MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2010) **10**, 177–180 doi: 10.1111/j.1755-0998.2009.02707.x

MOLECULAR DIAGNOSTICS AND DNA TAXONOMY

Identification of single copy nuclear DNA markers for North American pit vipers

H. LISLE GIBBS and JOSE DIAZ
Department of Evolution, Ecology and Organismal Biology, Ohio State University, Columbus, OH 43210-1293, USA

Abstract

We describe 18 single copy nuclear DNA loci (10 loci cloned from a *Sistrurus catenatus catenatus* genomic library and eight intron-based loci amplified using conserved primers) that detect sequence variation in species from all genera (*Sistrurus*, *Agkistrodon* and *Crotalus*) of North American pit viper snakes. These loci (mean size in bp ± SE: 433 ± 51) show large numbers of phylogenetically informative sites across species (mean + SE: 10.2 ± 1.5), but limited variation within subspecies suggesting that they will be most useful for multilocus species-level phylogenetic analyses in these snakes.

Keywords: *Agkistrodon*, *Crotalus*, North American pit vipers, single copy nuclear DNA markers, *Sistrurus*

Received 25 November 2008; revision accepted 21 March 2009

Introduction

The use of data from multiple DNA loci promises to increase the accuracy with which population genetic and phylogenetic parameters are estimated because the statistical power to make accurate estimates increases with the number of loci used (Brumfield *et al.* 2003). However, in non-model organisms, the potential for this advance is dependant on the availability of conserved primer sets that can amplify multiple homologous loci in different species. With the goal of increasing the number of such loci in one understudied yet important group of vertebrates (snakes), we describe primer sets which amplify 18 single copy nuclear DNA markers in North American pit vipers.

Materials and Methods

We identified single copy nDNA loci in viperid snakes using two approaches: isolation of anonymous loci from a genomic DNA library generated from a single *Sistrurus catenatus catenatus* or screening for variable loci using previously published conserved intron-spanning primer sets that amplified presumed homologous regions in birds, amphibians and reptiles (see references below).

Correspondence: H. Lisle Gibbs, Fax: 614 292 2030;
E-mail: gibbs.128@osu.edu

For the first approach, we generated a small insert plasmid library using genomic DNA extracted from blood sample from a single *S. c. catenatus*. Briefly, 10 μg of DNA was digested with *Mbo*I, run on a 1.5% agarose gel and fragments 400–600 bp in size excised. The DNA fraction was then purified using a phenol–chloroform extraction, ETOH precipitated and resuspended in water. Genomic DNA (50 ng) was then ligated into *Bam*HI precut pUC-19 vector, transformed into competent DH5α *Escherichia coli* cells and plated on LB plates. Two-hundred colonies were then picked, grown overnight in LB media and inserts sequenced using an ABI 3100 DNA Analyser. Based on sequence clarity and lack of repetitive DNA, we selected 14 sequences for further analyses and designed *Sistrurus*-specific primers to amplify these regions using Primer 3 (Rozen & Skaletsky 2000). We amplified these sequences from genomic DNA in 10 μL reactions consisting of ~20 ng snake DNA, 1x PCR buffer (50 mM KCl, 20 mM Tris–Hcl pH 8.4), 0.2 mM each dNTP, between 1.5 and 3.5 mM $MgCl_2$, 0.05 U *Taq* polymerase (Invitrogen, Life technologies) and 0.25 mM of each of the forward and reverse primers. Thermal cycling was carried out in a MJ PTC-200 thermal cycler (Bio-Rad); conditions included an initial denaturation step at 94 °C for 3 min, followed by 30 cycles at 94 °C for 20 s, 53 °C for 20 s and 72 °C for 60 s with a final extension step at 72 °C for 5 min. Products were then sequenced using an ABI 3100 DNA Analyser.

Our second approach involved screening a large number of previously identified conserved primer sets that had successfully amplified sequences in other nonmammalian vertebrates. Overall, we screened a total of 30 sets of primers based on sequences described in Friesen *et al.* (1997, 1999 and unpublished), Dolman & Phillips (2004) and Creer *et al.* (2005). For initial screening, we used a standard set of conditions (as above but with an annealing temperature of 50 °C, 2 mM $MgCl_2$ and denaturation, annealing and extension times of 60 s) and attempted to amplify DNA from multiple individuals from species in each of *Sistrurus* and *Agkistrodon*. When a crisp, clear band of DNA of approximately the expected size was obtained from a given amplification, it was sequenced.

