TECHNICAL NOTE

New primers for the amplification and sequencing of nuclear loci in a taxonomically wide set of reptiles and amphibians

Catarina Pinho · Sara Rocha · Bruno M. Carvalho · Susana Lopes · Sofia Mourão · Marcelo Vallinoto · Tuliana O. Brunes · Célio F. B. Haddad · Helena Gonçalves · Fernando Sequeira · Nuno Ferrand

Received: 8 October 2009 / Accepted: 9 October 2009 / Published online: 14 November 2009 © Springer Science+Business Media B.V. 2009

Abstract We report new primers for the amplification and sequencing of 11 nuclear markers in squamate reptiles and anuran amphibians (five in squamates, six in anurans). Ten out of the 11 loci are introns (three of which are linked) that were amplified using an exon-primed, introncrossing (EPIC) PCR strategy, whereas an eleventh locus spans part of a protein-coding gene. Squamate and anuran primers were initially developed for *Lacerta schreiberi* (Squamata: Lacertidae) and *Pelodytes* spp. (Anura: Pelodytidae), respectively. Cross-species amplification of the squamate markers was evaluated in four genera representing two additional families, whereas for anurans three genera corresponding to three additional families were

tested. Three out of the five loci were successfully sequenced in all squamate taxa tested. Cross-amplification of the six anuran markers had lower, but still significant, success. We predict these markers will be of great utility for both population genetics and phylogenetic studies.

Keywords Anura · Squamata · Nuclear sequence markers · Introns

Reptiles and amphibians are becoming important model organisms for evolutionary studies. Their low dispersal capabilities and usually strong dependency on environmental conditions make these taxa particularly susceptible to climate change and appropriate models for studying the influence of climatic or geologic phenomena on the distribution of genetic diversity. Additionally, amphibians and reptiles are also amongst the most problematic animals in terms of conservation, with well documented world-wide declines in recent years (Stuart et al. 2004; Gibbons et al. 2000).

Most studies addressing evolutionary questions in amphibians and reptiles have relied solely on mitochondrial DNA (mtDNA) variation. MtDNA displays a set of properties that makes it a useful marker in phylogeographic and phylogenetic studies (Avise 2000); however, such single-locus studies are prone to erroneous interpretations, both because of random lineage sorting (e.g., Irwin 2002) and natural selection (e.g., Ballard and Whitlock 2004), and usually provide low resolution for the estimation of parameters of interest (Edwards and Beerli 2000). In recent years, the acknowledgement of these limitations and the improvement of analytical methods especially suited for multilocus data sets have led researchers to investigate the nuclear genome.

C. Pinho () · S. Rocha · B. M. Carvalho · S. Lopes ·

S. Mourão · M. Vallinoto · T. O. Brunes · H. Gonçalves ·

F. Sequeira · N. Ferrand

CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos. Campus Agrário de Vairão, Universidade do Porto, 4485-661 Vairão, Portugal e-mail: catarina@mail.icav.up.pt

S. Rocha · N. Ferrand

Departamento de Zoologia e Antropologia, Faculdade de Ciências da Universidade do Porto, R. do Campo Alegre, s/n, 4169-007 Porto, Portugal

S. Rocha

Departamento de Bioquímica, Genética e Inmunología, Facultad de Biología, Universidad de Vigo, Vigo 36310, Spain

M. Vallinoto

Universidade Federal do Pará, Campus de Bragança, Alameda Leandro Ribeiro s/n, Bragança, Pará 68.000-000, Brazil

C. F. B. Haddad

Departamento de Zoologia, Instituto de Biociências, Universidade Estadual Paulista, Rio Claro, SP 13506-900, Brazil



In reptiles and amphibians, however, this has proven to be a difficult task, since the pace of development of genomic resources (e.g. Expressed Sequence Tags (EST) databases, whole-genome sequencing projects, etc.) has been slow when compared to other vertebrate groups. For most cases, whenever primers for the amplification of nuclear genes were developed, these corresponded either to slow-evolving protein coding genes (e.g., Saint et al. 1998; Hoegg et al. 2004) with low utility for population genetics and phylogeography, or to species-specific efforts with low probability of cross-amplification in distant taxa (e.g., Dolman and Phillips 2004).

