Characterization of Microsatellite Loci for an Australian Epiphytic Orchid, Dendrobium calamiforme, Using Illumina Sequencing

Authors: Trapnell, Dorset W., Beasley, Rochelle R., Lance, Stacey L., Field, Ashley R., and Jones, Kenneth L.

Source: Applications in Plant Sciences, 3(6)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1500016

Applications in Plant Sciences

Applications in Plant Sciences 2015 3(6): 1500016

Primer Note

Characterization of microsatellite loci for an Australian epiphytic orchid, *Dendrobium calamiforme*, using Illumina sequencing[1]

Dorset W. Trapnell[2,7], Rochelle R. Beasley[3], Stacey L. Lance[3], Ashley R. Field[4,5], and Kenneth L. Jones[6]

[2]Plant Biology Department, University of Georgia, Athens, Georgia 30602 USA; [3]Savannah River Ecology Laboratory, University of Georgia, Aiken, South Carolina 29802 USA; [4]Queensland Herbarium, Department of Science, Information Technology and Innovation, Toowong, Queensland 4066, Australia; [5]Australian Tropical Herbarium, James Cook University, Smithfield, Queensland 4878, Australia; and [6]Department of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, Aurora, Colorado 80045 USA

- *Premise of the study:* Microsatellite loci were developed for the epiphytic pencil orchid *Dendrobium calamiforme* for population genetic and phylogeographic investigation of this Australian taxon.
- *Methods and Results:* Nineteen microsatellite loci were identified from an Illumina paired-end shotgun library of *D. calamiforme*. Polymorphism and genetic diversity were assessed in 24 individuals from five populations separated by a maximum distance of ~80 km. All loci were polymorphic with two to 14 alleles per locus, expected heterozygosity ranging from 0.486 to 0.902, and probability of identity values ranging from 0.018 to 0.380.
- *Conclusions:* These novel markers will serve as valuable tools for investigation of levels of genetic diversity as well as patterns of gene flow, genetic structure, and phylogeographic history.

Key words: *Dendrobium calamiforme*; *Dockrillia calamiformis*; genetic diversity; Orchidaceae; phylogeography; simple sequence repeat (SSR) markers.

Molecular phylogeographic approaches can provide potent tests of historical biogeographic hypotheses, such as the influence of historical barriers to gene flow on evolutionary diversification. The tropical rainforests of northeastern Australia harbor a diverse flora rich in basal angiosperm lineages that has long been thought to have been assembled principally through ecological filtering of relict Gondwanan stock and exchange of lineages with Malesia and Southeast Asia (e.g., Webb and Tracey, 1981; Crayn et al., 2015). However, the role of in situ diversification in this old biome may be underappreciated. Within this biome, a congruent genetic discontinuity has been found in various fauna groups (e.g., Schneider et al., 1998) and tree species (Rossetto et al., 2009) across the biogeographic barrier known as the Black Mountain Corridor (BMC), located between Cairns and Cape Tribulation. To better understand the processes that gave rise to this pattern, and the significance of in situ diversification to the origins and maintenance of tropical rainforest diversity, we aim to determine the phylogeographic structure of a codistributed epiphytic orchid. These orchids release tiny, wind-borne seeds high in the air column, where they can be picked up by wind currents and potentially transported great distances.

Dendrobium calamiforme Lodd. ex Lindl., commonly known as the pencil orchid in reference to the long, terete leaves, had been renamed *Dockrillia calamiformis* (Lodd. ex Lindl.) M. A. Clem. & D. L. Jones (Clements and Jones, 1996); however, this was rejected by Adams (2011). This orchid is indigenous to coastal tropical Queensland, Australia, ranging from Badu Island in the Torres Strait to Mount Elliott near Townsville, with nearly continuous distribution in its habitat across its range. It is a canopy and subcanopy epiphyte that grows in vine forest, swamp forest, beach forest, and riparian forest but is uncommon in ever-wet closed canopy rainforest. Although it can occur on large boulders, populations reach their highest density in large mature trees and can be locally abundant. Individuals become reproductive within five years and can live for several decades. *Dendrobium calamiforme* flowers in the dry season (July to September) and, while the pollination syndrome has not been verified, Hymenoptera, Coleoptera, and birds have been observed visiting flowering plants.

The development of highly polymorphic microsatellite markers will allow insights into the levels and partitioning of neutral genetic variation in this common epiphytic orchid. With these markers, patterns of seed dispersal, colonization, and genetic connectivity across the BMC will be investigated. Based on the dispersal ability of *D. calamiforme*, we predict low genetic

[1]Manuscript received 19 February 2015; revision accepted 7 April 2015.

