades. Our study demonstrated that *V. orientalis* originated in the continental area of northeastern Asia and, during the last glacial period, has spread southward through the Korean Peninsula across the Japanese Islands. This finding supports the previously proposed evolutionary hypothesis regarding the origin and migration routes of the Mansen elements.

Keywords DdRAD-seq · Genetic structure · Grassland · Japanese flora · Migration history · Population genetics

Introduction

Global climate oscillations during the Quaternary Period caused significant changes to the distributions of many plant species (Liu et al. 2012; Meng et al. 2015; Qiu et al. 2011; Soltis et al. 1997). Although East Asia was primarily free of ice sheets during the last glacial period (approximately 115,000–11,700 years ago) (Batchelor et al. 2019; Shi et al. 1986), climatic oscillations during the Quaternary have influenced the distribution of vegetation in this region (Axelrod et al. 1996; Harrison et al. 2001; Qiu et al. 2011). The

✉ Noriyuki Fujii
nfujii@kumamoto-u.ac.jp

1 Graduate School of Science and Technology, Kumamoto University, 2-39-1 Kurokami, Chuo, Kumamoto, Kumamoto 860-8555, Japan

2 Natural Science Division, Faculty of Core Research, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan

3 Institute of Plant Science and Resources, Okayama University, 2-20-1 Chuo, Kurashiki, Okayama 710-0046, Japan

4 Faculty of Advanced Science and Technology, Kumamoto University, 2-39-1 Kurokami, Chuo, Kumamoto, Kumamoto 860-8555, Japan

5 Federal Scientific Center of the East Asia Terrestrial Biodiversity, Far Eastern Branch of the Russian Academy of Sciences, Vladivostok, Russia

6 Department of Biological Science, Chonnam National University, Gwangju 61186, South Korea

7 East Coast Research Institute of Life Science, Gangneung-Wonju National University, Jukheon-gil 7, Gangneung city, Gangwon-do 25457, South Korea

8 Botanic Garden, Field Science Center for Northern Biosphere, Hokkaido University, N3W8, Chuo-ku, Sapporo, Hokkaido 060-0003, Japan

9 Faculty of Agriculture, Ryukoku University, 1-5 Yokotani, Seta Oe-cho, Otsu, Shiga 520-2194, Japan

10 Course of Biological Science, Faculty of Science, Kumamoto University, 2-39-1 Kurokami, Chuo, Kumamoto, Kumamoto 860-8555, Japan

historical vegetation dynamic has been traditionally explored by pollen (Ooi 2016; Tsukada 1983; Yasuda and Miyoshi 1998) and plant macrofossils (Momohara 2016, 2018). Recently, genetic variation and structure implemented by phylogeographic studies enabled us to infer a more detailed range dynamic history of each species (Fujii and Senni 2006; Ikeda et al. 2016, 2018; Liu et al. 2012; Lu et al. 2020; Qiu et al. 2011; Sakaguchi et al. 2018; Xu et al. 2010; Zhao et al. 2019).

The Japanese archipelago extends 3000 km from the northeast to southwest in the eastern end of the Asian continent, covering a wide range of climatic zones, from subarctic to subtropical. The archipelago is inhabited by approximately 5000 vascular plants (Union of Japanese Societies for Systematic Biology 2003), representing relatively high species diversity and endemism (Mittermeier et al. 2004, 2011). The Japanese flora is supposed to be strongly related to that of the Asian continent (Hotta 1974; Maekawa 1998) and is considered part of the Sino-Japanese floristic region (Good 1974; Takhtajan et al. 1986). Recent phylogeographic studies have demonstrated past connections between the Japanese Islands and the Asian continent in several types of species, including alpine plants (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 plant species with relict disjunct distribution (Li et al. 2008; Qiu et al. 2009a, b). Most of these studies suggest that the Japanese flora generally originated on the Asian continent, and after expanding their distribution to the Japanese archipelago, an endemic lineage evolved and diversified in Japan. However, these studies have not sufficiently examined whether the origin of the plant species is on the continent or in Japan itself, which is an unresolved issue. Several studies suggest a Japanese origin rather than a continental origin for these species (Ikeda et al. 2018, 2020; Xia et al. 2021).

Temperate grassland plants in Japan are usually distributed throughout northeastern China, Far East Russia, and the Korean Peninsula. In Japan, most of these plants inhabit the temperate southwestern parts of the archipelago and are not found on the northernmost large island, Hokkaido (Hotta 1974; Koizumi 1931; Murata 1988). Koizumi (1931) termed such indigenous plants as the "Mansen elements" in reference to the geographical names of these continental regions (Man-shu and Cho-sen in Japanese). These 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 the species in the Mansen elements originated in the Asian continent and migrated to Japan via the Korean Peninsula under cold climate in the Pleistocene (Hotta 1974; Kitamura 1957; Murata 1988; Tabata 1997). Furthermore, given the small morphological differences of species in Mansen elements between the continental and Japanese populations, their migration presumably occurred during the recent cool and dry period (Murata 1977, 1988). However, few phylogeographic studies have been conducted using the Mansen elements (except for Takaishi et al. (2019), discussed below). Therefore, the hypotheses regarding the origin and migration history in the plant group were not adequately tested.

A previous study investigated the genetic diversity and structure of *Pulsatilla cernua* (Thnub.) Berchtold et J.Presl (Ranunculaceae), one of the Mansen elements species, using variations in cpDNA and microsatellites of nuclear DNA (Takaishi et al. 2019) to assess this biogeographic hypothesis. A common cpDNA haplotype and its satellite haplotypes were observed across the Japanese Islands and the Russian populations. Furthermore, microsatellite analyses did not show genetic structure, either throughout the Japanese or the continental populations. Thus, the previous study suggested a rapid expansion of *P. cernua* in Japan. However, the authors did not address the continental origin of the species of the Mansen elements and their recent migration history.

In this study, we attempted to assess the continental origin of the species in the Mansen elements by a phylogeographic survey of *Viola orientalis* (Maxim.) W. Becker (Violaceae). *Viola orientalis* is a perennial herb native to sunny grasslands of low mountains in Japan, the Korean Peninsula, northeastern China, and the Primorsky Krai located in the Far East region of Russia (Bezdeleva 1987; Noda 1971; Oh and Pak 2001) (Fig. 1). In Japan, this species is sparsely distributed across the main islands of the archipelago: Honshu (west of the Tokai region), Shikoku, and Kyushu (Akiyama et al. 1999; Hama 2002; Kadota 2016; Ohwi 1953, 1983) but is absent from northeastern Honshu and Hokkaido. Based on such a distribution pattern, this species is supposed to have immigrated to Japan from the west through the Korean Peninsula (Igari 2004; Kitamura 1957; Murata 1977, 1988), being considered a typical example of the Mansen elements. In addition, because most *Viola* plants have diplochorous seeds, dispersed by explosive ejection from capsules and ants conveyance (Beattie and Lyons 1975), it is difficult for them to migrate across seas. Therefore, *V. orientalis* is a suitable species to test the hypothesis regarding the origin and migration history of the Mansen elements.

Together with chloroplast DNA sequences, we used double-digest restriction site-associated DNA sequencing (ddRAD-seq), a powerful tool for single-nucleotide polymorphisms (SNPs) genotyping of hundreds of individuals of any species at a time (Andrews et al. 2016; Peterson et al. 2012). With such techniques, we were able to reveal the range-wide genetic structure of *V. orientalis*. In addition, by estimating demographic histories, we examined (1) whether *V. orientalis* migrated into Japan from the Asian continent or vice versa and (2) infer the migration event temporal framework. The elucidation of the origin and migration history of the Mansen elements will contribute to understanding the

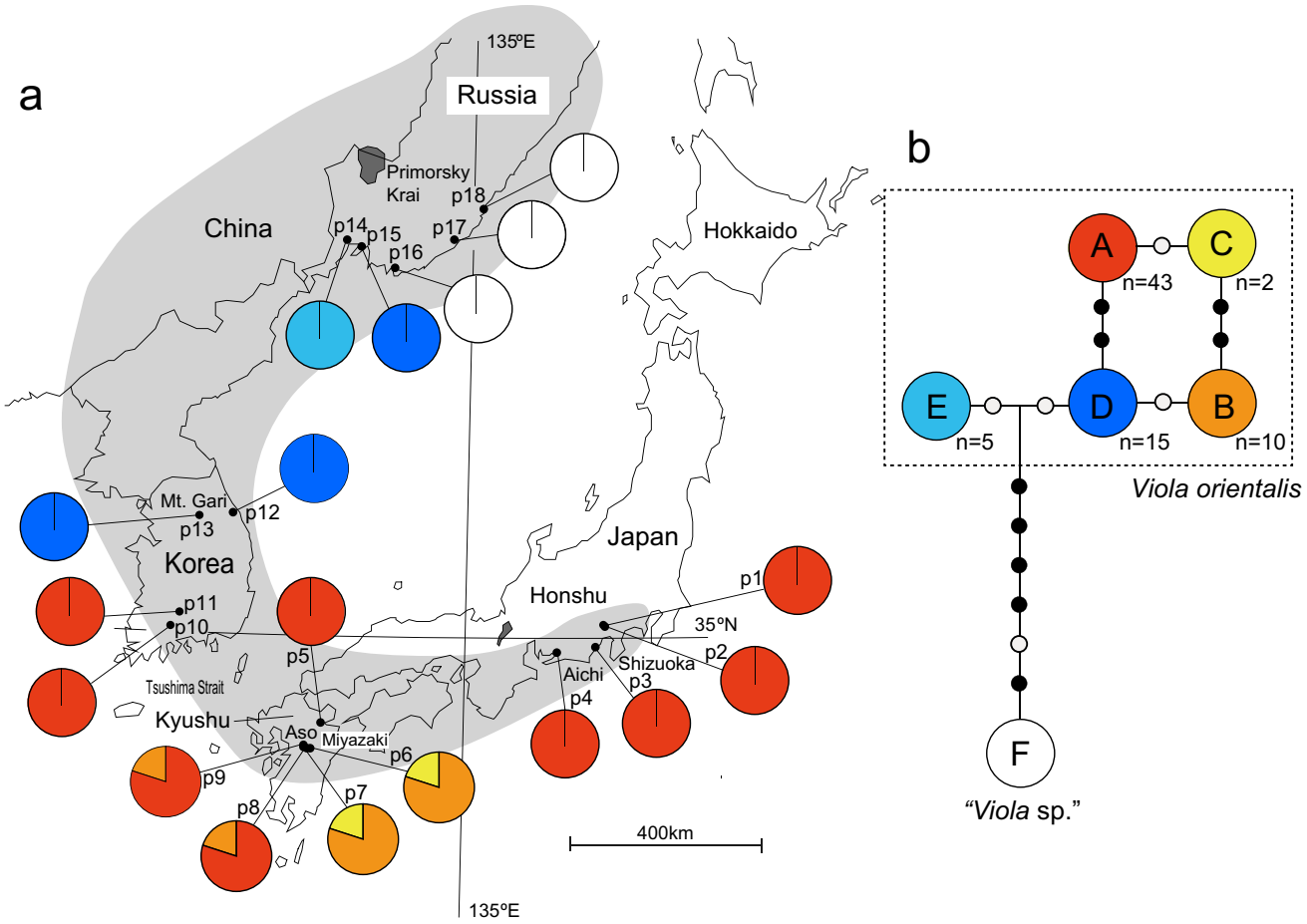

Fig. 1 Geographic distribution of the chloroplast DNA haplotypes of *Viola orientalis* (**a**), and median-joining (MJ) network among the haplotypes (**b**). Population numbers correspond to those shown in Table 1. The size of each circle in the left figure represents five individuals per population. In the MJ network, nucleotide substitutions and insertions/deletions detected are represented by solid circles and open circles, respectively

evolutionary history of grassland plants in Northeast Asia, including the Japanese Islands.