Here, we describe the development of nuclear sequence markers in several squamate reptiles and anuran amphibians occurring in a variety of regions including Southern Europe, North Africa, South America and Indian Ocean islands. The original target species of this work were Schreiber's green lizard (Lacerta schreiberi) and the Parsley frog (Pelodytes ibericus/punctatus group). L. schreiberi is an Iberian endemic species which comprises two well-defined phylogroups inferred from both mtDNA (4.67% divergence in cytochrome b; Paulo et al. 2001) and nuclear genealogies (Godinho et al. 2006). Iberian Pelodytes have been recently divided in two species (Sánchez-Herraíz et al. 2000) although mtDNA variation evidences the existence of four similarly distinct phylogroups showing 3.23-4.16% divergence in cytochrome b (M. Tejedo, personal communication).

Most of the loci we report were amplified following an exon-primed, intron-crossing (EPIC) polymerase chain reaction (PCR) methodology. We targeted genes for which mRNA sequences of either a reptile or an amphibian were available in GenBank. Resource databases (namely the Ambystoma tigrinum EST database—http://salamander.uky. edu—and the *Xenopus* genomic database—www.xenbase. org) were also used to choose putative target genes. For each gene our alignments typically further included human, mouse, chicken and zebrafish mRNA sequences (when available) and were visually inspected for highly conserved regions appropriate for primer anchorage. Whenever possible, primers were designed at least 50 bp away from the splicing site in order to ensure a sufficiently long exonic sequence for molecular confirmation of gene identity. For the locus MC1R, primers were designed based on an alignment of several squamate sequences obtained by Rosenblum et al. (2004), and chosen in order to amplify a fragment spanning ~ 700 bp of the gene.

A total of 24 markers was tested in 2–5 individuals of both *L. schreiberi* and *Pelodytes* spp. Initial essaying conditions were the same for all the markers and for both species, although some adjustments had to be made prior to sequencing: PCRs were carried out in 10 μ l volume containing 1× PCR buffer (50 mm Tris–HCl, 50 mm NaCl,

pH 8.5); 3 mM MgCl₂; 0.4 mM each dNTPs, 0.5U of GoTaq DNA polymerase (Promega), 0.3 μM each primer and approximately 50 ng of genomic DNA. Amplification conditions consisted of a pre-denaturing step of 5 min at 92°C followed by 40 cycles of a denaturing step of 30 s at 92°C, annealing at 50°C for 30 s and extension at 72°C for 90 s. The final extension was accomplished at 72°C for 5 min. For successful amplifications, PCR products were enzymatically purified (with the ExoSap-IT purifying kit, Amersham-Pharmacia Biotech) and sequenced from both ends using BigDye 3.1 on an Applied Biosystems 3100xl DNA sequencer.

From the tested markers, five (CTSDint8, MC1R, NFY-Cint16, PKM2int5 and RELNint61) were successfully amplified and sequenced in L. schreiberi, and another four (CHERPint7, PPPCAint4, RPL3int5 and RPL9int5) had similar success in Pelodytes. Two additional loci, corresponding to introns 4 and 6 of the RPL9 gene (and hence in close linkage to RPL9int5) were also amplified in the latter species, but only with marginal success (most individuals did not amplify). From these two markers, good quality sequences were obtained only for RPL9int4.

The sequences of the primers developed in this study are presented in Table 1.

Cross-amplification tests for the loci developed for L. schreiberi were performed in Podarcis (Lacertidae; species used: P. bocagei and P. carbonelli, 12.6% divergence in cytochrome b (Pinho et al. 2006)), Mabuya (Scincidae; species used: M. seychellensis and M. wrightii, \sim 8% divergence in cytochrome b (S. Rocha, unpublished data)), Phelsuma (Gekkonidae; species used: P. sundbergi and P. astriata, $\sim 13\%$ divergence in cytochrome b (S. Rocha, unpublished data)) and Urocotyledon inexpectata (Gekkonidae; two intraspecific phylogroups showing a divergence of $\sim 8\%$ in cytochrome b (S. Rocha, unpublished data)). For anurans, we essayed cross-amplification of *Pelodytes* primers in *Alytes* (Discoglossidae; taxa used: A. cisternasii, A. maurus, A. dickhilleni, A. muletensis and 4 subspecies of A. obstetricans, with divergence ranging from 0.29 to 15.5% for cytochrome b (Martínez-Solano et al. 2004)), Rhinella marina (Bufonidae; three phylogroups with divergences ranging from 1.2 to 2.3% in cytochrome b and control region combined (M. Vallinoto, unpublished data)) and Phyllomedusa (Hylidae; species used: P. burmeisteri and P. iheringii showing a divergence of 7.7% in ND2 (T. O. Brunes, unpublished data)). We used the initially developed amplification protocol with minor modifications for each species.

