Short communication

Title

Development and characterization of nuclear microsatellite markers in *Aphananthe aspera* (Thunb.) Planch. (Cannabaceae)

Authors

Taishi Sakagami[1], Shota Sakaguchi[2,4], Yuji Isagi[3], Hiroaki Setoguchi[2, 4]

Affiliations

[1]Faculty of Integrated Human Studies, Kyoto University, Kyoto, Japan; [2]Graduate School of Human and Environmental Studies, Kyoto University, Kyoto, Japan; [3]Division of Forest and Biomaterials Science, Graduate School of Agriculture, Kyoto University, Kyoto, Japan

Email addresses: TS: skgt0404@gmail.com, SS: sakaguchi.shota.6a@kyoto-u.ac.jp, YI: isagiy@kais.kyoto-u.ac.jp, HS: setoguchi.hiroaki.2c@kyoto-u.ac.jp

[4]Author for correspondence: setoguchi.hiroaki.2c@kyoto-u.ac.jp, sakaguci54@gmail.com

19

20

21 **Running head:** *Aphananthe* SSR markers

22 **Keywords**: *Aphananthe aspera*, Cannabaceae, genetic diversity, microsatellite markers

Abstract

Nuclear microsatellite markers were developed for *Aphananthe aspera* (Thunb.) Planch. (Cannabaceae), a deciduous canopy tree species distributed in East Asia, to evaluate the genetic diversity and genetic structure of *A. aspera* populations in remnant forest fragments in urbanized areas of Japan. A total of 94 primer pairs were designed based on genomic sequence data. Of the 25 primer pairs which showed clear microsatellite peaks, 20 pairs showed allelic polymorphisms in 57 individuals collected from two distant populations. The length of PCR products ranged from 120 to 482 bp, and expected heterozygosity for the 20 microsatellite markers ranged from 0.017 to 0.768. These newly developed SSR markers will be used in population genetic studies of *A. aspera* to evaluate genetic diversity and the extent of genetic isolation of the fragmented populations in urban areas.

Introduction

Habitat fragmentation can have major genetic and demographic consequences on populations (Honnay et al. 1999; Oostermeijer et al. 2003; Fahrig 2017). It is predicted that reduction in effective population size and the decreased population connectivity can lead to negative effects such as greater inbreeding (Keller and Waller 2002), restricted gene flow (Browne and Karubian 2018), and reduced immigration rates (Couvet 2002; Dubreuil et al. 2010). While the effects of habitat fragmentation on tree species may be buffered by their longevity and greater capabilities of gene dispersals, empirical studies showed the consequences vary by species and context (Lowe et al. 2015), requiring further examinations in different study systems.

Aphananthe aspera (Thunb.) (Cannabaceae) is a temperate broad-leaved deciduous tree species widely distributed in Japan (from subtropical Okinawa Island to temperate Kanto region), extending to the Korean peninsula, China, and Taiwan. It is a long-lived (in some cases up to > 500 years) and wind-pollinated canopy tree. Ripe berries of *A. aspera* are black purple in color in autumn, and mainly dispersed by birds (Yoshikawa and Kikuzawa 2009). It sometimes becomes dominant in warm-temperate forests in alluvial lowlands, but currently forms small, scattered populations due to human activities especially in urbanized area in Japan (Kimura et al. 2019). Although

the species is ecologically important in terms of dominance and interaction with avian and animal species in warm-temperate forests, there have been no studies on the species' genetic diversity and genetic differentiation of the fragmented *A. aspera* populations. Microsatellite markers (or simple sequence repeat, SSR) are highly polymorphic, codominant, and useful to evaluate population genetic characteristics (Jones and Ardren 2003). To investigate the negative effects on gene diversity and genetic structure by fragmentation, microsatellite markers that can be applied to *A. aspera* are needed. In this study, we developed nuclear microsatellite markers from *A. aspera* using genomic sequence data and evaluated their polymorphisms.

