ORIGINS AND MAINTENANCE OF DIVERSITY

IN A VERTEBRATE PARTHENOGEN

by

ALEXANDER SCOTT HALL

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy at
The University of Texas at Arlington
December, 2016

Arlington, Texas

Supervising Committee:

Matthew K. Fujita, Supervising Professor
Paul T. Chippindale
Todd A. Castoe
Eric N. Smith
Esther Betrán

ABSTRACT

Origins and Maintenance of Diversity

in a Vertebrate Parthenogen

Alexander Scott Hall, Ph.D.

The University of Texas at Arlington, 2016

Supervising Professor: Matthew K. Fujita

Clonal lineages are expected to experience a decline in fitness over time due to their inability to decouple deleterious and beneficial alleles. Individually, however, asexual populations pass twice as much of their DNA per generation as sexual lineages. This balancing act between Muller's Ratchet and the lack of the two-fold cost of sex in asexuals brings significant scientific intrigue to instances of asexual populations. Among vertebrates, squamates (and mainly lizards) are uniquely capable of true parthenogenesis whereby females reproduce without males. These lineages are always hybrid, as is the case with parthenogenetic *Aspidoscelis* lizards (Teiidae). In particular, the checkered whiptail lizard *A. tesselata* exhibits phenotypic diversity in color and scalation. This could be caused by multiple hybrid origins, genetic input from other clades (hybridization or lateral gene transfer), gene conversion, epigenetics, or mutation. *De novo* mutations are generally called upon as the main source of variation in clonal parthenoforms, but this hypothesis remains largely untested *A. tesselata*. Here, we investigate three topics pertaining to diversity in clonal *A. tesselata* and other whiptails: 1) *A. tesselata* may have arisen by multiple hybrid origin events; 2) *A. dixoni*, a close relative of *A.*

tesselata, is a post-formational lineage derived from *A. tesselata*, with a description of a new population from northern Texas; and 3) using a newly sequenced and facultatively parthenogenetic *Aspidoscelis* genome, homogeneous regions of the genome (i.e., isochores) in *Aspidoscelis* are intermediate in diversity between genomes with high and low isochore diversity. Using ddRADseq, mitochondrial genomes, and ecological niche modeling, we found *A. tesselata* originated from limited origins – estimated one or two – and that this occurred 134.3–588.2 thousand years before present (95% HPD). Additionally, diploid *A. tesselata* are paraphyletic without including *A. dixoni*. Finally, we recover a spread of genomic heterogeneity across four groups of vertebrates and found intermediate isochoric diversity within *Aspidoscelis* compared to the *Anolis carolinensis* and *Homo sapiens* genomes.

ACKNOWLEDGEMENTS

A dissertation is an adventure, and I would like to thank Matt Fujita for guiding me through it. As new faculty at UTA, he boldly supported this work at all stages and often believed in the project when I struggled to find signal amidst the noise. I also thank Paul Chippindale for recruiting me into UTA's biology program. We've become close friends and his influence on my work will extend far beyond co-authorship and as a committee member. Todd Castoe, Eric Smith, and Esther Betrán supported me as dissertation committee members and offered critical feedback when I needed it most. Eric reminded me to always think of the fossil record. Todd prodded me to consider population-level diversity, even in clonal species. Esther reminded me to forever think about recombination in new ways: gene conversion, of course!

Each project in my dissertation is co-authored with colleagues and friends, but this represents only a fraction of the number of people who have contributed their own time and energy towards my research. In the museum, Carl Franklin was a tireless supporter of the whiptail work and his efforts to improve my specimen preservation skills will (I hope) pay great dividends in years to come. Travis LaDuc, Carl Lieb, Tom Giermakowski, Toby Hibbitts, and Jay Cole provided access to valuable tissue and specimen collections. Without their support, and the museums to back them, this project may not have been possible.

In the field, Jessica Majid was my one field assistant. As we both figured out how to catch seemingly uncatchable lizards, she offered her patience and levity. Few people have put as much faith in me as she did, and for that I am thankful. Greta McGaugh and Jessica helped to identify many of the original collecting localities that we would eventually sample. Caleb Hinojos, Sarah Young, Ian Doan, and Paul Nguyen helped with projects related to my

dissertation that did not make the final cut. For many of them, catching lizards in Big Bend was their first time camping and doing field work. I can't think of any better way to get started.

In the lab, the Fujita, Castoe, Chippindale, and Smith labs all contributed to an emerging knowledge of genome-scale data at a price point graduate students could manage. In particular, Kyle O'Connell has been a great supporter when troubleshooting protocols, debugging code, and reading drafts of grants and papers. Jeff Streicher mentored me as a new graduate student and helped me run my first projects in the Smith and Fujita labs. Peter and Diana Baumann, Bill Neaves, and Araceley Newton generously provided their time during my visits to the Stowers Institute. They provided me perspective on the utility of studying parthenogenetic lizards other than for its own sake (that isn't enough?).

As part of an extension to the work presented here, I traveled to Mexico several times to visit museums and collect whiptail tissues. I thank Luna Sánchez-Reyes for hosting me during my first "real" time in Mexico (cruises don't count!). It was fun to explore Mexico City together! I also thank my collaborators Norma Manríquez Morán at UAEH and Adrián Nieto at UNAM. I look forward to our future work on *Aspidoscelis* population genetics, systematics, and ever-distracting side-projects. More specific credits and permits are listed in the acknowledgements sections of Chapters 2–4.

Several sources funded the projects in this dissertation. A Rosemary Grant from the Society for the Study of Evolution, two Phi Sigma research grants from UTA's Beta Phi chapter, and a travel grant from UTA's graduate student senate to ASH funded the majority of Chapters 2 and 3. Startup funds to Dr. Fujita supplemented field work and lab supplies for chapter 2. Danielle Rivera was funded by the National Science Foundation's (NSF) Graduate Research Fellowship Program in support of Chapter 2. An NSF Doctoral Dissertation Improvement Grant

(NSF-1501675) funded aspects of Chapters 2 and 3. Chapter 4 was supported, in part, by the Stowers Institute for Medical Research. Travel for presenting Chapters 2 and 3 were supported by several travel grants, including from the American Genetics Association, the Society for the Study of Evolution, UTA Graduate Student Senate, Phi Sigma at UTA, and the UTA Biology Department. Matching funds were provided by the UTA college of Science and the UTA Dean's Office. This dissertation originally included projects ultimately not included in the final version. These projects were funded by Phi Sigma at UTA, the Texas Herpetological Society, and the Texas Academy of Sciences. Finally, a UTA Office of Graduate Studies Dissertation Fellowship funded a summer of writing and finalizing my dissertation.

DEDICATION

I dedicate this work to my family, who have been supportive every step of the way. In particular, I thank my sister, Regina Hall. As kids, it's no secret that she was the outdoorsy one between us. Her unfiltered passion for reptiles and amphibians kindled my own interest in these cool critters. My parents, Mary Beth and David Hall, supported me in countless ways throughout my decision to go to graduate school, even when they were totally confused by what, exactly, I was doing there.

I also dedicate this work to my mentors, past and present. Having a good mentor is the difference between knowing how to hit a ball and knowing where to hit it. In particular, I thank my undergraduate mentors Ben Pierce and Romi Burks for their continued interest in my scientific career. As a new graduate student unfamiliar with teaching, Coleman Sheehy and Matt Nelson gave selflessly of their time on innumerable occasions in teaching and managing the zoology lab. I went on to teach this lab for five years, and their confidence in me extended to my confidence in teaching advanced material.

Finally, I dedicate this work to friends and family who passed away before they could see it finished: my grandfather Joe Albert Hall, my grandmother Barbara "B.G." Reilly, my great-aunt Shirley Armour, fellow graduate student Chad Watkins, and close friend Kelly Llewellyn. Even my former pet dog, Patches. Each of you touched my life in special ways, and I hope you were at peace in the end. I especially hope that my grandmother B.G., who left her psychology Ph.D. program ABD, would be proud of me as the first in our family to finish a doctoral dissertation.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION: DISCOVERY AND INTERPRETATION OF VERTEBRATE PARTHENOGENESIS AND CLONAL DIVERSITY

A Brief History of Vertebrate Parthenogenesis

Since the initial discovery of vertebrate clonality (Hubbs and Hubbs 1932), many additional instances of unisexual vertebrates have been described and the topic has received sustained attention within the scientific community (Hubbs 1955; Maslin 1968; Cole 1975; Dawley and Bogart 1989; Avise 2008; Schön et al. 2009; Booth and Schuett 2015). Lizards provide the most-studied taxon of unisexual vertebrate in terms of numbers of described unisexual biotypes (Vrijenhoek et al. 1989) and overall volume of information related to their ecology and evolution. Notably, several lizards exhibit true parthenogenesis, making them unique among vertebrates (Uzzell 1970). In true (obligate) parthenogenesis, the female passes on its entire genome to its offspring without male involvement. No sperm are required to stimulate development. This process stands in contrast to a diversity of alternative mating schemes that can involve clonality: gynogenesis (sperm used only to stimulate egg development), kleptogenesis (sperm stealing with occasional genome incorporation), androgenesis (obliteration of female germ line), hybridogenesis (genome swapping with retention) and facultative parthenogenesis (ability for sex and clonality; Dawley 1989). Two lizard families in particular, the Teiidae and Lacertidae, have produced most examples of vertebrate parthenogenesis and possess deep ties to the study of vertebrate clonality. I will focus on the Teiidae and in particular the genus *Aspidoscelis* (formerly *Cnemidophorus* [Reeder et al. 2002]).

Throughout the 1950s and 1960s, initial taxonomic descriptions of species and subspecies within the *Aspidoscelis* placed heavy emphases on geographic range and phenotypic distinctiveness (Maslin 1968). Occasionally, taxonomists described unisexual taxa based on ecology, range, and morphology before realizing their unisexual nature, as in the case of *A. neomexicanus* (Lowe and Zweifel 1952) and *A. exsanguis* (Lowe 1956). The first evidence of

parthenogenesis in the *Aspidoscelis* (i.e., *Cnemidophorus*) came from the independent discoveries by Sherman Minton (1958), Donald Tinkle (1959), and Paul Maslin (1962) that *A. tesselata* lacked males. Minton (1958) simply states "All specimens are females. I have never seen a male of [*Aspidoscelis tesselata*] from any part of the range." Tinkle (1959) acknowledged Minton's discovery of missing males in *A. tesselata* and reaffirmed that among a series of 65 specimens, he could find no male lizards. Tinkle hesitated to comment further on a mechanism for this observation; hypothesizing that different field collecting conditions might yet reveal cryptic male individuals. The first formal association between parthenogenesis and *Aspidoscelis* came from a *Science* article by Maslin (1962), who discovered a lack of females from a survey of museum specimens distributed among 74 then-recognized species and forms (*sensu lato*). Maslin's inquiry began with the auspicious discovery of an all-female population of '*Cnemidophorus deppei cozumelus*' (a member of the *A. cozumela* complex). His survey of museum specimens revealed five additional forms to be completely or nearly all female: *A. costatus exsanguis* (now called *A. exsanguis*), a western *A. inornata ssp.* (most likely not *A. inornata*, as this is a bisexual species), *A. perplexus* (generally meaning *A. neomexicanus*, but see Wright and Lowe 1967b), *A. velox*, and *A. tesselata*. Maslin carefully notes that among 223 *A. tesselata* specimens examined, he found one male (though this may have been a misidentification). Further discoveries related to parthenogenesis in *Aspidoscelis* arose from technological advances resulting in four general techniques: karyology (e.g., Lowe and Wright 1966a), tissue grafts (Maslin 1967; Cuellar 1977), allozyme analysis (e.g., Neaves 1969; Dessauer and Cole 1984), and mitochondrial DNA restriction enzyme analysis (e.g., Brown and Wright 1979).

Emerging Technologies Revealed Clonal Diversity

As first hypothesized by Maslin (1962), genetics have been suspected to play a role in squamate parthenogenesis since their initial discovery. Lowe and Wright (1966a) published the first karyotypic data for eight parthenogenetic and nine sexual *Aspidoscelis* species. This brief paper first identified triploid asexual species and the hybrid origin of diploid asexual *Aspidoscelis* (Lowe and Wright 1966a). These authors also first hypothesized the origins of triploid asexual *Aspidoscelis* as resulting from hybridization between female, diploid, parthenogenetic intermediate and male from a sexual species (Lowe and Wright 1966a, b). A further inquiry into the karyotypes of the six pattern classes of *A. tesselata* (Fig. 1.1) identified by Zweifel (1965) revealed pattern classes C through F to be allodiploid (2n = 46) but pattern classes A and B to be allotriploid (3n = 69; Wright and Lowe 1967a). Interestingly, some members of pattern classes C and D possessed 'modified allodiploid' conditions where duplications or centric fissions led to the addition of an extra chromosome (thus, 2n = 46 + 1). This modified condition apparently carried over to all allotriploid lizards examined, and thus all observed pattern class A and B *A. tesselata* had 3n = 69 + 1 chromosomes (Wright and Lowe 1967a). Careful study of the karyotypes in this study revealed the diploid *A. tesselata* pattern classes to be hybrids between *A. tigris (marmorata)* and *A. (gularis) septemvittata* (Fig. 1.2). The triploid *A. tesselata* pattern classes (A and B; Fig. 1.1) were hybrids between diploid *A. tesselata* and *A. sexlineata (viridis)*. Walker et al. (1997) later described the triploid variants as a separate species, *A. neotesselata*.

In testing a prediction of parthenogenesis in *Aspidoscelis*, Paul Maslin (1967) adapted a skin graft technique (May 1923; Kallman 1962a, b; Whimster 1962, 1965) to purported parthenogenetic *Aspidoscelis* lizards, namely *A. tesselata*. In this technique, animals are

anaesthetized and ~1 cm^2 patches of skin are removed and transplanted to either the same animal (i.e., control) or another animal. The paradigm is that organisms will not reject tissues from genetically similar conspecifics, but poor histocompatibility between genetically distinct 'species' will cause tissue grafts to be rejected. Maslin (1967) found that different animals within populations of *A. tesselata* accepted grafts from donor lizards, leading him to suggest that these populations are histocompatible and thus genetically homogeneous. These findings agreed with Zweifel's (1965) observations that color patterning and scale counts within populations are remarkably constant, though variable across the species. Similar skin graft studies on *Aspidoscelis* lineages confirmed parthenogenesis as first identified in 1967 by Wright and Lowe (Cuellar and McKinney 1976; Cuellar 1977).

At approximately the same time Maslin (1967) published his first skin graft studies on parthenogenetic *Aspidoscelis*, William Neaves and Park Gerald (1968) reported one of the first allozyme datasets in *Science*. In their study, parthenogenetic *A. tesselata* and *A. neomexicanus* displayed heterozygous phenotypes at the lactate dehydrogenase (LDH) *b*-locus. Both parthenogenetic phenotypes were complex intermediates between their hypothesized sexual progenitors, indicating that each species arose from a hybridization event and that each genome is transcribed and translated (Neaves and Gerald 1968). A related study (also published in *Science*) demonstrated that triploid *A. tesselata* (i.e., *A. neotesselata*) exhibited a heterozygous phenotype at the LDH *b*-locus, with two copies of *b* and one copy of *b'* estimated (Neaves and Gerald 1969). Neaves (1969) quickly published additional allozyme data (adenosine deaminase [ADA] in this study) for five parthenogenetic and five sexual *Aspidoscelis*. Again, heterozygous phenotypes for the ADA locus are observed in parthenogenetic *Aspidoscelis*, providing further evidence of their hybrid origins. Before DNA sequencing became routine, expanded allozyme

studies revealed exceptionally high heterozygosity (~0.4) in parthenogenetic *Aspidoscelis* (Dessauer and Cole 1984, 1986) and verified complex origins of triploid *Aspidoscelis exsanguis* (Good and Wright 1984; Dessauer and Cole 1986).

