Genes Genet. Syst. (2021) **96**, p. 1–6

Development and characterization of nuclear microsatellite markers to reveal the neutral demographic background of flower color polymorphism in *Geranium thunbergii* (Geraniaceae)

Seikan Kurata[1*], Shota Sakaguchi[2], Hitomi Mishima[1], Takashi Tsuchimatsu[3] and Motomi Ito[1]

[1]*Department of General Systems Studies, Graduate School of Arts and Sciences, University of Tokyo, Meguro, Tokyo 153-8902, Japan*
[2]*Department Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Kyoto 606-8501, Japan*
[3]*Department of Biological Sciences, Graduate School of Science, University of Tokyo, Bunkyo, Tokyo 113-0033, Japan*

(Received 11 November 2020, accepted 15 January 2021; J-STAGE Advance published date: 22 April 2021)

Nuclear microsatellite markers were developed for *Geranium thunbergii*, an herbaceous plant characterized by petal color polymorphism. Utilizing RNA sequencing data obtained by next-generation sequencing techniques, we developed and characterized 19 polymorphic microsatellite markers with two to 12 alleles in the nuclear genome. These markers will be used to reveal the genetic structure and demographic history of *G. thunbergii* in the Japanese archipelago, which will elucidate the genetic background of flower color polymorphism among populations.

Key words: flower color polymorphism, genetic diversity, *Geranium thunbergii*, microsatellite marker, population genetics

Plants are often characterized by petal color variation within and among populations; this phenomenon is defined as flower color polymorphism. Flower color polymorphism has long been of interest to ecologists and evolutionists (Darwin, 1862; Forsman, 2016). For species with intraspecific floral color variation [e.g., *Antirrhinum majus* L. (Plantaginaceae), *Aquilegia coerulea* E. James (Ranunculaceae), *Dactylorhiza sambucina* (L.) Soó (Orchidaceae), *Lysimachia arvensis* (L.) U. Manns & Anderb. (Primulaceae) and *Raphanus sativus* L. (Brassicaceae)], the mechanisms underlying the maintenance of floral diversity have been rigorously investigated (Gigord et al., 2001; Jones and Reithel, 2001; Arista et al., 2013; McCall et al., 2013; Thairu and Brunet, 2015). Based on these investigations, biotic and abiotic elements (e.g., negative frequency-dependent selection, pollinator preference and temperature) are suggested to be important determinants of flower color polymorphism (Narbona et al., 2018). Selectively neutral factors, i.e., genetic drift and gene flow, have also been implicated in the maintenance of flower color polymorphism (Wright, 1978; Narbona et al., 2018). Therefore, it is important to evaluate the relative contributions of different evolutionary forces in maintaining floral color polymorphism within and among populations.

Geranium thunbergii Siebold ex Lindl. & Paxton (Geraniaceae) is a perennial herb distributed throughout the Japanese archipelago. The habitats of this species are lowland forest edges and grassy areas. The species shows a geographic cline in floral color: individuals having purple petals (i.e., purple flowers) are distributed on the western side of the Japanese archipelago, while those having white or pale pink petals (i.e., considered as "white flowers") are distributed on the eastern side (Akiyama, 2001; Kadota, 2016). Both flower colors (i.e., purple and white) co-occur in central Honshu, the geographically intermediate distribution area of the species. As such, *G. thunbergii* has floral color variation both within and

Edited by Yoshihiko Tsumura
* Corresponding author. seikan.kurata@gmail.com
DOI: https://doi.org/10.1266/ggs.20-00062

among populations, and is therefore a suitable plant species for exploring the formation and maintenance mechanisms of flower color polymorphism. In particular, since the geographic distribution pattern of floral color may be related to past neutral demographic history, phylogeographic and population genetics approaches are necessary to reveal the geographic pattern of flower color polymorphism. Tsuchimatsu et al. (2014) indicated that selective pressure by herbivores on white flowers and the anthocyanin-mediated herbivore defense of purple flowers are associated with the flower color polymorphism of *G. thunbergii*, but a more detailed study (e.g., including

Table 1. Characteristics of 19 microsatellite markers developed for *G. thunbergii*

