Transcriptome-Derived Microsatellite Markers for Dioon (Zamiaceae) Cycad Species

Authors: Prado, Alberto, Cervantes-Díaz, Fret, Perez-Zavala, Francisco G., González-Astorga, Jorge, Bede, Jacqueline C., et al.

Source: Applications in Plant Sciences, 4(2)

Published By: Botanical Society of America

URL: https://doi.org/10.3732/apps.1500087

Applications in Plant Sciences 2016 4(2): 1500087

Primer Note

Transcriptome-derived microsatellite markers for *Dioon* (Zamiaceae) cycad species[1]

Alberto Prado[2], Fret Cervantes-Díaz[3], Francisco G. Perez-Zavala[3], Jorge González-Astorga[4], Jacqueline C. Bede[2], and Angélica Cibrián-Jaramillo[3,5]

[2]Department of Plant Science, McGill University, 21,111 Lakeshore, Ste-Anne-de-Bellevue, Québec, Canada; [3]Laboratorio Nacional de Genómica para la Biodiversidad (Langebio), Unidad de Genómica Avanzada, Cinvestav, Km 9.6 Libramiento Norte Carr. Irapuato-León, 36821 Irapuato, Mexico; and [4]Laboratorio de Genética de Poblaciones, Red de Biología Evolutiva, Instituto de Ecología A.C., Carretera Antigua a Coatepec 351, El Haya, Xalapa 91070, Veracruz, Mexico

- *Premise of the study: Dioon* (Zamiaceae) is an endangered North American cycad genus of evolutionary and ornamental value. We designed and validated a set of microsatellite markers from *D. edule* that can be used for population-level and conservation studies, and that transferred successfully to *D. angustifolium*, *D. spinulosum*, and *D. holmgrenii.*
- *Methods and Results:* We tested 50 primers from 80 microsatellite candidate loci in the OneKP *D. edule* transcriptome. Genotypes from 21 loci in 20 *D. edule* individuals revealed up to 14 alleles per locus and observed heterozygosity from 0.15 to 0.92; one locus was monomorphic. Seven of those 21 loci were polymorphic in *D. angustifolium*, *D. spinulosum*, and *D. holmgrenii*, with up to seven alleles, and an observed heterozygosity up to 0.89.
- *Conclusions:* The transcriptome-derived microsatellites generated here will serve as tools to advance population genetic studies and inform conservation strategies of *Dioon*, including the identification and origin of illegal plants in the cycad trade.

Key words: cycads; *Dioon edule*; microsatellites; transcriptome; Zamiaceae.

Of Permian origin, cycads (Cycadales) are dioecious gymnosperms distributed in tropical and subtropical regions (Norstog and Nicholls, 1997). *Dioon* Lindl. is composed of 14 species, 13 of which are distributed in Mexico and one in Honduras (Osborne et al., 2008). *Dioon* is of economic importance as an ornamental plant, as well as an alternative food source to main crops, and it has cultural value throughout its distribution area (Bonta and Osborne, 2008). *Dioon edule* Lindl. is a medium-sized cycad with an erect trunk of 1.5 m and a rigid crown of 15–25 long, blue-green leaves that is endemic to eastern Mexico, growing mainly in tropical deciduous thorn forests and oak forests (Octavio-Aguilar et al., 2008). They can live up to 2000 yr and have slow growth rates with long reproductive cycles (Vovides, 1990). Despite legal protection in the Norma Oficial Mexicana of the Secretaría de Medio Ambiente y Recursos Naturales (Nom-059 SEMARNAT-2010) and the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES; 2008 Appendix II), poaching of *Dioon* species is still common and wild populations are disappearing.

Allozyme variation in *D. edule* has shown that there is unusually high genetic diversity in this genus (González-Astorga et al., 2003; Cabrera-Toledo et al., 2010, 2012), but these markers are limited in their scale of analysis and in their use for conservation, as they lack individual-level resolution and are dominant markers. Independent laboratories (J. C. Bede, A. Cibrián-Jaramillo, D. Cabrera-Toledo, and L. Yañez-Espinosa, personal communication) have been unable to replicate microsatellites previously developed for *D. edule* (Moynihan et al., 2007). Therefore, there is a need to develop robust genetic markers to understand population genetic history and to inform conservation strategies in *Dioon*.

