Conservation Genet Resour (2011) 3:65–67
DOI 10.1007/s12686-010-9291-5

TECHNICAL NOTE

Identification of 12 polymorphic microsatellite loci for the eastern box turtle (*Terrapene carolina carolina*)

S. J. A. Kimble · J. A. Fike · O. E. Rhodes Jr · R. N. Williams

Received: 14 July 2010 / Accepted: 16 July 2010 / Published online: 28 July 2010

Abstract We developed 12 highly polymorphic microsatellite markers for the eastern box turtle (*Terrapene carolina carolina*). The number of alleles per locus ranged from 5 to 27 (mean: 15.2 per locus) and the observed heterozygosities ranged from 0.371 to 0.943 among 35 individuals. Two loci exhibited significant deviations from Hardy–Weinberg equilibrium. Five loci exhibited putative null alleles but this is probably due to the high number of alleles relative to the sample size. There was no evidence of linkage disequilibrium between pairs of loci. This library should facilitate research on the population structure and ecology of this threatened species.

Keywords Population genetics · Terrestrial turtle · Genetic diversity · Power of exlusion · Probability of identity · Population decline

The eastern box turtle (*Terrapene carolina carolina*) is a predominantly terrestrial turtle that ranges across many eastern North American hardwood forests. It is experiencing population declines throughout much of its range, likely due to a combination of road mortality, habitat loss and fragmentation, and collection for the pet trade. It is considered threatened or endangered in several states (Dodd 2001). A greater understanding of the genetic consequences of population declines and habitat fragmentation are needed for successful management of this long-lived species. Currently, no subspecies-specific library exists for the eastern box turtle. While some confamilial and conspecific markers have been shown to cross-amplify in *T. c. carolina*, the levels of allelic richness were low and neither the powers of exclusion nor the probabilities of identity were reported.

We extracted DNA from blood samples from two eastern box turtles using a modified ammonium acetate protocol (modified from PUREGENE kit, Gentra Systems). We constructed dinucleotide [$(GT)_{12}$, $(CT)_{12}$] and tetranucleotide [$(GATA)_7$, $(GACA)_7$ and $(GATC)_7$] genomic libraries using a cloning protocol based on Hamilton and colleagues (1999) and modified as described previously (Beheler et al. 2004; Williams and DeWoody 2004). We transformed enriched DNA into competent *E. coli* cells for cloning and selected 530 colonies for amplification and sequencing. We chose 44 simple sequence repeat (SSR) sequences with suitable flanking regions in Sequencher 4.1 (Gene Codes Corporation). We designed primers with the aid of Primer3 (Rozen and Skaletsky 2000). We optimized 13 of these loci for PCR in 10 μL reactions with ~100 ng of DNA template, 10 mM Tris–HCL, 50 mM KCL, 0.05 mg/mL BSA, 1 mM $MgCl_2$, 0.2 mM of each dNTP, 0.3 μM of each end-labeled primer and 1 U of *Taq* polymerase (NEB). Thermocycler conditions were 94°C for 2 min, 94°C for 30 s, primer-specific annealing temperature (Table 1) for 30 s, 72°C for 30 s for 30 cycles, 72°C for 10 min and a final extension at 60°C for 45 min. All 13 loci produced bands of the expected size when visualized on 2% agarose gels for 35 individuals from a single population in Indiana. We then genotyped all samples on a 3730 automated sequencer (Applied Biosystems). We assigned genotypes with GeneMapper 3.7 (Applied Biosystems). Of the 13 loci tested, 12 loci demonstrated

S. J. A. Kimble (✉) · J. A. Fike · O. E. Rhodes Jr · R. N. Williams
Department of Forestry and Natural Resources, Purdue University, 715 West State Street, West Lafayette, IN 47907, USA
e-mail: skimble@purdue.edu

Table 1 Characterization of 12 hypervariable nuclear loci for the eastern box turtle *Terrapene carolina carolina* including locus name, GenBank accession number, primer sequences, repeat motif, PCR product size (bp), allelic richness (A), annealing temperature (T_A), expected heterozygosity (H_e) and observed heterozygosity (H_o)