Materials and Methods

We assembled the genomic sequence reads of *A. aspera* (GenBank accession numbers of raw sequence reads; Bioproject: PRJDB5250, Submission: DRA005224, Experiment: DRX069505 and Sample: DRS055811) (Sakaguchi et al. 2017), using CLC Genomics Workbench 7.5.1 (CLC bio, Aarhus, Denmark) with the parameter settings of word size = 19, bubble size = 161, and minimum contig length = 100. Then the contigs including microsatellite regions, for ≥ 6 tri-nucleotide repeats, ≥ 8 di-nucleotide repeats and ≥ 15 mono-nucleotide repeats, were searched with

MSATCOMMANDER (Faircloth 2008). In total, 4,220 microsatellite motifs were found, and 94 of them were selected to design primer pairs using Primer 3 (Rozen and Skaletsky 2000). For all the primer pairs, the forward primers were synthesized with one of four different universal sequences (5'-CACGACGTTGTAAAACGAC3', 5'-TGTGGAATTGTGAGCGG-3', CGGAGAGCCGAGAGGTG-3', or 5'-CTATAGGGCACGCGTGGT-3' (Blacket et al. 2012)) and the reverse primers were tagged with PIG-tail sequences (5'-GTTTCTT-3') (Brownstein et al. 1996). BLASTN search for each contig was performed with a threshold of E-05 for significant hit. 57 individuals of *A. aspera* from each of the two populations [29 individuals from Shimogamo Shrine, Kyoto, Japan (35°02'N/135°46'E), 28 individuals from Kawaketa-mikabe Shrine, Shiga, Japan (35°06'N/136°13'E)] were used to evaluate the polymorphisms of the target microsatellite loci. All sampled leaves were dried in silica gel, polysaccharides removed with HEPES buffer (pH 8.0) (Setoguchi and Ohba 1995), and then genomic DNA was isolated using the CTAB method (Doyle and Doyle 1987). The PCR reaction was performed in a 10μL reaction volume containing 10–20 ng of DNA, 5 μL of Multiplex PCR Master Mix (Qiagen, Hilden, Germany), 0.01 μM of forward primer, 0.2 μM of reverse primer and 0.1 μM of fluorescently labeled M13 primer. PCR amplification for all primer pairs started with 3 min at 95°C for initial

denaturation, 35 cycles of denaturation at 95°C for 30 s, primer annealing at 60°C for 3 min, and extension at 68°C for 1 min, and a final extension for 20 min at 68°C. PCR product was loaded onto an auto sequencer (3130 Genetic Analyzer; Applied Biosystems, Carlsbad, California, USA) to assess fragment lengths using GeneMapper software (Applied Biosystems). To evaluate the utility of developed genetic markers, genetic diversity indices (number of alleles, observed heterozygosity and expected heterozygosity) were calculated by using GenAlEx ver. 6.5.2 (Peakall and Smouse 2012). Frequency of null alleles were estimated for each combination of locus and population using CERVUS 3.0 (Marshall et al. 1998). Significance of Hardy-Weinberg equilibrium and linkage disequilibrium between loci were tested by GenAlEx ver. 6.5.2.

Results and discussion

Of 94 primer pairs tested, 25 pairs showed clear microsatellite peaks (Table 1) and 20 pairs displayed allelic polymorphisms in the two populations of *A. aspera* (Table 2). The corresponding sequences of these regions were deposited in DDBJ Center: accession numbers LC360729- LC360744 (Table 1). In the populations of *A. aspera*, the number of alleles per locus ranged from 1 to 8, the observed heterozygosity

ranged from 0.000 to 0.724 and the expected heterozygosity ranged from 0.000 to 0.783 (Table 2). One locus (Aa_105260) showed consistently significant deviations from Hardy-Weinberg equilibrium in both populations ($P < 0.001$), and relatively high frequencies of null alleles were estimated (0.563 for Shimogamo Shrine and 0.673 for Kawaketa-mikabe Shrine, respectively). Hence, care must be taken when applying the marker (Aa_105260) in any future genetic research. No locus pairs showed significant linkage disequilibrium ($P > 0.05$). In conclusion, the novel 20 nuclear microsatellite markers were polymorphic and reliably scorable in *A. aspera*. These markers will be useful in future studies investigating the population genetics, and evolutionary history and genetic interactions within populations of the species.

Acknowledgements

We are grateful to S. Mori, R. Nakamasu for collecting plant materials, and to two anonymous reviewers and coordinating editor for their useful comments on this manuscript Research permissions were provided by Shimogamo Shrine and Kawaketa-mikabe Shrine. This research was supported by the Japan Society for the Promotion of Science Grant-in-Aid for Scientific Research (JSPS KAKENHI 16H04827, 16H04831 and 25291085) and the Environmental Research and Technology Development Fund of the Ministry of the Environment (4-1702, 4-1902),

126 and SICORP Program of the Japan Science and Technology Agency (grant no.
127 4-1403).