Locus	Primer sequence (5'–3')	Size range (bp)	Motif	BLASTX top hit description [species]	*E*-value	ONEKP ID
VKGP_202	F: CAGCAAGCACCATGTTACCC R: AGTGAAGCTCAAGAGAAGCG	150–153	$(AG)_9$	UDP-galactose transporter 2 isoform X1 [*Carex littledalei*]	2.00E-63	VKGP-2000202
VKGP_2694	F: TCTCCTCCTTCCATTGCCTG R: CGACCTTCACACAGCAATCC	318–339	$(AG)_{10}$	kinesin-like protein KIN-4A isoform X1 [*Ricinus communis*]	0	VKGP-2002694
VKGP_3319	F: CAGAAACAGAGACCATAGCGTC R: TGTCTGCGAAGAGAGTACGG	167–169	$(AG)_9$	hypothetical protein EZV62_019291 [*Acer yangbiense*]	5.00E-109	VKGP-2003319
VKGP_4952	F: AACCAGACCCTTGTAGCTCC R: CTTTGGAGCTCATTTGAACGTG	326–330	$(AC)_9$	No hit		VKGP-2004952
VKGP_10356	F: TCCTTTCTTCCTTGGTTTCCTG R: TCTCCATGCACAACTCCTCC	183–197	$(AG)_9$	protease Do-like 7 [*Ziziphus jujuba*]	0	VKGP-2010356
VKGP_14936	F: TAGACCCAATTCAGCCTCGG R: CTCACCAGTTTCCGATTCGC	269–281	$(AAC)_{10}$	F-box/LRR-repeat protein 4 [*Syzygium oleosum*]	3.00E-20	VKGP-2014936
VKGP_19219	F: ATGCGAAGGTGGAGAAGACG R: TCTCTCGGCCTGATACAGTG	257–266	$(AG)_8$	DUF502 domain-containing protein [*Cephalotus follicularis*]	5.00E-11	VKGP-2019219
VKGP_25965	F: GTTAAGAATGCGGGCGGTAG R: AGCAAAGCGAATGTCTCACG	334–337	$(ATC)_8$	hypothetical protein FH972_002063 [*Carpinus fangiana*]	8.00E-112	VKGP-2025965
VKGP_26672[*1]	F: TGTTTCTGTTCCGTTGACCC R: AAGCTCCCATCTCCGATTCC	102–124	$(AG)_9$	probable glutathione peroxidase 4 [*Rhodamnia argentea*]	9.00E-95	VKGP-2026672
VKGP_29328	F: GCATTCCTACACAGCATCGG R: ATCCCAGAGGTGCAGACAAG	211–214	$(ATC)_{17}$	hypothetical protein CISIN_1g018444mg [*Citrus sinensis*]	4.00E-87	VKGP-2029328
VKGP_31098	F: GCAGATTGGAATGTTGGTGC R: TTGCAAAGCCATCACCCATG	370–376	$(AGC)_9$	No hit		VKGP-2031098
VKGP_31943	F: GCATTACGTACACTGGCTGG R: GGATCCGACCTCCCAAATCC	263–266	$(AAC)_8$	ataxin-2 homolog [*Malus domestica*]	7.00E-06	VKGP-2031943
VKGP_32221	F: GAGTGAGCAGAGTCTCGAGG R: AGACGGAGACAGAGCTTCTC	399–427	$(AG)_9$	two-component response regulator ARR12-like [*Juglans regia*]	0	VKGP-2032221
VKGP_87603[*1]	F: CCGACAGAGAAGCTACGAAC R: TCGTGACTCAGTGACCTTCC	123–134	$(AG)_9$	No hit		VKGP-2087603
VKGP_92431[*1, 2]	F: AAGCAGAGAGGTCGATCGAG R: AGTGTGTGAGAGACTGTACGG	131–141	$(AG)_9$	No hit		VKGP-2092431
VKGP_108374[*1]	F: CAGACGCGGACAAAGCTAAG R: TGAACAGCGGGTAAAGAGAG	168–172	$(AG)_9$	No hit		VKGP-2108374
VKGP_108676[*1]	F: GAGCAGGAGAGAGAAGCAATC R: AGCAGTTCGTGTACATTGCG	129–141	$(AG)_9$	hypothetical protein [*Gossypium harknessii*]	7.00E-44	VKGP-2108676
YGCX_17221	F: AGAGGGACCAAACCACTGTC R: AGGTCAGTGCATGTAGAGGC	160–181	$(AAC)_9$	hypothetical protein GH714_022102 [*Hevea brasiliensis*]	3.00E-49	YGCK-2017221
YGCX_28878	F: ACACTCCTTCCCATGATCCG R: TCTTCTACGCCAACCACCTC	395–419	$(ATC)_9$	uncharacterized protein LOC109022334 [*Juglans regia*]	6.00E-09	YGCX-2028878

[*1] These primers were previously reported as monomorphic microsatellite primers of *G. soboliferum* var. *kiusianum* (Kurata et al., 2017).

