



OPEN

Genome-wide data implicate terminal fusion automixis in king cobra facultative parthenogenesis

Daren C. Card^{1,2,3}, Freek J. Vonk^{4,5}, Sterrin Smalbrugge⁶, Nicholas R. Casewell⁷, Wolfgang Wüster^{8,9}, Todd A. Castoe¹, Gordon W. Schuett^{9,10} & Warren Booth^{9,11}✉

Facultative parthenogenesis (FP) is widespread in the animal kingdom. In vertebrates it was first described in poultry nearly 70 years ago, and since then reports involving other taxa have increased considerably. In the last two decades, numerous reports of FP have emerged in elasmobranch fishes and squamate reptiles (lizards and snakes), including documentation in wild populations of both clades. When considered in concert with recent evidence of reproductive competence, the accumulating data suggest that the significance of FP in vertebrate evolution has been largely underestimated. Several fundamental questions regarding developmental mechanisms, nonetheless, remain unanswered. Specifically, what is the type of automixis that underlies the production of progeny and how does this impact the genomic diversity of the resulting parthenogens? Here, we addressed these questions through the application of next-generation sequencing to investigate a suspected case of parthenogenesis in a king cobra (*Ophiophagus hannah*). Our results provide the first evidence of FP in this species, and provide novel evidence that rejects gametic duplication and supports terminal fusion as a mechanism underlying parthenogenesis in snakes. Moreover, we precisely estimated heterozygosity in parthenogenetic offspring and found appreciable retained genetic diversity that suggests that FP in vertebrates has underappreciated evolutionary significance.

Asexual reproduction in otherwise sexually reproducing species—a phenomenon termed facultative parthenogenesis (FP)¹—has been reported in three major lineages of vertebrates: chondrichthyan fishes^{2–4}, birds^{5,6}, and squamate reptiles^{7,8}. Originally limited to a few instances in captivity, FP has been described as an outcome of reproductive error^{1,9}. However, reports of vertebrate FP continue to increase both in number and taxonomic diversity, including cases from the wild^{3,10}. With this accumulating information, our understanding of this reproductive mode has expanded greatly⁸, and commonalities have been detected across all vertebrate lineages shown capable of FP. Important commonalities include: (i) constraints that the sex chromosome system itself places on the sex of parthenogens^{8,11}, (ii) the capacity to switch between sexual and asexual modes^{8,12}, and (iii) the ability to produce consecutive (i.e., multiple) FP births^{5,13–18}. Recently, second generation FP (i.e., parthenogenetic reproduction by a parthenogen) has been documented in the white-spotted bamboo shark, *Chiloscyllium plagiosum*¹⁹, and observed in the royal python, *Python regius* [W. Booth, unpubl. data]. Furthermore, a parthenogenetic boa constrictor, *Boa imperator*, was found to be capable of sexual reproduction [W. Booth, unpubl. data]. Combined, these diverse data strengthen the view that FP is not an uncommon mode of reproduction in certain vertebrate taxa, and suggest that its significance in vertebrate evolution has been underestimated⁸.

The developmental mechanism responsible for nearly all cases of vertebrate FP has been ascribed to automixis, with restoration of diploidy through the fusion or duplication of meiotic products. Of the automictic modes, terminal fusion (TF) is the most commonly cited mechanism^{1,8}. Under this mode, the egg nucleus fuses

¹Department of Biology, The University of Texas Arlington, Arlington, TX, USA. ²Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA. ³Museum of Comparative Zoology, Harvard University, Cambridge, MA, USA. ⁴Naturalis Biodiversity Center, Leiden, The Netherlands. ⁵Amsterdam Institute of Molecular and Life Sciences, Division of BioAnalytical Chemistry, Department of Chemistry and Pharmaceutical Sciences, Faculty of Science, Vrije Universiteit Amsterdam, 1081HV Amsterdam, The Netherlands. ⁶Wildlife Ecology and Conservation Groups, Wageningen University, Wageningen, The Netherlands. ⁷Centre for Snakebite Research and Interventions, Liverpool School of Tropical Medicine, Liverpool, UK. ⁸Molecular Ecology and Evolution Group, School of Biological Sciences, Bangor University, Bangor, UK. ⁹Chiricahua Desert Museum, Rodeo, NM, USA. ¹⁰Department of Biology, Neuroscience Institute, Georgia State University, Atlanta, GA, USA. ¹¹Department of Biological Science, The University of Tulsa, Tulsa, OK, USA. ✉email: warren-booth@utulsa.edu

