**Article Category:**

Original article

**Title:**

Extremely low level of genetic diversity in *Gentiana yakushimensis*, an endangered species in Yakushima Island, Japan

Naohiro I. Ishii[1,2*], Shun K. Hirota[1], Yoshihiro Tsunamoto[1,3], Ayumi Matsuo[1], Harue Abe[4], Yoshihisa Suyama[1]

[1]Field Science Center, Graduate School of Agricultural Science, Tohoku University, 232-3 Yomogida, Naruko-onsen, Osaki, Miyagi, 989-6711, Japan

[2]Graduate School of Environment and Information Sciences, Yokohama National University, 79-7 Tokiwadai, Hodogaya, Yokohama, Kanagawa, 240-8501, Japan

[3]Research Institute of Energy, Environment, and Geology, Hokkaido Research Organization, Kita19-Nishi12, Kita-ku, Sapporo, Hokkaido, 060-0819, Japan

[4]Sado Station, Sado Island Center for Ecological Sustainability, Niigata University, 94-2 Koda, Sado, Niigata 952-2206, Japan

**[*]Correspondence to:**

Naohiro I. Ishii

Graduate School of Environment and Information Sciences, Yokohama National University, 79-7 Tokiwadai, Hodogaya, Yokohama, Kanagawa 240-8501, Japan

E-mail: na0hiro1shi1@gmail.com

## Abstract

Rare or endangered species largely contributing to global biodiversity are essential components of ecosystems. Because genetic diversity depends not only on apparent population size but also demographic processes, the conservation priority of endangered species based on population genetics needs to be assessed by comparison with common congeners and demographic events. In this study, we performed population genetic analysis for *Gentiana yakushimensis*, a rare and endangered plant species with an isolated distribution on Yakushima Island, Japan. We also compared genetic diversity with the congener (*Gentiana triflora*) and demographic inferences based on genome-wide single nucleotide polymorphisms. The result was that *G. yakushimensis* had an extremely low genetic diversity level and a high inbreeding coefficient level compared to those of the congener. The population of *G. yakushimensis* was estimated to experience gentle expansion during the Holocene and a bottleneck over the past several decades by demographic inference. Therefore, the low level of genetic diversity of this species might have resulted from the impact of the recent bottleneck and the long-term maintenance of a small population size. Due to the long lifespan of the genus *Gentiana*, the current level of genetic diversity probably does not reflect a large part of recent demographic events, which requires long-term monitoring of changes in genetic diversity to identify the time lag between the reduction in apparent population size and genetic diversity as a future outlook.

## Keywords

Genetic differentiation, genetic diversity, genome-wide SNPs, Gentiana

## Introduction

Rare or endangered species, largely contributing to global biodiversity, are essential components of various ecosystems. Most previous studies have demonstrated that rare species

with limited geographic distributions tend to have lower levels of genetic diversity than common species because of their small effective population size (Broadhurst et al., 2017; Chung et al., 2018; Gibson, Rice, & Stucke, 2008; Hamrick & Godt, 1996; Nybom & Bartish, 2000). The low level of genetic diversity in a species or a population can trigger its vulnerability because it could generally result in the reduction in adaptive potential to environmental changes and resilience or resistance to environmental extremes (Reusch, Ehlers, Hammerli, & Worm, 2005; Willi, Van Buskirk, & Hoffmann, 2006). Thus, it is reasonable to globally evaluate the conservation priority of rare or endangered species based on apparent population size and genetic diversity (Garner, Hoban, & Luikart, 2020). Accumulation of genetic information for thousands of endangered species can lay the foundation for biodiversity conservation and conservation prioritization.

The genetic diversity of rare species is not always low, and it can be species-specific because it depends on apparent population size and demographic processes or life history traits (Ellegren & Galtier, 2016; Hamrick & Godt, 1996; Nybom & Bartish, 2000; Thiel-Egenter et al., 2009). For example, previous studies have demonstrated that species with a long lifespan tend to have higher levels of genetic diversity than those with short lifespans (Kort et al., 2021; Reisch & Rosbakh, 2020). Past and current demographic events, such as population bottlenecks or expansion due to environmental and anthropogenic fluctuations, can affect the current genetic diversity while comparing closely related species with similar life history traits (Ellegren & Galtier, 2016). For example, the species widely distributed in the past can become rare or endangered due to disturbances (Li et al., 2014). Conversely, in isolated habitats (e.g. islands), a small population size can be maintained due to local adaptation with limited immigrants (Gonzalez, Cerón-Souza, Mateo, & Zardoya, 2014). Therefore, estimating demographic history allows us to examine which current level of genetic diversity is determined by historical or recent events.

Alpine plants, including many rare or endangered species, are the targets for which conservation prioritization based on genetic aspects is required (Noroozi et al., 2018; Perrigo, Hoorn, & Antonelli, 2020), as these plants are vulnerable to climate change and anthropogenic disturbances (Guisan & Theurillat, 2000; Thuiller, Lavorel, Araújo, Sykes, & Prentice, 2005). Previous studies have demonstrated that the current levels of genetic diversity of alpine plants are strongly affected by several demographic events, such as expansion or reduction of distribution range after the Last Glacial Maximum (LGM; 21,000 years ago) (Loïc Pellissier et al., 2016), geographical isolation among populations by the distribution patterns restricted to high-altitude areas (Reisch & Rosbakh, 2020), and direct human activities such as trampling by tourists and illegal digging (Banks et al., 2013; Feng, Liu, Chiang, & Gong, 2017). Therefore, inferring historical demography allows us to examine the influence of historical demographic events on current genetic diversity levels and accurately assess the conservation status of rare alpine species.

*Gentiana yakushimensis* (Figure 1) is a rare and endangered plant species in Japan (Japanese Ministry of Environment, 2020). *G. yakushimensis* is endemic to Yakushima Island, Japan (Iwatsuki, Yamazaki, Boufford, & Ohba, 1993). On Yakushima Island, designated a UNESCO World Heritage Site, this species grows only in limited alpine zones covered by granite rocks with low vegetation cover in the central mountains (Figure 2). Illegal digging has drastically reduced its population (Japanese Ministry of Environment, 2020). *G. yakushimensis*, or its ancestor, was estimated to colonize Japan 4–1 million years ago with the Qinghai–Tibet Plateau as the origin (Appendix S3 in Favre et al., 2016). The taxonomic position of *G. yakushimensis* is discussed. Recently, phylogenetic analysis based on chloroplast DNA regions suggested that *G. yakushimensis* is close to the section *Pneumonanthe* (Mishiba et al., 2009). Moreover, phylogenetic analysis covering all *Gentiana* sections supported the inclusion of *G. yakushimensis* in *the Pneumonanthe*/*Cruciata* lineage (Favre et al., 2016).

101 The habitat area of *G. yakushimensis* was assumed to be isolated and small because a
102 warm-temperature evergreen broadleaf forest surrounded it during and after the LGM (Tsukada,
103 1985). Thus, *G. yakushimensis* might maintain a small population size from the LGM, although
104 the center of the habitat area is expected to move up and down due to fluctuations in the forest
105 limit caused by changes in temperature. Because of the limited habitat in high-altitude areas,
106 the population size of *G. yakushimensis* may also be affected by the rise of trees and vegetation
107 lines, which is assumed to be similar to what globally occurred with warming after the
108 beginning of the Holocene and/or in recent years ( Harsch, Hulme, McGlone, & Duncan, 2009;
109 Tinner & Kaltenrieder, 2005). Thus, by estimating the population demography of *G.*
110 *yakushimensis*, we can examine whether the current level of genetic diversity resulted from
111 recent and/or historical demographic events.

112 Here, we assessed the genetic diversity and demographic history of *G. yakushimensis*
113 based on genome-wide single nucleotide polymorphisms (SNPs) to identify their conservation
114 status. We performed 1) the comparison of genetic diversity with a congener (*Gentiana triflora*),
115 2) the evaluation of genetic differentiation across Yakushima Island, and 3) the demographic
116 inference of *G. yakushimensis* to identify whether the determinant of the current level of genetic
117 diversity is recent or historical, aiming to provide the genetic information of the endangered
118 species, supporting decision-making in conservation implications, and prioritizing.

119

120 **Materials and Methods**

121 ***Studied species***

122 *G. yakushimensis* (Gentianaceae) is a perennial herb endemic to Yakushima Island,
123 Japan (Chen & Wang, 1999; Ho & Liu, 2001; Iwatsuki et al., 1993). Although Ho (1988)
124 recorded the occurrence of *G. yakushimensis* in Taiwan based on a single specimen (PE No.
125 81854; https://www.cvh.ac.cn/spms/detail.php?id=ec90580f), Chen and Wang (1999) denied

126 the distribution of this species in Taiwan because the specimen lacks collection labels and no
127 other specimens from Taiwan exist. *G. yakushimensis* has a minimal and patchy distribution; it
128 frequently appears on granite rocks around the tops of the central mountains, at altitudes ranging
129 from 1,700 to 1,900 m (Figure 2). *G. yakushimensis* was pollinated by bees (H. Abe, personal
130 observation), and its seeds with no wings were assumed to be dispersed by gravity and wind.
131 The population size of *G. yakushimensis* has decreased due to illegal digging in recent years
132 (Japanese Ministry of the Environment, 2014). In addition, feeding damage by sika deer
133 (*Cervus nippon yakushimae*) might also be a factor in population size reduction. In Yakushima
134 Island, grazing by sika deer has affected vegetation from lowlands to subalpine zones
135 (Takatsuki, 1990; Tsujino & Yumoto, 2004). Takatsuki (1990) demonstrated that sika deer relied
136 on forbs for about 5 % of their total diet in the subalpine area of Yakushima Island, indicating
137 that *G. yakushimensis* could be affected by grazing. Thus, *G. yakushimensis* has been listed in
138 the Japanese Red List as Endangered (EN) (Japanese Ministry of Environment, 2020) and
139 conserved under the "Act on Conservation of Endangered Species of Wild Fauna and Flora"
140 since 2016.

141 *Gentiana triflora* var. *japonica* was used to compare the genetic diversity of *G.*
142 *yakushimensis. G. triflora* var. *japonica* is a relatively common species widely distributed in
143 the mountainous area from Hokkaido to the Kinki region of Japan (Figure 2). This species is
144 one of the most closely related to *G. yakushimensis* based on the chloroplast DNA region (Favre
145 et al., 2016; Mishiba et al., 2009).