[*2] Primer pair VKGP_92431 amplifies two independent loci, and we confirmed that the two loci are in linkage disequilibrium. However, only one locus was adopted in this study; the other locus often showed a null allele, which could have been derived from PCR amplification error.

migration history) is required to elucidate the geographic pattern of flower color polymorphism in *G. thunbergii*. In this study, we developed polymorphic microsatellite markers for *G. thunbergii* to reveal its population genetics and phylogeographic history.

Assembled RNA sequencing data of *G. carolinianum* L. and *G. maculatum* Dum. Cours. were obtained from the ONEKP: BLAST for 1,000 Plants repository (https://db.cngb.org/onekp/). A similarity search of the contigs against the National Center for Biotechnology Information non-redundant protein database was conducted using the BLASTX algorithm (Altschul et al., 1990) with an *E*-value cutoff of 1.0E-5. We screened the sequences with microsatellite regions for ≥ 8 dinucleotide repeats and ≥ 8 trinucleotide repeats using MSATCOMMANDER (Faircloth, 2008), and designed primers using Primer3 software (Rozen and Skaletsky, 2000). A total of 369 primer pairs bordering microsatellites were designed and 141 pairs were selected for PCR amplification trials, using eight individuals from eight populations (Kodaira, Hokkaido Prefecture (Pref.); Saku, Nagano Pref.; Hachioji, Tokyo Pref.; Katsuura, Chiba Pref.; Taki, Mie Pref.; Okayama, Okayama Pref.; Amakusa, Kumamoto Pref.; and Yakushima, Kagoshima Pref.). For all loci, the forward primer was synthesized with one of three different tag sequences (5′-CACGACGTTGTAAAACGAC-3′, 5′-TGTGGAATTGTGAGCGG-3′ or 5′-CTATAGGGCACGCGTGGT-3′) and the reverse primer was tagged with a pig-tail sequence (5′-GTTTCTT-3′) (Brownstein et al., 1996). Genomic DNA was extracted from dried leaves using a modified CTAB method (Milligan, 1992). PCR amplification was carried out following the standard protocol of the Qiagen Multiplex PCR Kit (Qiagen, Hilden, Germany) in a final reaction volume of 10 µl containing approximately 5 ng of DNA, 5 µl of 2× Multiplex PCR Master Mix, 0.01 µM forward primer, 0.2 µM reverse primer, and 0.1 µM M13 primer (fluorescently labeled with Beckman Dye; Beckman Coulter, Brea, CA, USA). The PCR thermal profile involved denaturation at 95 °C for 3 min, followed by 35 cycles of 95 °C for 30 s, 54 °C for 3 min, and 68 °C for 1 min, and a final 20-min extension step at 68 °C. PCR products were loaded onto an auto sequencer (GenomeLab GeXP; Beckman Coulter) to assess fragment lengths using Fragment Analysis Software ver. 8.0 (Beckman Coulter).