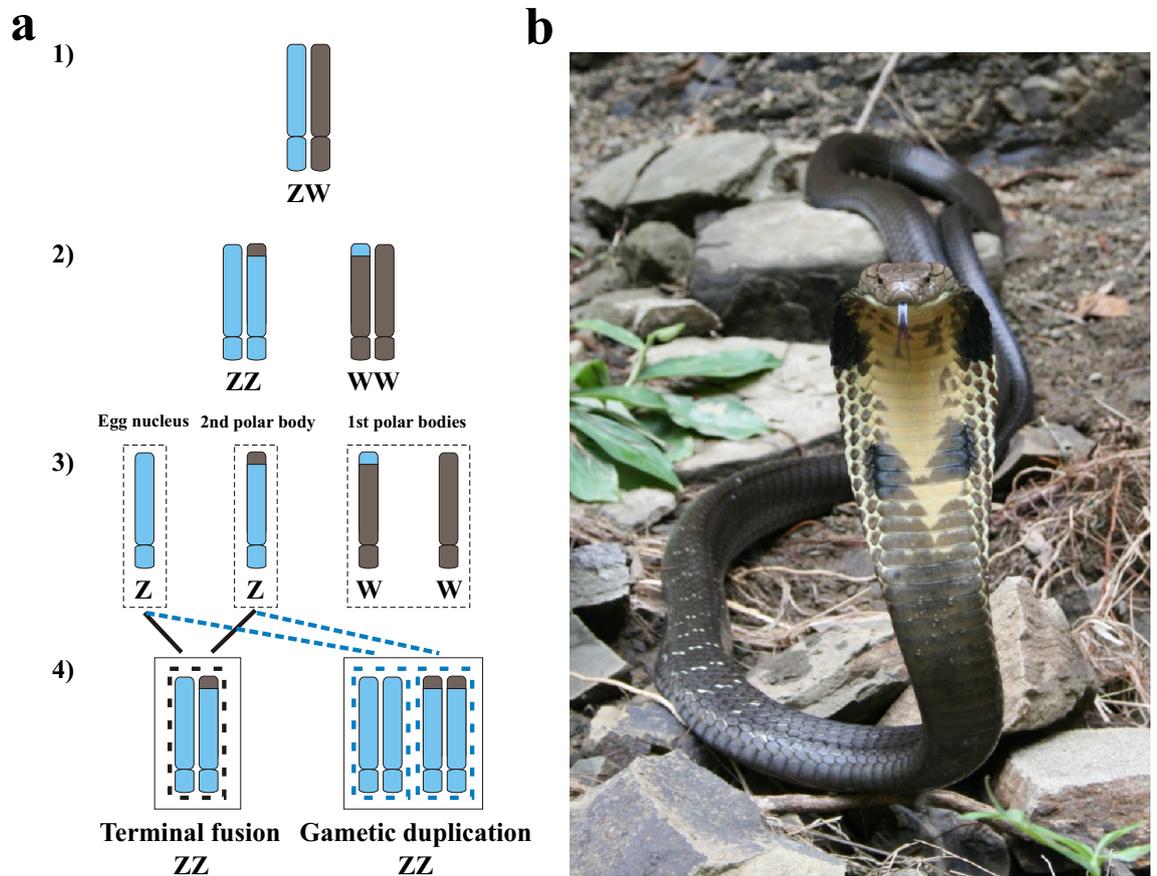


Figure 1. (a) Proposed mechanisms of automixis in snakes. (1) Primordial germ cell. (2) Meiotic products following DNA replication and recombination during the first round of cell division. (3) Meiotic products following the second round of cell division. (4) Potential sex chromosomal arrangements following terminal fusion and gametic duplication (here depicted for one Z chromosome). Note that parthenogens with a WW sex chromosome arrangement are not, at present, known to be viable. (Modified from⁵⁸). (b) Adult king cobra (*Ophiophagus hannah*). Photo courtesy of Freek Vonk.

with its second polar body, resulting in the retention of reduced levels of heterozygosity limited to the telomeric regions of chromosomes^{4,8} (Fig. 1a). Few studies, however, provide supporting evidence, i.e., heterozygosity in FP progeny^{17,18,20}. Gametic duplication (GD), in contrast, is predicted to result in genome-wide homozygosity^{1,8} (Fig. 1a). Accurately assigning which mechanism underlies FP in a given species has proven challenging, largely owing to limitations of the molecular markers employed (mainly microsatellites). In such studies, a dramatic reduction in heterozygosity has been shown in parthenogens, relative to their mother, with most parthenogens exhibiting homozygosity across all loci surveyed^{2,3,7,12–17,21–25}. Consequently, due to levels of heterozygosity in the parthenogens, some authors infer GD²⁵, whereas others propose TF^{2,8,17,18}. A key issue with the conclusions of these studies is that there is limited power to detect low levels of retained heterozygosity expected under TF with only a few markers employed. Thus, high false positive rates may result for identifying GD over TF. Unfortunately, most published studies on the phenomenon of FP have not been sufficiently robust to accurately differentiate between these two developmental mechanisms and provide little opportunity to examine secondary questions focused on how genetic variation is inherited in parthenogenetic offspring, and the impact parthenogenesis may have on the evolutionary trajectories of natural populations^{8,25}.

Arguably, with such indeterminate outcomes, understanding which mechanism of automixis is operating has important implications on the levels of genetic diversity retained in parthenogenetic offspring, which in turn has major ramifications for understanding the impact that these parthenogens might have on the genetic diversity of natural populations. This consideration is particularly significant when the status of a species, or population, is threatened or endangered^{3,7}. These concerns are further amplified by the findings that FP may be heritable^{5,6}, which could lead to greater incidence of this phenomenon in small populations as a result of inbreeding³. Despite advances that recent studies have made towards understanding FP in vertebrates⁹, the fundamental questions of: (i) resolving which automictic mechanism(s) underlie FP, and (ii) how FP impacts individual-level genetic diversity on a genome-scale, remain unresolved.

Compared to other vertebrate lineages, recent studies in snakes have generated novel and field-advancing results that have significantly contributed to our understanding of FP^{8,11,26}; thus, snakes represent an ideal model system with which to address outstanding questions. Following the publication of its genome²⁷, the king cobra,

| Sample | Mapped PE reads | SNPs | Mean (\pm SD) heterozygosity | | Mean (\pm SD) pairwise relatedness | | |
|---------------------------|-----------------|------|---------------------------------|---------------------|---------------------------------------|----------------------|----------------------|
| | | | OH | HL | Mother | Offspring #1 | Offspring #2 |
| Mother (Cobra UK) | 28,844,748 | 401 | 0.011 (\pm 0.004) | 0.30 (\pm 0.027) | – | 0.99 (\pm 0.0004) | 0.99 (\pm 0.0004) |
| Offspring #1 (Cobra211) | 2,696,748 | 373 | 0.007 (\pm 0.005) | 0.53 (\pm 0.027) | 0.99 (\pm 0.0004) | – | 0.99 (\pm 0.0005) |
| Offspring #2 (CobraFV807) | 24,027,342 | 328 | 0.007 (\pm 0.001) | 0.52 (\pm 0.029) | 0.99 (\pm 0.0004) | 0.99 (\pm 0.0005) | – |

Table 1. Measures of the number of paired-end reads mapped to the king cobra, *O. hannah*, reference genome, the number of SNPs vs. the reference genome in final thinned genotype dataset, mean heterozygosity (OH observed heterozygosity, HL homozygosity by loci), and mean pairwise relatedness for all samples. Values above the diagonal in the relatedness matrix represent B_{xy} , while values below the diagonal represent M_{xy} . SD standard deviation.

