Applications in Plant Sciences 2017 5(10): 1700086

Primer Note

Development and characterization of genomic SSR markers for *Anneslea fragrans* (Pentaphylacaceae)[1]

Lijing Sun[2], Kaikai Meng[3], Boyong Liao[3], Chunmei Li[3], Yue Zhang[4], Wenbo Liao[3,5], and Sufang Chen[3,5]

[2]Guangdong Polytechnic of Science and Trade, Guangzhou 510430, People's Republic of China; [3]State Key Laboratory of Biocontrol and Guangdong Key Laboratory of Plant Resources, Sun Yat-sen University, Guangzhou 510275, People's Republic of China; and [4]Management Burea of Guangdong Xiangtoushan National Nature Reserve, Huizhou 516003, People's Republic of China

- *Premise of the study:* The genus *Anneslea* (Pentaphylacaceae) contains four species and six varieties, most of which are locally endemic. Here, simple sequence repeat (SSR) markers were developed for the conservation of these species.
- *Methods and Results:* The genome of *A. fragrans* was sequenced and de novo assembled into 445,162 contigs, of which 30,409 SSR loci were detected. Primers for 100 SSR loci were validated with PCR amplification in three populations of *A. fragrans*. Seventy-nine loci successfully amplified, and 30 were polymorphic. The mean number of alleles, observed heterozygosity, and expected heterozygosity were 7.01 ± 1.60, 0.817 ± 0.241, and 0.796 ± 0.145, respectively. Most primers could be amplified in *Ternstroemia gymnanthera*, *T. kwangtungensis*, and *Cleyera pachyphylla*.
- *Conclusions:* Our study demonstrated that shotgun genome sequencing is an efficient way to develop genomic SSR markers for nonmodel species. These genomic SSR loci will be valuable in population genetic studies in *Anneslea* and its relatives.

Key words: *Anneslea*; genomic SSRs; Illumina sequencing technology; Pentaphylacaceae; shotgun genome sequencing.

The genus *Anneslea* Wall. (Pentaphylacaceae, previously Theaceae) contains only four species: *A. fragrans* Wall. occurs in China and Southeast Asia and is the most widespread of these species, while *A. donnaiensis* (Gagnep.) Kobuski and *A. paradoxa* H. H. Nguyen & Yakovlev are found only in Vietnam, and *A. steenisii* Kobuski is observed only in Sumatra (Angiosperm Phylogeny Group, 2016; Hassler, 2017). For *A. fragrans,* six varieties have been reported, with four distributed in China and two in Malaysia and Vietnam (Hassler, 2017). *Anneslea* was listed as a relict genus in tropical Asia (Liao and Jin, 2014), and its current status remains unknown and calls for scientific attention. Using "*Anneslea*" as a keyword, only two papers were found in a search of the Web of Science database (http://apps.webofknowledge.com), both of which analyzed the chemical constituents extracted from *A. fragrans*.

Anneslea fragrans, an evergreen tree or shrub (3–15 m in height), is an important species (Min and Bruce, 2007). Its Chinese name, which translates to "tea pear" in English, was given because of its reddish, *Camellia*-like flowers and edible, pear-like fruits. It has strong ecological adaptability to extreme environments, grows quickly, and is highly resistant to pests (Shen and Wang, 2009). It is planted as a garden tree in China.

In this study, we shotgun sequenced the *A. fragrans* genome with Illumina sequencing technology. Based on the assembled contigs, 30 polymorphic genomic simple sequence repeat (SSR) loci were developed and characterized in three populations of this species. The transferability of these markers was tested in *Ternstroemia gymnanthera* (Wight & Arn.) Bedd., *T. kwangtungensis* Merr., and *Cleyera pachyphylla* Chun ex H. T. Chang, which were previously listed in Theaceae, and later ascribed to Pentaphylacaceae (Min and Bruce, 2007; Angiosperm Phylogeny Group, 2016).

METHODS AND RESULTS

A seedling of *A. fragrans* was collected from the South China Botanical Garden, Guangdong, China (23°11′19.09″N, 113°22′22.51″E), and planted in the greenhouse of Sun Yat-sen University. Total genomic DNA was extracted from the fresh leaves using the modified cetyltrimethylammonium bromide (CTAB) method (Doyle and Doyle, 1987). A DNA library was constructed following the Illumina protocol and sequenced with HiSeq X Ten System (Illumina, San Diego, California, USA). The raw data were filtered with NGSQCToolkit_v2.3.3 (Patel and Jain, 2012), where low-quality reads containing unknown "N" bases or more than 10% bases with a Q value <20 were removed. Finally, a total of 25.4 million cleaned 145-bp paired reads were obtained and de novo assembled into 445,162 contigs with Edena v3.131028 (Hernandez et al., 2008). The mean length and the N50 value were 433 bp and 455 bp, respectively, and the longest contig was 39,694 bp. The cleaned raw data and the assembled contigs

[1]Manuscript received 9 August 2017; revision accepted 28 August 2017.