Results and Discussion

Based on the procedures described earlier, we identified primer sets for 10 anonymous loci and eight intron-based loci that reliably yielded clear, unambiguous hence presumably single copy sequence when used to amplify DNA from North American viperid snakes (Table 1). To

Table 1 Primer sequences for (a) anonymous loci isolated from a *Sistrurus catenatus catenatus* genomic library and (b) Introns amplified using conserved primers described in the literature

Locus	Primer sequences (5′–3′)	Reference
(a)		
A	F: AGA ATT GAG CTC CCG TCC TTT	This study
	R: GGG AGC AAT GCC TAG ACC AAG	
1	F: GAA GCT GCC CCA GTT GAA	This study
	R: CTC TTA AAT CAG GCT GCA GTC	
4	F: GGG GAT CTC CTC CGT GTT A	This study
	R: AGC GGT TAC CAC AGC AAC AA	
11	F: TCC TTA CTG AGT GAG CAC C	This study
	R: GCA AAG TCA ATG GAG AAA G	
25	F: ACC TTT CTC TTT TGT TCA GCA	This study
	R: ATG TCT CTG TTT CCC AAA ATG	
31	F: GGA GAA GTT TCA TAC CCA	This study
	R: TGA ATC TAA GCA AGA ACT TGA	
41	F: CAA ACA ACC AAT GAG GAT GTA	This study
	R: TTT TGA GGG TAC TTC CAT GTG	
51	F: ACT TGC CTT CAG AAA TCA TG	This study
	R: GGG ATC AAA GGT TTA AAG AA	
61	F: ACT CTT CTA ATT GTC ATG GCC	This study
	R: TTT TAG GCT TGA CAG ATG AAT	
63	F: ATT AGC CCA GAA CTG TGC TTA	This study
	R: AAA GAT TCT GGG AAG CCA AA	
(b)		
TATA box-binding protein (TBP)	F: CCT TTA CCA GGA ACC ACA CC	1
	R: CGA AGG GCA ATG GTT TTT AG	
Cytoplasmic β-actin (CBA)	F: GGC ACC ACA CTT TCT ACA ATG AGC	4
	R: TAC TAA CTT TGG CTT GCT C	
Ornithine decarboxlyase (OD)	F: GAC TCC AAA GCA GTT TGT CGT CTC AGT GT	2
	R: TCT TCA GAG CCA GGG AAG CCA CCA CCA AT	
Ets oncogene (ETS)	F: CCA TCA ACA GAC ACA CAG G	3, 5
	R: GTC TGC TTT TTA CTT TGC G	
Elongation factor 1 (EF)	F: AAG CGA ACC ATC GAG AAG TTC GAG AAG G	5
	R: AGT AAT CAT GTT CTT GAT GAA GTC CCT GTG	
Glyceraldehyde-3-phosphate dehydrogenase (GAPD)	F: ACC TTT AAT GCG GGT GCT GGC ATT GC	3
	R: CAT CAA GTC CAC AAC ACG GTT GCT GTA	
Lamin (LAM)	F: CCA AGA AGC AGC TGC AGG ATG AGA TGC	4
	R: CTG CCG CCC GTT GTC GAT CTC CAC CAG	
β-Fibrinogen (FGB)	F: AGA GAC AAT GAT GGA TGG TAA G	1
	R: CCT TTT GGG ATC TGG GTG TA	

References that are sources for intron primer sequences are as follows: (1) Creer *et al.* (2005); (2) Friesen *et al.* (1999); (3) Dolman & Phillips (2004); (4) Friesen *et al.* (1997); (5) Friesen *et al.*, unpublished manuscript.