Materials and methods

Plant materials

We sampled a total of 79 individuals from 15 populations representing the entire geographic range of *V. orientalis* (75 plants for cpDNA analyses, 79 plants for ddRAD-seq analyses, Fig. 1 and Table 1). In each population, leaf samples were collected from individuals as far as possible apart from each other and preserved in silica gel. We collected samples from three populations in the Primorsky Krai Russian region (p16–p18) whose leaf shape was slightly different from the remaining populations. The leaves of the three populations were rather large, with cordate bases. Hence, we distinguished the plants of these populations as "*Viola* sp." Additionally, we collected four outgroup species: *V. alliariifolia* Nakai (Sect. *Chamaemelanium* Ging.), *V. brevistipulata* (Franch. et Sav.) W. Becker (Sect. *Chamaemelanium*), *V. yubariana* Nakai (Sect. *Chamaemelanium*), and *V. biflora* L. (Sect. *Dischidium* Ging.). We used a single individual in each outgroup species (Table 1). The outgroups were selected based on taxonomic reports (Akiyama et al. 1999; Hama 2002; Kadota 2016) and a molecular phylogenetic study on Korean *Viola* species (Yoo and Jang 2010). The section names followed Kadota (2016). Voucher specimens of the samples used in this study were deposited in the herbaria of the Faculty of Science, Kumamoto University (KUMA); College of Natural Sciences, Kangwon National University (KWNU); and the Botanic Garden, Field Science Center for Northern Biosphere, Hokkaido University (SAPT).

Table1 The materials and their sources analyzed for RAD-seq and cpDNA variation

Population code and taxon	Locality, latitude/longitude and voucher	No. of plants	
		cpDNA	RAD-seq
Viola orientalis (Maxim.) W. Becker (Sect. *Chamaemelanium* Ging.)			
Japan			
p1	Asagiri-Highland, Fujinomiya, Shizuoka, Honshu, 35° 23′ N/138° 35′ E, N. Fujii F03418 (KUMA)	5	5
p2	Asagiri-Highland, Fujinomiya, Shizuoka, Honshu, 35° 25′ N/138° 35′ E, N. Fujii F03417 (KUMA)	5	5
p3	Mt. Takakusa, Yaizu, Shizuoka, Honshu, 34° 54′ N/138° 19′ E, N. Fujii F03421 (KUMA)	5	5
p4	Iwasaki and Owaki, Toyohashi, Aichi, Honshu, 34° 44′ N/137° 27′ E, N. Fujii F03432 (KUMA)	5	5
p5	Mt. Yufu, Beppu, Oita, Kyushu, 33° 16′ N/131° 23′ E, N. Fujii F03407 (KUMA)	5	5
p6	Gokasho-Highland, Takachiho, Miyazaki, Kyushu, 32°48′ N/131° 17′ E, N. Fujii F03436 (KUMA)	5	5
p7	Takamori, Aso, Kumamoto, Kyushu, 32° 50′ N/131° 14′ E, N. Fujii F03410 (KUMA)	5	5
p8	Mt. Okamado, Aso, Kumamoto, Kyushu, 32° 51′ N/131° 03′ E, N. Fujii F03409 (KUMA)	5	5
p9	Kario, Aso, Kumamoto,Kyushu, 32° 57′ N/130° 60′ E, N. Fujii F03408 (KUMA)	5	5
Korea			
p10	Jeongryeong-chi, Gurye, Jeollabuk, 35° 12′ N/127° 27′ E, S.-K. Jang, KWNU92084 (KWNU)	5	6
p11	Mt. Gwaegwan-san, Hamyang, Gyeongsangnam, 35° 36′ N/127° 41′ E, T. K. Im F03422 (KUMA)	5	6
p12	Neungkyeong Peak, Gangneung, Gangwon, 37° 45′ N/128° 52′ E, S.-K. Jang, KWNU92082 (KWNU)	5	6
p13	Mt. Gari, Guan-eum, Hongcheon, Gangwon, 37° 42′ N/127° 54′ E, S.-K. Jang, KWNU92093 (KWNU)	5	6
Russia			
p14	40 km southwest of Ussurijisk, Primorsky Krai, 43° 26′ N/131° 44′ E, N. Fujii F03427 (KUMA)	5	5
p15	15 km northeast of Vladivostok, Primorsky Krai, 43° 13′ N/132° 04′ E, N. Fujii F03426 (KUMA)	5	5
"*Viola* sp."[1]			
p16	5 km east of Nahodka, Primorsky Krai, 42° 49′ N/132° 60′ E, N. Fujii F03433 (KUMA)	5	5
p17	10 km east of Milogradovo, Primorsky Krai, 43° 13′ N/134° 28′ E, N. Fujii F03434 (KUMA)	5	5
p18	20 km noutheast of Ol'ga, Primorsky Krai, 43° 54′ N/135° 27′ E, N. Fujii F03435 (KUMA)	5	5
Outgroups			
Viola brevistipulata (Franch. et Sav.) W.Becker (Sect. *Chamaemelanium* Ging.)			
	Higashiyuri, Yurihonjou, Akita, Honshu, Japan, 39° 21′ N/140° 13′ E, T. Azuma 2804 (SAPT)	1	1
Viola alliariaefolia Nakai (Sect. *Chamaemelanium* Ging.)			
	Taisetsu Mtns., Kamikawa, Kamikawa, Japan, 43° 39′ N/142° 54′ E, T. Azuma 2799 (SAPT)	1	0
Viola yubariana Nakai (Sect. *Chamaemelanium* Ging.)			
	Mt.Yubari, Yubari, Hokkaido, Japan, 43° 06′ N/142° 14′ E, T. Azuma 2795 (SAPT)	1	0
Viola biflora L. (Sect. *Dischidium* Ging.)			
	Along Unabetsu River, Shari, Shari, Hokkaido, Japan, 43° 54′ N/144° 47′ E, T. Azuma 2802 (SAPT)	1	1

[1]These three populations in the Primorsky Krai had been identified as *V. orientalis* in Bezdeleva (1987), however, we treated them as "*Viola* sp." in the present study, because the characteristics of leaves of these populations were slightly different from those of *V. orientalis* (see the text for details)

Outline of DNA analyses

For the cpDNA analysis, the *rpl16* intron was sequenced for 75 accessions of *V. orientalis* and 15 accessions from three populations of "*Viola* sp." in the Primorsky Krai region to determine cpDNA haplotypes (for a detailed procedure, see below). This region, the *rpl16* intron, was the most variable in preliminary experiments, including three other cpDNA regions [*trnL–trnF*, *atpB–rbcL*, and *trnS–trnC* (Kim et al. 2005)]. Two additional noncoding regions (*trnL–trnF* and *atpB–rbcL*) were sequenced for a single accession of each population of *V. orientalis* and "*Viola* sp." as well as the outgroup accessions to elucidate relationships among cpDNA haplotypes/populations further. For the ddRAD-seq analysis, all 79 accessions of *V. orientalis*, 15 accessions of "*Viola* sp.", and two from the outgroup species (*V. biflora* and *V. brevistipulata*) were assessed (Table 1).

DNA isolation

After powdering dry leaves using Sea Sand (Wako Pure Chemical Industries, Ltd., Osaka, Japan), total genomic DNA was extracted using the DNeasy Plant Mini Kit (Qiagen Inc., Hilden, Germany) following the manufacturer's protocol. For ddRAD-seq analysis, the concentration of each template DNA sample was quantified on Quantus Fluorometer using QuantiFluor ONE dsDNA Dye (Promega, Madison, WI) and was subsequently adjusted to 10–20 ng μL^{-1}.

Direct sequencing analysis of cpDNA

Three noncoding regions of cpDNA were amplified and sequenced: the *rpl16* intron using the primer pair "F71" and "R1661" (Jordan et al. 1996), the *trnL–trnF* region using the primer pair "c" and "f" (Taberlet et al. 1991), and the *atpB–rbcL* region using the primer pair "atpB" and "rbcL" (Terachi 1993). The PCR reaction mixtures contained 50–100 ng template DNA, 5 µL 10 × PCR buffer, 4 µL 2.5 mM of each deoxyribonucleotide, 2.5 µL 0.5 µM of each of the primer pair, and 0.25 µL ExTaq DNA polymerase (Takara Bio Inc., Tokyo) in a total reaction volume of 50 µL. The PCR program ran for 5 min at 94 ℃ for initial denaturalization, followed by 30 cycles of denaturation at 94 ℃ for 1 min, primer annealing at 50 ℃ for 1 min and extension at 72 ℃ for 2 min. The reactions were finally extended by 7 min at 72 ℃. After DNA amplification was confirmed, we used the Illustra ExoProStar (GE Healthcare, Tokyo) to eliminate the remaining primers and nucleotides. For sequencing analyses, the same primers were used (final concentration 9.6 pM). We ordered an analysis of DNA sequencing from Eurofins Genomics Corp. (Tokyo, Japan). The obtained sequences were assembled using ChromasPro ver. 1.74 (Technelysium Pty Ltd., South Brisbane, Australia) and then aligned automatically using MEGA software ver. 6.06 (Tamura et al. 2013) with the ClustalW alignment option (Thompson et al. 1994). Parts of gap characters (insertions/deletions [indels]) were corrected manually so as to maximize the number of matching nucleotides in the corresponding sequences. The nucleotide sequence data reported in this paper will appear in the DDBJ, EMBL, and GenBank DNA databases under accession numbers LC548058–LC548092.

Data analyses using cpDNA sequences

CpDNA haplotypes of *V. orientalis* based on *rpl16* intron were determined by the DnaSP ver. 6.12.01 (Rozas et al. 2017). Haplotypes were discriminated using both nucleotide substitutions and indel characters. In this data set, we estimated nucleotide diversity (π) and haplotype diversity (*Hd*) among the Japanese populations and the continental populations using DnaSP software. Additionally, analyses of molecular variance (AMOVA) with 10,000 permutations were performed to assess genetic differentiation within and between geographic regions using Arlequin ver. 3.5.1.2 (Excoffier and Lischer 2010). Furthermore, we constructed a median-joining (MJ) network among the haplotypes using the software Network ver. 5.0.0.1 (Bandelt et al. 1999). MJ network was post-processed with the maximum-parsimony (MP) algorithm to remove unnecessary linkages and median vectors (Polzin and Daneshmand 2003). Phylogenetic relationships among *V. orientalis* populations using the three cpDNA regions (*rpl16* intron, *trnL–trnF*, and *atpB–rbcL*) were inferred by the maximum likelihood (ML) method using RAxML ver. 8.2.12 (Stamatakis 2014). The best-fit model for each cpDNA marker was determined using jModelTest ver. 2.1.10 (Darriba et al. 2012; Guindon and Gascuel 2003) by Akaike information criterion (AIC) (Akaike 1974). For *V. orientalis*, the optimal models were TVM, F81, and TPM1uf, respectively, for *rpl16* intron, *trnL–trnF,* and *atpB–rbcL*. Clade supports were estimated using a bootstrap analysis based on 1000 replicates (Felsenstein 1985). The phylogenetic tree was visualized and refined using FigTree ver. 1.4.4 (Rambaut 2009). For rooting the tree, an accession of *V. biflora* was selected as an outgroup (Yoo and Jang 2010).