Clonal Vertebrates Complicate Species Concepts

As previously mentioned, early taxonomic descriptions of the *Aspidoscelis* placed heavy emphases on geographic range and phenotypic distinctiveness, rather than shared evolutionary history (Maslin 1968). Amid the technological advancements previously listed, the number of named parthenogenetic subspecies based on specific ancestral hybridizations sharply increased; thus concerning some taxonomists. Zweifel (1965) was among the first to display active avoidance of name inflation for unisexual vertebrates when he referred to populations of *A. tesselata* differing in color pattern and scale counts as classes with alphabetical designations A – F (Fig. 1.1; recall pattern classes A and B are triploid *A. neotesselata*; Walker et al. 1997). Walker (1986) reflected Zweifel's (1965) caution by urging reducing the number of named parthenogenetic *Aspidoscelis* species. He attempted to assign many named parthenogenetic subspecies to alphabetically designated variants under the name of the species complex (see Table 4 of Walker 1986). In my experience and through surveying the literature, few appear to have adopted Walker's (1986) reductive nomenclature.

Before Sanger sequencing (Sanger et al. 1977) became widely feasible, several studies utilized restriction endonuclease of mitochondrial DNA to infer the progenitors and relative ages of parthenogenetic *Aspidoscelis*. In another study appearing in *Science*, Wesley Brown and John Wright (1979) utilized electrophoretic gels and electron microscopy to investigate restriction sites in the mitochondrial genomes, and thus the maternal ancestry, of parthenogenetic *A. tesselatus* and *A. neomexicanus* and their sexual progenitors. *Aspidoscelis tesselata* and *A.*

neomexicanus share the same *Eco* RI and *Hind* III mtDNA recognition sites as their maternal progenitor, *A. tigris marmorata* (Brown and Wright 1979). Similar techniques were utilized to determine the maternal origin of *A. laredoensis* (Wright et al. 1983) and several *A. sexlineata* group species (Densmore et al. 1989). An early parsimony phylogeny of sexual *Aspidoscelis* by Moritz et al. (1992) utilized restriction sites as characters. This study demonstrated the behavior within the *Aspidoscelis* genus for parthenogenetic lineages to arise from hybrids between distantly related species within the genus (Moritz et al. 1992). Subsequent mtDNA sequence analysis by Reeder et al. (2002) expanded upon work by Moritz et al. (1992) and provides the most complete and most current treatment of *Aspidoscelis* phylogenetics.

The most recent and thorough investigation of the phylogeny of *Aspidoscelis* uncovered that the genus previously considered "*Cnemidophorus*" was paraphyletic (Reeder et al. 2002). Thus, most North American lizards formerly placed in the genus *Cnemidophorus* are actually a part of a revived genus *Aspidoscelis* (Fitzinger 1843; Reeder et al. 2002). This study uncovered phylogenetic relationships within "*Cnemidophorus*" lizards using a combination of mtDNA sequence analysis, morphology, and allozymes. The traditional *Cnemidophorus* genus was found to be paraphyletic, and thus the name *Aspidoscelis* (Fitzinger 1843) was resurrected to reflect a monophyletic group. Most lizards in the genus *Aspidoscelis* occur only in North America and all share the following characters: an absent basal tongue sheath, posterior portion of tongue clearly forked, smooth ventral scutes, eight rows of ventral scutes at midbody, absence of anal spurs in males, mesoptychial scales abruptly enlarged over scales of gular fold (more anterior mesoptychials becoming smaller), three parietal scales, and three or four supraocular scales on each side (Reeder et al. 2002). As Lowe and Wright (1966a) first investigated, karyotypic evolution is quite pronounced within the *Aspidoscelis* clade. In summary, most diploid

Aspidoscelis possess 2n = 46 chromosomes though the clade including *A. deppii* possess 2n = 52 chromosomes. Note that *A. dixoni* falls within the *A. tesselata* complex (Scudday 1973); though Cordes and Walker (2006) challenged the validity of *A. dixoni* based on complete skin graft histocompatibility between *A. dixoni* A, B, and C and *A. tesselata* E.

New Information and This Dissertation

The remainder of my dissertation consists of three data-driven chapters and a brief conclusion. Each data chapter is being considered for publication, and thus the literature cited and figures follow each chapter rather than in one large section at the end of the dissertation. This work combines field, lab, and bioinformatic investigation into diversity within parthenogenetic vertebrates.

My goal was to describe variation in *Aspidoscelis tesselata*. Previous investigators have largely assumed that post-formational divergence caused *A. tesselata*'s pattern classes (Fig. 1.1). In chapter 2, I address this question using genome-scale data for the first time. Danielle Rivera contributed ecological niche modeling and Jose Maldonado recovered the mitochondrial genomes used. Matthew Fujita provided the first sequence data for this project from samples provided by Peter Baumann and Travis LaDuc. Extending on these findings, in chapter 3 I present preliminary findings on the genetic nature and origin of *A. dixoni*, also referred to as *A. tesselata* pattern class F. The taxonomy of this array or lineage is contentious, and this chapter aims to provide an update to the rationale to sustain or sink the species. James Walker and James Cordes provided a new and crucial specimen of *A. dixoni* and considerably improved the manuscript through our many conversations. This chapter of my dissertation was, in large part, inspired by their own dissertations (Walker 1966; Cordes 1991). While pursuing these projects, I was invited to collaborate on the publication of the first whiptail genome for *A. marmorata*.

Interestingly, the animal that was sequenced was apparently a facultatively parthenogenetic individual. The final publication will comment further on the implications of this chance discovery. This paper will not be published before my dissertation; nevertheless, in Chapter 4 I expand upon my involvement. Primarily, I focus on my investigation of genomic heterogeneity and isochores in the first Teiid genome with a comparative analysis across all sequenced vertebrate genomes. Matt Fujita collaborated on all work presented here in conceiving these projects, collecting and interpreting data, and suggesting the final format of the dissertation.

Literature Cited

Avise, J. 2008. *Clonality: The Genetics, Ecology, and Evolution of Sexual Abstinence in Vertebrate Animals*. Oxford University Press, USA.

Booth, W., and G. W. Schuett. 2015. The emerging phylogenetic pattern of parthenogenesis in snakes. *Biological Journal of the Linnean Society*, 118:172–186.

Brown, W. M., and J. W. Wright. 1979. Mitochondrial DNA analysis and the origin and relative age of parthenogenetic lizards (genus *Cnemidophorus*). *Science*, 203:1247–1249.

Cole, C. J. 1975. Evolution of parthenogenetic species of reptiles. In *Intersexuality in the Animal Kingdom*, R. Reinboth (ed.), pp 340–355. Heidelberg: Springer.

Cordes, J. E. 1991. Biology of two sympatric clonal complexes of parthenogenetic whiptail lizards [TES-G (2n) and TES-H (2n)] in the *Cnemidophorus tesselatus* complex and identification of hybrid TES-G (2n) x *Cnemidophorus gularis septemvittatus* in the Chinati Mountains in trans-Pecos Texas. Thesis (Ph.D.). University of Arkansas.

Cordes, J. E., and J. M. Walker. 2006. Evolutionary and systematic implications of skin histocompatibility among parthenogenetic Teiid lizards: Three color pattern classes of *Aspidoscelis dixoni* and one of *Aspidoscelis tesselata*. *Copeia*, 2006:14–26.

Cuellar, O. 1977. Genetic homogeneity and speciation in the parthenogenetic lizards *Cnemidophorus velox* an *C. neomexicanus*: Evidence from intraspecific histocompatibility. *Evolution*, 31:24–31.

Cuellar, O., and C. O. McKinney. 1976. Natural hybridization between parthenogenetic and bisexual lizards: Detection of uniparental source by skin grafting. *Journal of Experimental Zoology*. 196:341–350.

Dawley, R. M. 1989. An introduction to unisexual vertebrates. In *Evolution and Ecology of Unisexual Vertebrates*, R. Dawley and J. P. Bogart, pp. 1–18. New York State Museum Bulletin.

Dawley, R., and J. P. Bogart (eds). 1989. *Evolution and Ecology of Unisexual Vertebrates*. New York State Museum Bulletin.

Densmore, L. D. C. Moritz, J. W. Wright, and W. M. Brown. 1989. Mitochondrial DNA analyses and the origin and relative age of parthenogenetic lizards (genus *Cnemidophorus*). IV. Nine *sexlineatus* group parthenoforms. *Evolution*, 43:969–983.

Dessauer, H. C., and C. J. Cole. 1984. Influence of gene dosage on electrophoretic phenotypes of proteins from lizards of the genus *Cnemidophorus*. *Comparative Biochemistry and Physiology*, *Part B: Biochemistry & Molecular Biology*, 77:181–189.

Dessauer, H. C., and C. J. Cole. 1986. Clonal inheritance in parthenogenetic whiptail lizards: Biochemical evidence. *Journal of Heredity*, 77:8–12.

Fitzinger, L. 1843. *Systema Reptilium*. Vienna: Vindobonae, 106 pp.

Good, D. A., and J. W. Wright. 1984. Allozymes and the hybrid origin of the parthenogenetic lizard *Cnemidophorus exsanguis*. *Experientia*, 40:1012–1014.

Hubbs, C. L. 1955. Hybridization between fish species in nature. *Systematic Zoology*, 4:1–20.

Hubbs, C. L., and L. C. Hubbs. 1932. Apparent parthenogenesis in nature, in a form of fish of hybrid origin. *Science*, 76:628–630.

Kallman, K. D. 1962a. Gynogenesis in the teleost, *Mollienesia formosa* (Girard), with a discussion of the detection of parthenogenesis in vertebrates by tissue transplantation. *Journal of Genetics*, 58:7–21.

Kallman, K. D. 1962b. Population genetics of the gynogenetic teleost, *Mollienesia formosa* (Girard). *Evolution*, 16:497–504.

Lowe, C. H., Jr. 1956. A new species and a new subspecies of whiptailed lizards (genus *Cnemidophorus*) of the inland Southwest. *Bulletin of the Chicago Academy of Sciences*, 10:137–150.

Lowe, C. H., Jr., and R. G. Zweifel. 1952. A new species of whiptailed lizard (genus *Cnemidophorus*) from New Mexico. *Bulletin of the Chicago Academy of Sciences*, 9:229–247.

Lowe, C. H., and J. W. Wright. 1966a. Chromosomes and karyotypes of cnemidophorine Teiid lizards. *Mammalian Chromosomes Newsletter*, 22:199–200.

Lowe, C. H., and J. W. Wright. 1966b. Evolution of parthenogenetic species of *Cnemidophorus* (whiptail lizards) in Western North America. *Journal of the Arizona Academy of Science*, 4:81–87.

Maslin, T. P. 1962. All female species of the lizard genus *Cnemidophorus*, Teiidae. *Science*, 135:212–213.

Maslin, T. P. 1967. Skin grafting in the bisexual Teiid lizard *Cnemidophorus sexlineatus* and in the unisexual *C. tesselatus*. *Journal of Experimental Zoology*, 166:137–150.

Maslin, T. P. 1968. Taxonomic problems in parthenogenetic vertebrates. *Systematic Zoology*, 17:219–231.

May, R. M. 1923. Skin grafts in the lizard *Anolis carolinensis*. *The British Journal of Experimental Biology*, 1:539–555.

Minton, S. A., Jr. 1958. Observations on amphibians and reptiles of the Big Bend Region of Texas. *The Southwestern Naturalist*, 3:28–54.

Moritz, C., J. W. Wright, and W. M. Brown. 1992. Mitochondrial DNA analyses and the origin and relative age of parthenogenetic *Cnemidophorus*: phylogenetic constraints on hybrid origins. *Evolution* 46:184–192.

Neaves, W. 1969. Adenosine deaminase phenotypes among sexual and parthenogenetic lizards in the genus *Cnemidophorus* (Teiidae). *Journal of Experimental Zoology*, 171:175–184.

Neaves, W. B., and P. S. Gerald. 1968. Lactate dehydrogenase isozymes in parthenogenetic Teiid lizards (*Cnemidophorus*). *Science*, 160:1004–1005.

Neaves, W. B., and P. S. Gerald. 1969. Gene dosage at the lactate dehydrogenase *b* locus in triploid and diploid Teiid lizards. *Science*, 164:557–559.

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 Novitates*, 3365:1–64.

Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences*, 74:5463–5467.

Schön, I., K. Martens, and P. van Dijk (eds). 2009. *Lost Sex: The Evolutionary Biology of Parthenogenesis*. Springer.

Scudday, J. F. 1973. A new species of lizard of the *Cnemidophorus tesselatus* group from Texas. *Journal of Herpetology*, 7:363–371.

Tinkle, D. W. 1959. Observations on the lizards *Cnemidophorus tigris*, *Cnemidophorus tesselatus* and *Crotaphytus wislizeni*. *The Southwestern Naturalist*, 4:195–200.

Uzzell, T. 1970. Meiotic mechanisms of naturally occurring unisexual vertebrates. *The American Naturalist*, 104:433–445.

Vrijenhoek, R. C., R. M. Dawley, C. J. Cole, and J. P. Bogart. 1989. A list of the known unisexual vertebrates. In *Evolution and Ecology of Unisexual Vertebrates*, R. Dawley and J. P. Bogart, pp. 19–23. New York State Museum Bulletin.

Walker, J. M. 1966. Morphological variation in the Teiid lizard *Cnemidophorus gularis*. Thesis (Ph.D.). University of Colorado.

Walker, J. M. 1986. The taxonomy of parthenogenetic species of hybrid origin: Clone hybrid populations of *Cnemidophorus* (Sauria: Teiidae). *Systematic Zoology*, 35:427–440.

Walker, J. M., J. E. Cordes, and H. L. Taylor. 1997. Parthenogenetic *Cnemidophorus tesselatus* complex (Sauria: Teiidae): A neotype for diploid *C. tesselatus* (Say, 1823), Redescription of the taxon, and description of a new triploid species. *Herpetologica*, 53:233–259.

Whimster, I. W. 1962. The mosaic nature of pigmentary change in diseased skin. *Annali Italiani di Dermatologia Allergologica Clinica e Sperimentale*, 16:357–384.

Whimster, I. W. 1965. An experimental approach to the problem of spottiness. *British Journal of Dermatology*. 77:397–420.

Wright, J. W., C. Spolsky, and W. M. Brown. 1983. The origins of the parthenogenetic lizard *Cnemidophorus laredoensis* inferred from mitochondrial DNA analysis. *Herpetologica*, 39:410–416.

Wright, J. W., and C. H. Lowe. 1967a. Evolution of the alloploid parthenospecies *Cnemidophorus tesselatus* (Say). *Mammalian Chromosome Newsletter* 8:95–96.

Wright, J. W., and C. H. Lowe. 1967. Hybridization in nature between parthenogenetic and bisexual species of whiptail lizards (genus *Cnemidophorus*). *American Museum Novitates*, 2286:1–36.

Zweifel, R. G. 1965. Variation in and distribution of the unisexual lizard *Cnemidophorus tesselatus*. *American Museum Novitates*, 2235:1–35.

Figure Legends

Figure 1.1. Clonal diversity seen in *Aspidoscelis tesselata* pattern classes A, B, C, D, E, and F as originally described by Zweifel (1965). Pattern classes C, D, E, and F are diploid and hybrids between *A. marmorata* and *A. gularis septemvittata*. Pattern class F was renamed *A. dixoni* (Scudday 1973). Pattern classes A and B are triploid and the result of a second cross between *A. tesselata* and *A. sexlineata*. Photo taken and provided by William Neaves. Figure and legend modified from Neaves (2014). Reprinted with permission from Taylor and Francis Group LLC Books.

Figure 1.2. Bisexual parental ancestral species of diploid *Cnemidophorus tesselatus* as proposed (Wright and Lowe 1967). Left to Right: *C. septemvittatus*, parental ancestral species, captured 8/8/1968 in Brewster County, Texas; *C. tesselatus*, color pattern class E, captured 8/22/1968 in De Baca County, New Mexico; *C. tigris*, parental ancestral species, captured 8/4/1968 in Luna County, New Mexico. Figure and legend modified from Neaves (2014). Reprinted with permission from Taylor and Francis Group LLC Books.