Success in amplification and sequencing was obtained for all five squamate markers in *Podarcis*, in four markers (all but *CTSDint8*) in the two gekkonid genera and three markers in *Mabuya* (*MC1R*, *PKM2int5* and *RELNint61*). In contrast, in anurans almost non-overlapping subsets of the



Table 1 Primers successfully used in this study for the amplification of nuclear markers

Group	Gene	Intron ^a	Locus abbreviation	Primer name	Primer utility ^b	Primer sequence 5'-3'		
Squamates	Cathepsin D	8	CTSDint8	CTSD8F	A	TACATGCTTCCATGTGATAAGCTGT		
				CTSD9R	A	ATGAAGACATCTCCCAGGATCCA		
	Melanocortin receptor 1	-	MC1R	MC1RF	A	GGCNGCCATYGTCAAGAACCGGAACC		
				MC1RR	A	CTCCGRAAGGCRTAAATGATGGGGTCCAC		
	Nuclear transcription factor Y, gamma	16	NFYCint16	NFYC16F1	В	GTCCAGCARGGACAGCAGCAGTTCAGC		
				NFYC16F2	C	GCARGGACAGCAGCAGTTCAGCCAGTT		
				NFYC17R1	D	GGCATRGTSACTTGCTGRATCTGGTA		
				NFYC17R2	E	GCWGGCATRGTSACTTGCTGRATCTGG		
				NFYCLscF	F	GTTGGATTAAAAGAGGATCAGACA		
	Pyruvate kinase, muscle	5	PKM2int5	PKM25F	G	AGTGGYACAGCAGARGTGGAGCTCAA		
				PKM26R	G	TTYTCAATYTTGCTGATRATCTTGATG		
				PKSQF	A	ACCAAAGTTGTWGATGTTGGCAGC		
				PKSQR	A	ATGAAGGAAGCAAACACCATGTC		
	Reelin	61	RELNint61	RELN61F	A	GAGTMACTGAAATAAACTGGGAAAC		
				RELN62R	A	GCCATGTAATYCCATTATTTACACTG		
Anurans	Calcium homeostasis endoplasmic reticulum protein	7	CHERPint7	CHERP7F1	Н	GCTCTGGGARAAGAAYGGCTACTT		
				CHERP7F2	I	CTGGGARAAGAAYGGCTACTTYGATGA		
				CHERP8R	J	CTGNAKCTGCTGCTGGAAKGCCA		
	Protein phosphatase 3, catalytic subunit, alpha isoform	4	PPP3CAint4	PPP3CA4F1	K	CTGTAYTTGTGGGCCTTGAAAATTC		
				PPP3CA4F2	G	CTWCGTGGRAATCATGAATGTAGACAT		
				PPP3CA5R1	L	AAGGCATCCATGCAGGCATCATATA		
				PPP3CA5R2	Н	GGCAGTCAAAGGCATCCATGCAGGC		
				PPP3CAPelF	I	CGTGGGAATCATGAATGTAGACATCT		
	Ribosomal protein L3	5	RPL3int5	RPL35F	M	AAGAAGTCYCACCTCATGGAGAT		
				RPL36R	J	TTRCGKGGCAGTTTCTTTGTGTGCCA		
				RPL36RA	N	AGTTTCTTTGTGTGCCAACGGCTAG		
				RPL3intF	P	AGTCTTTGGCCAGGATGAAATG		
				RPL3intR	P	TCACACCTAGGAGGGATAATG		
	Ribosomal protein L9	4	RPL9int4	RPL94F	O	CGTGTKGACAAATGGTGGGGTAA		
				RPL95R	O	ATGGGAAAGTGAGCRTACACAGA		
		5	RPL9int5	RPL95F	J	TCTGTGTAYGCTCACTTTCCCAT		
				RPL96R	J	AGAATCAGYTCRTCTTTCTGGGCTTG		
		6	RPL9int6	RPL96F	J	CAGAAAGAYGARCTGATTCTTGAAGG		
				RPL97R	J	TACWGTGGTGGCYTGCTGGATCAAGG		
				RPL96intF	N	TGTACAGGTCAAGTGTTATC		
				RPL96intR	N	ATGCCAGTTAAAAATCAGACC		

