Conservation Genet Resour (2012) 4:807–810
DOI 10.1007/s12686-012-9649-y

TECHNICAL NOTE

Using next-generation DNA sequencing for rapid microsatellite discovery in Mexican leaf-toed geckos (*Phyllodactylus tuberculosus*)

Christopher Blair · Victor H. Jiménez-Arcos · Fausto R. Méndez de la Cruz · Robert W. Murphy

Received: 19 March 2012 / Accepted: 3 April 2012 / Published online: 19 April 2012
© Springer Science+Business Media B.V. 2012

Abstract We use a combination of 454 shotgun pyrosequencing and traditional molecular cloning to isolate microsatellite loci for the Mexican yellow-bellied leaf-toed gecko (*Phyllodactulus tuberculosus*). A pyrosequencing run on a 454 GS Junior yields 62 Mbp of data composed of 150 k fragments with an average length of 412 bp. Among the fragments, 18 144 (12 %) contain a suitable microsatellite. Dinucleotides and tetranucleotides are most frequently encountered. We then genotype 30 individuals of *P. tuberculosus* using 12 polymorphic loci. We find moderate allelic diversity and heterozygosity. No loci are in linkage disequilibrium and only two loci show deviations from Hardy–Weinberg expectations. Our results further highlight the utility of next-generation sequencing for discovering molecular markers in non-model taxa and suggest that sequencing on a 454 GS Junior system is a rapid and cost-effective approach for microsatellite characterization.

Keywords Evolution · Gecko · Genomics · Pyrosequencing

C. Blair (✉) · R. W. Murphy
Department of Ecology and Evolutionary Biology, University of Toronto, 25 Willcocks Street, Toronto, ON M5S 3B2, Canada
e-mail: Christopher.Blair@utoronto.ca

C. Blair · R. W. Murphy
Department of Natural History, Royal Ontario Museum, 100 Queen's Park, Toronto, ON M5S 2C6, Canada

V. H. Jiménez-Arcos · F. R. M. de la Cruz
Laboratorio de Herpetologia, Instituto de Biología, Universidad Nacional Autónoma de México, A.P. 70-153, C.P. 04510 Mexico, Mexico

Leaf-toed geckos (*Phyllodactylus)* occur in arid to semi-arid regions of the Neotropics (Dixon 1964). The yellow-bellied gecko, *Phyllodactylus tuberculosus*, ranges from southern Sonora, Mexico to northern Costa Rica. In Mexico, species are commonly encountered throughout tropical dry forest (TDF) habitat. TDF is considered one of Earth's biodiversity hotspots (Myers et al. 2000; García 2006) that is experiencing rapid rates of deforestation (Trejo and Dirzo 2000). However, little is known about how fragmentation will influence population connectivity for species inhabiting this ecosystem. Further, species within this genus pose a conservation concern as local people actively kill species of the genus as they are assumed to be venomous (pers. obs.). At present, the species has yet to be assessed by the IUCN. We combine molecular cloning with NGS to isolate and characterize polymorphic microsatellite loci for *P. tuberculosus*. The microsatellite loci obtained will be used for subsequent studies that examine fine-scale population genetic structure of geckos within Mexico's TDF.

In 2008 we sampled 30 individuals of *P. tuberculosus* from a single population adjacent to the Río Cuchujaqui near Alamos, Sonora. Tail tips were collected for genetic material and preserved in the field with 95 % ethanol. DNA was extracted using standard phenol–chloroform procedures. Extracts were precipitated and resuspended in distilled water.

We first used standard molecular cloning methods to isolate microsatellite loci (Hamilton et al. 1999). Genomic DNA was digested with the restriction enzymes NheI, RsaI and XmnI. Following digestion, SNX linkers were ligated to the digested fragments for amplification via PCR. Following amplification, we enriched for the following microsatellite motifs: $(CGG)_4$, $(AAAC)_6$, $(AAT)_{12}$, $(AAAT)_8$. Fragments containing the microsatellite motifs

Table 1 Characterization of twelve microsatellite loci for the leaf-toed gecko *Phyllodactylus tuberculosus* (n = 30) based on molecular cloning and 454 shotgun sequencing