Figure 1.1

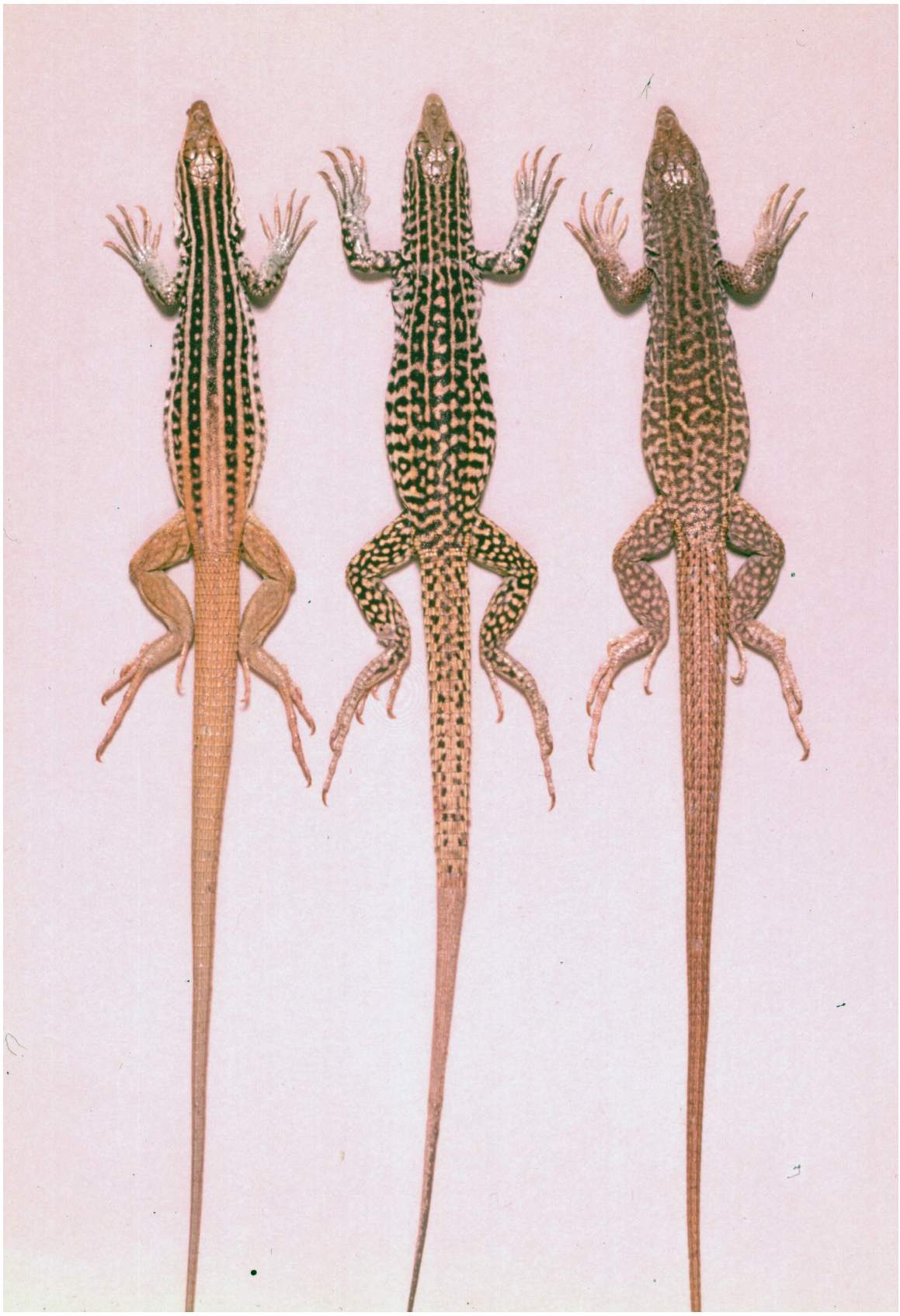

Figure 1.2

CHAPTER 2

MULTIPLE ORIGINS OF A UNISEXUAL HYBRID LIZARD, ASPIDOSCELIS TESSELATA

MULTIPLE ORIGINS OF A UNISEXUAL HYBRID LIZARD, *ASPIDOSCELIS TESSELATA*[1]

Alexander S. Hall[1], Danielle Rivera[1],

Jose A. Maldonado[1], and Matthew K. Fujita[1,2]

[1]This article is in preparation for submission to the journal *Evolution*.

Abstract

Asexually reproducing organisms are assumed to evolve with mutation and recombination driving genetic – and therefore phenotypic – diversity. Lacking the ability to rapidly generate new allelic combinations, asexually reproducing organisms rely on *de novo* mutations to produce new haplotypes. In the case of parthenogenetic vertebrates, however, obligate parthenogenetic species are typically hybrids. Each hybrid genome contains two or more distinct genomes with different evolutionary histories. In this way, multiple origins of a hybrid vertebrate parthenoform can increase potential for novel alleles and haplotypes by intergenomic crossing over, gene conversion, and mutation. In the first such study at a genomic level, we tested for multiple origins in a parthenogenetic lizard complex, *Aspidoscelis tesselata* (Squamata: Teiidae). We used genome-wide SNPs collected by ddRADseq (N = 192) and entire mitochondrial genomes (N = 37) to reconstruct the phylogeographic histories of the two species that hybridized to form *A. tesselata*. We found that *A. tesselata*, (and by extension the closely related species *A. dixoni*) may have arisen more than once. We estimated *A. tesselata* came to be 134.3–588.2 thousand years ago. Finally, we provide evidence from niche modeling and genomic SNPs that *A. tesselata* likely originated in northern Coahuila, Mexico and near Peloncillo Mountains in southwestern New Mexico.

Introduction

Mechanisms of generating genetic and phenotypic diversity in unisexual vertebrates are poorly understood, limiting our ability to accurately describe evolutionary processes and infer phylogenies (Avise 2008; Fujita and Moritz 2010; Grismer et al. 2014). Unisexual reproduction in vertebrates occurs nonrandomly and true all-female parthenogenesis only occurs in squamates (Kearney et al. 2009; Neaves and Baumann 2011). Existing parthenogenetic lizard species exhibit an array of color patterns, different skin histocompatibilities, and several mitochondrial haplotypes (Morán 2007). Such diversity could arise from a mixture of processes: mutation, genetic input from other clades, genome dynamics such as gene conversion, and multiple origins (Fujita and Moritz 2010). In this study, we explore for the first time at a genomic scale how multiple origins of a parthenogenetic vertebrate may contribute to its observed phenotypic diversity.

Except in the case of vertebrate facultative (i.e., not obligate) parthenogenesis (e.g., Chapman et al. 2007; Booth et al. 2011a; Booth et al. 2011b), unisexual vertebrate lineages are hybrids between sexual (bisexual) species (Kearney et al. 2009; Lutes et al. 2011). Hybrid F1 offspring reproduce independently of their parent lineages (Lutes et al. 2011), except in the cases of further hybridization events (Neaves 1971; Walker et al. 1997; Lutes et al. 2011; Moritz and Bi 2011; Walker et al. 2012; Cole et al. 2014). Thus, the possibility of multiple F1 origins in unisexual vertebrates may explain some diversity seen at the phenotypic level in moderately diverged parthenogenetic lineages.

Among all parthenogenetic squamate lineages, perhaps no system has received more scientific attention than whiptail lizards in the Teiid genus *Aspidoscelis* (formerly *Cnemidophorus*; Reeder et al. 2002). Between a third and half of all *Aspidoscelis* species arose

from hybridization and reproduce via true parthenogenesis (Reeder et al. 2002; Cole et al. 2014). Tinkle (1959) and Maslin (1962; 1966) described true vertebrate parthenogenesis for the first time from within *Aspidoscelis* using the common checkered whiptail, *A. tesselata*. Soon after, Zweifel (1965) thoroughly described *A. tesselata* phenotypic diversity and assigned populations to six pattern classes A–F, establishing a framework against which hypotheses regarding unisexual speciation could be tested. For these reasons, *A. tesselata* represents an exceptional system for investigating mechanisms of diversity in unisexual vertebrates.

As previously mentioned, several processes could generate this observed phenotypic diversity. Further hybridization between diploid *A. tesselata* and *A. sexlineata* explains two of Zweifel's (1965) pattern classes A and B, now known to be triploid *A. neotesselata* (Walker et al. 1997; Walker et al. 2012). The remaining four pattern classes C–F arose through hybridization between female *A. marmorata* and male *A. gularis* (Fig. 2.1). One pattern class, F, was described as *A. dixoni* by Scudday (1973), though the species is broadly recognized as a variant of *A. tesselata* (Cordes 1991; Walker et al. 1994; Cordes and Walker 2006). Hereafter, *A. tesselata* refers to *A. dixoni* and *A. tesselata* C–E unless otherwise stated. It is unknown whether *A. tesselata* originated from a single hybridization event (Fig. 2.1A) or multiple hybridization events (e.g., Fig. 2.1B). In the absence of further hybridization or horizontal gene transfer, the unisexual lineage must work from its available genome and de novo mutations to produce genetic diversity (Muller 1932; Parker and Selander 1976). By pulling from different populations within 'parent' species, the unisexual arrays can be thought to, collectively, accept genetic diversity beyond that expected from a single origin.

Generally, previous investigations assert that all *A. tesselata* arose from one F_1 zygote (Maslin 1967; Cordes and Walker 2003, 2006; Taylor et al. 2003). This has been restated in

several papers (Taylor et al. 2005; Taylor et al. 2006; Paulissen et al. 2006; Taylor et al. 2012). Taylor et al. (2003) state quite clearly "[…] that all color pattern classes, morphological subgroups, and genotypic clones of *A. tesselata* can be traced back to a single ancestral F_1 hybrid zygote." Indeed, a single origin and postformational divergence best describes the current model of *A. tesselata* clonal diversity (Fig. 2.1A). If *A. tesselata*, instead, arose from more than one hybridization event, a substantial portion of recent study on unisexual vertebrates could be considered phylogenetically naïve.

Since unisexual *Aspidoscelis* (and most other unisexual vertebrates) are hybrids, their maternal ancestry can be tracked using phylogenetic analysis of maternally inherited mitochondrial sequence data (Grismer et al. 2014). Extensive studies by Densmore and colleagues (Brown and Wright 1979; Wright et al. 1983; Densmore et al. 1989a, 1989b; Moritz et al. 1989) followed this reasoning using variation between mitochondrial restriction sites to determine maternal ancestry of many then-described North American *Aspidoscelis* lineages. This series of studies confirmed *A. marmorata* as the maternal ancestor of *A. tesselata*. Thus far, paternal ancestry has been more difficult to infer. Karyotypic analysis of hybrid unisexuals generally suffices to determine paternal ancestry to species. Within unisexual lineages, creative approaches – such as skin histocompatibility measured from reciprocal skin transplants (Maslin 1967) – serve as a proxy for examining relatedness (Cuellar 1977; Cordes and Walker 2006). Beyond skin histocompatibility, allozymes have generally been used as evidence of a species' contribution to a hybrid (e.g., Neaves and Gerald 1968; Neaves 1969). This approach generally precludes the resolution necessary for population-level description of the origins of a parthenogenetic lineage, however.

Recent developments in the use of subsampled genomic sequence data allow high resolution insight into the evolutionary histories of recently diverged species complexes (Baird et al. 2008; Streicher et al. 2014; Schield et al. 2015). Combined with projected ecological niche models derived from over one thousand museum records and sequenced mitochondrial genomes, we aim to utilize genomic DNA to improve the estimation of hybrid origin inference. Overall, this study aims to answer if multiple hybrid origins of a parthenogenetic hybrid lineage are detectable using genomic single nucleotide polymorphisms (SNPs). We hypothesize that what are referred to as *A. tesselata* pattern classes C–E (*sensu* Zweifel 1965) and *A. dixoni* (i.e., *A. tesselata* F; Scudday 1973) arose from multiple hybridization events (Fig. 2.1B). In this study, we compare mitochondrial genome data and nuclear genomic SNPs in gene tree and model-free structural phylogenetic methods to uncover the checkered population genetics of the species generating the checkered whiptail, *A. tesselata*.

Materials and Methods

Data Collection

Sampling.- The species under investigation are found throughout North America, so to cover the broadest area for each species we collected animals in Texas and New Mexico and received tissue loans from museums and individuals (Fig. 2). We collected 93 whiptails from March through August of 2013 and al l vouchers are deposited at the Amphibian and Reptile Diversity Research Center at UT Arlington. Of 272 available samples (179 loans + 93 our collection), 37 were used for mtDNA sequencing and 192 were used for ddRADseq (Appendix A).

Whiptail DNA was extracted from blood, liver, leg muscle, or tail. Tissues were extracted in 100–300 µL of cell lysis buffer (100 mM NaCl, 100 mM Tris-Cl pH 8.0, 25 mM EDTA pH 8.0, and 0.5% SDS), 5–15 µL proteinase K (20 mg/mL), and 5–15 µL RNAse A (4 mg/mL) and

gently agitated at 55°C for 1–8 h. Extractions were cleaned with a 15 min incubation with a 1.8X volume of freshly prepared Serapure beads (a substitute for AMPure XP, but using Sera-mag SpeedBeads; [Fisher Scientific, Pittsburgh, PA, USA; Rohland and Reich 2012]), twice rinsed with 70% ethanol while on magnets, allowed to air-dry for 5 min, and diluted in 100 µL 10 mM Tris-HCl, pH 8.0.

For building mitochondrial genome alignments, we downloaded the most closely related squamate mitochondrial genome from GenBank, a Lacertid wall lizard *Podarcis muralis* from Austria (GenBank Version: NC_011607.1; Podnar et al. 2009).

Mitochondrial Genomes.- Without a reference *Aspidoscelis* mitochondrial genome, we sequenced mitochondrial genomes by first digesting linear DNA and amplifying remaining circular DNA. We followed the protocol from an upcoming publication by Fujita et al. to prepare Illumina libraries of mitochondrial genomes from whole DNA. For constructing Illumina libraries, we used 300 ng of DNA per sample as measured by a Qubit® 2.0 Fluorometer (Invitrogen) and corresponding Qubit® dsDNA BR Assay Kit (Invitrogen). To sequence multiple individuals at once, we multiplexed samples using adapter barcodes and Illumina PCR primers. We verified successful library preparations (grouped by Illumina primers) using an Agilent BioAnalyzer 2100 and Agilent DNA 7500 Kit and measured DNA quantity via Qubit. Equal amounts of each library were combined, cleaned with 1.0X Serapure, and again checked using the BioAnalyzer and Qubit. Multiplexed samples were sequenced on UT Arlington's Illumina MiSeq producing 300 bp paired-end reads.

Genomic Subsampling.- To collect nuclear loci, we utilized double digest restriction-site associated DNA sequencing (ddRADseq) largely following Peterson et al. (2012). Briefly, we used the two restriction enzymes *MspI* and *SbfI* (New England Biology) to fragment 100–1000

ng of DNA for each individual. We ligated custom Illumina adaptor sequences to the DNA fragments. We then pooled 8–12 individuals into separate pools and size-selected fragments between 435–535 bp using a Pippin Prep™ (Sage Science). We followed size-selection with 12 cycles of library amplification which also adds Illumina indices to each library. The combination of adapters and Illumina indices allowed identification of each individual after 100-bp paired-end sequencing on two lanes of an Illumina HiSeq 2500 at The University of Texas Southwestern Medical Center.

Ecological Niche Modeling.- To construct ecological niche models (ENMs), we compiled a dataset of locality records for *Aspidoscelis gularis* (*sensu lato*), *A. marmorata*, and *A. tesselata*. These data were primarily obtained from VertNet2 by filtering results from the search terms: "Cnemidophorus," "Aspidoscelis," "tesselata OR tesselatus OR tessellatus," "gularis," and "marmorata OR marmoratus." We also downloaded research-grade observations with locality data from iNaturalist using the search term "Aspidoscelis" that we then filtered to species of interest. Finally, we included records from our own capture efforts and other loaned specimens with locality data not already covered by these datasets. In total, this dataset included 7626 records.

Given that *A. gularis* is likely a complex of species currently divided by phenotype (Walker 1966; del Carmen 2012; Iván 2014), we partitioned our records assigned to *A. gularis* based on a non-overlapping set of subspecies corresponding to phenotype: largely according to Walker (1966). We are aware of extensive reports of hybridization under this model of diversity within the *A. gularis* complex (e.g., Cordes 1991; Reeder et al. 2002; del Carmen 2012; Iván 2014). Provided our incomplete knowledge of species level diversity within this complex, we excluded known or suspected hybrids and assigned all individuals to only one subspecies.

Further work on species level relationships within *A. gularis* may afford a more realistic partitioning of diversity, but until then, our partitioning logic comes with the desirable property of being entirely repeatable.

