Signatures of Genome Evolution in Two Divergent Lizard Species from Texas

By

Jose A. Maldonado

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

August 2022

Acknowledgments

This dissertation was made possible with the help of past and current UTA Biology graduate students. I especially want to thank my graduate advisor, Dr. Matthew Fujita, for his support during my graduate studies and all the members of the Fujita lab who helped me. I thank all the UTA faculty that helped and taught me during my time as a graduate student. I thank the undergrads that helped me collect data and do fieldwork. I want to thank UTA Phi-Sigma Chapter, NSF GRP, and Rosemary Grant from the SSE for all funding my research. I want to thank other graduate students at other universities for any help that they gave me. Lastly, I want to thank the museum collections that provided me with tissue and DNA samples from areas I could not collect myself.

Abstract

The continual reduction of the cost of high-throughput sequencing is now making it feasible to sequence large genomic and transcriptomic datasets for non-model organisms. Many of these under study non-model organisms have no available molecular data available. Yet, many of these organisms are of interest for their usefulness in addressing long-standing unanswered questions in the field of evolutionary biology or of conservational concern and need immediate intervention. Now we can quickly generate large datasets for these organisms, and here we generate two large datasets for lizards that are both of in interest in the genus *Holbrookia* and *Aspidoscelis.* In chapter 1 we sequence the transcriptome for three lizard species in the genus *Holbrookia* who are of conservation interest as their populations are in decline. We aimed to generate the needed molecular needed for further conservation studies and identify adaptive loci. In chapter 2 we generated an extensive whole mitochondrial genome dataset for multiple lizards in the genus *Aspidoscelis*, whose genus contains multiple asexual and sexual reproducing lizards. In this chapter, we examined how the absence of sex influences the mitochondrial genome by comparing the asexual lizards' mitochondrial genomes to their sexual reproducing counterpart. In chapter 3 we begin developing further resources beyond only shot-gun sequencing genomes but develop a protocol to culture fibroblast cells from the tail tissues of lizards successfully. With additional resources beyond shot-gun sequence data, we can better address questions found in chapters 1 and 2.

LIST OF TABLES

CHAPTER 1

CHAPTER 2

LIST OF FIGURES

Chapter 1

CHAPTER 2

Chapter 1

Transcriptome sequencing reveals signatures of positive selection in the spot-tailed earless lizard

Jose Maldonado[1], Thomas J. Firneno Jr[1], Corey Roelke[1], Nathan Rains[2], Juliet Mwgiri[1]

Matthew K. Fujita[1,3]

[1]Amphibian and Reptile Diversity Research Center, Department of Biology, The University of Texas at Arlington, Arlington, TX 76019

[2] Texas Parks and Wildlife Department

4200 Smith School Rd

Austin, Texas 78744,

Abstract

The reduced cost of high-throughput sequencing technologies now provides opportunities to investigate non-model organisms. Many under-study non-model organisms have little to no genetic data available. Yet, many of these under-study species are of extreme interest as they can be useful to address long-standing unanswered questions in the field of evolutionary biology or organism that are of conservational interest. The low cost of high-throughput sequencing makes it feasible to generate large genomic and transcriptomic datasets for these under-study organisms and begin addressing some of the questions. Here we generated high-throughput sequencing datasets for lizards in the genus *Aspidoscelis* and *Holbrookia.* Here we examine in chapter 1 how generating a transcriptome for lizards in the genus *Holbrookia*, lizards that populations are in decline can help the conservational efforts by developing the molecular resources for future conservation studies and characterizing their transcriptome. In chapter 2 we investigate the genomic consequences of asexuality by developing an extensive dataset of whole mitochondrial genomes for lizards in the genus *Aspidoscelis* to analyze how the lack of sex influences the mitochondrial genome. Finally, in chapter 3 we develop a protocol to culture fibroblast cells from tail tissues from three different species in the genus *Aspidosceli* by developing additional resources to address further questions examined in chapters 1 and 2.

INTRODUCTION

The continuing reduction of per-base costs of high-throughput sequencing (HTS) methodologies has provided new opportunities to generate large-scale genomic and transcriptomic resources. This is especially beneficial for data deficient, non-model organisms that may require conservation action, are commercially valuable, or are excellent systems to investigate ecological and evolutionary questions [1–4]. The field of conservation genetics has historically relied on the use of few dozen genetic markers (allozymes, microsatellites, mitochondrial DNA, and a few nuclear genes) to understand population structure, viability, and evolutionary processes (e.g. genetic drift, selection, and migration) [5,6]. With the low cost of HTS, it is now possible to rapidly generate hundreds to thousands of genetic markers from multiple individuals and populations to assay genetic diversity for virtually any species [7]. These larger datasets can overcome some of the limitations of traditional methods used in conservation genetics that yield only a few variable loci [8]. For example, with hundreds to thousands of variable sites, HTS datasets can be beneficial for species management by providing high resolution and accurate inferences of important parameters such as genetic diversity, inbreeding depression, effective population size [9,10], as well as historical demography and local adaptations. All of this can provide insight into resolving taxonomic uncertainties to determine which species need immediate attention and protection [8,11]. Using RNA-Seq methodologies to address the types of questions for species with no available genomic resources is becoming increasingly favorable [12–14]. As long as appropriately preserved tissue is available, it is possible to sequence the expressed genes by sequencing complementary DNA (cDNA) libraries. By using millions of short reads generated by massive parallel sequencing of cDNA libraries and robust assembly methods [15,16] , one can generate a high coverage *de novo*

transcriptome assembly without the need for a reference genome. Because transcriptome sequencing is versatile, it is a desirable option for developing conservation genetics tools because it largely circumvents the time-consuming process of identifying and optimizing genetic markers (e.g. primer development and testing) [12,17]. Furthermore, transcriptome sequencing captures protein coding regions with functional significance [2,18]. The usage of gene expression data for conservation biology is an emerging field that will significantly benefit wildlife management [19].

Our focus in this paper is the development of genomic resources for the Northern and Southern Spot-tailed Earless Lizard (*Holbrookia lacerata* and *Holbrookia subcaudalis)*, whose historic range extends from central to southern Texas, and into northeastern Mexico. For the past several decades*, H. lacerata* has faced taxonomic uncertainties and confusion. Initially, there was a single species, *H. lacerata*, that included the currently-recognized *H. lacerata* as well as its sister species, *H. maculata* [20]. *Holbrookia lacerata* was then reduced to a subspecies under *H. maculata* [21], and subsequently elevated back to full species status [22]. Subsequently, two subspecies of *H. lacerata* were recognized, *H. l. lacerata* and *H. l. subcaudalis*, that are geographically separated by the Balcones Escarpment fault line. Renewed interest in resolving their taxonomic status revealed that the disjunct populations are reciprocally monophyletic with an approximately 4% mitochondrial sequence divergence [3]. Most recently, Hibbits *et al*. (2019) took an integrative approach and used morphology, mitochondrial, and nuclear data to elevate each subspecies to full species status [23].

Both species have experienced a sharp decline in distribution and abundance through their historical range and are labeled as near threatened by the International Union for Conservation of Nature organization. The most notable decrease has occurred in southern Texas,

where *H. subcaudalis* populations have become increasingly fragmented and isolated, with fewer observations being made throughout its historic range. Hypotheses for the decline of *H. lacerata* include pesticides and agriculture practices [24], and urbanization, exotic grasses, and invasive fauna are additional factors contributing to their decline [25]. The conservation concerns of both species have led to recent studies utilizing *H. lacerata* as a focal organism for better land management practices in Texas [25–27], and efforts are being made to protect both species under the Endangered Species Act [28].

The decline in *H. lacerata* and *H. subcaudalis* abundance can have a substantial impact on population viability. Small and fragmented populations can lead to an increase in homozygosity, the disappearance of genetic diversity, and an increase in the frequency deleterious variants become fixed can lead to inbreeding depression, and thus a reduction in individual fitness. The goal of this study is to provide a detailed characterization of a transcriptome as a means to generate molecular resources for conservation studies of *H. lacerata* and *H. subcaudalis*. We provide an annotated liver transcriptome, identify adaptive loci, and estimate genetic distance, all of which are of value for the conservation efforts of the North and Southern Spot-tailed Earless Lizard.

MATERIALS AND METHODS

Sampling

Samples for this study were collected in summer 2015 from three different localities. Lizards were caught by hand or loops and humanely euthanized under our IACUC protocol (#A16.010) approved by the University of Texas at Arlington. We harvested skeletal muscle, liver, heart, blood, and integument from one individual per species and stored the tissue in RNAlater (Sigma-Aldrich, St. Louis, MO). We preserved the specimens with 10% formalin and

deposited them in the Amphibian and Reptile Diversity Research Center (ARDRC) at the University of Texas at Arlington (see supplementary table 1 for locality data and field numbers).

RNA extraction and RNA-Seq library preparation for sequencing

Total RNA was extracted from liver tissue using a Promega SV Total RNA Isolation kit (Promega, Madison, WI) following the manufacturer's protocol. We quantified our RNA extractions on the Qubit 2.0 Fluorometer (Life Technologies, Carlsbad, CA), and assessed RNA quality and size distribution on an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA). All RNA extractions yielded high-quality RNA, with all samples having an RNA integrity number (RIN) >8. RNA isolates from each sample were sent to the Brigham Young University DNA Sequencing Center to generate cDNA libraries using the Kapa Biosystems RNA depletion kit (Kapa Biosystems, Wilmington, MA) and sequenced on the Illumina HiSeq 2500 (Illumina, San Diego, CA) generating 100bp paired-end sequences.

***De novo* Transcriptome Assembly and Quality Assessments**

The data were processed to remove low quality reads using Trimmomatic v.32 [29]. We used the following parameters to trim and remove failed reads: a 4-base sliding window trimming nucleotides with a Q score <5 and discarding reads <25bp long [30]. To ensure successful quality trimming and filtering, we ran the processed reads through FASTQC v0.10.1 (Babraham Bioinformatics) to evaluate read quality, length, and the number of reads retained. With no reference genomes for any *Holbrookia* species and given that previous results have demonstrated that guided transcriptome assembly methods for diverged species typically perform worse than *de novo* assembly [31,32] , we carried out *de novo* assemblies using Trinity short-read assembler V2.2.1 [15] for each sample.

To measure the completeness of our assembled transcripts, we performed two quality assessments as suggested by the Trinity package. First, to evaluate the assembly quality of each transcriptome, we mapped the input processed RNA-Seq reads back to their corresponding assemblies for each species of *Holbrookia* and quantified the number of input reads represented in our *de novo* transcriptome assemblies. Reads were mapped back to the transcriptome using the short read aligner Bowtie 2 v2.3.4 using the –local and –no-unal options [33]. Second, we ran our *Holbrookia* protein coding transcript set (see below for protein coding transcript set generation) through CD-Hit v4.8.1 (90% sequence identity threshold with all other parameters set to default) [34] to produce a non-redundant protein coding transcript set for each individual. We compared our non-redundant transcripts with a set of 3,950 highly conserved single-copy tetrapod orthologs using the BUSCO (benchmarking universal single-copy orthologs) v3 pipeline [35]. Figure 1 is an overview of our transcriptome assembly and analysis pipeline described below.

Identification of Orthologs and Paralogs and Pairwise Distance

To identify candidate coding genes from our assembled transcript sets with the longest open reading frames (ORFs), we used the TransDecoder v3.0.1 pipeline [36]. To maximize the number of ORFs captured and to ensure we did not lose any potential coding genes, we ran TransDecoder optional homology search against the PFAM database [37]. We ran the TransDecoder pipeline on our three *de novo Holbrookia* transcriptomes, the *Anolis carolinensis* transcriptome downloaded from NCBI server (GCF_000090745.1), and the publicly available *Sceloporus undulatus* transcriptome [38].

We used OrthoFinder v2.2.3to identify orthologous genes between all five species from the protein coding transcript set generated by TransDecoder. OrthoFinder uses a BLAST all-vs-

all algorithm and performs reciprocal best-hit BLAST searches that normalize the bit scores to overcome transcript length bias in the ortholog detection [39]. We extracted all of the 1-to-1 orthologs found between all five species into Fasta files and performed an amino-acid guided alignment for each ortholog using MACSE v1.2 [40]. We inspected our multiple species codon alignments and removed ten orthologs for having a poor alignment. We used Geneious R9 (https://www.geneious.com) to calculate the uncorrected pairwise distance across the remaining orthologous alignments.

To identify paralogs (duplicated single-copy orthologs) in our *Holbrookia* transcripts, we used our non-redundant protein coding transcript set generated by CD-Hit to reduce false positive detection of paralogs. We searched against the tetrapod single-copy ortholog database to determine the number of transcripts that are likely paralogs using the BUSCO pipeline [35] and OrthoFinder all-vs-all BLAST [39] algorithm.

DAVID Functional Analysis

We submitted all *A. carolinensis* gene IDs from the 1-to-1 orthologs set to the bioinformatics database DAVID [41] to identify genes in our orthologous set, cluster these genes into their biological processes, and their inclusion in biological pathways. We ran DAVID using all default parameters for enrichment and pathway analysis.

Selection Test

To identify positive selection in *Holbrookia*, we estimated dN/dS in all 1-to-1 orthologous genes found by OrthoFinder across all five individuals by following the recommendations of Yang *et al*. (2006) [42]. We used the branch-sites model (CODEML: M2 and NSites2) from the PAML v4.9 package [43], which requires an *a priori* phylogenetic tree to test for positive selection in foreground branches. We utilized a phylogenetic tree generated by

OrthoFinder, by reconciling over a thousand gene trees into a species tree and designated the *Holbrookia* lineage as the foreground branches. The remaining *Anolis* and *Sceloporus* branches on the species tree were labeled as background branches. We compared the likelihood of two different models for each orthologous gene: **1)** our alternative model that allows for a proportion of the sites to be under positive selection ($\omega_2 \geq 1$) along the *Holbrookia* branch, and the background branches having a proportion of sites being under purifying selection ($\omega_1<1$) or neutrally evolving ($\omega_0=1$); and **2)** the simplistic null model that has the ω_2 fixed at 1 on the foreground branches, and all other branches having $\omega_0=1$ and $\omega_1<1$ [43,44]. We obtained likelihood values after running each transcript under two different models and carried out likelihood ratio tests ($2\times\Delta\ln L$) between the models to evaluate whether the alternative model outperforms the null model. We performed a Bonferroni correction to account for multiple comparisons. A significant result from the branch-sites model is indicative that a subset of the sites in the coding gene has gone through episodic positive selection, with the selected sites providing an advantage to *Holbrookia* lineage.

