

**ANONYMOUS AND EST-BASED MICROSATELLITE
DNA MARKERS THAT TRANSFER BROADLY ACROSS THE
FIG TREE GENUS (*FICUS*, MORACEAE)¹**

KATRIN HEER^{2,7}, CARLOS A. MACHADO³, ANNA G. HIMLER⁴, EDWARD ALLEN HERRE⁵,
ELISABETH K. V. KALKO^{2,5,8}, AND CHRISTOPHER W. DICK^{5,6}

²Institute of Experimental Ecology, University of Ulm, Ulm, Germany; ³Department of Biology, University of Maryland, College Park, Maryland 20742 USA; ⁴Department of Entomology, University of Arizona, Tucson, Arizona 85721 USA; ⁵Smithsonian Tropical Research Institute, Balboa, Panama; and ⁶Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109 USA

- *Premise of the study:* We developed a set of microsatellite markers for broad utility across the species-rich pantropical tree genus *Ficus* (fig trees). The markers were developed to study population structure, hybridization, and gene flow in neotropical species.
- *Methods and Results:* We developed seven novel primer sets from expressed sequence tag (EST) libraries of *F. citrifolia* and *F. popenoei* (subgen. *Urostigma* sect. *Americana*) and optimized five previously developed anonymous loci for cross-species amplification. The markers were successfully tested on four species from the basal subgenus *Pharmacosycea* sect. *Pharmacosycea* (*F. insipida*, *F. maxima*, *F. tonduzii*, and *F. yoponensis*) and seven species of the derived subgenus *Urostigma* (*F. citrifolia*, *F. colubrinae*, *F. costaricana*, *F. nymphaeifolia*, *F. obtusifolia*, *F. pertusa*, and *F. popenoei*). The 12 markers amplified consistently and displayed polymorphism in all the species.
- *Conclusions:* This set of microsatellite markers is transferable across the phylogenetic breadth of *Ficus*, and should therefore be useful for studies of population structure and gene flow in approximately 750 fig species worldwide.

Key words: *Ficus insipida*; hybridization; Moraceae; spatial genetic structure; tropical forests.

Worldwide, there are more than 750 species of *Ficus* L., which are classified into four subgenera and 20 sections (Rønsted et al., 2005). Fig trees are well known for the complex mutualism with the species-specific pollinators, the fig wasps (Agaonidae) (Herre et al., 2008). They are also considered keystone species in many tropical forests because of their year-round production of calcium-rich fruits, which are consumed by many vertebrates such as birds, bats, and monkeys. Despite the broad literature on the ecology and evolutionary biology of figs, there are still many unresolved questions regarding their complex interactions with pollen and seed dispersers. For example, despite the tight coevolutionary relationship between figs and their pollinating wasps, there is some evidence for hybridization and introgression that needs to be investigated in depth (Herre et al., 2008). Furthermore, despite numerous studies of the mechanisms of seed dispersal, fig seed

dispersal distances are mostly unknown. Highly variable molecular markers such as microsatellites may help to further our understanding of these aspects of *Ficus* biology.

Neotropical figs belong to two distantly related subgenera, *Pharmacosycea* (Miq.) Miq. and *Urostigma* (Gasp.) Miq. While the former comprises the so-called free-standing figs and is the most basal subgenus, the latter consists of strangler figs belonging to the section *Americana* Miq. These two subgenera are thought to share a common ancestor at least 60 Ma and represent the greatest phylogenetic distance within *Ficus* (Rønsted et al., 2005). So far, most microsatellites for *Ficus* have been developed for Asian and African figs that belong to subgenera that do not occur in the neotropics. Some of these markers were shown to cross-amplify with the neotropical strangler fig *F. citrifolia* Mill. (subgen. *Urostigma*) and were highly variable (Crozier et al., 2007; Nazareno et al., 2009). Vignes et al. (2006) published primers for the common neotropical species *F. insipida* Willd. (subgen. *Pharmacosycea*), but most of these loci presented few alleles. Here we present an additional set of microsatellite markers that transfer reliably across species within the strangler and free-standing figs, and should therefore amplify in all 750 fig species that share the most recent common ancestor of the subgenera *Urostigma* and *Pharmacosycea*.