Ophiophagus hannah (Fig. 1b), has emerged as a new model system for studying multiple aspects of snake evolution, development, and venom. This iconic species, distributed widely across South and Southeast Asia, is the world's longest venomous snake, attaining lengths up to 5 m. Owing to human persecution and habitat destruction, it is of high conservation concern throughout its natural range, and IUCN-listed as 'Vulnerable'²⁸. In an effort to address the outstanding questions in vertebrate FP, we used the king cobra as a model to explore genetic patterns of tens of thousands of loci from a virgin king cobra female and her two male offspring suspected of being parthenogens. Restriction Associated DNA sequencing (RADseq) data provide an unprecedented, genome-scale view of the impact of FP on genetic diversity and inheritance, confidently providing support for a terminal fusion-based mechanism of development that results in the maintenance of appreciable levels of heterozygosity in parthenogenetic offspring.

Materials and methods

Subject history and sampling.

The mother (Cobra UK) was an adult, originally imported to Germany from Sumatra as a 70 cm juvenile in September 2012. Kept in isolation, on 3 August 2015, at ~200 cm, she laid 24 eggs. Two eggs were viable, determined through the detection of blood vessels following egg candling, and were artificially incubated in a Jaeger incubator at 27–28 °C in a substrate consisting of moist vermiculite. One egg was opened on 26 October 2015, revealing a male offspring (offspring #1—Cobra211), which died two days later. The second egg, opened on 12 November 2015, also contained a male offspring (offspring #2—CobraFV807), which died within three days. Both fetuses were sexed based on morphology and were undersized (offspring 1—27.7 cm; offspring 2—20.7 cm, Supplementary Figure 1a) compared to an average expected hatchling size of ~50 to 55 cm²⁹. In cases of facultative parthenogenesis (FP) involving several lineages of colubroid snakes, particularly natricines (e.g., genera *Nerodia* and *Thamnophis*) developmental abnormalities to both the head region and or body are not uncommon^{8,17,21,30}. These deformities include, but are not limited to, incomplete development of the brain, impartial fusion of the cranium, uni- or bilateral absence of one of both eyes (anophthalmia), loss or the addition of body scales, scoliosis of the vertebral column. The craniofacial abnormalities observed in the neonate king cobras are similar to those described for the checkered garter snake (Supplementary Figure 1b⁸). Allen et al.²⁰ reported similar post-cranial developmental abnormalities (e.g., vertebral scoliosis) associated with FP in an elapid species, the death adder (*Acanthophis antarcticus*). DNA was extracted from shed skin from the mother and muscle tissue from the offspring using standard phenol–chloroform methods. All extractions were quantified using Qubit broad-range DNA assays (Thermo Fisher Scientific) following the manufacturer's instructions. *Ophiophagus hannah* is a widespread species with high morphological and genetic variability^{31,32}, and likely constitutes a species complex. A taxonomic revision is in progress (P. Gowri Shankar et al., in prep). To future-proof the identifiability of our specimen in relation to future taxonomic changes, we Sanger-sequenced parts of the mitochondrial genes for NADH dehydrogenase subunit 4 (ND4) and 16S ribosomal RNA of the mother (as used by Gowri Shankar et al.) as barcode vouchers. For primers and PCR conditions see Maddock et al.³³.

Genomic library generation and data processing.

A modified version of the Peterson et al.³⁴ protocol was used to prepare double digest RADseq libraries. Genomic DNA was simultaneously cut with both rare *Pst*I (6 bp) and common *Sau*3AI (4 bp) restriction enzymes, estimated to target approximately 200,000 genomic loci per individual. To allow for hierarchical pooling and multiplexing of samples, barcoded Illumina adapter oligonucleotides were ligated to the ends of digested DNA. These adapters also included 8 bp unique molecular identifiers (UMIs; 8 consecutive random nucleotides prior to the ligation site). Following adapter ligation, samples were pooled, and these pools were size selected for a range of 430–600 bp using a Blue Pippin (Sage Science). After size selection, samples were PCR-amplified with pool-specific indexed primers, and amplification products were further pooled into a single sample based on molarity calculations from analysis on a Bioanalyzer (Agilent Technologies) using a DNA 7500 chip. The final pooled libraries were sequenced on an Illumina HiSeq 2000 lane, resulting in at least 2 million reads per sample (Table 1).

Read processing and genotyping.

Raw Illumina sequence data were filtered to remove PCR clones using the adapter UMIs and the clone_filter tool from the Stacks v. 1.35 analysis pipeline^{35,36}. Samples were parsed using the process_radtags tool from Stacks, with restriction sites and barcodes rescued using default

settings, and Trimmomatic v. 0.32³⁷ was used to quality filter the resulting data using the settings LEADING:10 TRAILING:10 SLIDINGWINDOW:4:15 MINLEN:36. An iterative strategy was used to focus on genetic variation sampled from high-quality RAD loci present in all individuals. Quality-trimmed reads were mapped to the *Ophiophagus hannah* genome (NCBI version OphHan1.0²⁷) using the mem algorithm in BWA v. 0.7.12-r1039³⁸. SAMtools v. 1.3.1^{39,40} and Picard v. 1.106 were used to process mapping files for each sample and merge mappings for downstream analyses. GATK v. 3.8-0-ge9d806836^{41–43} was used to perform realignment around indels with default settings. The mapping data was then used to isolate the coordinates of RAD loci present in all individuals by identifying high quality (Phred ≥ 30) sites where read depth was at least 5 per individual, merging adjacent sites together to form loci (allowing up to 10 bp low coverage “gaps” in coordinates), and filtering away loci shorter than 100 bp. Using these RAD loci footprint coordinates, a new reference sequence from the *O. hannah* reference genome was exported using BEDtools v. 2.29.0⁴⁴ that only contained the RAD locus regions.