Gene Annotation

We used the Blast2Go v5 pipeline [45] to annotate and assign functions to our complete protein coding gene set produced by TransDecoder: 34,214 transcripts for *H. maculata*, 33,379 transcripts for *H. lacerata*, and 29,149 transcripts for *H. subcaudalis*. We used the Blastp function to blast our transcript set against the NCBI NR-protein database using an e-value cutoff of 1e-5 and all other parameters set to default. Annotation was performed using an e-value cutoff of 1e-3, an annotation score of 45, and a GO weight of 5. To generate biological processes, molecular functions, and cellular components graphs we first grouped our Go annotation into

GO-slim terms to simplify the input and filtered out nodes containing >10 sequences cellular component, nodes containing >10 sequences were filtered out.

RESULTS

Sequencing and *De novo* transcriptome assembly

Our sequenced cDNA libraries on the Illumina Hi-Seq 2500 yielded a total of 223 million paired-end reads between all three samples, and we retained 99.5% of our reads after filtering out low-quality reads from our raw data set. FastQC verified that only high-quality reads with a Q score >30 were kept and assembled by the short-read assembler Trinity. The total number of assembled transcripts for *H. maculata, H.* lacerata, and *H. subcaudalis* are respectively 107,863, 99,821, and 91,278. The average maximum transcript length between all individuals was ~18,340 bp. Assembly statics for all three *Holbrookia* samples used in this study are listed in Table 1.

Transcriptome sequencing and assembly statistics	*H. maculata*	*H. l. lacerata*	*H. subcaudalis*
Raw paired-end reads	93,688,798	73,668,784	66,629,872
Cleaned paired-end reads	93464743	73492888	66508242
N50	2,407	2,288	2,056
Number of Transcripts	107,385	99,821	91,278
Max Transcript length	17,863 bp	20,377 bp	16,777 bp
Min Transcript length	224 bp	224 bp	224 bp
Mean Transcript length	1,061 bp	1,037 bp	953 bp
GC content	43%	43%	42%

Table 1.

Assembly statistics for *H. maculata, H. lacerata*, and *H. subcaudalis*.

Assembly Quality Assessment

Our first assembly quality check involved mapping the processed reads with Bowtie2 back to their assembled transcriptome. Across all samples, an average of 97% of the reads

mapped back to their respective assemblies. If ≥70% of the input reads mapped back to their assembly, it is indicative of a robust transcriptome assembly by Trinity [15]. To further quantify the completeness of our *Holbrookia* transcript set, we ran our protein coding gene set for each individual through CD-HIT to remove redundant transcripts from the assemblies. Our *Holbrookia* protein coding transcript set (96,742) decreased by a factor of one-third to generate a non-redundant transcript list of 63,957 sequences. We ran our three non-redundant *Holbrookia* transcript sets against a conserved set of 3,950 universal tetrapod single-copy orthologs using the BUSCO pipeline [35,46]. We recovered 60-70% complete and 8-11% partial orthologs from the BUSCO tetrapod database. Between 831 to 1,113 orthologs were classified as missing from our transcript set. The high number of complete conserved single-copy orthologs present in each transcript set is indicative of high coverage and high-quality protein coding transcriptome assemblies for each Holbrookia species. All BUSCO statistics are detailed in Table 2.

BUSCO statistics	*H. maculata*	*H. lacerata*	*H. subcaudalis*
Complete Buscos	2,764	2,727	2,372
Complete-single-copy BUSCOs	2,627	2,604	2,259
Complete-duplicated BUSCOs	137	123	113
Fragmented BUSCOs	345	392	465
Missing BUSCO	841	831	1,113

Table 2.

Benchmarking Universal Single-Copy Orthologs (BUSCO) summary of complete, duplicated, fragmented, and missing orthologs search against the 3950 single-copy orthologs.

Orthologs and Paralogs Identification

To identify the 1-to-1 orthologous groups among our coding genes identified by TransDecoder between the three *Holbrookia* species, *Anolis carolinensis*, and *Sceloporus*

undulatus, we ran all five individuals through OrthoFinder. TransDecoder found between 29,149 to 35,594 protein coding genes with the longest open reading frame from our assembled transcript set. We submitted a total of 160,639 coding transcripts to OrthoFinder to identify orthologous groups. OrthoFinder identified 19,401 orthogroups (defined as containing both orthologs and paralogs) containing 125,845 transcripts. We found that 43% (8,243) of the orthogroups had all five individuals present, and 56% (10,916) of the total orthogroups had at least two individuals present (figure 2). We found 6,805 transcripts (1,361 orthologs) that were true 1-to-1 orthologs between all five individuals, and a low proportion of our transcripts that were unique to their given transcriptome assemblies.

To identify paralogs in our *Holbrookia* transcriptome, we ran our protein coding set through CD-Hit [34] to remove redundancy in our transcript set for each of our *Holbrookia* samples. BUSCO found that between 3% to 3.5% of the 3950 single-copy tetrapod orthologs that we searched against were detected as duplicated in our transcript assemblies. Our second method to identify paralogs in the *Holbrookia* transcriptome was using the OrthoFinder BLAST algorithm only on our *Holbrookia* protein coding gene list, by counting the number of self-BLAST hits identified between each transcript. We discovered 4590 transcripts in *H. maculata*, 4797 transcripts in *H. lacerata*, and 4271 transcripts in *H. subcaudalis* as paralogous transcripts within each species.

DAVID Functional Analysis and Selection Test

DAVID clustered our orthologous genes into separate biological processes, 409 genes into metabolic processes, 107 genes associated with stress response processes, 190 genes that play a role in gene expression, and 216 genes that are involved in cellular component organization. DAVID could not associate 838 genes with any biological process as they did not

meet the enrichment threshold (P-values < .1). DAVID identified 156 genes involved in different metabolic pathways (purine/pyrimidine metabolism, amino acid metabolism, and the Citric acid cycle). DAVID identified thirteen genes associated with the nucleotide excision repair pathway, six genes linked to nucleotide mismatch repair mechanism, and seven genes that have a role in DNA replication.

To determine if any of the 1-to-1 orthologous genes have undergone positive selection in *Holbrookia*, we carried out selection tests using PAML [44] and performed likelihood-ratio tests to assess significance. We adjusted our P-values by performing a Bonferroni correction test to account for multiple comparisons and committing a type I error. We found twelve genes from our ortholog set that have undergone positive selection with a significance value of < .05, after adjusting our P-values. Table 3 has a complete list of all twelve genes with the alternative and null likelihood values and the Bayes Empirical Bayes (BEB) score for sites under selection.

Gene	Alternative Log Likelihood	Null Log Likelihood	Unadjusted P-Values	Adjusted P-Values	BEB scores for positive sites
BNIP3L	-1236.019200	-1245.697904	1.083979E-05	7.7000E-04	N 0.996 N 0.986 N 0.986
TBL2	-3224.401186	-3232.500738	5.702108E-05	4.3890E-03	P 0.939 Q 0.938
RP2	-2022.830549	-2036.859674	1.177182E-07	9.0860E-06	G 0.975 E 0.950 E 0.995
PHGDH	-3332.163360	-3342.054215	8.680941E-06	6.6836E-04	Q 0.980 A 0.921
KRT18	-2900.331070	-2907.152926	0.0002209801	1.6940E-02	G 0.947 S 0.942 N 0.977
LURAP1	-1161.470425	-1170.150278	3.093145E-05	2.3793E-03	P 0.861 A 0.867
CDO1	-1337.215619	-1344.793392	9.900707E-05	7.6230E-03	Q 0.925
LBHD2	-761.257900	-788.429241	1.68404E-13	1.2936E-11	Q 0.988 I 0.999 C 0.999 V 0.999 D 0.999 A 0.993 C 0.994 T 0.999
GRSF1	-2724.823688	-2731.690040	0.0002107525	1.6170E-02	39 T 0.914 40 Q 0.970
SLC9A3R1	-1661.883638	-1669.365305	0.0001096206	8.4700E-03	E 0.975
RNF81	-1070.037763	-1079.230221	1.804811E-05	1.5400E-03	P 0.972 S 0.985
MPHOSPH6	-1077.189311	-1086.384783	1.79911E-05	1.3860E-03	A 0.921 R 0.953 P 0.951

Table 3.

The alternative and null model likelihood values for the twelve genes that show footprints of positive selection in the *Holbrookia* lineage.

Pairwise Distance

We calculated the uncorrected pairwise distance using the complete single copy ortholog set. The most significant sequence divergence was found between *A. carolinensis* and all other

species being around ~11%. The split between *Anolis* (family Dactyloidae) and the four other species in the family Phrynosomatidae occurred ~72MYA and accounts for the vast amount of genetic divergence [47]. There is an average 5.57% sequence divergence between *S. undulatus* and all *Holbrookia* species, with the pairwise distance between *H. maculata* and *H. lacerata* and *H. subcaudalis* is ~1.5%. The genetic distance between *H. lacerata* and *H. subcaudalis* 0.73% (Table 4).

Pairwise Distance	*A. carolinensis*	*S. undulatus*	*H. maculata*	*H. lacerata*	*H. subcaudalis*
A. carolinensis	0				
S. undulatus	11.29%	0			
H. maculata	11.10%	5.58%	0		
H. lacerata	11.27%	5.56%	1.53%	0	
H. subcaudalis	11.29%	5.57%	1.50%	0.73%	0

Table 4.

The uncorrected pairwise distance calculated across all orthologous gene sets between all species.

Annotations

Out of the 96,742 transcripts submitted to BLAST2GO, around ~85% were successfully blasted against the NCBI NR-protein database, with the top BLAST hit being *A. carolinensis* with the remaining orthologs blasting to other reptile species (e.g. *Python bivittatus, Gecko japonicus*). We successfully annotated a total of 71,314 transcripts into 55 functional categories using GO-Slim assignments within the three categories of the GO classification system biological process, cellular component, and molecular function. The primary categories our coding genes were clustered into are cellular metabolic processes (30,145), nitrogen compound

metabolic process (28,816), primary metabolic process (23,249), and ion binding (21,228). A breakdown of GO terms for each category for each *Holbrookia* sample is shown in figure 3.

DISCUSSION

Genetic data are extremely useful in conservation studies because they can provide estimates of genetic diversity, population size (and viability), and local adaptations for imperiled populations and species. However, many taxa of conservation significance have little to no genetic data available, impeding studies that can inform conservation planning and action. The speed and ease of generating molecular data for the use of species management and policy have quickly increased with the transition to HTS technologies [8,48]. The decline of Northern and Southern Spot-Tailed Earless Lizard led us to generate and profile the liver transcriptome for each species. By doing so, we establish a genomic resource for future genome-scale studies on *Holbrookia* as well as provide initial insights into the divergence and evolution of *H. lacerata* and *H. subcaudalis.* Because divergence and adaptation are important components of conservation genetics, we (1) quantify the genetic distance between *H. lacerata* and *H. subcaudalis,* and (2) identify potential genes under selection in *H. lacerata* and *H. subcaudalis*.

Transcriptome Annotation

Our gene ontology analyses revealed that the biological functions of our liver transcripts closely follow similar patterns of other liver transcriptomes in squamates [49]. A large number of genes expressed in the liver play a role in metabolism, proteolysis, and nitrogen compound synthesis and breakdown. Oxidative damage can lead to DNA lesions and strand breaks, and are caused by endogenous reactive oxygen species produced during normal cellular metabolism [50]. Here we found both heath shock and glutathione peroxidase proteins in our transcriptome set that are well documented as having roles in protecting against oxidative damage [51].

Pairwise Distances

The U.S. Fish and Wildlife Service (FWS) is currently reviewing if *H. lacerata and H. subcaudalis* require protection under the endangered species act [28]. Using their historical classification as subspecies, the FWS can place either the Northern or Southern Spot-Tailed Earless Lizard in need of protection, or both. Here we have found significantly more divergence (0.73%) between the Northern and Southern Spot-Tailed Earless Lizard using conserved orthologous sequences than previous study has estimated between two allopatric species using nuclear genes [52] . The results in this study, alongside Roelke *et al.* (2018) finding that both the northern and southern lizards are genetically distinct using mitochondrial genomes, all further support the Hibbitts *et al.* (2019) recent reclassification of their taxonomy. The elevation of the Northern and Southern Spot-Tailed Earless Lizard from subspecies to full species status requires the FWS consider them separately for protection, as each species faces unique threats [23,24]. The southern *H. subcaudalis* has two isolated populations that were once part of a larger distribution across southern Texas, while the more robust northern species has more observations in multiple surveys of the region [23].

DAVID Analysis of Orthologs and Positive Selection

Established methods to find signatures of adaptive evolution across the genome coupled with the increased number of molecular markers captured with HTS techniques are facilitating the discovery of adaptive variation. Identifying populations from at-risk species that show signs of local adaptive genetic variation and subsequently maintaining it are essential parts of any conservation strategy [53]. Adaptive variation can inform conservation managers on how to perform genetic rescues or assisted gene flow to raise the fitness of small populations that lack genetic diversity [54]. As the use of large scale genomic data becomes more widely adopted in

the field of conservation biology, it will have a significant role in conservation management and policy.

The bulk of nonsynonymous mutations are generally considered deleterious as they alter protein structure and function, which can reduce an organism's fitness, and purifying selection should remove these mutations. Genes such as *RAD23* protein [55] and Damage-specific DNA Binding protein I [56] are involved in the repair of DNA lesions, and were clustered by DAVID into the nucleotide repair pathways that had no fingerprints of positive selection. Seldomly occurring, some amino acid replacements can lead to changes in proteins that are advantageous in new environments and selected for and kept in the species lineage.

Our branch-sites model selection test identified 12 positively selected genes in the *Holbrookia* lineage that have roles in immunity, development, and metabolism. From the twelve positively selected genes, three genes, *BNIP3L*, *LBHD2*, and *PHGDH,* are of particular interest for their potential role in shaping morphology, resistance against viruses, and amino acid production. The *LBH* (Limb-bud-and-heart) Domain-2 is part of a protein family that are key transcription regulators in embryonic development. The *LBH* gene is expressed early in embryogenesis with proteins playing a role in the development of limb buds and heart formation in a mice-model system [57]. A previous study has identified a single nonsynonymous mutation in an *LBH* homolog [58] associated with an adaptive variation, and here we have found multiple amino acid replacements selected for in the *Holbrookia* lineage. While we do not know the functional consequences of these amino acid changes, they may play an important role in local adaptation to desert and grassland environments that *Holbrookia* inhabit.