The extracted DNA of 23 individuals from Hachioji

Table 2. Genetic diversity of the 19 polymorphic markers for *G. thunbergii*

Locus	Hachioji (n = 23) [white]				Mifune (n = 27) [purple]				Nantan (n = 30) [purple/white]				Total			
	A	Ar	H_e	H_o	A	Ar	H_e	H_o	A	Ar	H_e	H_o	A	Ar	H_e	H_o
VKGP_202	2	1.948	0.091	0.000***	2	1.839	0.071	0.000***	2	2.000	0.391	0.000***	3	2.357	0.184	0.000
VKGP_2694	4	3.454	0.279	0.136*	6	4.455	0.239	0.111***	5	4.024	0.453	0.444	8	5.584	0.324	0.231
VKGP_3319	1	1.000	0.000	0.000	1	1.000	0.000	0.000	1	1.000	0.000	0.000	2	2.000	0.000	0.000
VKGP_4952	2	1.727	0.044	0.046	2	1.839	0.071	0.074	2	1.533	0.033	0.033	2	1.600	0.050	0.051
VKGP_10356	2	2.000	0.500	1.000***	2	2.000	0.458	0.708**	3	2.552	0.355	0.310	4	3.169	0.438	0.673
VKGP_14936	1	1.000	0.000	0.000	2	1.857	0.074	0.000***	2	2.000	0.500	1.000***	3	2.373	0.191	0.333
VKGP_19219	2	1.948	0.091	0.000***	4	3.329	0.180	0.192	4	3.523	0.222	0.103***	6	3.511	0.164	0.099
VKGP_25965	1	1.000	0.000	0.000	1	1.000	0.000	0.000	2	2.000	0.444	0.133***	2	2.000	0.148	0.044
VKGP_26672	2	1.930	0.087	0.000***	3	2.185	0.072	0.074	3	2.862	0.213	0.233	5	3.647	0.124	0.103
VKGP_29328	1	1.000	0.000	0.000	1	1.000	0.000	0.000	2	1.786	0.064	0.067	2	2.000	0.022	0.022
VKGP_31098	2	1.696	0.043	0.044	2	1.839	0.071	0.074	2	1.533	0.033	0.033	4	2.722	0.049	0.050
VKGP_31943	1	1.000	0.000	0.000	1	1.000	0.000	0.000	2	2.000	0.346	0.000***	2	2.000	0.115	0.000
VKGP_32221	4	3.588	0.210	0.045***	5	3.703	0.181	0.115**	7	5.376	0.484	0.400	12	5.391	0.292	0.187
VKGP_87603	1	1.000	0.000	0.000	1	1.000	0.000	0.000	4	3.983	0.634	0.333*	5	4.743	0.211	0.111
VKGP_92431	3	2.608	0.124	0.130	3	2.531	0.139	0.074*	3	2.615	0.292	0.269	5	3.266	0.185	0.158
VKGP_108374	2	1.727	0.044	0.046	4	2.846	0.111	0.115	2	1.533	0.033	0.033	4	1.911	0.063	0.065
VKGP_108676	3	2.391	0.084	0.087	2	1.938	0.105	0.111	4	3.531	0.242	0.233	5	3.550	0.144	0.144
YGCX_17221	1	1.000	0.000	0.000	1	1.000	0.000	0.000	4	3.916	0.684	0.276***	6	5.461	0.228	0.092
YGCX_28878	4	3.390	0.268	0.217	4	2.778	0.107	0.111	6	4.318	0.215	0.231	9	3.498	0.197	0.186
Average	2.1	1.864	0.098	0.092	2.5	2.060	0.099	0.093	3.2	2.741	0.297	0.217	4.7	3.199	0.165	0.134

A, number of alleles per locus; Ar, allelic richness; H_e, expected heterozygosity; H_o, observed heterozygosity; n, number of individuals genotyped. Asterisks denote significant departure from Hardy–Weinberg equilibrium ($^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$). Petal color is indicated with population name.

(Tokyo Pref.), 27 from Mifune (Kumamoto Pref.) and 30 from Nantan (Kyoto Pref.) was used to evaluate allelic polymorphisms. To characterize each microsatellite marker (= 19 markers), four summary statistics were calculated using FSTAT v.2.9.3 (Goudet, 1995) and GenAlEx v.6.501 (Peakall and Smouse, 2012): the number of alleles (A), allelic richness (Ar), expected heterozygosity (H_e) and observed heterozygosity (H_o). These summary statistics were calculated for each locus and population. The significance of Hardy–Weinberg equilibrium and genotypic equilibrium were tested by chi-squared test using GenAlEx v.6.501. In addition, the F_{ST} index (Weir and Cockerham,

Table 3. A further 20 primer pairs that were monomorphic or fixed to a heterozygote genotype for *G. thunbergii*