This work was supported by the Basic Work Special Project of the National Ministry of Science and Technology of China (2013FY111500); the Fourth National Survey on Chinese Traditional Medicine Resources (20147002); the Natural Science Foundation of Guangdong Province, China (2016A030313326); the Science and Technology Planning Project of Guangdong Province, China (2015A030302020); and Chang Hungta Science Foundation of Sun Yat-sen University

[5]Author for correspondence: lsslwb@mail.sysu.edu.cn (W.L.); chsuf@mail.sysu.edu.cn (S.C.)

doi:10.3732/apps.1700086

Table 1. Characteristics of 30 polymorphic genomic SSR loci developed for *Anneslea fragrans*.[a]

Locus	Primer sequences (5′–3′)	Repeat motif	Expected allele size (bp)	Allele number (size range), bp	Putative function	*E*-value	GenBank accession no.
CL6	F: TTCCCCAAATACAGCAGTCC R: ATGAACCATCCAGGCAGAAG	$(GA)_{23}$	161	10 (133–157)	Pectin lyase-like superfamily protein [*Arabidopsis thaliana*]	8e-11	MF579145
CL11	F: AATGCACATGCAATGCCTTA R: TAAAAGAGGTGTGCTCGGCT	$(CT)_{21}$	266	10 (254–274)	None	—	MF579149
CL13	F: ACATGTGCGGGACAATTTTT R: CCCTTTCCCCTCTACTTTGG	$(AG)_{21}$	182	15 (142–176)	None	—	MF579150
CL16	F: TCACCAAGTAGCAGCAGTCG R: TGGTTTCTTTGGCAGATTCC	$(GA)_{21}$	192	12 (156–190)	None	—	MF579152
CL22	F: AAGCAAGGTGAATCCACAGG R: TTTCCCATTGCAAGAGTGTG	$(AG)_{20}$	139	11 (111–137)	None	—	MF579157
CL31	F: CAAAATTAAAAATTAGGTGTTATGTCA R: TGGCCTATTTATGAGGATGGA	$(TAA)_{13}$	231	8 (207–231)	None	—	MF579164
CL34	F: AGCTGATTTCCTCCCCAAAT R: TTCTGAATTATTTCGAATGGTCTC	$(AAC)_{13}$	280	9 (244–271)	None	—	MF579166
CL35	F: GCATCTATAATTGCTACCTGAGCAT R: GAAGCTGGAATAAATAAACATAACCA	$(CTT)_{13}$	252	9 (234–258)	None	—	MF579167
CL36	F: TGCGAACGAGTTCATCTTTG R: ATGCCATAACATGGTAGCCC	$(AAG)_{13}$	234	8 (201–222)	None	—	MF579168
CL39	F: GACTGATGCCAAAGTCATCG R: GGACGTCCAAATCCAAAAGA	$(TCA)_{12}$	172	8 (142–163)	None	—	MF579170
CL40	F: CATAGCAGCCGTGAAGCATA R: ACCAGCAGACCAGGTACGAC	$(AAG)_{12}$	122	7 (113–131)	Unnamed protein product [*Vitis vinifera*]	4e-12	MF579171
CL41	F: TGCTTTGAATGACACCCTTG R: AATGCTTTCCCCAAACAGTG	$(CTT)_{12}$	206	6 (179–194)	Nitrous oxide reductase, N-terminal [*Medicago truncatula*]	6e-08	MF579172
CL44	F: CACTTGAAGGAAAGGGCAAA R: GCTCCAGAAGCCTTTCTGTC	$(AAG)_{12}$	216	8 (189–216)	None	—	MF579175
CL48	F: GGTTTCCAACGAGTAACCGA R: GCATGGCCTGTTAGAGTTCC	$(TTC)_{12}$	219	7 (207–225)	Ovate protein [*Lycopersicon esculentum*]	4e-24	MF579177
CL52	F: TTTGCAAAGAAATTAGAGGGAAA R: CAAGATCCAGTAGAAGAGGGAGA	$(TC)_{20}$	248	11 (206–230)	None	—	MF579180
CL53	F: AAACAACGCGAACCAAGAAT R: TAAGGGTTGAAATGGTGGGA	$(AG)_{20}$	253	10 (217–235)	None	—	MF579181
CL56	F: AATCAAAGCTGCAAACCACA R: AGCCTCCTACAAATTGGCCT	$(GA)_{19}$	242	10 (216–240)	FRIGIDA-LIKE 2 [*Arabidopsis thaliana*]	1e-09	MF579182
CL58	F: GCCAAATCAAACCGAAATGA R: CCAAACCATTTTTGTTTCCC	$(AG)_{19}$	270	9 (250–270)	None	—	MF579184
CL60	F: TGAATCTGCATACTCAAAGAGAAA R: ACCGATGACTCCCTTCAATG	$(GA)_{19}$	155	10 (133–155)	None	—	MF579186
CL61	F: CAACATCCAGCAACGTCATC R: TACACCAATCACAGCTTGGC	$(TC)_{19}$	252	8 (238–258)	Unknown protein [*Arabidopsis thaliana*]	3e-21	MF579187
CL62	F: CAGTGTTGGACAGCTCTGGA R: GAGATCATTGAACATTTTCAGCA	$(CT)_{19}$	202	7 (174–186)	None	—	MF579188
CL64	F: CAGTTGTCCTGTGAAAGCCA R: GCTGCCCTTGATGACCTTTA	$(AG)_{19}$	211	9 (193–211)	None	—	MF579190
CL66	F: CCACGAAATAGTAAATGCCG R: CGCACCGGTTCTTCTTAAAT	$(CA)_{19}$	279	8 (265–279)	None	—	MF579192
CL70	F: TCTCCCAACACCCATTTTGT R: AACTTGGGCTGCATGGTAAC	$(AC)_{19}$	256	9 (234–254)	None	—	MF579196