128

References

Blacket M, Robin C, Good R, Lee S, Miller A (2012) Universal primers for fluorescent labelling of PCR fragments—an efficient and cost-effective approach to genotyping by fluorescence. Molecular ecology resources 12:456-463

Browne L, Karubian J (2018) Habitat loss and fragmentation reduce effective gene flow by disrupting seed dispersal in a neotropical palm. Mol Ecol 27:3055-3069

Brownstein MJ, Carpten JD, Smith JR (1996) Modulation of non-templated nucleotide addition by taq DNA polymerase: Primer modifications that facilitate genotyping. Biotechniques 20:1004-+

Couvet D (2002) Deleterious effects of restricted gene flow in fragmented populations. Conservation Biology 16:369-376

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem. Bull Bot Soc Am 19

Dubreuil M, Riba M, González-Martínez SC, Vendramin GG, Sebastiani F, Mayol M (2010) Genetic effects of chronic habitat fragmentation revisited: Strong genetic structure in a temperate tree, Taxus baccata (Taxaceae), with great dispersal capability. American Journal of Botany 97:303-310

147 Fahrig L (2017) Ecological Responses to Habitat Fragmentation Per Se. Annual
148 Review of Ecology, Evolution, and Systematics 48:1-23
149 Faircloth BC (2008) MSATCOMMANDER: detection of microsatellite repeat arrays
150 and automated, locus-specific primer design. Molecular Ecology Resources
151 8:92-94
152 Honnay O, Hermy M, Coppin P (1999) Effects of area, age and diversity of forest
153 patches in Belgium on plant species richness, and implications for conservation
154 and reforestation. Biological Conservation 87:73-84
155 Jones AG, Ardren WR (2003) Methods of parentage analysis in natural populations.
156 Molecular Ecology 12:2511-2523
157 Keller LF, Waller DM (2002) Inbreeding effects in wild populations. Trends in
158 Ecology & Evolution 17:230-241
159 Kimura M, Imanishi J, Fukamachi K, Sakamoto K, Shibata S (2019) Remnant pattern
160 of woods and scattered trees of *Celtis sinensis*, *Aphananthe aspera* and *Zelkova*
161 *serrata* in the urban area of Kyoto for 30 years. Journal of The Japanese
162 Institute of Landscape Architecture 82:697-702
163 Lowe AJ, Cavers S, Boshier D, Breed MF, Hollingsworth PM (2015) The resilience of
164 forest fragmentation genetics—no longer a paradox—we were just looking in

the wrong place. Heredity 115:97

Marshall T, Slate J, Kruuk L, Pemberton J (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Molecular ecology 7:639-655

Oostermeijer J, Luijten S, Den Nijs J (2003) Integrating demographic and genetic approaches in plant conservation. Biological Conservation 113:389-398

Peakall R, Smouse PE (2012) GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. Bioinformatics 28:2537-2539

Rozen S, Skaletsky HJ (2000) Primer3 on the www for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics Methods and Protocols: Methods in Molecular Biology. Humana Press, Totowa, NJ

Sakaguchi S, Ueno S, Tsumura Y, Setoguchi H, Ito M, Hattori C, Nozoe S, Takahashi D, Nakamasu R, Sakagami T, Lannuzel G, Fogliani B, Wulff AS, L'Huillier L, Isagi Y (2017) Application of a simplified method of chloroplast enrichment to small amounts of tissue for chloroplast genome sequencing. Applications in Plant Sciences 5:1700002

Setoguchi H, Ohba H (1995) Phylogenetic relationships in *Crossostylis*

(Rhizophoraceae) inferred from restriction site variation of chloroplast DNA. Journal of Plant Research 108:87-92

Yoshikawa T, Kikuzawa K (2009) Pre-dispersal Seed Predation by a Granivorous Bird, the Masked Grosbeak (*Eophona personata*), in Two Bird-dispersed Ulmaceae Species. Journal of Ecology and Environment 32:137-143

Table 1

Nuclear microsatellite markers for Aphananthe aspera (Thunb.) Planch.. Shown for each primer pair are repeat motif, forward and reverse primer sequence (with tag sequence), the results of BLASTN database search with E-value and the accession number in DDBJ, and allele size range.