For genotyping, a second round of mapping was conducted and variant calling performed using the GATK best-practices guidelines^{41–43}. BWA was used to map quality-trimmed reads to the RAD locus reference and SAMtools, Picard, and GATK were used to quality control and realign around indels, as above. Variants were called using GATK HaplotypeCaller using default settings by first producing GVCF variant files for individuals before performing joint genotyping across all samples. Variants were filtered using BCFtools and VCFtools v. 0.1.15⁴⁵ as follows: (1) indels were excluded; (2) genotype calls for individuals with a read depth of less than 5 were set as missing data; (3) SNPs within 3 bp of indels and clusters of indels separated by 10 bp or less were excluded; (4) variants with a Phred quality score below 30 were excluded; (5) SNPs with significant statistical biases were removed using the hard filter ‘QD < 2 || FS > 60.0 || MQ < 40.0 || MQRankSum < - 12.5 || ReadPosRankSum < - 8’; (6) SNPs with a total depth greater than $2 \times$ average depth (= 100) and less than $0.5 \times$ average depth (= 25) were excluded; and (7) non-biallelic SNPs were excluded. We further thinned SNPs to avoid the potential effects of linkage by randomly selecting one variant per 50 kb region of the king cobra genome.

Genomic analyses of patterns of parthenogenesis. Filtering of loci (by quality, representation across individuals, and by linkage) resulted in a dataset of 20,562 independent (i.e., unlinked) RAD loci with genotypes called in the mother and both offspring. Given a genome assembly size of approximately 1.6 Gb, we expect a RAD locus density of one marker per 80 kb. The Rhh⁴⁶ package in R v. 3.4.3⁴⁷ was used to calculate two measures of genome-wide heterozygosity for each individual: proportion observed heterozygosity (OH; 0 is complete homozygosity and 1 is complete heterozygosity) and homozygosity by loci (HL⁴⁸; 0 is complete heterozygosity and 1 is complete homozygosity). Two measures of pairwise relatedness were calculated between all samples using the R package Demerelate⁴⁹: the shared alleles index (B_{xy} ⁵⁰) and genotype sharing (M_{xy} ⁵¹)—both measures vary between 0 (no relatedness) and 1 (complete relatedness or identical genotypes). A custom R function available from <https://github.com/darencard/parthenogenesis> was used to sample 100 bootstrap datasets to measure variance in each measure. The proportions of retained and lost maternal heterozygosity in each offspring was quantified and we used a binomial test to determine whether there were equal proportions of loci that are homozygous for the reference (0/0) or alternative (1/1) allele at loci where heterozygosity was lost in offspring, as is expected under a model of independent inheritance of unlinked loci. We estimated the Jaccard index to quantify patterns of retained heterozygosity shared in both offspring in the empirical dataset compared to randomly permuted datasets based on the formula $J(A, B) = \frac{|A \cap B|}{|A \cup B|}$ where A, and B represent sets of loci in the two offspring.

Results and discussion

Here we provide the first in-depth, genome-wide assessment of genetic diversity for a squamate reptile reproducing via FP, through the comparison of two parthenogenetic male king cobra siblings with their mother. The analysis of thousands of loci from throughout the genome indicates that both offspring had significantly reduced levels of heterozygosity relative to the mother, though offspring #2 retained slightly higher amounts than offspring #1 (Fig. 2a, Table 1). Both measures of relatedness fell between approximately 0.99 and 1 (Fig. 2b, Table 1), which is far above the expectation of 0.5 in pairwise comparisons between a parent and sexually produced progeny. The highly reduced level of heterozygosity, inflated values of pairwise relatedness, and offspring sex (i.e., male being the expected sex of a parthenogen in species with ZZ/ZW sex determination; see^{8,11}), strongly support the conclusion that these offspring are the product of FP, and provide the first evidence of this mode of reproduction in the king cobra.

Automictic mechanisms and evolutionary consequences of FP. Resolving the automictic mechanism underlying the development of progeny via FP in previous studies has relied on the level of homozygosity observed in the progeny. For example, terminal fusion automixis is expected to produce offspring that exhibit highly elevated levels of homozygosity relative to the mother, with the majority of heterozygous loci in the mother becoming differentially homozygous in the offspring. In contrast, central fusion automixis is predicted to produce parthenogens that retain levels of heterozygosity comparable to the mother. However, resolving the mechanism between terminal fusion and gametic duplication has proven difficult, for the latter results from the duplication of the gametes chromosomes and hence 100% homozygosity (Fig. 1a). As terminal fusion is assumed to result in low levels of heterozygosity retained at the telomeric regions, detection of heterozygosity relies on having markers in these specific regions of recombination. Previous studies of vertebrate FP have largely utilized microsatellite markers for genetic confirmation. However, the limited numbers of markers employed in each study and the low likelihood that these markers will be located in regions of recombination has resulted in parthenogens exhibiting complete homozygosity at the mother’s heterozygous loci in all but two studies^{17,18}. Accordingly, while TF has been cited as the most likely developmental mechanism⁸, GD also has been proposed²⁵.