TABLE 1. Continued.

Locus	Primer sequences (5′–3′)	Repeat motif	Expected allele size (bp)	Allele number (size range), bp	Putative function	*E*-value	GenBank accession no.
CL73	F: ACCTTTGGTTGAAACGATGC	$(CT)_{18}$	181	7 (131–143)	None	—	MF579198
	R: ACGTATTTGAACGTGGGGAA						
CL77	F: TTTTGAAATGTGGATTTTAGCAA	$(AAG)_{12}$	188	8 (146–167)	None	—	MF579201
	R: TGATTCAAGATGCCACTCGT						
CL82	F: TCGAGTCTTGGGCTATGCT	$(GAG)_{12}$	204	7 (171–189)	None	—	MF579205
	R: ATTCCCGCGTGAAGATGTAG						
CL89	F: GGAGTGATGGGTCAAAGGAA	$(GAA)_{11}$	184	6 (160–178)	Peroxidase, putative [*Arabidopsis thaliana*]	5e-13	MF579210
	R: GCTCCTCATTTCCAAAACCA						
CL90	F: GCAAGCTTGAAGGTCTACGG	$(CTT)_{11}$	150	6 (126–141)	None	—	MF579211
	R: AATGAAACCCCCACCAATTT						
CL98	F: CACAATGTTAAGACGTTGAGTGG	$(ATT)_{11}$	214	9 (187–211)	None	—	MF579218
	R: TTATCGAGAGATACTCTTTAGGAAGG						

[a] Annealing temperature was 52°C for all loci.

were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA; SRR5880481) and Transcriptome Shotgun Assembly (TSA; GFTZ00000000) databases.

Applying the perl script MISA (Thiel et al., 2003) with the default parameters except that the settings for mononucleotide repeats were removed from the analysis, a total of 30,409 SSRs were detected in 25,855 contigs. Among these SSR loci, dinucleotide repeats were the most common (80.4%), followed by trinucleotide (15.3%), tetranucleotide (3.0%), pentanucleotide (0.8%), and hexanucleotide repeats (0.5%).

Using the online perl scripts p3-in and p3-out (http://pgrc.ipk-gatersleben.de/misa/primer3.html) and Primer3 (Rozen and Skaletsky, 1999), we successfully designed paired primers for 20,179 SSR loci with an expected PCR product size of 100–280 bp and melting temperature of 60°C. The paired primers designed for the 20,179 SSR loci and their characterization are provided in Appendix S1. Among them, primers for the 100 SSR loci containing the largest number of dinucleotide or trinucleotide repeats were screened for polymorphism in the following experiments.

