A suite of potentially amplifiable microsatellite loci for ten reptiles of conservation concern from Africa and Asia

Article *in* Conservation Genetics Resources · April 2016

DOI: 10.1007/s12686-016-0534-y

CITATIONS
2

READS
176

6 authors, including:

Kyle J Shaney
Universidad Nacional Autónoma de México
24 PUBLICATIONS **293** CITATIONS

SEE PROFILE

Richard H Adams
Florida Atlantic University
42 PUBLICATIONS **281** CITATIONS

SEE PROFILE

Nia Kurniawan
Brawijaya University
59 PUBLICATIONS **211** CITATIONS

SEE PROFILE

Amir Hamidy
Indonesian Institute of Sciences
144 PUBLICATIONS **497** CITATIONS

SEE PROFILE

Conservation Genet Resour
DOI 10.1007/s12686-016-0534-y

METHODS AND RESOURCES ARTICLE

A suite of potentially amplifiable microsatellite loci for ten reptiles of conservation concern from Africa and Asia

Kyle J. Shaney[1] · Rich Adams[1] · Nia Kurniawan[3] · Amir Hamidy[2] · Eric N. Smith[1] · Todd A. Castoe[1]

Received: 25 January 2016 / Accepted: 31 March 2016

Abstract Here we document the addition of thousands of potentially amplifiable microsatellite loci (PALs) and associated primer sequences for public use. We conducted whole-genome shotgun sequencing to obtain ultra-low coverage, random genomic sampling from ten African and Asian squamate reptile species (representing ten genera). We used unique genomic processing methods and generated PALs for the following species: *Acrochordus granulatus*, *Ahaetulla prasina*, *Cerberus rhynchops*, *Gonocephalus kuhlii*, *Ophiophagus hannah*, *Python bivittatus*, *Tribolonotus gracilis*, *Trimeresurus sabahi (Popeia sabahi)*, *Uromastyx geryi* and *Varanus exanthematicus*. All taxa included, as well as other related taxa not included in our study, are exploited heavily by the international skin and pet trades, yet researchers and conservation agencies currently lack substantial genetic resources for guiding conservation and management. Using stringent filtering methods, we generated between 467 and 8641 PALs for each of the ten species (52,164 PALs total), yielding a rich database of microsatellite loci and amplification primers for these taxa. In addition to the ten species targeted in our study, microsatellite markers provided in this database can likely be applied to a variety of closely related taxa that are also of conservation and commercial interest.

Keywords Conservation genetics · Indonesia · PALs · Simple sequence repeats · Squamate reptiles

Introduction

The trade of reptiles is now a billion dollar industry worldwide and the exploitation of reptiles for meat, pets, and skins has led to the decline of numerous species globally (Natusch and Lyons 2012; Koch et al. 2013). Because reptilian diversity is poorly understood in many regions of the world, assessing population health and regulating harvest quotas for species of conservation concern can be quite challenging (Shaney et al. in press). Exploitation of reptiles is particularly an issue in African and Asian countries where few resources are available for conservation and wildlife management initiatives. Many species collected and traded from these regions are often erroneously labeled as "farmed" or "captive bred", despite being wild caught (Koch et al. 2013). Additionally, rare and "at-risk" species are commonly misrepresented and traded under common species designations, and limits on permitted collections are often greatly exceeded and rarely enforced (Shaney et al. in press). One major hurdle in addressing these problems is the inability of regulatory agencies to identify the precise geographic location and population origins, or even the precise species designation, of traded animals.

In recent years, researchers have applied molecular 'conservation forensic' approaches that employ population genetic markers to assist in the detection of illegally traded wildlife (Iyengar 2014). For example, Wasser et al. (2004)

✉ Kyle J. Shaney
kjshaney@uta.edu

[1] The Amphibian and Reptile Diversity Research Center and Department of Biology, University of Texas at Arlington, 501 S. Nedderman Drive, Arlington, TX 76010, USA

[2] Laboratory of Herpetology, Museum Zoologicum Bogoriense, Research Center for Biology, Indonesian Institute of Sciences-LIPI, Jl. Raya Jakarta Bogor km 46, Cibinong, West Java 16911, Indonesia