Locus name	GenBank accession number	Primer sequences	Repeat motif	PCR product size range (bp)	A	T_A (°C)	H_e	H_o
TCC_di_045[b]	HM622274	F: GCACTTCAGTCCTAGACAATATGC	$(CA)_{10}$	245–281	15	58.0	0.883	0.657
		R: CCTGATGCCTTAACAATGAGC						
TCC_di_082	HM622275	F: TCCCTGAATGCATCAGTAACC	$(CT)_{12}$	104–134	12	58.0	0.834	0.771
		R:CCAAATTGGGCTTATTTTGAA						
TCC_di_189[b]	HM622276	F: CCATCATCCCCAGTAAATGC	$(CT)_{13}$	152–162	5	60.0	0.526	0.371
		R: AGGAGTGGCAGGAAATGTTG						
TCC_di_300	HM622277	F: TGCTCTCTGGGTACCTGCAC	$(CA)_{14}$	168–232	18	57.0	0.918	0.914
		R: AGCTATGTTGGCGCTTAGGTC						
TCC_di_318	HM622278	F: CTATGGAACTCACTGCCACAA	$(TG)_{26}$	152–216	26	57.0	0.957	0.943
		R: TATGCCTTGAAGTGCGACA						
TCC_di_345[a,b]	HM622279	F: GTGACTTGTGGCCTGGCTAT	$(TG)_{23}(GA)_5(CA)_5$	146–196	27	63.0	0.959	0.771
		R: GGACAGAGTGGGACTTGCTC						
TCC_di_352	HM622280	F: GATCTGCTTGCAGAGGTTCC	$(GA)_{18}$	180–214	13	57.0	0.892	0.829
		R: GAAATGATACATCTTCTAAGTGTGTGG						
TCC_di_366	HM622281	F: ACAACAGATGTTGGCTGGAA	$(TG)_{11}$	158–174	8	57.0	0.817	0.857
		R: TGGTCAACAAGAGCTAACTTGG						
TCC_tetra_012/342[b]	HM622282	F: TGCACCTACTAAAACCAAAATCC	$(TATC)_{17}$	196–254	15	60.0	0.825	0.686
		R: GCACAGTTCTGACCCCTTAAA						
TCC_tetra_043	HM622283	F: TGCTATGCAGTATCTAATAATCACG	$(TATC)_{10}$	168–220	13	60.0	0.874	0.857
		R: AAGTCAAGGAAGTCTGAAGGATT						
TCC_tetra_070	HM622284	F: TTGATGAGTGTGCATGATGA	$(GATA)_{14}$	178–250	15	58.0	0.931	0.914
		R: CTTCCAGTCACCTAATTGATCTT						
TCC_tetra_309[a,b]	HM622285	F: TGTTTGAGCACATGGGTTGT	$(TATC)_{19}$	162–234	15	58.0	0.910	0.514
		R: CTCATGACCTGACCCATGATT						

[a] Significant deviation from Hardy–Weinberg equilibrium ($\alpha = 0.0042$)

[b] Putative null alleles frequency >0.05

polymorphism with an average of 15.2 alleles per locus ranging from 104 to 281 bp in size.

We used CERVUS (Marshall et al. 1998) to calculate expected heterozygosity (H_e) and observed heterozygosity (H_o, Table 1), the average exclusion probability and probability of identity. The average exclusion probability with one known parent was 99.9989% and with two known parents was 99.9999%. The probability of identity was 1.81×10^{-19}. Using sequential Bonferroni corrections, we calculated deviations from Hardy–Weinberg equilibrium (HWE, $\alpha = 0.00417$) and linkage disequilibria ($\alpha = 0.000758$) using GENEPOP 4.0.10 (Raymond and Rousset 1995). Loci TCC_di_345 and TCC_tetra_309 were found to be out of HWE. No loci were found to be out of linkage equilibrium. We calculated the probabilities of null alleles with MICRO-CHECKER (van Oosterhout et al. 2004). Significant probabilities of null alleles were found for TCC_di_045, TCC_di_189, TCC_di_345, tetra_012/342 and TCC_tetra_309, but are likely due to the high allelic richness compared to the sample size.

Acknowledgments We thank Shem Unger, Brad Bumgardner, Katie Smith and all who collected samples. Turtles were collected under Indiana Department of Natural Resources permits and as approved by Purdue University Animal Care and Use Committee. Funding for this research was provided by the Indiana Department of Natural Resources Division of Fish and Wildlife, Wildlife Diversity Section SWIG #E2-08-WDS15.

References

Beheler AS, Fike JA, Murfitt LM, Rhodes OE Jr, Serfass TS (2004) Development of the polymorphic loci for North American river otters (*Lontra canadensis*) and amplification in related Mustelids. Mol Ecol Notes 4:56–58

Dodd CK (2001) North American box turtles: a natural history. University of Oklahoma Press, Norman, OK, USA

Hamilton MB, Pincus EL, DiFiore A, Fleischer RC (1999) Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. BioTechniques 27(3): 500–507

Marshall TC, Slate J, Kruuk LEB, Pemberton JM (1998) Statistical confidence for likelihood-based paternity inference in natural populations. Mol Ecol 7:639–655

Raymond M, Rousset F (1995) GENEPOP version 1.2. Population genetics software for exact tests and ecuminicism. J Hered 86:248–249

Rozen S, Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, NJ, pp 365–396

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

Williams RN, DeWoody JA (2004) Fluorescent dUTP helps characterize 10 novel tetranucleotide microsatellites from an enriched salamander (*Ambystoma texanum*) genomic library. Mol Ecol Notes 4:17–19