Locus	Primer sequence (5'–3')	Predicted size (bp)	Motif	BLASTX top hit description [species]	*E*-value	ONEKP ID
VKGP_4891	F: CTCCAAGGTACGAGGTGGTC R: GATGTCACGACGTTCACAGC	345	$(AAC)_8$	transcription repressor MYB4-like [*Rhodamnia argentea*]	6.00E-89	VKGP-2004891
VKGP_10465	F: CAGCATCGTTCTTTCCCACC R: ACGTGCTGTTACAAACTGGG	345	$(AAG)_{13}$	LOW QUALITY PROTEIN: probable serine/threonine protein kinase IRE4 [*Pistacia vera*]	3.00E-68	VKGP-2010465
VKGP_12870	F: GGTTGGTTCTGTTTCTGGGC R: CAACCAGCTCACACCTCAAC	276	$(AG)_{11}$	hypothetical protein CMV_011704 [*Castanea mollissima*]	0	VKGP-2012870
VKGP_22151	F: TCATTGTGGCGAGCAAGTTC R: TTGCCCGGGTTCTCTTATCC	172	$(AG)_9$	No hits		VKGP-2022151
VKGP_22086	F: CCAAATCACTGACCTCCACG R: TCACAATCTCGTTCTCATCACC	206	$(AT)_9$	NAC domain-containing protein 43-like isoform X2 [*Hevea brasiliensis*]	3.00E-21	VKGP-2022086
VKGP_23492	F: CAACTTGATCATGCACTTGTGC R: AGTCAGTGCTGGACAAGGAG	289	$(AAC)_8$	hypothetical protein G4B88_004180 [*Cannabis sativa*]	6.00E-129	VKGP-2023492
VKGP_23939	F: TTCGTCTGATTCGGCATTGC R: TCGGCCATGGAAGGTAGAAG	396, 417	$(AAG)_8$	mechanosensitive ion channel protein 10-like [*Herrania umbratica*]	0	VKGP-2023939
VKGP_25427	F: AAATAGAGGGAACAAGGCGC R: AGAGTATACGCCTCCATCGC	188	$(AG)_9$	AMP deaminase/myoadenylate deaminase, putative isoform 1 [*Theobroma cacao*]	0	VKGP-2025427
VKGP_29076	F: ATCAGCCACCTCATCACCTC R: TGGGCATGACGATATCCTGG	230	$(AAC)_{11}$	No hits		VKGP-2029076
VKGP_31603	F: GAGAACACAATCCTCGTCGC R: AGCTCTCCTCCACTTCTTGC	333	$(AG)_9$	hypothetical protein [*Gossypium lobatum*]	2.00E-27	VKGP-2031603
VKGP_76520	F: TCTCACCGACCTTTCCCATC R: CAGTCTCTAGTTGCTCATCAGG	132	$(AC)_8$	No hits		VKGP-2076520
VKGP_78116*	F: GAGAGGCTTGCGATGGAGAG R: AAAGCTCCACTCAACAACGC	118	$(AG)_9$	No hits		VKGP-2078116
YGCX_1175	F: GATTCTGCTTCTCGTGACCC R: GAAGCTCACTGTCTCGTTGC	176	$(AAG)_9$	aldo-keto reductase family 4 member c9-like protein [*Trifolium pratense*]	5.00E-06	YGCX-2001175
YGCX_3345	F: TCCTCCTGTATCGCCGAAAG R: CCCGAATCCATTTGAGGTGC	225	$(AGC)_8$	unnamed protein product [*Prunus armeniaca*]	2.00E-51	YGCX-2003345
YGCX_6368	F: CCCTTCCAACAAGTGCATGG R: AGCTTCTGTGAGGGAGGAAC	313	$(AAC)_8$	probable transcription factor PosF21 [*Cicer arietinum*]	2.00E-08	YGCX-2006368
YGCX_22751	F: TCCTCTGAGCTATGGTGTCAC R: ATCCCTCTCACAATCTGGCC	270	$(AAG)_9$	senescence/dehydration-associated protein At4g35985, chloroplastic-like isoform X1 [*Rosa chinensis*]	1.00E-114	YGCX-2022751
YGCX_22772	F: CTGATGAACTTGGACGACGC R: ATGTGGAGAGGATCATGGCC	392	$(AAC)_9$	AP2-like ethylene-responsive transcription factor ANT [*Herrania umbratica*]	6.00E-36	YGCX-2022772
YGCX_23325	F: TTGAGCCGGAACAGAGTCAG R: CGAGAATGTCACCGAACTGC	265	$(ACC)_8$	RNA-binding protein 38 isoform X3 [*Prosopis alba*]	2.00E-44	YGCX-2023325
YGCX_25909	F: GCCACTACAACTGGACTTGC R: ATCTGCCCTATGAGCTCCAG	404	$(ATC)_9$	No hits		YGCX-2025909
YGCX_28044	F: ACCATCAATTTGCGGGACAC R: GCACCAACATCATCCCTCTC	208	$(AAG)_{14}$	protein ENHANCED DISEASE RESISTANCE 2-like [*Carica papaya*]	4.00E-22	YGCX-2028044