Immune genes are known hotspots for selection to act on, having persistent selective pressures by pathogens that elicit immune responses. *BNIP3L* promotes apoptotic activity,

targeting dysfunctional mitochondria organelles. Specific viruses produced anti-apoptotic proteins that suppress the apoptosis of virally infected cells, allowing viruses to replicate and multiply. *BNIP3L* proteins interact with viral and cellular anti-apoptosis proteins and overcome the suppression to initiate apoptosis [59,60].

Amino acids are the essential building blocks to protein, and here we found the *PHGDH* (D-3-phosphoglycerate dehydrogenase) gene that has a role in amino synthesis has experienced positive selection with multiple amino acid replacements. The protein encoded by this gene is responsible for producing the non-essential amino acid L-serine. While described as non-essential, L-serine is a critical precursor required for the synthesis of D-serine, amino acids, and other metabolites in animal cells [61]. Deficiencies in the D-3-phosphoglycerate dehydrogenase protein results in metabolic defects affecting the nervous systems [62,63]. The kidneys produce the majority of L-serine under normal conditions [61]. When dietary protein is limited, the liver becomes the primary production of the metabolite. While this has been shown only in mammals, it is of interest to see if a similar pattern will be observed in *Holbrookia* species as the alteration of grassland for agriculture purposes and the use of pesticides, can impact the invertebrate populations and reduce their food availability.

Conclusions

While there are still problems in conservation genetics that can still be answered successfully using conventional conservation genetics techniques, the low cost of HTS technologies allows us to address questions of genetic diversity, adaptation, and taxonomy. High-throughput sequencing technologies are improving on traditional approaches generating extensive molecular resources for organisms of conservation interest. Scaling up from just a few loci to genomics and transcriptomics allows for better inferences and conservation practices.

Here we have generated the first transcriptome for three lizards in the genus *Holbrookia* and identified multiple genes under positive selection. These transcriptomes have already provided insight into potentially adaptive loci in *Holbrookia* and will continue to contribute to future population-based and systematics studies of iguanian lizards.

ACKNOWLEDGEMENTS

We thank the biologists of the Texas Parks and Wildlife Department. We thank the many landowners and land managers, both public and private, who granted us access to their land for lizard surveys. This work was funded by a Texas Parks & Wildlife Section Six Grant awarded to Corey E. Roelke and Matthew K. Fujita (TPWD 474241).

Data Archive

The raw read will be submitted to the NCBI Sequence Read Archive, and transcriptome assemblies and PAML selections test output files will be uploaded to the digital repository Dryad. https://datadryad.org/stash/dataset/doi:10.5061%2Fdryad.kwh70rz1m

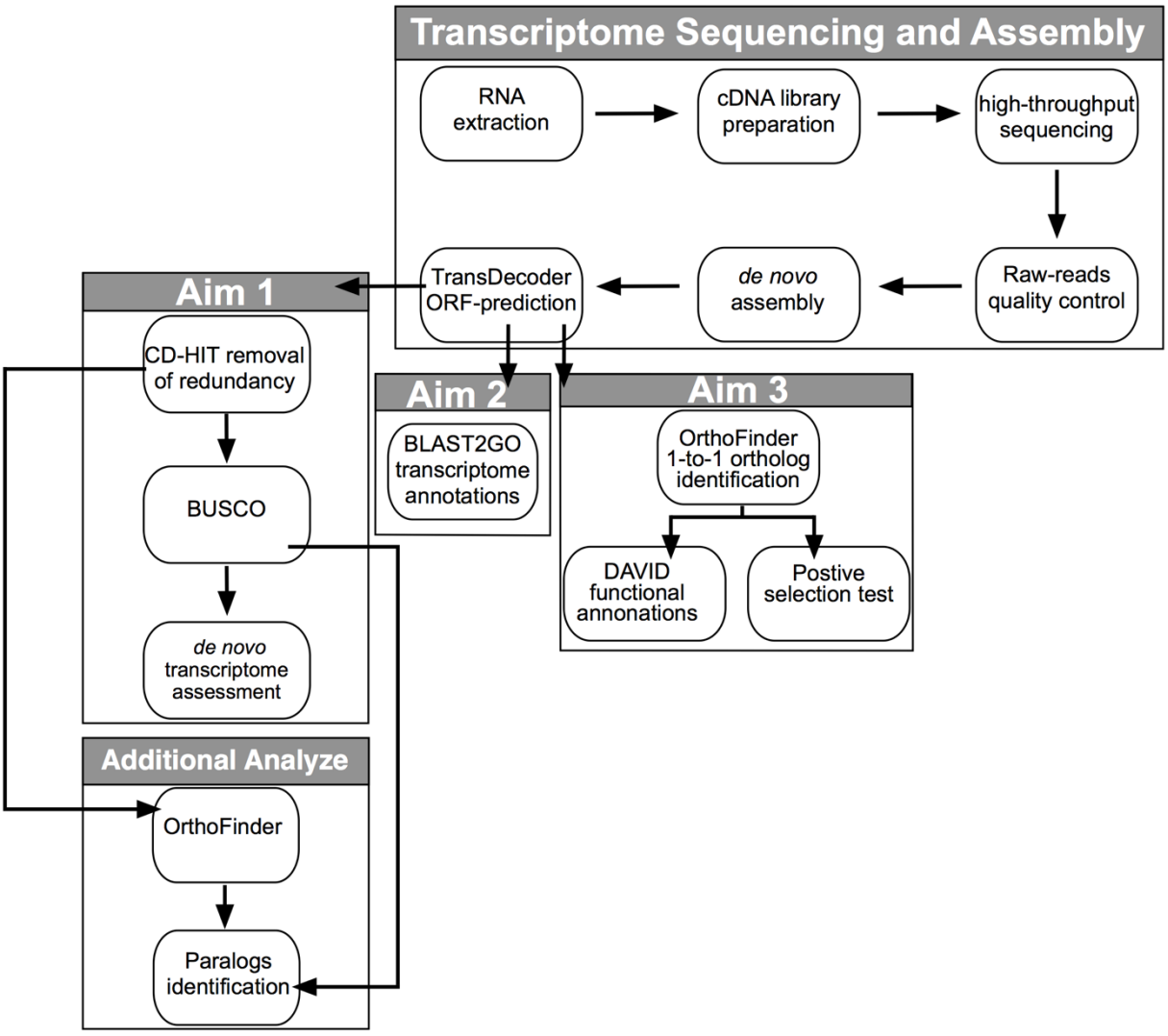

Figure 1.

Schematic overview of our workflow to assemble the transcriptome of *H. lacerata, H. subcaudalis,* and *H. maculata,* and analysis carried out in this study.

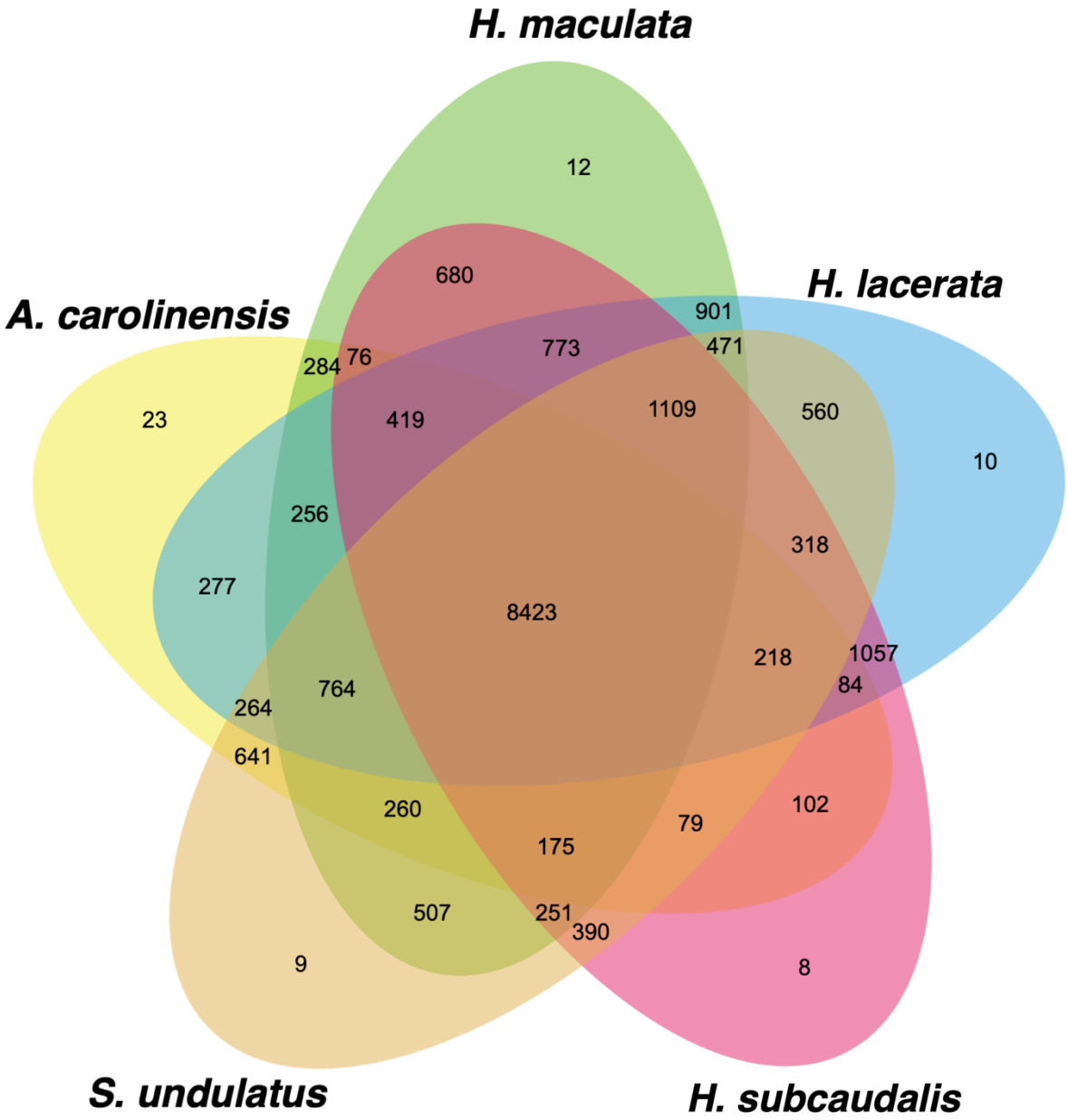

Figure 2.

Venn diagram showing the number of shared orthologous groups identified by OrthoFinder between all four species assembled transcriptome.

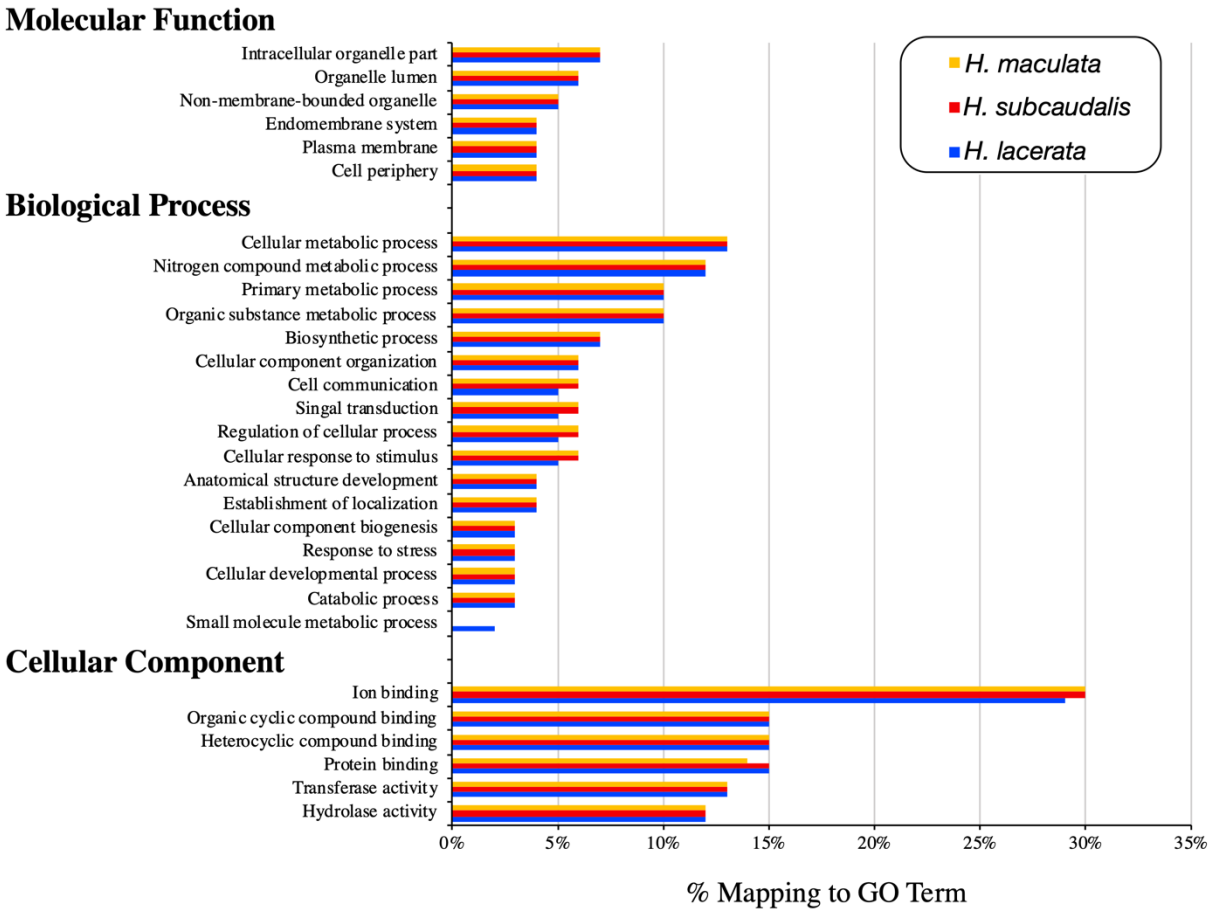

Figure 3.

Gene ontology annotations for the liver transcriptome of *H. maculata* (yellow), *H. subcaudalis* (red), *and H. lacerata* (blue).

References

1. Tollis M, DeNardo DF, Cornelius JA, Dolby GA, Edwards T, Henen BT, et al. The Agassiz's desert tortoise genome provides a resource for the conservation of a threatened species. PLOS ONE. 2017. p. e0177708. doi:10.1371/journal.pone.0177708

2. Cheng T, Fu B, Wu Y, Long R, Liu C, Xia Q. Transcriptome sequencing and positive selected genes analysis of Bombyx mandarina. PLoS One. 2015;10: e0122837.