^a Intron number is based on the nomenclature used in humans

markers developed for *Pelodytes* were amplified in other genera: *PPP3CAint4* and *RPL9int4* in *Alytes*, *RPL3int5* in *Rhinella* and *Phyllomedusa*, and *RPL9int6* in *Rhinella*. In

some cases, in order to optimize amplification and/or sequencing, new primers, specific for the genus in question, were designed (also reported in Table 1).



b Legend for primer utility: A used for routine amplification and sequencing in all squamates, B used for amplification only in Phelsuma and Urocotyledon (also amplifies Lacerta and Podarcis but was not used routinely), C used for routine amplification in Lacerta and Podarcis and for sequencing in Urocotyledon and Phelsuma, D internal primer used only for sequencing, E used for amplification, F internal primer used for sequencing in Lacerta and Podarcis, G used for some amplifications and sequences, but not routinely, H used for amplification only (in Pelodytes), I used only for sequencing in Pelodytes (also amplification and sequencing in Alytes, L used only for sequencing in Pelodytes and for both amplification and sequencing in Alytes, M used for amplification and sequencing in Pelodytes and Rhinella, N used routinely for amplification and sequencing in Rhinella, O used routinely for amplification and sequencing in Pelodytes and Alytes, P used routinely for amplification and sequencing in Phyllomedusa

Table 2 Summary statistics for the markers developed in this study

Group	Genus	Marker	N (per species or phylogroup) ^a	Length (bp)	S	π	h	Hd	Rm	Indels (size)
Squamates	Lacerta	CTSDint8	9 (5/4)	688–689	6	0.00177	7	0.793	0	1 (M)
		MC1R	10 (6/4)	614	3	0.00185	3	0.416	0	0
		NFYCint16	12 (4/8)	903-919	10	0.00387	7	0.793	1	1 (16)
		PKM2int5	12 (7/5)	376-401	2	0.00065	3	0.236	0	2 (23/2)
		RELNint61	17 (9/8)	721-723	10	0.00559	6	0.738	1	1 (2)
	Podarcis	CTSDint8	10 (5/5)	688-695	30	0.01423	14	0.968	9	4 (1/6/4/2)
		MC1R	9 (6/3)	649	18	0.00677	10	0.876	2	0
		NFYCint16	5 (5/0)	797	3	0.00196	4	0.778	1	0
		PKM2int5	8 (5/3)	352	12	0.00999	10	0.942	1	0
		RELNint61	7 (4/3)	860-897	34	0.01034	11	0.967	3	6 (12/1/22/M ^b /M ^b /1)
	Mabuya	MC1R	10 (5/5)	633	15	0.01148	6	0.805	1	0
		PKM2int5	13 (13/0)	418	3	0.00169	3	0.514	0	0
		RELNint61	10 (5/5)	914-921	11	0.00394	7	0.832	0	5 (7/5/11/3/1)
	Urocotyledon	MC1R	10 (5/5)	621	6	0.00305	8	0.836	1	0
		NFYCint16	6 (2/4)	220	17	0.02624	9	0.955	3	0
		PKM2int5	5 (2/3)	735–736	9	0.00577	5	0.822	1	1 (1)
		RELNint61	10 (5/5)	813-816	39	0.00950	18	0.989	4	4 (1/2/2/1)
	Phelsuma	MC1R	10 (5/5)	617	13	0.00715	12	0.937	2	0
		NFYCint16	5 (4/1)	172	2	0.00329	3	0.511	0	$1 (M^b)$
		PKM2int5	10 (5/5)	520-521	14	0.00673	10	0.837	0	1 (1)
		RELNint61	10 (5/5)	875-877	24	0.00613	14	0.947	1	$3 (M^b/1/1)$
Anurans	Pelodytes	CHERPint7	8 (3/2/1/2)	676-698	13	0.00875	5	0.808	0	2 (1/22)
		PPP3CAint4	14 (5/4/3/1)	956–1,428	34	0.01085	10	0.881	0	11 (17/115/23/20/288/M ^b / 22/1/7/4/1)
		RPL3int5	9 (2/2/3/2)	636-639	20	0.00959	10	0.902	1	2 (1/2)
		RPL9int4	3 (1/1/0/1)	660-681	15	0.00835	3	0.733	0	8 (11/2/2/4/1/1/2/4)
		RPL9int5	8 (2/2/3/1)	571-575	14	0.01019	7	0.75	2	2 (2/2)
	Alytes	PPP3CAint4	8 (1/1/1/1/1/1/1/1)	691-700	68	0.02700	9	0.942	2	11 (4/1/4/4/2/3/1/2/7/1/1)
		RPL9int4	8 (1/1/1/1/1/1/1)	491–497	26	0.01873	8	0.933	2	$8 (1/1/1/2/2/M^b/4/5)^c$
	Rhinella	RPL3int5	8 (2/6/0)	673–676	6	0.00370	5	0.775	0	1 (3)
		RPL9int6	6 (2/3/1)	497-501	12	0.00924	6	0.848	0	1 (4)
	Phyllomedusa	RPL3int5	5(3/2)	530-532	33	0.02935	6	0.889	0	1 (2)