Locus	Method	Repeat unit	Primer sequence (5′–3′)	Label	T_a (°C)	MgCl2 (mM)	Alleles	Range (bp)	H_O	H_E	GenBank accession
G2_22	Cloning	$(GAA)_{19}$	F: TGCCAGAAGTGGAAAACTAACC R: ATTGTGGAACTGGCATACCC	FAM	60	1.50	14	231–282	0.9000	0.9136	JQ906186
G2_37	Cloning	$(CTT)_{12}$	F: GAGAGGAAGGGAAGGCAACT R: GCCGCCTGATACTCTTTTTG	FAM	55	1.50	13	249–303	0.7000	0.8616	JQ906187
G2_96	Cloning	$(CAAA)_{11}$	F: GATTTGTAAGCCGCCTCAGT R: TCCAAGGACCTTTGTTGTTTG	FAM	55	1.50	8	197–229	0.7667	0.7373	JQ906188
G2_85	Cloning	$(CAAA)_{6}$	F: TGTGTTCTGCAAGGGAGAAA R: GGAACCATGTGAAGCCATCT	FAM	55	2.00	5	243–259	0.7333	0.6531	JQ906189
G2_59	Cloning	$(CCG)_{3}$	F: CGACGAAGACTCCAGAGACC R: TGGATGTTGCTGCAATTTGT	FAM	60	2.00	3	259–265	0.6207	0.5668	JQ906190
Pt1	454 sequencing	$(AAAT)_{10}$	F: ATTTTCCCTGGCTTTCCTTT R: CTTCCATTGTCAACGGGACT	FAM	55	2.50	9	231–267	0.6818	0.8277	JQ906191
Pt2	454 sequencing	$(AAAT)_{10}$	F: CCCCCTCTGGTGGTTATATTT R: CATGTCTGGCCAGTGTTCTG	FAM	53	1.50	6	229–265	0.7667	0.7418	JQ906192
Pt6	454 sequencing	$(AATG)_{11}$	F: GCAGTTGTTTTACCGCATCA R: AAGGCAGGACAAACCCTACC	FAM	55	1.50	5	198–214	0.8333	0.6915	JQ906193
Pt7	454 sequencing	$(AATG)_{13}$	F: TTATATTGTATGCCGCCTTGG R: AGCTGGAACGTGGTTCAGAG	FAM	60	2.00	9	231–263	0.6897	0.8361	JQ906194
Pt12	454 sequencing	$(ATCC)_{10}$	F: TGGCACTGAAATGGATGCTA R: AGGGTGGGCTTTTAGATGGT	FAM	60	2.00	8	206–246	0.8148	0.8162	JQ906195
Pt15	454 sequencing	$(CATT)_{10}$	F: CCCTGCTTTCCAAAACATCT R: AATGTGAAGCGGGAGGGTAT	FAM	55	1.50	7	213–237	0.7931	0.8312	JQ906196
Pt19	454 sequencing	$(GGAT)_{15}$	F: AGCCATCCTGAGATGTTCGT R: GATTGTTGACATGCAAAGCTAGT	FAM	54	2.00	7	183–219	0.8667	0.8395	JQ906197

All forward primers were labelled with an M13 tail and amplified using the three primer method of Schuelke (2000). T_a annealing temperature, $MgCl_2$ concentration of $MgCl_2$ in each reaction, H_O observed heterozygosity, H_E expected heterozygosity

were captured by streptavidin beads. The linker-ligated enriched DNA was then amplified, inserted into a plasmid vector, and cloned into *Escherichia coli* cells. Positive clones were then amplified using M13 forward and reverse primers and directly sequenced using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems). Sequences were then read on an ABI 3730 DNA sequencer (Applied Biosystems) and edited using BioEdit version 7.0.9 (Hall 1999). A single individual was also sent to Trent University's Natural Resources DNA Profiling and Forensics Centre for 454 shotgun sequencing using a full run on a 454 GS Junior. For genomic sequencing, we extracted DNA using a DNeasy Tissue Kit (QIAGEN) to obtain a cleaner product. This extract was then sent to Trent for library construction, emulsion PCR, and 454 sequencing.

MsatCommander (Faircloth 2008) was used to search for 454 fragments ($\sim$150 k) containing microsatellites. When suitable clones and 454 fragments were found, we used Primer3 (Rozen and Skaletsky 2000) to develop primers. Our total data set consisted of 12 loci (5 from cloning and 7 from 454 sequencing). We chose loci from the 454 run containing tetranucleotide repeats to minimize potential scoring errors due to stutter.

PCR amplification was performed using a 12.5 μL reaction volume containing 1.25 μL 10× PCR buffer containing 1.5 mM $MgCl_2$, 0.1 μL BSA, 1 μL 10 mM of each primer, 0.4 μL 10 mM dNTPs and 0.1 μL *taq* DNA polymerase (New England Biolabs). Water and $MgCl_2$ concentrations were adjusted to obtain optimal PCR products (Table 1). Each forward primer was fluorescently labelled with a 6-FAM M13 tag following Schuelke (2000). All PCRs were performed on an Eppendorf AG 5345 thermal cycler using the following conditions: initial denaturation at 95 °C for 2 min, followed by 35 cycles of 95 °C for 30 s, annealing at the optimal temperature for 45 s (Table 1), and extension at 72 °C for 45 s, with a final extension temperature of 72 °C for 6 min. Genotyping was performed by mixing 2 μL of each PCR product with a mixture of Hi-Di and Genescan 500 LIZ size standard (Applied Biosystems). Samples were read on an ABI 3730 DNA sequencer and alleles were scored using the software GeneMarker version 1.9.1 (SoftGenetics).