3. Roelke CE, Maldonado JA, Pope BW, Firneno TJ, Laduc TJ, Hibbitts TJ, et al. Mitochondrial genetic variation within and between *Holbrookia lacerata lacerata* and *Holbrookia lacerata subcaudalis*, the spot-tailed earless lizards of Texas. J Nat Hist. 2018;52: 1017–1027.

4. Carruthers M, Yurchenko AA, Augley JJ, Adams CE, Herzyk P, Elmer KR. De novo transcriptome assembly, annotation and comparison of four ecological and evolutionary model salmonid fish species. BMC Genomics. 2018;19: 32.

5. Frankham R, Briscoe DA, Ballou JD. Introduction to Conservation Genetics. Cambridge University Press; 2002.

6. Avise JC, Hamrick JL. Conservation genetics: case histories from nature. 1996.

7. Allendorf FW, Hohenlohe PA, Luikart G. Genomics and the future of conservation genetics. Nat Rev Genet. 2010;11: 697–709.

8. Funk WC, McKay JK, Hohenlohe PA, Allendorf FW. Harnessing genomics for delineating conservation units. Trends Ecol Evol. 2012;27: 489–496.

9. Höglund J. Evolutionary Conservation Genetics. Oxford University Press; 2009.

10. Primmer CR. From conservation genetics to conservation genomics. Ann N Y Acad Sci. 2009;1162: 357–368.

11. Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E. Towards next-generation biodiversity assessment using DNA metabarcoding. Mol Ecol. 2012;21: 2045–2050.

12. Vera JC, Wheat CW, Fescemyer HW, Frilander MJ, Crawford DL, Hanski I, et al. Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. Mol Ecol. 2008;17: 1636–1647.

13. Miller HC, Biggs PJ, Voelckel C, Nelson NJ. De novo sequence assembly and characterisation of a partial transcriptome for an evolutionarily distinct reptile, the tuatara (*Sphenodon punctatus*). BMC Genomics. 2012;13: 439.

14. Schunter C, Vollmer SV, Macpherson E, Pascual M. Transcriptome analyses and differential gene expression in a non-model fish species with alternative mating tactics. BMC Genomics. 2014;15: 167.

15. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat Biotechnol. 2011;29: 644–652.

16. Henschel R, Nista PM, Lieber M, Haas BJ, Wu L-S, LeDuc RD. Trinity RNA-Seq assembler performance optimization. Proceedings of the 1st Conference of the Extreme Science and Engineering Discovery Environment on Bridging from the eXtreme to the

campus and beyond - XSEDE '12. 2012. doi:10.1145/2335755.2335842

17. Parchman TL, Geist KS, Grahnen JA, Benkman CW, Buerkle CA. Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. BMC Genomics. 2010;11: 180.

18. Renaut S, Nolte AW, Bernatchez L. Mining transcriptome sequences towards identifying adaptive single nucleotide polymorphisms in lake whitefish species pairs (*Coregonus spp.* Salmonidae). Mol Ecol. 2010;19 Suppl 1: 115–131.

19. Kristensen TN, Sørensen P, Pedersen KS, Kruhøffer M, Loeschcke V. Inbreeding by environmental interactions affect gene expression in Drosophila melanogaster. Genetics. 2006;173: 1329–1336.

20. Cope ED. The Crocodilians, Lizards, and Snakes of North America. Smithsonian; 1900.

21. Stejneger L. Annotated List of Reptiles and Batrachians collected by Dr. C. Hart Merriam and party in Idaho, 1890. N Am Fauna. 1891; 109–114.

22. Axtell RW. A solution to the long neglected *Holbrookia lacerata* problem, and the description of two new subspecies of *Holbrookia*. Chicago Academy of Sciences; 1956. Available: http://webapps.fhsu.edu/ksherp/bibFiles/854.pdf

23. Hibbitts TJ, Ryberg WA, Harvey JA, Voelker G, Lawing AM, Adams CS, et al. Phylogenetic structure of *Holbrookia lacerata* (Cope 1880) (Squamata: Phrynosomatidae): one species or two? Zootaxa. 2019;4619: zootaxa.4619.1.6.

24. Duran M, Axtell RW. A Rangewide Inventory and Habitat Model for the Spot-tailed

Earless Lizard, Holbrookia lacerata. Report submitted to Texas Parks and Wildlife Department. 2010.

25. Wolaver BD, Pierre JP, Labay BJ, LaDuc TJ, Duran CM, Ryberg WA, et al. An approach for evaluating changes in land-use from energy sprawl and other anthropogenic activities with implications for biotic resource management. Environ Earth Sci. 2018;77: 171.

26. Pierre JP, Wolaver BD, Labay BJ, LaDuc TJ, Duran CM, Ryberg WA, et al. Comparison of Recent Oil and Gas, Wind Energy, and Other Anthropogenic Landscape Alteration Factors in Texas Through 2014. Environ Manage. 2018;61: 805–818.

27. Wolaver BD, Pierre JP, Ikonnikova SA, Andrews JR, McDaid G, Ryberg WA, et al. An Improved Approach for Forecasting Ecological Impacts from Future Drilling in Unconventional Shale Oil and Gas Plays. Environ Manage. 2018;62: 323–333.

28. Ingram M. The Endangered Species Act in Texas: a survey and history. Austin (Texas): Texas Policy Foundation. 2017. Available: https://www.texaspolicy.com/library/doclib/2017-06-study-endangeredspeciessurvey-acee-mingram-2.pdf

29. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinforma. Oxf. Engl. 30, 2114--2120. 2014.

30. Macmanes MD. On the optimal trimming of high-throughput mRNA sequence data. Front Genet. 2014;5: 13.

31. Visser EA, Wegrzyn JL, Steenkmap ET, Myburg AA, Naidoo S. Combined de novo and

genome guided assembly and annotation of the Pinus patula juvenile shoot transcriptome. BMC Genomics. 2015;16: 1057.

32. Chopra R, Burow G, Farmer A, Mudge J, Simpson CE, Burow MD. Comparisons of de novo transcriptome assemblers in diploid and polyploid species using peanut (*Arachis spp.)* RNA-Seq data. PLoS One. 2014;9: e115055.

33. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9: 357–359.

34. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics. 2006;22: 1658–1659.

35. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics. 2015;31: 3210–3212.

36. Haas B, Papanicolaou A, Others. TransDecoder (find coding regions within transcripts). Github, nd https://github com/TransDecoder/TransDecoder (accessed May 17, 2018). 2015.

37. Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 2016;44: D279–85.

38. McGaugh SE, Bronikowski AM, Kuo C-H, Reding DM, Addis EA, Flagel LE, et al. Rapid molecular evolution across amniotes of the IIS/TOR network. Proc Natl Acad Sci U S A. 2015;112: 7055–7060.

39. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol. 2015;16: 157.

40. Ranwez V, Harispe S, Delsuc F, Douzery EJP. MACSE: Multiple Alignment of Coding SEquences accounting for frameshifts and stop codons. PLoS One. 2011;6: e22594.

41. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4: 44–57.

42. Yang Z. Computational Molecular Evolution: Oxford University Press. New York. 2006.

43. Zhang J. Evaluation of an Improved Branch-Site Likelihood Method for Detecting Positive Selection at the Molecular Level. Molecular Biology and Evolution. 2005. pp. 2472–2479. doi:10.1093/molbev/msi237

44. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 2007;24: 1586–1591.

45. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 2005;21: 3674–3676.

46. Waterhouse RM, Seppey M, Simão FA, Manni M, Ioannidis P, Klioutchnikov G, et al. BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. Mol Biol Evol. 2018;35: 543–548.

47. Blankers T, Townsend TM, Pepe K, Reeder TW, Wiens JJ. Contrasting global-scale

evolutionary radiations: phylogeny, diversification, and morphological evolution in the major clades of iguanian lizards. Biol J Linn Soc Lond. 2013;108: 127–143.

48. Garner BA, Hand BK, Amish SJ, Bernatchez L, Foster JT, Miller KM, et al. Genomics in Conservation: Case Studies and Bridging the Gap between Data and Application. Trends in ecology & evolution. 2016. pp. 81–83.

49. Duan J, Sanggaard KW, Schauser L, Lauridsen SE, Enghild JJ, Schierup MH, et al. Transcriptome analysis of the response of Burmese python to digestion. Gigascience. 2017;6: 1–18.

50. Cooke MS, Evans MD, Dizdaroglu M, Lunec J. Oxidative DNA damage: mechanisms, mutation, and disease. FASEB J. 2003;17: 1195–1214.

51. Pamplona R, Costantini D. Molecular and structural antioxidant defenses against oxidative stress in animals. Am J Physiol Regul Integr Comp Physiol. 2011;301: R843–63.

52. Sullivan JP, Lundberg JG, Hardman M. A phylogenetic analysis of the major groups of catfishes (Teleostei: Siluriformes) using rag1 and rag2 nuclear gene sequences. Mol Phylogenet Evol. 2006;41: 636–662.

53. Crandall KA, Bininda-Emonds OR, Mace GM, Wayne RK. Considering evolutionary processes in conservation biology. Trends Ecol Evol. 2000;15: 290–295.

54. Flanagan SP, Forester BR, Latch EK, Aitken SN, Hoban S. Guidelines for planning genomic assessment and monitoring of locally adaptive variation to inform species conservation. Evol Appl. 2018;11: 1035–1052.

55. Watkins JF, Sung P, Prakash L, Prakash S. The Saccharomyces cerevisiae DNA repair gene RAD23 encodes a nuclear protein containing a ubiquitin-like domain required for biological function. Mol Cell Biol. 1993;13: 7757–7765.

56. Iovine B, Iannella ML, Bevilacqua MA. Damage-specific DNA binding protein 1 (DDB1): a protein with a wide range of functions. Int J Biochem Cell Biol. 2011;43: 1664–1667.

57. Briegel KJ, Joyner AL. Identification and characterization of Lbh, a novel conserved nuclear protein expressed during early limb and heart development. Dev Biol. 2001;233: 291–304.

58. Powder KE, Cousin H, McLinden GP, Craig Albertson R. A nonsynonymous mutation in the transcriptional regulator lbh is associated with cichlid craniofacial adaptation and neural crest cell development. Mol Biol Evol. 2014;31: 3113–3124.

59. Ney PA. Mitochondrial autophagy: Origins, significance, and role of BNIP3 and NIX. Biochim Biophys Acta. 2015;1853: 2775–2783.

60. Imazu T, Shimizu S, Tagami S, Matsushima M, Nakamura Y, Miki T, et al. Bcl-2/E1B 19 kDa-interacting protein 3-like protein (Bnip3L) interacts with Bcl-2/Bcl-xL and induces apoptosis by altering mitochondrial membrane permeability. Oncogene. 1999. pp. 4523–4529. doi:10.1038/sj.onc.1202722

61. Grant GA. D-3-Phosphoglycerate Dehydrogenase. Front Mol Biosci. 2018;5: 110.

62. Tabatabaie L, de Koning TJ, A J J, van den Berg IET, Berger R, Klomp LWJ. Novel mutations in 3-phosphoglycerate dehydrogenase (PHGDH) are distributed throughout the

protein and result in altered enzyme kinetics. Human Mutation. 2009. pp. 749–756. doi:10.1002/humu.20934

63. Fuchs SA, Dorland L, de Sain-van der Velden MG, Hendriks M, Klomp LWJ, Berger R, et al. D-serine in the developing human central nervous system. Ann Neurol. 2006;60: 476–480.

Chapter 2

Parthenogenesis doubles the rate of amino acid substitution in whiptail mitochondria

Jose Maldonado[1,2], Thomas J. Firneno Jr.[1], and Matthew Fujita[1,2]

[1]Amphibian and Reptile Diversity Research Center, Department of Biology, The University of Texas at Arlington, Arlington, TX 76019

Abstract

Sexual reproduction is ubiquitous in the natural world, suggesting that sex must have extensive benefits to overcome the cost of males compared to asexual reproduction. One hypothesized advantage of sex with strong theoretical support is that sex plays a role in removing deleterious mutations from the genome. Theory predicts that transitions to asexuality should lead to the suppression of recombination and segregation and, in turn, weakened natural selection, allowing for the accumulation of slightly deleterious mutations. We tested this prediction by estimating the d*N*/d*S* ratios in asexual vertebrate lineages in the genus *Aspidoscelis* using whole mitochondrial genomes from seven asexual and five sexual species. We found higher d*N*/d*S* ratios in asexual *Aspidoscelis* species, indicating that asexual whiptails accumulate non-synonymous substitutions due to weaker purifying selection. Additionally, we estimated nucleotide diversity and found that asexuals harbor significantly less diversity. Thus, despite their recent origins, slightly deleterious mutations accumulated rapidly enough in asexual lineages to be detected. We provided empirical evidence to corroborate the connection between asexuality and increased amino acid substitutions in asexual vertebrate lineages.

Introduction

The recognition of the paradox of sex has spearheaded efforts to identify the advantages of sex that make up for its high cost (Williams and Mitton 1973; Maynard-Smith 1978; Otto and Lenormand 2002). Theory predicts that the per-capita birth rate of asexual populations should rapidly cause the displacement of sexual populations. This is because asexual females only invest in daughters, unlike sexual individuals that must also invest in males; this is what John Maynard-Smith coined the "cost of males" (Maynard-Smith 1971, 1978). Yet, asexual reproduction is rare in nature while sexual reproduction is pervasive, a pattern that has led to the long-standing and unresolved mystery that has been labeled as the queen of problems in evolutionary biology: "Why is there sex?" (Bell 1982; Barton and Charlesworth 1998).

The mutation-recombination model is one theory for the prevalence of sexual reproduction and is also supported by a substantial body of theoretical work (Otto and Gerstein 2006). Transitioning to asexuality abolishes the benefits of recombination and segregation and thrusts the whole genome into complete linkage (Hill and Robertson 1966). The tight association between all loci reduces the effective population size of linked genes, which weakens the efficacy of natural selection to act independently on individual loci and strengthens the influence of genetic drift on the fate of *de novo* mutations resulting in an irreversible accumulation of mildly deleterious mutations (Muller's ratchet), a decline of individual fitness, and eventual population extinction (Muller 1964). In contrast to asexual reproduction, sex breaks the linkage between genes via recombination and segregation, thereby increasing the efficiency of selection to remove deleterious variants and increase the frequency of beneficial mutations.