Fresh leaves were collected from three populations of *A. fragrans* in China (population Eshan [ES] located in Yunnan Province, population Yangchun [YC] located in Guangdong Province, and population Jinggangshan [JGS] located in Jiangxi Province; Appendix 1) and then stored in plastic bags with silica gel. Total DNA was extracted with the modified CTAB method. For the first trial experiment, two individuals were randomly selected from each population, PCR amplifications were performed for the selected 100 paired primers following the procedure used in Fan et al. (2013), and agarose gel electrophoresis (1%) was used to check amplification. Seventy-nine loci successfully amplified in the six individuals with expected product size (Table 1, Appendix 2). The assembled sequences for these 79 SSR loci were deposited in the NCBI GenBank database (accession no.: MF579141–MF579219).

The PCR products were further inspected with the Fragment Analyzer Automated CE System (Advanced Analytical Technologies [AATI], Ames, Iowa, USA), in which the Quant-iT PicoGreen dsDNA reagent kit (35–500 bp; Invitrogen, Carlsbad, California, USA) was used. Finally, the raw data were analyzed using PROSize version 2.0 software (AATI); these results showed that among these 79 SSR loci, 30 were polymorphic among the six individuals (Table 1).

Polymorphism of the 30 loci was checked in 54 individuals collected from the three populations. PCR products were inspected to calculate the polymorphism level using the above-mentioned procedures. GenAlEx version 6.5 (Peakall and Smouse, 2012) was used to calculate linkage disequilibrium, deviation from Hardy–Weinberg equilibrium (HWE), average number of alleles per locus, observed heterozygosity, and expected heterozygosity. The tests for linkage disequilibrium showed that 59 of the 435 tests were significant ($P < 0.05$; Appendix 3), indicating that some paired loci may be linked with each other. The number of alleles per locus ranged from four to 10 (7.01 ± 1.60), the observed heterozygosity values ranged from 0.053 to 1.000 (0.817 ± 0.241), and the expected heterozygosity values ranged from 0.572 to 1.000 (0.796 ± 0.145). HWE testing showed that 14, 22, and 20 loci demonstrated significant deviation from HWE in the populations ES, YC, and JGS, respectively (Table 2).

Transferability of the 30 loci was tested in four to six individuals of three related species: *T. gymnanthera*, *T. kwangtungensis*, and *C. pachyphylla* (Appendix 1). Results showed that 22, 21, and 19 paired primers amplified in *T. gymnanthera*, *T. kwangtungensis*, and *C. pachyphylla*, respectively, and 19 amplified in all three species (Table 3).

CONCLUSIONS

In this study, we developed 30 new polymorphic genomic SSR markers based on whole-genome shotgun sequencing of *A. fragrans*. Our study showed that shotgun sequencing is an efficient way to develop highly polymorphic genomic SSR markers. These SSR markers are valuable in population genetic studies of *Anneslea* and its relatives.

LITERATURE CITED

ANGIOSPERM PHYLOGENY GROUP. 2016. An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV. *Botanical Journal of the Linnean Society* 181: 1–20.

Table 2. Genetic characterization of 30 polymorphic microsatellites of *Anneslea fragrans*.[a]