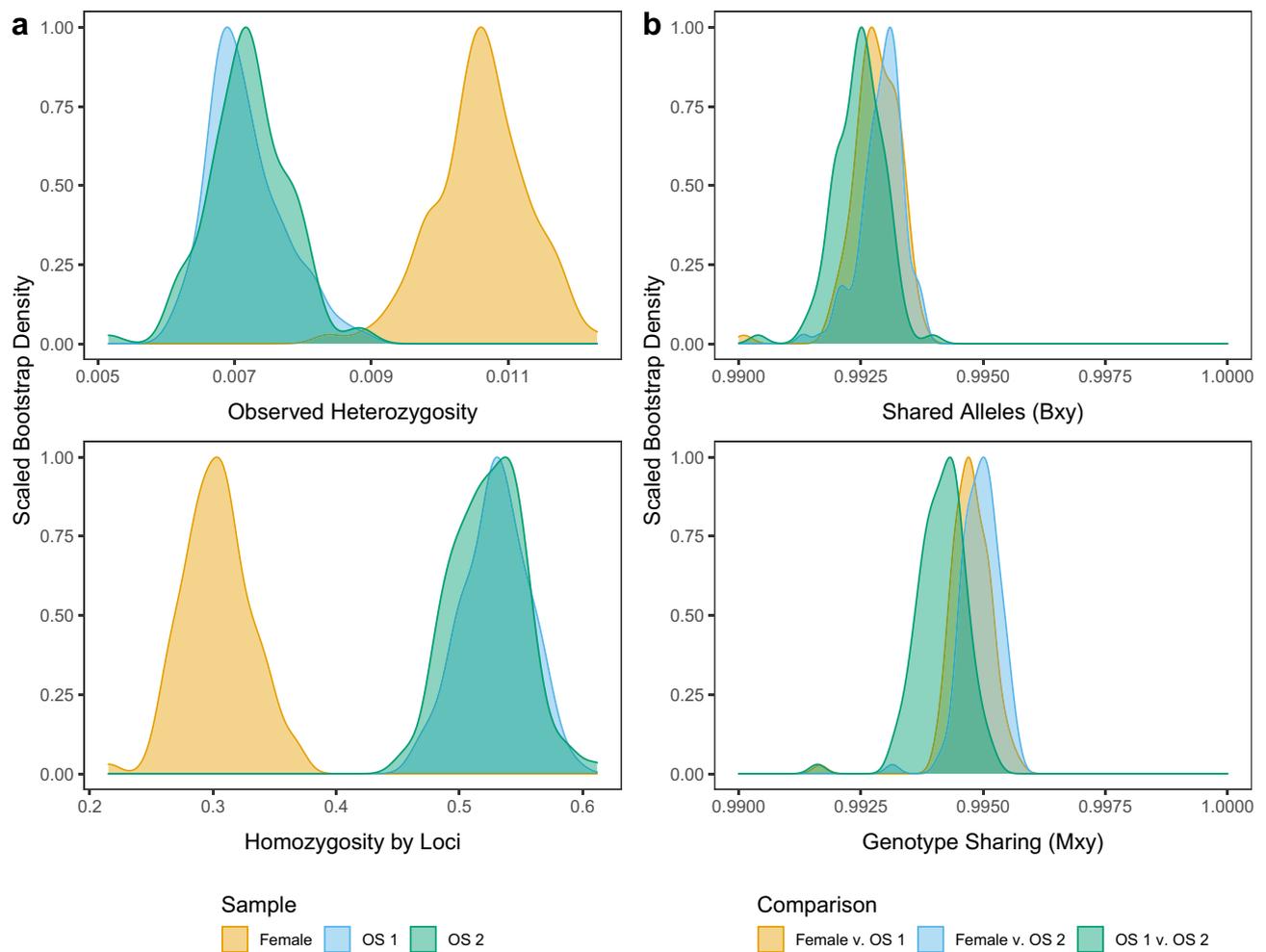


Figure 2. (a) Bootstrap densities of two measures of individual heterozygosity for each member of the cobra family (female = mother, OS 1 = offspring 1, OS 2 = offspring 2) based on 100 bootstrap replicates of 20,562 independent RAD loci: (top) Proportion of heterozygous sites (Observed Heterozygosity), and (bottom) the standardized level of homozygosity (Homozygosity by Loci). (b) Bootstrap densities of two measures of pairwise relatedness between the mother and each parthenogenetic offspring based on 100 bootstrap replicates sampled variants: (top) shared alleles index (B_{xy}) and (bottom) genotype sharing index (M_{xy}). Sexual reproduction would produce measures at 0.5.

Using RADseq, we conclusively show, through the presence of retained heterozygosity in FP offspring, that the developmental mechanism of automixis is TF over GD. Here, the detection of low, but clearly identifiable levels of heterozygosity in each offspring (Fig. 2a), is in agreement with TF. Notably, approximately a quarter of maternal heterozygosity is retained in these offspring ($N = 278$ maternal heterozygous loci; 23% retained in offspring #1 and 27% retained in offspring #2), which is greater than previous estimates of retained heterozygosity based on microsatellite markers (typically 0% retained heterozygosity^{21,22,25,52}). This discrepancy may stem from technical differences between SNP and microsatellite markers but given that non-trivial retained heterozygosity is also observed in a far more stringently filtered set of SNPs (Phred quality score > 1000; $N = 86$ maternal heterozygous loci; 20–22% retained heterozygosity), we speculate that previous studies based on small numbers of microsatellite loci had reduced power to detect retained heterozygosity. Further exploration of FP using genomic datasets is necessary to better understand how much maternal heterozygosity is retained in parthenogenetic offspring. Moreover, affirming a pattern of significant retention of heterozygosity will require reconciling our current understanding of the impact of FP on genetic diversity and the evolutionary trajectories of natural populations.

Patterns of inheritance of genome-wide heterozygosity. Genome-wide patterns of inheritance of heterozygosity have never been quantitatively compared across multiple offspring due to low numbers of profiled markers or a lack of sibling samples necessary for evaluating inheritance. Under a model of terminal fusion automixis, we expected that the modest amount of heterozygosity retained only in the telomeric regions of chromosomes would result in modest amounts of shared retained heterozygosity between the two offspring. In accordance with our expectations, while each of the offspring in this study does possess unique genomic regions with retained heterozygosity, a large percentage of retained heterozygosity is shared between offspring (38%),