146 We collected leaf samples from all individuals of *G. yakushimensis* (560 individuals),
147 which had grown in accessible habitats, for DNA analysis at Yakushima Island in August from
148 2009 to 2010. Small individuals or individuals with few leaves were not collected to conserve
149 *G. yakushimensis*. When multiple individuals grew at the same coordinates, one individual was
150 selected for DNA analysis. Leaf samples were collected with permission from the prefecture

151 and ministries. The leaf samples of *G. triflora* were collected as congeners from an area of
152 approximately 100 $m^2$ in Osaki, Miyagi Prefecture, in the northern area of Japan (Figure 2).
153

154 ***DNA extraction, amplification, and sequencing***

155 Total genomic DNA was extracted from frozen leaf samples using the modified
156 cetyltrimethylammonium bromide (CTAB) method (Doyle, 1991). For DNA analysis, we used
157 353 and 9 samples of *G. yakushimensis* and *G. triflora*, respectively (Table 1). The *G.*
158 *yakushimensis* samples for the analysis were at least 3 m away from each other.

159 To evaluate the genetic diversity of *G. yakushimensis* and *G. triflora*, we performed *de*
160 *novo* single-nucleotide polymorphism (SNP) detection using multiplexed inter-simple sequence
161 repeat (ISSR) genotyping by sequencing (MIG-seq) (Suyama et al., 2021; Suyama & Matsuki,
162 2015). We prepared the MIG-seq library and sequenced it following the protocol described by
163 Suyama et al. (2021). Briefly, ISSRs were amplified using primer set-1 (Suyama & Matsuki,
164 2015) in the first PCR, and then the index and adaptor sequences were added to both ends of
165 the 1st PCR products in the second PCR. We sequenced 80 base pair-end reads using the
166 Illumina MiSeq platform and MiSeq Reagent Kit v3 (150 cycles) (Illumina, San Diego, CA,
167 USA). We skipped detection of the first 17 bases of Reads 1 and 2 using the "DarkCycle" option.
168

169 ***Read assembling and SNP marker filtering***

170 One base at each end of the processed reads was trimmed, and the reads were filtered
171 based on the quality of reads and the adaptor sequences using Trimmomatic v.0.3.2 (Bolger,
172 Lohse, & Usadel, 2014). Therefore, we performed the subsequent analysis using filtered reads
173 with a length of 78 bp. Stacks v2.52 (Catchen, Amores, Hohenlohe, Cresko, & Postlethwait,
174 2011) which was used to process the filtered reads with the default parameter settings: the
175 minimum number of identical reads required to create a stack (m = 3), the nucleotide

176 mismatches between loci within a single individual (M = 2), and mismatches between loci when
177 building the catalog (n = 1). The SNP genotype for each individual was exported with the
178 following parameter setting using the 'populations' command in Stacks v2.52: the maximum
179 observed heterozygosity, --max-obs-het = 0.6; minimum minor allele frequency, --min-maf = 2
180 / the number of individuals. For the dataset of both species (common loci) and individual
181 species (common and species-specific loci; hereafter, referred to as "all loci"), the SNP
182 genotype was exported by setting r (the parameter of the minimum sharing rate of SNP among
183 individuals per population) at 0.5, 0.6, and 0.7, in addition to the above parameter settings to
184 compare the genetic diversity of *G. yakushimensis* and *G. triflora*. We set up each species as a
185 population to meet the r threshold for each species (p = 2 for the dataset of common loci and p
186 = 1 for all loci in the 'populations' command). The comparison with the different values of the
187 parameter r was performed because the number of SNPs of the common loci was low (less than
188 100) when the parameter r was 0.7. The estimation of genetic indices with low parameter r
189 values leads to increase in missing data but ensures accuracy by increasing the number of SNPs
190 (Hodel et al., 2017).
191
192 ***Comparison of genetic diversity between G. yakushimensis and G. triflora***
193 Genetic diversity was compared between *G. yakushimensis* and *G. triflora* for both the
194 common and all loci. We calculated nucleotide diversity ($\pi$) based on fixed and variant sites
195 using the 'populations' command of Stacks v2.52. The other indices, observed and expected
196 heterozygosity ($H_O$ and $H_E$), and inbreeding coefficient ($F_{IS}$), were determined using GenAlEx
197 v6.51 (Peakall & Smouse, 2006). *t*-tests were performed on these indices to test the significance
198 of the difference between the species using the R package BSDA (Arnholt & Evans, 2017) in
199 R v3.6.1.
200

201 ***Genetic differentiation of G. yakushimensis***

202 Subsequent analyses were performed using the dataset of all *G. yakushimensis* loci.
203 Kosman & Leonard's (2005) pairwise genetic distance between *G. yakushimensis* individuals
204 was calculated using the R package PopGenReport (Adamack & Gruber, 2014). Based on
205 genetic and geographic distance matrices among individuals, we assessed spatial
206 autocorrelation (isolation by distance, IBD) across the distribution area using the
207 'mantel.correlog' function in the R package vegan (Oksanen et al., 2020). To test for the
208 presence of genetic structure, we performed principal coordinate analysis (PCoA) based on the
209 genetic distance matrix using the 'pcoa' function in the R package ape (Paradis & Schliep,
210 2019).

211

212 ***Inference of population demographic history***

213 Finally, we inferred the population demographic history of *G. yakushimensis*. In the
214 'populations' command in Stacks v2.52, we exported SNP genotypes for the dataset of the all
215 loci of *G. yakushimensis* with the addition of the following parameter setting: the restriction of
216 output to a random single SNP per locus, --write-random-snp. To avoid the effects of linkage
217 disequilibrium, we removed one out of a pair of SNPs with an $R^2$ value > 0.5 using PLINK
218 v1.90 (Purcell et al., 2007). As a result of the aforementioned filtering, 362, 302, and 232 SNPs
219 were obtained when the parameter r was 0.5, 0.6, and 0.7, respectively.

220 The demography of *G. yakushimensis* was inferred by DIYABC Random Forest v1.0
221 (Cornuet et al., 2014), an approximate Bayesian computation method of inferring historical
222 changes in population size by the supervised machine learning method. We assumed that the
223 lifespan (generation time) of *G. yakushimensis* is 15 years, according to which the lifespan was
224 at least 15 years in a congener, *G. pneumonanthe* (Oostermeijer, Veer, & Nijs, 1994). In this
225 analysis, we estimated that the change in population size occurred twice. The first is around the

226 beginning of the Holocene, including the Bolling-Allerod period and the Younger Dryas period
227 (10,000–15,000 years ago), which coincides with the onset of warming (Stuiver, Grootes, &
228 Braziunas, 1995); and the population size was expected to be reduced via the vegetation uplift
229 caused by warming. Second is within the last 100 years, when the population size may have
230 been reduced by recent events (e.g. recent vegetation uplift, grazing by sika deer, and illegal
231 digging). We examined four demographic population scenarios (Figure S1 and Table S1). In
232 these scenarios, t# represents the time scale measured in the number of generations, and N#
233 represents the effective population size of the corresponding populations during the relevant
234 time period (0–t1, t1–t2, and t2–). We performed 100,000 simulations per scenario by setting
235 N#, t1 (corresponding to the recent 100 years), and t2 (corresponding to around the beginning
236 of the Holocene) to 10–10000 individuals, and 1–10 and 650–1000 generations, respectively
237 (Figure S5). All the parameters were assumed to have a uniform prior distribution. To produce
238 a posterior distribution, we simulated 100,000 datasets, and 1000 trees were used to grow a
239 random forest classification for scenario selection. The compatibility between scenarios,
240 associated priors, and observed data was assessed using linear discriminant analysis (LDA).
241

242 **Results**

243 The mean number of filtered reads per sample of *G. yakushimensis* and *G. triflora* was
244 92,960 and 58,900, respectively. The number of SNPs in the common loci between the two
245 species was 72, 51, and 34, with r = 0.5, 0.6, and 0.7, respectively (Table 1). In the common
246 and species-specific (all) loci of *G. yakushimensis*, the number of SNPs was 784, 562, and 402
247 with r = 0.5, 0.6, and 0.7, respectively (Table 1). On the other hand, the number of SNPs in all
248 loci of *G. triflora* was 401, 297, and 202, with r = 0.5, 0.6, and 0.7, respectively (Table 1).

249 Compared to *G. triflora*, the genetic diversity indices ($\pi$, $H_O$, and $H_E$) of *G.*
250 *yakushimensis* were significantly lower in all parameter settings (Table 1). When comparing *G.*

251 *yakushimensis* with *G. triflora* for common loci with setting r = 0.7 in the 'populations'
252 command, the former was 0.05 times for π and $H_E$, and 0.03 times for $H_O$ as much as the latter
253 (Table 1). The inbreeding coefficients ($F_{IS}$) of *G. yakushimensis* were significantly greater than
254 zero for all parameter settings (Table 1). Also, it was significantly greater than those of *G.*
255 *triflora* except for the common loci with setting r = 0.5 in the 'populations' command.

256 We observed a significant positive correlation between genetic and geographic
257 distances among the individuals of *G. yakushimensis* (Mantel correlation coefficient, $r$ = 0.25–
258 0.27, $P < 0.001$) (Figures 3a, S2a, and S3a). However, an increase in genetic differentiation
259 among the individuals was not detected when the pairwise geographic distance was over
260 approximately 500 m (Figure 3a). As a result of the Mantel correlogram, calculating Mantel
261 correlation coefficients for each distance class, consistently positive correlations were shown
262 up to approximately 140–160 m (Figures 3b, S2b, and S3b). PCoA demonstrated gradual
263 genetic differentiation among the individuals across the distribution area, although the first
264 three axes of PCoA explained only 4–6 % of the whole genetic variation of *G. yakushimensis*
265 (Figure 4, S4, and S5).

266 As a result of the demographic inference by the ABC method, scenario 3 was selected
267 with the highest probability from 0.52 to 0.76 (Figure 5 and Table S2). Subsequently, scenario
268 4 was selected with the second-highest probability from 0.24 to 0.30. The selection probability
269 of scenario 3 was higher than that of scenario 4 for all the settings of parameter r. LDA results
270 showed that the simulated data based on scenarios 3 and 4 were consistent with the observed
271 data, especially on axis 1 of LDA (Figure S6). These scenarios have a characteristic in common
272 that they estimated the increase in effective population size after t2 corresponding to the LGM.
273 On the other hand, scenarios 3 and 4 estimated the different changes in population size after t1.
274 Based on the higher probability of scenario 3, it was presumed that the Yakushima population
275 of *G. yakushimensis* experienced a slight reduction in the effective population size over the

276 recent years. In scenario 3, the median of N0, N1, and N2 was estimated to be 461, 4864, and
277 1071, respectively, when the parameter r was 0.5 (Figure 5 and Table S3). Similar results were
278 obtained with the other parameter r settings (Table S3).
279

280 **Discussion**

281 In this study, we revealed 1) the extremely low level of genetic diversity of *G.*
282 *yakushimensis*, 2) the gentle genetic differentiation across the Yakushima population despite
283 the patchy distribution, and 3) the small population size due to the recent bottleneck and no
284 dramatic increase in population size over the long term.