	Eshan County (n = 16)			Yangchun (n = 19)			Jinggangshan (n = 19)		
Locus	A	H_o	H_e[b]	A	H_o	H_e[b]	A	H_o	H_e[b]
CL6	10	1.000	0.852	9	1.000	0.870	8	0.947	0.744
CL11	9	0.875	0.836	9	1.000	0.852	9	0.895	0.803*
CL13	8	0.875	0.811*	10	0.842	0.864***	8	0.947	0.789
CL16	10	0.875	0.854***	8	0.895	0.810	9	0.895	0.823**
CL22	6	1.000	0.762***	9	0.947	0.839***	10	1.000	0.860
CL31	7	0.813	0.824	7	0.947	0.830	7	1.000	0.769***
CL34	7	0.813	0.824	7	1.000	0.792***	6	0.684	0.795***
CL35	9	1.000	0.848	7	1.000	0.795*	5	0.895	0.752***
CL36	7	1.000	0.818*	6	1.000	0.794***	6	0.947	0.752**
CL39	5	0.875	0.703**	4	0.895	0.716***	6	0.947	0.729***
CL40	5	1.000	0.662	6	0.947	0.730*	6	0.947	0.681
CL41	5	1.000	0.729**	6	1.000	0.780**	4	0.895	0.702*
CL44	8	1.000	0.818	7	0.947	0.733*	8	0.947	0.741
CL48	6	0.875	0.803	5	0.895	0.717***	6	0.947	0.785**
CL52	8	0.813	0.830	8	1.000	0.805	6	0.579	0.781***
CL53	8	0.938	0.832	10	0.947	0.860**	9	1.000	0.819
CL56	8	1.000	0.805	8	1.000	0.848***	8	1.000	0.781***
CL58	9	0.438	0.861***	7	0.474	0.846***	8	0.579	0.806*
CL60	8	0.875	0.799	9	0.947	0.834	8	0.947	0.802
CL61	7	0.375	0.791**	6	0.211	0.698***	6	0.211	0.741***
CL62	7	0.750	0.826*	4	0.368	0.586	7	0.947	0.751***
CL64	8	0.750	0.766**	7	0.474	0.803**	5	0.105	0.702***
CL66	7	0.688	0.834**	7	0.158	0.733***	7	0.368	0.823***
CL70	9	0.938	0.846	7	1.000	0.828**	7	0.789	0.771
CL73	7	0.875	0.771	6	0.737	0.785	7	0.842	0.825*
CL77	5	0.250	0.695**	4	0.053	0.572***	5	0.737	0.627
CL82	6	1.000	0.742	5	0.789	0.668**	6	0.789	0.763
CL89	4	0.875	0.736***	6	0.421	0.712**	6	1.000	0.654***
CL90	5	0.563	0.684*	4	1.000	0.705***	6	0.947	0.717**
CL98	9	1.000	0.867	8	0.947	0.780***	9	0.895	0.842*

Note: A = number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; n = number of individuals sampled.
[a]Locality and voucher information are provided in Appendix 1.
[b]Asterisks indicate signficant deviation from Hardy–Weinberg equilibrium (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$).

Doyle, J. J., and J. L. Doyle. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

Fan, Q., S. F. Chen, M. W. Li, S. Y. He, R. C. Zhou, and W. B. Liao. 2013. Development and characterization of microsatellite markers from the transcriptome of *Firmiana danxiaensis* (Malvaceae s.l.). *Applications in Plant Sciences* 1: 1300047.

Hassler, M. 2017. World Plants: Synonymic checklists of the vascular plants of the world (Version May 2017). *In* Y. Roskov, L. Abucay, T. Orrell, D. Nicolson, N. Bailly, P. M. Kirk, T. Bourgoin, et al. [eds.], Species 2000 & ITIS Catalogue of Life. Species 2000, Naturalis, Leiden, The Netherlands.

Hernandez, D., P. Francois, L. Farinelli, M. Osteras, and J. Schrenzel. 2008. De novo bacterial genome sequencing: Millions of very short reads assembled on a desktop computer. *Genome Research* 18: 802–809.

Liao, W. B., and J. H. Jin. 2014. Relic plants in JGS region and their biological refuge characteristics. *In* W. B. Liao, Y. Y. Wang, Z. Li, S. L. Peng, C. Q. Chen, Q. Fan, F. L. Jia, et al. [eds.], Integrated study on biodiversity of Mount Jinggangshan regions in China. Science Press, Beijing, China.

Min, T. L., and B. Bruce. 2007. Theaceae. *In* C.Y. Wu, P. H. Raven, and D. Y. Hong [eds.], Flora of China, vol. 12. Science Press, Beijing, China, and Missouri Botanical Garden Press, St. Louis, Missouri, USA.

Patel, R. K., and M. Jain. 2012. NGS QC Toolkit: A toolkit for quality control of next generation sequencing data. *PLoS ONE* 7: e30619.

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

Rozen, S., and H. Skaletsky. 1999. Primer3 on the WWW for general users and for biologist programmers. *In* S. Misener and S. A. Krawetz [eds.], Methods in molecular biology, vol. 132: Bioinformatics: Methods and protocols, 365–386. Humana Press, Totowa, New Jersey, USA.

Shen, S. K., and Y. H. Wang. 2009. Seed germination characteristics and their significance in artificial propagation of native species *Anneslea fragrans. Shengtaixue Zazhi* 28: 2165–2170.

Thiel, T., W. Michalek, R. K. Varshney, and A. Graner. 2003. Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare* L.). *Theoretical and Applied Genetics* 106: 411–422.

TABLE 3. Cross-amplification of 30 *Anneslea fragrans* genomic SSR markers in three related species.