The authors thank Dat Hoang for laboratory assistance, the Biostatistics/Bioinformatics Shared Resource at the University of Colorado Cancer Center (5P30CA046934) for bioinformatics support, Darren Crayn for valuable feedback on the manuscript, and an anonymous reviewer. Funding was provided by a University of Georgia Faculty Research Grant (D.W.T.), the University of Georgia Office of the Vice President of Research (D.W.T.), and by the U.S. Department of Energy under Award Number DE-FC09-07SR22506 to the University of Georgia Research Foundation.

[7]Author for correspondence: dorset@uga.edu

doi:10.3732/apps.1500016

structure among populations straddling the BMC; however, biogeographic disjunctions have been found in Costa Rican epiphytic orchids (Trapnell and Hamrick, 2004; Kartzinel et al., 2013; Trapnell et al., unpublished) that appear to be maintained by cryptic processes.

METHODS AND RESULTS

Total DNA was extracted from one individual of *D. calamiforme*, following the cetyltrimethylammonium bromide (CTAB) protocol of Doyle and Doyle (1990). After shearing 1 µg of genomic DNA with a Covaris S220 Focused-ultrasonicator (Covaris, Woburn, Massachusetts, USA), a paired-end shotgun library was prepared with the Illumina TruSeq DNA Library Kit (Illumina, San Diego, California, USA). During library preparation, a multiplex identified adapter was incorporated as multiple species were run together on an Illumina MiSeq using 100-bp paired-end reads. The program *PAL_FINDER_v0.02.03* (Castoe et al., 2012) was used to examine 5 million reads and identify those containing microsatellite repeats; positive reads were targeted for primer design using Primer3 (version 2.0.0; Rozen and Skaletsky, 1999). The frequency of designed primer sequences in all reads was assessed by the software, and only primers whose sequences occurred ≤2 times were selected to avoid duplicated loci. Forty-eight loci of the 5412 that met this criterion were chosen. To use a three-primer PCR with one universally labeled primer (CAG tag 5′-CAGTCGGGCGTCATCA-3′), one primer from each pair was modified at the 5′ end with the addition of the CAG tag sequence and the 5′ end of the second primer from each pair was modified with the addition of GTTT.

The selected 48 primer pairs were tested for clean amplification and polymorphism across four individuals. The PCR amplifications were performed in a 12.5-µL volume (10 mM Tris [pH 8.4], 50 mM KCl, 0.25 µg bovine serum albumin [BSA], 0.4 µM unlabeled primer, 0.04 µM tag-labeled primer, 0.36 µM universal dye-labeled primer, 3.0 mM $MgCl_2$, 0.8 mM dNTPs [Thermo Scientific, Waltham, Massachusetts, USA], 0.5 units AmpliTaq Gold Polymerase [Life Technologies, Carlsbad, California, USA], and 20 ng DNA template) using an Applied Biosystems GeneAmp PCR System 9700 (Life Technologies). For all loci, a touchdown thermal cycling program (Don et al., 1991) was used that had a 10°C span of annealing temperatures ranging between 65–55°C (TD65). The cycling profile consisted of an initial denaturation step of 5 min at 95°C followed by 20 cycles of 95°C for 30 s, highest annealing temperature (decreased 0.5°C per cycle) for 30 s, and 72°C for 30 s; and 20 cycles of 95°C for 30 s, lowest annealing temperature for 30 s, and 72°C for 30 s; and a final extension at 72°C for 5 min. After amplification, all PCR products were genotyped by running on an ABI-3130xl sequencer (Life Technologies) and using a size standard prepared according to DeWoody et al. (2004), with the exception that primers that were not fluorescently labeled had GTTT added to their 5′ ends. Results of fragment analysis were analyzed using GeneMapper version 3.7 (Life Technologies). Of the 48 loci tested, 19 yielded high-quality polymorphic PCR products and were characterized by trinucleotide (12 loci)

Table 1. Characteristics of 19 polymorphic microsatellite loci developed for *Dendrobium calamiforme*.[a]