[3] Department of Biology, Universitas Brawijaya, Jl. Veteran, Malang, East Java 65145, Indonesia

Published online: 05 April 2016

used genotypic data to infer the origin of illegally traded ivory in Africa, and Ghobrial et al. (2010) utilized molecular markers to trace the origins of illegally hunted chimpanzees in Cameroon. These conservation forensic methods thereby provide important management and enforcement tools that allow officials to determine the precise identification and origins of harvested animals. Importantly, population genetic data (such as microsatellites) provide an invaluable resource for assessing genetic diversity and health of target species, which can be further used to inform wildlife management guidelines and conservation legislation. Such molecular approaches have not been effectively applied to many reptilian species of conservation concern, in part due to a lack of genetic resources and molecular markers available for reptile species of interest.

Here, we document and provide thousands of potentially amplifiable microsatellite loci (PALs) and associated primer sets designed from whole-genome data for ten squamate reptile species that represent African and Asian lineages commonly collected for the reptile trade and/or are of conservation concern. These species include: *Acrochordus granulatus* (Little File Snake, from India through southeast Asia, IUCN = Least Concern), *Ahaetulla prasina* (Asian Green Vine Snake, from throughout the Sunda Region of southeast Asia, IUCN = Least Concern), *Cerberus rhynchops* (Bockadam Water Snake/Dog-faced Water Snake, from Australia and southeast Asia, IUCN = Least Concern), *Gonocephalus kuhlii* (Kuhl's Forest Dragon, from Indonesia, IUCN = Not Assessed), *Ophiophagus hannah* (King Cobra, from southeast Asia, IUCN = Vulnerable), *Python bivittatus* (Burmese Python, from southeast Asia, IUCN = Vulnerable), *Tribolonotus gracilis* (Red-Eyed Crocodile Skink, from Indonesia and Papua New Guinea, IUCN = Least Concern), *Trimeresurus sabahi (or Popeia sabihi,* Asian Pitviper, from throughout southeast Asia, IUCN = Least Concern), *Uromastyx benti* (Spiny-tailed Lizards, from north Africa and the Middle-East, IUCN = Least Concern), and *Varanus exanthematicus* (Savannah Monitor, from central Africa, IUCN = Least Concern). These ten species include multiple "at-risk" taxa and five CITES Appendix II species (*C. rhynchops, O. hannah, P. bivittatus, U. benti* and *V. exanthematicus)*–all of which have been largely overlooked by conservation initiatives and wildlife management programs. These ten species are traded globally and are illegally traded in many instances, yet our current understanding of population viability of these taxa is either completely lacking or minimal at best (Knapp 2004; Schoppe 2009; Koch et al. 2013, http://trade.cites.org/). The five species mentioned which are not currently listed by CITES are traded in high volume and are likely experiencing similar conservation issues, but have yet to receive assessment by CITES. Despite the high economic value and vulnerability of these species, very little data is currently available that can be used for conservation efforts. Our aim here was to provide a large database of molecular markers (microsatellite loci and associated primer pairs) that we have identified through analysis of genome-scale sequencing efforts that can be used for population genetic monitoring and molecular forensic identification programs for these and related taxa.

Beyond application to these ten target species, previous studies indicate that these markers will be further applicable to other related species of reptiles that are relatively closely related to the target species sampled. For example, Castoe et al. (2012a, b) tested microsatellite applicability across closely related coral snake species of the genus *Micrurus* and showed that nearly half of their tested PALs (using the same methods described here) were readily amplifiable across multiple species within the same genus. Other studies have investigated the application of microsatellites across taxa and found that a high proportion of loci successfully cross-amplified between related species (Alacs et al. 2009) and even related genera (Zucoloto et al. 2006). Therefore, we expect that many PALs provided in our dataset are likely to work well across multiple related species of interest and have broader utility for a variety of evolutionary and ecological studies. Additionally, recent studies have highlighted extensive and rapid evolution of genomic microsatellite composition across squamate lineages (Adams et al. in press; Castoe et al. 2011, 2013). Thus, microsatellite loci obtained from these ten species will not only provide valuable resources for conservation research, but also contribute to the growing set of genomic surveys of microsatellites for reptiles and vertebrate genomes in general.