METHODS AND RESULTS

Genomic DNA was isolated from 20 randomly chosen samples out of 40 *D. edule* individuals, representing four populations found in the states of San Luis Potosí and Veracruz, Mexico (Appendix 1). Leaflets were ground in liquid nitrogen and sieved through a fine 0.5-mm mesh to remove cuticle and fiber particles. DNA was extracted using the DNeasy Plant Mini Kit (QIAGEN, Valencia, California, USA) following the manufacturer's protocols after adjustment for the extraction of 200 mg of tissue. The *D. edule* transcriptome was accessed through the OneKP project (www.onekp.com), and a total of 121,771 contigs were analyzed for tandem repeats using the algorithm mreps version 2.5 (Kolpakov and Kucherov, 1999) available at the Mobyle Portal (http://mobyle.pasteur.fr/cgi-bin/portal.py?#forms::mreps). We targeted five or more tandem nucleotide repeats prioritizing di-, tri-, and tetranucleotide repeats with adjacent 5′ and 3′ 15–30-nucleotide sequences for primer design. Primers were designed with Primer3Plus (Untergasser et al., 2007), and self-annealing and heterodimer formation was tested with OligoAnalyzer version 3.1 (Integrated DNA Technologies, Coralville, Iowa, USA). Eighty of 150 microsatellites identified in the *D. edule* transcriptome had candidate primer sites. We designed primers for 80 loci and tested 50 of them in three randomly selected *D. edule* samples using a PCR

[1]Manuscript received 31 July 2016; revision accepted 28 September 2015.

The authors thank Dr. Laura Yañez-Espinosa (Instituto de Investigaciones de Zonas Desérticas [IIZD]) for her advice, the plant collectors who contributed samples to the OneKP initiative, as well as Dennis W. Stevenson and Gane Ka-Shu Wong for facilitating access to the *Dioon edule* transcriptome. Funding for this research was provided by the Consejo Nacional de Ciencia y Tecnología (CONACyT) (to A.P., and no. 169701 to A.C.J.), the Natural Sciences and Engineering Research Council of Canada (J.C.B.), and Instituto de Ecología A.C. (2003/10776 to J.G.A.).

[5]Author for correspondence: acibrian@langebio.cinvestav.mx

doi:10.3732/apps.1500087

Table 1. Characteristics of 21 microsatellite loci developed for *Dioon edule*.