¹Manuscript received 23 January 2012; revision accepted 20 March 2012.

The German National Academic Foundation provided financial support to K.H. C.W.D. acknowledges support from the National Science Foundation (DEB 0640379). We are grateful to Autoridad Nacional del Ambiente (ANAM) (Permit No. SC/AP-2-10 and SEX/AP-8-10) and Ministerio del Ambiente, Energía y Telecomunicaciones (MINET) (Permit No. 047-2010_SINAC) for granting research and export permits. The authors thank Juanita Choo, Ashley Thomson, and Wei Na for helpful advice with the laboratory work.

⁷Author for correspondence: katrin.heer@uni-ulm.de

⁸Deceased (1962–2011).

doi:10.3732/ajb.1200032

METHODS AND RESULTS

Leaf samples of eleven *Ficus* species were collected in the Barro Colorado Natural Monument (9°10'N, 79°51'W) in Panama and in La Tirimbina,

TABLE 1. Characteristics of 12 microsatellite markers for neotropical *Ficus* species.^a

Primer	Primer sequences (5'–3')	Repeat motif	T _a (°C)	Best match (BLAST)	GenBank accession no. or reference
FC14	F: CTCATCCCTTGCTTACCTTA R: CCAAATTGCACTTGAAATAA	(TC) ₁₁	54	XP_002304206.1 hypothetical protein POPTRDRAFT_712066 [<i>Populus trichocarpa</i>]. E value: 2e-70	JK730349
FC22	F: GATTTCAGAGGTCATTCCAA R: CAAACTACATGGATCAAGCA	(TA) ₈	54	XP_002285437.1 protein PROTON GRADIENT REGULATION 5, chloroplastic isoform 1 [<i>Vitis vinifera</i>]. E value: 6e-71	JK730350
FC27	F: CTGGTCATGTGGGAAGTAGT R: ATAAATGTGGAAGGCTCAA	(AG) ₁₀	55	ACE00235.1 cytoplasmic male sterility protein 60 [<i>Brassica oleracea</i> var. <i>capitata</i>]. E value: 9e-07	JK730351
FP21	F: AGACGAACCAGAAGACGTTA R: ATATGAACCAGCTAGGCGCT	(CA) ₂₁	55	XP_003603706.1 hypothetical protein MTR_3g111360 [<i>Medicago truncatula</i>]. E value: 4e-20	JK730352
FP22	F: AGAATGGACTTTGAAGCTGA R: CGAAATAGGAGACGAAGTTG	(AAG) ₁₀	55	XP_002300690.1 predicted protein [<i>Populus trichocarpa</i>]. E value: 4e-105	JK730353
FP25	F: GAAGCGTCACATTTAACTCC R: ACAAATTCGATGCATGAC	(TC) ₁₂	54	XP_002530128.1 conserved hypothetical protein [<i>Ricinus communis</i>]. E value: 2e-29	JK730354
FP57	F: AAATATCATCACTGCCTCAA R: GAAACGACGTAGTAAATGGC	(TC) ₈	54	XP_002510808.1 steroid-binding protein, putative [<i>Ricinus communis</i>]. E value: 6e-98	JK730355
FinsA1	F: AATCCCCGTACTTCACTTG R: AGAACTTATTGCACGGACAG	(CT) ₁₂	55	anonymous region	Vignes et al., 2006
FinsH5	F: GACCGTATAGATGATTTGGG R: CATCCTGTGAACGACACTT	(AT) ₅ GTAT(GT) ₁₁	54	anonymous region	Vignes et al., 2006
FinsI12	F: GAACCTTCAACCTCAATCAA R: CTCCCCTTTCCTAGTCCTTA	(TC) ₅ (CT) ₁₁	55	anonymous region	Vignes et al., 2006
Frub29	F: CCACTTTGGAATGTCACCTGGGA R: TGAACACGCCCACTGAGAATG	(AG) ₂₄	49	anonymous region	Crozier et al., 2007
Frub38	F: ACACGTGCAGTGTCTGCTGA R: ACAGCTGCCCAATTCCTTGA	(AG) ₈ AAC(GA) ₁₃	49	anonymous region	Crozier et al., 2007

Note: T_a = annealing temperature.