To partition individuals by subspecies, we plotted all localities in Google Earth and drew non-overlapping polygons corresponding to *A. gularis* subspecies. We used maps from the following reports: Zweifel (1961), Duellman and Zweifel (1962), Dixon et al. (1971), Walker (1981a), Walker (1981b), Walker et al. (2001), del Carmen (2012), and Iván (2014). By combining the results of these papers, we assigned individuals to the following nonoverlapping *A. gularis* subspecies groups: *A. g. gularis*, *A. g. pallidus*, *A. g. ruani* (*sensu* Walker 1966), *A. g. scalaris*, *A. g. semiannulatus*, *A. g. semifasciatus*, *A. g. ssp.* (*sensu* Walker 1981a), and *A. g. septemvittata*. When subspecies accounts were ambiguous or overlapped with another subspecies, we deferred to Walker (1966) due to its relatively complete account of *A. gularis* phenotypic diversity across the entire range of the complex. Using these non-overlapping polygons and Boolean logic, localities were assigned to one of these eight subspecies. In practice, we found reports of *A. g. ruani* and *A. g. colossus* to be largely indistinguishable based primarily on locality, so we combined these subspecies' records.

To estimate suitable niche space, ENMs were generated using 19 bioclimatic variables available from Worldclim.org in the program Maxent (Phillips et al. 2006). Occurrence data were spatially rarefied (10 km distance) to reduce the effects of spatial autocorrelation (Boria et al. 2014), and a minimum training presence threshold was applied. All model parameters were determined using the spatial jackknifing method in SDMToolbox available in ArcGIS v 10.2 (ESRI 2013; Brown 2014). Combinations of models were tested with regularization multiplier values from 0.5 to 5 in 0.5 increments and five feature class combinations. Parameters for each

species' model are listed in Table S2. Models were then projected to the past, including the mid-Holocene (~6 ka), the last glacial maximum (~21 ka), and the last interglacial (~120 ka), and overlap of suitable areas for *Aspidoscelis marmorata*, *A. tesselata*, and *A. gularis* subtypes were visualized in ArcGIS.

Data Processing

Mitochondrial Genomes.- Using FASTX-Toolkit v 0.0.13.2 (http://hannonlab.cshl.edu/fastx_toolkit/index.html), we demultiplexed Illumina MiSeq data into libraries corresponding to Illumina indices. We then removed the 8 bp unique molecular identifier, demultiplexed individual libraries based on their unique 5 bp barcode, removed the 5 bp barcode and an extra 1 bp (added during dA-tailing), and removed low-quality reads (≥ 90% of bases needed Phred scores ≥20 to be retained). To generate a *de novo* reference whiptail mitochondrial genome, we used AbySS (Simpson et al. 2009) and CAP3 (Huang and Madan 1999) on an *A. gularis* dataset with a k-mer size of 64 and otherwise default options. We found that Velvet did not produce good final assemblies and Geneious R7 to be too slow for our needs, even with fairly small expected assembly size. The AbySS/CAP3 reference assembly could then be indexed and used by BWA-MEM (Li 2013) as a reference for other whiptail datasets. Using the default BWA-MEM parameters, output in SAM format was then converted back to paired FASTQ format for use by AbySS and CAP3. This process of guided de novo assembly improved assembly length and reduced time to assembly when compared to using AbySS and CAP3 without first using BWA-MEM.

We used the online version NCBI's nr/nt blastn to verify that the longest recovered contig belonged to a squamate (rather than bacterial contamination). After verification, we excised this longest contig into a new FASTA file. Next, we annotated the reference *A. gularis* genome by

uploading the AbySS/CAP3 assembly to the Mitos WebServer (mitos.bioinf.uni-leipzig.de/index.py; Bernt et al. 2013). This returns a GFF file of gene annotations, to which we appended the FASTA-format sequence used in the Mitos query. This annotated *A. gularis* genome served as our reference *Aspidoscelis* mitochondrial genome.

At this point, we selected 37 individuals to be used in the current study. We recovered each annotated genome as just described. We aligned these genomes MUSCLE (Edgar 2004) and gaps and poorly aligned sequences removed using Gblocks 0.91b (Castresana 2000). We then aligned the remaining nucleotides to the new *Aspidoscelis* mitochondrial genome reference using the Geneious aligner in Geneious 6.1.8 (Kearse et al. 2012). At this point, we used the gene annotations from the whiptail reference to generate data partitions for PartitionFinder 1.1.1 (Guindon et al. 2010; Lanfear et al. 2012). We partitioned the data into 13 protein-coding genes (each split by codon), 20 tRNAs, 2 rRNA subunits, and the origin of lagging strand replication. PartitionFinder recovered 11 partitions: 10 partitions and one extra partition for loci not included in the other 10 partitions (Table S1). We ran PartitionFinder while including all available models of molecular evolution and once again using only models implemented in MrBayes. We compiled the data, partitions, and MrBayes-only models of molecular evolution into a nexus file and analyzed these data using MrBayes 3.2.6 (Ronquist et al. 2012). Model parameters were unlinked between partitions and rates were allowed to vary under a flat Dirichlet prior. In MrBayes, we set up four runs with two chains each to run for 100,000,000 generations and sampled every 1,000 generations. The first 25% of these data were discarded as burn-in. After verifying stationarity (ESS > 200) of runs in Tracer 1.6 (Rambaut et al. 2014), we combined remaining trees from all parallel runs. We used TreeAnnotator 1.8.3 (Drummond et al. 2012) to generate a majority rule consensus tree from the combined runs.

To infer divergence times of *A. tesselata* from *A. marmorata* from molecular evidence, we constructed two time-calibrated trees using BEAST 2.3.2 (Bouckaert et al. 2014) – one using the Yule model and another using the birth-death model tree prior. For each model, we selected a relaxed clock and used the same partition scheme as in the MrBayes analysis; however, we used the more complete model set provided to PartitionFinder. We ran each analysis for 100,000,000 generations and sampled every 1,000 generations.

In calibrating the time trees, we incorporated four priors based on fossil and molecular evidence. (1) We set the divergence between *Podarcis* and *Aspidoscelis* using a normal distribution at 116.9 Ma ± 4.2 SD to match the divergence of the Laterata estimated from a calibrated phylogeny of the Lacertidae (Node 2 of Fig. 2 in Hipsley et al. 2009). (2) Based on fossils assigned to the Teiidae from the late Campanian (80–70 Ma), Mulcahy et al. (2012) conservatively estimated the Teiioidae crown group to be at least 70 Myr old, (Node 7 of Fig. 5 in Mulcahy et al. 2012). In justifying this node age, the authors admit older fossils (as catalogued on Paleobiology Database; Alroy et al. 2009) have been attributed to the Teiidae, up to 112 Myr old. In combining this information, we applied a log normal prior at 2.0 Ma ± 0.89 SD, and offset 70.0 Ma to a clade including all Teiids in our analysis (i.e., all but the *Podarcis muralis* sample). These authors also estimated the divergence time between Laterata and Teiioidae using ND2 data and BEAST: 163.5–117.6 Ma (estimated node F in Appendix I in Mulcahy et al. 2012). This provides further support for our first calibration point. (3) Additionally, Holman 1979 describes a fossil of "*A. sexlineata*" (University of Michigan Museum of Paleontology V61051) from a lower jaw dated to the 'late Aftonian' (i.e., Pre-Illinoian Stage: ~2.5–0.5 Ma). To represent this, we also included a log normal prior at 0.0 ± 1.0 with no offset for a clade including our *A. inornata* and *A. sexlineata* samples. (4) Finally, Brattstrom (1964) described "*A.*

tigris" jaw bones from a Quaternary gypsum deposit in remains dated to be 8–10,000 y old, effectively placing a maximum age on *A. marmorata* at the beginning of the Quaternary (~2.588 Ma). We incorporated this information with a uniform prior with bounds of 0–2.588 Ma and no offset. BEAST and MrBayes analyses were run on the CIPRES portal (Miller et al. 2010).

ddRADseq.- UT Southwestern Medical Center demultiplexed Illumina data into the sequence libraries corresponding to Illumina indices. Only forward-reads were used in subsequent analyses. We removed the 8 bp unique molecular identifier from each sequence using fastx trimmer (Gordon 2014). Then, we demultiplexed individual libraries based on the 8 bp barcode, which we then removed along with the 4 bp *MspI* restriction sequence; thus, yielding 80 bp reads. We removed FASTQ files with fewer than 10,000 reads (i.e., 20 out of 177 samples; 10,000 reads was ~0.2% of the sample with the most reads) as to avoid including poor loci data that could confound downstream analyses. To maximize the recovery of loci, we mapped RAD loci to a de novo *Aspidoscelis marmorata* reference genome to be described by Peter Baumann and colleagues (see Chapter 4). We used bowtie2 to index the *A. marmorata* genome. Each input file was aligned using a fairly strict and sensitive parameter set: -D 20 –R 3 –N 1 –L 18 –i S,1,0.5. This procedure output aligned SAM files which we then processed using the Stacks 1.37 (Catchen et al. 2011, 2013) pipeline as follows: *pstacks*, *cstacks*, *sstacks*, and *populations*. We investigated many options of the last three of these programs by filtering for reads recovered per individual, different population assignments provided to *populations*, and site coverage across and within individuals. In all runs, we compressed multi-allelic loci recovered by *sstacks* to be biallelic. As SNPs within a stack are closely linked, we used only the first SNP per stack. Using a random SNP from within a stack generally produced similar results. In cases where a locus had a sample with more than two haplotypes (i.e., sequencing error or poor stack assembly) the locus

was entirely removed. Finally, we used custom python scripts to transform processed Stacks output into various file formats used for phylogenetics.

To assess population structure, we used the Bayesian clustering program STRUCTURE (Pritchard et al. 2000) and discrete analysis of principle components (DAPC; Jombart et al. 2010) as implemented in *adegenet* (Jombart and Ahmed 2011). In STRUCTURE, we first included biallelic SNPs from the *Aspidoscelis gularis*, *A. marmorata*, and *A. tesselata* species complexes. We would not expect to recover population signal within a species given the evidence for recent divergence times between *A. gularis* and *A. marmorata* species (Densmore et al. 1989a, b; Reeder et al. 2002; Zheng and Wiens 2016). Therefore, we further divided genomic data in separate runs of STRUCTURE based on the results of the first evaluation. To test the tradeoff between low missing data and number of SNPs available to the model, we tested each population using the 'r' flag in *populations* with the values 0, 50, 70, 80, 90, and 100. We then ran STRUCTURE on the respective runs that had the highest stringency that allowed for ≥200 loci: for all three species, r=50; for *A. gularis*, r=70; for *A. marmorata*, r=90; and for *A. tesselata*, r=80. We also filtered the data by allowing a 50% missing data tolerance per animal. Filtering the data in this way produced four datasets: 1) all three species – 104 individuals and 1191 sites; 2) *A. gularis* – 47 individuals and 695 sites; 3) *A. marmorata* – 40 individuals and 466 sites; and 4) *A. tesselata* – 40 individuals and 665 sites. Using StrAuto v1.0 (Chhatre and Emerson 2016) to set up analyses, subpopulations (K) of 1 through 10 were analyzed using an MCMC chain with 200,000 generations and the first 50,000 discarded as burnin. We evaluated results using the Evanno method (Evanno et al. 2005) as implemented in PopHelper (Francis 2016). Finally, we generated consensus ancestry across five independent runs of each analysis using CLUMPP (Jakobsson and Rosenberg 2007) and generated ancestral bar plots using PopHelper.

As a complement to STRUCTURE analyses for assessing underlying population structure, we used DAPC. Recent studies have utilized DAPC and STRUCTURE and recovered similar results (e.g., Dailianis et al. 2011; Kanno et al. 2011; Henry et al. 2012). Similar to principal components analysis, DAPC transforms data (in this case genotypes) into uncorrelated principal components. A discriminant analysis then applies to these principal components to maximize variation among groups and minimize variation between them. Notably, DAPC is a model-free approach; thus, it makes no assumptions about Hardy-Weinberg equilibrium (such as in STRUCTURE). This is appropriate in the context of this study, especially in the case of a parthenogenetic and hybrid species. In our DAPC, we analyzed dataset 1 to potentially uncover populations of *A. tesselata* not recovered by STRUCTURE. Using the *find.clusters* function, we saw a fairly clear penalty for fewer than five or greater than ten clusters. Six clusters provided the highest information content, so we modeled this number in our analysis. To avoid over-fitting the model, we included only the first four of >100 principal components. This allowed us to explain 77% of the genetic variance in our data set.

We produced a maximum likelihood phylogeny using the SSE3 parallelized version of RaxML 8.2.9 (Stamatakis 2014). This dataset included all bisexual whiptail species analyzed separately with *cstacks*, *sstacks*, and *populations* in Stacks. We used dataset 1 from the STRUCTURE analysis and further trimmed the data resulting in a dataset of 79 individuals and 399 SNPs with at most 25% missing data per sample. The process ran using the rapid hill climbing algorithm implemented in Stamatakis et al. (2007) and with 1000 bootstrap pseudoreplicates.

We ran EEMS (Petkova et al. 2016) on each species complex – datasets 2–4 from STRUCTURE analyses. EEMS assigns individuals to demes and uses a dissimilarity matrix to

estimate effective migration rates between demes. To do this, we first converted adegenet input files (coded with all four bases) to binary .bed files using PLINK 1.07 (Purcell et al. 2007). We then calculated genetic dissimilarities between samples using the bed2diffs program included within EEMS. In EEMS, deme size is arbitrarily set, but can run the risk of excessive partitioning or summarization. To check this balance, for each species we tested 200 and 400 demes. By providing EEMS with the sample's coordinates and an outer boundary (based roughly on Fig. 2.2), we ran EEMS using three MCMC chains each with 8,000,000 generations, 2,000,000 discarded as burnin, and thin iterations set to 9,999. Using the rEEMSplots package in R, we combined MCMC chains per species and mapped results.

Results

Mitochondrial Origins of Aspidoscelis tesselata

We sequenced, assembled, and annotated 37 *Aspidoscelis* mitochondrial genomes using 300 bp paired-end reads. Average coverage exceeded 100X per sample (Fig. 2.3). Assembled mitochondrial genomes had a median length of 14,879 bp. Other than five assemblies shorter than 14 kb and one greater than 15 kb, all assemblies were within 1–3 bp of the median assembly length. MrBayes and both BEAST mitochondrial phylogenies recovered the same topology. *Aspidoscelis sexlineata* and *A. tigris* clades were recovered as in Reeder et al. (2002). As expected, mitochondrial evidence grouped *A. tesselata* together and within the *A. marmorata* clade (arrows in Fig. 2.3). One *A. tesselata* sample (ASH 102; UTA R-62291) was initially recovered separately within *A. marmorata*. Suspecting contamination or mislabeling, we resequenced this animal twice. The replicates group together and within *A. tesselata* (Fig. 2.3A). In these data, the last common ancestor of *A. tesselata* shares a node with a small clade of *A. m. marmorata* collected in the Trans-Pecos of west Texas along the Rio Grande. Our BEAST

analysis estimated this node's age, and thus the age of *A. tesselata*, to be 0.1343–0.5882 Ma (95% HPD) using a birth-death tree prior and a very similar estimate of 0.1344–0.5505 Ma using a Yule tree prior.

ddRADseq and Population Structure

As mitochondria are cytoplasmically inherited, nuclear data (ddRADseq) were necessary to comment further on the phylogenetic origins of *A. tesselata*. Of the 192 animals sequenced with ddRADseq, 157 passed our data filtering steps. In analyzing *A. tesselata*, *A. gularis*, and *A. marmorata* together in STRUCTURE, the Evanno method suggested a K=3 being most informative. When the population ancestry data were graphed with K=2, *A. tesselata* easily revealed its hybrid origins (Fig. 2.4). Subsetting these data by species allowed further insight into each species' population structure. In *A. marmorata*, a K=2 was most informative, though setting K=3 isolated the one *A. tigris* (*sensu lato*) in our dataset. Across all examined pattern classes of *A. tesselata* and *A. dixoni*, a K=2 was most informative, though a K=1 would likely be most appropriate. Two individuals showed ancestry at least 40% attributed to a second population. Both were attributed to *A. tesselata* pattern class E and collected from within 20 km of Van Horn in far west Texas. Examining both specimens confirmed them to be *A. tesselata* and pattern class E. Outwardly, they did not present obvious evidence of being backcrosses with a parental species, though this is an intriguing possibility (but see arguments in Falush et al. 2016). Finally, sampled *A. gularis* were best explained with K=3. Populations of *A. gularis* split by nominal subspecies except two *A. g. gularis* from Querétaro, Mexico which grouped as a third population.