Locus	Primer sequences (5′–3′)	Repeat motif	Allele size range (bp)[b]	TD[c]
Doca5	F: *GAAGGTGGTAGTGCCAGAGC	ATC	252–297	TD65
	R: AACTTGAATAACACACCCAAGC			
Doca6	F: *AGTTGTAAGCAATGTGCTAGGC	ATT	219–273	TD65
	R: AAGGTTCAATTATTGCTCATAGGC			
Doca10	F: *TTGCTCTCTCTCTTCCTACAAATGC	ATCT	212–222	TD65
	R: AGAGCGATAGAGAGGGTCTAAGATAGC			
Doca11	F: *GTCTTTGCTGACCAAGGTGC	ATT	194–209	TD65
	R: AGGAGACGAGTCCAGGATGG			
Doca13	F: *CCACCTGCACCCATCTATCC	TTC	175–211	TD65
	R: CAACGCAAGGAAAGTCTCCG			
Doca14	F: *AATGCAATGACCATAAAGTGGC	ATT	251–293	TD65
	R: CAACTAATCTACCATGCCTTCAAGC			
Doca15	F: *GGAAGCTGTGGGATTTCTGC	TTC	184–206	TD65
	R: CATGCTTCACAGCCCATCC			
Doca16	F: *CATTTGACGATATGGCGGG	ATT	145–181	TD65
	R: CCAAAGACCGCTCTTGAAGG			
Doca18	F: *CATAATGGAGTCTTTCCTTCACCC	ATCT	260–338	TD65
	R: CTGTTGGCGACCTCAGACC			
Doca19	F: *GACGACAGAATGTATGGCCG	TTC	222–271	TD65
	R: GACATAGATGCCGGTGAAGC			
Doca25	F: *GACCTAAACTTTCAACTTATACTCATAGCC	ATT	219–264	TD65
	R: GCTCCTGATGCACAAATAAGAGC			
Doca27	F: *CTTCAATTACCCGACGAGCC	ATCT	173–175	TD65
	R: GGAGACTGAGGTGAGACCAGG			
Doca28	F: *TGCATTTGCTATACAAACATCCC	TTC	283–313	TD65
	R: GGCTTCACAGGGATTGAACC			
Doca33	F: *CATATAAGATCGATAACTGATAAGACCG	AATC	204–212	TD65
	R: TACCATACCAGCATAGCCGC			
Doca37	F: *GGCAGACAAAGAGAGAATAAGGG	ATCT	195–299	TD65
	R: TCTCTCAAACCCTCCACTTCC			
Doca38	F: *GAGAGAGACACAAAGCAAGGG	ATCT	216–276	TD65
	R: TCTCTCTATCCCTCCACTGGC			
Doca39	F: *AGAGAAAGCGCGAGGAGG	ATCT	244–267	TD65
	R: TCTCTTTATCCCTTCCCTTGC			
Doca40	F: *GGATATTAGAGTATGAACAAGGCATGG	ATT	402–477	TD65
	R: TCTTCTTTATAGCCATCAATCAATAGG			
Doca41	F: *CGCTTGAAAGCACTAAATGCC	ATC	272–299	TD65
	R: TGAAACAGGGCTCCAATTCC			

[a] The GenBank accession number for all loci is SAMN03437177.
[b] Includes the length of the CAG tag.
[c] Touchdown protocol used for PCR (see Methods and Results section).
* Indicates CAG tag (5′-CAGTCGGGCGTCATCA-3′) label.

TABLE 2. Genetic diversity values for five populations of *Dendrobium calamiforme* in Queensland, Australia, using 19 newly developed polymorphic microsatellite loci.[a]