285 We demonstrated that *G. yakushimensis* had extremely low levels of genetic diversity
286 and a high level of inbreeding coefficient compared to those of its congener (Table 1).
287 Surprisingly, the genetic diversity across only nine individuals from a single population of *G.*
288 *triflora* exceeded that across 353 individuals of *G. yakushimensis*. This result was consistent
289 with that of previous studies that showed that relatively rare species have lower genetic diversity
290 than that of common species (Broadhurst et al., 2017; Chung et al., 2018; Gibson et al., 2008;
291 Hamrick & Godt, 1996; Nybom & Bartish, 2000).

292 There was less genetic differentiation among the fragmented patches of *G.*
293 *yakushimensis* across Yakushima Island. IBD was significant only across a narrow range (up to
294 140–160 m), and the constant pattern of pairwise genetic distance was detected over this range,
295 suggesting that gene flow occurs across the 140–160 m range. The granitic rocks, the habitats
296 of *G. yakushimensis*, are densely distributed in the center of Yakushima Island. Thus, there is
297 gene flow between adjacent habitats, and the patchy habitat network of *G. yakushimensis* might
298 be maintained. These results indicate that *G. yakushimensis* on Yakushima Island comprises
299 almost a genetically single population. *Gentiana yakushimensis* was mainly pollinated by bees
300 (H. Abe, personal observation). The foraging range of bees was at least 100 m for species with

301 the smallest body size (about 1 mm of intertegular span) (Greenleaf, Williams, Winfree, &
302 Kremen, 2007). Generally, bees prioritize the foraging of neighboring flower patches after
303 foraging within a patch (Woodgate, Makinson, Lim, Reynolds, & Chittka, 2017). Thus, the
304 pollination behavior by bees might gently connect patchy habitats and maintain a low level of
305 genetic differentiation.

306 The Yakushima population of *G. yakushimensis* was likely to expand during the
307 Holocene, based on demographic inferences (Figure 5). Although we expected a reduction in
308 population size via a decrease in habitat area, which was associated with the rise of forest and
309 vegetation lines by warming after the beginning of the Holocene (Tinner & Kaltenrieder, 2005),
310 the scenario of an increase of numbers of *G. yakushimensis* after the beginning of the Holocene
311 was supported by ABC analysis. However, because the maximum population size was
312 approximately 5,000 at N1 (Figure 5), one of the possible reasons for low genetic diversity is
313 the long-term maintenance of a small population size. Warming might have provided new
314 habitats at higher altitudes for *G. yakushimensis* (You et al., 2018). A possible explanation is
315 that the center and upper distribution area limits of *G. yakushimensis* during the LGM were at
316 lower altitudes than those in recent years. The large proportion of Yakushima Island is
317 composed of granite (Anma, Yamamoto, Geshi, Nanayama, & Nakagawa, 2014), which
318 provides potential habitats now covered by vegetation. Because the temperature during the
319 LGM was expected to be 5–6 °C lower than the current temperature (Tsukada, 1983), the center
320 of the distribution area is assumed to be located at a lower altitude than that in recent years. In
321 addition, the highest region of Yakushima Island might not have been suitable for *G.*
322 *yakushimensis* before the Holocene. As a result, the habitat expansion to high altitudes during
323 the Holocene might be more significant than the habitat shrinkage by vegetation uplift,
324 including that in the *Cryptomeria japonica* forest. This can be explained by the fact that the
325 range of expansion of herbaceous plants is faster than that of woody plants, which is similar to

the first-order succession pattern of vegetation after deglaciation. Alternatively, the degradation of *C. japonica* forest at an altitude of at least 1,000–1,200 m was caused by the pyroclastic flow from the Kikai Caldera located 50 km northwest of Yakushima Island approximately 6,300 years ago (Kimura, Ooi, & Suzuki, 1996), which could lead to habitat expansion of *G. yakushimensis* to low altitudes despite the temporary decrease in population size by the pyroclastic flow.

The reduction in the apparent number of individuals was likely to be reflected in the current level of the effective population estimated by demographic inference (Figure 5). Therefore, the low level of genetic diversity in *G. yakushimensis* was assumed to be due to the significant impact of the recent bottleneck. There are three possible reasons for this result: illegal digging (Japanese Ministry of Environment, 2020), grazing by sika deer (Takatsuki, 1990, 2009; Tsujino & Yumoto, 2004), and a habitat area reduction by the uplift of forest and vegetation lines by warming (Japanese Ministry of the Environment, 2017). In Yakushima Island, because the forest cover and density of *C. japonica* increased in the high-altitude area from 1963 to 2014 (Japanese Ministry of the Environment, 2017), it is inferred that the increase in the vegetation limit of shrubs and dwarf bamboos reduces the habitat area of *G. yakushimensis*. While the scenario with the recent bottleneck was the most supported, the recent expansion scenario was also selected with a certain probability (Table S2), which might be attributed to the long lifespan of the genus *Gentiana* (more than 15 years) (Oostermeijer et al., 1994). The longevity generally causes a time lag between reducing the apparent population size and the decline in genetic diversity (Fuller & Doyle, 2018; Reisch et al., 2017). This time lag might lead to a further reduction in the genetic diversity of *G. yakushimensis* in the near future.

*G. yakushimensis* has an extraordinarily low level of genetic diversity, resulting in vulnerability to environmental changes and human disturbances (Reusch et al., 2005; Willi et al., 2006). Moreover, the effective range of gene flow in this species is extremely narrow. In

species with limited dispersal ability, a high level of patch connectivity can enhance gene flow, decreasing the extinction probability (Dornier & Cheptou, 2012). Thus, the loss of small patches may threaten the maintenance of connectivity and genetic diversity in the entire population. Therefore, we emphasize the conservation of large and small patches of *G. yakushimensis*. In the present study, we provided information on the genetic diversity and population demography of *G. yakushimensis*, which is useful for conservation prioritization of endangered species (Garner et al., 2020). For an outlook, we require the assessment of population stability by long-term monitoring for the changes in genetic diversity and population size to identify the factors influencing the reduction in genetic diversity and the length of the time lag between the decrease in apparent population size and effective population size, since a drastic reduction in effective population size can occur in the near future.

**Acknowledgments**

The authors are grateful to Yuji Isagi, the project leader of the Environment Research and Technology Development Fund. The authors would like to thank Kenshi Tetsuka, Arata Yoichi, Toshihiro Saito, Tatsuko Tetsuka, and Kengo Fuse for sampling plant materials. We also thank our laboratory colleagues for helping with the analysis and an earlier draft of this manuscript, especially Kenji Seiwa, Yu Fukasawa, Kimiyo Matsukura, Takuya Konuma, and Rimpei Nagaoka. This research was supported by the Environment Research and Technology Development Fund (F-091, 4-1605 and 4-1902) of the Environmental Restoration and Conservation Agency of Japan. Permission to conduct research in the study area was provided by Kagoshima Prefecture, the Japanese Ministry of Environment, the Japanese Ministry of Education, Culture, Sports, Science and Technology, and the Japanese Forestry Agency.

## Conflict of Interest

The authors have no conflicts of interest directly relevant to the content of this article.

## Availability of data

Raw sequence data will be deposited at the DDBJ Sequence Read Archive (DRA) after acceptance.

## Author's Contributions

N.I.I. performed the molecular experiments and analyses and led the writing of the manuscript; Y.T., A.M., and H.A. contributed to the molecular experiments and analyses; S.K.H. and Y.S. contributed to the analyses and writing; and N.I.I., H.A., and Y.S. designed the research.

## References


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

Anma, R., Yamamoto, Y., Geshi, N., Nanayama, F., & Nakagawa, S. (2014). Geology and development of the Yakushima island: a world heritage site. *The Journal of the Geological Society of Japan*, *120*, 101–125. doi: 10.5575/geosoc.2014.0026

Arnholt, A. T., & Evans, B. (2017). *Package "BSDA."* Retrieved from https://cran.r-project.org/web/packages/BSDA/index.html

Banks, S. C., Cary, G. J., Smith, A. L., Davies, I. D., Driscoll, D. A., Gill, A. M., … Peakall, R. (2013). How does ecological disturbance influence genetic diversity? *Trends in Ecology and Evolution*, *28*, 670–679. doi: 10.1016/j.tree.2013.08.005

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

Broadhurst, L., Breed, M., Lowe, A., Bragg, J., Catullo, R., Coates, D., … Byrne, M. (2017). Genetic diversity and structure of the Australian flora. *Diversity and Distributions*, *23*, 41–52. doi: 10.1111/ddi.12505

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., & Postlethwait, J. H. (2011). Stacks: Building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, *1*, 171–182. doi: 10.1534/g3.111.000240

Chen, C.-H., & Wang, J.-C. (1999). Revision of the genus Gentiana L. (Gentianaceae) in Taiwan. *Botanical Bulletin of Academia Sinica*, *40*, 9–38.

Chung, M. Y., López-Pujol, J., Son, S., Suh, G. U., Yukawa, T., & Chung, M. G. (2018). Patterns of genetic diversity in rare and common orchids focusing on the Korean peninsula: Implications for conservation. *Botanical Review*, *84*, 1–25. doi: 10.1007/s12229-017-9190-5

Cornuet, J. M., Pudlo, P., Veyssier, J., Dehne-Garcia, A., Gautier, M., Leblois, R., … 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. doi: 10.1093/bioinformatics/btt763

Dornier, A., & Cheptou, P. O. (2012). Determinants of extinction in fragmented plant populations: Crepis sancta (asteraceae) in urban environments. *Oecologia*, *169*, 703–712. doi: 10.1007/s00442-011-2229-0

Doyle, J. (1991). DNA protocols for plants. *Molecular Techniques in Taxonomy*, *57*, 283–293. doi: 10.1007/978-3-642-83962-7_18

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*, 422–433. doi: 10.1038/nrg.2016.58

Favre, A., Michalak, I., Chen, C. H., Wang, J. C., Pringle, J. S., Matuszak, S., … Muellner-Riehl, A. N. (2016). Out-of-Tibet: The spatio-temporal evolution of Gentiana (Gentianaceae). *Journal of Biogeography*, *43*, 1967–1978. doi: 10.1111/jbi.12840

Feng, X., Liu, J., Chiang, Y. C., & Gong, X. (2017). Investigating the genetic diversity, population differentiation and population dynamics of Cycas segmentifida (cycadaceae) endemic to southwest china by multiple molecular markers. *Frontiers in Plant Science*, *8*, 1–14. doi: 10.3389/fpls.2017.00839

Fuller, M. R., & Doyle, M. W. (2018). Gene flow simulations demonstrate resistance of long-lived species to genetic erosion from habitat fragmentation. *Conservation Genetics*, *19*, 1439–1448. doi: 10.1007/s10592-018-1112-5

Garner, B. A., Hoban, S., & Luikart, G. (2020). IUCN Red List and the value of integrating genetics. *Conservation Genetics*, *21*, 795–801. doi: 10.1007/s10592-020-01301-6

Gibson, J. P., Rice, S. A., & Stucke, C. M. (2008). Comparison of population genetic diversity between a rare, narrowly distributed species and a common, widespread species of Alnus (Betulaceae). *American Journal of Botany*, *95*, 588–596. doi: 10.3732/ajb.2007316

Gonzalez, E. G., Cerón-Souza, I., Mateo, J. A., & Zardoya, R. (2014). Island survivors: Population genetic structure and demography of the critically endangered giant lizard of La Gomera, Gallotia bravoana. *BMC Genetics*, *15*. doi: 10.1186/s12863-014-0121-8

Greenleaf, S. S., Williams, N. M., Winfree, R., & Kremen, C. (2007). Bee foraging ranges and their relationship to body size. *Oecologia*, *153*, 589–596. doi: 10.1007/s00442-007-0752-9

Guisan, A., & Theurillat, J.-P. (2000). Assessing alpine plant vulnerability to climate change: a modeling perspective. *Integrated Assessment*, *1*, 307–320. doi: 10.1023/A:1018912114948

Hamrick, J. L., & Godt, M. J. W. (1996). Effects of life history traits on genetic diversity in plant species. *Philosophical Transactions of the Royal Society in London Series B*, *351*, 1291–1298. doi: 10.1098/rstb.1996.0112

Harsch, M. A., Hulme, P. E., McGlone, M. S., & Duncan, R. P. (2009). Are treelines advancing? A global meta-analysis of treeline response to climate warming. *Ecology Letters*, *12*, 1040–1049. doi: 10.1111/j.1461-0248.2009.01355.x

Ho, T. N. (1988). *Flora reipublicae popularis sinicae*. Beijing: Science Press.