DdRAD-seq analysis

A double-digest restriction site-associated DNA (ddRAD) library was prepared following the previously reported protocol (Peterson et al. 2012) with slight modifications to correspond to the method reported by Sakaguchi et al. (2017). The library was sequenced as 51-bp single-end reads in one lane using an Illumina HiSeq2000 (Illumina, San Diego,

CA, USA). After read trimming using TRIMMOMATIC ver. 0.32 software (using the commands LEADING:19, TRAILING:19, SLIDING WINDOW:30:20, AVGQUAL:20, and MINLEN:51) (Bolger et al. 2014), Stacks ver. 1.37 software (Catchen et al. 2011) was used to process the ddRAD-seq reads with the following parameter settings: the minimum number of identical reads required to create a stack ($m = 3$), the nucleotide mismatches between loci within a single individual ($M = 2$), and the mismatches between loci when building the catalog ($n = 4$).

SNP marker filtering

SNP genotypes for each individual were exported with a minimum stack depth set to 5 and a maximum observed heterozygosity cutoff of 0.6. Only the first SNP in each catalog was retrieved to exclude highly linked SNPs from the dataset. We extracted only SNPs having a frequency higher than 0.1 in each of the four pre-defined population groups: Japan, Korea, Russia, and outgroup taxa (*V. brevistipulata* and *V. biflora*). The filtering was performed using the "populations" command. Then, the exported genotype data was processed using PLINK ver. 1.07 software (Purcell et al. 2007); markers with locus having a missing individual rate of > 0.5 and a minor allele frequency of < 0.01, as well as the individuals with a missing locus rate of > 0.6, were filtered out. In this filtering process, only one individual from p12 was removed. This first SNP dataset (Data_1) was used for the STRUCTURE and NeighborNet analyses. Furthermore, to accurately assess the genetic diversity of *V. orientalis* and "*Viola* sp." stated in Table 1, we excluded two outgroup species (*V. brevistipulata* and *V. biflora*) and minor SNPs with an allele frequency of < 0.03 were filtered out. This second SNP dataset was used for genetic diversity and DIYABC analyses (Data_2). Conversely, all SNPs (not only first SNPs) in the same catalogs were concatenated into a single sequence dataset for all individuals (Data_3). This third dataset was used for the subsequent phylogenetic analysis.

Genetic structure analyses

Genetic structure analyses were examined by both model- and distance-based methods using 2790 SNPs from 95 individuals including two outgroup accessions (Data_1). In the model-based method, Bayesian clustering analysis was conducted using STRUCTURE ver. 2.2.4 (Pritchard et al. 2000). The number of clusters (K) was estimated without a priori population assignment of individuals under the admixture F model for correlated allele frequencies (Falush et al. 2003) with 100,000 Markov chain Monte Carlo (MCMC) steps and a burn-in of 100,000 iterations. We conducted 25 independent runs for the whole dataset ($K = 1$–17). Then, we determined the most meaningful number of genetic clusters (K) on the basis of a combination of the mean estimated Ln probability of the data [LnP(K)] (Pritchard et al., 2000) and the second-order rate of change in the log probability of the data (K) (Evanno et al. 2005). The ΔK values were calculated using STRUCTURE HARVESTER ver. 0.6.94 (Earl and vonHoldt 2012). Both the expected heterozygosity within each cluster and the genetic diversity of each cluster were calculated in STRUCTURE. The F_{ST} value for each cluster and the amount of drift from a common ancestral population for each cluster (analogous to traditional F_{ST} values between each cluster and a common ancestral population) were also calculated in STRUCTURE (see the documentation for the STRUCTURE program; Pritchard et al. 2010).

Pairwise genetic distances were calculated as binary Genetic Distances (GD) among individuals using GenAlEx ver. 6.501 (Peakall and Smouse 2012). A phylogenetic network using the NeighborNet method (Bryant and Moulton 2004) was reconstructed using SplitsTree ver. 4.10 (Huson and Bryant 2006) on the basis of the calculated GD.

Genetic diversity analyses

Genetic diversity parameters for each population: mean number of different alleles per locus (Na), mean number of different alleles unique to a single population per locus—private alleles (PA), mean number of locally common alleles (LCA) per locus found in 25% or fewer populations, observed heterozygosity (H_O), expected heterozygosity (H_E), and unbiased expected heterozygosity (uH_E) were estimated using GenAlEx ver. 6.501 (Peakall and Smouse 2012) on Data_2 including 1,445 SNPs from 93 individuals. Unbiased expected heterozygosity (uH_E) was calculated using the formula $(2\,N/(2\,N-1)) \times H_E$ by GenAlEx to standardize the genetic diversity considering different number of samples across the populations. Both allelic richness (AR) and fixation index (F_{IS}) per population were calculated using the R package "diveRsity" (Keenan et al. 2013), and their bootstrapped 95% CIs values (10,000 repeats) were estimated accordingly. The AR in the rarefaction framework was also estimated to diminish the bias caused by the sample size difference. Moreover, departure from Hardy–Weinberg equilibrium (HWE) was tested for each population with 9,999 MCMC iterations using the R package "diveRsity" (Keenan et al. 2013).

Phylogenetic analysis using ddRAD-seq data

A ML tree was inferred using RAxML ver. 8.2.12 with 1,000 bootstrap replicates based on the concatenated sequences of 8542 SNPs on 2,790 catalogs from 95 individuals (Data_3). The GTRGAMMA model of nucleotide substitution with a correction for ascertainment bias (-m GTRGAMMA–asc-corr = lewis) (Leaché et al. 2015) was used. Default settings

were used for other parameters. The program FigTree ver. 1.4.4 was utilized to visualize the phylogenetic tree.

Approximate Bayesian computation (ABC) analysis for inference of population demographic history

ABC methods were implemented with DIYABC ver. 2.1 (Cornuet et al. 2014) to examine the intraspecific divergence process of *V. orientalis*. Considering the genetic structure in STRUCTURE and NeighborNet analyses, we classified the 18 studied populations into the following five regional population groups: POP1, all Japanese populations (p1–p9); POP2, most of the Korean populations (p10–p12, because this study focuses on the relationship between Japanese and continental populations, the three Korean populations were grouped together in order to simplify the model); POP3, the northern unique Korean population (p13), which diverged genetically from the other Korean populations; POP4, the southern Russian populations (p14, p15); and POP5, the populations of "*Viola* sp." (p16–p18). After removing 126 loci absent in at least one population from Data_2, 1,319 SNPs across 93 individuals from 18 populations were used in the analysis.

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 analyses using various demographic models, seven models were assumed to estimate the demographic histories (Fig. 6). These models were characterized by several demographic parameters, including divergence times in generation (t1, t2, t3, and t4) and effective population sizes of POP1, POP2, POP3, POP4, and POP5 (N1, N2, N3, N4, and N5) and the ancestral population (Nanc). Models 1–5 assume that POP5 ("*Viola* sp.") first diverged from the others, and then POP4 diverged from the remaining POP1–3. Models 6 and 7 assume that POP4 diverged from POP5 (POP4 is included in the "*Viola* sp." lineage). Model 1 considered a scenario in which this species reached Japan with a continental origin, branching from Russia to Korea and then to Japan. On the other hand, Model 3 assumed a scenario in which Japan (POP1) diverged from Russia (POP4), followed by reverse migration into Korea from Japan. Models 2, 5, and 6 additionally assumed a common ancestral population from Japan and Korea (the effective population size is Njk), and they diverged from it in various divergence histories. Model 4 is a modified version of model 3 and assumes a scenario in which the common ancestral population of Korea (the effective population size is Njk) was formed by a mixture of Japanese (POP1) and Russian (POP4) populations. The admixture ratio is indicated by the additional parameters "ra" and "1-ra," representing the genetic contribution of the ancestral population POP 1 and POP5, respectively.

The settings for prior of each parameter are listed in Table S1. We used five kinds of summary statistics for each population as well as for each population pair as follows: the proportion of zero values and mean of complete distribution in genetic diversities; the proportion of zero values, mean of non-zero values, and mean of complete distribution in F_{ST} distances. We simulated 1,000,000 data sets for each scenario. Pre-evaluation of each scenario was performed by PCA (principal component analysis) within DIYABC. We compared the different scenarios by calculating their relative posterior probabilities using a logistic regression method from the 1% of the simulated data sets most closely resembling the observed data set without "linear discriminant analysis on SS" option. Confidence of the scenario choice was assessed by using DIYABC's function "evaluate the confidence in scenario choice" with logistic regression for estimating type I (false positives) and type II (false negative) errors based on computing 1,000 data sets without the "linear discriminant analysis on SS" option. Under the most likely scenario as determined by the logit transformation of parameters, the posterior distributions of original parameters were estimated on the 1% of the simulated data sets most closely resembling the observed data set. We also used the option "model checking" with PCA using DIYABC to assess the goodness of fit of the most likely scenario. This option can be used to evaluate the consistency of the observed data with the posterior predictive distribution of the model for the best scenario.

Table 2 The haplotype diversity (*Hd*) and nucleotide diversity (π) of the populations of *Viola orientalis* based on chloroplast DNA variations

Regions[1]	Nos. of populations	Nos. of haplotypes[2]	*Hd*[3]	π
Honshu (p1–p4)	4	1 (1)	0	0
Kyushu (p5–p9)	5	3 (2)	0.5867 (0.500)	0.00105
Korea (p10–p13)	4	2 (2)	0.5263 (0.5263)	0.0011
Russia (p14, p15)	2	2 (1)	0.5556 (0)	0
Japan (p1–p9)	9	3 (2)	0.4202 (0.354)	0.00074
Continent (p10–p15)	6	3 (2)	0.6322 (0.4598)	0.00096
Total	15	5 (2)	0.6166 (0.486)	0.00102

[1]The population numbers are shown in parentheses, see Table 1 and Fig. 1

[2]The values in parentheses indicate numbers of haplotypes excluding indel characters

[3]The values in parentheses indicate Hd excluding indel characters

Results

Patterns of variability in cpDNA

The length of the *rpl16* intron was 952–985 bp through the accessions, including outgroups. Thirteen site changes and seven indels were detected among all accessions, whereas two site changes and two indels were detected within *Viola orientalis* (excluding "*Viola* sp."). On the basis of the polymorphic characters, five distinct cpDNA haplotypes (Types A–E) were recognized within the species, one or two steps separated from each other in the MJ network (Fig. 1). Haplotype A was widely distributed in the populations from Honshu and Kyushu, as well as from the south of the Korean Peninsula (p1–p5 and p8–p11, respectively). Haplotypes B and C were found only in the Aso populations from Kyushu Island (p6–p9). Haplotype D was recorded in the populations either from the central part of the Korean Peninsula or from Primorsky Krai (p12, p13, and p15), whereas haplotype E was scored in only one population from Primorsky Krai (p14). Most populations were monomorphic; however, four populations (p6–p9) from Kyushu Island showed intrapopulation variation. All three populations of "*Viola* sp." (p16–p18) showed haplotype F, which was separated from haplotypes D and E of *V. orientalis* by seven steps.