	Font Hill Station (n = 5)				Bruce Hwy. at Toogood Rd. (n = 5)				Bruce Hwy. at Mulgrave River (n = 4)				Barron River (n = 4)				Centenary Lakes (n = 6)			
Locus	A	H_o	H_e	P_{ID}	A	H_o	H_e	P_{ID}	A	H_o	H_e	P_{ID}	A	H_o	H_e	P_{ID}	A	H_o	H_e	P_{ID}
Doca5†	3	0.800	0.540	0.285	4	0.400	0.640	0.188	4	0.500	0.656	0.169	4	0.250	0.656	0.169	4	0.500	0.653	0.174
Doca6	2	0.600	0.420	0.425	6	0.800	0.760	0.088	4	0.750	0.656	0.169	6	1.000	0.813	0.062	7	1.000	0.806	0.062
Doca10†	2	0.00	0.480	0.386	4	1.000	0.660	0.165	3	0.500	0.406	0.388	2	0.333	0.500	0.375	6	0.667	0.778	0.082
Doca11	2	0.600	0.420	0.425	3	0.200	0.620	0.217	4	1.000	0.719	0.130	6	0.750	0.813	0.062	5	0.500	0.736	0.114
Doca13†	3	1.000	0.620	0.217	5	0.600	0.600	0.190	3	0.250	0.406	0.388	4	1.000	0.667	0.157	3	1.000	0.569	0.278
Doca14†	3	0.600	0.660	0.189	4	0.400	0.720	0.126	4	0.250	0.719	0.130	4	0.250	0.656	0.169	4	0.167	0.681	0.161
Doca15†	3	0.200	0.580	0.265	4	0.400	0.580	0.221	4	0.500	0.656	0.169	4	0.250	0.719	0.130	5	0.667	0.764	0.094
Doca16	3	0.200	0.540	0.285	5	0.800	0.740	0.106	6	1.000	0.813	0.062	4	0.750	0.656	0.169	6	0.833	0.792	0.072
Doca18	4	1.000	0.688	0.155	6	1.000	0.780	0.080	5	0.750	0.750	0.101	4	1.000	0.719	0.130	8	1.000	0.847	0.041
Doca19†	2	0.600	0.420	0.425	5	0.400	0.680	0.140	4	0.750	0.688	0.155	3	0.500	0.625	0.211	6	0.833	0.778	0.082
Doca25†	4	0.200	0.700	0.145	4	0.400	0.580	0.221	3	0.000	0.625	0.211	3	0.250	0.531	0.283	6	0.167	0.792	0.072
Doca27†	1	0.000	0.000	1.000	2	0.000	0.320	0.514	1	0.000	0.000	1.000	1	0.000	0.000	1.000	2	0.333	0.444	0.407
Doca28†	2	0.000	0.480	0.386	2	0.000	0.375	0.461	2	0.000	0.375	0.461	2	0.000	0.375	0.461	4	0.167	0.514	0.274
Doca33†	2	0.000	0.480	0.386	2	0.200	0.180	0.689	2	0.250	0.469	0.392	3	0.250	0.594	0.248	2	0.333	0.278	0.560
Doca37	2	0.400	0.320	0.514	4	0.800	0.580	0.221	4	0.500	0.656	0.169	1	0.000	0.000	1.000	3	0.333	0.569	0.278
Doca38†	2	0.400	0.320	0.514	5	0.600	0.780	0.084	3	0.000	0.625	0.211	3	0.500	0.625	0.211	7	0.667	0.806	0.062
Doca39†	2	0.000	0.480	0.386	3	0.200	0.340	0.461	3	0.500	0.625	0.211	2	0.000	0.500	0.375	4	0.333	0.722	0.128
Doca40†	3	1.000	0.620	0.217	6	0.600	0.800	0.068	4	0.500	0.656	0.169	6	0.500	0.813	0.062	7	0.667	0.833	0.049
Doca41†	2	0.600	0.420	0.425	6	1.000	0.780	0.080	4	1.000	0.719	0.130	5	0.750	0.781	0.083	6	1.000	0.806	0.066
Mean	2.5	0.432	0.484	0.370	4.2	0.516	0.606	0.227	3.5	0.474	0.590	0.253	3.5	0.439	0.581	0.282	5.0	0.588	0.693	0.161

Note: A = number of alleles identified; H_e = expected heterozygosity; H_o = observed heterozygosity; n = number of individuals genotyped; P_{ID} = probability of identity.

[a] See Appendix 1 for locality and voucher information.

† Indicates significant deviations from Hardy–Weinberg expectations for all 24 samples after Bonferroni corrections.

and tetranucleotide (7 loci) repeat motifs (Table 1). The remaining 29 loci did not amplify well and therefore were not used.

We assessed the variability of these 19 loci in 24 specimens of *D. calamiforme* collected from five sites, spanning a distance of 79.7 km (Appendix 1). Vouchers from each site were deposited at the Australian Tropical Herbarium (CNS) (Appendix 1). Each site consisted of a small number of *D. calamiforme* individuals in each of two to five host trees. We used GenAlEx version 6.4 (Peakall and Smouse, 2006) to estimate the number of alleles per locus (A), observed heterozygosity (H_o), expected heterozygosity (H_e), and the probability of identity (P_{ID}). To test for deviations from Hardy–Weinberg equilibrium (HWE) and for linkage disequilibrium, GENEPOP version 4.0 (Rousset, 2008) was used.

These 19 loci were highly polymorphic with mean per locus values of A = 8.4 (range = 2–14), H_e = 0.754 (0.486–0.902), and H_o = 0.496 (0.043–0.957). Mean population values were A = 3.7 (range = 2.5–5.0), H_e = 0.591 (0.484–0.693), H_o = 0.489 (0.432–0.588), and P_{ID} = 0.259 (0.161–0.370) (Table 2). After Bonferroni correction for multiple comparisons, 14 loci showed significant deviation from expectations under HWE (Table 2). Linkage disequilibrium was detected for 61 of the 171 paired loci comparisons, which is not surprising considering that our samples came from multiple small populations.