Among vertebrates, only squamates (lizards and snakes) have multiple species that are obligate parthenogens, which exhibit clonal reproduction via meiosis (Vrijenhoek et al. 1989;

Kearney et al. 2009). The hypothesized cellular mechanisms to produce genetically identical daughters via meiosis in obligately parthenogenetic whiptails requires a second doubling of DNA before entering meiosis (premeiotic endoreplication), allowing for the pairing of sister chromosomes to maintain heterozygosity, and finally the production of unreduced eggs during meiosis II (Cuellar 1971; Lutes et al. 2010; Newton et al. 2016). Obligately parthenogenetic lizards and snakes have arisen multiple times independently across nine families but represent <1% of squamate diversity (Fujita et al. 2020). Whiptail lizards in the genus *Aspidoscelis* have 15 known all-female asexual species that have hybrid origins: hybridization events between two distinct extant sexual species resulted in initial diploid parthenogens, with backcrossing or even further hybridization with a third sexual species producing triploids (Neaves and Baumann 2011). Different diploid asexual whiptail species have independent origins from one another, though it is still unclear whether single or multiple hybridization events occurred within a species (Reeder et al. 2002).

Sex chromosomes in mammals and birds provide insight into the long-term (>10 million years) evolutionary trajectory of the genome in the absence of recombination (Graves 2006). However, obligately parthenogenetic lizards provide an opportunity to examine mutation accumulation at much more recent timescales (<1 million years). Furthermore, parthenogenetic whiptail lizards are exceptional systems to address questions surrounding the evolution of sex because they serve as a contrasting model to how sexual reproduction influences genome evolution (Kearney et al. 2009; Fujita et al. 2020). Due to their high abundance and wide distribution throughout the southwestern U.S. and northern Mexico, asexual whiptail lizards (*Aspidoscelis*) are excellent model systems to investigate the fundamental cellular mechanisms of parthenogenesis and the genomic consequence of asexuality (Cuellar 1971; Lutes et al. 2010;

Newton et al. 2016). A few studies have begun leveraging parthenogenetic genomes to test whether weakened selection and mutation accumulation are characteristic of asexual vertebrate populations (Moritz 1991; Fujita et al. 2007; Boussau et al. 2011).

Most studies investigating the genetic consequences of asexuality using DNA sequence data have primarily focused on asexual invertebrates and plants. These studies have yielded mixed results when examining whether or not mutation accumulation occurs in asexual lineages. Two studies using asexual stick insects in the genus *Timema* found a higher rate of non-synonymous to synonymous mutations in both mitochondrial and nuclear coding genes of asexual populations (Henry et al. 2012; Bast et al. 2018). The New Zealand snail *Potamopyrgus antipodarum* has mixed populations of sexual and asexual individuals. The results demonstrated an increased point mutation accumulation in mitochondrial coding genes, with asexual snails having a d*N*/d*S* ratio two-times higher than sexuals (Neiman et al. 2010). Transcriptome- and genome-wide studies respectively found that two distinct asexual plant species exhibited a higher load of deleterious mutations than sexual plant species (Hollister et al. 2015; Lovell et al. 2017). On the other hand, recent studies using whole genomic- and transcriptomic-wide data have found evidence that is contrary to the expectation of mutation accumulation in asexual populations. Transcriptome-wide data from oribatid mites revealed that purifying selection was more effective in the extremely old asexual lineages (~10 million years old) than in sexual lineages (Brandt et al. 2017, 2021). Non-recombining portions of the genome are expected to accumulate transposable elements (TE) and other repetitive content that leads to genomic degradation, such as in the Y-chromosomes (Junakovic et al. 1998; Bachtrog 2013). Yet, genome-wide data from several distinct asexual crustacean species had no evidence of increased TE content (Bast et al. 2016). Similarly, asexual aphid lineages show no evidence of increased mutation accumulation

in mitochondrial genes, and only one lineage had increased mutation accumulation in nuclear genes (Normark and Moran 2000). The mixed results between studies show that the expected effect of suppressing recombination across the whole genome does not always lead to the detrimental genetic consequences we expect based on our predicted theory (Jaron et al. 2021). The differences among asexual species, such as the cellular mechanism of asexuality, the age of the asexual lineage, and the method of transitioning to asexuality, could affect the genetic diversity across distinct asexual species.

Our study uses whole mitochondrial genome data from asexual and sexual whiptail lizards to investigate our prediction that parthenogenetic lineages accumulate mutations faster than sexual lineages. We calculated d*N*/d*S* ratios between asexual and sexual lineages, carried out selection tests on asexual lineages to see if these lineages have undergone positive selection, estimated nucleotide diversity levels between asexual and sexual groups, and estimated the age of asexual lineages.

Methods

We sampled multiple populations of both asexual and sexual whiptail species throughout the southwestern United States and received additional tissue samples from the Burke Museum and American Museum of Natural History collections. We collected individuals by hand, lasso, or using rubber bands in accordance with our IACUC protocol (A13.010). We humanely euthanized each individual and collected liver and skeletal tissue, and stored them in RNAlater (Thermo Fisher Scientific, Waltham, MA). Specimens were preserved using 10% formalin, stored in ethanol, and deposited at the Amphibian & Reptile Diversity Research Center at the University of Texas at Arlington. We sequenced 86 individuals: 18 *A. tesselata*, 14 *A.*

neotesselata, 14 *A. marmorata*, 10 *A. inornata*, 8 *A. gularis*, 7 *A. exsanguis*, 5 *A. sexlineata*, 4 *A. neomexicana*, 2 *A. dixoni* and *A. laredoensis*, 1 *A. velox,* and *A. scalaris*. Supplementary table 1 has the locality data associated with each sequenced sample.

We extracted DNA from liver tissue using a standard phenol-chloroform protocol and measured the concentrations using a Qubit 2.0 (Invitrogen, Carlsbad, CA). To sequence the whole mitochondrial genome, we followed the protocol from (Roelke et al. 2018). Briefly, we used exonucleases that specifically remove double-stranded linear DNA. The remaining intact circular mitochondrial genomes are then isolated using SeraPure beads (Rohland and Reich 2012) and resuspended in fresh buffer (10 mM Tris, pH 8.0). We amplified the mitochondrial genome using Φ29 DNA polymerase before following standard Illumina library preparation to generate genomic libraries. Each individual was dual-labeled with inline barcodes and standard Illumina indices before sequencing on the Illumina MiSeq (Illumina) to generate 300 bp paired-end reads.

Our sequenced reads were filtered using the FASTX-Toolkit to remove low-quality reads and reads shorter than 50bp long. We assembled the genomes using the CLC Genomics WB 9 (Qiagen, Germantown, MD) short-read assembler. We used the Mitos web server (Bernt et al. 2013) to annotate the mitochondrial genome for protein-coding, RNA, and tRNA genes. Because the origins of parthenogenesis in vertebrates are essentially founder events, we expect to see limited genetic diversity in parthenogenetic whiptails compared to sexual lineages. To examine this, we calculated nucleotide diversity (π) for each whiptail species using DNAsp v.6 (Librado and Rozas 2009).

We used the Geneious R6 aligner to align each individual protein-coding gene and both rRNA genes. We then used a concatenated alignment to identify the best partitioning scheme and

evolutionary models with PartitionFinder 2 (Lanfear et al. 2017) using the following parameters: linked branch lengths, the greedy algorithm, BIC for our selection model, and the following models of evolution that can be implemented in BEAST. To estimate divergence times and to obtain a rooted tree, we used a Bayesian statistical framework with BEAST v2.6.6 (Bouckaert et al. 2014) with *Ameiva undulata*, *Teius teyou*, *Gymnophthalmus speciosus*, and *Podarcis muralis* used as outgroups. We used two calibration points with uniform priors: (1) fossil of an ancestor for living cnemidophorines (GHUNLPam21745, 9.0-10.0 Ma) with lower and upper bounds of 9 and 86 Ma, respectively (Albino et al. 2013; Tucker et al. 2017) ; and (2) the node of Gymnophthalmoidea (Teiidae + Gymnophthalmidae) with lower and upper bounds of 70 and 86 Ma, based on previous squamate studies (Vidal and Hedges 2009; Pyron 2010; Mulcahy et al. 2012; Tucker et al. 2017). We performed the analysis on our concatenated mitochondrial dataset since the mitochondrial genes are linked together, using a GTR+I nucleotide substitution model (as predicted by PartitionFinder), a log-normal relaxed molecular clock, and a birth-death model. We ran the analysis of three independent runs with 20,000,000 generations, sampling every 1,000 generations, producing a total of 20,000 trees. The runs were assessed using Tracer v1.6 (Rambaut 2009) to examine convergence and runs were combined in LogCombiner v2.6.6. A burn-in of 10% was discarded, and a maximum clade credibility (MCC) tree with median heights was created from the remaining trees. To produce an unrooted phylogenetic tree needed for carrying out the d*N*/d*S* analysis, we used MrBayes and PartitionFinder 2 with the same parameters described above (Figure 1). We first removed three samples that failed to assemble into the complete ~16kb genome as we could not allow missing data for our d*N*/d*S* analysis. Additionally, we removed the outgroups from the concatenated dataset because it can only be classified as sexual (Neiman et al. 2010).

We carried out our d*N*/d*S* analysis on our concatenated alignment of 13 protein-coding genes using the branch model implemented in the codeml program from the PAML V4.9 package (Yang 2007). Codeml works by calculating the fitness of our specified model of evolution to a phylogenetic data set and uses likelihood ratio tests to assess which model of evolution provides the best fit to the data. We used codeml to compare the goodness of fit of our specified evolutionary models to our unrooted phylogenetic tree and protein-coding sequence data. These ranged from the simplest models where the entire whiptail phylogeny experienced the same rate of evolution to more complex models that allow sexual and asexual whiptail lineages to evolve at different rates. We first ran codeml's one-ratio maximum likelihood model (M0: one omega ratio for all branches) to estimate a single d*N*/d*S* value for our unrooted whiptail phylogeny. The second model we ran was the two-ratio model, where sexual and asexual lineages were allowed to have different d*N*/d*S* values (M2: two or more d*N*/d*S* ratios for branches). We then carried out a likelihood ratio test to assess whether the two-ratio or the one-ratio model provided a better fit to the data. An elevated d*N*/d*S* value for the asexual lineages and a significantly better fit of the two-ratio model to the data would imply that asexual whiptails accumulate harmful mutations faster than sexual whiptails (Neiman et al. 2010). Additionally, we estimated the d*N*/d*S* ratio per gene using the M0 and M2 model to test if we see elevated non-synonymous substitutions in asexuals.

The young age of asexual whiptails lineages (Cullum 1997; Birky 2010) can confound the differences between sexual and asexual d*N*/d*S* due to differences in branch length between sexual and asexual whiptail lineages. The time lag between the origin of new mutations and their removal by purifying selection results in a higher estimate of mutation accumulation in terminal branches (Neiman et al. 2010). To address this issue, we used codeml to evaluate the fitness of

two additional models to our data that allowed the terminal and internal branches to have different evolutionary rates (terminal vs. internal branches). The additional model we ran was similar to Neiman et al. 2010, where we categorized branches in our unrooted phylogenetic tree as terminal sexual, terminal asexual, internal sexual, and internal asexual (four-ratio model). We used likelihood ratio tests to compare the fit of the model of the terminal vs. internal two ratio-model to our one-ratio model and compare the four-ratio model to both two-ratio models (asexual vs. sexual and terminal vs. internal) and the one-ratio model. As a simple test to make sure neutral evolution did not produce our empirical results, we generated a simulated nucleotide data set using the Evolver codon setting from the PAML V4.9 package (Yang 2007), and we reran the one-ratio vs. two-ratio model and compared the results from our simulation to our empirical dataset. Evolver generated our simulated DNA sequences by using our unrooted tree, and using the Monte Carlo simulation model, and we kept all other parameters at default.

If asexual whiptails lineages exhibit a higher d*N*/d*S* ratio than their sexual counterpart, we had to test that the inflated ratio was not caused by positive selection but by an increase of non-synonymous substitutions (d*N*). To do so, we ran the RELAX and aBSREL selection test on our data set found in the Hypothesis Testing using Phylogenies V2.3 software package (Pond et al. 2005). The RELAX selection test asks whether the strength of natural selection has been intensified or relaxed along specified set test branches. In our RELAX selection test, we labeled asexual branches as our test branches and sexual branches as references branches. RELAX uses the parameter k, which is the selection intensity parameter, where $k > 1$ indicates selection has been intensified and $k < 1$ selection has been relaxed along our specified test branches (Wertheim et al. 2015). The aBSREL model tested if asexual branches in our phylogeny have evolved under positive selection. To do so, aBSREL tests whether a proportion of sites have

evolved under positive selection in our specified branches of interest. aBSREL then runs Likelihood Ratio Tests to compare the adaptive model ($\omega > 1$) to the null model (Smith et al. 2015).

Results

De novo Assembly, gene annotations, and phylogenetic analysis

We sequenced 86 whiptail mitochondrial genomes. Three samples were missing several genes and failed to assemble into the complete ~16kb mitochondrial genome. Mitos successfully annotated protein-coding genes and identified the start and stop codons, both rRNA genes, and all tRNA genes. To conduct the d*N*/d*S* analysis, we removed the three individuals with incomplete mitochondrial genome assemblies to generate an unrooted tree (Figure 1). The divergence dating analysis of our concatenated dataset strongly supports an Eocene divergence (36.9 Mya [95% HPD 28.0-46.8 Mya]) between the *Aspidoscelis marmorata* group (including *marmorata* and the parthenogenetic *tesselata*, *neotesselata*, and *neomexicana*) from the remaining *Aspidoscelis* in our study. The *marmorata* clade dates to within the Pleistocene (1.26 Mya [95% HPD 0.54-1.77 Mya]), as does the origin of *A. tesselata* (0.36 Mya [95% HPD 0.1-0.4 Mya]. Other parthenogenetic lineages (*exsanguis* and *laredoensis*) have Pleistocene origins. The ESS values for our combined runs were all ≥200.

d*N*/d*S* analysis and nucleotide diversity

We found that the two-ratio model (M2) fit our data better when comparing the two-ratio to a one-ratio model (M0) that only allows for one d*N*/d*S* for the entire tree. PAML showed that asexuals have ~2x higher d*N*/d*S* than sexual whiptail lineages (d*N*/d*S* = 0.0725 for asexuals vs.