Methods

Tissue samples were collected from muscle and/or liver tissue and DNA was extracted using the Phenol-Chloroform-Isoamyl alcohol (PCI) approach. Genomic DNA was fragmented using the nebulization method (Sambrook and Russell 2006). Genomic shotgun libraries were constructed using the NEBNext Illumina DNA library Prep Kit, which includes end-repairing, poly(A)-tailing, and barcoding the samples for pooled sequencing. Libraries were size-selected using the BluePippen prep for a length of 450–550 bp, and size-selected libraries were amplified via PCR. Genomic libraries were sequenced on the Illumina MiSeq platform using paired-end reads with a forward read length of 300–370 bp and reverse read length of 200–250 bp.

Post sequencing, raw genomic reads were computationally overlapped and merged into single, longer reads

Table 1 Sample handling information

Genus	Voucher	Sample details	Tissue	Geographic origin	Lat	Long
Acrochordus	UTA 61631	One adult	Muscle	Negros Island, Philippines	NA	NA
Ahaetulla	UTA 63181	One adult	Liver	Unknown	NA	NA
Cerberus	ENS 7539	One adult	Muscle	Lampung, Sumatra, Indo.	−5.573333	105.2005
Gonocphalus	ENS 7655	One adult	Muscle	Sumatera Selatan, Sumatra, Indo.	−5.2787	104.56198
Tribolonotus	UTA 63184	One adult	Liver	Unknown	NA	NA
Trimeresurus	MBH 5564	One adult	Liver	Singawang, West Kalimantan, Indo.	NA	NA
Uromastyx	UTA 63183	One adult	Liver	Unknown	NA	NA
Varanus	UTA 63182	One adult	Liver	Unknown	NA	NA

Table 2 Illumina sequencing results, including total reads after sequencing, file names, the number of reads before and after filtering and average read length (respectively)

Genus	Total combined reads	Filename_R1	Filename_R2	Read post-merge/trim	Avg. read length post trim
Acrochordus	2,698,706	A_gran_R1.fastq.gz	A_gran_R2.fastq.gz	1,290,852	453
Ahaetulla	3,487,320	A_pras_R1.fastq.gz	A_pras_R2.fastq.gz	156,346	318
Cerberus	2,368,578	C_rhyn_R1.fastq.gz	C_rhyn_R2.fastq.gz	871,869	443
Gonocephalus	2,037,934	G_kuhl_R1.fastq.gz	G_kuhl_R2.fastq.gz	30,893	346
Tribolonotus	2,604,866	T_grac_R1.fastq.gz	T_grac_R2.fastq.gz	1,079,445	444
Trimeresurus	2,367,988	T_saba_R1.fastq.gz	T_saba_R2.fastq.gz	56,294	320
Uromastyx	1,998,784	U_geyr_R1.fastq.gz	U_geyr_R2.fastq.gz	869,660	452
Varanus	2,905,680	V_exan_R1.fastq.gz	V_exan_R2.fastq.gz	1,376,950	452

Table 3 "Single Copy PALs" identified from each species sorted by microsatellite length type

Genus	2mer	3mer	4mer	5mer	6mer	Total single copy PALs	Filename
Acrochordus	779	77	1470	98	1	2425	A_gran.txt
Ahaetulla	444	98	402	214	7	1165	A_pras.txt
Cerberus	4090	726	2979	804	42	8641	C_rhyn.txt
Gonocephalus	189	100	166	9	3	467	G_kuhl.txt
Ophiophagus	3533	1175	2175	642	217	7742	O_hann.txt
Python	546	743	4045	1334	528	7196	P_bivi.txt
Tribolonotus	5868	462	1005	11	7	7353	T_grac.txt
Popeia	1114	270	667	108	3	2162	T_saba.txt
Uromastyx	2383	1223	3302	239	36	7183	U_geyr.txt
Varanus	3750	844	2557	598	81	7830	V_exan.txt

Total "Single Copy PALs" and filename identifiers are provided for each respective species

using CLC Genomics Workbench 8.01, which was also used to quality filter reads and remove adapter sequences. For quality filtering, we set the Phred score threshold at 33+ and reads were removed from the analysis if quality was below this threshold, averaged across the read. For downstream analyses, filtered and trimmed reads were concatenated into a single file for each species, which included both merged and unmerged reads. In addition to the eight species for which we generated new genomic shotgun data for, we also identified microsatellite loci from two snake genomes that have been published previously: the King cobra (Vonk et al. 2013, NCBI AZIM00000000.1) and the Burmese python (Castoe et al. 2013, NCBI AEQU00000000.2).