Ho, T. N., & Liu, S. W. (2001). *A worldwide monograph of Gentiana*. Beijing: Science Press.

Hodel, R. G. J., Chen, S., Payton, A. C., McDaniel, S. F., Soltis, P., & Soltis, D. E. (2017). Adding loci improves phylogeographic resolution in red mangroves despite increased missing data: Comparing microsatellites and RAD-Seq and investigating loci filtering. *Scientific Reports*, *7*, 1–14. doi: 10.1038/s41598-017-16810-7

Iwatsuki, K., Yamazaki, T., Boufford, D. E., & Ohba, H. (1993). *Flora of Japan vol. IIIa*. Tokyo: Kodansha Ltd.

Japanese Ministry of the Environment. (2014). *Japanese ministry of the environment red data book*. 272. Retrieved from https://ikilog.biodic.go.jp/rdbdata/files/envpdf/植物I_270.pdf

Japanese Ministry of the Environment. (2017). *Consideration of adaptation strategies to climate change in the World Natural Heritage area of Yakushima Island*. Retrieved from https://www.env.go.jp/park/yakushima/ywhcc/wh/kagaku/17/180217-2-4-2.pdf

Japanese Ministry of the Environment. (2020). *Japanese ministry of the environment red list*. Retrieved from http://www.env.go.jp/press/files/jp/113667.pdf

Kimura, K., Ooi, N., & Suzuki, S. (1996). Evidence of vegetation recovery on Yakushima Island after the major Holocene eruption at the Kikai Caldera, as revealed by the pollen record of buried soils under the old-growth Cryptomeria japonica forest. *Japanese Journal of Historical Botany*, *4*, 13–23. doi: 10.34596/hisbot.4.1_13

Kort, H. De, Prunier, J. G., Ducatez, S., Honnay, O., Baguette, M., Stevens, V. M., & Blanchet, S. (2021). Life history, climate and biogeography interactively affect worldwide genetic diversity of plant and animal populations. *Nature Communications*,

*12*. doi: 10.1038/s41467-021-20958-2

Li, S., Li, B., Cheng, C., Xiong, Z., Liu, Q., Lai, J., … Yan, J. (2014). Genomic signatures of near-extinction and rebirth of the crested ibis and other endangered bird species. *Genome Biology*, *15*, 557. doi: 10.1186/s13059-014-0557-1

Loïc Pellissier, Eidesen, P. B., Ehrich, D., Descombes, P., Schönswetter, P., Tribsch, A., … Alsos, I. G. (2016). Past climate-driven range shifts and population genetic diversity in arctic plants. *Journal of Biogeography*, *43*, 461–470. doi: 10.1111/jbi.12657

Mishiba, K. I., Yamane, K., Nakatsuka, T., Nakano, Y., Yamamura, S., Abe, J., … Nishihara, M. (2009). Genetic relationships in the genus Gentiana based on chloroplast DNA sequence data and nuclear DNA content. *Breeding Science*, *59*, 119–127. doi: 10.1270/jsbbs.59.119

Noroozi, J., Talebi, A., Doostmohammadi, M., Rumpf, S. B., Linder, H. P., & Schneeweiss, G. M. (2018). Hotspots within a global biodiversity hotspot-areas of endemism are associated with high mountain ranges. *Scientific Reports*, *8*, 1–10. doi: 10.1038/s41598-018-28504-9

Nybom, H., & Bartish, I. V. (2000). Effects of life history traits and sampling strategies on genetic diversity estimates obtained with RAPD markers in plants. *Perspectives in Plant Ecology, Evolution and Systematics*, *3*, 93–114. doi: 10.1078/1433-8319-00006

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., … Maintainer, H. W. (2020). Package "vegan." Retrieved from https://cran.ism.ac.jp/web/packages/vegan/vegan.pdf

Oostermeijer, J. G. B., Veer, R. V., & Nijs, J. C. M. Den. (1994). Population structure of the rare, long-lived perennial Gentiana pneumonanthe in relation to vegetation and management in the Netherlands. *The Journal of Applied Ecology*, *31*, 428. doi: 10.2307/2404440

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

Peakall, R., & Smouse, P. E. (2006). GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, *6*, 288–295. doi: 10.1111/j.1471-8286.2005.01155.x

Perrigo, A., Hoorn, C., & Antonelli, A. (2020). Why mountains matter for biodiversity. *Journal of Biogeography*, *47*, 315–325. doi: 10.1111/jbi.13731

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., … Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, *81*, 559–575. doi: 10.1086/519795

Reisch, C., & Rosbakh, S. (2020). Patterns of genetic variation in European plant species depend on altitude. *Diversity and Distributions*, *27*, 157–163. doi: 10.1111/ddi.13177

Reisch, C., Schmidkonz, S., Meier, K., Schöpplein, Q., Meyer, C., Hums, C., … Schmid, C. (2017). Genetic diversity of calcareous grassland plant species depends on historical landscape configuration. *BMC Ecology*, *17*, 1–13. doi: 10.1186/s12898-017-0129-9

Reusch, T. B. H., Ehlers, A., Hammerli, A., & Worm, B. (2005). Ecosystem recovery after climatic extremes enhanced by genotypic diversity. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 2826–2831. doi: 10.1073/pnas.0500008102

Stuiver, M., Grootes, P. M., & Braziunas, T. F. (1995). The GISP2 δ18O climate record of the past 16,500 years and the role of the sun, ocean, and volcanoes. *Quaternary Research*, *44*, 341–354. doi: 10.1006/qres.1995.1079

Suyama, Y., Hirota, S. K., Matsuo, A., Tsunamoto, Y., Mitsuyuki, C., Shimura, A., & Okano, K. (2021). Complementary combination of multiplex high-throughput DNA sequencing for molecular phylogeny. *Ecological Research*, 1–11. doi: 10.1111/1440-1703.12270

Suyama, Y., & Matsuki, Y. (2015). MIG-seq: An effective PCR-based method for genome-wide single-nucleotide polymorphism genotyping using the next-generation sequencing platform. *Scientific Reports*, *5*, 1–12. doi: 10.1038/srep16963

Takatsuki, S. (1990). Summer dietary compositions of sika deer on Yakushima Island, southern Japan. *Ecological Research*, *5*, 253–260. doi: 10.1007/BF02346995

Takatsuki, S. (2009). Effects of sika deer on vegetation in Japan: A review. *Biological Conservation*, *142*, 1922–1929. doi: 10.1016/j.biocon.2009.02.011

Thiel-Egenter, C., Gugerli, F., Alvarez, N., Brodbeck, S., Cieślak, E., Colli, L., … Holderegger, R. (2009). Effects of species traits on the genetic diversity of high-mountain plants: A multi-species study across the Alps and the Carpathians. *Global Ecology and Biogeography*, *18*, 78–87. doi: 10.1111/j.1466-8238.2008.00421.x

Thuiller, W., Lavorel, S., Araújo, M. B., Sykes, M. T., & Prentice, I. C. (2005). Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 8245–8250. doi: 10.1073/pnas.0409902102

Tinner, W., & Kaltenrieder, P. (2005). Rapid responses of high-mountain vegetation to early Holocene environmental changes in the Swiss Alps. *Journal of Ecology*, *93*, 936–947. doi: 10.1111/j.1365-2745.2005.01023.x

Tsujino, R., & Yumoto, T. (2004). Effects of sika deer on tree seedlings in a warm temperate forest on Yakushima Island, Japan. *Ecological Research*, *19*, 291–300. doi: 10.1111/j.1440-1703.2004.00638.x

Tsukada, M. (1983). Vegetation and climate during the last glacial maximum in Japan.

*Quaternary Research*, *19*, 212–235. doi: 10.1016/0033-5894(83)90006-6

Tsukada, M. (1985). Map of vegetation during the last glacial maximum in Japan. *Quaternary Research*, *23*, 369–381. doi: 10.1016/0033-5894(85)90041-9

Willi, Y., Van Buskirk, J., & Hoffmann, A. A. (2006). Limits to the adaptive potential of small populations. *Annual Review of Ecology, Evolution, and Systematics*, *37*, 433–458. doi: 10.1146/annurev.ecolsys.37.091305.110145

Woodgate, J. L., Makinson, J. C., Lim, K. S., Reynolds, A. M., & Chittka, L. (2017). Continuous radar tracking illustrates the development of multi-destination routes of bumblebees. *Scientific Reports*, *7*, 1–15. doi: 10.1038/s41598-017-17553-1

You, J., Qin, X., Ranjitkar, S., Lougheed, S. C., Wang, M., Zhou, W., … Song, Z. (2018). Response to climate change of montane herbaceous plants in the genus Rhodiola predicted by ecological niche modelling. *Scientific Reports*, *8*, 1–12. doi: 10.1038/s41598-018-24360-9

**FIGURE 1**. Photograph of (a) *Gentiana yakushimensis* and (b-c) its habitat. The white arrows in the panels (b) and (c) point to *Gentiana yakushimensis*.