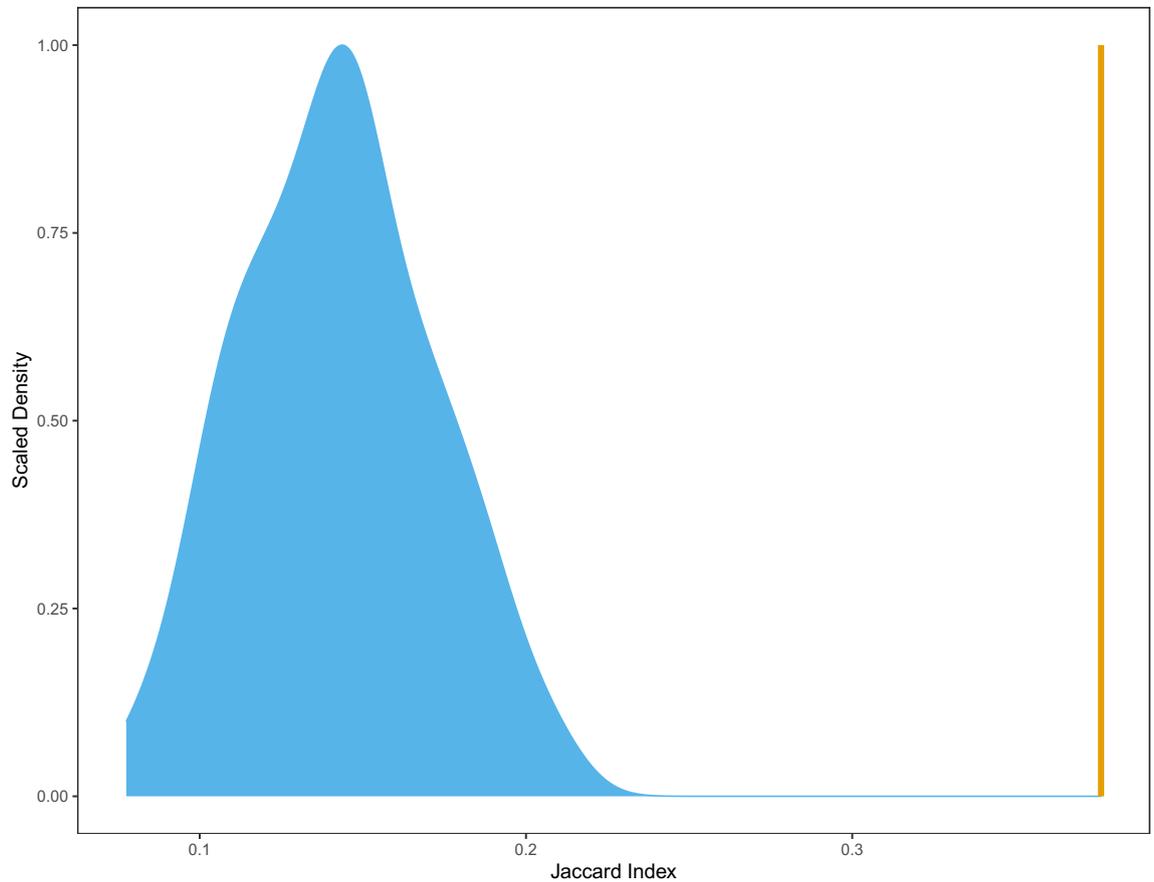


Figure 3. Scaled density of Jaccard index measures of the overlap in heterozygous sites in the two offspring, showing that the degree of overlap in the empirical dataset (in gold) is nonrandom and significantly greater than expected based on measures from 100 randomly permuted datasets (in blue).

which is greater than expected by chance (Fig. 3). The degree of shared heterozygosity between offspring confidently rules out false positive signals of retained heterozygosity due to genotyping error and is likely explained by localization of heterozygosity to similar genomic regions—likely the distal ends of chromosome arms, which are predicted to represent regions of retained heterozygosity in TF. Unfortunately, it is not possible to directly localize regions of retained offspring heterozygosity and confirm their location in the telomeric regions of chromosomes due to the low contiguity of the current king cobra genome assembly²⁷. However, more indirect measures of genomic composition may provide information on the locations of regions of retained heterozygosity, providing further support for a model of terminal fusion (TF) automixis in this species. Under the TF automixis model, GC content is expected to be inflated around areas of retained heterozygosity due to higher rates of recombination and GC-biased gene conversion in telomeric regions (Fig. 1a)^{53–55}. However, we find that GC content in regions surrounding retained heterozygosity in both offspring is statistically indistinguishable from GC content in regions that lose heterozygosity in the offspring and from permuted datasets produced from regions homozygous in the mother and both offspring (Supplementary Materials).

Terminal fusion automixis does result in the retention of maternal heterozygosity in offspring, but the majority of heterozygosity found in centromeric regions ends up purged in the genomes of offspring (Fig. 1a). Under any form of automixis, roughly equal numbers of loci that are heterozygous in the mother will segregate as homozygous genotypes of one of the parental alleles in the offspring^{1,8}. Yet, in our dataset, we observed biased patterns of inheritance in both of the offspring, where the reference allele is preferentially inherited (Table 2, Supplementary Figure 2). While these results are not immediately predicted by any existing model of automixis, this observation may be explained by some form of biased gene conversion^{53,56,57}, a process which has not been thoroughly investigated in snakes. Recent studies on the evolution of mammalian genomic landscapes supports such non-Mendelian transmission of alleles to occur in regions located near recombination sites that yield a GC bias, and thus result in an evolutionary advantage for GC over AT alleles⁵⁷. While we did not observe a bias in GC content in regions of retained heterozygosity in the king cobra genome (Supplementary Materials, Supplementary Figure 3), this may be due to the limitations of the dataset and these findings do not refute TF as the developmental mechanism underlying facultative parthenogenesis in snakes⁸. Other unknown mechanisms or technical artifacts may therefore explain the biased pattern of inheritance in the two offspring. Indeed, the large bias towards inheritance of the reference allele (i.e., the allele present in the reference genome) in both offspring is not expected under GC-biased gene conversion and is logical in a situation where there is reduced power to call alternative SNPs (alleles) due to low sequencing coverage or sample numbers. This pattern may therefore be

| Offspring | No. of 0/0 loci | No. of 0/1 loci | No. of 1/1 loci | Binomial test <i>p</i> -value |
|--------------|-----------------|-----------------|-----------------|-------------------------------|
| Offspring #1 | 150 | 64 | 64 | 3.9×10^{-9} |
| Offspring #2 | 144 | 75 | 59 | 2.2×10^{-9} |

Table 2. A tally of the fate of maternal heterozygous loci ($N = 278$) in both offspring. 0/0, 0/1, and 1/1 encode homozygous and heterozygous locus genotypes based on the reference genome (0) and alternative alleles (1). The Binomial Test *p*-value column represents the results of two-sided binomial tests for each offspring that compared the proportions of 0/0 vs. 1/1 genotypes under the assumption of an expected proportion of 0.5 for each with random inheritance of unlinked loci.

a result of a statistical artifact, but additional studies that leverage high-coverage, genome-wide data in larger numbers of samples should provide a better opportunity to further explore the biased patterns of inheritance observed in this study.

Conclusion

In summary, here we used a genome-scale dataset to study parthenogenesis in a non-traditional model vertebrate and find conclusive support for terminal fusion automixis-mediated facultative parthenogenesis in the king cobra. When considered alongside genomic data recently reported in two species of Australian elapid²⁰, and microsatellite data presented in a North American garter snake¹⁷, this likely represents the underlying developmental mechanism in all species of snakes presently reported to have reproduced through facultative parthenogenesis⁸. Our genomic dataset also illuminates patterns of inheritance during TF, including the first precise estimates of appreciable retained heterozygosity in parthenogenetic offspring that have important biological and evolutionary implications. These results advance our understanding of FP in vertebrates and identify new areas of research for which snakes, and likely other squamate reptiles, represent a robust model system.