*This primer pair was previously reported as a monomorphic microsatellite primer pair of *G. soboliferum* var. *kiusianum* (Kurata et al., 2017).

1984) between populations was calculated using FSTAT v.2.9.3 for elucidating genetic differentiation among the three populations. Cross-amplification trials of the 19 markers were also performed for the related species *G. wilfordii* Maxim. (16 individuals from two populations: Onneyu, Hokkaido Pref. and Oshino, Yamanashi Pref.) and *G. sibiricum* L. (16 individuals from two populations: Nakatonbetsu and Onneyu, Hokkaido Pref.).

For the first primer screening using the auto sequencer, 39 of 141 primer pairs successfully amplified DNA fragments of the predicted size, while the remaining 102 pairs amplified fragments of unpredicted size, produced multiple bands, or failed to amplify any fragment. For the 39 reliable primer pairs that showed clear microsatellite peaks of the predicted fragment size, we conducted a second PCR trial using 80 individuals from three populations. We found that 19 loci were polymorphic across the three populations (Table 1), ranging from two to 12 alleles with H_e and H_o values ranging from 0.0 to 0.684 and 0.0 to 1.000, respectively (Table 2). Among these 19 loci, five markers (Table 1, Table 2) were found to be identical to sequences that were previously reported as monomorphic microsatellite primers of *G. soboliferum* var. *kiusianum* (Kurata et al., 2017). Genetic diversity was highest in the Nantan population (Ar = 2.741, H_e = 0.297), followed by the populations of Mifune (Ar = 2.060, H_e = 0.099) and Hachioji (Ar = 1.864, H_e = 0.098) (Table 2). In the Nantan population, individuals of both flower colors are distributed, while only purple and white flowers are distributed in Mifune and Hachioji, respectively. We confirmed significant departures from Hardy–Weinberg equilibrium at some loci in particular populations. Specifically, we detected significant deviations at VKGP_202, VKGP_2694, VKGP_10356, VKGP_14936, VKGP_19219 and VKGP_32221 in two or three populations in each locus, which may indicate the presence of null alleles at these loci. Significant genetic differentiation among the three *G. thunbergii* populations was detected using these markers (i.e., Hachioji–Mifune, F_{ST} = 0.841; Hachioji–Nantan, F_{ST} = 0.574; Mifune–Nantan, F_{ST} = 0.647). The genotyping error rate of the 19 markers was 2.18% based on 24 individuals arbitrarily selected from three populations. Note that the remaining 20 primers were inappropriate for performing population genetic studies because all individuals were fixed to an allele (i.e., monomorphic) or a heterozygote genotype (Table 3).

Table 4. Cross-amplification results of the 19 markers for *G. wilfordii* and *G. sibiricum*