CONCLUSIONS

The 19 novel microsatellites developed for *D. calamiforme* revealed high levels of polymorphism and genetic diversity and thus should prove valuable for elucidating levels and patterns of genetic variation in future population genetic and phylogeographic investigations of this species in northeastern Australia. These highly variable markers may also be useful for discerning species boundaries among *D. calamiforme* and the putative taxa *D. baseyanum* St. Cloud and *D.* ×*foederatum* St. Cloud, which have in the past been recognized as occurring in the Cairns area of northeastern Australia (Field and Zich, 2012).

LITERATURE CITED

ADAMS, P. B. 2011. Systematics of Dendrobiinae (Orchidaceae), with special reference to Australian taxa. *Botanical Journal of the Linnean Society* 166: 105–126.

CASTOE, T. A., A. W. POOLE, A. P. J. DE KONING, K. L. JONES, D. F. TOMBACK, S. J. OYLER-MCCANCE, J. A. FIKE, ET AL. 2012. Rapid microsatellite identification from Illumina paired-end genomic sequencing in two birds and a snake. *PLoS ONE* 7: e30953.

CLEMENTS, M. A., AND D. L. JONES. 1996. New species of Dendrobiinae (Orchidaceae) from Papua New Guinea. *Lasianthera* 1: 8–25.

CRAYN, D. M., C. COSTION, AND M. G. HARRINGTON. 2015. The Sahul–Sunda floristic exchange: Dated molecular phylogenies document Cenozoic intercontinental dispersal dynamics. *Journal of Biogeography* 42: 11–24.

DEWOODY, A. J., J. SCHUPP, L. KENEFIC, J. BUSCH, L. MURFITT, AND P. KEIM. 2004. Universal method for producing ROX-labeled size standards suitable for automated genotyping. *BioTechniques* 37: 348–350.

DON, R. H., P. T. COX, B. J. WAINWRIGHT, K. BAKER, AND J. S. MATTICK. 1991. 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research* 19: 4008.

DOYLE, J. J., AND J. L. DOYLE. 1990. Isolation of plant DNA from fresh tissue. *Focus (San Francisco, Calif.)* 12: 13–15.

FIELD, A. R., AND F. A. ZICH. 2012. Types of enigmatic north-Queensland orchids from the Dockrill herbarium. *Austrobaileya* 8: 696–698.

KARTZINEL, T. R., R. P. SHEFFERSON, AND D. W. TRAPNELL. 2013. Relative importance of pollen and seed dispersal across a Neotropical mountain landscape for an epiphytic orchid. *Molecular Ecology* 22: 6048–6059.

PEAKALL, R., AND P. E. SMOUSE. 2006. GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.

ROSSETTO, M., D. CRAYN, A. FORD, R. MELLICK, AND K. SOMMERVILLE. 2009. The influence of environment and life-history traits on the distribution of genes and individuals: A comparative study of 11 rainforest trees. *Molecular Ecology* 18: 1422–1438.

ROUSSET, F. 2008. GENEPOP'007: A complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources* 8: 103–106.

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.

Schneider, C. J., M. Cunningham, and C. Moritz. 1998. Comparative phylogeography and the history of endemic vertebrates in the Wet Tropics rainforests of Australia. *Molecular Ecology* 7: 487–498.

Trapnell, D. W., and J. L. Hamrick. 2004. Partitioning nuclear and chloroplast variation at multiple spatial scales in the Neotropical epiphytic orchid, *Laelia rubescens*. *Molecular Ecology* 13: 2655–2666.

Webb, L. J., and J. G. Tracey. 1981. Australian rainforests: Patterns and change. *In* A. Keast [ed.], Ecological biogeography of Australia, vol. 1, 605–694. W. Junk, The Hague, The Netherlands.

Appendix 1. Geographic locations and voucher information for *Dendrobium calamiforme* samples collected from five sites in Queensland, Australia, and deposited in the Australian Tropical Herbarium (CNS) by Ashley R. Field (ARF).

Geographic coordinates	Site description	CNS primary collector no.
16°40′52.6″S, 145°10′47.0″E	Font Hill Station in Baker and Blue Mountain Range	ARF4151
16°57′02.1″S, 145°44′36.8″E	Bruce Hwy. and Toogood Rd. intersection	ARF4154
17°06′00.3″S, 145°47′12.0″E	Bruce Hwy. crossing of Mulgrave River	ARF4155
16°52′29.7″S, 145°40′40.8″E	Barron River off of Stony Creek Rd.	ARF5152
16°54′08.6″S, 145°45′08.3″E	Centenary Lakes, on track to Flecker Botanical Garden	ARF4150