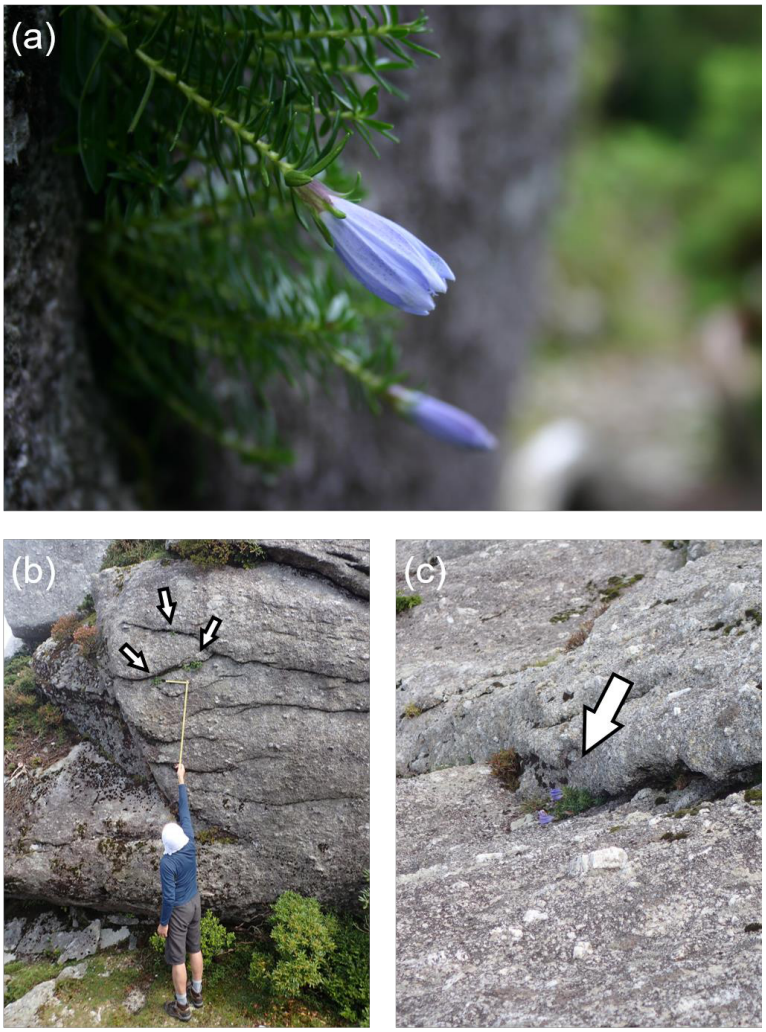

**FIGURE 2**. Map of Yakushima Island and the distribution area of *Gentiana yakushimensis*. The gray part in the right map denotes the area above 1,700 m above sea level. In the left map, the gray part denotes the distribution area of *G. triflora*. The contour interval is 50 m.

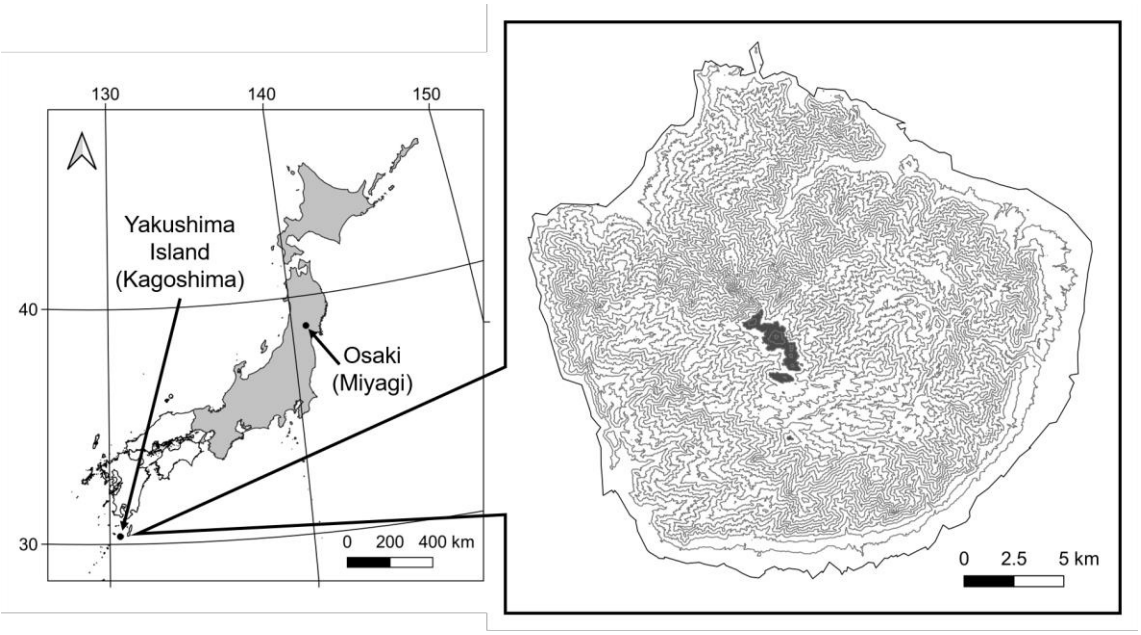

**TABLE 1**. Indices of genetic diversity and inbreeding coefficient parameters of *Gentiana yakushimensis* and *G. triflora*. r, the parameter of the minimum sharing rate among individuals in the 'populations' command of Stacks v2.52; π, nucleotide diversity; $H_O$, observed heterozygosity; $H_E$, expected heterozygosity; $F_{IS}$, inbreeding coefficient. *P*-values were based on pairwise *t*-test. $P_0$ denotes *P*-value of one-sample *t*-test whether $F_{IS}$ was significantly different from zero. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

| Common/ all loci | Species | No. of samples | r | No. of loci | No. of polymorphic loci | No. of Sites | No. of polymorphic sites | π | | | $H_O$ | | | $H_E$ | | | $F_{IS}$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Mean | SE | *P* | Mean | SE | *P* | Mean | SE | *P* | Mean | SE | $P_0$ | *P* |
| Common | *G. yakushimensis* | 353 | 0.5 | 89 | 47 | 7,237 | 72 | 0.00020 | 0.00008 | *** | 0.013 | 0.005 | *** | 0.020 | 0.008 | *** | 0.222 | 0.045 | *** | |
| Common | *G. triflora* | 9 | | 89 | 47 | 7,237 | 72 | 0.00135 | 0.00030 | | 0.121 | 0.030 | | 0.125 | 0.023 | | 0.048 | 0.076 | | |
| Common | *G. yakushimensis* | 353 | 0.6 | 69 | 35 | 5,623 | 51 | 0.00009 | 0.00002 | ** | 0.006 | 0.002 | ** | 0.009 | 0.002 | *** | 0.263 | 0.054 | *** | * |
| Common | *G. triflora* | 9 | | 69 | 35 | 5,623 | 51 | 0.00103 | 0.00028 | | 0.100 | 0.030 | | 0.105 | 0.025 | | 0.058 | 0.075 | | |
| Common | *G. yakushimensis* | 353 | 0.7 | 51 | 25 | 4,150 | 34 | 0.00005 | 0.00002 | ** | 0.004 | 0.001 | * | 0.006 | 0.002 | ** | 0.274 | 0.069 | *** | * |
| Common | *G. triflora* | 9 | | 51 | 25 | 4,150 | 34 | 0.00095 | 0.00031 | | 0.113 | 0.041 | | 0.108 | 0.031 | | -0.004 | 0.098 | | |
| All | *G. yakushimensis* | 353 | 0.5 | 2,124 | 487 | 172,094 | 784 | 0.00055 | 0.00003 | *** | 0.084 | 0.004 | *** | 0.121 | 0.005 | *** | 0.341 | 0.014 | *** | *** |
| All | *G. triflora* | 9 | | 1,475 | 302 | 116,730 | 401 | 0.00162 | 0.00008 | | 0.329 | 0.009 | | 0.436 | 0.002 | | 0.230 | 0.022 | *** | |
| All | *G. yakushimensis* | 353 | 0.6 | 1,758 | 387 | 142,510 | 562 | 0.00041 | 0.00003 | *** | 0.070 | 0.005 | *** | 0.103 | 0.005 | *** | 0.358 | 0.017 | *** | *** |
| All | *G. triflora* | 9 | | 1,181 | 222 | 93,622 | 297 | 0.00147 | 0.00009 | | 0.311 | 0.009 | | 0.430 | 0.003 | | 0.263 | 0.023 | *** | |
| All | *G. yakushimensis* | 353 | 0.7 | 1,429 | 292 | 115,877 | 402 | 0.00033 | 0.00003 | *** | 0.058 | 0.005 | *** | 0.095 | 0.006 | *** | 0.387 | 0.019 | *** | *** |
| All | *G. triflora* | 9 | | 945 | 155 | 74,977 | 202 | 0.00125 | 0.00009 | | 0.315 | 0.011 | | 0.435 | 0.004 | | 0.264 | 0.027 | *** | |

**FIGURE 3**. Relationship between pairwise genetic and geographic distance. (a) Relationship between individual-based pairwise genetic and geographic distance of *Gentiana yakushimensis* when the parameter r is 0.5. Mantel's *r* and its significance are denoted. (b) Mantel correlograms for the association between individual-based genetic and geographic distance. Filled symbols in the figure denote statistical significance.