Table 2 Characterization of variation in nuclear loci across *Sistrurus* and *Agkistrodon* species and subspecies listed in Methods. Results are also given for taxon (*Sistrurus catenatus catenatus*) with the largest sample size as an estimate of levels of intraspecific variation

Locus	Approximate size (bp)	*N*	No. haplotypes (E cat only)	Accession number	π ± SD (E cat only)	No. indels	No. PI sites	Divergence (d_{xy} ± SD) S. cat vs S. mil	Sis vs. Agk
A	296	30	6 (2)	FJ660324–FJ660353	0.012 ± 0.0037 (0.0016 ± 0.00048)	3	17	0.011 ± 0.0034	0.055 ± 0.022
1	220	30	6 (1)	FJ660056–FJ660085	0.011 + 0.0024 0	1	9	0.021 ± 0.0082	0.031 ± 0.0017
4	267	30	5 (1)	FJ660086–FJ660115	0.0053 ± 0.0015 0	3	5	0.0079 ± 0.0031	0.020 ± 0.011
11	420	29	7 (1)	FJ660116–FJ660144	0.0073 ± 0.0022 0	0	13	0.011 ± 0.0054	0.026 ± 0.014
25	262	30	8 (2)	FJ660145–FJ660174	0.014 ± 0.0029 (0.00075 + 0.00060)	1	12	0.014 ± 0.0047	0.045 ± 0.017
31	256	30	4 (1)	FJ660175–FJ660204	0.0031 ± 0.00081 0	2	4	0.0055 ± 0.0023	0.0095 ± 0.0048
41	274	30	3 (1)	FJ660205–FJ660234	0.0019 ± 0.00091 0	1	4	0.0032 + 0.0030	0.0045 ± 0.0027
51	260	29	10 (1)	FJ660235–FJ660263	0.017 ± 0.0037 0	1	17	0.020 ± 0.0060	0.056 ± 0.019
61	194	30	3 (1)	FJ660264–FJ660293	0.0041 ± 0.0026 0	0	5	0	0.032 ± 0.023
63	471	30	6 (1)	FJ660294–FJ660323	0.0052 ± 0.0022 0	0	14	0.0021 ± 0.00087	0.030 ± 0.014
TBP	796	29	10 (1)	FJ660383–FJ660411	0.0080 + 0.0014 0	6	23	0.0088 ± 0.0026	0.022 ± 0.0081
CBA	525	28	9 (1)	FJ659889–FJ659916	0.0059 ± 0.0013 0	3	6	0.0041 ± 0.0014	0.0060 ± 0.0016
OD	522	29	6 (1)	FJ660354–FJ660382	0.011 ± 0.00083 0	2	14	0.013 ± 0.0048	0.016 ± 0.0070
ETS	830	29	5 (1)	FJ659944–FJ659972	0.0066 ± 0.0018 0	2	20	0.0064 ± 0.0024	0.025 ± 0.011
EF	447	27	4 (1)	FJ659917–FJ659943	0.0042 + 0.0022 0	0	12	0.00048 ± 0.00034	0.026 ± 0.016
GAPD	260	29	3 (2)	FJ660001–FJ660029	0.0028 + 0.0015 (0.0015 + 0.0039)	4	3	0	0.017 ± 0.013
LAM	684	26	9 (2)	FJ660030–FJ660055	0.0038 ± 0.0013 (0.0016 ± 0.0003)	3	10	0.0030 ± 0.0013	0.0022 ± 0.0010
FGB	730	28	9 (1)	FJ659973–FJ660000	0.0048 ± 0.0020 0	5	6	0.0036 ± 0.0013	0.033 ± 0.018
ATP	665	29	13 (4)	FJ659860–FJ659888	0.12 + 0.013 (0.0015 + 0.00063)	0	201	0.20 ± 0.045	0.28 ± 0.081

Values provided for approximate size given in base pairs; number of sites with indels (range of indel sizes); nucleotide diversity (π) estimated using DnaSP; mean uncorrected sequence divergence between sequences in all taxa and number of distinct sequences observed across all taxa. In addition, the following loci were successfully amplified and sequenced in the following *Crotalus* species: Loci 1, 4, 25, 31, 41, 61, 63, A, OD [*cerastes* (c), *viridis* (v), *horridus* (h); Locus 25 (c, v, h) and Locus A (v, h only)].

assay the variation in these loci at both intra- and interspecific levels in these snakes, we genotyped a maximum of 21 individuals from three named subspecies of *Sistrurus catenatus* [*catenatus* ($n = 11$); *tergeminus* (6) and *edwardsi* (6)], seven individuals from three subspecies of *Sistrurus miliarius* [*barbouri* (3); *streckeri* (2) and *miliarius* (2)] and single individuals of *Agkistrodon piscivorous* and *Agkistrodon contortrix* at all 18 loci. We also sequenced a subset of loci in single individuals from three species of *Crotalus* (*Crotalus cerastes*, *Crotalus viridis* and *Crotalus horridus*). For comparison, we also sequenced all *Sistrurus* and *Agkistrodon* individuals for a 665-bp fragment of the mtDNA genes ATP 6-8 using primers described in Douglas *et al.* (2002). Sequences were compiled and variation characterized using the program DnaSP ver. 4.5 (Rozas *et al.* 2003) (Table 2).