The program *Palfinder v.2.7.0* (Castoe et al. 2012a, b, *Palfinder* hereafter) was used to identify microsatellite loci, and to identify PCR primer sequences flanking microsatellite loci (http://sourceforge.net/projects/palfinder/). Microsatellite

loci for which flanking primer sequences were identified are referred to as Potentially Amplifiable Loci (or PALs). We used previously published *Palfinder* thresholds (Castoe et al. 2010, 2012a, b), which included identification of 2mer tandem repeat loci if repeated for at least 6 units (12 bp total), 3mers if repeated at least 4 units (12 bp), 4mers if repeated at least 3 units (12 bp), 5mers If repeated at least 3 units (15 bp), and 6mers if repeated at least 3 units (18 bp total). Next, we sorted the *Palfinder* output file to include only PALs that are likely to target a single locus in the genome by selecting PALs in which both the forward and reverse primer sequences occurred only a single time in our genomic read data—we refer to these single-copy loci as "Single Copy PALs". These ten microsatellite primer datasets have been submitted to the Dryad Data repository. For each species, the *Palfinder* output file is sorted based on the number of times each primer pair (associated with matching PAL) was found in the raw sequence reads. Each file is further sorted based on number of tandem repeats, repeat motif size, and repeat motif sequence, respectively. Sorting these output files further may allow users to design primer sets for specific purposes, such as sorting and selecting potential PALs based on repeat length, microsatellite length type (i.e., 2mer vs. 3mer microsatellites), and other characteristics that may be desirable for population genetic markers (Table 1).

Results

We recovered a total of 52,083 "Single Copy PALs" and associated primer sets across all ten samples (Summary information available in Tables 2, 3). We included PALs primer sets designed from 2mer, 3mer, 4mer, 5mer and 6mer microsatellite loci. The numbers of total PALs recovered per microsatellite length class are listed in Table 3. Although we only present the number of highest quality markers ("Single Copy PALs", Table 3) the data files provided online include thousands of additional PALs and associated primer sets that were eliminated using the sorting methods described above, but may be useful for marker screening.

Data files were submitted to Dryad (doi:10.5061/dryad.4gb62) and are sorted based on the criteria described above. Raw read files are provided in fastq format and microsatellite files are provided as tab-delimited text files (Tables 2, 3). The latter data files report the primer information for each identified PAL, which includes: repeat motif size and sequence, total number of repeats in amplicon, primer designed (1 = yes, 0 = no), forward primer name, forward primer sequence, reverse primer name, reverse primer sequence, and the occurrence of primer pairs in both the PAL sets and genomic reads.

Acknowledgments We are grateful to representatives of LIPI at the Museum Zoologicum Bogoriense for facilitating the study of specimens from Indonesia and providing field research permits, namely Boadi, M. Amir, R. Ubaidillah, I Sidik, and Ir. R. M. Marwoto. We are grateful to the Ministry of Research and Technology of the Republic of Indonesia (RISTEK), for coordinating and granting research permissions. Pak S. Wahyono (RISTEK) provided assistance throughout the permitting process. RISTEK and LIPI reviewed and approved our fieldwork in Indonesia and provided export permits for specimens to the United States for study and deposition at UTA. W. Trilaksono, I. Sidik, and A. Ryanto provided laboratory assistance at MZB. Mr. Widodo and Marwoto from the Faculty of Mathematicas and Natural Sciences of Universitas Brawijaya (MIPA-UB) kindly provided logistical support. The Forestry Department of Indonesia provided research permits for areas under their jurisdiction. We thank the local communities for their support, advice and kindness during our travels. We thank the members of the field expeditions throughout Java and Sumatra: G. Barraza (Broward College); U. Arifin (Institut Teknologi Bandung), W. Trilaksono (MZB); C. Franklin, K. O'Connell, U. Smart, E. Wostl (UTA), and A. M. Kadafi, D. R. Wulandari, R. Darmawan, K. I. Nawie, A. Dharasa, and S. Pratassi (MIPA- UB). Funding was provided by a National Science Foundation Grant (DEB-1146324) to E. N. Smith and M. B. Harvey, and startup funds to T. A. Castoe.