Table 2 presents the haplotype diversity (*Hd*) and the nucleotide diversity (π) within each region (Honshu, Kyushu, Korea, and Russia). Haplotype diversity values, including gap characters, ranged from 0 to 0.5867, and those of nucleotide diversity ranged from 0 to 0.0011. The Japanese populations ($Hd = 0.4202$, $\pi = 0.00074$) were found less genetically variable than the continental population ($Hd = 0.6322$, $\pi = 0.00096$).

In the comparison among four regions (Honshu, Kyushu, Korea, and Russia), AMOVA revealed that the highest percentage of variation (54.58%, $P < 0.001$) were distributed among populations within groups (Table 3). If Japanese and continental populations were compared, the same AMOVA parameter reached 72.23%. Conversely, in the comparison between the groups defined by a boundary in the Korean Peninsula (p1–p11 vs. p12–p15), the highest variation was recorded among groups (50.41%, $P < 0.001$).

Phylogenetic analysis based on cpDNA markers

In 22 accessions including outgroups, the length of the *trnL–trnF* region was 682–692 bp, and that of the *atpB–rbcL* region was 678–692 bp. The total length of the combined data set after multiple alignments of the three regions was 2,415 bp; 37 site changes and 16 gaps were found among all accessions including outgroups, whereas 20 site changes and six indels were detected across the accessions of *V. orientalis*. Based on these additional sequences and *rpl16* intron, a phylogenetic tree was constructed using the ML method (Fig. 2). All accessions of *V. orientalis* populations (p1–p15) constituted a monophyletic group with 100% bootstrap probability (BP), although the relationships among these populations have remained unclear. Conversely, the three populations of "*Viola* sp." (p16–p18), made a clade both *V. yubariana* and *V. brevistipulata* accessions, which was supported by 73% BP.

Table 3 Analysis of molecular variance (AMOVA) of *Viola orientalis* based on sequences of the *rpl16* intron of chloroplast DNA

Source of variation	*df*	Sum of squares	Variance components	Percentage of variation (%)	*P* value
Among four regions (Honshu, Kyushu, Korea, and Russia)					
Among groups	3	33.007	0.41550	34.39	< 0.001
Among populations within groups	11	37.740	0.65952	54.58	< 0.001
Within populations	60	8.000	0.13333	11.03	< 0.001
Total	74	78.747	1.20835		
Japanese populations vs. Continental population					
Among groups	1	11.891	0.20455	16.81	< 0.001
Among populations within groups	13	58.856	0.87880	72.23	< 0.001
Within populations	60	8.000	0.13333	10.96	0.023
Total	74	78.747	1.21669		
Between population groups defined by a boundary in the Korean Peninsula[1]					
Among groups	1	26.751	0.79660	50.41	< 0.001
Among populations within groups	13	43.995	0.65019	41.15	< 0.001
Within populations	60	8.000	0.13333	8.44	0.008
Total	74	78.747	1.58012		

[1]The groups were defined according to the cpDNA haplotype distribution (see Fig. 1)

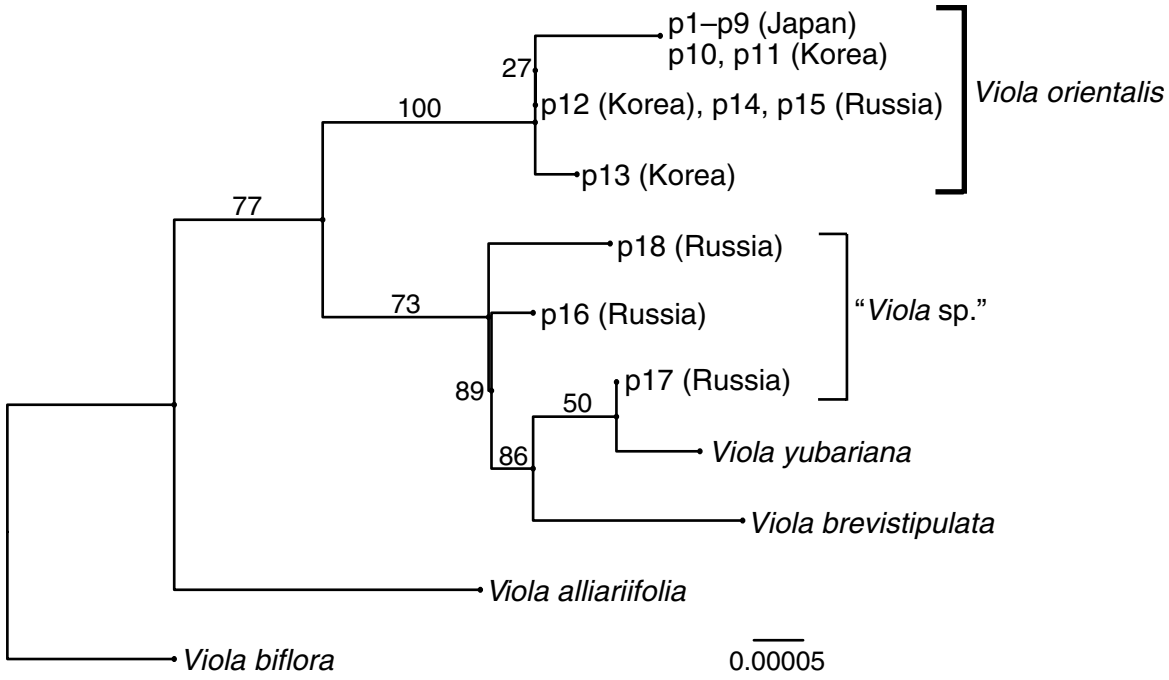

Fig. 2 Phylogenetic tree based on the maximum likelihood (ML) analysis among the populations of *Viola orientalis* using the three plastid regions: *rpl16* intron, *trnL–trnF*, and *atpB–rbcL*. Numbers along the branches indicate bootstrap probabilities (%) based on 1,000 replicates

Genetic diversity estimated by ddRAD-seq data

After SNP filtering for the evaluation of genetic diversity, a total of 78 samples from 15 populations of *V. orientalis* and 15 samples from the three populations of "*Viola* sp." were genotyped as 1,445 biallelic SNPs (Data_2), mean genotyping rate was 0.794 by ddRAD-sequencing. In the filtering procedure, one accession of the Korean population p12 was removed because of low genotyping rate. Table 4 shows the genetic diversity parameters of each population. The values of PA and LCA, indicating genetic uniqueness, were relatively high in populations p10–p15 distributed in South Korea and Russia (PA; 0.003–0.075, LCA; 0.039–0.158). Values of AR and heterozygosities (H_O, H_E, and uH_E), indicating genetic diversity, were found not significantly different among populations (AR; 0.881–1.147, *H*O; 0.036–0.081, *H*E; 0.028–0.079, and u*H*E; 0.035–0.093), although the Russian populations showed slightly less diversity. Inbreeding coefficient (*F*IS) values were relatively high in Russian populations (p14 and p15). However, no significant deviations from the HWE were recorded for any population.

Population genetic structure according to ddRAD-seq data

In the STRUCTURE analysis, the log-likelihood of the data, LnP(*K*), increased gradually when *K* increased from 2 to 6 (Fig. S1a), The ΔK showed the highest peak value at $K=2$ (Fig. S1b), in which a genetic differentiation was observed between the populations from Japan/Korea (p1–p12, red) and those from Russia, including "*Viola* sp." (p14–p18, blue) and outgroup species. The population from Mt. Gari in Korea (p13) was intermediate, in which both red and blue clusters were observed. At $K=3$, the third cluster appeared in the Korean and Russian populations (p13–p15, green) as well as in outgroup species. At $K=4$, the fourth cluster appeared in the Korean populations (p10–p13, purple). At $K=5$, we selected a lower iteration pattern (7/25), since a higher iteration pattern showed relatively low likelihood. This barplot showed that the p13 population from northern Korea (purple) was separated from other Korean populations (p10–p12, pink). At $K=6$, two outgroup species were recruited from an additional cluster (gray) in 17 of 25 iterations. The expected heterozygosity within each cluster, genetic diversity of each cluster, showed relatively high genetic diversity in the populations from Russia (p14, p15), and low values were recorded in the populations from Japan and Korea (p1–p12) (Fig. 3). On the other hand, the F_{ST} value of each cluster, the amount of genetic drift from a common ancestral population, showed relatively high genetic drift in the populations from Japan and Korea, and low values in the populations from Russia (Fig. 3).

The NeighborNet network detected the similar population substructures as inferred by STRUCTURE (Fig. 4). The Japanese (p1–p9) and most of the Korean populations (p10–p12) made a single cluster. Although most

Table 4 Genetic diversity of *Viola orientalis* based on 1,445 SNPs obtained from RAD-seq analysis

Population	n	Mean genotyping rate	Na	PA	LCA	AR	95% CI[1] of AR	H_O	H_E	uH_E	F_{IS}	95% CI[1] of F_{IS}	HWE test[2]
Viola orientalis													
Japan													
p1	5	0.880	1.159	0.000	0.048	1.121	[1.061, 1.159]	0.071	0.061	0.071	−0.129	[−0.452, −0.129]	n.s
p2	5	0.906	1.149	0.001	0.030	1.113	[1.058, 1.149]	0.053	0.057	0.065	0.049	[−0.364, 0.058]	n.s
p3	5	0.885	1.109	0.002	0.025	1.076	[1.018, 1.109]	0.040	0.043	0.050	0.059	[−0.365, 0.087]	n.s
p4	5	0.879	1.123	0.002	0.036	1.085	[1.020, 1.123]	0.053	0.067	0.063	0.026	[−0.429, 0.038]	n.s
p5	5	0.811	1.167	0.001	0.028	1.113	[0.998, 1.167]	0.066	0.066	0.079	−0.009	[−0.421, 0.002]	n.s
p6	5	0.711	1.033	0.001	0.018	0.984	[0.862, 1.033]	0.036	0.029	0.035	−0.183	[−0.646, −0.161]	n.s
p7	5	0.750	1.123	0.000	0.026	1.068	[0.873, 1.123]	0.061	0.053	0.064	−0.123	[−0.603, −0.123]	n.s
p8	5	0.777	1.165	0.001	0.035	1.104	[0.979, 1.165]	0.070	0.053	0.080	−0.040	[−0.433, 0.040]	n.s
p9	5	0.815	1.206	0.001	0.038	1.147	[1.046, 1.206]	0.081	0.079	0.093	−0.025	[−0.435, −0.025]	n.s
Korea													
p10	6	0.809	1.167	0.003	0.066	1.113	[1.022, 1.160]	0.071	0.067	0.077	−0.048	[−0.325, −0.063]	n.s
p11	6	0.822	1.159	0.003	0.063	1.108	[1.022, 1.154]	0.069	0.065	0.074	−0.047	[−0.411, −0.047]	n.s
p12	5	0.798	1.151	0.010	0.048	1.104	[0.992, 1.151]	0.070	0.059	0.069	−0.140	[−0.846, −0.140]	n.s
p13	6	0.809	1.098	0.075	0.039	1.064	[1.001, 1.091]	0.065	0.053	0.061	−0.186	[−0.467, −0.186]	n.s
Russia													
p14	5	0.605	0.947	0.015	0.141	0.881	[0.719, 0.947]	0.048	0.028	0.035	−0.369	[−0.892, −0.342]	n.s
p15	5	0.649	1.053	0.012	0.158	0.976	[0.842, 1.053]	0.064	0.044	0.054	−0.257	[−0.655, −0.257]	n.s
"*Viola* sp."													
p16	5	0.772	1.078	0.026	0.197	1.019	[0.893, 1.078]	0.053	0.045	0.053	−0.126	[−0.564, −0.119]	n.s
p17	5	0.685	0.997	0.010	0.207	0.953	[0.871, 0.997]	0.060	0.039	0.047	−0.359	[−0.758, −0.339]	n.s
p18	5	0.736	0.973	0.019	0.204	0.945	[0.880, 0.973]	0.052	0.030	0.035	−0.439	[−0.941, −0.403]	n.s