Ethics. Specimens were maintained in a private collection (see Acknowledgments), with tissues collected non-invasively (female shed skin) or post-mortem (offspring).

Data accessibility

Raw Illumina data are available from the NCBI SRA (accessions SRR13866658–SRR13866660). Genotype data (including VCF) and measures of heterozygosity, relatedness, and GC content are available from Figshare (<https://doi.org/10.6084/m9.figshare.14166677>) and R analysis functions for calculating heterozygosity and relatedness are available from <https://github.com/darencard/parthenogenesis>. Mitochondrial 16S and ND4 sequences are available from GenBank (accession MT946535-6).

Received: 14 September 2020; Accepted: 15 March 2021

Published online: 31 March 2021

References

- Lampert, K. P. Facultative parthenogenesis in vertebrates: Reproductive error or chance?. *Sex. Dev.* **2**, 290–301 (2008).
- Chapman, D. D. *et al.* Virgin birth in a hammerhead shark. *Biol. Lett.* **3**, 425–427 (2007).
- Fields, A. T., Feldheim, K. A., Poulakis, G. R. & Chapman, D. D. Facultative parthenogenesis in a critically endangered wild vertebrate. *Curr. Biol.* **25**, R446–R447 (2015).
- Harmon, T. S., Kamerman, T. Y., Corwin, A. L. & Sellas, A. B. Consecutive parthenogenetic births in a spotted eagle ray *Aetobatus narinari*. *J. Fish Biol.* **88**, 741–745 (2015).
- Olsen, M. W. *Avian Parthenogenesis* (Agricultural Research Service, U.S. Department of Agriculture, 1975).
- Ramachandran, R. & McDaniel, C. Parthenogenesis in birds: A review. *Reproduction* **155**, R245–R257 (2018).
- Watts, P. C. *et al.* Parthenogenesis in Komodo dragons. *Nature* **444**, 1021–1022 (2006).
- Booth, W. & Schuett, G. W. The emerging phylogenetic pattern of parthenogenesis in snakes. *Biol. J. Linn. Soc.* **118**, 172–186 (2016).
- van der Kooi, C. J. & Schwander, T. Parthenogenesis: Birth of a new lineage or reproductive accident?. *Curr. Biol.* **25**, R659–R661 (2015).
- Booth, W. *et al.* Facultative parthenogenesis discovered in wild vertebrates. *Biol. Lett.* **8**, 983–985 (2012).
- Gamble, T. *et al.* The discovery of XY sex chromosomes in a *Boa* and *Python*. *Curr. Biol.* **27**, 2148–2153 (2017).
- Dudgeon, C. L., Coulton, L., Bone, R., Ovenden, J. R. & Thomas, S. Switch from sexual to parthenogenetic reproduction in a zebra shark. *Sci. Rep.* **7**, 40537. <https://doi.org/10.1038/srep40537> (2017).
- Feldheim, K. A. *et al.* Shark virgin birth produces multiple, viable offspring. *J. Hered.* **101**, 374–377 (2010).
- Feldheim, K. A. *et al.* Multiple births by a captive swellshark *Cephaloscyllium ventriosum* via facultative parthenogenesis. *J. Fish Biol.* **90**, 1047–1053 (2017).
- Booth, W., Johnson, D. H., Moore, S., Schal, C. & Vargo, E. L. Evidence for viable, non-clonal but fatherless *Boa* constrictors. *Biol. Lett.* **7**, 253–256 (2010).
- Booth, W. *et al.* Consecutive virgin births in the new world boid snake, the Colombian Rainbow Boa, *Epicrates maurus*. *J. Hered.* **102**, 759–763 (2011).
- Reynolds, R. G., Booth, W., Schuett, G. W., Fitzpatrick, B. M. & Burghardt, G. M. Successive virgin births of viable male progeny in the checkered gartersnake, *Thamnophis marcianus*. *Biol. J. Linn. Soc.* **107**, 566–572 (2012).
- Robinson, D. P., Baverstock, W., Al-Jaru, A., Hyland, K. & Khazanehdari, K. A. Annually recurring parthenogenesis in a zebra shark *Stegostoma fasciatum*. *J. Fish Biol.* **79**, 1376–1382 (2011).
- Straube, N., Lampert, K. P., Geiger, M. F., Weiß, J. D. & Kirchhauser, J. X. First record of second-generation facultative parthenogenesis in a vertebrate species, the whitespotted bambooshark *Chiloscyllium plagiosum*. *J. Fish Biol.* **88**, 668–675 (2016).
- Allen, L., Sanders, K. L. & Thomson, V. A. Molecular evidence for the first records of facultative parthenogenesis in elapid snakes. *R. Soc. Open Sci.* **5**, 171901. <https://doi.org/10.1098/rsos.171901> (2018).