Using DAPC, we were able to explain 77% of the genetic variance in a dataset containing *A. tesselata*, *A. marmorata*, and *A. gularis* (Fig. 2.5). This analysis recovered 6 clusters. Modeled in this way, *A. gularis* grouped together and *A. marmorata* clustered as two closely related

groups. The two *A. marmorata* groups loosely split by subspecies, but some exceptions occurred. Unsurprising for a hybrid, *A. tesselata* clustered on the first discriminant function nearly equidistant between *A. marmorata* and *A. gularis*, but with a slightly greater affinity for *A. marmorata*. In contrast to the two bisexual species, *A. tesselata* was assigned to three groups overlapping on the first discriminant function but stratifying on the second. Populations were not assigned according to pattern classes and *A. dixoni* clustered within these groups rather than as a separate group. Three *A. tesselata* group together in an interesting manner (top-most cluster in the middle of Fig. 2.5), as these animals were collected from geographically distant locations: San Miguel and Chaves counties in New Mexico and Brewster County in Texas.

The RaxML phylogeny (Fig. 2.6) recovered the same clades within *A. gularis* as in STRUCTURE (Fig. 2.4). Subspecies within *A. marmorata* did not resolve so clearly, indicating incomplete lineage sorting between these subspecies, at least with these data. Bootstrap support throughout the tree was rather poor, likely due to including only a small number of SNPs. Nevertheless, the data should be randomly distributed throughout the genomes of these species; thus, the data are presumed to be unlinked and considered high quality. Due to the stringency of our data filtering for RaxML, no proper outgroup could be included. Allelic dropout of increasingly distant taxa likely caused this pattern. Thus, we rooted at the midpoint between *A. g. gularis* of Querétaro, and *A. marmorata* (*sensu lato*). This rooting process resulted in an *A. inornata heptagramma* falling within the *A. gularis* clade. Though *A. inornata* and *A. gularis* are both members of the *A. sexlineata* clade, we suspect that the paraphyly of *A. gularis* was an artifact of this rooting process and *A. inornata*'s relatively long branch length.

We successfully modeled effective migration and diversification rates for *Aspidoscelis gularis*, *A. marmorata*, and *A. tesselata* using EEMS (Fig. 2.7). Both 200 deme and 400 deme

models produced similar results; thus, we will focus on the 200 deme data. Models for *A. gularis* indicated this species easily traversed the plains of Texas and Oklahoma, but share fewer haplotypes in the more heterogeneous environments of the Trans-Pecos, southern Texas, and Mexico (Figs. 2.7A, D). As the distribution and habitat use of *A. marmorata* differs from *A. gularis*, it was unsurprising to recover different effective migration models for *A. marmorata* (Figs. 2.7C, F). In *A. marmorata*, the relatively uninhabited (by this species) middle third of New Mexico (Fig. 2.1) provided a barrier to effective migration. This corresponds to the area between the Rio Grande and Pecos River in the southern two thirds of the state. This lower effective migration surface extends at a lesser intensity through the Trans-Pecos which is where the greatest phenotypic diversity of this species is described (Hendricks and Dixon 1986). Interestingly, EEMS recovered a more patchy effective migration surface for *A. tesselata* and *A. dixoni* (Figs. 2.7B, E). Effective migration was much lower in the Trans-Pecos in between the Guadalupe and Chinati mountains and also the Big Bend area of the Rio Grande. Effective migration was highest in a corridor extending from the Peloncillo Mountains on the border of New Mexico and Arizona, eastward towards Roswell, NM, and then south to a relatively narrow range between and including the Davis and Chinati Mountains.

ENM Projections

We incorporated 7,626 points (1010 *A. tesselata*, 2,518 *A. marmorata*, and 4,098 *A. gularis* complex) into an ENM model and projected this model to the last glacial maximum (Fig. 2.6), mid-Holocene, and last interglacial. When projected to the last glacial maximum (Fig. 2.8), we recovered two major areas of overlap between *A. tesselata*, *A. marmorata*, and the *A. gularis* species cluster corresponding with *A. g. septemvittata* and *A. g. scalaris*. This clade of *A. gularis* corresponds to group 3 in ND2 and CytB analyses of Mexican *A. gularis* recovered by del

Carmen (2012). One overlapping region corresponds with the north central Chihuahuan desert near the Cochise filter barrier near the Peloncillo Mountains. The other, larger, region of overlap occurs on the northeastern edge of the Chihuahuan desert, mostly south of the Big Bend region of the Rio Grande River. Overlap between *A. gularis* and *A. marmorata* was similar to the overlap between three species, but overlap extended farther south in both regions.

Discussion

Previous investigators hypothesized a single origin of the hybrid and parthenogenetic lizard *Aspidoscelis tesselata* (Maslin 1967; Cordes and Walker 2003, 2006; Taylor et al. 2003). We recovered a single maternal origin of *A. tesselata* based on whole mitochondrial genomes (Fig. 2.3). Using many anonymous nuclear loci we also recovered evidence for genetic structure within *A. tesselata* not well explained by the traditional model of single origins followed by clonal reproduction (Figs. 2.4, 2.5). This structure did not strongly correlate with well-characterized pattern classes – phenotypic dorsal color patterns that differ between largely allopatric populations.

There was a slight geographical pattern of genetic diversity within *A. tesselata* (Fig. 2.5): one group consisted of three animals along an enormous transect between Big Bend National Park, TX and Sumner Lake State Park, NM; a second group included mostly animals from Guadalupe and Eddy counties in New Mexico, but also one *A. dixoni* of the Chinati Mountains in Texas and an *A. tesselata* from Caprock Canyons State Park, TX; and all other samples fit into third catch-all group of *A. tesselata*. Notably, despite being phenotypically distinct and geographically isolated, all three pattern classes of *A. dixoni* distributed evenly between the two more common *A. tesselata* pattern classes, and mostly fit into the ‘catch-all’ cluster (Figs. 2.4,

2.5). These data suggest, as has been previously hypothesized (e.g., Cordes and Walker 2006), that *A. dixoni* and *A. tesselata* are the same evolutionary unit.

Originally, Zweifel (1965) described *A. dixoni* as pattern class F of *A. tesselata*, refraining from recommending *A. tesselata* F as its own species. Despite all being diploid and a handful of studies suggesting that *A. tesselata* pattern classes C–F arose from the same pairing of species (perhaps the same clutch), Scudday (1973) recommended promoting *A. tesselata* F to species status as *A. dixoni*. This conclusion was based primarily on a morphological species concept after examining museum specimens and collections by Scudday (1971; also Scudday and Dixon 1973). Known populations of *A. dixoni* can be identified by dorsal pattern and distribution, but little else distinguishes it from *A. tesselata* (Walker et al. 1994). Scudday (1971, 1973) referenced unpublished karyotypes to be published by Wright, but these data remain unpublished.

The evolutionary and biological species concepts struggle to accommodate parthenogenetic vertebrates (Cole 1985; Frost and Wright 1988; Taylor et al. 2005; Winkler et al. 2007; Cole et al. 2014). Taken to an extreme, each individual could be considered an evolutionarily distinct unit, unable to exchange genes with others except in rare cases of hybridization. This interpretation is generally perceived to be untenable, as naming every individual as its own species is cumbersome at best but also is not in the spirit of a species concept (Mayr 1996; Winkler et al. 2007). Although this study did not set out to make taxonomic recommendations, it seems the species status of *A. dixoni* bears revisiting in light of our findings.

Our mitochondrial genome phylogeny of several whiptails places the maternal origins of the hybrid in middle Pleistocene (Fig. 2.3). To our knowledge, this is the first estimation of the timing of hybridization between *A. gularis* and *A. marmorata* for extant *A. tesselata*. Parker and

Selander (1976) discussed possibilities of very recent origins of *A. tesselata*, even as recently as 250 years ago based on overgrazing of short-grass prairie and *A. tesslelata*'s affinity for disturbed habitat. These authors also reference Axtell's conclusion that another parthenogenetic lizard *A. neomexicanus* may have arisen in the Wisconsin period (~12 ka). Our estimate is 2–4 orders of magnitude older than these previous estimations, but the first based on a time-calibrated phylogeny. A better sampled *Aspidoscelis* mitochondrial phylogeny and additional North American fossil Teiids would add further clarity to the timing of the origins of *A. tesselata*.

Of course, more than one hybridization event may have occurred to generate *A. tesselata*. In such a case, it seems likely that the progeny of those events have since gone extinct; thus, precluding us from sampling them. In more recent time, mapping the ecological niches of *A. gularis*, *A. tesselata*, and *A. marmorata* indicate two geographic areas of overlap (Fig. 2.8). Given the cyclical nature of glaciation in the Pleistocene, it may have been that there were sympatric populations of *A. marmorata* and *A. gularis* in one or both of these locations at the estimated time of origin for *A. tesselata*. Kearney (2005) suggested a geographical pattern of parthenogenesis associated with glaciation in the context of a hybrid's advantage in open environments. If this logic holds, *A. tesselata* originated from one or both of the following and migrated north: (1) south of the Peloncillo Mountains in northern Chihuahua, Mexico; and (2) southeast of Big Bend, TX, in Coahuila or Nuevo Leon, Mexico. These geographic origins agree with Parker and Selander (1976) who noted that all *A. tesselata* they sampled shared three alleles with *A. gularis septemvittatus* from Pinto Canyon near the Chinati Mountains. Further phylogenetic investigation into northern Mexico's *A. gularis* populations may yet recover a clade more closely related to the paternal ancestor of *A. tesselata* than has currently been accessible and studied.

DAPC recovered evidence of population structure in *A. tesselata* not well explained by pattern classes (Fig. 2.5). Pattern classes in *A. tesselata* were initially described on morphology and color pattern alone (Zweifel 1965). As this study produced the first sequence data for *A. tesselata*, directly comparing our results to previous work is challenging. For example, Parker and Selander (1976) used 21 allozyme loci to recover evidence of variation in *A. tesselata*. These authors' invoked mutation, recombination, and hybridization in explaining three patterns found at six variable loci. Concurrent with other investigators (Maslin 1967; Cordes and Walker 2003, 2006; Taylor et al. 2003), Parker and Selander (1976) did not suggest more than one hybrid origin of *A. tesselata* and their investigation appears to have been based on an assumption of a single origin. Additionally, in unisexual lizards the assumption has been that reciprocal transplants of skin (i.e., skin histocompatibility) lend evidence for pattern classes being the same evolutionary unit or not (Maslin 1967; Cuellar 1997; Cordes and Walker 2003, 2006). This test indirectly measures evolution at the MHC locus, as failed skin grafts should represent different MHC antigens produced by different MHC orthologs (Gould and Auchincloss 1999). In contrast to the single MHC locus, our dataset analyzing structure across *A. tesselata*, *A. gularis*, and *A. marmorata* included over 1000 loci anonymously collected throughout the genome. As increasing the number of phylogenetically informative loci generally improves one's phylogenetic resolution and support (e.g., Hillis et al. 1994; Delsuc et al. 2005), substantially increasing the number of characters such as by ddRADseq should orient our inferred phylogeny towards a most true representation.

Multiple origins of parthenogenetic species complexes have been imputed before – even in *Aspidoscelis*. Manríquez-Morán et al. (2014) inferred at least two hybridization events between *A. angusticeps* and *A. deppii* led to three unisexual species: *A. rodecki*, *A. maslini*, and *A.*

cozumela (see also Morán 2007). Furthermore, in creating a lab strain of *A. neavsii* Cole and colleagues artificially invoked multiple origins of a tetraploid parthenogenetic whiptail also discovered in the wild (Cole et al. 2014). We are, though, faced with reconciling a single mitochondrial origin, but three populations, of *A. tesselata*. Given the mitochondrial and nuclear data (Figs. 2.3–2.6) and ENM results (Fig. 2.8), our conclusion is that *A. tesselata* originated more than once, but we were unable to sample all *A. marmorata* lineages involved in producing *A. tesselata* (Fig. 2.1B). This could be simple sampling bias, as no samples of *A. marmorata* were collected south of the US-Mexico border. Alternatively, one or more contributing *A. marmorata* lineages have subsequently gone extinct and can no longer be sampled. Ultimately, there are limited origins of *A. tesselata* and heritable mechanisms other than multiple origins – genetic drift, mutation, hybridization, or mutation (e.g., Parker and Selander 1976) – likely explain the bulk of phenotypic diversity seen in *A. tesselata*.

Acknowledgements

We thank Jessica Majid for helping ASH catch animals. Corey Roelke, Travis LaDuc, Carl Lieb, Tom Giermakowski, Toby Hibbitts, Jay Cole, Jeff Streicher, Peter Baumann, and Christian Cox were especially instrumental by donating tissues or providing access to tissues. Jeff Streicher, Daniel Portik, James Titus-McQuillan, and Kyle O'Connell improved several aspects of the manuscript drafts through careful review and discussion. Daniel Portik provided code for automating the use of Stacks. This project was funded by an SSE Rosemary Grant to ASH, two Beta Phi chapter Phi Sigma research grants to ASH, a travel grant from the UT Arlington graduate student senate to ASH, startup funds to MKF from UT Arlington, an NSF GRFP to DR, and an NSF DDIG to ASH (NSF-1501675). Animals collected for this project were obtained under New Mexico Energy, Minerals and Natural Resources Department Research

Permit 2013-7; New Mexico Department of Game and Fish 3557; and Arizona Game and Fish Department #SP612813; and Texas Parks and Wildlife Department (TPWD) permits SPR-0707-1387 and SPR-0513-066. David Riskind of TPWD allowed collection within Texas State Parks. Collection at Black Gap WMA was authorized by permit BG 13-06 at Elephant Mountain WMA by permit EM 13-03. All activities in this project were allowed by UT Arlington's IACUC protocol number A13.010.

Literature Cited

Alroy, J., C. Marshall, and A. Miller. 2009. Paleobiology database. Available at https://paleobiodb.org. Accessed May 23, 2016.

Avise, J. C. 2008. Clonality: the genetics, ecology, and evolution of sexual abstinence in vertebrate animals. Oxford Univ. Press, Oxford, U.K.

Baird, N. A., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z. A. Lewis, E. U. Selker, et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS ONE*, 10:e3376.

Bernt, M., A. Donath, F. Jühling, F. Externbrink, C. Florentz, G. Fritzsch, J. Pütz, et al. 2013. MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, 69:313–319.

Booth, W., D. H. Johnson, S. Moore, C. Schal, and E. L. Vargo. 2011a. Evidence for viable, non-clonal but fatherless Boa constrictors. *Biology Letters*, 7:253–256.

Booth, W., L. Million, R. Graham Reynolds, G. M. Burghardt, E. L. Vargo, C. Schal, A. C. Tzika, et al. 2011b. Consecutive virgin births in the new world boid snake, the Colombian rainbow boa, *Epicrates maurus*. *Journal of Heredity*, 102:759–763.

Boria, R.A., Olson, L.E., Goodman, S.M., Anderson, R.P. 2014. Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling*, 275:73–77.

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 Computational Biology*, 10:e1003537.

Brattstrom, B. H. 1954. Amphibians and reptiles from Gypsum Cave, Nevada. *Bulletin of the Southern California Academy of Sciences*, 53:8–12.

Brown, J.L. 2014. SDMtoolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods in Ecology and Evolution*, 5:694–700.

Brown, W. M., and J. W. Wright. 1979. Mitochondrial DNA analyses and the origin and relative age of parthenogenetic lizards (genus *Cnemidophorus*). *Science*, 203:1247–1249.

Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution* 17:540–552.

Catchen, J., A. Amores, P. Hohenlohe, W. Cresko, and J. H. Postlethwait. 2011. *Stacks*: building and genotyping loci *de novo* from short-read sequences. *G3 Genes, Genomes, Genetics*, 1:171–182.

Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. *Molecular Ecology*, 22:3124–3140.

Chapman, D. D., M. S. Shivji, E. Louis, J. Sommer, H. Fletcher, and P. A. Prodhöhl. 2007. Virgin birth in a hammerhead shark. *Biology Letters*, 3:425–427.

Chhatre, V. E., and K. J. Emerson. 2016. StrAuto: automation and parallelization of STRUCTURE analysis. Available at http://strauto.popgen.org. Accessed September 20, 2016.