^a For markers developed from the EST library, the best match of the BLAST search and the GenBank accession number are indicated. Alternatively, the reference of the marker is given. Markers labeled FC were derived from the EST library of *F. citrifolia*, and FP from *F. popenoei*.

Sarapiqui, Costa Rica (10°24'N, 84°06'W; Appendix 1). Seven of the sampled species, namely *F. citrifolia*, *F. colubrinae* Standl., *F. costaricana* (Liebm.) Miq., *F. nymphaeifolia* Mill., *F. obtusifolia* Kunth, *F. pertusa* L.f., and *F. popenoei* Standl., belong to the subgenus *Urostigma* (sect. *Americana*), while the other four species (*F. insipida*, *F. maxima* Mill., *F. tonduzii* Standl., and *F. yoponensis* Desv.) are classified into the subgenus *Pharmacosycea*. Sampled leaves were dried in silica gel immediately upon collection.

To obtain a set of highly variable markers for our study species, we developed markers from expressed sequence tag (EST) libraries of *F. citrifolia* and

F. popenoei. RNA was extracted from leaves of *F. citrifolia* and *F. popenoei* using the Plant RNeasy kit (QIAGEN, Hilden, Germany), and mRNA was isolated using the Oligotex kit (QIAGEN). cDNA libraries for each species were constructed using the SMART cDNA library construction kit (Clontech, Mountain View, California, USA). First-strand synthesis was done following the kit's protocol with the Clontech Reverse Transcriptase. PCR amplification of first-strand cDNA was done with the Platinum-Pfx DNA polymerase (Invitrogen #11708-013, Carlsbad, California, USA). The protocol was the same as in the SMART cDNA kit except that it used Platinum-Pfx DNA polymerase and its

TABLE 2. Characterization of the novel and published markers with samples of four *Ficus* species collected in BCNM, Panama.

Marker ^a	<i>F. insipida</i>				<i>F. yoponensis</i>				<i>F. citrifolia</i>				<i>F. obtusifolia</i>			
	N	Size (bp) ^b	A	H _o ^c	N	Size (bp) ^b	A	H _o ^c	N	Size (bp) ^b	A	H _o ^c	N	Size (bp) ^b	A	H _o ^c
FC14	30	289–317	8	0.83	30	287–301	4	0.43	30	300–314	7	0.66	30	283–307	4	0.60
FC22	30	140–180	15	0.93	30	143–164	9	0.73	30	155–180	10	0.73	30	151–180	11	0.73
FC27	24	170	1	—	15	170	1	—	22	166–174	3	0.36	30	164–174	4	0.47*
FP21	30	273–290	4	0.43	30	271–296	8	0.73	30	275–294	7	0.87	24	284	1	—
FP22	30	217–241	7	0.53	30	214–235	7	0.63	30	214–232	6	0.77	30	217–235	5	0.57
FP25	24	146–150	4	0.29*	13	150–161	8	0.77*	30	158–174	6	0.63	—	NA	0	—
FP57	24	200–207	5	0.12	16	203–208	6	0.38	30	211–236	18	0.80	30	216–239	16	0.70**
FinsA1	30	226–244	6	0.73	30	231–247	12	0.97	30	224–237	7	0.70	30	224–241	4	0.47
FinsH5	30	283–295	5	0.30	30	280–299	10	0.83	30	274–319	13	0.73**	30	281–330	18	0.20***
FinsI12	30	167–173	3	0.10	30	167–171	3	0.33	22	168–169	2	0.41	30	168–177	5	0.77
Frub29	30	221–245	5	0.63	30	221–241	5	0.13	30	208–239	12	0.70	24	212	1	—
Frub38	30	197–238	5	0.80	30	221–232	4	0.80	30	207–236	9	0.90	30	219–238	4	0.50

Note: — = not available; A = number of alleles; H_o = observed heterozygosity; N = number of individuals tested; NA = no amplification.

^a Only markers used in the final study were tested with 30 individuals, while other markers have been tested with lower sample size.