Locus	Primer sequences (5′–3′)	Repeat motif	T_a (°C)	GenBank accession no.	Organism/Putative function/Accession no.[a]	*E*-value
2001304[b]	F: CTGGGCCTCGACATAACATT R: TCAAAATCATTCCGGCTTTC	$(ATAA)_5$	60	KT289910	*Picea sitchensis* EST BT123050	0
2001597[b]	F: CTTACAAGCGGCACCATTG R: AAGCAGGCCAGACTTCAGAC	$(TA)_{12}$	60	KT289911	None	—
2001955[b]	F: CTGCCGAGGAGGGACA R: CGCAGGGTTGGAGAGC	$(TGC)_8$	58	KT289912	*Pinus monticola* WRKY TF EU269755	3.00E-41
2002082[b]	F: TGACCTTGCCTTAGGTCAAAA R: AGATGTGGGTGACACGTCCT	$(GT)_{14}$	60	KT289913	None	—
2002349[b]	F: AGAGCTGCTTCCACGTTCAT R: GCGGAAACTTCTTCAACAGC	$(TCT)_6$	60	KT289914	None	—
2002757[b]	F: TGGGAAATGCACACCTAAAA R: ACCTGGGCCACTTGAGG	$(CA)_{20}$	58	KT289915	None	—
2003643[b]	F: CGAACTTGAAGACGATGACG R: CGGGTAGCACCAAAGATTGT	$(GCA)_7$	60	KT289916	None	—
2014158[b]	F: CACCGTGCCCGTCATT R: GCTGGCCCTGCAAAGA	$(TTC)_9$	60	KT289917	None	—
2016409[b]	F: AGCGCCGCTGTCATTC R: TTCGGCTGTGCATCTCAA	$(ATT)_9$	59	KT289918	None	—
2016799[b]	F: GAAGGGTGGGTATGGCACT R: TCCTGCCTGCAAAGCAC	$(GA)_{14}$	59	KT289919	None	—
2011473[b]	F: CATAGGAGGCCACCATGTAGA R: GCAAGTGCTTTGGATATGCTC	$(TG)_{15}$	60	KT289920	None	—
2017825[b]	F: GGGGCTGCATCCTTCC R: GATGGGGCAATGGGAAT	$(GA)_{10}$	60	KT289921	*Picea glauca* EST BT107925	0
2015001[b]	F: TTGTTTGGGCAATTCCTGA R: CAAAGCCAACAAAGACATGA	$(TG)_{10}$	58	KT289923	*Picea glauca* EST BT116353	0
2015907[b]	F: TGCCATCTTCCCATCACA R: CCCAGGCCCCAAGATT	$(AT)_{11}$	59	KT289924	*Oryza brachyantha* heat shock protein XM_006663065	0
2014311[c,e]	F: TTAGGGGGCCGAGGAG R: CCGACCGCCAAGAGAA	$(CCA)_6$	58	KT261421	*Zea mays* Predicted formin-like protein XM_00867779	1.00E-44
2015232[c,d,e]	F: TGGAGATCAACAACGACCAA R: CGAGACCCAGAGARCCTGAC	$(CAT)_6$	58	KT261422	None	—
2018276[c,e]	F: CACCCTGCCAAAGGTCAT R: GCACCCATTGTTGGACA	$(TG)_{14}$	58	KT261423	*Picea glaucea* EST BT106585	0
2019766[c,e]	F: CGTGCGACCAGCAAGA R: TCGCCGGTGAGAACAAG	$(TCG)_7$	58/55	KT261424	None	—
2018116[c,e]	F: GGCAGATTAGCTCCAGCAG R: CAGAGCTTCCAATTCCTTGC	$(GAC)_8$	58/55	KT261425	*Picea glauca* EST BT117927	9.00E-57
2018893[c,e]	F: GCGGTAGCTGGAGAGGTTC R: AGTCTGGGGCCTCATCAAC	$(GAA)_{10}$	58/55	KT261426	None	—
2019765[c,e]	F: CAAATTCCTGTGGGAGATGG R: GCAGGCAGTTTGGAAAGAAC	$(AGC)_8$	58/55	KT261427	*Picea sitchensis* EST WS02775_I20	7.00E-157

Note: T_a = annealing temperature in *D. edule*/other *Dioon* species.

[a] Search made with the online version of TBLASTX versus the nr/nt database with default parameters at NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=tblastx&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome).

[b] Forward primer contains a 19-bp M13(–29) infrared dye extension at the 5′ end.

[c] Forward primer contains an 18-bp M13(–21) fluorescent dye extension at the 5′ end.

[d] Monomorphic in *D. edule* but polymorphic in congeners.

[e] Loci transferable to *D. angustifolium*, *D. holmgrenii*, and *D. spinulosum*.