Cole, C. J. 1985. Taxonomy of parthenogenetic species of hybrid origin. *Systematic Zoology*, 34:359–363.

Cole, C. J., H. L. Taylor, D. P. Baumann, and P. Baumann. 2014. Neaves' whiptail lizard: the first known tetraploid parthenogenetic tetrapod (Reptilia: Squamata: Teiidae). *Breviora*, 539:1–20.

Cordes, J. E. 1991. Biology of two sympatric clonal complexes of parthenogenetic whiptail lizards [TES-G (2n) and TES-H (2n)] in the *Cnemidophorus tesselatus* complex and identification of hybrid TES-G (2n) x *Cnemidophorus gularis septemvittatus* in the Chinati Mountains in trans-Pecos Texas. Thesis (Ph.D.). University of Arkansas.

Cordes, J. E., and J. M. Walker. 2003. Skin histocompatibility between syntopic pattern classes C and D of parthenogenetic *Cnemidophorus tesselatus* in New Mexico. *Journal of Herpetology*, 37:185–188.

Cordes, J. E., and J. M. Walker. 2006. Evolutionary and systematic implications of skin histocompatibility among parthenogenetic Teiid lizards: three color pattern classes of *Aspidoscelis dixoni* and one of *Aspidoscelis tesselata*. *Copeia*, 2006:14–26.

Cuellar, O. 1977. Genetic homogeneity and speciation in the parthenogenetic lizards *Cnemidophorus velox* and *C. neomexicanus*: evidence from intraspecific histocompatibility. *Evolution*, 31:24–31.

Dailianis, T., C. S. Tsigenopoulous, C. Dounas, and E. Voultsiadou. 2011. Genetic diversity of the imperiled bath sponge *Spongia officinalis* Linnaeus, 1759 across the Mediterranean Sea: patterns of population differentiation and implications for taxonomy and conservation. *Molecular Ecology*, 50:3757–3772.

Del Carmen, G. C. M. 2012. Relaciones filogenéticas y delimitación de especies en el complejo *Aspidoscelis gularis* (Sauria: Teiidae). Thesis (M. S.). Universidad Autónoma del Estado de Hidalgo.

Delsuc, F. H. Brinkmann, H. Philippe. 2005. Phylogenomics and the reconstruction of the Tree of Life. *Nature Reviews Genetics*, 6:361–375.

Densmore, L. D., III, C. C. Moritz, J. W. Wright, and W. M. Brown. 1989a. Mitochondrial-DNA analyses and the origin and relative age of parthenogenetic lizards (genus *Cnemidophorus*). IV. Nine *sexlineatus*-group unisexuals. *Evolution*, 43:969–983.

Densmore, L. D., III, J. W. Wright, and W. M. Brown. 1989b. Mitochondrial-DNA analyses and the origin and relative age of parthenogenetic lizards (genus *Cnemidophorus*). II. *C. neomexicanus* and the *C. tesselatus* complex. *Evolution*, 43:943–957.

Dessauer, H. C., and C.J. Cole. 1986. Clonal inheritance in parthenogenetic whiptail lizards: biochemical evidence. *Journal of Heredity*, 77:8–12.

Dixon, J. R., C. S. Lieb, and C. A. Ketchersid. 1971. A new lizard of the genus *Cnemidophorus* (Teiidae) from Querétaro, Mexico. *Herpetologica*, 27:344–354.

Drummond, A. J., M. A. Suchard, D. Xie, and A. Rambaut. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, 29:1969–1973.

Duellman, W. E., and R. G. Zweifel. 1962. A synopsis of the lizards of the *sexlineatus* group (genus *Cnemidophorus*). *Bulletin of the American Museum of Natural History*, 123:155–210.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32:1792–1797.

Edwards, S. V., L. Liu, and D. K. Pearl. 2007. High-resolution species trees without concatenation. *Proceedings of the National Academy of Science USA*, 104:5936–5941.

ESRI, 2013. ArcGIS Desktop: Release 10.2. Environmental Systems Research Institute, Redlands, U.S.A.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14:2611–2620.

Falush, D., L. van Dorp, and D. J. Lawson. 2016. A tutorial on how (not) to over-interpret STRUCTURE/ADMIXTRUE bar plots. bioRxiv, DOI: http://dx.doi.org/10.1101/066431.

Fitzinger, L. 1843. *Systema Reptilium*. Vienna: Vindobonae, 106 pp.

Frost, D. R., and J. W. Wright. 1988. The taxonomy of uniparental species with special reference to parthenogenetic *Cnemidophorus* (Squamata: Teiidae). *Systematic Zoology*, 37:200–209.

Fujita, M. K., and C. Moritz. 2010. Origin and evolution of parthenogenetic genomes in lizards: current state and future directions. *Cytogenetic and Genome Research*, 127:261–272.

Gordon, A. 2014. FASTX-Toolkit. Available at http://hannonlab.cshl.edu/fastx_toolkit. Accessed March 20, 2016.

Gould, D. S., H. Auchincloss Jr. 1999. Direct and indirect recognition: the role of MHC antigens in graft rejection. *Immunology Today*, 20:77–82.

Grismer, J. L., A. M. Bauer, L. L. Grismer, K. Thirakhupt, A. Aowphol, J. R. Oaks, P. L. Wood, Jr., et al. 2014. Multiple origins of parthenogenesis, and a revised species phylogeny for the Southeast Asian butterfly lizards, *Leiolepis*. *Biological Journal of the Linnean Society*, 113:1080–1093.

Guindon, S., J.-F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, 59:307–321.

Harris, A. H., and J. S. Findley. 1964, Pleistocene-recent fauna of the Isleta Caves, Bernalillo County, New Mexico. *American Journal of Science*, 262:114–120.

Henry, P., Z. Sim, and M. A. Russello. 2012. Genetic evidence for restricted dispersal along continuous altitudinal gradients in a climate change-sensitive mammal: the American pika. *PloS ONE*, 7:e39077.

Hendricks, F. S., and J. R. Dixon. 1986. Systematics and biogeography of *Cnemidophorus marmoratus* (Sauria: Teiidae). *Texas Journal of Science*, 38:327–402.

Hillis, D. M., J. P. Huelsenbeck, and C. W. Cunningham. 1994. Application and accuracy of molecular phylogenies. *Science*, 29:671–677.

Hipsley, C. A., L. Himmelmann, D. Metzler, and J. Müller. 2009. Integration of Bayesian molecular clock methods and fossil-based soft bounds reveals early Cenozoic origin of African Lacertid lizards. *BMC Evolutionary Biology*, 9:151.

Holman, A. 1979. Herpetofauna of the Nash local fauna (Pleistocene: Aftonian) of Kansas. *Copeia*, 1979:747–749.

Huang, X., and A. Madan. 1999. CAP3: a sequence assembly program. *Genome Research*, 9:868–877.

ICZN. 1999. International code of zoological nomenclature, 4th ed. London: International Trust for Zoological Nomenclature.

Iván, G. S. O. 2014. El código de barras genético como herramienta para la identificación de especies en el complejo *Aspidoscelis gularis*. Thesis (M. S.). Universidad Autónoma del Estado de Hidalgo.

Jakobsson, M., and N. A. Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23:1801–1806.

Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured population. *BMC Genetics*, 11:94.

Jombart, T., and I. Ahmed. 2011. *Adegenet 1.3-1*: new tools for the analysis of genome-wide SNP data. *Bioinformatics*, 27:3070–3071.

Jones, L., and R. Lovich. 2009. Lizards of the American Southwest. Rio Nuevo Publishers, Tuscon, U.S.A.

Kanno, Y., J. C. Vokoun, B. H. Letcher. 2011. Fine-scale population structure and riverscape genetics of brook trout (*Salvelinus fontinalis*) distributed continuously along headwater channel networks. *Molecular Ecology*, 20:3711–3729.

Kearney, M. 2005. Hybridization, glaciation and geographical parthenogenesis. *Trends in Ecology and Evolution*, 20:495–502.

Kearney, M., M. K. Fujita, and J. Ridenour. 2009. Lost sex in the reptiles: constraints and correlations. Pp 447–474 *in* Schön, I., K. Martens, and P. Van Dijk, eds. Lost Sex. Springer, London, U.K.

Kearse, M., R. Moir, A. Wilson, S. Stones-Havas, M. Cheung, S. Sturrock, S. Buxton, A. Coper, S. Markowitz, C. Duran, T. Thierer, B. Ashton, P. Meintjes, and A. Drummond. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28:1647–1649.

Lanfear, R., B. Calcott, S.Y.W. Ho, and S. Guindon. 2012. PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29:1695–1701.

Li, H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997 http://arxiv.org/abs/1303.3997v2.

Lutes, A. A., D. P. Baumann, W. B. Neaves, and P. Baumann. 2011. Laboratory synthesis of an independently reproducing vertebrate species, *Proceedings of the National Academy of Science USA*, 108:9910–9915.

Manríquez-Morán, N. L., F. R. Méndez-de la Cruz, and R. Murphy. 2014. Genetic variation and origin of parthenogenesis in the *Aspidoscelis cozumela* complex: evidence from mitochondrial genes. *Zoological Science*, 31:14–19.

Maslin, T. P. 1962. All-female species of the lizard genus *Cnemidophorus*, Teiidae. *Science*, 135:212–213.

Maslin, T. P. 1966. The sex of hatchlings of five apparently unisexual species of whiptail lizards (*Cnemidophorus*, Teiidae). *American Midland Naturalist*, 76:369–378.

Maslin, T. P. 1967. Skin grafting in the bisexual Teiid lizard *Cnemidophorus sexlineatus* and in the unisexual *C. tesselatus*. *Journal of Experimental Zoology*, 166:137–149.

Maslin, T. P. 1968. Taxonomic problems in parthenogenetic vertebrates. *Systematic Zoology*, 17:219–231.

Mayr, E. 1996. What is a species and what is not? *Philosophy of Science*, 63:262–277.

Miller, M. M., W. Pfeifer, and T. Schwartz. 2010. Creating the CIPRES science gateway for inference of large phylogenetic trees. In *Proceedings of the gateway computing environments workshop (GCE)*: pp. 1–8, New Orleans.

Morán, N. M. 2007. Diversidad clonal en los lacertilios unisexuales del género *Aspidoscelis*. *Boletín de la Sociedad Herpetológica Mexicana*, 15:1–12.

Moritz, C., and K. Bi. 2011. Spontaneous speciation by ploidy elevation: laboratory synthesis of a new clonal vertebrate. *Proceedings of the National Academy of Science USA*, 108:9733–9734.

Moritz, C. C., J. W. Wright, and W. M. Brown. 1989. Mitochondrial-DNA analyses and the origin and relative age of parthenogenetic lizards (genus *Cnemidophorus*). III. *C. velox* and *C. exsanguis*. *Evolution*, 43:958–968.

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. *Molecular Phylogenetics and Evolution*, 65:974–991.

Muller, H. J. 1932. Some genetic aspects of sex. *American Naturalist*, 66:118–138.

Neaves, W. B., and P. S. Gerald. 1968. Lactate dehydrogenase isozymes in parthenogenetic Teiid lizards (*Cnemidophorus*). *Science*, 160:1004–1005.

Neaves, W. B. 1969. Adenosine deaminase phenotypes among sexual and parthenogenetic lizards in the genus *Cnemidophorus* (Teiidae). *Journal of Experimental Zoology*, 171:175–183.

Neaves, W. B. 1971. Tetraploidy in a hybrid lizard of the genus *Cnemidophorus* (Teiidae). *Breviora*, 381:1–25.

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

Parker, E. D., and R. K. Selander. 1976. The organization of genetic diversity in the parthenogenetic lizard *Cnemidophorus tesselatus*. *Genetics*, 84:791–805.

Paulissen, M. A., J. M. Walker, and H. L. Taylor. 2006. Diet of sympatric pattern classes C and E of the parthenogenetic whiptail lizard *Aspidoscelis tesselata* at Sumner Lake, De Baca County, New Mexico. *The Southwestern Naturalist*, 51:555–560.

Peterson, B. K., J. N. Weber, E. H. Kay, H. S. Fischer, and H. E. Hoekstra. 2012. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS ONE*, 7:e37135.

Petkova, D., J. Novembre, and M. Stephens. 2016. Visualizing spatial population structure with estimated effective migration surfaces. *Nature Genetics*, 48:94–100.

Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, 190:231–259.

Podnar, M., W. Pinsker, and W. Mayer. 2009. Complete mitochondrial genomes of three lizard species and the systematic position of the Lacertidae (Squamata). *Journal of Zoological Systematics and Evolutionary Research*, 47:35–41.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155:945–959.

Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, P. Sklar, P. I. W. de Bakker, M. J. Daly, and P. C. Sham. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American Journal of Human Genetics*, 81:559–575.

Rambaut, A., M. A. Suchard, D. Xie, and A. J. Drummond. 2014. Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer. Accessed September 20, 2016.

Reeder, T. W., C. J. Cole, and H. C. Dessauer. 2002. Phylogenetic relationships of whiptail lizards of the genus *Cnemidophorus* (Squamata: Teiidae). *American Museum Novitates*, 3365:1–61.

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

Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M. A. Suchard, and J. P. Huelsenbeck. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61:539–542.

Schield, D. R., D. C. Card, R. H. Adams, T. Jezkova, J. Reyes-Velasco, F. N. Proctor, C. L. Spencer, et al. 2015. Incipient speciation with biased gene flow between two lineages of the western diamondback rattlesnake (*Crotalus atrox*). *Molecular Phylogenetics and Evolution*, 83:213–223.

Scudday, J. F. 1971. The biogeography and some ecological aspects of the Teiid lizards (*Cnemidophorus*) of Trans-Pecos Texas. Thesis (Ph.D.). Texas A&M University.

Scudday, J. F. 1973. A new species of lizard of the *Cnemidophorus tesselatus* group from Texas. *Journal of Herpetology*, 7:363–371.

Scudday, J. F., and J. R. Dixon. 1973. Diet and feeding behavior of Teiid lizards from Trans-Pecos, Texas. *The Southwestern Naturalist*, 18:279–289.

Simpson, J. T., K. Wong, S. D. Jackman, J. E. Schein, S. J. M. Jones, and İ. Birol. 2009. ABySS: A parallel assembler for short read sequence data. *Genome Research*, 19:1117–1123.

Stamatakis, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30:1312–1313.

Stamatakis, A., F. Blagojevec, D. S. Nikolopoulos, and C. D. Antonopoulos. 2007. Exploring new search algorithms and hardware for phylogenetics: RAxML meets the IBM cell. *Journal of VLSI Signal Processing*, 48:271–286.

Streicher, J. W., T. J. Devitt, C. S. Goldberg, J. H. Malone, H. Blackmon, and M. K. Fujita. 2014. Diversification and asymmetrical gene flow across time and space: lineage sorting and hybridization in polytypic barking frogs. *Molecular Ecology*, 23:3273–3291.

Taylor, H. L., C. J. Cole, H. C. Dessauer, and E. D. Parker, Jr. 2003. Congruent patterns of genetic and morphological variation in the parthenogenetic lizard *Aspidoscelis tesselata* (Squamata: Teiidae) and the origins of color pattern classes and genotypic clones in eastern New Mexico. *American Museum Novitates*, 3424:1–40.

Taylor, H. L., C. J. Cole, G. J. Manning, J. E. Cordes, and J. M. Walker. 2012. Comparative meristic variability in whiptail lizards (Teiidae, *Aspidoscelis*): samples of parthenogenetic *A. tesselata* versus samples of sexually reproducing *A. sexlineata*, *A. marmorata*, and *A. gularis septemvittata*. *American Museum Novitates*, 3744:1–24.

Taylor, H. L., B. A. Droll, and J. M. Walker. 2006. Proximate causes of a phylogenetic constraint on clutch size in parthenogenetic *Aspidoscelis neotesselata* (Squamata: Teiidae) and range expansion opportunities provided by hybridity. *Journal of Herpetology*, 40:294–304.

Taylor, H. L., J. M. Walker, J. E. Cordes, and G., J. Manning. 2005. Application of the evolutionary species concept to parthenogenetic entities: comparison of postformational divergence in two clones of *Aspidoscelis tesselata* and between *Aspidoscelis cozumela* and *Aspidoscelis maslini* (Squamata: Teiidae). *Journal of Herpetology*, 39:266–277.