^b Fragment size includes 18 bp of the M13 adapter tail.

^c Deviations from Hardy–Weinberg equilibrium are indicated by asterisks (except for species with sample size <10). Levels of significance: *P < 0.05, **P < 0.01, ***P < 0.001.

TABLE 3. Characterization of the novel and published markers from seven sampled species of *Ficus*.

Locus	subgen. <i>Urostigma</i>						subgen. <i>Pharmacosyceae</i>							
	<i>F. costaricana</i> (n = 6)		<i>F. colubrinae</i> (n = 14) ^a		<i>F. nymphaeifolia</i> (n = 22)		<i>F. pertusa</i> (n = 4) ^b		<i>F. popenoei</i> (n = 22)		<i>F. maxima</i> (n = 22)		<i>F. tonduzii</i> (n = 20) ^c	
	A	H _o ^b	A	H _o ^b	A	H _o ^b	A	H _o ^b	A	H _o ^b	A	H _o ^b	A	H _o ^b
FC14	6	0.8	9	0.76	8	0.82	4	0.67	2	0.33	5	0.68	5	0.5
FC22	4	0.33	7	0.69	4	0.67	3	0.5	3	0.71	5	0.5	6	0.41*
FC27	3	0.4	3	0.36	5	0.71	5	0.75	5	0.76	3	0.1	4	0.25
FP21	5	0.5	3	0.33	9	0.82	4	1	10	0.71	4	0.43	2	0.11
FP22	4	0.5	4	0.62	4	0.5	4	0.75	5	0.45	5	0.33	9	0.8
FinsA1	5	0.17	0	—	4	0.64	4	0.25	3	0.57	9	0.82	3	0.26
FinsH5	5	0.33	7	0.33***	10	0.41***	3	0.5	6	0.33***	5	0.82	6	0.85
FinsI12	3	0.5	3	0.5	2	0.05	2	0	3	0.13*	4	0.55	1	—
FruB29	4	0.6	6	0.31*	10	0.3***	4	1	1	—	5	0.52	3	0.25
FruB38	3	0.17	3	0.43	6	1.00**	6	1	2	0.09	3	0.14	1	—

Note: — = not available; A = number of alleles; H_o = observed heterozygosity; NA = no amplification.

^aFragment size includes 18 bp of the M13 adapter tail.

^bDeviations from Hardy-Weinberg equilibrium are indicated by asterisks (except for species with sample size <10). Levels of significance: *P < 0.05, **P < 0.01, ***P < 0.001.

^cDenotes samples of species collected in La Tirimbina, Costa Rica; all other species were collected in BCNM, Panama.

corresponding buffers. The PCR cycle included an initial step of 2 min at 95°C followed by 25–30 cycles of 5 s at 95°C and 6 min at 68°C.

Amplified cDNAs were size selected after the second-strand synthesis to include inserts >500 bp as described in the SMART cDNA kit. Final fractions were precipitated as described in the kit but resuspended in 15 µL of water. Fragments were ligated to the PCR-Blunt II-TOPO vector (from the Zero Blunt TOPO PCR cloning kit; Invitrogen K2800-20), and transformation was done into TOP10 cells, ensuring that almost 100% of picked clones had an insert. Best ligation efficiency was achieved with a vector:insert ratio of 1:1. We used a 30-min ligation step (instead of 25 min) and electroporated TOP10 electrocompetent cells using a standard electroporation protocol (0.1 cm cuvettes and 1.8 Kv pulse). We checked for presence of inserts in a random selection of 20 clones per library by PCR using the following primers: 5' SMART PCR: 5'-AAGCAGTGGTATCAACGCAGAGT-3', 3' AMPLIMER 1: 5'-AGGCCGCCGACATGTTTTTTTTTTTTT-3'.