21. Booth, W. & Schuett, G. W. Molecular genetic evidence for alternative reproductive strategies in North American pitvipers (Serpentes: Viperidae): Long-term sperm storage and facultative parthenogenesis. *Biol. J. Linn. Soc.* **104**, 934–942 (2011).
22. Booth, W. *et al.* New insights on facultative parthenogenesis in pythons. *Biol. J. Linn. Soc.* **112**, 461–468 (2014).
23. Miller, K. L. *et al.* Parthenogenesis in a captive Asian water dragon (*Physignathus cocincinus*) identified with novel microsatellites. *PLoS ONE* **14**, e0217489. <https://doi.org/10.1371/journal.pone.0217489> (2019).
24. Seixas, F., Morinha, F., Luis, C., Alvura, N. & dos Anjos Pires, M. DNA-validated parthenogenesis: First case in a captive Cuban boa (*Chilabothrus angulifer*). *Salamandra* **56**, 83–86 (2020).
25. Shibata, H., Sakata, S., Hirano, Y., Nitasaka, E. & Sakabe, A. Facultative parthenogenesis validated by DNA analyses in the green anaconda (*Eunectes murinus*). *PLoS ONE* **12**, e0189654. <https://doi.org/10.1371/journal.pone.0189654> (2017).
26. Calvete, J. J. *et al.* Venom complexity in a pitviper produced by facultative parthenogenesis. *Sci. Rep.* **8**, 11539. <https://doi.org/10.1038/s41598-018-29791-y> (2018).
27. Vonk, F. J. *et al.* The king cobra genome reveals dynamic gene evolution and adaptation in the snake venom system. *Proc. Natl. Acad. Sci.* **110**, 20651–20656 (2013).
28. Stuart, B. *et al.* *Ophiophagushannah*. *The IUCN Red List of Threatened Species* <https://doi.org/10.2305/IUCN.UK.2012-1.RLTS.T1177540A1491874.en> (2012).
29. Whitaker, N., Shankar, P. G. & Whitaker, R. Nesting ecology of the King Cobra (*Ophiophagus hannah*) in India. *Hamadryad* **36**, 101–107 (2013).
30. Schuett, G., Fernandez, P., Chiszar, D. & Smith, H. Fatherless sons: A new type of parthenogenesis in snakes. *Fauna* **1**, 20–25 (1998).
31. Charlton, T. *King Cobra: Natural History and Captive Management* (Natural History Publication Borneo, 2018).
32. Suntrarachun, S., Chanhome, L. & Sumontha, M. Phylogenetic analysis of the king cobra, *Ophiophagus hannah* in Thailand based on mitochondrial DNA sequences. *Asian Biomed.* **8**, 269–274 (2014).
33. Maddock, S. T., Ellis, R. J., Doughty, P., Smith, L. A. & Wüster, W. A new species of death adder (Acanthophis: Serpentes: Elapidae) from north-western Australia. *Zootaxa* **4007**, 301–326 (2015).
34. Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. & Hoekstra, H. E. Double Digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* **7**, e37135. <https://doi.org/10.1371/journal.pone.0037135> (2012).
35. Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W. & Postlethwait, J. H. Stacks: Building and genotyping loci de novo from short-read sequences. *G3 Genes Genomes Genet.* **1**, 171–182 (2011).
36. Catchen, J. M., Hohenlohe, P. A., Bassham, S., Amores, A. & Cresko, W. A. Stacks: An analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124–3140 (2013).
37. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
38. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
39. Li, H. *et al.* The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
40. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
41. McKenna, A. *et al.* The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
42. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491 (2011).
43. Van der Auwera, G. A. *et al.* From FastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinform.* **43**, 11.10.1–11.10.33 (2013).
44. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
45. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
46. Alho, J. S., Välimäki, K. & Merilä, J. Rhh: An R extension for estimating multilocus heterozygosity and heterozygosity–heterozygosity correlation. *Mol. Ecol. Resour.* **10**, 720–722 (2010).
47. R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2020).
48. Aparicio, J. M., Ortego, J. & Cordero, P. J. What should we weigh to estimate heterozygosity, alleles or loci?. *Mol. Ecol.* **15**, 4659–4665 (2006).
49. Kraemer, P. & Gerlach, G. Demerelate: Calculating interindividual relatedness for kinship analysis based on codominant diploid genetic markers using R. *Mol. Ecol. Resour.* **17**, 1371–1377 (2017).
50. Li, C. C. & Horvitz, D. G. Some methods of estimating the inbreeding coefficient. *Am. J. Hum. Genet.* **5**, 107–117 (1953).
51. Blouin, M. S., Parsons, M., Lacaille, V. & Lotz, S. Use of microsatellite loci to classify individuals by relatedness. *Mol. Ecol.* **5**, 393–401 (1996).
52. Jordan, M. A., Perrine-Ripplinger, N. & Carter, E. T. An independent observation of facultative parthenogenesis in the copperhead (*Agkistrodon contortrix*). *J. Herpetol.* **49**, 118–121 (2015).
53. Galtier, N., Piganeau, G., Mouchiroud, D. & Duret, L. GC-content evolution in mammalian genomes: The biased gene conversion hypothesis. *Genetics* **159**, 907–911 (2001).
54. Meunier, J. & Duret, L. Recombination drives the evolution of GC-content in the human genome. *Mol. Biol. Evol.* **21**, 984–990 (2004).
55. Schield, D. R. *et al.* Snake recombination landscapes are concentrated in functional regions despite PRDM9. *Mol. Biol. Evol.* **37**, 1272–1294 (2020).
56. Marais, G. Biased gene conversion: Implications for genome and sex evolution. *Trends Genet.* **19**, 330–338 (2003).
57. Duret, L. & Galtier, N. Biased gene conversion and the evolution of mammalian genomic landscapes. *Annu. Rev. Genomics Hum. Genet.* **10**, 285–311 (2009).
58. Taylor, E. N. & Booth, W. Rattlesnakes as models for reproductive studies of vertebrates. In *Rattlesnakes of Arizona* Vol. 2 (eds Schuett, G. W. *et al.*) 123–157 (ECO Wear and Publishing, 2016).

Acknowledgements

The authors acknowledge Gert-Jan Klootwijk, a private herpetoculturist, for providing shed skin from the female king cobra and tissues from the deceased babies, analyzed in this study. Additionally, we acknowledge Romilly van den Bergh for assisting in the coordination of sample collection. Without the involvement of the private herpetocultural community, this and other advances in our understanding of vertebrate facultative parthenogenesis would not have been possible.

Author contributions

G.W.S., F.J.V., T.A.C., and W.B. conceived the study. W.B. and T.A.C. provided funding for this work. S.S., F.J.V., N.R.C., and W.W. provided samples for this study. D.C.C. collected and analyzed the genomic data. D.C.C.,

G.W.S., T.A.C. and W.B. wrote the manuscript. All authors agreed to be held responsible for the content of the manuscript and approved the final version of the manuscript.

Funding

Research was supported by an NSF Doctoral Dissertation Improvement Grant to D.C.C. (NSF DEB-1501747), and both a Summer Research fellowship and Faculty startup funds from The University of Tulsa to W.B.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-86373-1>.

Correspondence and requests for materials should be addressed to W.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021