Tinkle, D. W. 1959. Observations on the lizards *Cnemidophorus tigris*, *Cnemidophorus tessellatus* and *Crotaphytus wislizeni*. *The Southwestern Naturalist*, 4:195–200.

Tucker, D. B., G. R. Colli, L. G. Giugliano, S. B. Hedges, C. R. Hendry, E. M. Lemmon, A. R. Lemmon, J. W. Sites, Jr., and R. A. Pyron. 2016. Methodological congruence in

phylogenomic analyses with morphological support for Teiid lizards (Sauria: Teiidae). *Molecular Phylogenetics and Evolution*, 103:75–84.

Walker, J. M. 1966. Morphological variation in the Teiid lizard *Cnemidophorus gularis*. Thesis (Ph.D.). University of Colorado.

Walker, J. M. 1981a. Systematics of *Cnemidophorus gularis*. I. Reallocation of populations currently allocated to *Cnemidophorus gularis* and *Cnemidophorus scalaris* in Coahuila, México. *Copeia*, 1981:826–849.

Walker, J. M. 1981b. Systematics of *Cnemidophorus gularis*. II. Specific and subspecific identity of the Zacatecas whiptail (*Cnemidophorus gularis semiannulatus*). *Copeia*, 1981:850–868.

Walker, J. M., J. E. Cordes, C. C. Cohn, H. L. Taylor, R. V. Meyer, and L. Richard. 1994. Life history characteristics of three morphotypes in the parthenogenetic *Cnemidophorus dixoni* complex (Sauria: Teiidae) in Texas and New Mexico. *Texas Journal of Science*, 46:27–33

Walker, J. M., J. E. Cordes, and H. L. Taylor. 1997. Parthenogenetic *Cnemidophorus tesselatus* complex (Sauria: Teiidae): a neotype for diploid *C. tesselatus* (Say, 1823), redescription of the taxon, and description of a new triploid species. *Herpetologica*, 53:233–259.

Walker, J. M., J. A. Lemos-Espinal, J. E. Cordes, and H. L. Taylor. 2001. Allocation of populations of whiptail lizards to *septemvittatus* Cope, 1892 (genus *Cnemidophorus*) in Chihuahua, México, and the *scalaris* problem. *Copeia*, 2001:747–765.

Walker, J. M., H. L. Taylor, G. J. Manning, J. E. Cordes, C. E. Montgomery, L. J. Livo, S. Keefer, and C. Loeffler. 2012. Michelle's lizard: identity, relationships, and ecological

status of an array of parthenogenetic lizards (genus *Aspidoscelis*: Squamata: Teiidae) in Colorado, USA. *Herpetological Conservation and Biology*, 7:227–248.

Winkler, K., D. A. Rocque, T. M. Braile, and C. L. Pruett. 2007. Vainly beating the air: species-concept debates need not impede progress in science or conservation. *Ornithological Monographs*, 63:30–44.

Wright, J. W., C. Spolsky, and W. M. Brown. 1983. The origin of the parthenogenetic lizard *Cnemidophorus laredoensis* inferred from mitochondrial DNA analysis. *Herpetologica*, 39:410–416.

Zheng, Y., and J. J. Wiens. 2016. Combining phylogenomic and supermatrix approaches, and a time-calibrated phylogeny for squamate reptiles (lizards and snakes) based on 52 genes and 4162 species. *Molecular Phylogenetics and Evolution*, 94:537–547.

Zweifel, R. G. 1961. Relationship of two whiptail lizards (genus *Cnemidophorus*) in western Mexico. *Copeia*, 1:98–103.

Zweifel, R. G. 1965. Variation in and distribution of the unisexual lizard, *Cnemidophorus tesselatus*. *American Museum Novitates*, 2235:1–49.

Figure Legends

Figure 2.1. Postulated origins of phenotypic diversity in the hybrid parthenogenetic lizard *Aspidoscelis tesselata*. In large part, the prevailing hypothesis (A) has been that phenotypic diversity comes from heritable variation after the hybrid origin of the species between *A. marmorata* and *A. gularis*. Further study should investigate the different, not necessarily mutually exclusive, hypothesis (B) that there were multiple hybrid origins for *A. tesselata*. Diversity in *A. gularis* is currently allotted to seven subspecies, but has hardly been characterized using molecular evidence.

Figure 2.2. Sampling localities used in this study for mitochondrial genome and ddRAD sequencing. Maps on the right correspond to each species' range derived from Jones and Lovich (2009). Colors used for each species in these maps are consistent within and between figures.

Figure 2.3. One maternal origin recovered for parthenogenetic *Aspidoscelis tesselata* (see arrows). Bayesian analysis of 37 Teiid and one Lacertid (*Podarcis muralis*) mitochondrial genomes partitioned by gene and codon recovered this phylogeny (A) and cladogram (B). Branches are colored by species: *A. gularis* is red, *A. marmorata* is blue, *A. tesselata* is green, *A. sexlineata* is violet, and *A. inornata* is teal. One *A. tesselata* (ASH 102; UTA R-62291), initially did not group within other *A. tesselata*. We resequenced this animal and the replicates group with each other and within the *A. tesselata* clade, suggesting the original sample was mislabeled. The cladogram shows nodal Bayesian posterior probabilities below 0.95. The scale bar represents substitutions per site.

Figure 2.4. Population structure of hybrid, unisexual *Aspidoscelis tesselata* and its two parental species: *A. marmorata* and *A. gularis*. Inferences were derived from 1191 nuclear SNPs and 104 individuals across the whole dataset. Species-specific plots are based on independently recovered subsets of these data.

Figure 2.5. Discrete analysis of principal components (DAPC) using 1191 nuclear SNPs. Six clusters were inferred and they are colored based on their majority constituent species. 77% of the variance was explained by this model which includes four principal components. The first two discriminant functions are shown which, in turn, explain the majority of variance accounted for by this model.

Figure 2.6. RAxML phylogeny of bisexual whiptails based on 399 nuclear SNPs. Taxa are colored by species group: *Aspidoscelis gularis* are red, *A. marmorata* are blue, and other species are black. The tree was rooted at the midpoint between *A. marmorata* and *A. gularis*.

Figure 2.7. Estimated effective migration surface (EEMS) analysis in three whiptail species. The top row (A–C) shows results for effective migration rate, summarized as *m* on a logarithmic scale. The bottom row (D–F) shows results for effective diversity rate, summarized as *q* and also on a logarithmic scale. All 6 maps estimated parameters assumed 200 demes represented as gray triangles within blue boundaries. The boundaries differ because species distributions differ between species. Black dots are samples included in calculating the distance matrix underlying these analyses. The size of these dots corresponds to the number of samples at that vertex.

Figure 2.8. Mid-Holocene (~6 ka) projection of an ecological niche model for whiptail lizards related to *Aspidoscelis tesselata*. Two major areas of overlap between all three species

groups were recovered: one in the north-central Chihuahuan desert and the other largely south and west of the Rio Grande on the eastern edge of the Chihuahuan desert.

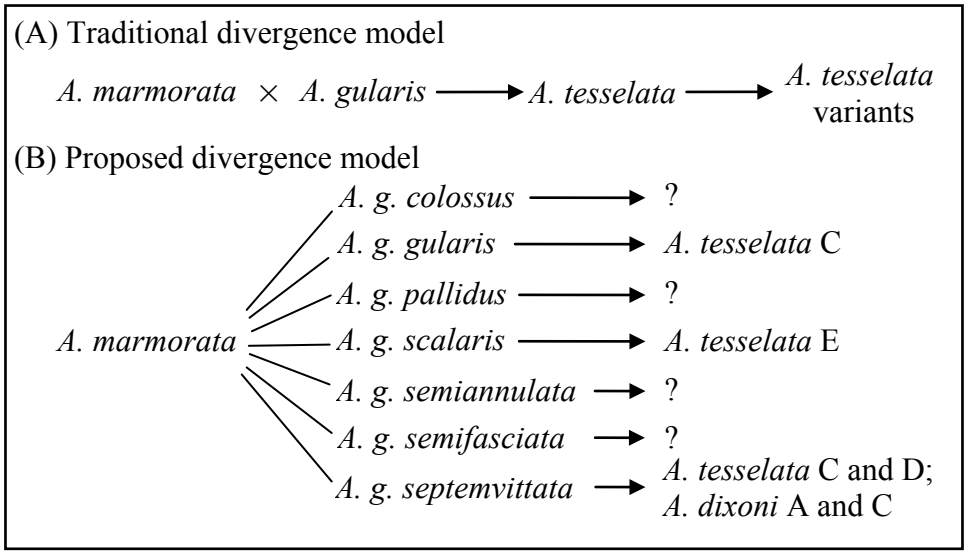
(A) Traditional divergence model
A. marmorata × A. gularis
A. tesselata
A. tesselata variants
(B) Proposed divergence model
A. marmorata
A. g. colossus
?
A. g. gularis
A. tesselata C
A. g. pallidus
?
A. g. scalaris
A. tesselata E
A. g. semiannulata
?
A. g. semifasciata
?
A. g. septemvittata
A. tesselata C and D;
A. dixoni A and C

Figure 2.1

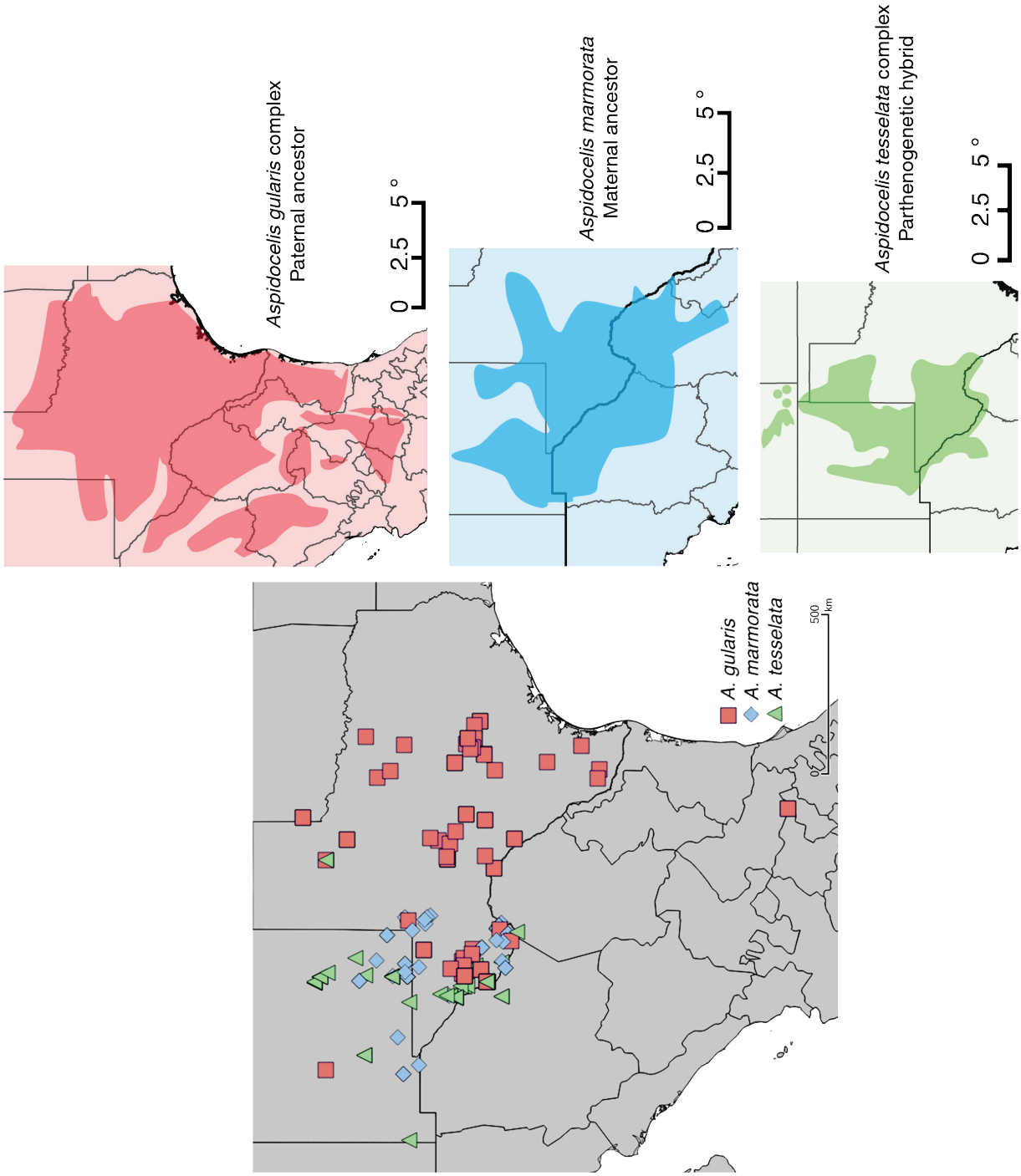
Aspidocelis gularis complex
Paternal ancestor
0 2.5 5 °
Aspidocelis marmorata
Maternal ancestor
0 2.5 5 °
Aspidocelis tesselata complex
Parthenogenetic hybrid
0 2.5 5 °
A. gularis
A. marmorata
A. tesselata
500
0
km

Figure 2.2

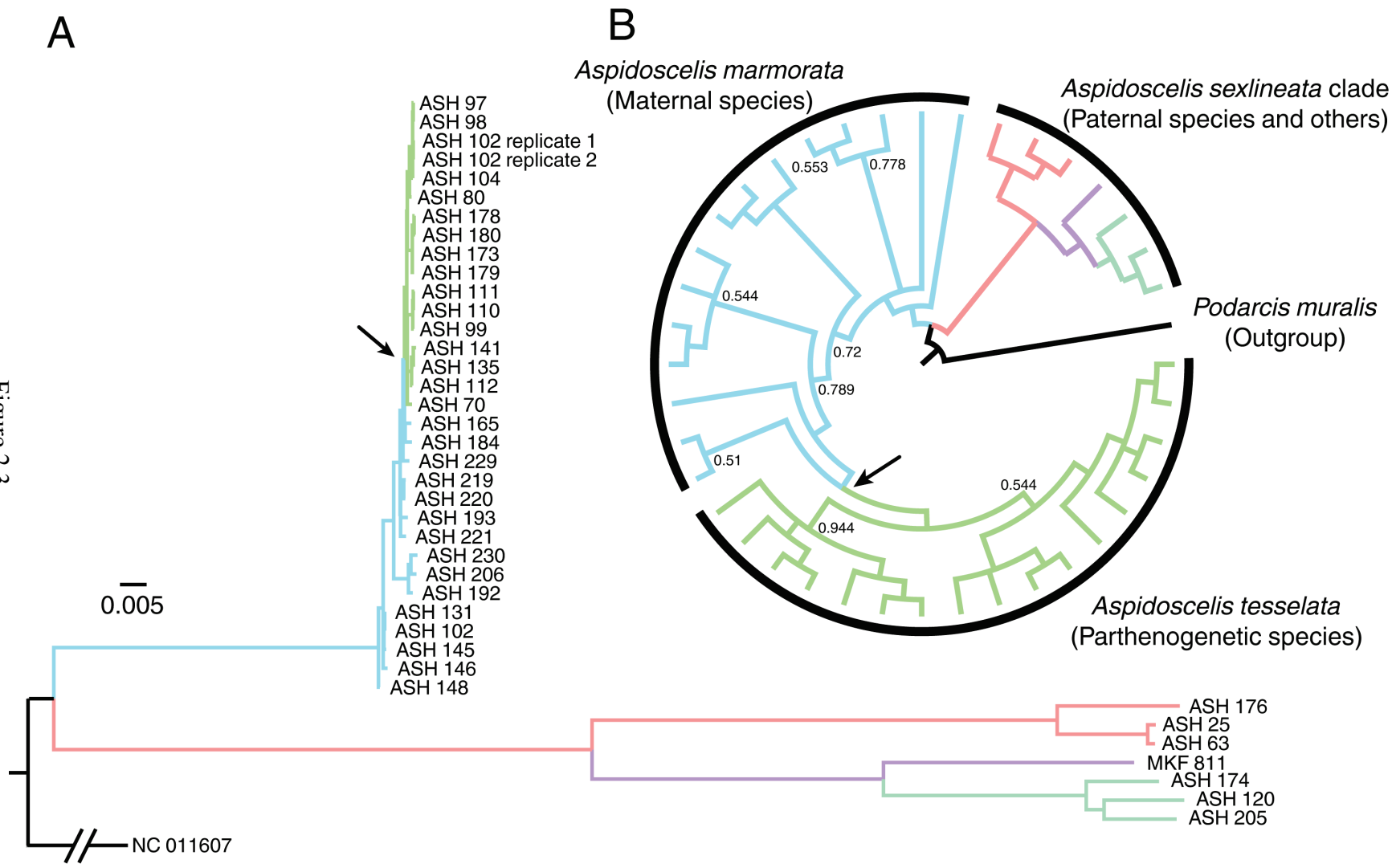
A
B
Aspidoscelis marmorata
(Maternal species)
Aspidoscelis sexlineata clade
(Paternal species and others)
Podarcis muralis
(Outgroup)
Aspidoscelis tesselata
(Parthenogenetic species)
0.553
0.778
0.544
0.72
0.789
0.51
0.544
0.944
ASH 97
ASH 98
ASH 102 replicate 1
ASH 102 replicate 2
ASH 104
ASH 80
ASH 178
ASH 180
ASH 173
ASH 179
ASH 111
ASH 110
ASH 99
ASH 141
ASH 135
ASH 112
ASH 70
ASH 165
ASH 184
ASH 229
ASH 219
ASH 220
ASH 193
ASH 221
ASH 230
ASH 206
ASH 192
ASH 131
ASH 102
ASH 145
ASH 146
ASH 148
0.005
ASH 176
ASH 25
ASH 63
MKF 811
ASH 174
ASH 120
ASH 205
NC 011607