Libraries were plated and up to 10000 colonies per library were picked and arrayed in 384-well plates using a colony-picking robot (QBot; Genetix, New Milton, Hampshire, United Kingdom). An average of 800 clones were sequenced per library at the Arizona Genomics Institute (AGI), University of Arizona. Sequencing was done unidirectionally using the 5' SMART PCR primer. EST sequences were BLASTed to determine gene identity and level of conservation. High quality sequence information was obtained for 815 different fig genes (57 genes were common between the two fig libraries). The simple sequence repeat (SSR) primer pipeline (Jewell et al., 2006) was used for surveying ESTs for microsatellites and designing primers for amplification. The pipeline first parses sequences to SPUTNIK (<http://espressoftware.com/sputnik/index.html>) to identify nucleotide repeats of length 2–5 on each EST. We used default parameters for SSR identification. The output of SPUTNIK is then parsed to Primer3 (Rozen and Skaletsky, 1999) for PCR primer design using a set of default parameters: primer pairs (21–23 bases long) are designed at least 10 bp distant from either side of the identified SSR, with optimum melting temperature (T_m) of 55°C and maximum GC content of 70%.

In addition to the EST-based microsatellites, we selected 15 published microsatellite primer pairs derived from anonymous nuclear genomic regions to test for cross-species amplification, including nine markers developed by Crozier et al. (2007), which proved to be variable in *F. citrifolia* (Nazareno et al., 2009) and the six most variable markers that were developed for *F. insipida* (Vignes et al., 2006).

DNA was extracted with a modified cetyltrimethylammonium bromide (CTAB) protocol (Doyle and Doyle, 1987). We tested whether the markers amplified reliably with DNA of four samples each of *F. insipida*, *F. citrifolia*, and *F. obtusifolia*. We ran PCRs of 10 µL containing 1× PCR buffer, 0.2 mM dNTPs, 2.5 mM MgCl₂, 0.5 U *Taq* polymerase, 0.1 µM forward and reverse primer, and 1 ng template DNA. The PCR conditions were set to 94°C for 5 min; 39 cycles at 94°C for 30 s, 56°C for 45 s, and 72°C for 45 s; and a final extension of 72°C for 10 min. Of the 71 primers we had selected from the EST library, 21 primer pairs, as well as nine of the 15 published markers, amplified consistently under the standard PCR conditions. These loci were tested for polymorphism with eight samples per species of *F. insipida*, *F. citrifolia*, and *F. obtusifolia* using dye-labeled M13 adapters (Schuelke, 2000). Initial PCR conditions were changed to 29 cycles with an annealing temperature of 56°C and complemented by 10 cycles at 53°C to facilitate the incorporation of the universal dye-labeled primer (M13 protocol). PCR products were calibrated with a ROX 500 size standard (Applied Biosystems, Carlsbad, California, USA) and analyzed with an ABI 3730 Sequencer. Alleles were scored with GeneMarker version 1.91.

The screening for polymorphism yielded seven EST-based and five anonymous microsatellite DNA markers that were variable and easy to score for at least one of the studied species (Table 1). For the final amplifications, we maintained the M13 dye-labeling protocol. The results presented in Table 2 and in the text are based on 30 samples of *F. citrifolia*, *F. insipida*, *F. obtusifolia*, and *F. yoponensis*. All markers except FP25 and FP57 were further tested with 14 to 22 samples of *F. colubrinae*, *F. maxima*, *F. nymphaeifolia*, *F. popenoei*, and *F. tonduzii* and four to six samples of the rare species *F. costaricana* and *F. pertusa* (Table 3). With few exceptions, the primers amplified and were polymorphic in all tested populations and species.

For each primer, we calculated the number of alleles (A), the observed heterozygosity (H_o), and tests for deviations from Hardy-Weinberg equilibrium (HWE) and linkage disequilibrium with GENEPOP 4.1 (Rousset, 2008). For the studied free-standing figs, the loci had on average 6.0 alleles ranging from one to 15 (Table 2). H_o ranged from 0.10 to 0.97 with a mean of 0.56. In strangler figs, mean number of alleles was 7.2, ranging from one to 18 with a mean H_o of 0.63 (range: 0.20–0.90). There were few significant deviations from

HWE, indicating undetected or low frequency of null alleles. After application of the Bonferroni correction, no significant linkage disequilibrium was detected. The overall paternity exclusion probability, calculated in Cervus version 3.0.3 (Kalinowski et al., 2007), was 0.96 for *F. insipida*.