protocol as shown for the M13(–21) fluorescent label. We chose 21 loci that produced bands consistently as evidenced by an agarose gel (Table 1). The issues with the remaining loci were double or multiple bands or lack of amplification. Forward primers for *D. edule* contained a 5′ extension of M13 following the protocol of Schuelke (2000). We used an infrared dye–labeled (LI-COR Biosciences, Lincoln, Nebraska, USA) M13(–29) sequence (CACGACGTTGTAAAACGAC) and a 6-FAM fluorescently labeled M13(–21) sequence (TGTAAAACGACGGCCAGT) (Sigma-Aldrich, St. Louis, Missouri, USA) (Table 1). We used 6-FAM to genotype other congeners: 10 individuals for *D. angustifolium* Miq. from one population, and 40 each for *D. holmgrenii* De Luca, Sabato & Vázq. Torres and *D. spinulosum* Dyer ex Eichler, as based on the availability of these individuals in the field. These three species are representatives of separate phylogenetic clades, with *D. spinulosum* being sister to the rest of *Dioon*, and *D. angustifolium* and *D. edule* sister to *D. holmgrenii*, according to González et al. (2008). The PCR mixture contained a minimum of 40 ng/μL of DNA template, 200 μM deoxynucleotides (New England Biolabs, Ipswich, Massachusetts, USA), 0.25 units of *Taq* DNA polymerase (New England Biolabs), 0.08 μM of forward primer with the 5′-M13 tail, and 0.2 μM reverse primer; we added 0.05 μM infrared dye–labeled primer for M13(–29) and 0.2 μM for the fluorescently labeled M13(–21) primer, plus 1 μL of 10× PCR buffer (Sambrook et al., 1989). PCR conditions for M13(–29) were: 94°C for 3 min; 16 cycles at 94°C for 30 s, 58–60°C for 1 min, and 72°C for 30 s; 10 cycles at 94°C for 30 s, 55°C for 1 min, and 72°C for 30 s; and 72°C for 2 min. PCR conditions for M13(–21) were: 95°C for 2 min; followed by 30 cycles at 95°C for 30 s, 55–58°C for 30 s, and 72°C for 1 min; then eight cycles at 95°C for 30 s, 53°C for 30 s, and 72°C for 1 min; and 70°C for 10 min. All protocols are modifications of Schuelke (2000). The annealing temperature for specific primer pairs is shown in Table 1. We included information on protein sequence matches to our loci according to a TBLASTX search against the nucleotide collection (nr/nt) with default parameters at the National Center for Biotechnology Information (NCBI). One of the 21 loci was monomorphic (locus 2015232), and 20 were polymorphic in 20 randomly chosen individuals from four *D. edule* localities (Table 2). Seven of those 21 loci, including the monomorphic locus 2015232, were consistently polymorphic in *D. angustifolium*, *D. spinulosum*, and *D. holmgrenii* (Table 3). The remaining 14 loci would require additional optimization to remove stutter bands. Amplicons for *D. edule* were separated on a 6.5% acrylamide gel on a NEN 4300 DNA Analyzer (LI-COR Biosciences) and compared to the LI-COR size standard 4200-44 (50–350 bp)

Table 2. Genetic diversity in 21 microsatellite loci developed in *Dioon edule*.[a]

Locus	Allele size range (bp)	A	H_o	H_e
2001304	183–203	4	0.619	0.604
2001597	197–211	6	0.6	0.713
2001955	202–214	4	0.35	0.411
2002082	240–256	7	0.15	0.798
2002349	368–386	7	0.75	0.725
2002757	256–312	14	0.6	0.879
2003643	225–246	6	0.65	0.703
2014158	164–176	3	0.667	0.647
2016409	237–270	7	0.6	0.822
2016799	250–272	11	0.688	0.877
2011473	237–269	8	0.8	0.825
2017825	248–272	9	0.867	0.838
2015001	194–200	5	0.923	0.754
2015907	211–223	5	0.733	0.638
2014311[b]	243–255	4	0.666	0.722
2015232[b,c]	249	1	—	—
2018276[b]	252–278	6	1	0.781
2019766[b]	231–252	3	0.571	0.540
2018116[b]	189–204	5	0.666	0.694
2018893[b]	255–261	3	0.666	0.611
2019765[b]	167–179	5	0.714	0.704

Note: A = number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity.

[a] Tested in 20 randomly chosen individuals from four *D. edule* localities; see Appendix 1.

[b] Transferred to three other species.

[c] Monomophic in *D. edule* but polymorphic in other species.

Table 3. Genetic diversity of seven microsatellite loci in related *Dioon* species.