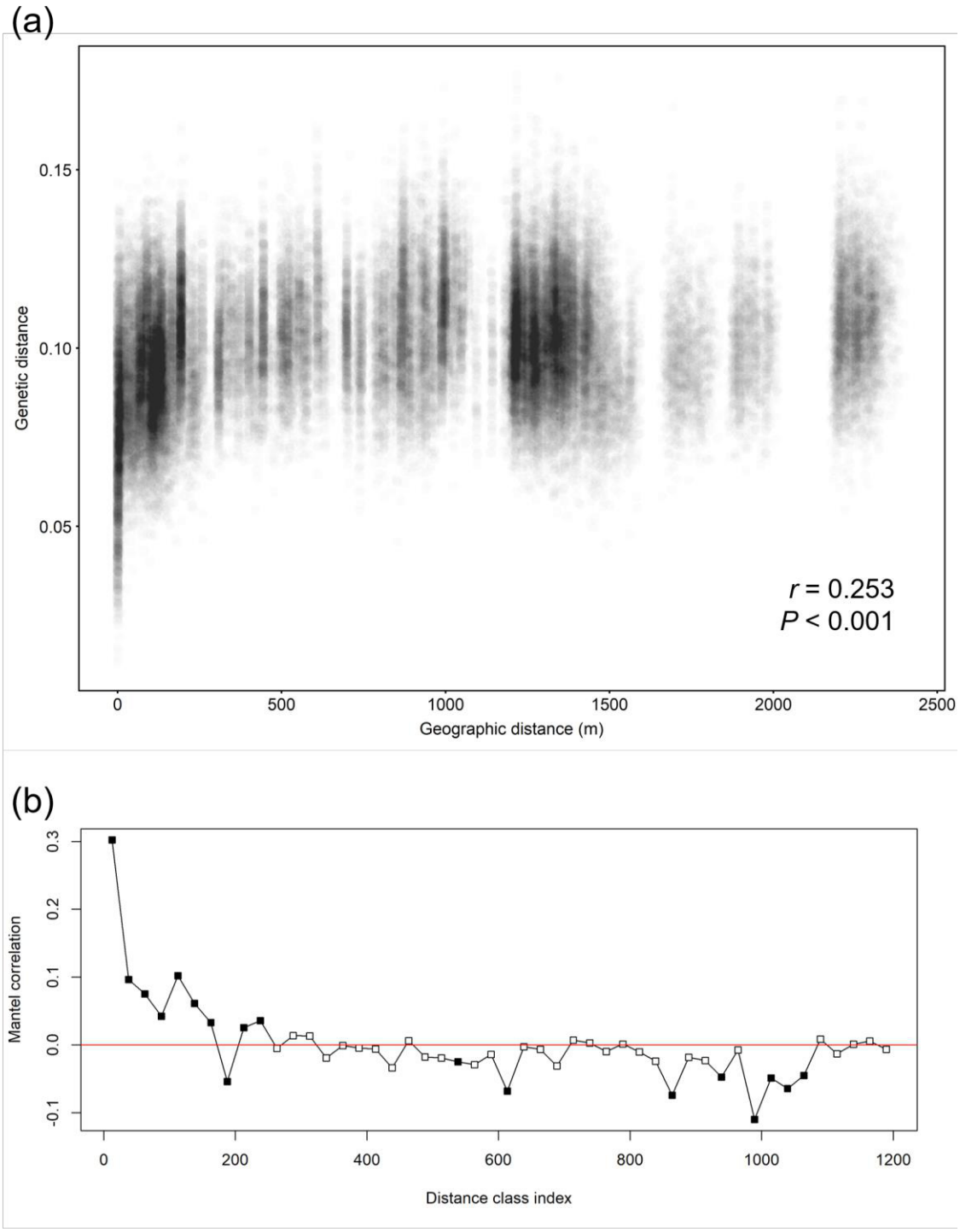

**FIGURE 4**. Principal coordinate analysis (PCoA) score plots based on individual-based genetic distance of *Gentiana yakushimensis* when the parameter r is 0.5. (a), axis 1-2; (b), axis 1-3. Colors denote the location (Lon: longitude) of each individual.

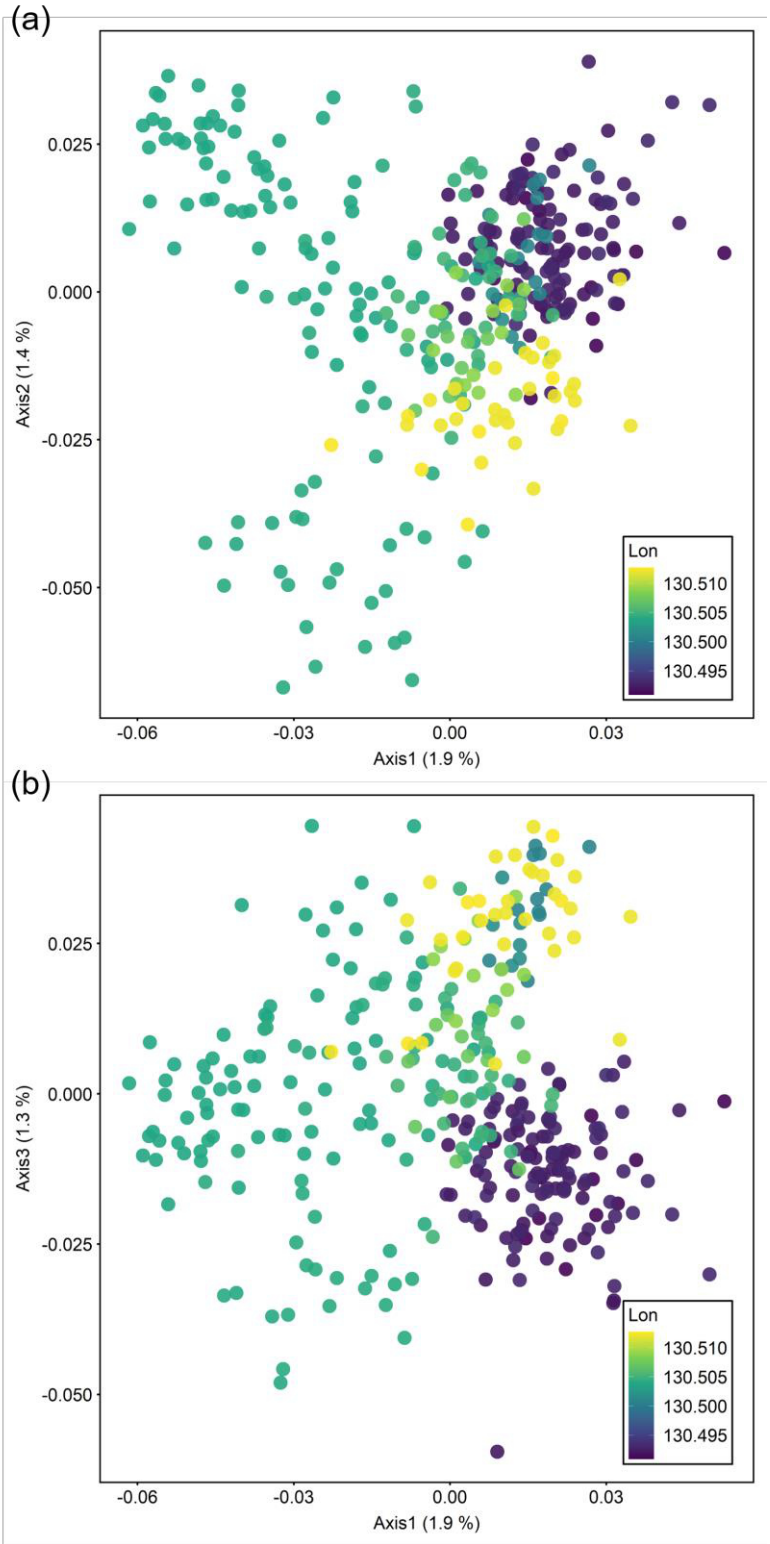

**FIGURE 5**. Parameter estimates of demographic scenarios when the parameter r is 0.5. In this figure, the selection probability of the scenario 3 and 4 are shown. The parameter N0, N1, and N2 (effective population size) of each scenario selected by the ABC analysis are denoted.

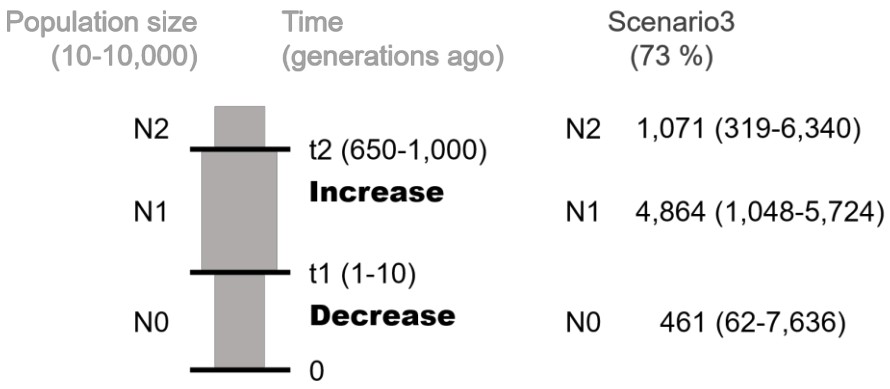

**FIGURE S1**. Parameters in the demographic inference by DIYABC Random Forest. In demographic scenarios, t# represents time scale measured in number of generations and N# represents effective population size of the corresponding populations during the relevant time period (0–t1, t1–t2, and t2–). We performed 100,000 simulations per scenario with setting N#, t1 (corresponding to the recent 100 years), and t2 (corresponding to LGM) to 10-10000 individuals, and 1-10 and 1000-2000 generations, respectively.

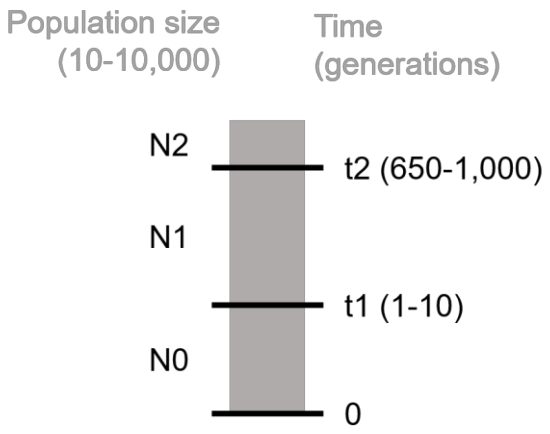

**Table S1**. Setting of each demographic scenario. Four scenarios were set up with the change of the parameter N# (effective population size) at t1 and t2 in figure S1 as increase or decrease. The scenario 5 (mentioned in results) is defined considering the result of scenario choice.

| Scenario | N2-N1 | N1-N0 |
|---|---|---|
| 1 | Decrease | Decrease |
| 2 | Decrease | Increase |
| 3 | Increase | Decrease |
| 4 | Increase | Increase |

**FIGURE S2**. Relationship between pairwise genetic and geographic distance. (a) Relationship between individual-based pairwise genetic and geographic distance of *Gentiana yakushimensis* when the parameter r is 0.6. Mantel's *r* and its significance are denoted. (b) Mantel correlograms for the association between individual-based genetic and geographic distance. Filled symbols in the figure denote statistical significance.