CONCLUSIONS

The set of highly variable microsatellite markers was transferable across species representing the most phylogenetically distant clades within *Ficus*, suggesting that these markers (or a subset of them) will be useful for ecological and evolutionary studies of any of the 750 *Ficus* species found in tropical regions around the world.

LITERATURE CITED

- CROZIER, Y. C., X. C. JIA, J. Y. YAO, A. R. FIELD, J. M. COOK, AND R. H. CROZIER. 2007. Microsatellite primers for *Ficus racemosa* and *Ficus rubiginosa*. *Molecular Ecology Notes* 7: 57–59.
- DOYLE, J., AND J. L. DOYLE. 1987. Genomic plant DNA preparation from fresh tissue-CTAB method. *Phytochemical Bulletin* 19: 11–15.
- HERRE, E. A., K. C. JANDER, AND C. A. MACHADO. 2008. Evolutionary ecology of figs and their associates: Recent progress and outstanding puzzles. *Annual Review of Ecology, Evolution, and Systematics* 39: 439–458.
- JEWELL, E., A. ROBINSON, D. SAVAGE, T. ERWIN, C. G. LOVE, G. A. C. LIM, X. LI, ET AL. 2006. *SSRPrimer* and *SSR Taxonomy Tree*: Biome SSR discovery. *Nucleic Acids Research* 34: W656–W659.
- KALINOWSKI, S. T., M. L. TAPER, AND T. C. MARSHALL. 2007. Revising how the computer program Cervus accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* 16: 1099–1106.
- NAZARENO, A. G., R. A. S. PEREIRA, J. M. FERES, M. A. MESTRINER, AND A. L. ALZATE-MARIN. 2009. Transferability and characterization of microsatellite markers in two Neotropical *Ficus* species. *Genetics and Molecular Biology* 32: 568–571.
- RØNSTED, N., G. D. WEIBLEN, J. M. COOK, N. SALAMIN, C. A. MACHADO, AND V. SAVOLAINEN. 2005. 60 million years of co-divergence in the fig-wasp symbiosis. *Proceedings of the Royal Society B. Biological Sciences* 272: 2593–2599.
- 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. J. SKALETSKY. 1999. Primer3 on the WWW for general users and for biologist programmers. Humana Press, Totowa, New Jersey, USA.
- SCHUELKE, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233–234.
- VIGNES, H., M. HOSSAERT-MCKEY, D. BEAUNE, D. FEVRE, M. C. ANSTETT, R. M. BORGES, F. KJELLBERG, AND M. H. CHEVALLIER. 2006. Development and characterization of microsatellite markers for a monoecious *Ficus* species, *Ficus insipida*, and cross-species amplification among different sections of *Ficus*. *Molecular Ecology Notes* 6: 792–795.

APPENDIX 1. Voucher information for *Ficus* taxa used in this study. All voucher specimens were collected at the Barro Colorado Natural Monument (BCNM), Panama (9°10'N, 79°51'W), and are deposited in the herbarium of the University of Panama. Information presented: species, country and locality, accession number.

subgen. *Urostigma* (sect. *Americana*)

- F. citrifolia* Mill.: Panama, BCNM, Península Bohio, Heer 002
- F. colubrinae* Standl.: Panama, BCNM, Península Bohio, Heer 003
- F. costaricana* (Liebm.) Miq.: Panama, BCNM, Península Bohio, Heer 004
- F. nymphaeifolia* Mill.: Panama, BCNM, Península Bohio, Heer 006
- F. obtusifolia* Kunth: Panama, BCNM, Península Bohio, Heer 011
- F. pertusa* L.f.: Panama, BCNM, Península Bohio, Heer 007

F. popenoei Standl.: Panama, BCNM, Península Bohio, Heer 008

subgen. *Pharmacosycea*

- F. insipida* Willd.: Panama, BCNM, Península Bohio, Heer 001
- F. maxima* Mill.: Panama, BCNM, Península Bohio, Heer 005
- F. tonduzii* Standl.: Panama, BCNM, Barro Colorado Island, Perez 2129
- F. yoponensis* Desv.: Panama, BCNM, Península Bohio, Heer 010
-