Locus	*D. angustifolium* (n = 10)				*D. holmgrenii* (n = 40)				*D. spinulosum* (n = 40)			
	Allele size range (bp)	A	H_o	H_e	Allele size range (bp)	A	H_o	H_e	Allele size range (bp)	A	H_o	H_e
2014311	246–252	4	0.75	0.656	240–252	4	0.333	0.597	240–243	2	0	0.5
2015232[a]	249–252	2	0	0.244	240–249	2	0.642	0.497	240–258	5	0.466	0.62
2018276	268–274	3	0.7	0.505	266–280	5	0.461	0.636	260–283	6	0.5	0.57
2019766	246	1	—	—	231–246	2	0.833	0.503	231–246	2	0.772	0.447
2018116	201–210	4	0.5	0.563	192–204	4	0.894	0.68	204–213	4	0.5	0.546
2018893	255–261	3	0.666	0.5	255	1	—	—	258	1	—	—
2019765	170–179	3	0.5	0.402	155–179	7	0.896	0.717	161–179	5	0.476	0.399

Note: A = number of alleles; H_e = expected heterozygosity; H_o = observed heterozygosity; n = number of individuals sampled.

[a] Locus 2015232 was monomorphic in *D. edule*.

(LI-COR Biosciences). Bands were scored using SAGA[GT] version 3.3 software (LI-COR Biosciences). The seven loci tested in *Dioon* congeners were genotyped in a separate run with an ABI 3730xl sequencer with GeneScan 500 LIZ size standard (Applied Biosystems/Thermo Fisher Scientific, Waltham, Massachusetts, USA) and interpreted with Geneious version 8.1.3 software (Biomatters, http://www.geneious.com/). Scoring errors that may result from stuttering, large allele drop out, or null alleles were identified using MICRO-CHECKER version 2.2.3 (van Oosterhout et al., 2004). Observed (H_o) and expected (H_e) heterozygosities were calculated using the R package Adegenet (Jombart, 2008). No evidence of scoring errors due to peak stuttering or large allele dropout was observed. H_o and H_e of the 21 microsatellite markers in *D. edule* ranged from 0.15 to 0.92 and 0.41 to 0.87, respectively (Table 2). Loci 2002082 and 2002757 had a homozygote excess that was not evenly distributed across all homozygote classes, which could be indicative of null alleles (van Oosterhout et al., 2004). PCR amplification with lower temperatures (52–56°C) did not recover any additional alleles for these loci. The observed heterozygosity confirms previous allozyme studies and is congruent with *Dioon*'s mating system (González-Astorga et al., 2003; Cabrera-Toledo et al., 2010). The number of alleles for the transferred loci ranged from one to seven, and H_o and H_e ranged from 0.33 to 0.89 and 0.24 to 0.71, respectively (Table 3), which suggests variability in other species. Vouchers were deposited at the Jardín Botánico Francisco Javier Clavijero in Xalapa, Veracruz, Mexico, and at the McGill University Herbarium (MTMG), Québec, Canada.

CONCLUSIONS

We identified and validated 21 new microsatellite loci in *D. edule*, one being monomorphic for this species. Seven of these 21 markers are polymorphic in the congeners *D. angustifolium*, *D. holmgrenii*, and *D. spinulosum*, which are representatives of divergent phylogenetic clades. This suggests that these markers are likely transferable to additional *Dioon* species. Our loci are useful for *Dioon* population genetics and have great potential to be used in in situ and ex situ conservation strategies, including as a means to help authorities identify the origins of illegal plant material.

LITERATURE CITED

Bonta, M., and R. Osborne. 2008. Cycads in the vernacular: A compendium of local names (Nombres vernaculares de Cycadales—Un compendio). *In* A. Vovides, D. W. Stevenson, and R. Osborne [eds.], Proceedings of CYCAD 2005, 7th International Conference on Cycad Biology. Memoirs of the New York Botanical Garden, 147–175. The New York Botanical Garden Press, Bronx, New York, USA.

Cabrera-Toledo, D., J. González-Astorga, F. Nicolalde-Morejón, F. Vergara-Silva, and A. P. Vovides. 2010. Allozyme diversity levels in two congeneric *Dioon* spp. (Zamiaceae, Cycadales) with contrasting rarities. *Plant Systematics and Evolution* 290: 115–125.