References

Adams R, Blackman H, Reyes-Velasco J, Schield D, et al (2016) Microsatellite landscape evolutionary dynamics across 450 million years of vertebrate genome evolution. Genome

Alacs EA, Hillyer MJ, Georges A, Fitzsimmons JM et al (2009) Development of microsatellite markers in the Australasian snake-necked turtle *Chelodina rugusa* and cross-species amplification. Mol Ecol Resour 9:350–353

Castoe TA, Poole AW, Gu W, de Koning APJ et al (2010) Rapid identification of thousands of microsatellite loci for the copperhead snake (*Agkistrodon contortrix*) from modest amounts of 454 shotgun genome sequence. Mol Ecol Resour 10:341–347

Castoe TA, Hall K, Guibotsy Mboulas ML, Gu W, de Koning APJ et al (2011) Discovery of highly divergent repeat landscapes in snake genomes using high throughput sequencing. Genome Biol Evolu 3:641–653

Castoe TA, Streicher JW, Meik JM et al (2012a) Thousands of microsatellite loci from the venomous coralsnake (*Micrurus fulvius*) and variability of select loci across populations and related species. Mol Ecol Resour 12:1105–1113

Castoe TA, Poole AW, de Koning APJ et al (2012b) Rapid microsatellite identification from Illumina paired-end genomic sequencing in two birds and a snake. PLoS ONE 7:e30953

Castoe TA, de Koning APJ, Hall KT, Card DC et al (2013) The Burmese Python genome reveals the molecular basis for extreme adaptation in snakes. Proc Natl Acad Sci 110:20645–20650

Ghobrial L, Lankester F, Kiyang JA, Akih AE et al (2010) Tracing the orgins of rescued chimpanzees reveals widespread chimpanzee hunting in Cameroon. BMC Ecol. doi:10.1186/1472-6785-10-2

Iyengar A (2014) Forensic DNA analysis for animal protection and biodiversity conservation: A review. J Nat Conserv 22:195–205

Knapp A (2004) An assessment of the international trade in Spiny-tailed Lizards Uromastyx with a focus on the role of the European Union. TRAFFIC Europe Report

Koch A, Ziegler T, Bohme W, Auliya M (2013) Pressing problems: distribution, threats, and conservation status of the monitor lizards (Varanidae: *Varanus* spp.) of Southeast Asia and the Indo-Australian archipelago. Herpetol Conserv Biol 8:1–62

Natusch DD, Lyons J (2012) Exploited for pets: the harvest and trade of amphibians and reptiles from Indonesian New Guinea. Biodivers Conserv 21:2899–2911. doi:10.1007/s10531-012-0345-8

Sambrook J, Russell DW (2006) Fragmentation of DNA aby nebulization. In: Sambrook J, Russell DW (eds) "Commonly Used Techniques in Molecular Cloning," Appendix 8, in molecular cloning, vol 3, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor **(2001)**

Schoppe S (2009) Status trade dynamics and management of the Southeast Asian box turtle in Indonesia. TRAFFIC Southeast Asia

Shaney KJ, Wostl E, Hamidy A, Kurniawan N, et al (2016) Conservation challenges regarding species status assessments in biogeographically complex regions: examples from overexploited reptiles of Indonesia. Oryx

Vonk FJ, Casewel NR, Henkel CV, Heimberg A et al (2013) The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. Proc Natl Acad Sci 10:20651–20656

Wasser SK, Shedlock AM, Comstock K, Ostrander EA et al (2004) Assigning African elephant DNA to geographic region of origin: Applications to the ivory trade. Proc Natl Acad Sci 101:14847–14852. doi:10.1073/pnas.0403170101

Zucoloto RB, Villela PMS, Verdade LM, Coutinho LL (2006) Cross-species microsatellite amplification in South American Caimans (*Caiman spp* and *Paleosuchus palpebrosus*). Genetics and Molecular Biology 29:75–78