0.0395 for sexuals; Table 1). The four-ratio model allows four d*N*/d*S* estimates between internal and terminal branches for both asexuals and sexual whiptails. The four-ratio model had a significantly better fit to the data than the two-ratio model (Asexual vs. Sexual) and one-ratio model (Table 1). Our d*N*/d*S* estimates per gene found that six out of the thirteen mitochondrial coding genes in asexual lineages had a significantly higher d*N*/d*S* than their sexual lineages (Table 2). We found that the other seven genes had no differences in d*N*/d*S* between both groups. While asexual whiptail lineages had an overall elevated d*N*/d*S*, we generally found higher levels of polymorphisms in sexual whiptail species. *Aspidoscelis sexlineata* is the only sexual whiptail species with a slightly lower nucleotide diversity of $\pi = 0.0012$ than the asexual *A. tesselata* species with $\pi = 0.0017$. We found the sexual species with the highest nucleotide diversity was *A. gularis* ($\pi = 0.0160$), while the asexual species with the lowest nucleotide diversity was the triploid *A. neotesselata* ($\pi = 0.0006$) (Table 3). We ran both M0 and M2 models on our simulated nucleotide dataset generated by Evolver and found no significant differences in d*N*/d*S* estimate between asexual and sexual whiptail lizards, indicating that our results deviate from the null expectation of our simulations.

Selection Test

To ensure that the elevated d*N*/d*S* we found in asexual lineages was due to weakened purifying selection and not positive selection, we carried out selection tests to determine if our asexual branches underwent positive selection. Both RELAX and aBSREL selection tests found no evidence of positive selection in asexuals. Our test of relaxation or intensification among asexual branches using the RELAX model found significant (p-value < 0.05) relaxation in asexual lineages with a $k = 0.39$. aBSREL tested our branches of interest for diversifying

selection and found no significant evidence of episodic diversifying selection in our asexual branches. To determine significance aBSREL assessed the alternative to the null model using the Likelihood Ratio Test at a threshold of $p \leq 0.05$, after correcting for multiple testing.

Discussion

Our study used whole mitochondrial genomes to study the genomic consequence of asexuality in parthenogenetic lizards. We found that the transition to asexuality led to relaxed natural selection in parthenogenetic lizards and the build-up of non-synonymous mutations. Our findings support theoretical predictions that the loss of sex should lead to an irreversible build-up of deleterious mutations due to a reduction in the efficiency of purifying selection, and sex facilitates the removal of harmful mutations.

Our empirical data and results are similar to those of other younger and ancient asexual lineages (Neiman et al. 2010; Bast et al. 2018) and other non-recombining genomes, such as sex chromosomes. Non-recombining sex chromosomes have evolved multiple times and accumulated both point and structural mutations, and many have degenerated to house just a few protein-coding genes (Vicoso 2019). The degradation of the mammalian Y and avian W sex chromosomes from the lack of recombination provided some of the best evidence for the long-term consequences of clonal inheritance (Fujita et al. 2020). Even the recently formed neo-Y sex chromosome that occurred one million years ago (Bachtrog and Charlesworth 2002) in *Drosophila miranda* has rapidly accumulated mutations and repetitive DNA (Bachtrog et al. 2008). In whiptails, studies have identified structural mutations in the mitochondrial genome in several asexual whiptail species that exhibit tandem duplications that are adjacent to tRNA, rRNA, or protein-coding genes (Moritz and Brown 1987; Stanton et al. 1994). Our results of

increased rates of non-synonymous mutations (mutation accumulation) in parthenogenetic asexual whiptail mitochondrial genomes are consistent with the findings of these other studies.

Our findings reveal that asexual whiptails exhibit lower levels of polymorphism compared to sexual species. Factors that can generate a disparity of genetic diversity between asexual and sexual populations include: the age of the asexual lineage, the number of hybridization events that occurred during the formation of the parthenogenetic lineages (the severity of the bottleneck during the origin of parthenogenesis), and the heterozygosity generated by those hybridization events.

Our phylogenetic analysis can address the first of these three factors. It is important to note that, while we used similar methods to Tucker et al. 2017 to estimate divergence dates, their analysis places the divergence between the cnemidophorines and *Teius* (a node which they denote as Teiinae) at ~35 Mya, whereas our analysis places this split at ~62.5 Mya, which is closer to their estimated node of Teiidae. This may be due to our using only two nodes for calibration within this group and/or less taxonomic coverage throughout this group. Nevertheless, we were still able to show that the origins of asexual populations are much younger than those of the sexual populations. For example, parthenogenetic *Aspidoscelis tesselata* have a recent origin well within the Pleistocene (0.36 [0.1-0.4] Mya), which is the general pattern with regard to the age of asexual lineages in vertebrates (Fujita et al. 2020). Further evidence for the recency of *A. tesselata* is that the maternal lineage, *A. marmorata*, is still extant.

Based on our sampling across its distribution, we inferred a single origin of *A. tesselata* (or at most, a few origins). With so few origins, we expect such a bottleneck to result in low nucleotide diversity in *A. tesselata*. Compared to *A. marmorata*, this is exactly what we see; *A.*

marmorata has nearly three times as much nucleotide diversity (θ_W) than *A. tesselata*. Furthermore, *A. neotesselata* has a nuclear diversity nearly ten times less than *A. marmorata*, which is consistent with an even more recent origin than *A. tesselata*. These patterns of reduced diversity and recent hybrid origins that we see in *A. tesselata* contrasts with some other parthenogenetic systems, such as the darwinulid ostracods and several lineages of asexual oribatid mites. In these latter systems, divergence has accumulated via asexual speciation (William Birky and Barraclough 2009), with the existence of ancient (>10 million years) parthenogenetic lineages implying that they have avoided the long-term consequences of Muller's Ratchet (Heethoff et al. 2009; Schön et al. 2009). Given the contrasting origins of ancient asexuals and parthenogenetic lizards, it is difficult to directly compare the evolutionary mechanisms that generate genetic diversity between the two groups.

High heterozygosity is often a signature of genomes from asexual organisms compared to their sexual relatives (Fujita et al. 2020). There are two main mechanisms that can produce this high heterozygosity in asexual lineages compared to sexual lineages. First, in the absence of recombination, haplotypes can diverge in what is known as the Meselson effect in ancient asexuals (Birky 1996; Welch et al. 2000). Recent genomic data has demonstrated the Meselson effect and the long-term evolution in oribatid mites (Brandt et al. 2021). Second, hybridization between divergent sexual progenitors instantaneously produces asexual lineages with high heterozygosity; this is the mechanism seen in nearly all parthenogenetic lizards (Fujita et al. 2020). Our mitochondrial data are unable to provide insight into whether the Meselson effect can be seen in parthenogenetic lizards, though we expect that any heterozygosity contributed by such divergence will be low compared to that contributed by hybrid origins.

Previous studies have found that even young asexual lineages (from ~100,000 to ~500,000 years old) have experienced an increased rate of non-synonymous substitutions in mitochondrial genes compared to sexual lineages (Johnson and Howard 2007; Neiman et al. 2010; Henry et al. 2012). Our findings are in line with the theory that purifying selection in the mitochondrial genomes acts less effectively in asexual populations than in their sexual counterparts. We carried out several d*N*/d*S* analyses to test for mutation accumulation in protein-coding genes between asexual and sexual whiptail lizards. Using all thirteen protein-coding genes, we found that asexual lineages accumulate non-synonymous substitutions at a faster rate, despite the relatively young age of these asexual lineages. Further, we found that this rate is due to relaxed selection rather than positive selection. While the uniparentally inherited and clonally reproducing mitochondrial genome exhibits no recombination, it can still serve as a valuable marker to investigate the genomic consequences of asexuality. In sexual populations, the nuclear genome and the mitochondrial genome still segregate from each other, while a transition to asexuality will force the nuclear and mitogenome into complete linkage. The tight linkage between the nuclear and mitochondrial genomes in asexuals should reduce their effective population sizes, resulting in less efficient selection and eventually an accumulation of non-synonymous variants. Additionally, the mitochondrial genomes can be a useful marker to test the genomic consequences of asexuality, since mitochondria DNA evolves at a higher rate than nuclear DNA (Brown et al. 1979; Xia 2012).

Sex facilitates in the long-term survival of metazoan lineages by allowing alleles to separate between generations (segregation) and by shuffling genetic variation (crossing over and independent assortment); thus, sex increases both genetic diversity and the efficiency of natural selection, allowing populations to purge deleterious alleles and to bring together beneficial

variants. On the other hand, asexuality is often viewed as an evolutionary dead-end because of the long-term consequences of accumulating harmful mutations and the reduced capabilities to adapt to novel environments (Vrijenhoek 1998; Butlin 2002). While some studies have seen possible evidence of adaptive evolution in asexual organisms (Jaron et al. 2021), our study is more consistent with the expectations that reduced effective population sizes (due to asexuality and complete linkage) results in less effective natural selection, leading to the accumulation of mutations.

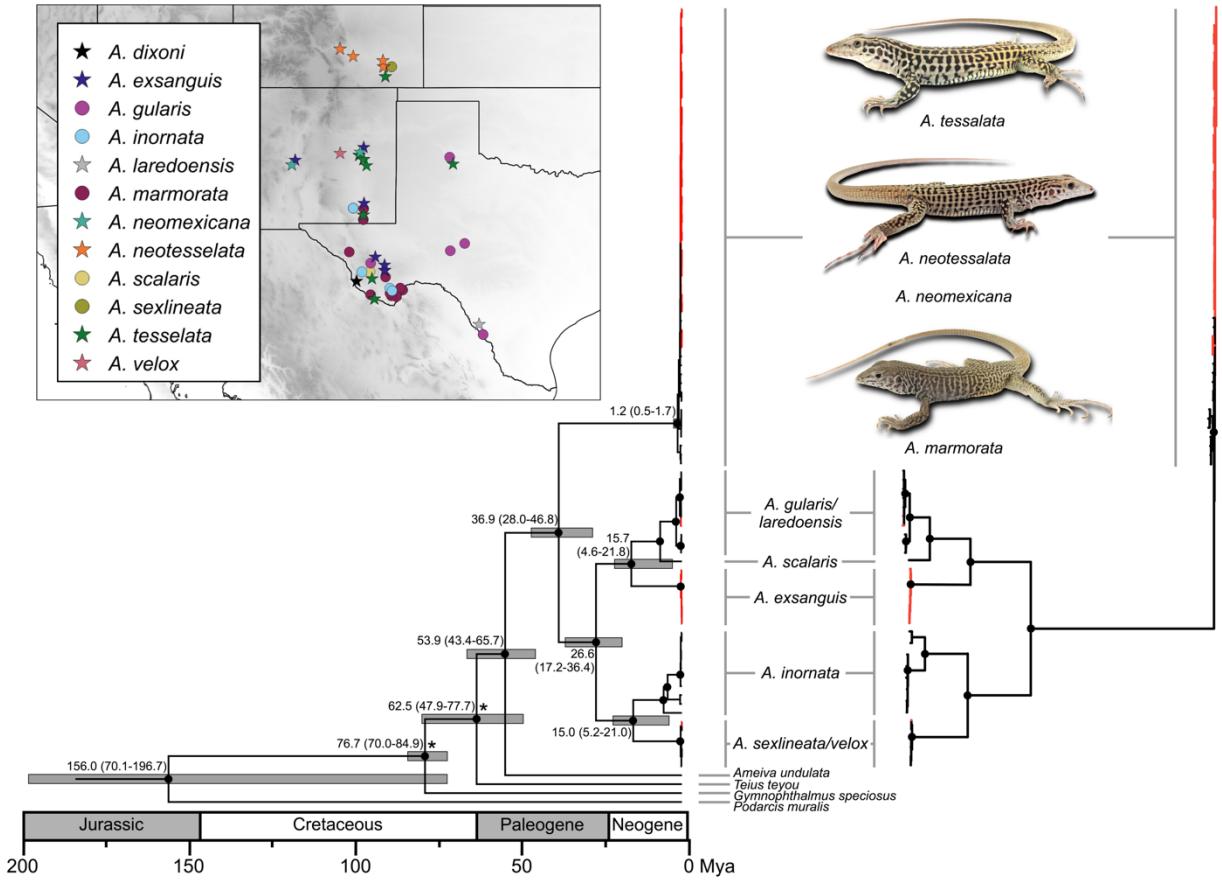

Figure 1.

Sampling map of all individuals used in genetic analyses (asexuals denoted with stars; sexuals denoted with circles). Samples of *A. neotesselata* (ADL 4935-4938) from Washington are not shown in the map, but were used for all analyses. Bayesian chronogram (left) and unrooted Bayesian phylogram (right) were made from our concatenated dataset. Red branches on both trees indicate asexual individuals, whereas black branches represent sexual individuals. Node/fossil calibrations are indicated by *. Median ages, 95% highest probability densities (HPD), and bars representing the 95% HPDs are provided. Nodes on both trees with high support (posterior probability ≥ 0.95) are indicated by black dots. The unrooted Bayesian phylogram had posterior probabilities >0.50 and the tree topology was used for all d*N*/d*S* ratio analyses.

Coding genes	Number of free parameters	ln(L)	*p*-values	d*N*/d*S*
One Ratio	142	−36392.0908	NA	0.0429
Two Ratio (sex vs. asex)	143	−36380.4102	vs. one ratio: 1.343e-06	Sex: 0.03925
				Asex: 0.07251
Two Ratio (T vs. I)	143	−36359.4881	vs. one ratio: 6.7494e-16	I: 0.03750
				T: 0.1018
Four Ratio (SexT, AsexT, SexI, AsexI)	146	−36344.8913	vs. one ratio: 1.530e-19	SexT: 0.0912
			vs. two ratio (A vs. S): 2.559e-15	AsexT: 0.1903
			vs. two (T vs. I): 2.039e-06	SexI: 0.0380
				AsexI: 0.0613

"T" indicates terminal branches and "I" indicates internal branches.

Table 1.

Summary of PAML analyses and comparisons of d*N*/d*S* Ratios in coding mitochondrial DNA genes in Sexual versus Asexual whiptails.

Gene	One ratio d*N*/d*S*	Two ratio d*N*/d*S*	*p*-values
COB	0.0173	Sex: 0.0140 Asex: 0.0517	0.0079
COX III	0.0271	Sex: 0.0227 Asex: 0.0719	0.0413
NAD I	0.0200	Sex: 0.0160 Asex: 0.0532	0.0132
NAD II	0.0485	Sex: 0.0428 Asex: 0.0909	0.0296
NAD IV	0.0536	Sex: 0.0461 Asex: 0.1035	0.0084
NAD V	0.0508	Sex: 0.0461 Asex: 0.0960	0.0223

Table 2.

Six out of the thirteen protein-coding genes had a higher d*N*/d*S* in asexual lineages.

Species	No. of individuals	Nucleotide diversity π	θW
A. neotesselata	14	0.0006	6E-04
A. exsanguis	7	0.001	0.001
A. neomexicana	4	0.0018	0.002
A. sexlineata	5	0.0012	0.001
A. tesselata	17	0.0017	0.002
A. marmorata	13	0.0047	0.006
A. inornata	8	0.015	0.014
A. gularis	8	0.016	0.019

Table 3.