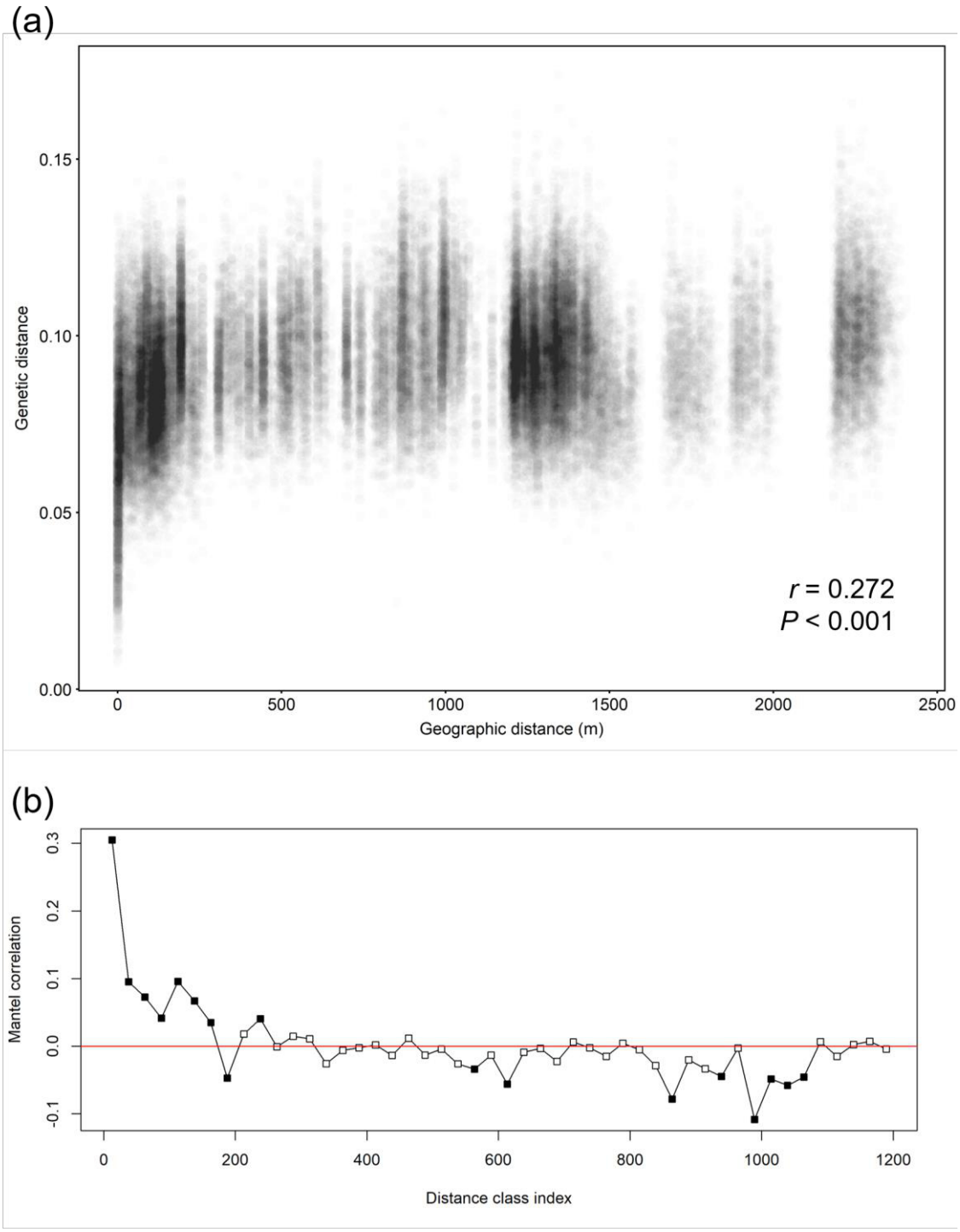

**FIGURE S3**. Relationship between pairwise genetic and geographic distance. (a) Relationship between individual-based pairwise genetic and geographic distance of *Gentiana yakushimensis* when the parameter r is 0.7. Mantel's *r* and its significance are denoted. (b) Mantel correlograms for the association between individual-based genetic and geographic distance. Filled symbols in the figure denote statistical significance.

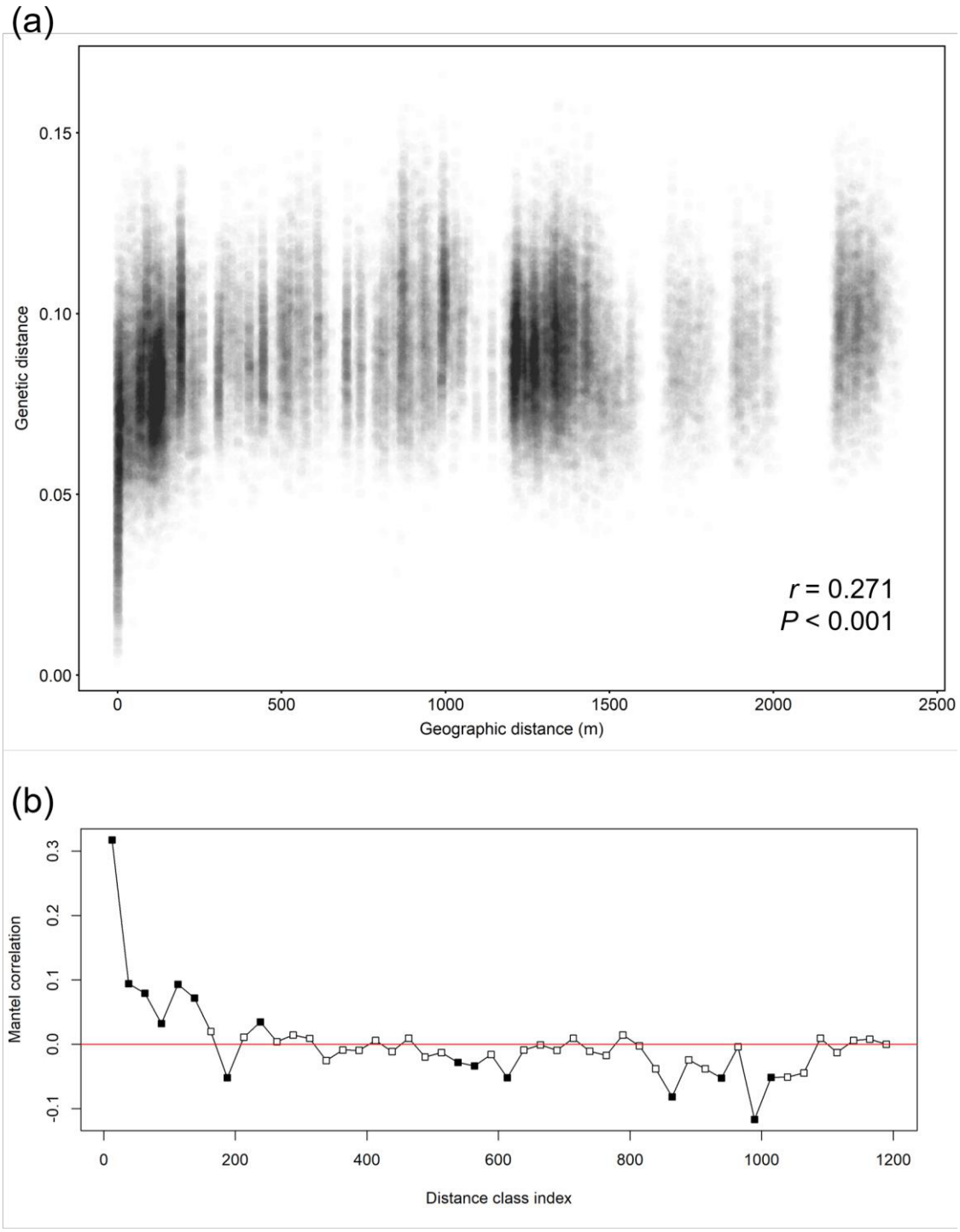

**FIGURE S4**. Principal coordinate analysis (PCoA) score plots based on individual-based genetic distance of *Gentiana yakushimensis* when the parameter r is 0.6. (a), axis 1-2; (b), axis 1-3. Colors denote the location (Lon: longitude) of each individual.

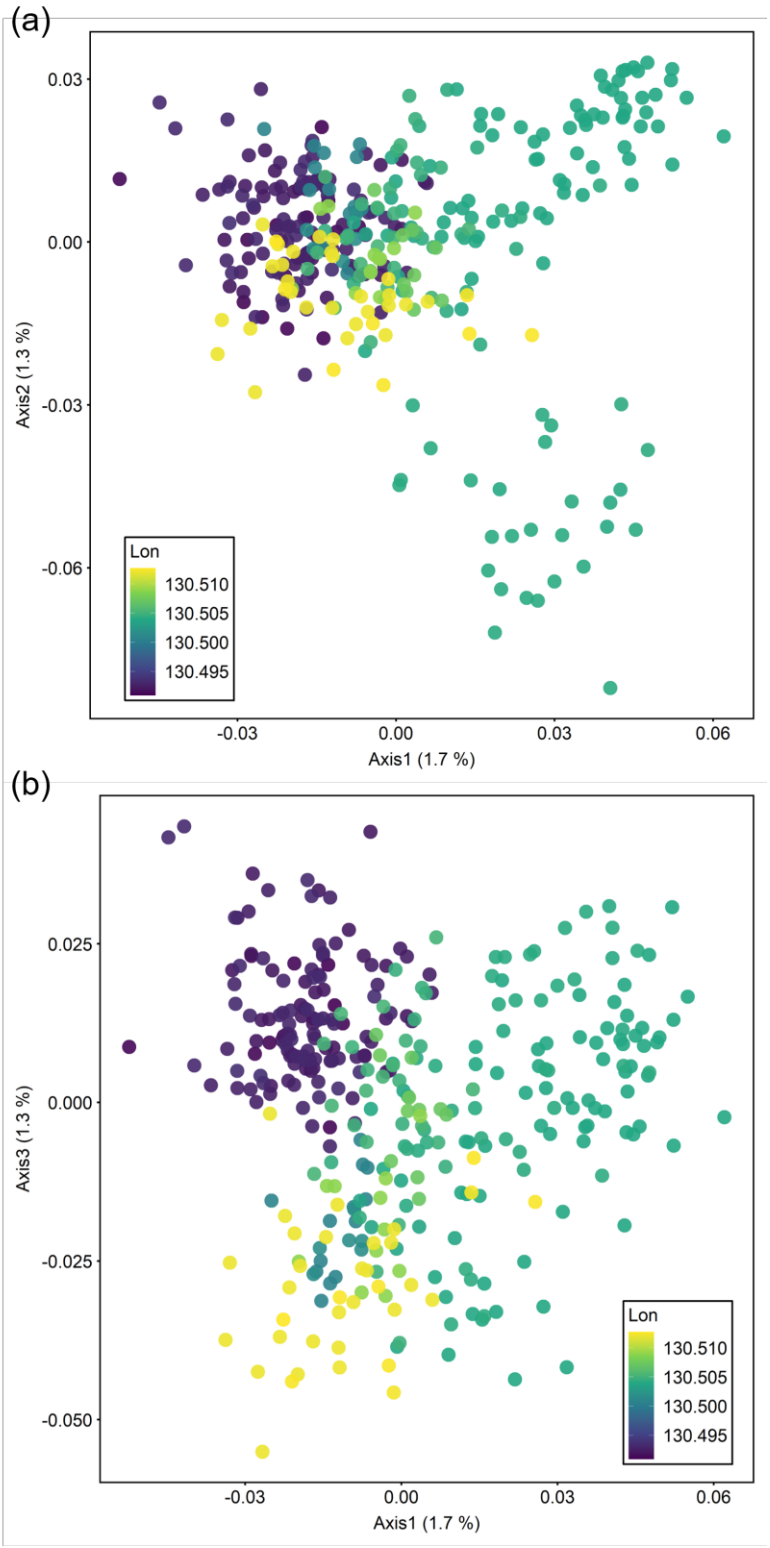

**FIGURE S5**. Principal coordinate analysis (PCoA) score plots based on individual-based genetic distance of *Gentiana yakushimensis* when the parameter r is 0.7. (a), axis 1-2; (b), axis 1-3. Colors denote the location (Lon: longitude) of each individual.