Cabrera-Toledo, D., J. González-Astorga, and J. C. Flores-Vázquez. 2012. Fine-scale spatial genetic structure in two Mexican cycad species *Dioon caputoi* and *Dioon merolae* (Zamiaceae, Cycadales): Implications for conservation. *Biochemical Systematics and Ecology* 40: 43–48.

González, D., A. P. Vovides, and C. Barcenas. 2008. Phylogenetic relationships of the neotropical genus *Dioon* (Cycadales, Zamiaceae) based on nuclear and chloroplast DNA sequence data. *Systematic Botany* 33: 229–236.

González-Astorga, J., A. P. Vovides, M. M. Ferrer, and C. Iglesias. 2003. Population genetics of *Dioon edule* Lindl. (Zamiaceae, Cycadales): Biogeographical and evolutionary implications. *Biological Journal of the Linnean Society* 80: 457–467.

Jombart, T. 2008. Adegenet: An R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403–1405.

Kolpakov, R., and G. Kucherov. 1999. Finding maximal repetitions in a word in linear time. Proceedings of the 40th Annual Symposium on Foundations of Computer Science (FOCS), 596–604. IEEE Computer Society, New York, New York, USA.

Moynihan, J., A. W. Meerow, and J. Francisco-Ortega. 2007. Isolation, characterization and cross-species amplification of microsatellite loci in the cycad genus *Dioon* (Zamiaceae): Potential utilization in population genetics studies of *Dioon edule*. *Molecular Ecology Notes* 7: 72–74.

Norstog, K. J., and T. J. Nicholls. 1997. The biology of the cycads. Cornell University Press, Ithaca, New York, USA.

Octavio-Aguilar, P., J. González-Astorga, and A. P. Vovides. 2008. Population dynamics of the Mexican cycad *Dioon edule* Lindl. (Zamiaceae): Life history stages and management impact. *Botanical Journal of the Linnean Society* 157: 381–391.

Osborne, R., M. A. Calonje, K. D. Hill, L. Stanberg, and D. W. Stevenson. 2008. The world list of cycads. Proceedings of the 8th International Conference on Cycad Biology, CYCAD 2008, Panama City, Panama.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning: A laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.

Schuelke, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233–234.

Untergasser, A., H. Nijveen, X. Rao, T. Bisseling, and R. Geurts. 2007. Primer3Plus, an enhanced Web interface to Primer3. *Nucleic Acids Research* 35: W71–W74.

van Oosterhout, C., W. F. Hutchinson, D. P. Wills, and P. Shipley. 2004. MICRO-CHECKER: Software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4: 535–538.

Vovides, A. P. 1990. Spatial distribution, survival, and fecundity of *Dioon edule* (Zamiaceae) in a tropical deciduous forest in Veracruz, Mexico, with notes on its habitat. *American Journal of Botany* 77: 1532–1543.

Appendix 1. Voucher information for *Dioon* species used in this study.

Species	Voucher no.[a]	Collection locality[b]	*N*
D. edule	2001-114	Actopan, Veracruz	10
D. edule	2002-013	Cd. Valles, San Luis Potosí	10
D. edule	2011-AGA21 (MTMG)	Agua de Gamotes, San Luis Potosí	10
D. edule	2011-GRAP43 (MTMG)	Los Pocitos, San Luis Potosí	10
D. angustifolium	JGA 10-2003	Aldama, Tamaulipas	10
D. holmgrenii	JGA 15-2005	Rancho El Limon, Oaxaca	20
D. holmgrenii	JGA 16-2005	San Bartolome Loxicha, Oaxaca	20
D. spinulosum	JGA 23-2007	Soyaltepec, Oaxaca	20
D. spinulosum	JGA 24-2007	Cerro Bola, Oaxaca	20

Note: N = number of individuals.

[a]Vouchers are deposited at the Jardín Botánico Francisco Javier Clavijero in Xalapa, Veracruz, Mexico, and at the McGill University Herbarium (MTMG), Québec, Canada; JGA = Jorge González-Astorga.

[b]Locality in Mexico; exact coordinates are not listed to protect endangered populations but are available upon request.