Species	CL6	CL11	CL13	CL16	CL22	CL31	CL34	CL35	CL36	CL39	CL40	CL41	CL44	CL48	CL52	CL53	CL56	CL58	CL60	CL61	CL62	CL64	CL66	CL70	CL73	CL77	CL82	CL89	CL90	CL98
Ternstroemia gymnanthera	—	+	—	—	—	—	—	—	+	+	—	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
T. kwangtungensis	—	+	—	—	—	—	—	—	+	+	—	+	+	+	+	+	+	—	+	+	+	+	+	+	+	+	+	+	+	+
Cleyera pachyphylla	—	+	—	—	—	—	—	—	+	+	—	+	+	+	—	+	+	—	+	+	+	+	—	+	+	+	+	+	+	+

Note: + = primers successfully amplified in all individuals; — = primers did not amplify in any individual.

Appendix 1. Geographic locations and voucher specimen information for *Anneslea fragrans* and three related species used in this study.[a]

Speies	Population	*N*	Collection locality	Geographic coordinates	Voucher no.
Anneslea fragrans Wall. var. *rubriflora* (H. H. Hu & H. T. Chang) L. K. Ling	YC	19	Yangchun, Guangdong, China	21°52′48.73″N, 111°25′21.20″E	*Sun XJD-004*
A. fragrans var. *fragrans*	ES	16	Eshan County, Yunnan, China	24°05′56.08″N, 102°11′46.00″E	*Q. Fan 15005*
A. fragrans var. *fragrans*	JGS	19	Jinggangshan, Jiangxi, China	26°38′55.06″N, 114°25′40.75″E	*Liu LXP-13-22966*
Ternstroemia gymnanthera (Wight & Arn.) Bedd.	—	6	Jinggangshan, Jiangxi, China	263°80′03.00″N,114°09′19.00″E	*Liao LXP-13-15320*
T. kwangtungensis Merr.	—	4	Jinggangshan, Jiangxi, China	26°30′30.06″N, 114°08′26.14″E	*Liao LXP-13-05933*
Cleyera pachyphylla Chun ex H. T. Chang	—	5	Ji'an, Jiangxi, China	26°02′12.98″N, 114°10′13.81″E	*Zhao LXP-13-16596*

Note: *N* = number of individuals.

[a]All voucher specimens are deposited in the Herbarium of Sun Yat-sen University (SYS), Guangzhou, Guangdong, China.

Appendix 2. Characteristics of 49 monomorphic SSR markers developed for *Anneslea fragrans*.