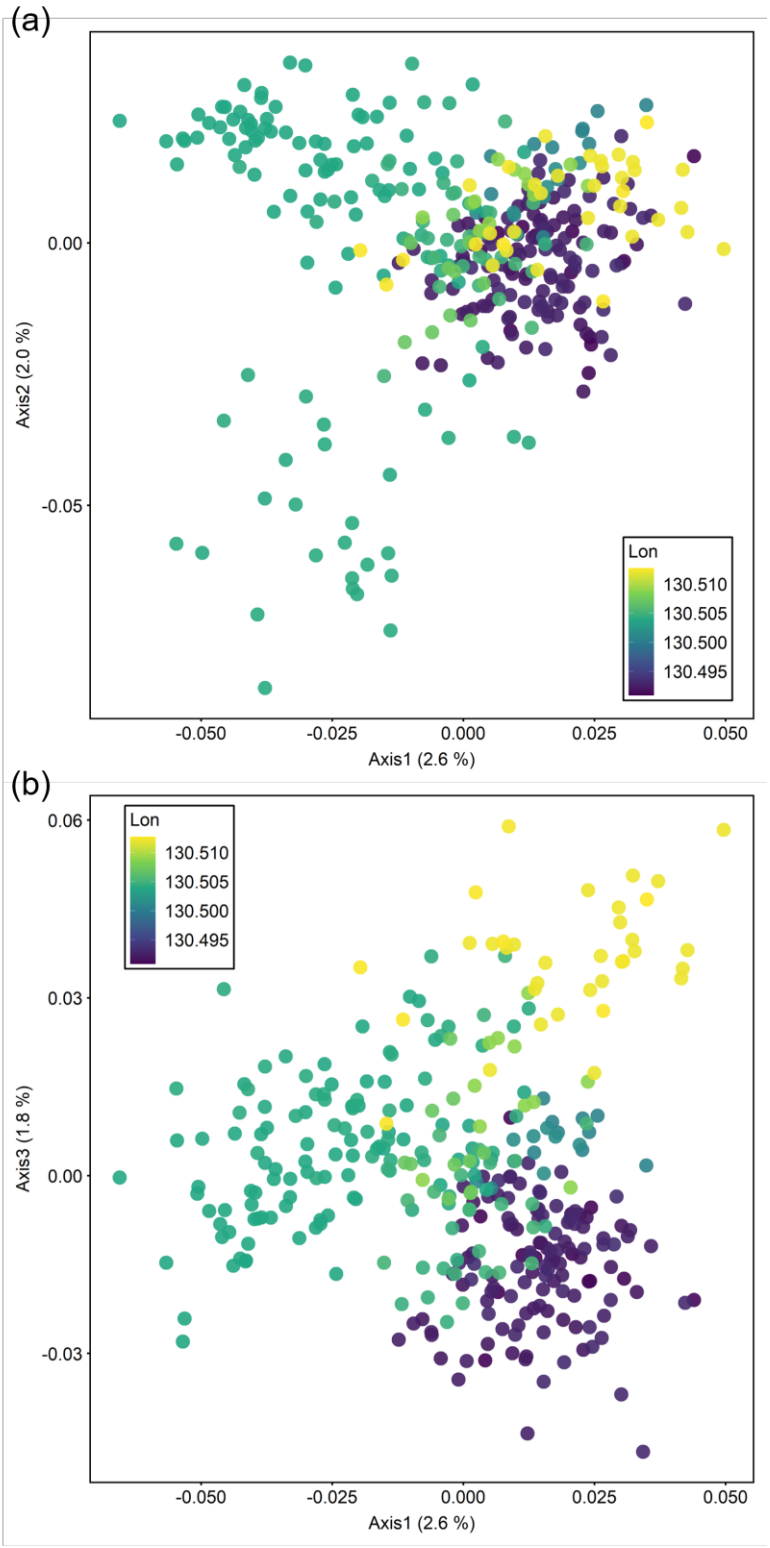

**Table S2**. Scenario choice of demographic inference by ABC analysis. The probability that each scenario was selected by 1000 regression trees in the random forest analysis.

| r | Scenario1 | Scenario2 | Scenario3 | Scenario4 |
|---|---|---|---|---|
| 0.5 | 0.004 | 0.001 | 0.730 | 0.265 |
| 0.6 | 0.000 | 0.000 | 0.701 | 0.299 |
| 0.7 | 0.000 | 0.000 | 0.763 | 0.237 |

**Figure S6**. Results of linear discriminant analysis (LDA) in demographic scenario choice. The panel (a)-(c) correspond the result with the setting of the parameter r is 0.5, 0.6, and 0.7, respectively. Simulated data based on scenarios that is consistent with the observed data is plotted closer to the observed data in these figures.

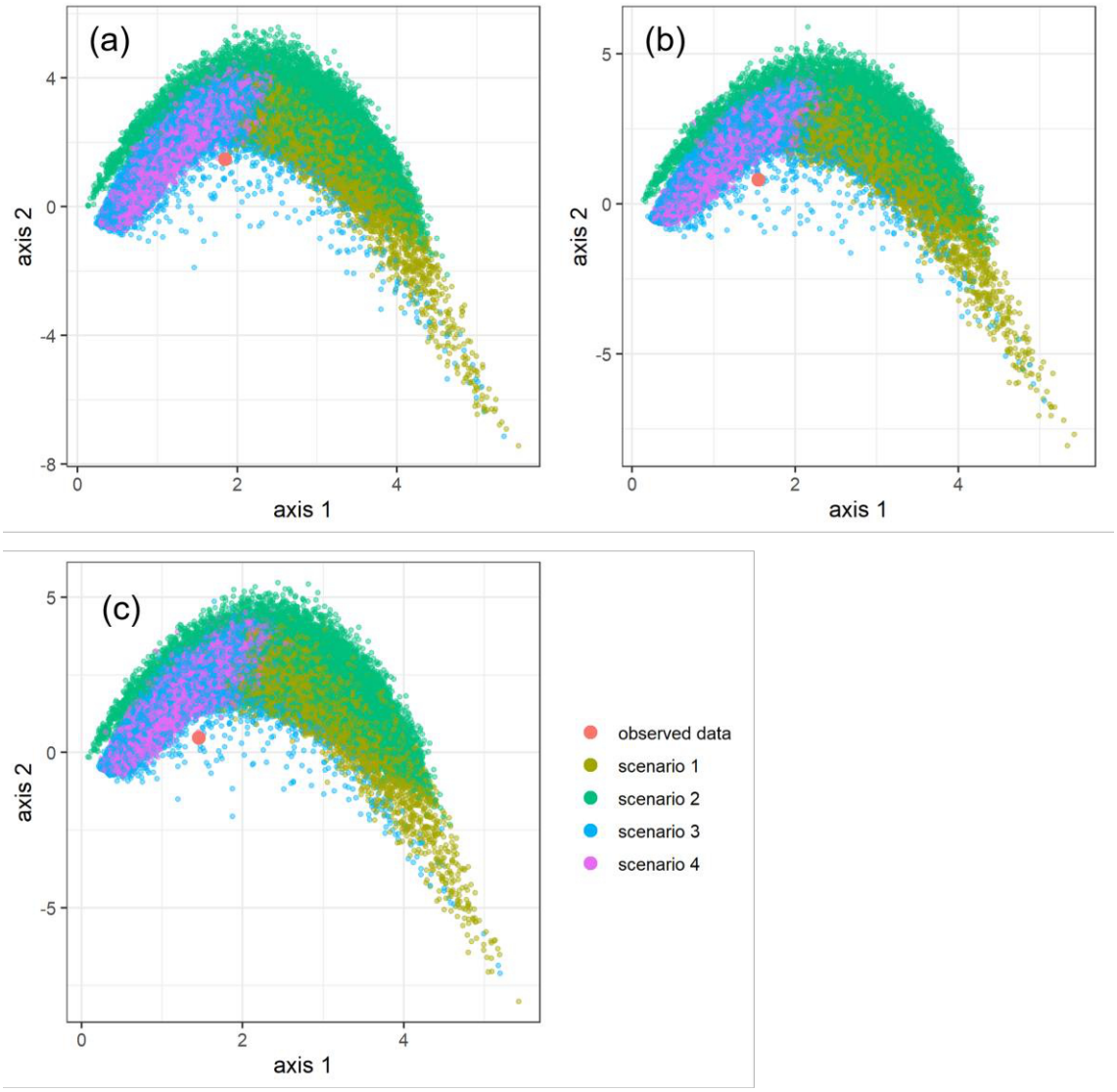

**Table S3**. Parameter estimates of demographic scenarios. The parameter r is the minimum sharing rate of SNP among individuals in the 'populations' command in Stacks.

| r | Scenario | N0 | | | N1 | | | N2 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Mean | Median | 95% CI | Mean | Median | 95% CI | Mean | Median | 95% CI |
| 0.5 | 3 | 1662.2 | 461.0 | 62.0-7636.0 | 5241.8 | 4864.0 | 1047.74-9688.0 | 2019.8 | 1071.0 | 319.1-6340.0 |
| 0.5 | 4 | 5805.2 | 5900.1 | 1597.7-9661.0 | 2789.3 | 1937.0 | 734.2-7082.0 | 2434.9 | 1817.0 | 399.0-5724.3 |
| 0.6 | 3 | 1129.2 | 453.0 | 64.3-4694.0 | 5147.6 | 4850.8 | 1554.3-9791.0 | 1051.6 | 835.0 | 180.0-2913.0 |
| 0.6 | 4 | 4407.4 | 3680.5 | 1191.7-9208.0 | 2006.8 | 1767.0 | 825.9-4378.0 | 1001.4 | 819.0 | 95.0-2659.0 |
| 0.7 | 3 | 1713.1 | 848.7 | 88.0-6308.0 | 4906.1 | 4558.0 | 1173.2-9433.0 | 1151.8 | 782.8 | 169.0-3422.1 |
| 0.7 | 4 | 5312.9 | 4992.7 | 2025.4-9495.0 | 2608.1 | 1889.0 | 816.1-7164.3 | 1254.7 | 947.0 | 175.0-3481.0 |

| r | Scenario | t1 | | | t2 | | |
|---|---|---|---|---|---|---|---|
| | | Mean | Median | 95% CI | Mean | Median | 95% CI |
| 0.5 | 3 | 5.8 | 6.0 | 1.0-10.0 | 824.7 | 824.8 | 668.0-981.0 |
| 0.5 | 4 | 4.8 | 5.0 | 1.0-9.0 | 851.7 | 869.0 | 669.0-999.0 |
| 0.6 | 3 | 5.2 | 5.0 | 1.0-9.0 | 836.9 | 853.0 | 672.0-990.0 |
| 0.6 | 4 | 5.8 | 6.0 | 1.0-10.0 | 829.4 | 826.0 | 665.0-986.3 |
| 0.7 | 3 | 5.6 | 5.0 | 1.0-10.0 | 839.6 | 838.0 | 671.0-993.0 |
| 0.7 | 4 | 4.9 | 4.0 | 1.0-9.0 | 853.3 | 863.0 | 684.0-987.0 |