Na: mean number of different alleles, PA: mean number of alleles unique to a single population, LCA: mean number of locally common alleles found in 25% or fewer populations, AR: allelic richness, H_O: observed heterozygosity, H_E: expected heterozygosity, uH_E: unbiased expected heterozygosity, F_{IS}: inbreeding coefficient

[1]95% CI 95% confidence interval estimated by 10,000 bootstraps

[2]HWE test were carried out by 9,999 MCMC replications

accessions belonging to a same population were shown to be closely related, three Korean accessions (p10_2, p12_6, and p13_5) were included in the Japanese Kyushu cluster. The northern Korean population (p13) and the Russian populations (p14 and p15) were differentiated among the populations, respectively. The populations of "*Viola* sp." (p16–p18) formed their own cluster, whereas the outgroup species, *V. brevistipulata* and *V. biflora*, were connected on the basis of the cluster.

Phylogenetic analysis using ddRAD-seq data

Phylogenetic analyses using RAxML identified two well-supported clades within the accessions (Fig. 5); one clade included the accessions of *V. orientalis*, whereas the other clade included those of *V. brevistipulata* and *"Viola* sp." (p16–p18). In the clade constructed with *V. orientalis*, the accessions of two Russian populations (p14 and p15) were located at the base of the clade supported by 100% BP. Subsequently, the accessions in northern Korean populations (p12 and p13) diverged, and those of the southern Korean populations (p10 and p11) formed a clade with those of all Japanese populations with 96% BP. In this clade, the accessions of Honshu (p1–p4) and Kyushu populations (p5–p9) were shown to be monophyletic with 75% and 44% BP, respectively. Each population constituted a monophyletic group except for Shizuoka populations (p1 and p2).

DIYABC analysis using ddRAD-seq data

By comparison of the seven models suggested, the highest values of posterior probability were obtained for the model 1 (0.9887; 95% CI 0.9824–0.9950). The posterior probability value did not overlap with the 95% CI of other models (Table S2). Most of the 40 summary statistics of simulated data for the model 1 did not significantly

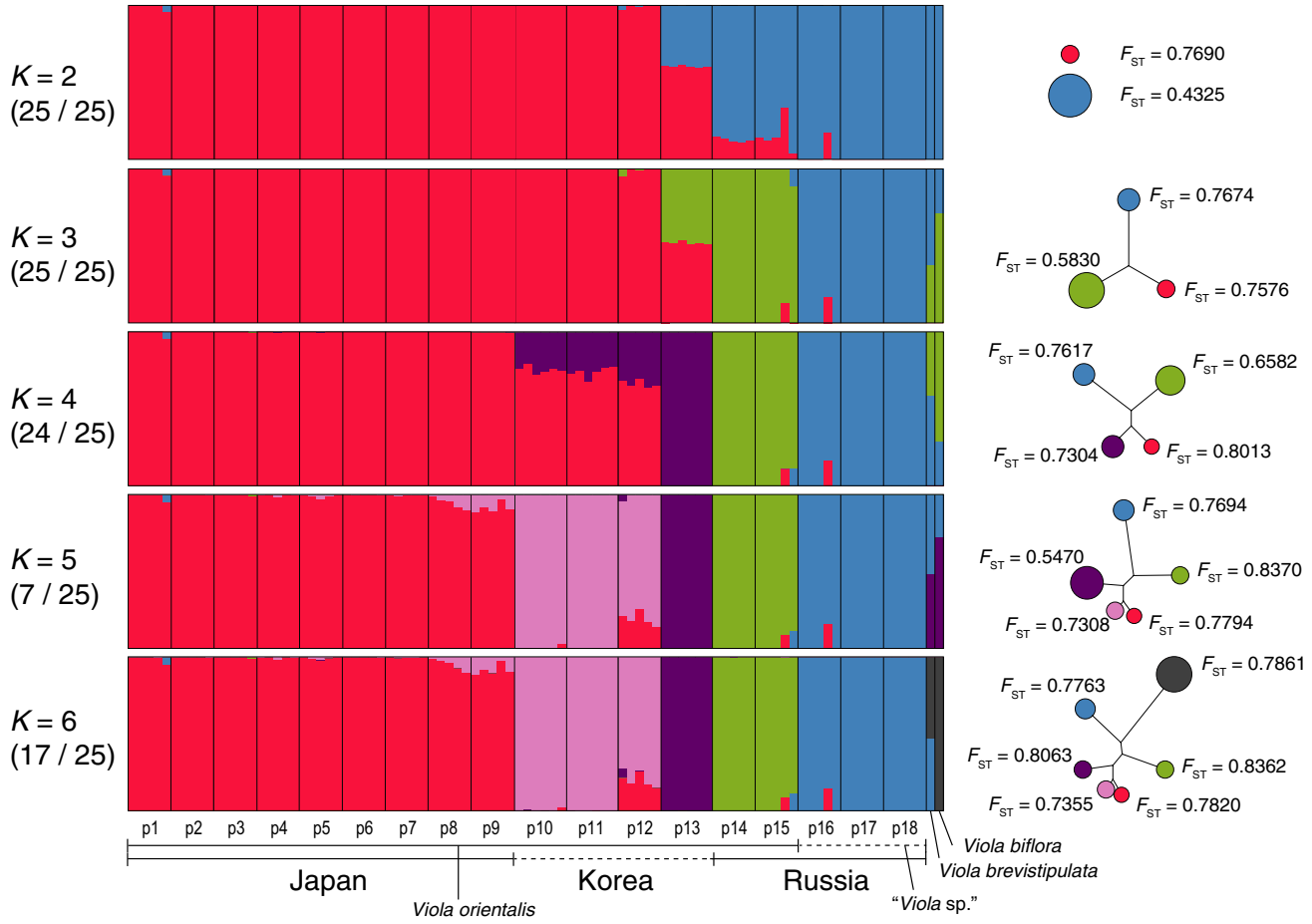

Fig. 3 The bar plots show the probabilities of ancestral clusters of each sample as estimated by STRUCTURE analyses using 2,790 SNPs. Populations and their geographic areas are shown below the bar plot (p1–p18). On the right side of the bar plots, the F_{ST} value of each cluster, the amount of drift of each cluster from a common ancestral population, were indicated. Circle size represents an expected heterozygosity value within each cluster

deviate from those of observed data (Table S3) and the position of the observed data is in close proximity to the simulated data cluster in the PCA scatterplot (Fig. S2). The type I error for the model 1 was 0.0372, whereas the type II errors for this model compared to the other six models (models 2–7) were 0.5020. In the model 1, the median values of the effective population size of N1 (Japan), N2 (Korea), N3 (northern Korea), N4 (Southern Russia), N5 ("*Viola* sp."), and Nanc (putative ancestral population) were 2.07×10^4 (95% CI = 1.18×10^4–2.86×10^4), 2.79×10^4 (95% CI = 2.36×10^4–2.98×10^4), 1.59×10^4 (95% CI = 7.18×10^3–2.99×10^4), 2.70×10^4 (95% CI = 2.03×10^4–2.97×10^4), 2.48×10^4 (95% CI = 1.57×10^4–2.94×10^4), and 5.03×10^4 (95% CI = 9.54×10^3–7.26×10^4), respectively (Table S4 and Fig. S3). The divergence times when N1 split from N2 (t1), N2 split from N3 (t2), N3 split from N4 (t3), and N4 split from N5 (t4) were estimated to be: 9.80×10^3 (95% CI = 5.86×10^3–1.38×10^4), 4.31×10^4 (95% CI = 3.13×10^4–4.90×10^4), 6.31×10^4 (95% CI = 4.00×10^4–8.79×10^4), and 7.39×10^4 (95% CI = 3.89×10^4–1.28×10^5) generations ago, respectively (Table S4 and Fig.S3). The generation time of the *Viola* genus is still unresolved; however, most species of this genus are small herbs, and the period from a seed to flowering is at most 2–3 years (Ballard et al. 2013; Hama 2002). By assuming 3–6 years as the generation time of *V. orientalis*, the scaled divergence time of t1, t2, t3, and t4 were 29,400–58,800 (95% CI = 17,580–82,800), 129,300–258,600 (95% CI = 93,900–294,000), 189,300–378,600 (95% CI = 120,000–527,400), and 221,700–443,400 (95% CI = 116,700–768,000) years ago, respectively. The time divergence of the continental populations (t2 and t3) were approximately 100,000 years earlier than that of the Japanese population (t1).