We used Microsatellite Analyser version 4.05 (Dieringer and Schlötterer 2003) to calculate the number of alleles per locus, allelic range, and observed and expected heterozygosities (H_O and H_E, respectively). Hardy–Weinberg and linkage equilibrium (log likelihood ratio statistic) were tested using Genepop version 4.0.10 (Rousset 2008) using default Markov chain parameters. We used Micro-Checker version 2.2.3 (van Oosterhout et al. 2004) to check for scoring errors due to stuttering, long allele dropout, or null alleles.

We were able to obtain reliable genotypes for 12 loci (three trinucleotide and nine tetranucleotide). Diversity was moderate, with the number of alleles per locus ranging between three and 14. Observed heterozygosity ranged between 0.6207 and 0.9000 and expected heterozygosity ranged from 0.5668 to 0.9136 (Table 1). No loci showed signs of linkage disequilibrium after Bonferroni correction ($P > 0.05$). Loci G2_37 and Pt1 showed deviations from Hardy–Weinberg expectations with an excess of homozygotes ($P < 0.01$ and $P = 0.0224$, respectively). No loci showed evidence of stutter or long allele dropout. However, G2_37 showed evidence of null alleles (0.0835). This study adds to a small body of literature that uses NGS to isolate polymorphic microsatellite markers in reptiles (e.g. Castoe et al. 2009; Metzger et al. 2011; Smith et al. 2011) and suggests that the 454 GS Junior is a rapid and cost-effective means to isolate microsatellite loci.

Acknowledgments We kindly thank the generation sequencing facilities available at Trent University's Natural Resources DNA Profiling and Forensics Centre for performing the 454 sequencing for this study. We thank Martin Gabriel Figueroa for help in the field. Stephanie Meyer provided logistic support for potential field sites. For laboratory assistance we thank Oliver Haddrath, Amy Lathrop, and Kristen Choffe at the Royal Ontario Museum. We thank Brant Faircloth for assistance with msatcommander. This work was supported by the Natural Sciences and Engineering Research Council (NSERC) Discovery Grant A3148 (RWM) and the Theodore Roosevelt Memorial Grant through the American Museum of Natural History (CB).

References

Castoe TA, Poole AW, Gu W, de Koning APJ, Daza JM, Smith EN, Pollock DD (2009) Rapid identification of thousands of copperhead snake (Agkistrodon contortrix) microsatellite loci from modest amounts of 454 shotgun genome sequence. Mol Ecol Resour 10:341–347

Dieringer D, Schlötterer C (2003) Microsatellite Analyser (MSA) a platform independent analysis tool for large microsatellite data sets. Mol Ecol Resour 3:167–169

Dixon JR (1964) The systematics and distribution of lizards of the genus Phyllodactylus in North and Central America. New Mexico State Univ Sci Bull 64:1–139

Faircloth BC (2008) Msatcommander: detection of microsatellite repeat arrays and automated, locus-specific primer design. Mol Ecol Resour 8:92–94

García A (2006) Using ecological niche modelling to identify diversity hotspots for the herpetofauna of Pacific lowlands and adjacent interior valleys of Mexico. Biol Conserv 130:25–46

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser 41:95–98

Hamilton MB, Pincus EL, Di Fiore A, Fleischer RC (1999) Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. BioTech 27:500–507

Metzger C, Ferchaud A, Geiser C, Ursenbacher S (2011) New polymorphic microsatellite markers of the endangered meadow viper (Vipera ursinii) identified by 454 high-throughput sequencing: when innovation meets conservation. Cons Genet Resour 3:589–592

Myers N, Mittermeier RA, Mittermeier CG, da Fonseca GAB, Kent J (2000) Biodiversity hotspots for conservation priorities. Nature 403:853–858

Rousset F (2008) GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. Mol Ecol Resourc 8:103–106

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S, Krawetz S (eds) Miseners. Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, NJ, pp 365–386

Schuelke M (2000) An economic method for the fluorescent labelling of PCR fragments. Nat Biotechnol 18:233–234

Smith A, Gardner M, Bull M, Driscoll D (2011) Primers for novel microsatellite markers in "fire-specialist" lizards (Amphibolurus norrisi, Ctenotus atlas and Nephrurus stellatus) and their performance across multiple populations. Cons Genet Resour 3:345–350

Trejo I, Dirzo R (2000) Deforestation of seasonally dry tropical forest: a national and local analysis in Mexico. Biol Conserv 94:133–142

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538

View publication stats