Locus	Repeat motif	Sequence	BLASTN	E-value	GenBank Accession No.	Allele size range (bp)
Aa_115573	$(GT)_{12}$	F:CACGACGTTGTAAAACGACTAATGGAGCCGCATCACTTG R:GTTTCTTACAGCTAGCTATTTCGTGGAAG	No significant hits	-	LC360719	122-128
Aa_84574	$(AAT)_{10}$	F:CGGAGAGCCGAGAGGTGTCGGCACCTTACACATTCTTG R:GTTTCTTCGTTAAGATGGGCAGACAGG	No significant hits	-	LC360720	171-180
Aa_105260	$(AC)_{14}$	F:CACGACGTTGTAAAACGACAAATTCACAGGAGCACATTTGG R:GTTTCTTGCTGCATGGCCATCTAACAG	No significant hits	-	LC360721	200-224
Aa_96067	$(AG)_{12}$	F:CTATAGGGCACGCGTGGTCGGGCTCGTTTGAAGAAGAC R:GTTTCTTGCCGTTGACTCAGACTACCC	No significant hits	-	LC360722	250-262
Aa_103887	$(GA)_{11}$	F:CACGACGTTGTAAAACGACTGCGGGTGAGAGATTCTAGC R:GTTTCTTGGAAGGTCAAAGCCCAAAGG	No significant hits	-	LC360723	356-366
Aa_95198	$(TG)_{13}$	F:CGGAGAGCCGAGAGGTGCCACGTCCCACTAGCTTCTC R:GTTTCTTGAACATGGCCGTGGTGTAG	No significant hits	-	LC360724	425-437

Aa_119647	$(GA)_{10}$	F:TGTGGAATTGTGAGCGGGGCTGTCTTAAGATGGGAGAAG R:GTTTCTTGTATTACAGCCCTGCACGAC	*Vitis vinifera*, whole genome shotgun sequence, contig VV78X118074.3, clone ENTAV 115	2.00E-12	LC360725	120-144
Aa_10258	$(TCT)_{27}$	F:TGTGGAATTGTGAGCGGTGGCCAGAAGTACCCTTGTC R:GTTTCTTCTCCTCGTCCGTATCCTTGG	No significant hits	-	LC360726	421
Aa_27760	$(CA)_{11}$	F:CGGAGAGCCGAGAGGTGACACACCGGAAGAAGAAAGC R:GTTTCTTGGACCACCTGCATACAAGAG	PREDICTED: *Ziziphus jujuba* myb-related protein Myb4-like (LOC107416961), mRNA	2.00E-15	LC360727	460-482
Aa_105963	$(GA)_{12}$	F:CTATAGGGCACGCGTGGTACTTTGCACAGATGAACAGAAC R:GTTTCTTTATTGCTGCTGGAGGATGGC	PREDICTED: *Prunus avium* ultraviolet-B receptor UVR8 (LOC110758246), mRNA	1.00E-14	LC360728	127-143
Aa_104223	$(GT)_{9}$	F:TGTGGAATTGTGAGCGGTGGGAATTTCAAATCCTGGCAG R:GTTTCTTACAAGAAGAGTCAAGCGCAG	No significant hits	-	LC360729	187-197
Aa_50376	$(TC)_{11}$	F:CTATAGGGCACGCGTGGTGGTGAACTTGTTGGGAGCAC R:GTTTCTTCATCCCACCCAATTTCCACC	No significant hits	-	LC360730	225-231
Aa_104956	$(GA)_{9}$	F:CACGACGTTGTAAAACGACGTCCTTGTCACGTGGCTTTC R:GTTTCTTTTTGCCAGAGATTGCAATTGG	No significant hits	-	LC360731	240-276
Aa_44791	$(AC)_{10}$	F:CACGACGTTGTAAAACGACTCGTAGGAATTTCGCACGTG R:GTTTCTTTCAGGATAATGGTGTCAAGCG	No significant hits	-	LC360732	423-439
Aa_104111	$(AG)_{11}$	F:CTATAGGGCACGCGTGGTACCAAACTCAACCACACCAG R:GTTTCTTGAAAGGATTGGCGTCGTTCC	No significant hits	-	LC360733	444-460
Aa_10240	$(TC)_{11}$	F:CGGAGAGCCGAGAGGTGAGTTTGTTTCCTCTTCTTTCTG R:GTTTCTTGTCATTATTGCCTAAGTTGCCC	No significant hits	-	LC360734	145-159
Aa_13484	$(AG)_{12}$	F:CTATAGGGCACGCGTGGTGATAAGGCGGGAGGAGTACG	No significant hits	-	LC360735	193-205