Nucleotide diversity (π and θW) was estimated from using 13 protein-coding genes for asexual and sexual whiptails.

Supplementary Table 1.

A complete list of our samples we sequenced with their locality data.

https://datadryad.org/stash/dataset/doi:10.5061%2Fdryad.7wm37pvw4

Acknowledgment

We would like to thank the A. Leaché and the Burke Museum, American Museum of Natural History, UTEP Biodiversity Collections, and Texas A&M Biodiversity Research and Teaching Collections for providing us with additional samples we needed to conduct this study. We would also like to thank Maurine Neiman for providing us with the information to properly run the Dn/Ds analysis. I would also like to acknowledge TJF, MKF, and Kathleen Currie for taking photographs of whiptail lizards used in figure 1.

The genetic sequences used in the study will be uploaded and made publicly available on Genbank once the manual script has been accepted into the journal. Additional output files generated from estimating Dn/Ds will be uploaded and made publicly available on Dryad for the public to access our data files.

Jose A. Maldonado conducted the lab work and d*N*/d*S* analyses and wrote the manuscript. Thomas J. Firneno Jr. contributed the time tree analysis and helped with editing and revising the writing. Along with Maldonado, Matthew K. Fujita conceived of the project, funded this work using his lab startup funds, and edited and revised the manuscript.

References

Albino, A. M., C. I. Montalvo, and S. Brizuela. 2013. New Records of Squamates from the Upper Miocene of South America. J. Herpetol. 47:590–598. Allen Press.

Bachtrog, D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. Nat. Rev. Genet. 14:113–124.

Bachtrog, D., and B. Charlesworth. 2002. Reduced adaptation of a non-recombining neo-Y chromosome. Nature 416:323–326.

Bachtrog, D., E. Hom, K. M. Wong, X. Maside, and P. de Jong. 2008. Genomic degradation of a young Y chromosome in *Drosophila miranda*. Genome Biol. 9:R30.

Barton, N. H., and B. Charlesworth. 1998. Why sex and recombination? Science 281:1986–1990. science.sciencemag.org.

Bast, J., D. J. Parker, Z. Dumas, K. M. Jalvingh, P. Tran Van, K. S. Jaron, E. Figuet, A. Brandt, N. Galtier, and T. Schwander. 2018. Consequences of Asexuality in Natural Populations: Insights from Stick Insects. Mol. Biol. Evol. 35:1668–1677.

Bast, J., I. Schaefer, T. Schwander, M. Maraun, S. Scheu, and K. Kraaijeveld. 2016. No Accumulation of Transposable Elements in Asexual Arthropods. Mol. Biol. Evol. 33:697–706.

Bell, G. 1982. The Masterpiece of Nature: The Evolution and Genetics of Sexuality. Routledge.

Bernt, M., A. Donath, F. Jühling, F. Externbrink, C. Florentz, G. Fritzsch, J. Pütz, M. Middendorf, and P. F. Stadler. 2013. MITOS: improved de novo metazoan mitochondrial genome annotation. Mol. Phylogenet. Evol. 69:313–319.

Birky, C. W., Jr. 1996. Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. Genetics 144:427–437.

Birky, C. W., Jr. 2010. Positively negative evidence for asexuality. J. Hered. 101 Suppl 1:S42–5.

Bouckaert, R., J. Heled, D. Kühnert, T. Vaughan, C.-H. Wu, D. Xie, M. A. Suchard, A. Rambaut, and A. J. Drummond. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10:e1003537.

Boussau, B., J. M. Brown, and M. K. Fujita. 2011. Nonadaptive Evolution of Mitochondrial Genome Size. Evolution 65:2706–2711.

Brandt, A., I. Schaefer, J. Glanz, T. Schwander, M. Maraun, S. Scheu, and J. Bast. 2017. Effective purifying selection in ancient asexual oribatid mites. Nat. Commun. 8:873.

Brandt, A., P. Tran Van, C. Bluhm, Y. Anselmetti, Z. Dumas, E. Figuet, C. M. François, N. Galtier, B. Heimburger, K. S. Jaron, M. Labédan, M. Maraun, D. J. Parker, M. Robinson-Rechavi, I. Schaefer, P. Simion, S. Scheu, T. Schwander, and J. Bast. 2021. Haplotype divergence supports long-term asexuality in the oribatid mite *Oppiella nova*. Proc. Natl. Acad. Sci. U. S. A. 118.

Brown, W. M., M. George, and A. C. Wilson. 1979. Rapid evolution of animal mitochondrial DNA. Proceedings of the National Academy of Sciences 76:1967–1971.

Butlin, R. 2002. Evolution of sex: The costs and benefits of sex: new insights from old asexual lineages. Nat. Rev. Genet. 3:311–317.

Cuellar, O. 1971. Reproduction and the mechanism of meiotic restitution in the parthenogenetic lizard Cnemidophorus uniparens. J. Morphol. 133:139–165. Wiley Online Library.

Cullum, A. J. 1997. Comparisons of Physiological Performance in Sexual and Asexual Whiptail Lizards (Genus *Cnemidophorus*): Implications for the Role of Heterozygosity. Am. Nat. 150:24–47. [The University of Chicago Press, The American Society of Naturalists].

Fujita, M. K., J. L. Boore, and C. Moritz. 2007. Multiple origins and rapid evolution of

duplicated mitochondrial genes in parthenogenetic geckos (*Heteronotia binoei*; Squamata, Gekkonidae). Mol. Biol. Evol. 24:2775–2786.

Fujita, M. K., S. Singhal, T. O. Brunes, and J. A. Maldonado. 2020. Evolutionary Dynamics and Consequences of Parthenogenesis in Vertebrates. Annu. Rev. Ecol. Evol. Syst. 51:191–214. Annual Reviews.

Graves, J. A. M. 2006. Sex chromosome specialization and degeneration in mammals. Cell 124:901–914.

Heethoff, M., R. A. Norton, S. Scheu, and M. Maraun. 2009. Parthenogenesis in Oribatid Mites (Acari, Oribatida): Evolution Without Sex. Pp. 241–257 *in* I. Schön, K. Martens, and P. Dijk, eds. Lost Sex: The Evolutionary Biology of Parthenogenesis. Springer Netherlands, Dordrecht.

Henry, L., T. Schwander, and B. J. Crespi. 2012. Deleterious mutation accumulation in asexual *Timema* stick insects. Mol. Biol. Evol. 29:401–408.

Hill, W. G., and A. Robertson. 1966. The effect of linkage on limits to artificial selection.

Hollister, J. D., S. Greiner, W. Wang, J. Wang, Y. Zhang, G. K.-S. Wong, S. I. Wright, and M. T. J. Johnson. 2015. Recurrent loss of sex is associated with accumulation of deleterious mutations in Oenothera. Mol. Biol. Evol. 32:896–905.

Jaron, K. S., J. Bast, R. W. Nowell, T. R. Ranallo-Benavidez, M. Robinson-Rechavi, and T. Schwander. 2021. Genomic Features of Parthenogenetic Animals. J. Hered. 112:19–33.

Johnson, S. G., and R. S. Howard. 2007. Contrasting patterns of synonymous and nonsynonymous sequence evolution in asexual and sexual freshwater snail lineages. Evolution 61:2728–2735.

Junakovic, N., A. Terrinoni, C. Di Franco, C. Vieira, and C. Loevenbruck. 1998. Accumulation

of transposable elements in the heterochromatin and on the Y chromosome of *Drosophila simulans* and *Drosophila melanogaster*. J. Mol. Evol. 46:661–668.

Kearney, M., M. K. Fujita, J. Ridenour, I. Schön, K. Martens, and P. van Dijk. 2009. Lost sex: the evolutionary biology of parthenogenesis.

Lanfear, R., P. B. Frandsen, A. M. Wright, T. Senfeld, and B. Calcott. 2017. PartitionFinder 2: New Methods for Selecting Partitioned Models of Evolution for Molecular and Morphological Phylogenetic Analyses. Mol. Biol. Evol. 34:772–773. academic.oup.com.

Librado, P., and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452.

Lovell, J. T., R. J. Williamson, S. I. Wright, J. K. McKay, and T. F. Sharbel. 2017. Mutation Accumulation in an Asexual Relative of Arabidopsis. PLoS Genet. 13:e1006550.

Lutes, A. A., W. B. Neaves, D. P. Baumann, W. Wiegraebe, and P. Baumann. 2010. Sister chromosome pairing maintains heterozygosity in parthenogenetic lizards. Nature 464:283–286.

Maldonado, J., Firneno Jr, T., Hall, A., Fujita, M. (2022). Parthenogenesis Doubles the Rate of Amino Acid Substitution in Whiptail Mitochondria. Dryad, Dataset, https://doi.org/10.5061/dryad.7wm37pvw4.

Maynard-Smith, J. 1978. The evolution of sex. Cambridge University Press Cambridge.

Maynard-Smith, J. 1971. The origin and maintenance of sex In Williams GC (Ed.), Group selection (pp. 163--175). Chicago, IL: Aldine Atherton.

Moritz, C. 1991. The origin and evolution of parthenogenesis in *Heteronotia binoei* (Gekkonidae): evidence for recent and localized origins of widespread clones. Genetics 129:211–219.

Moritz, C., and W. M. Brown. 1987. Tandem duplications in animal mitochondrial DNAs: variation in incidence and gene content among lizards. Proc. Natl. Acad. Sci. U. S. A. 84:7183–7187.

Mulcahy, D. G., B. P. Noonan, T. Moss, T. M. Townsend, T. W. Reeder, J. W. Sites Jr, and J. J. Wiens. 2012. Estimating divergence dates and evaluating dating methods using phylogenomic and mitochondrial data in squamate reptiles. Mol. Phylogenet. Evol. 65:974–991.

Muller, H. J. 1964. The Relation of Recombination to Mutational Advance. Mutat. Res. 106:2–9.

Neaves, W. B., and P. Baumann. 2011. Unisexual reproduction among vertebrates. Trends Genet. 27:81–88.

Neiman, M., G. Hehman, J. T. Miller, J. M. Logsdon Jr, and D. R. Taylor. 2010. Accelerated mutation accumulation in asexual lineages of a freshwater snail. Mol. Biol. Evol. 27:954–963.

Newton, A. A., R. R. Schnittker, Z. Yu, S. S. Munday, D. P. Baumann, W. B. Neaves, and P. Baumann. 2016. Widespread failure to complete meiosis does not impair fecundity in parthenogenetic whiptail lizards. Development 143:4486–4494.

Normark, B. B., and N. A. Moran. 2000. Testing for the accumulation of deleterious mutations in asexual eukaryote genomes using molecular sequences. J. Nat. Hist. 34:1719–1729. Taylor & Francis.

Otto, S. P., and A. C. Gerstein. 2006. Why have sex? The population genetics of sex and recombination. Biochem. Soc. Trans. 34:519–522.

Otto, S. P., and T. Lenormand. 2002. Resolving the paradox of sex and recombination. Nat. Rev. Genet. 3:252–261.

Pond, S. L. K., S. D. W. Frost, and S. V. Muse. 2005. HyPhy: hypothesis testing using phylogenies. Bioinformatics 21:676–679.

Pyron, R. A. 2010. A likelihood method for assessing molecular divergence time estimates and the placement of fossil calibrations. Syst. Biol. 59:185–194.

Rambaut, A. 2009. TRACER v1. 5. http://tree.bio.ed.ac.uk/software/tracer/.ci.nii.ac.jp.

Reeder, T. W., C. J. Cole, and H. C. Dessauer. 2002. Phylogenetic Relationships of Whiptail Lizards of the Genus *Cnemidophorus* (Squamata, Teiidae): A Test of Monophyly, Reevaluation of Karyotypic Evolution, and Review of Hybrid Origins. American Museum of Natural History.

Roelke, C. E., J. A. Maldonado, B. W. Pope, T. J. Firneno, T. J. Laduc, T. J. Hibbitts, W. A. Ryberg, N. D. Rains, and M. K. Fujita. 2018. Mitochondrial genetic variation within and between *Holbrookia lacerata lacerata* and *Holbrookia lacerata subcaudalis*, the spot-tailed earless lizards of Texas. J. Nat. Hist. 52:1017–1027. Taylor & Francis.

Rohland, N., and D. Reich. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Res. 22:939–946. genome.cshlp.org.

Schön, I., G. Rossetti, and K. Martens. 2009. Darwinulid Ostracods: Ancient Asexual Scandals or Scandalous Gossip? Pp. 217–240 *in* I. Schön, K. Martens, and P. Dijk, eds. Lost Sex: The Evolutionary Biology of Parthenogenesis. Springer Netherlands, Dordrecht.

Smith, M. D., J. O. Wertheim, S. Weaver, B. Murrell, K. Scheffler, and S. L. Kosakovsky Pond. 2015. Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. Mol. Biol. Evol. 32:1342–1353.

Stanton, D. J., L. L. Daehler, C. C. Moritz, and W. M. Brown. 1994. Sequences with the potential to form stem-and-loop structures are associated with coding-region duplications in

animal mitochondrial DNA. Genetics 137:233–241.

Tucker, D. B., S. B. Hedges, G. R. Colli, R. A. Pyron, and J. W. Sites Jr. 2017. Genomic timetree and historical biogeography of Caribbean island ameiva lizards (*Pholidoscelis*: Teiidae). Ecol. Evol. 7:7080–7090.

Vicoso, B. 2019. Molecular and evolutionary dynamics of animal sex-chromosome turnover. Nat Ecol Evol 3:1632–1641.

Vidal, N., and S. B. Hedges. 2009. The molecular evolutionary tree of lizards, snakes, and amphisbaenians. C. R. Biol. 332:129–139.

Vrijenhoek, R. C. 1998. Animal Clones and Diversity. Bioscience 48:617–628. [American Institute of Biological Sciences, Oxford University Press].

Vrijenhoek, R. C., R. M. Dawley, C. J. Cole, and J. P. Bogart. 1989. A list of the known unisexual vertebrates. Evolution and ecology of unisexual vertebrates 466:19–23. New York State Museum USA.

Welch, D. B. M., D. B. Mark Welch, and M. Meselson. 2000. Evidence for the Evolution of Bdelloid Rotifers Without Sexual Reproduction or Genetic Exchange.

Wertheim, J. O., B. Murrell, M. D. Smith, S. L. Kosakovsky Pond, and K. Scheffler. 2015. RELAX: detecting relaxed selection in a phylogenetic framework. Mol. Biol. Evol. 32:820–832.

William Birky, C., and T. G. Barraclough. 2009. Asexual Speciation. Pp. 201–216 *in* I. Schön, K. Martens, and P. Dijk, eds. Lost Sex: The Evolutionary Biology of Parthenogenesis. Springer Netherlands, Dordrecht.