Most variations at these loci were detected at the interspecific level with only limited variation present among the relatively large range-wide sample of *S. c. catenatus* individuals (Table 2). Across all samples, three to 10 distinct haplotypes (mean ± SE: 6.1 ± 0.56) were detected (excluding LAM), whereas within *S. c. catenatus* there was a maximum of two different sequences detected at any locus. Indels were found at most loci (14 of 18; 78%) and the mean (± SE) estimate of nucleotide diversity per site (π) was 0.0077 ± 0.0011 and was about twice as high for the anonymous (0.0081 ± 0.0016) loci compared to the intron-based (0.0041 ± 0.0014) loci. As expected, these loci also showed higher levels of divergence when comparisons were made between genera (mean d_{xy} across all loci for *Sistrurus* vs. *Agkistrodon*: 0.025 ± 0.0037) vs. within genera (*S. miilarius* vs. *S. catenatus*: 0.0073 ± 0.0015). Finally, as predicted from differences in population genetic characteristics of mitochondrial vs. nuclear DNA, the ATP sequences for these snakes were substantially more variable than the nuclear loci: more distinct haplotypes were detected among the ATP sequences [13 vs. 6.1 (mean)] and nucleotide diversity was more than an order of magnitude higher [0.12 vs. 0.0077 (mean)].

The limited levels of intraspecific variation in *S. c. catenatus* compared with the much higher levels of variation at the species and generic levels suggest that these loci will be most useful for species-level phylogenetic analyses in these snakes (see Kubatko & Gibbs unpublished manuscript).

Acknowledgements

We thank all those who assisted in collection of blood and tissue samples (see Kubatko & Gibbs unpublished manuscript for more complete list), V. Friesen for providing unpublished primer sequences and Ohio State University and Ohio Department of Natural Resources—Division of Wildlife for funding.

References

Brumfield RT, Beerli P, Nickerson DA *et al.* (2003) The utility of single nucleotide polymorphism is inferences of population history. *Trends in Ecology and Evolution*, **18**, 249–256.

Creer S, Malhotra A, Thorpe RS *et al.* (2005) Targeting optimal introns for phylogenetic analyses in non-model taxa: experimental results in Asian pitvipers. *Cladistics*, **21**, 390–395.

Dolman G, Phillips B (2004) Single copy nuclear DNA markers characterized for comparative phylogeography in Australian wet tropics rainforest skinks. *Molecular Ecology Notes*, **4**, 185–187.

Douglas ME, Douglas MR, Schuett GW *et al.* (2002) Phylogeography of the Western Rattlesnake (*Crotalus viridis*) complex, with emphasis on the Colorado Plateau. In: *Biology of the Vipers* (eds Schuett GW, Höggren M, Douglas ME, Greene HW), pp. 11–50. Eagle Mountain Publishing, Eagle Mountain, UT.

Friesen VL, Congdon BC, Kidd MG *et al.* (1997) Intron variation in marbled murrelets detected using analyses of single-stranded conformation polymorphisms. *Molecular Ecology*, **6**, 1047–1058.

Friesen VL, Congdon BC, Kidd MG *et al.* (1999) Polymerase chain reaction (PCR) primers for the amplification of five nuclear introns in vertebrates. *Molecular Ecology*, **8**, 2141–2152.

Rozas J, Sánchez-DelBanco JC, Messeguer X *et al.* (2003) DnaSP, DNA polymorphism analysis by the coalescent, and other methods. *Bioinformatics*, **19**, 2496–2497.

Rozen S, Skaletsky HJ (2000) Primer 3 on the www for general users and biologist programmers. In: *Bioinformatics Methods and Protocols: Methods in Molecular Biology* (eds Krawetz S, Misener S), pp. 365–386. Humana Press, Totowa, NJ. Available from http://frodo.wi.mit.edu/