Locus	*G. wilfordii*								*G. sibiricum*							
	Onneyu (n = 8)				Oshino (n = 8)				Nakatonbetsu (n = 8)				Onneyu (n = 8)			
	A	Ar	H_e	H_o	A	Ar	H_e	H_o	A	Ar	H_e	H_o	A	Ar	H_e	H_o
VKGP_202	2	1.992	0.219	0.000	1	1.000	0.000	0.000	1	1.000	0.000	0.000	1	1.000	0.000	0.000
VKGP_2694	3	2.867	0.320	0.250	3	2.875	0.398	0.125	2	1.875	0.117	0.125	1	1.000	0.000	0.000
VKGP_3319	3	3.000	0.586	0.125	2	2.000	0.375	0.250	1	1.000	0.000	0.000	2	2.000	0.245	0.000
VKGP_4952	2	1.875	0.117	0.125	3	2.750	0.227	0.250	2	1.875	0.117	0.125	1	1.000	0.000	0.000
VKGP_10356	2	1.875	0.117	0.125	3	2.750	0.227	0.250	1	1.000	0.000	0.000	3	2.875	0.398	0.250
VKGP_14936	2	1.875	0.117	0.125	1	1.000	0.000	0.000	2	2.000	0.500	1.000	2	2.000	0.500	1.000
VKGP_19219	3	3.000	0.255	0.286	2	1.875	0.117	0.125	2	2.000	0.245	0.000	2	2.000	0.133	0.143
VKGP_25965	1	1.000	0.000	0.000	1	1.000	0.000	0.000	1	1.000	0.000	0.000	–	–	–	–
VKGP_26672	2	1.992	0.219	0.000	2	1.875	0.117	0.125	1	1.000	0.000	0.000	2	1.875	0.117	0.125
VKGP_29328	2	1.875	0.117	0.125	3	3.000	0.633	0.250	1	1.000	0.000	0.000	1	1.000	0.000	0.000
VKGP_31098	3	2.867	0.320	0.375	4	3.625	0.328	0.375	3	2.750	0.227	0.250	3	2.750	0.227	0.250
VKGP_31943	2	2.000	0.430	0.125	2	2.000	0.305	0.125	1	1.000	0.000	0.000	1	1.000	0.000	0.000
VKGP_32221	2	2.000	0.500	1.000	2	2.000	0.500	1.000	1	1.000	0.000	0.000	2	1.875	0.117	0.125
VKGP_87603	2	2.000	0.305	0.125	2	2.000	0.492	0.125	1	1.000	0.000	0.000	1	1.000	0.000	0.000
VKGP_92431	3	2.867	0.320	0.125	1	1.000	0.000	0.000	2	1.992	0.219	0.000	1	1.000	0.000	0.000
VKGP_108374	2	1.992	0.219	0.250	2	1.875	0.117	0.125	2	1.875	0.117	0.125	2	1.875	0.117	0.125
VKGP_108676	1	1.000	0.000	0.000	4	3.858	0.492	0.375	2	1.992	0.219	0.000	3	2.867	0.320	0.375
YGCX_17221	2	1.875	0.117	0.125	1	1.000	0.000	0.000	1	1.000	0.000	0.000	1	1.000	0.000	0.000
YGCX_28878	2	1.992	0.219	0.000	2	1.875	0.117	0.125	2	1.875	0.117	0.125	2	1.875	0.117	0.125
Average	2.2	2.102	0.237	0.173	2.2	2.071	0.234	0.191	1.5	1.486	0.099	0.092	1.7	1.666	0.127	0.140

A, number of alleles per locus; Ar, allelic richness; H_e, expected heterozygosity; H_o, observed heterozygosity; n, number of individuals genotyped. One marker, VKGP_25965, failed to PCR-amplify any fragments for *G. sibiricum*.

Moreover, we checked for cross-amplification of these polymorphic markers in *G. wilfordii* and *G. sibiricum*. In *G. wilfordii*, although one locus (VKGP_25965) was monomorphic, the other 18 loci were polymorphic across the two populations, ranging from two to four alleles with H_e and H_o values ranging from 0.0 to 0.633 and 0.0 to 1.000, respectively (Table 4). Allelic richness (*Ar*) ranged from 1.000 to 3.858 (Table 4). In *G. sibiricum*, 13 markers showed polymorphisms across the two populations, with the number of alleles ranging from two to three and H_e and H_o values ranging from 0.0 to 0.500 and 0.0 to 1.000, respectively (Table 4). Allelic richness (*Ar*) ranged from 1.000 to 2.875 (Table 4). However, six loci (i.e., VKGP_202, VKGP_25965, VKGP_29328, VKGP_31943, VKGP_87603 and YGCX_17221) were monomorphic, and one marker, VKGP_25965, failed to PCR-amplify any fragments for *G. sibiricum*.

Overall, the microsatellite markers developed here will be useful to reveal the genetic structure and demographic history of *G. thunbergii* in the Japanese archipelago, which will elucidate the genetic background of flower color polymorphism among populations.