Locus[a]	Primer sequences (5′–3′)	Repeat motif	T_a (°C)	Expected allele size (bp)	GenBank accession no.
CL1	F: GAAGCCCTCAATGCCATAAA	$(GA)_{24}$	52	224	MF579141
	R: TGCATTGTGTGTGGTACACG				
CL2	F: CACGGAAGACCATAGAGGGA	$(AG)_{24}$	52	263	MF579142
	R: CAGAGGAAAGCTGAAATGGC				
CL3	F: CTCAACCATCCATTCCTTTCA	$(GA)_{23}$	52	269	MF579143
	R: ATGTATGGACAGGACAGGGG				
CL4	F: TGGACACCATGGAATCATAAA	$(AC)_{23}$	52	229	MF579144
	R: CGCTTAGTGCATAACGTCAA				
CL7	F: CCTGAGGTTTGGGAATTCAA	$(TC)_{22}$	52	226	MF579146
	R: TAAAATGCCCATCCTGTTCC				
CL8	F: TTTGCTGACTGTTGGCTGAC	$(TC)_{22}$	52	227	MF579147
	R: AAAGCAAACTTAATTCTAACCCTCAA				
CL9	F: TCCTTGGAGGAGGACTTCAA	$(CA)_{22}$	52	203	MF579148
	R: ACCACTTCCAAACCAACTCG				
CL14	F: GCATTATATTTCCCTTTCGGC	$(TC)_{21}$	52	244	MF579151
	R: GCTCCAAAGCCTTGTGAAAG				
CL18	F: ATGGCGAGGTAAAAGGTGTG	$(GT)_{21}$	52	275	MF579153
	R: TCGGTCCATCTTGAAGCTCT				
CL19	F: GACAACTATGCCACCTGCCT	$(TC)_{20}$	52	210	MF579154
	R: GGGTCATCCAAAGCACTGAT				
CL20	F: GGAGCTCACTAAACAATGCCA	$(TC)_{20}$	52	261	MF579155
	R: TTCCAATCAAATGAAAGCCC				
CL21	F: AAGTTTCTCTTGCACACGTACATC	$(GA)_{20}$	52	236	MF579156
	R: GGATTACGGGTCGGATAGGT				
CL23	F: AGTGTGCTGCTAATGTGCCA	$(TC)_{20}$	52	225	MF579158
	R: TTTGGCCAACTTCTGAATCC				
CL24	F: TGCGAAAGATCGAAAGTGAA	$(TC)_{20}$	52	231	MF579159
	R: TCGATCCACAGGGAATATGG				
CL26	F: GAGCTGAAACAACCTTCTTAACG	$(GAA)_{17}$	52	131	MF579160
	R: TTTCATGCTTGCCTTTCCTT				
CL28	F: TAGGCTGGTGTAAGATGCGT	$(TTC)_{15}$	52	135	MF579161
	R: GCAGTCATCTTGGCACTTTAAGA				
CL29	F: GCTTCCTCAAGCTTCAATGG	$(TTC)_{14}$	52	155	MF579162
	R: CCGGGTAAGGAAGGAAAGAG				
CL30	F: TTGATTCTTGATGATGCATGG	$(TTC)_{14}$	55	250	MF579163
	R: GATTGGAAGGCAAATCCAAA				
CL32	F: ATTGTTGGAGTCGTTGCCTC	$(CAC)_{13}$	55	168	MF579165
	R: AACCCTAGCCACCAAGTGTG				
CL37	F: CCTCTTCCCAAGCTTCTTCA	$(TTC)_{13}$	52	214	MF579169
	R: TGCCCATGGTCACATATCAG				
CL42	F: TCACTCATCGGGTAAGGAGG	$(GAA)_{12}$	52	154	MF579173
	R: GCTTCCTCAAGCTTCAATGG				
CL43	F: TCCCACAAGCCTCATTTTTC	$(TAT)_{12}$	52	265	MF579174
	R: GGCAAATAAATGACAAAGACCC				
CL45	F: GAAACTGAGGAGAAAGAACCGA	$(AAG)_{12}$	52	265	MF579176
	R: CTAAGAAGGGCACAAGGACG				
CL50	F: CGAGCATGGAGTGTGTTTTG	$(GA)_{20}$	55	260	MF579178
	R: AGCTTGGATGTTGTCCCATC				
CL51	F: ACGAGAGCTCGATGAAGACA	$(TG)_{20}$	52	234	MF579179
	R: TGGTAAAGATGTTGCTTGGG				
CL57	F: GGAACCAGACCAGAAATGGA	$(AG)_{19}$	52	247	MF579183
	R: GTAAGGACAGACTCCAGGCG				

Appendix 2. Continued.