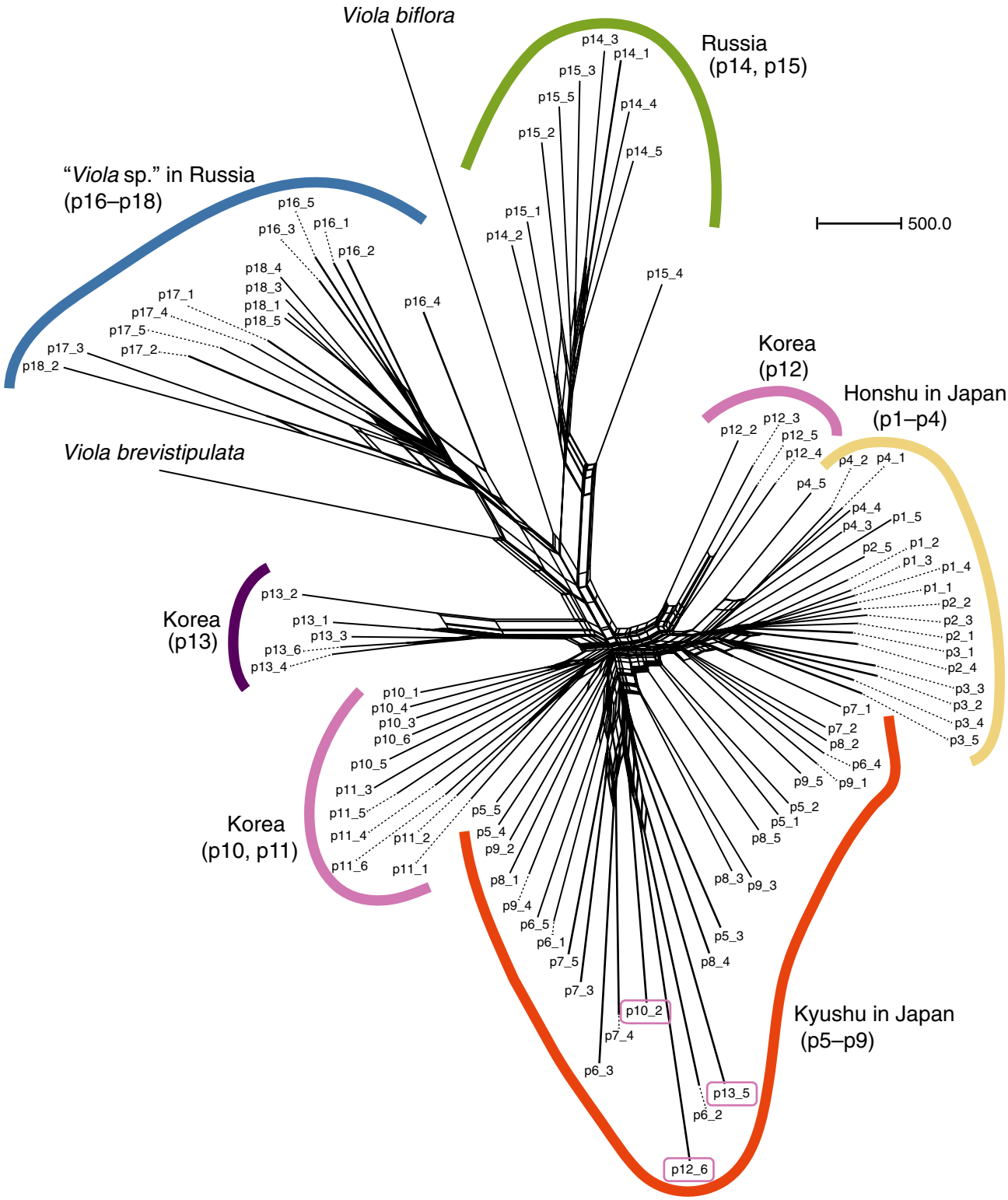

Fig. 4 NeighborNet dendrogram of 15 populations of *Viola orientalis* and three other species constructed using SplitsTree. The population number corresponded to those shown in Table 1. Pairwise genetic distances among individuals were calculated as binary Genetic Distances (GD)

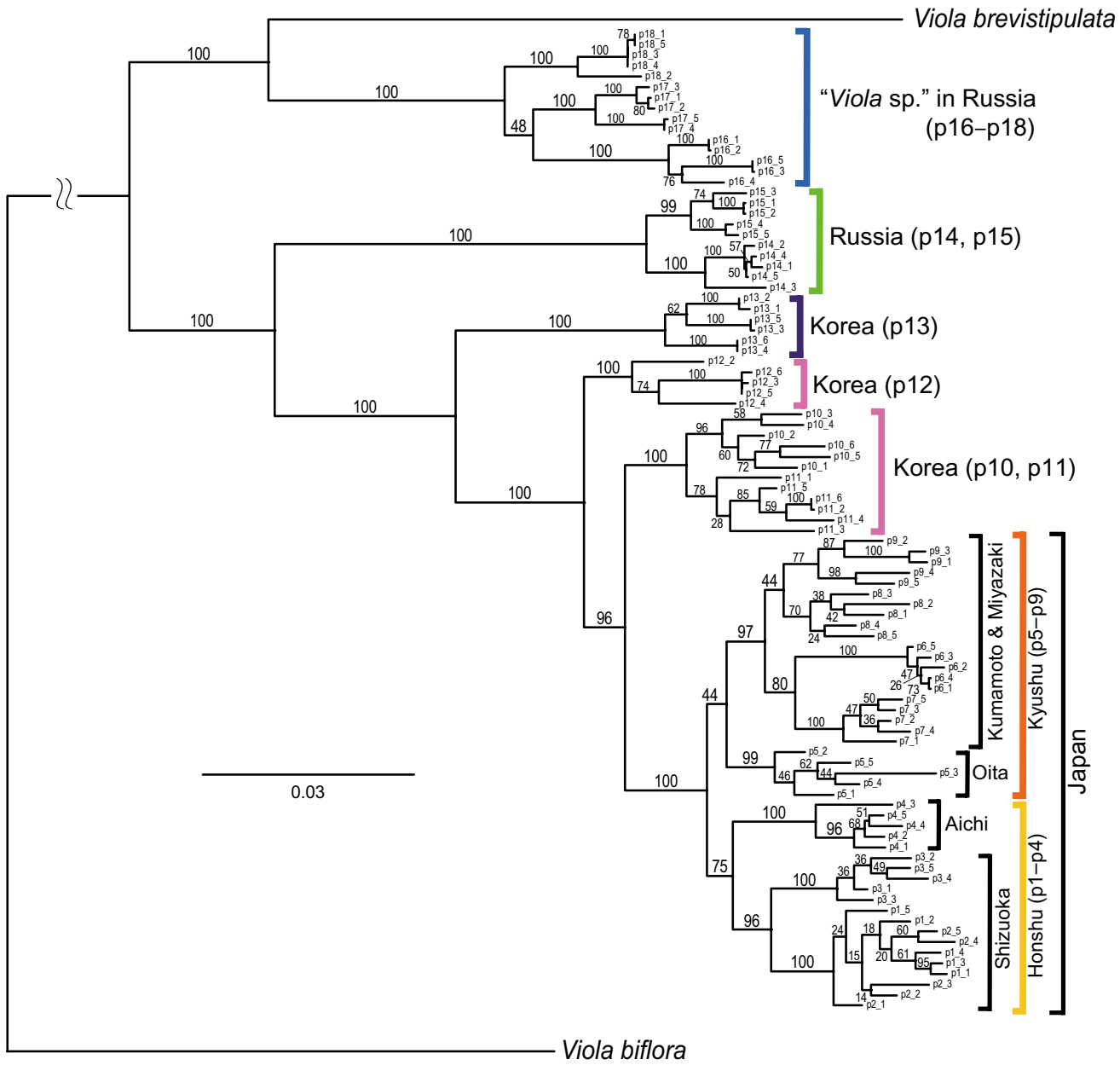

Fig. 5 Phylogenetic tree of *Viola orientalis* and other three species accessions constructed with RAxML from 8,542 SNPs based on maximum likelihood (ML) reconstruction. The numbers along the branches indicate bootstrap values (%) with 1,000 replicates. The scale bar shows the number of substitutions per site. The population numbers corresponded to those shown in Table 1

Discussion

Genetic structure of *V. orientalis*

The cpDNA analyses showed the genetic differentiation between the populations of *V. orientalis* in Japan/southern Korea (p1–p11) bearing haplotypes A–C and northern Korea/Russia (p12–p15) bearing haplotypes D and E (Fig. 1). This genetic differentiation was also supported by AMOVA (Table 3). Additionally, the ddRAD-Seq data analyses showed a similar differentiation pattern in STRUCTURE analysis, NeighborNet analysis, and RAxML tree (Figs. 3, 4, 5). However, the Korean p12 population showed a slightly different genetic composition from both the Japanese and other Korean populations. Thus, both data consistently exhibited the genetic closeness between the southern Korean population (p10 and p11) and all Japanese ones (p1–p9). This result suggests that the ancestral polymorphisms between southern Korean and Japanese populations have retained after a recent vicarience event of their

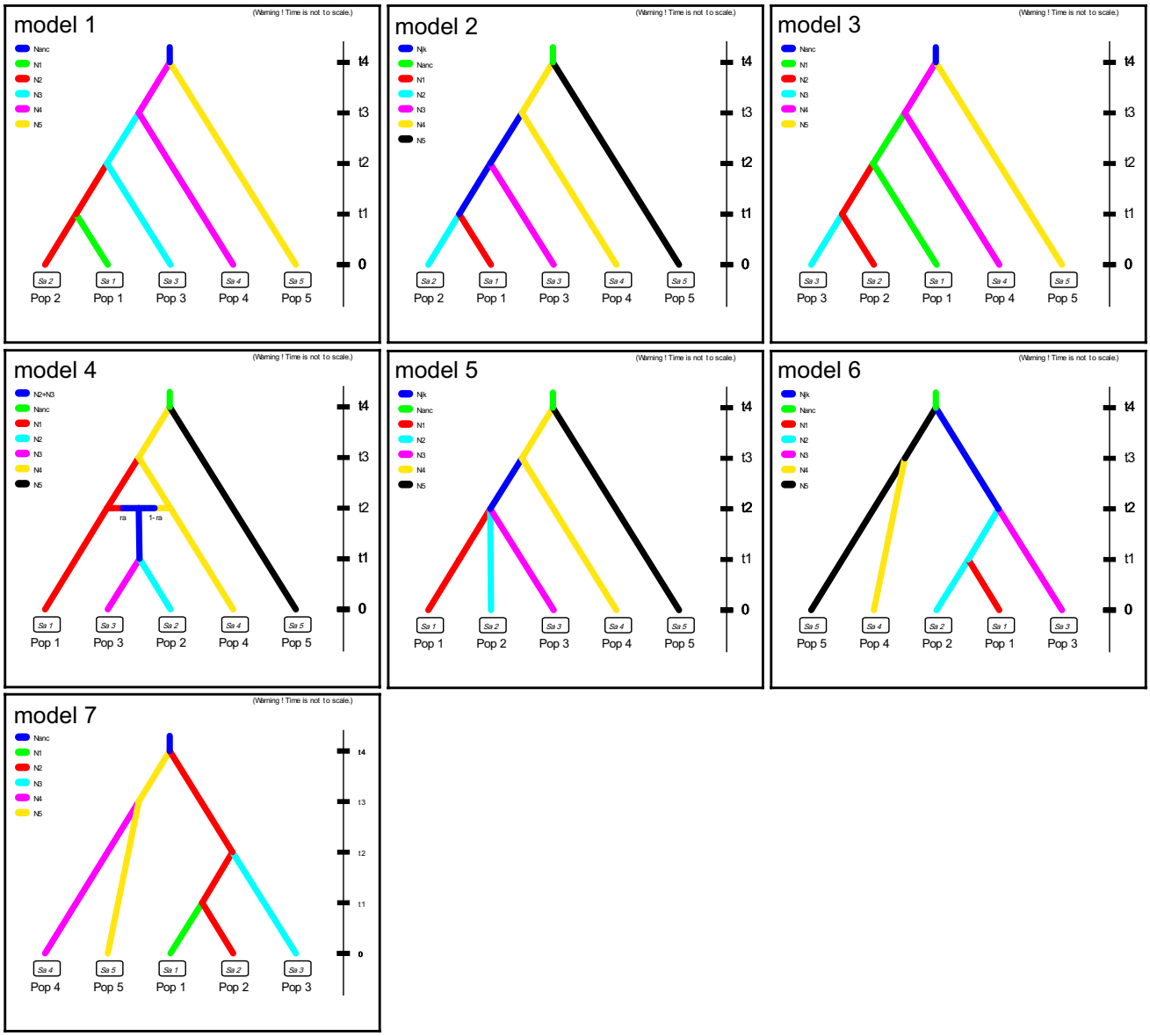

Fig. 6 Seven competing models were designed for inferring the populations' demographic histories using DIYABC analysis. Considering the results of the STRUCTURE, NeighborNet, and phylogenetic analyses, we classified the 18 populations into the following five large regional population groups: POP1 consists of the Japanese populations (p1–p9); POP2 included the Korean populations (p10–p12); POP3 covered a northern Korean population (p13); POP4 involved the Russian populations (p14 and p15); POP5 consisted of "*Viola* sp." populations (p16–p18)

widespread distribution, or that there has been gene flow between them in the near past, regardless of their current geographic separation by the Tsushima Strait (Fig. 1). In addition, this genetic structure may indicate that Japan and the Korean Peninsula formed a land bridge in the near past (see below).