From left to right: N, number of individuals analysed in this study; Length, size of the sequences in the final alignment used; S number of segregating sites, π nucleotide diversity, h number of haplotypes, Hd nucleotide diversity, Rm minimum number of recombination events, Indels number of insertion/deletion polymorphisms found. Summary statistics were calculated excluding indels

After preliminary amplification and sequencing essays, we resequenced at least five individuals from each genus in order to evaluate diversity levels of the newly developed markers across taxa. All the sequences generated for this study were submitted to GenBank (accession numbers GU180920–GU181194). We used both statistical and

laboratorial procedures to determine haplotype phase: for a first assessment, we used the Bayesian algorithm implemented in the software PHASE v2.1 (Stephens et al. 2001). Runs consisted of 1,000 iterations of burn-in and 1,000 main iterations, with a thinning interval of 1. For each data set, we repeated the runs five times with different random



^a The number of individuals per phylogroup or species is represented in parenthesis; these correspond to phylogroups in the case of *L. schreiberi*, *U. inexpectata*, *Pelodytes* and *R. marina* and distinct species in all other cases, except *Alytes* where both distinct species and subspecies within *A. obstetricans* were analysed

^b M means a repeat region consisting of the repetition of a single base and where actual indel length was variable across the data set; we consider two such motifs in tandem as a single repeat region. Because in these regions it is difficult to evaluate repeat number, the reported sequence length considers only the larger apparent size for the repeat region observed

^c This locus also included one inversion spanning 25 bp

seeds. We used a threshold of 0.80 posterior probability to accept a given reconstruction. Second, for a subset of the samples for which PHASE could not reconstruct haplotypes, PCR fragments were cloned using the pGEM-T-Easy Vector Systems kit (Promega). A minimum of 5 clones per sample were sequenced. Finally, we re-ran PHASE using the information from the reconstructed haplotypes. Phased data sets were trimmed and aligned by eye using BioEdit v7.0.9.0 (Hall 1999) and imported to DNAsp v5.00.7 (Librado and Rozas 2009) for polymorphism analysis. We calculated standard summary statistics and the minimum number of recombination events (Rm, Hudson and Kaplan 1985) (see Table 2).

All the markers, even short introns, showed polymorphism in every genus, with some particular genes showing remarkably high levels of variability. We may conclude that the markers we developed can be used within a broad taxonomic scope. Moreover, they offer substantial polymorphism levels for both population genetics and phylogenetic approaches. We expect that these markers can therefore provide an important tool for the study of the evolutionary history and conservation genetics of reptiles and amphibians.