		R:GTTTCTTCCCATTTGCCCGGTTCTTTC				
Aa_17675	$(A)_{19}$	F:CGGAGAGCCGAGAGGTGGTATGTGTTTATGACCTTGTGC	No significant hits	-	LC360736	250
		R:GTTTCTTTCGTAATGTTACCTCGCTAAGC				
Aa_13178	$(AGT)_7$	F:CGGAGAGCCGAGAGGTGAAAGAAGGTTGAAGGCTGCG	No significant hits	-	LC360737	187-199
		R:GTTTCTTAGATCATCTTCTAATTCGCCAC				
Aa_12869	$(T)_{18}$	F:CTATAGGGCACGCGTGGTGCTAGGTCAAACTATGGGCC	No significant hits	-	LC360738	221
		R:GTTTCTTTTTCCCTCGACGAGCTATGG				
Aa_13711	$(AAT)_6$	F:TGTGGAATTGTGAGCGGATTGTGGGCCTCGACCTTAG	No significant hits	-	LC360739	327
		R:GTTTCTTAGCACCCAGTAACATCATGTG				
Aa_19942	$(A)_{18}$	F:CGGAGAGCCGAGAGGTGAGGGTGAGCTGTCCTGTAAG	No significant hits	-	LC360740	195-196
		R:GTTTCTTTGAGATGGTTCGGAGGTCTG				
Aa_10150	$(ATA)_7$	F:CGGAGAGCCGAGAGGTGGGCTGGATTGTTTGCATTGC	No significant hits	-	LC360741	234-273
		R:GTTTCTTTACGAACAGAGACGTGGTGG				
Aa_4179	$(AAT)_6$	F:CACGACGTTGTAAAACGACCCCAGATTAGCTAGATGTCAC	*Cucumis melo* genomic chromosome, chr_1	2.00E-16	LC360743	191-197
		R:GTTTCTTAACCAGCAATCCGAGCATTC				
Aa_15273	$(GA)_9$	F:CTATAGGGCACGCGTGGTACCAGGCCGATGAGAGTTTC	No significant hits	-	LC360744	244-246
		R:GTTTCTTGGCCCGTCATTTAACCAAGG				

Table 2

Genetic diversity statistics for two populations of *Aphananthe aspera* (Thunb.) Planch. based on the newly developed polymorphic genomic microsatellite markers. Abbreviations are as follow; N_A: number of alleles, H_O: observed heterozygosity, H_E: expected heterozygosity, Null: null allele frequency estimate. Deviations from Hardy-Weinberg equilibrium are noted by asterisks (** $P < 0.001$).

Locus	Shimogamo Shrine (n = 29)				Kawaketa-mikabe Shrine (n = 28)				Total (n = 57)		
	N_A	H_O	H_E	Null	N_A	H_O	H_E	Null	N_A	H_O	H_E
Aa_4179	2	0.034	0.034	-0.002	1	0.000	0.000	-0.002	2	0.018	0.017
Aa_10150	5	0.556	0.719	0.201	5	0.704	0.783	0.125	6	0.630**	0.768
Aa_10240	4	0.586	0.612	0.050	4	0.714	0.631	0.050	4	0.649	0.647
Aa_13178	3	0.241	0.216	-0.060	2	0.500	0.497	-0.060	3	0.368	0.416
Aa_13484	4	0.483	0.594	0.141	4	0.607	0.578	0.141	4	0.544	0.606
Aa_19942	2	0.448	0.348	-0.274	2	0.464	0.357	-0.286	2	0.456	0.352
Aa_27760	4	0.655	0.631	-0.007	4	0.714	0.682	-0.007	4	0.684	0.663
Aa_44791	3	0.500	0.498	0.002	2	0.464	0.469	0.002	3	0.482	0.508
Aa_50376	3	0.172	0.161	-0.038	3	0.286	0.283	-0.038	3	0.228	0.223
Aa_84574	3	0.552	0.565	0.024	2	0.393	0.392	0.024	3	0.474	0.492
Aa_95198	4	0.448	0.521	0.057	3	0.571	0.579	0.057	4	0.509	0.554
Aa_96067	4	0.483	0.508	0.037	4	0.679	0.658	0.037	4	0.579	0.604
Aa_103887	4	0.310	0.278	-0.068	3	0.464	0.424	-0.068	4	0.386	0.357
Aa_104111	5	0.724	0.736	0.001	5	0.679	0.642	0.001	5	0.702	0.699
Aa_104223	3	0.379	0.355	-0.025	2	0.143	0.133	-0.026	3	0.263	0.263
Aa_104956	8	0.621	0.737	0.116	4	0.643	0.600	0.116	8	0.632	0.695
Aa_105260	7	0.345**	0.675	0.563	4	0.071**	0.226	0.673	7	0.211**	0.495
Aa_105963	3	0.517	0.450	-0.043	3	0.393	0.456	-0.043	3	0.456	0.453
Aa_115573	3	0.690	0.651	-0.019	3	0.607	0.548	-0.019	3	0.649	0.614
Aa_119647	3	0.414	0.424	0.044	3	0.536	0.506	0.044	4	0.474	0.474
Average	3.9	0.458	0.486	0.035	3.2	0.482	0.472	0.036	4.0	0.470	0.495

06
07
08