Origin and migration history of *V. orientalis*

The present phylogenetic analysis using SNP data showed that the continental accessions, Russian and Korean populations, were located at the tree base, whereas the Japanese accessions were positioned derivatively (Fig. 5). This divergent process was supported as the most likely model of demographic history in DIYABC analysis (model 1 in Fig. 6). The STRUCTURE analyses inferred that the Japanese (p1–p9) and southern Korean populations (p10–p12) showed low genetic diversity and high genetic drift, whereas the northern Korean population and Russian populations (p13–p15) were characterized by high genetic diversity and low genetic drift (Fig. 3). Additionally, the network analysis among cpDNA haplotypes showed that the haplotypes

in northern Korea/Russia populations (D and E) connected them to haplotype F, considered to originate from a closely related species (see below for details). Such finding suggests that populations in northern Korea/Russia populations are relatively ancestral among the populations of *V. orientalis*. Hence, these results indicate that *V. orientalis* originated in the continental area and migrated to Japan via the Korean Peninsula.

The scaled divergence time by DIYABC analysis showed that the Russian and Korean/Japanese populations diverged approximately 189,300–378,600 years ago, and Korean and Japanese populations diverged about 29,400–58,800 years ago (Table S4). These temporal frameworks correspond to the middle to late Pleistocene, in which there were four glacial-interglacial cycles in the past 400,000 years (Lisiecki and Raymo 2005). Thus, these cycles of climate change would have a significant impact on the genetic divergence and range dynamics of *V. orientalis*. Given that the Japanese archipelago and the Korean Peninsula were almost connected during the cold glacial period due to a lowering of the sea level (Saito et al. 2006; Tsutsumi 2014), the land bridge may have enabled *V. orientalis* to disperse into Japan during the last glacial period.

Previous phytogeographic studies hypothesized that the species in the Mansen elements originated in the Asian continent and have migrated to Japan via the Korean Peninsula under cold climate in the Pleistocene (Hotta 1974; Kitamura 1957; Murata 1988; Tabata 1997). The current study provided evidence for a continental origin of *V. orientalis* and their range expansion to Japan in the Late Pleistocene. Thus, our finding is the first evidence supporting the earlier hypothesis on the origin and range dynamics of Mansen elements. This study will provide essential insights into the establishment process of grassland flora in Northeast Asia. Further comparative studies including more species would provide the general patterns of origin and migration history in the Mansen elements.

In phylogeographic studies of the Japanese archipelago, there is a relatively large number of studies that inferred migration history through the Sakhalin and the Kuril Islands in the north (Fujii and Senni 2006; Hata et al. 2017; Ikeda et al. 2018, 2020) and the Ryukyu Islands in the south (Nakamura et al. 2010; Setoguchi et al. 2006, 2008). However, few studies dealt with grassland plants with a migration history through the Korean Peninsula, such as the Mansen elements. Here, we found that *V. orientalis* originated in the Asian continent and expanded its distribution to Japan via the Korean Peninsula. Therefore, this study significantly elucidates the formation process of the flora from the Japanese archipelago.

Taxonomy of Russian samples of *V. orientalis*

According to Bezdeleva (1987), samples from the Primorsky Krai region (p16–p18) were assigned to *V. orientalis*. However, the present cpDNA and ddRAD-seq data suggests that these samples were not *V. orientalis* but should be treated as "*Viola* sp.." The phylogenetic relationships revealed that the plants of "*Viola* sp." are related to *Viola yubariana* and *V. brevistipulata* rather than to *V. orientalis* (Figs. 2, 5). *Viola yubariana* and *V. brevistipulata* as well as *V. orientalis* are included in the Sect. *Chamaemelanium* (Akiyama et al. 1999; Hama 2002; Kadota 2016). The former species is endemic to Mt. Yubari in Hokkaido, Japan. The latter is also endemic to Japan and is distributed on the Sea of Japan side from the western Honshu to southwestern parts of the Hokkaido (Kadota 2016). Further studies are required for the taxonomic reexamination of "*Viola* sp." in the northern Primorsky Krai region.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10265-021-01339-8.

Acknowledgements We express our sincere thanks to M. Igari, S. Sei, K. Takayama, M. Watanabe, and Y. Yuasa (alphabetical order, title omitted) for sample collection of plant materials. We are grateful for the public institutions that permitted us permission to collect samples. We thank S. Kondo for assistance with ddRAD-seq experiments. We also extend our appreciation for the members of our laboratory in supporting our research. We would like to thank Enago (https://www.enago.jp/) for English language editing. Our cordial thanks are going to the referee who kindly gave us advices to revise our manuscript. This study was supported by JSPS KAKENHI Grant Number JP17H03721 and a grant from the Institute of Plant Science and Resources, Okayama University (3037, 3140, & R236).

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat Control 19:716–723. https://doi.org/10.1109/TAC.1974.1100705

Akiyama S, Ohba H, Tabuchi S (1999) Violaceae. In: Iwatsuki K, Boufford DE, Ohba H (eds) Flora of Japan 2c. Kodansha, Tokyo, pp 161–190

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA (2016) Harnessing the power of RADseq for ecological and evolutionary genomics. Nat Rev Genet 17:81–92. https://doi.org/10.1038/nrg.2015.28

Axelrod DI, Al-Shehbaz I, Raven PH (1996) History of the modern flora of China. In: Zhang AL, Wu SG (eds) Floristic characteristics and diversity of East Asian Plants. Springer, NewYork, pp 43–55

Ballard HE, Paula-Souza Jd, Wahlert GA (2013) Violaceae. In: Kubitzki K (ed) Flowering Plants 11 eudicots: malpighiales. Springer Science & Business Media, Berlin, pp 303–322

Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. Mol Bio Evol 16:37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036

Batchelor CL, Margold M, Krapp M, Murton DK, Dalton AS, Gibbard PL, Stokes CR, Murton JB, Manica A (2019) The

configuration of Northern Hemisphere ice sheets through the Quaternary. Nat Commun 10:3713. https://doi.org/10.1038/s41467-019-11601-2

Beattie AJ, Lyons N (1975) Seed dispersal in *Viola* (Violaceae): adaptations and strategies. Am J Bot 62:714–722. https://doi.org/10.1002/j.1537-2197.1975.tb14104.x

Bezdeleva (1987) *Viola orientalis* (Maxim.) W. Beck. In: Kharkev (ed) Pl Vasc Or Extr Soviet, vol 2. Nauka, St.-Petersburg, pp 129–130 **(In Russian)**

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Bryant D, Moulton V (2004) Neighbor-net: an agglomerative method for the construction of phylogenetic networks. Mol Bio Evol 21:255–265. https://doi.org/10.1093/molbev/msh018

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci de novo from short-read sequences. G3 (bethesda) 1:171–182. https://doi.org/10.1534/g3.111.000240

Cornuet J-M, Pudlo P, Veyssier J, Dehne-Garcia A, Gautier M, Leblois R, Marin J-M, Estoup A (2014) DIYABC v2.0: a software to make approximate Bayesian computation inferences about population history using single nucleotide polymorphism, DNA Sequence and Microsatellite Data. Bioinformatics 30:1187–1189. https://doi.org/10.1093/bioinformatics/btt763

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. Nat Methods 9:772–772. https://doi.org/10.1038/nmeth.2109

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conserv Genet Resour 4:359–361. https://doi.org/10.1007/s12686-011-9548-7

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol Ecol 14:2611–2620. https://doi.org/10.1111/j.1365-294X.2005.02553.x

Excoffier L, Lischer HE (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. https://doi.org/10.1111/j.1755-0998.2010.02847.x

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

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x

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

Fujii N, Ueda K, Watano Y, Shimizu T (1997) Intraspecific sequence variation of chloroplast DNA in *Pedicularis chamissonis* Steven (Scrophulariaceae) and geographic structuring of the Japanese "Alpine" plants. J Plant Res 110:195–207. https://doi.org/10.1007/BF02509308

Good R (1974) The geography of the flowering plants, 4th edn. Longman, London

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 52:696–704. https://doi.org/10.1080/10635150390235520

Hama E (2002) The wild violets of Japan in color, enlarged. Seibundo-Shinko-sha, Tokyo (**In Japanese**)

Harrison SP, Yu G, Takahara H, Prentice IC (2001) Diversity of temperate plants in east Asia. Nature 413:129–130. https://doi.org/10.1038/35093166

Hata D, Higashi H, Yakubov V, Barkalov V, Ikeda H, Setoguchi H (2017) Phylogeographical insight into the Aleutian flora inferred from the historical range shifts of the alpine shrub *Therorhodion camtschaticum* (Pall.) Small (Ericaceae). J Biogeogr 44:283–293. https://doi.org/10.1111/jbi.12876

Hotta M (1974) History and geography of plants. Sanseido, Tokyo (**In Japanese**)

Huson DH, Bryant D (2006) Application of phylogenetic networks in evolutionary studies. Mol Biol Evol 23:254–267. https://doi.org/10.1093/molbev/msj030

Igari M (2004) Wild violets of Japan, enlarged and, revised. Yama-to-Keikoku-sha, Tokyo (**In Japanese**)

Ikeda H, Yakubov V, Barkalov V, Setoguchi H (2014) Molecular evidence for ancient relicts of arctic-alpine plants in East Asia. New Phytol 203:980–988. https://doi.org/10.1111/nph.12863

Ikeda H, Sakaguchi S, Yakubov V, Barkalov V, Setoguchi H (2016) Importance of demographic history for phylogeographic inference on the arctic-alpine plant *Phyllodoce caerulea* in East Asia. Heredity 116:232. https://doi.org/10.1038/hdy.2015.95

Ikeda H, Yakubov V, Barkalov V, Setoguchi H (2018) Post-glacial East Asian origin of the alpine shrub *Phyllodoce aleutica* (Ericaceae) in Beringia. J Biogeogr 45:1261–1274. https://doi.org/10.1111/jbi.13230

Ikeda H, Yakubov V, Barkalov V, Sato K, Fujii N (2020) East Asian origin of the widespread alpine snow-bed herb, *Primula cuneifolia* (Primulaceae), in the northern Pacific region. J Biogeogr 47:1–13. https://doi.org/10.1111/jbi.13918

Jordan WC, Courtney MW, Neigel JE (1996) Low levels of intraspecific genetic variation at a rapidly evolving chloroplast DNA locus in North American duckweeds (Lemnaceae). Am J Bot 83:430–439. https://doi.org/10.1002/j.1537-2197.1996.tb12724.x