Acknowledgments This work was partially financed by Fundação para a Ciência e a Tecnologia (FCT) research projects POCI/BIA-BDE/60911/2004, PTDC/BIA-BDE/69769/2006 and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and FCT joint project CAPES/FCT 244/09. CP, HG and FS are supported by post-doctoral fellowships SFRH/BPD/28869/2006, SFRH/BPD/26555/2006 and SFRH/BPD/27134/2006, respectively and SR is supported by pre-doctoral grant SFRH/BD/17541/2004, all from FCT. BMC and TOB are supported by the research grants (BI/PTDC/BIA-BDE/69769/2006 and BI/PTDC/BIA-QOR/71492/2006, respectively). MV is supported by a post-doctoral grant CAPES 1362-07-0. CFBH thanks Fundação de Amparo à Pesquisa do Estado de S. Paulo and Conselho Nacional de Desenvolvimento Científico e Tecnológico for financial support. We further thank Raquel Godinho and Miguel Tejedo for providing tissue samples.

References

- Avise JC (2000) Phylogeography: the history and formation of species. Harvard University Press, Cambridge
- Ballard JWO, Whitlock MC (2004) The incomplete natural history of mitochondria. Mol Ecol 13:729–744

- Dolman G, Phillips B (2004) Single copy nuclear DNA markers characterized for comparative phylogeography in Australian Wet Tropics rainforest skinks. Mol Ecol Notes 4:185–187
- Edwards SV, Beerli P (2000) Perspective: gene divergence, population divergence, and the variance in coalescence time in phylogeographic studies. Evolution 54:1839–1854
- Gibbons JW, Scott DE, Ryan TJ, Buhlmann KA, Tuberville TD, Metts BS, Greene JL, Mills T, Leiden Y, Poppy S, Winne CT (2000) The global decline of reptiles, déjà vu amphibians. Bioscience 50:653–666
- Godinho R, Mendonça B, Crespo EG, Ferrand N (2006) Genealogy of the nuclear b-fibrinogen locus in a highly structured lizard species: comparison with mtDNA and evidence for intragenic recombination in the hybrid zone. Heredity 16:454–463
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp Ser 41:95–98
- Hoegg S, Vences M, Brinkmann H, Meyer A (2004) Phylogeny and comparative substitution rates of frogs inferred from sequences of three nuclear genes. Mol Biol Evol 21:1188–1200
- Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 111:147–164
- Irwin DE (2002) Phylogeographic breaks without geographic barriers to gene flow. Evolution 56:2383–2394
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451– 1452
- Martínez-Solano I, Gonçalves H, Arntzen JW, García-París M (2004) Phylogenetic relationships and biogeography of midwife toads (Discoglossidae: Alytes). J Biogeogr 31:603–618
- Paulo S, Dias C, Bruford MW, Jordan WC, Nichols RA (2001) The persistence of Pliocene populations through the Pleistocene climatic cycles: evidence from the phylogeography of an Iberian lizard. Proc R Soc B 268:1625–1630
- Pinho C, Ferrand N, Harris DJ (2006) Reexamination of the Iberian and North African Podarcis (Squamata: Lacertidae) phylogeny based on increased mitochondrial DNA sequencing. Mol Phylogenet Evol 38:266–273
- Rosenblum EB, Hoekstra HE, Nachman MW (2004) Adaptive reptile color variation and the evolution of the McIr gene. Evolution 58:1794–1808
- Saint KM, Austin CC, Donnellan SC, Hutchinson MN (1998) C-mos, a nuclear marker useful for squamate phylogenetic analysis. Mol Phylogenet Evol 10:259–263
- Sánchez-Herraíz MJ, Barbadillo LJ, Machordom A, Sanchiz B (2000) A new species of pelodytid frog from the Iberian Peninsula. Herpetologica 56:105–118
- Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 68:978–989
- Stuart SN, Chanson JS, Cox NA, Young BE, Rodrigues ASL, Fischman DL, Waller RW (2004) Status and trends of amphibian declines and extinctions worldwide. Science 306:1783–1786