Williams, G. C., and J. B. Mitton. 1973. Why reproduce sexually? J. Theor. Biol. 39:545–554.

Xia, X. 2012. Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. U. S. A.

1979;76(4):1967-1971.

Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. 24:1586–1591.

Chapter 3

Protocol for Culturing Fibroblast Cells of Whiptail Lizards Genus *Aspidoscelis*

Jose Maldonado[1,] Douja Chamseddine[1], Mark Pellegrino[1], and Matthew Fujita[1,2]

[1]The University of Texas at Arlington, Arlington, TX 76019

[2]E-mail: Jose.Maldonado@uta.edu mkfujita@uta.edu

Abstract

The development of high-throughput chromosome conformation capture to capture the organization of chromatin in three dimensions within the nucleus of cells has allowed us to investigate how genomic sequences that are far apart from each other at a linear distance can be closer to each other in the 3D space depending on how the chromatin is folded. To capture the chromatin in its native state, our goal was to successfully establish primary fibroblast cell lines cells from tail tissue of whiptail lizards (Genus: Aspidoscelis) for three different species one asexual (*Aspidoscelis tesselata*) and two sexual (*Aspidoscelis gularis* and *Aspidoscelis marmorata*). By having fibroblast cell lines, we can carry out conformation capture sequencing to have chromosome-level genome assemblies, identify structural rearrangements, and gene regulatory studies.

Introduction

Traditional whole shot-gun genome sequencing is able to produce millions of short genetic sequence reads that are reconstructed back into large contigs and have been vital for different types of studies in biological science [1]. While traditional shot-gun sequencing can provide a myriad of information, some information becomes lost from the act of extracting the DNA from cells and fragmenting the DNA into smaller pieces to be able to sequence them and generate those small reads. Specifically, the information that is lost is how the DNA is spatially arranged within the cells.

With the development of chromosomes conformation capture sequencing technique, it not only allows as to capture the genetic sequences data but as well as how the chromatin is folded within the cells and allows us to investigate how genetic loci that can be megabases apart in a linear sequence and may be in close contact in the 3-dimensional structure [2]. How the

genome is organized within the cells can provide insights into identifying large structural variation, have chromosome level genome assemblies, transcriptional regulation, gene silencing, cell cycle regulation, and disease [3-5].

To investigate how asexuality influence the evolution of the asexual nuclear genome we generated primary fibroblast cell lines we established primary cell line cultures of tail and toe clips of 3 species of Whiptails Lizards: the asexual species *Aspidoscelis tesselata* and its two sexual sister species (*Aspidoscelis gularis* and *Aspidoscelis marmorata*). Sex chromosomes in mammals and birds provide insight into the long-term deterioration of isolated non-recombining chromosomes, the evolutionary genomic changes that occur in asexual genomes over short timescales remain largely unexplored. The recent transition to parthenogenesis in the asexual lizards gives us a window into understanding the genomic consequences of asexuality over short timescales [6-7]. Previous data revealed that the asexual gecko *Heteronotia binoei* exhibits mitochondrial structural variations, accumulating large, tandem duplications that are absent in sexual gecko sister species [8-9]. By using chromosomes conformation capture sequencing technique and generating a chrome level assembly we can find large structural rearrangements in the nuclear genome, something that is not easily to do with traditional shot-gun sequencing technique that only produce short reads and tend to have a high false discovery rate for structural variation [10-11].

We used dermal tissue as is the most abundant and useful for establishing primary cell cultures, containing mitotically active fibroblast cells. Fibroblast cells are plentiful, easy to grow, remain viable for several months, and can be stored in liquid nitrogen, thawed, and re-cultured [1]. Cell lines from all species were propagated continuously up to five passages, cryopreserved, and successfully cultured after being preserved at -80 °C. The method is likely to apply to a

range of species in the genus *Aspidoscelis*. Such cell lines provide an unlimited source of material for cytogenetic, evolutionary, and genomic studies

Methods

Animals

All whiptails were collected using pitfall traps at the Indio Mountains Research Station in El Paso, Texas. Live whiptail lizards were transported to UT Arlington, housed under our approved IACUC protocol (# A15.004).

Reagents and culture medium

Washing Buffer: 1X PBS (Phosphate-Buffered Saline, 1X without calcium and magnesium, pH 7.4 ± 0.1, Corning). Collection Medium (CME): was prepared inside a laminar hood by supplementing plain DMEM (Gibco) with 1 mg/mL kanamycin (Sigma), 40 lg/mL chloramphenicol (Sigma), 60 lg/mL penicillin (Gibco), 100 lg/mL streptomycin (Gibco). This mix was sterilized by passing through a 0.2 lm filter (Millipore GP). Subsequently, amphotericin B (Sigma) and tetracycline chloride (Sigma) were added to a final concentration of 20–100 lg/mL, respectively. Culture Medium with Antibiotics (CMA): Amniomax C-100 (basal medium and supplement, Gibco), 100 U/mL penicillin G, sodium salt (Gibco), and 100 lg/mL streptomycin sulfate (Gibco). Freezing Medium (FM): 20% foetal bovine serum (Invitrogen) and 10% DMSO (Sigma) in plain DMEM.

Tissue

We sterilized the tail tips and toes of the lizards by wiping them down with gauze soaked in 70% ethanol. Approximately 5–10 mm of the tail tip and three millimeters of the fourth toe of the hind foot were collected from each specimen using sterile scissors. The tissues were immediately transferred to a beaker containing 70% ethanol and left in the solution for five

minutes. Tail and toe clips were then transferred to 5 mL of CME and incubated at room temperature for 24hr under a fume hood.

Culturing and freezing cultures

All explants for cultures were set up in an aseptic environment inside a laminar hood using sterilized equipment following Ezaz et al. (2009) published protocol. Tissue pieces were placed on a petri dish containing two drops of culture medium (CMA) and minced with sterilized scissors. Once the tissues were sufficiently minced, we added 10ml of CMA to the petri dish and transferred it to a humidified incubator at 37 °C with 5% CO2. The culture medium was changed every four days, the old medium was aspirated, and cells were washed twice with 1x PBS. The washing buffer was removed, and fresh CMA was added to the petri dish. Once cells took fibroblastic morphology, the tissue pieces were manually removed from the petri dish with sterilized forceps.

After removing tissue and cells had reached ~75% to 80% confluency, cells were passaged. The culture medium was aspirated, and cells were manually scrapped with a sterilized cells scrapper. Then 3ml of culture medium was added to the petri dish and pipetted up ten times. The culture medium was then transferred to a 15ml conical tube, and the cell suspension was centrifuged at 250g in a swivel rotor centrifuge for 5 min to pellet the cells. The supernatant was removed, only leaving the cell pellet. Cells were responded by adding 1ml of culture medium and pipping down. The cell suspension was split between two petri dishes, and 10ml of fresh culture medium was added to both petri dishes. For freezing, a cell pellet was obtained as above, resuspended in 1.5 mL FM, transferred to a cryovial, and stored at -80 °C.

Results and Discussion

Here we present a protocol adopted by Ezaz et al. (2009) to successfully culture fibroblast cells from the tail and toe tissue of whiptail lizards. We found optimum cell growth when cultured in Amniomax medium and when cultures were incubated at 37°C for all whiptail species. Cultures from all specimens reached fibroblastic morphology within a week of the culture being started. After the cultures were started, it took 1–3 weeks to get 50–70% confluency, and after each passage, 70–80% confluency was reached within 1–3 days. Our primary fibroblast cell lines were maintained and propagated and passages up to five times and finally cryopreserved successfully. Here we present a technique to establish, propagate, maintain, and cryopreserved fibroblastic cell lines from toe and tail clip explant of whiptail lizards. Now that we have established a protocol that can successfully establish primary fibroblast cell cultures, we can begin to generate the molecular dataset using chromosome conformation capture sequencing techniques to begin addressing similar question found in chapter 2. If the transition to asexuality abolishes the genetic advantages of recombination and segregation, and natural selection is weekend we expect to see an overabundance of large structural mutations accumulating within asexual individuals (Muller's Ratchet) [16] compared to sexual species. Additionally, this sequence tool has already been adopted by conservation biologist as a tool for making conservation decisions for threatened species [17-18]. Until the majority of conservation studies have relied on a few microsatellites or mitochondrial sequence data to make conservation decisions, and reduced representation sequencing has begun to close the gap for threatened species that previously had few to no molecular dataset available [19-26]. By developing a successful protocol to culture fibroblast cells for reptiles, future studies can use this resource to better investigate non-model organism to address questions in both evolutionary and conservation biology.

Acknowledgments

We thank the Pellegrino lab for allowing me to use his equipment to grow cells. We thank Douja Chamseddine for teaching us proper cells techniques. We would like to thank the students that collected whiptails lizards at UTEP Indio Mountain Research Station.

References

1. Giani, A. M., Gallo, G. R., Gianfranceschi, L., & Formenti, G. (2020). Long walk to genomics: History and current approaches to genome sequencing and assembly. *Computational and Structural Biotechnology Journal*, *18*, 9-19.
2. Belaghzal, H., Dekker, J., & Gibcus, J. H. (2017). Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods*, *123*, 56-65.
3. Gong, H., Yang, Y., Zhang, S., Li, M., & Zhang, X. (2021). Application of Hi-C and other omics data analysis in human cancer and cell differentiation research. *Computational and Structural Biotechnology Journal*, *19*, 2070-2083.
4. Kim, K., Kim, M., Kim, Y., Lee, D., & Jung, I. (2021, May). Hi-C as a molecular rangefinder to examine genomic rearrangements. In *Seminars in Cell & Developmental Biology*. Academic Press.
5. Liang, Z., Li, G., Wang, Z., Djekidel, M. N., Li, Y., Qian, M. P., ... & Chen, Y. (2017). BL-Hi-C is an efficient and sensitive approach for capturing structural and regulatory chromatin interactions. *Nature communications*, *8*(1), 1-7.
6. Abbott, J. K., Nordén, A. K., & Hansson, B. (2017). Sex chromosome evolution: historical insights and future perspectives. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1854), 20162806.
7. Bachtrog, D. (2006). A dynamic view of sex chromosome evolution. *Current opinion in genetics & development*, *16*(6), 578-585.
8. Radtkey, R. R., Donnellan, S. C., Fisher, R. N., Moritz, C., Hanley, K. A., & Case, T. J. (1995). When species collide: the origin and spread of an asexual species of

gecko. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *259*(1355), 145-152.

9. Fujita, M. K., Boore, J. L., & Moritz, C. (2007). Multiple origins and rapid evolution of duplicated mitochondrial genes in parthenogenetic geckos (Heteronotia binoei; Squamata, Gekkonidae). *Molecular Biology and Evolution*, *24*(12), 2775-2786.
10. Sethi, R., Becker, J., Graaf, J. D., Löwer, M., Suchan, M., Sahin, U., & Weber, D. (2020). Integrative analysis of structural variations using short-reads and linked-reads yields highly specific and sensitive predictions. *PLoS computational biology*, *16*(11), e1008397.
11. Ho, S. S., Urban, A. E., & Mills, R. E. (2020). Structural variation in the sequencing era. Nature Reviews Genetics, 21(3), 171-189. Chicago
12. Takashima A. Establishment of fibroblast cultures. Current Protocols in Cell Biology. 2001;Chapter 2: Unit 2.1.
13. Waters PD, Marshall Graves JA, Thompson K, Sankovic N, Ezaz T. Identification of cryptic sex chromosomes and isolation of X- and Y-borne genes. *Methods Mol Biol*ogy. 2008;422: 239–251.
14. Ezaz T, O'Meally D, Quinn AE, Sarre SD, Georges A, Marshall Graves JA. A simple non-invasive protocol to establish primary cell lines from tail and toe explants for cytogenetic studies in Australian dragon lizards (Squamata: Agamidae). *Cytotechnology*. 2008;58: 135–139.
15. Helgason, C. D., & Miller, C. L. (2005). *Basic cell culture protocols*. Totowa, NJ.: Humana Press.
16. Muller, H. J. (1964). The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *1*(1), 2-9.

17. Yang, L., Wei, F., Zhan, X., Fan, H., Zhao, P., Huang, G., ... & Hu, Y. (2022). Evolutionary conservation genomics reveals recent speciation and local adaptation in threatened takins. *Molecular Biology and Evolution*, *39*(6), msac111.
18. Armstrong, E. E., Taylor, R. W., Miller, D. E., Kaelin, C. B., Barsh, G. S., Hadly, E. A., & Petrov, D. (2020). Long live the king: chromosome-level assembly of the lion (Panthera leo) using linked-read, Hi-C, and long-read data. BMC biology, 18(1), 1-14.
19. Dubach JM, Briggs MB, White PA, Ament BA, Patterson BD. Genetic perspectives on "Lion Conservation Units" in Eastern and Southern Africa. Conserv Genet. 2013;14:741–55 Springer.
20. Dubach J, Patterson BD, Briggs MB, Venzke K, Flamand J, Stander P, et al. Molecular genetic variation across the southern and eastern geographic ranges of the African lion, *Panthera leo*. Conserv Genet. 2005;6:15–24 Springer.
21. Antunes A, Troyer JL, Roelke ME, Pecon-Slattery J, Packer C, Winterbach C, et al. The evolutionary dynamics of the lion Panthera leo revealed by host and viral population genomics. PLoS Genet. 2008;4:e1000251 journals.plos.org.
22. Singh A, Shailaja K, Gaur A, Singh L. Development and characterization of novel microsatellite markers in the Asiatic lion (Panthera leo persica). Mol Ecol Notes. 2002;2:542–3 Wiley Online Library.
23. Bruche S, Gusset M, Lippold S, Barnett R, Eulenberger K, Junhold J, et al. A genetically distinct lion (Panthera leo) population from Ethiopia. Eur J Wildl Res. 2013;59:215–25 Springer.

24. Miller SM, Harper CK, Bloomer P, Hofmeyr J, Funston PJ. Evaluation of microsatellite markers for populations studies and forensic identification of African lions (Panthera leo). J Hered. 2014;105:762–72 academic.oup.com.

25. Gaur A, Shailaja K, Singh A, Arunabala V, Satyarebala B, Singh L. Twenty polymorphic microsatellite markers in the Asiatic lion (Panthera leo persica). Conserv Genet. 2006;7:1005–8 Springer.

26. Maldonado, J. A., Firneno Jr, T. J., Roelke, C. E., Rains, N. D., Mwgiri, J., & Fujita, M. K. (2020). Transcriptome sequencing reveals signatures of positive selection in the Spot-Tailed Earless Lizard. PloS one, 15(6), e0234504. Chicago