Locus[a]	Primer sequences (5′–3′)	Repeat motif	T_a (°C)	Expected allele size (bp)	GenBank accession no.
CL59	F: TGGCCAGTGTTCAAAAGTTG	$(GT)_{19}$	55	273	MF579185
	R: CACATCTGTCAGGTCTCCCA				
CL63	F: TCCAAGTGCTTTGTTGCTTG	$(GT)_{19}$	52	254	MF579189
	R: CGAAGAAGAGCTCGAGATGG				
CL65	F: CTGATGATGGAGAGGGCATT	$(TC)_{19}$	52	199	MF579191
	R: CGGTTCCTTTCACGATCAAT				
CL67	F: TGCTTCTTCACACGCATCTC	$(CT)_{19}$	52	144	MF579193
	R: TATCATTTTCCCAAACCCCA				
CL68	F: CCCATTTACAGATTGAAAAACGA	$(AC)_{19}$	55	219	MF579194
	R: ACATGTTGATGGCGAAGTGA				
CL69	F: CCCACAAAAAGCTGAAGGAG	$(GA)_{19}$	52	199	MF579195
	R: CATACTCCTCAGGGAGCCAA				
CL71	F: TGGCTAAGAGGGCAAAAAGA	$(AG)_{19}$	52	254	MF579197
	R: GCTAGCAACGTTCTTCCCTG				
CL74	F: AAGGTTGTGTTCCCTTCACG	$(GA)_{18}$	52	259	MF579199
	R: CAGAATTCAGATTTTGCTGTCAA				
CL76	F: GCCTAAGCCAAAACCCTACC	$(GAA)_{12}$	55	138	MF579200
	R: CACCAAAAGCATAAAAAGGCA				
CL78	F: GGCAGTTTTCTTCTTGCGAC	$(TCC)_{12}$	52	230	MF579202
	R: TATACGAGGAGGCCGATGAG				
CL79	F: GCTTCCTCAAGCTTCAATGG	$(TTC)_{12}$	52	268	MF579203
	R: GAGATCCCGAACCTCCTTTC				
CL81	F: TCTGCTACTGCCGCTAACTTT	$(CTT)_{12}$	52	153	MF579204
	R: CACTGGTTTTTCCGACAACA				
CL83	F: TCCCAGTGGTAAAAGTGATGC	$(TCC)_{11}$	52	222	MF579206
	R: ACCAACTTGTGGGTTGGTTT				
CL85	F: CGTGCTGCTGCAGGATAATA	$(CCA)_{11}$	52	207	MF579207
	R: ATTGTTGCAAGCTGCCTTTT				
CL86	F: TAACCAAAAGGCGAAACCAG	$(AGA)_{11}$	52	241	MF579208
	R: AAGGGACCCATTTTGAGGTT				
CL87	F: CCTGTGGTCACACATGAAGG	$(AAG)_{11}$	52	234	MF579209
	R: TCCTAACTCGGCCCTTATCA				
CL91	F: GGCATATGGATGGACTTTGC	$(TTG)_{11}$	55	140	MF579212
	R: GGCGCTCTATCGAACTTGAG				
CL93	F: CCGGGTAAGGAAGGAAAGAG	$(AAG)_{11}$	52	161	MF579213
	R: GCTTCCTCAAGCTTCAATGG				
CL94	F: CACCCACACAATCACCTCAC	$(AGA)_{11}$	52	166	MF579214
	R: AAGCATCAACAATGGCTTCC				
CL95	F: CACCCTCCTCCTCATTTCTTC	$(TTC)_{11}$	52	181	MF579215
	R: AAAACCTGAGAGGGAAAATAAAAA				
CL96	F: CCAAGACCTAGTCGTGCTCT	$(ACA)_{11}$	55	139	MF579216
	R: GGTTGCTTCAATTGCTCCAT				
CL97	F: ACGCCCAGTAGTCATCTTGG	$(AAG)_{11}$	52	216	MF579217
	R: ATGACATCCCAAGGGTTCTG				
CL99	F: AACTTGGTGGCTCATTCGAC	$(AAT)_{10}$	52	184	MF579219
	R: CGGAGAGCCCAAATTAGTGA				

Note: T_a = annealing temperature.

Appendix 3. Linkage disequilibrium tests of the 30 polymorphic loci developed for *Anneslea fragrans* among all populations.

Locus 1	Locus 2	*P* value[a]
CL35	CL39	0.000
CL48	CL90	0.000
CL36	CL89	0.001
CL16	CL98	0.001
CL35	CL58	0.002
CL13	CL89	0.002
CL39	CL48	0.003
CL52	CL56	0.004
CL35	CL61	0.004
CL41	CL53	0.007
CL6	CL41	0.009
CL35	CL60	0.009
CL40	CL58	0.010
CL22	CL31	0.012
CL58	CL70	0.012
CL35	CL48	0.013
CL16	CL35	0.013
CL34	CL77	0.013
CL61	CL89	0.013
CL6	CL40	0.014
CL36	CL90	0.014
CL6	CL48	0.015
CL31	CL35	0.016
CL41	CL61	0.017
CL35	CL66	0.017
CL31	CL39	0.018
CL13	CL56	0.018
CL48	CL66	0.019
CL35	CL90	0.020
CL16	CL52	0.020
CL35	CL41	0.020
CL44	CL73	0.020
CL35	CL89	0.021
CL34	CL39	0.022
CL60	CL90	0.023
CL53	CL89	0.026
CL66	CL82	0.026
CL16	CL34	0.027
CL35	CL44	0.028
CL56	CL89	0.029
CL31	CL40	0.029
CL52	CL77	0.030
CL52	CL82	0.032
CL13	CL31	0.032
CL44	CL62	0.032
CL44	CL90	0.034
CL56	CL62	0.036
CL73	CL90	0.036
CL36	CL53	0.038
CL53	CL73	0.038
CL22	CL44	0.040
CL98	CL40	0.040
CL40	CL82	0.041
CL40	CL48	0.043
CL34	CL52	0.043
CL39	CL52	0.044
CL52	CL90	0.044
CL13	CL36	0.047
CL31	CL34	0.048

[a] Only P values <0.05 are displayed.