Kadota Y (2016) Violaceae. In: Ohashi H, Kadota Y, Murata J, Yonekura K, Kihara H (eds) Wild Flowers od Japan, Rosaceae-Meliaceae, vol 3. Heibonsha, Tokyo, pp 209–227

Keenan K, McGinnity P, Cross TF, Crozier WW, Prodöhl PA (2013) diveRsity: an R package for the estimation and exploration of population genetics parameters and their associated errors. Methods Ecol Evol 4:782–788. https://doi.org/10.1111/2041-210X.12067

Kim K-J, Choi K-S, Jansen RK (2005) Two chloroplast DNA inversions originated simultaneously during the early evolution of the sunflower family (Asteraceae). Mol Biol Evol 22:1783–1792. https://doi.org/10.1093/molbev/msi174

Kitamura S, Murata G, Hori M (1957) Colored illustrations of herbaceous plants of Japan, vol 1 (Sympetalae). Hoikusha, Osaka (**In Japanese**)

Koidzumi G (1931) Zengen (Preface). Florula austro-higoensis. Private publication (**In Japanese**)

Leache AD, Banbury BL, Felsenstein J, de Oca AN, Stamatakis A (2015) Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. Syst Biol 64:1032–1047. https://doi.org/10.1093/sysbio/syv053

Lee JH, Lee DH, Choi BH (2013) Phylogeography and genetic diversity of East Asian *Neolitsea sericea* (Lauraceae) based on variations in chloroplast DNA sequences. J Plant Res 126:193–202. https://doi.org/10.1007/s10265-012-0519-1

Li EX, Yi S, Qiu Y-X, Guo J-T, Comes HP, Fu C-X (2008) Phylogeography of two East Asian species in *Croomia* (Stemonaceae) inferred from chloroplast DNA and ISSR fingerprinting variation. Mol Phylogenet Evol 49:702–714. https://doi.org/10.1016/j.ympev.2008.09.012

Lisiecki LE, Raymo ME (2005) A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. Paleoceanography 20:PA1003. https://doi.org/10.1029/2004PA001071

Liu J-Q, Sun Y-S, Ge X-J, Gao L-M, Qiu Y-X (2012) Phylogeographic studies of plants in China: advances in the past and directions in the future. J Syst Evol 50:267–275. https://doi.org/10.1111/j.1759-6831.2012.00214.x

Lu R-S, Chen Y, Tamaki I, Sakaguchi S, Ding Y-Q, Takahashi D, Li P, Isaji Y, Chen J, Qiu Y-X (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. https://doi.org/10.1016/j.ympev.2019.106693

Maekawa F (1998) The origin of plants (Syokubutsu-no-kitamichi). Yasaka syobou, Tokyo (**In Japanese**)

Meng H-H, Gao X-Y, Huang J-F, Zhang M-L (2015) Plant phylogeography in arid Northwest China: retrospectives and perspectives. J Syst Evol 53:33–46. https://doi.org/10.1111/jse.12088

Mittermeier RA, Robles Gil P, Hoffmann M, Pilgrim JT, Brooks T, Mittermeier CG, Lamoureux J, da Fonseca GAB (2004) Hotspots revisited: earth's biologically wealthiest and most threatened ecosystems. CEMEX, Mexico City

Mittermeier RA, Turner WR, Larsen FW, Brooks TM, Gascon C (2011) Global biodiversity conservation: the critical role of hotspots. In: Zachos F, Habel J (eds) Biodiversity hotspots. Springer, Berlin, pp 3–22

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. https://doi.org/10.1016/j.quaint.2015.03.008

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

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

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**)

Nakamura K, Denda T, Kokubugata G, Suwa R, Yang T, Peng CI, Yokota M (2010) Phylogeography of *Ophiorrhiza japonica* (Rubiaceae) in continental islands, the Ryukyu Archipelago, Japan. J Biogeogr 37:1907–1918

Noda M (1971) Flora of the North-East Province (Manchuria) of China. Kazama Shobo, Tokyo (**In Japanese**)

Oh S-Y, Pak J-H (2001) Distribution maps of vascular plants in Korea. Academy Book Publishing Co., Seoul (**In Korean**)

Ohwi J (1953) Flora of Japan. Shibundo, Tokyo

Ohwi J (1983) Flora of Japan, revised. Shibundo, Tokyo (**In Japanese**)

Ooi N (2016) Vegetation history of Japan since the last glacial based on palynological data. Jpn J Hist Bot 25:1–101. https://doi.org/10.34596/hisbot.25.1-2_1

Peakall R, Smouse P (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research–an update. Bioinformatics 28:2537–2539. https://doi.org/10.1093/bioinformatics/bts460

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLoS ONE 7:e37135. https://doi.org/10.1371/journal.pone.0037135

Polzin T, Daneshmand SV (2003) On Steiner trees and minimum spanning trees in hypergraphs. Oper Res Lett 31:12–20. https://doi.org/10.1016/S0167-6377(02)00185-2

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

Pritchard JK, Wen W, Falush D (2010) Documentation for STRUCTURE software: Version 2. University of Chicago, Chicago

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, De Bakker PI, Daly MJ (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81:559–575. https://doi.org/10.1086/519795

Qiu Y-X, Qi X-S, Jin X-F, Tao X-Y, Fu C-X, Naiki A, Comes HP (2009a) Population genetic structure, phylogeography, and demographic history of *Platycrater arguta* (Hydrangeaceae) endemic to East China and South Japan, inferred from chloroplast DNA sequence variation. Taxon 1226–1241. http://www.jstor.org/stable/27757014

Qiu Y-X, Sun Y, Zhang X-P, Lee J, Fu C-X, Comes HP (2009b) Molecular phylogeography of East Asian *Kirengeshoma* (Hydrangeaceae) in relation to Quaternary climate change and landbridge configurations. New Phytol 183:480–495. https://doi.org/10.1111/j.1469-8137.2009.02876.x

Qiu Y-X, Fu C-X, Comes HP (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. https://doi.org/10.1016/j.ympev.2011.01.012

Rambaut A (2009) FigTree ver. 1.4.4, http://tree.bio.ed.ac.uk/software/figtree/

Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sánchez-Gracia A (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol 34:3299–3302. https://doi.org/10.1093/molbev/msx248

Saito Y, Yokoyama K, Tsutsumi Y, Tanimura Y (2006) Formation of the Japanese Islands. In: National Science Museum (ed) Natural History of the Japanese Islands. Tokai University Press, pp 23–40 **(In Japanese)**

Sakaguchi S, Qiu Y-X, Liu Y-H, Qi X-S, Kim S-H, Han J, Takeuchi Y, Worth JR, Yamasaki M, Sakurai S, Isagi Y (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. https://doi.org/10.1111/j.1365-294X.2012.05652.x

Sakaguchi S, Horie K, Ishikawa N, Nagano AJ, Yasugi M, Kudoh H, Ito M (2017) Simultaneous evaluation of the effects of geographic, environmental and temporal isolation in ecotypic populations of *Solidago virgaurea*. New Phytol 216:1268–1280. https://doi.org/10.1111/nph.14744

Sakaguchi S, Kimura T, Kyan R, Maki M, Nishino T, Ishikawa N, Nagano AJ, Honjo MN, Yasugi M, Kudoh H, Li P, Choi HJ, Chernyagina OA, Ito M (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. https://doi.org/10.1093/aob/mcx182

Setoguchi H, Yukawa T, Tokuoka T, Momohara A, Sogo A, Takaso T, Peng CI (2006) Phylogeography of the genus *Cardiandra* based on genetic variation in cpDNA sequences. J Plant Res 119:401–405. https://doi.org/10.1007/s10265-006-0283-1

Setoguchi H, Watanabe W, Maeda Y, Peng CI (2008) Molecular phylogeny of the genus *Pieris* (Ericaceae) with special reference to phylogenetic relationships of insular plants on the Ryukyu Islands. Plant Syst Evol 270:217–230. https://doi.org/10.1007/s00606-007-0600-8

Shi Y, Ren B, Wang J, Derbyshire E (1986) Quaternary glaciation in China. Quat Sci Rev 5:503–507. https://doi.org/10.1016/0277-3791(86)90217-9

Soltis DE, Gitzendanner MA, Strenge DD, Soltis PS (1997) Chloroplast DNA intraspecific phylogeography of plants from the Pacific Northwest of North America. Plant Syst Evol 206:353–373. https://doi.org/10.1007/BF00987957

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313. https://doi.org/10.1093/bioinformatics/btu033

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

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

Takaishi A, Kozhevnikov AE, Kozhevnikova ZV, Ikeda H, Fujii N, Soejima A (2019) Phylogeography of *Pulsatilla cernua* (Ranunculaceae), a grassland species, in Japan. Ecol Evol 9:7262–7272. https://doi.org/10.1002/ece3.5298

Takhtajan A, Crovello TJ, Cronquist A (1986) Floristic regions of the world. University of California Press, Berkeley

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729. https://doi.org/10.1093/molbev/mst197

Terachi T (1993) Structural alterations of chloroplast genome and their significance to the higher plant evolution. Bull Inst Nati Land Util Developm Kyoto Sangyo Univ 14:138–148

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680. https://doi.org/10.1093/nar/22.22.4673

Tsukada M (1983) Vegetation and climate during the last glacial maximum in Japan. Quat Res 19:212–235. https://doi.org/10.1016/0033-5894(83)90006-6

Tsutsumi Y (2014) An illustrated guide to birth of the Japanese Islands. Kodansha, Tokyo (**In Japanese**)

Union of Japanese Societies for Systematic Biology (2003) Japanese biota species number survey, 1st edn. http://ujssb.org/biospnum/search.php

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

Xia MQ, Liao RY, Zhou JT, Lin HY, Li JH, Li P, Fu CX, Qiu YX (2021) Phylogenomics and biogeography of *Wisteria*: implications on plastome evolution among inverted repeat-lacking clade (IRLC) legumes. J Syst Evol. https://doi.org/10.1111/jse.12733

Xu T, Abbott RJ, Milne RI, Mao K, Du FK, Wu G, Ciren Z, Miehe G, Liu J (2010) Phylogeography and allopatric divergence of cypress species (*Cupressus* L.) in the Qinghai–Tibetan Plateau and adjacent regions. BMC Evol Biol 10:1–10. https://doi.org/10.1186/1471-2148-10-194

Yasuda Y, Miyoshi N (1998) Illustrated vegetation history of Japanese archipelago. Asakura-shoten, Tokyo (**In Japanese**)

Yoo KO, Jang SK (2010) Infrageneric relationships of Korean *Viola* based on eight chloroplast markers. J Syst Evol 48:474–481. https://doi.org/10.1111/j.1759-6831.2010.00102.x

Zeng YF, Wang WT, Liao WJ, Wang HF, Zhang DY (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. https://doi.org/10.1111/mec.13408

Zhao KK, Landrein S, Barrett RL, Sakaguchi S, Maki M, Mu WX, Yang T, Zhu ZX, Liu H, Wang H-F (2019) Phylogeographic analysis and genetic structure of an endemic Sino-Japanese disjunctive genus *Diabelia* (Caprifoliaceae). Front Plant Sci 10:913. https://doi.org/10.3389/fpls.2019.00913