The authors thank Daiki Takahashi, Kazutoshi Masuda and Koki Nagasawa for their great help with the sampling. The authors are grateful to Drs. Atsushi Ohwaki and Kenji Horie for granting access to their collection of materials (*G. wilfordii* and *G. sibiricum*). We are grateful to the Ashiu Forest Research Station (Kyoto University) for granting us permission to perform field surveys. We are also grateful to the ONEKP: BLAST for 1,000 Plants for kindly allowing us to access the sequence data. This work was supported by JSPS KAKENHI under Grant No. JP16H04827; the Agency for Medical Research and Development National BioResource Project under Grant No. 18km0210136j0002; and Fujiwara Natural History Foundation.

REFERENCES

Akiyama, S. (2001) Geraniaceae. *In* Flora of Japan, Vol. IIb. (eds.: Iwatsuki, K., Boufford, D. E., and Ohba, H.), p. 289. Kodansha Ltd., Tokyo, Japan.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. J. Mol. Biol. **215**, 403–410.

Arista, M., Talavera, M., Berjano, R., and Ortiz, P. L. (2013) Abiotic factors may explain the geographical distribution of flower colour morphs and the maintenance of colour polymorphism in the scarlet pimpernel. J. Ecol. **101**, 1613–1622.

Brownstein, M. J., Carpten, J. D., and Smith, J. R. (1996) Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. Biotechniques **20**, 1004–1010.

Darwin, C. (1862) On the Various Contrivances by Which British and Foreign Orchids are Fertilised by Insects, and on the Good Effects of Intercrossing. John Murray, London, UK.

Faircloth, B. C. (2008) MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. Mol. Ecol. Resour. **8**, 92–94.

Forsman, A. (2016) Is colour polymorphism advantageous to populations and species? Mol. Ecol. **25**, 2693–2698.

Gigord, L. D. B., Macnair, M. R., and Smithson, A. (2001) Negative frequency-dependent selection maintains a dramatic flower color polymorphism in the rewardless orchid *Dactylorhiza sambucina* (L.) Soó. Proc. Natl. Acad. Sci. USA **98**, 6253–6255.

Goudet, J. (1995) FSTAT (Version 1.2): A computer program to calculate *F*-statistics. J. Hered. **86**, 485–486.

Jones, K. N., and Reithel, J. S. (2001) Pollinator-mediated selection on a flower color polymorphism in experimental populations of *Antirrhinum* (Scrophulariaceae). Am. J. Bot. **88**, 447–454.

Kadota, Y. (2016) Geraniaceae. *In* Wild flowers of Japan, Vol. 3. (eds.: Ohashi, H., Kadota, Y., Murata, J., Yonekura, K., and Kihara, H.), pp. 252–253. Heibonsha, Tokyo, Japan (in Japanese).

Kurata, S., Sakaguchi, S., and Ito, M. (2017) Development of nuclear microsatellite markers for the threatened wetland plant *Geranium soboliferum* var. *kiusianum* (Geraniaceae). Plant Species Biol. **32**, 466–470.

McCall, A. C., Murphy, S. J., Venner, C., and Brown, M. (2013) Florivores prefer white versus pink petal color morphs in wild radish, *Raphanus sativus*. Oecologia **172**, 189–195.

Milligan, B. (1992) Plant DNA isolation. *In* Molecular Genetic Analysis of Populations: A Practical Approach, 2nd ed. (ed.: Hoelzel, A. R.), pp. 59–88. IRL Press, Oxford, UK.

Narbona, E., Wang, H., Ortiz, P. L., Arista, M., and Imbert, E. (2018) Flower colour polymorphism in the Mediterranean Basin: occurrence, maintenance and implications for speciation. Plant Biol. **20**, 8–20.

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

Rozen, S., and Skaletsky, H. (2000) Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. **132**, 365–386.

Thairu, M. W., and Brunet, J. (2015) The role of pollinators in maintaining variation in flower colour in the Rocky Mountain columbine, *Aquilegia coerulea*. Ann. Bot. **115**, 971–979.

Tsuchimatsu, T., Yoshitake, H., and Ito, M. (2014) Herbivore pressure by weevils associated with flower color polymorphism in *Geranium thunbergii* (Geraniaceae). J. Plant Res. **127**, 265–273.

Weir, B. S., and Cockerham, C. C. (1984) Estimating *F*-statistics for the analysis of population structure. Evolution **38**, 1358–1370.

Wright, S. (1978) Evolution and the Genetics of Populations, Vol. 4. Variability Within and Among Natural Populations. University of Chicago Press, Chicago, USA.