Figure 2.3

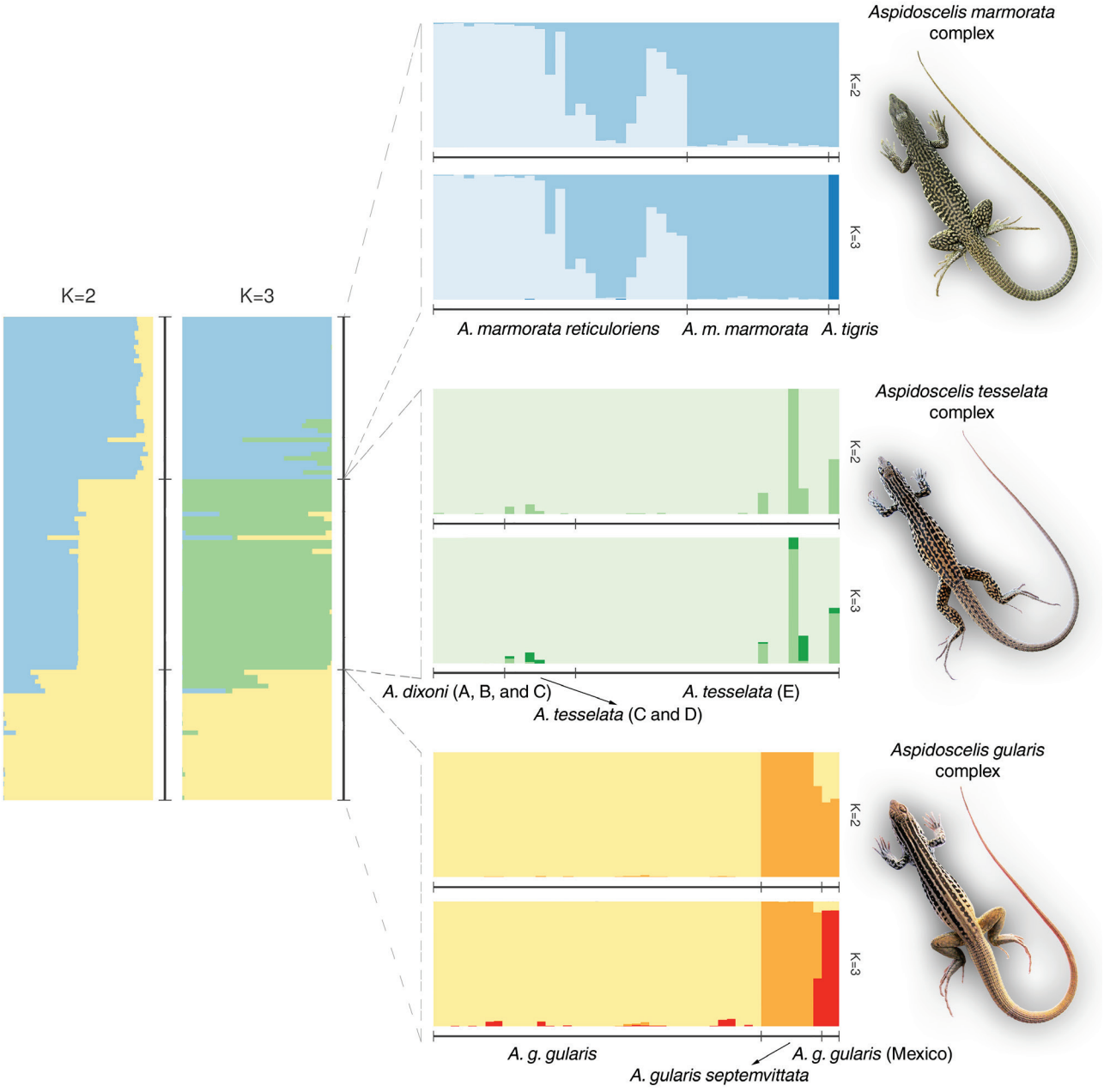
K=2
K=3
Aspidoscelis marmorata
complex
K=2
K=3
A. marmorata reticuloriens
A. m. marmorata
A. tigris
Aspidoscelis tesselata
complex
K=2
K=3
A. dixoni (A, B, and C)
A. tesselata (E)
A. tesselata (C and D)
Aspidoscelis gularis
complex
K=2
K=3
A. g. gularis
A. g. gularis (Mexico)
A. gularis septemvittata

Figure 2.2

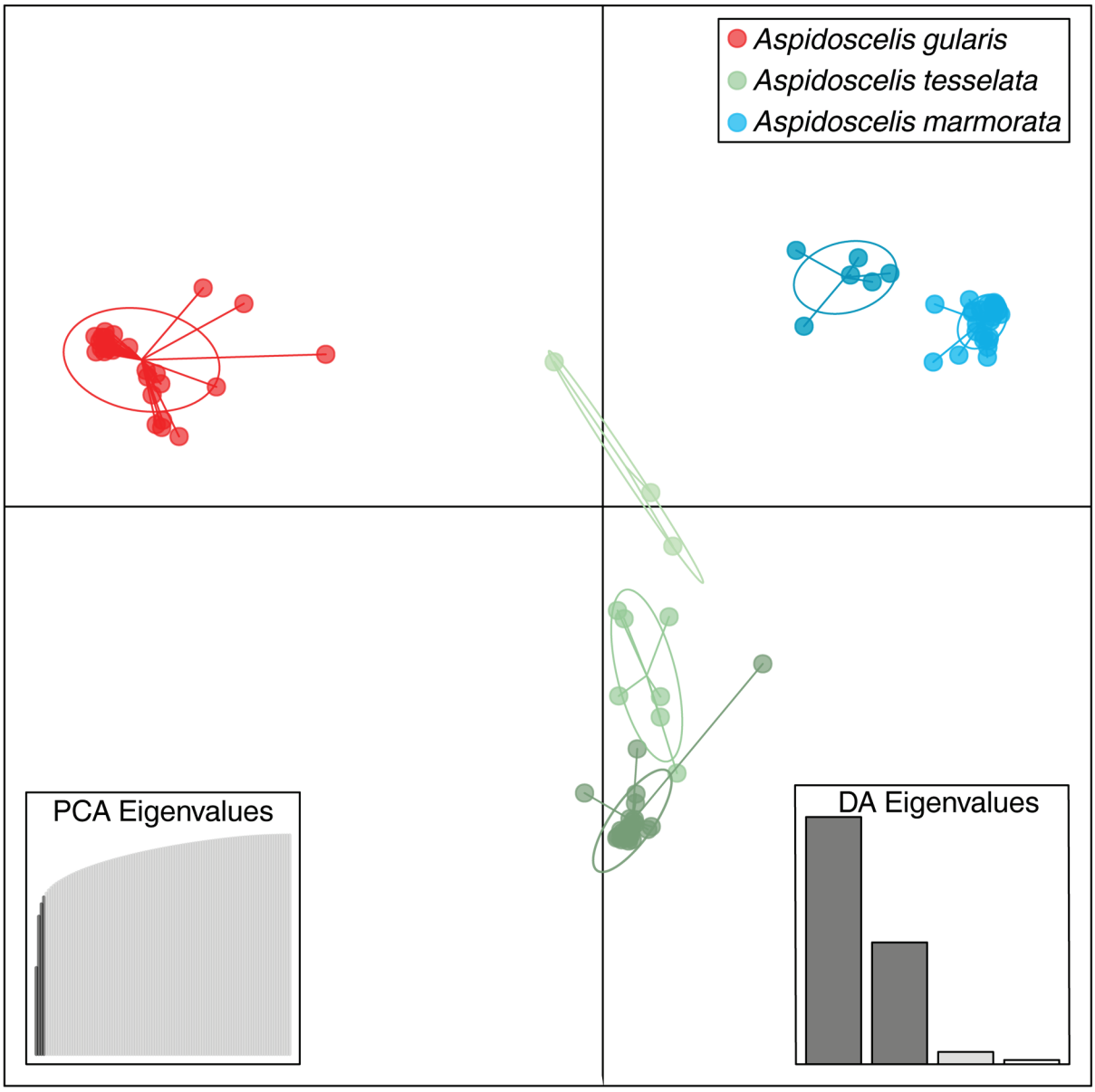
Aspidoscelis gularis
Aspidoscelis tesselata
Aspidoscelis marmorata
PCA Eigenvalues
DA Eigenvalues

Figure 2.3

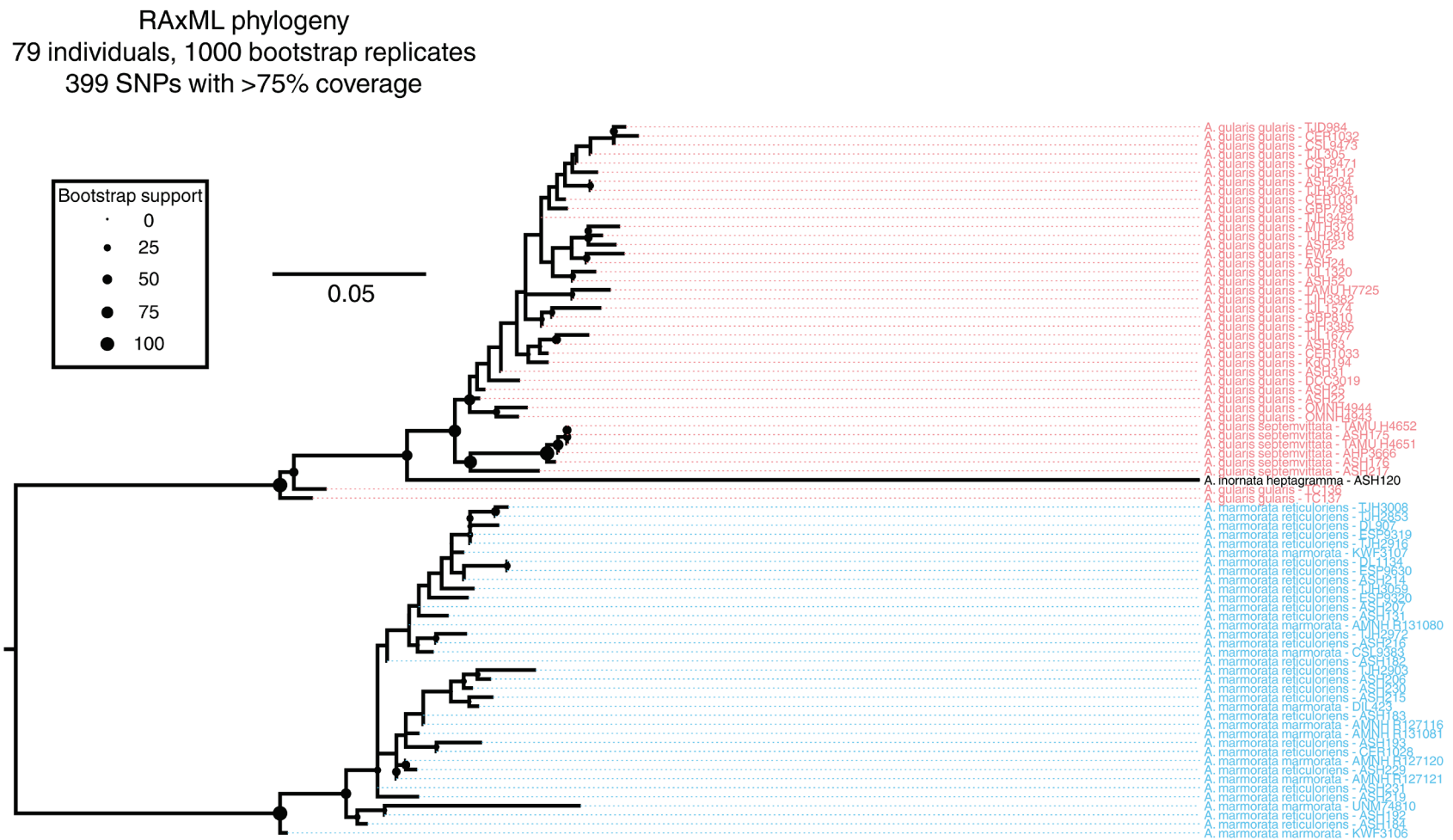
RAxML phylogeny
79 individuals, 1000 bootstrap replicates
399 SNPs with >75% coverage
Bootstrap support
0
25
50
75
100
0.05
A. gularis gularis - TJD984
A. gularis gularis - CER1032
A. gularis gularis - CSL9473
A. gularis gularis - TJD305
A. gularis gularis - CSL9471
A. gularis gularis - TJH2112
A. gularis gularis - ASH234
A. gularis gularis - TJH3055
A. gularis gularis - CER1031
A. gularis gularis - GBP789
A. gularis gularis - TJH3454
A. gularis gularis - MTJ370
A. gularis gularis - TJH2818
A. gularis gularis - ASH23
A. gularis gularis - EW2
A. gularis gularis - ASH24
A. gularis gularis - TJL1320
A. gularis gularis - ASH22
A. gularis gularis - TAMUH7725
A. gularis gularis - TJH3382
A. gularis gularis - TJL1574
A. gularis gularis - GBP810
A. gularis gularis - TJH3385
A. gularis gularis - TJL1677
A. gularis gularis - ASH63
A. gularis gularis - CER1033
A. gularis gularis - KQC194
A. gularis gularis - ASH31
A. gularis gularis - DCC3019
A. gularis gularis - ASH25
A. gularis gularis - ASH53
A. gularis gularis - OMNH4944
A. gularis gularis - OMNH4943
A. gularis septemvittata - TAMUH4652
A. gularis septemvittata - ASH175
A. gularis septemvittata - TAMUH4651
A. gularis septemvittata - AHP3666
A. gularis septemvittata - ASH179
A. gularis septemvittata - ASH177
A. inornata heptagramma - ASH120
A. gularis gularis - TCL137
A. gularis gularis - TCL97
A. marmorata reticuloriens - TJH3008
A. marmorata reticuloriens - TJH2853
A. marmorata reticuloriens - DL907
A. marmorata reticuloriens - ESP9319
A. marmorata reticuloriens - TJH2916
A. marmorata marmorata - KWF3107
A. marmorata reticuloriens - DL1134
A. marmorata reticuloriens - ESP9630
A. marmorata reticuloriens - ASH214
A. marmorata reticuloriens - TJH3059
A. marmorata reticuloriens - ESP9320
A. marmorata reticuloriens - ASH207
A. marmorata reticuloriens - ASH131
A. marmorata marmorata - AMNH R131080
A. marmorata reticuloriens - TJH2972
A. marmorata reticuloriens - ASH216
A. marmorata marmorata - CSL9383
A. marmorata reticuloriens - ASH182
A. marmorata reticuloriens - TJH2903
A. marmorata reticuloriens - ASH206
A. marmorata reticuloriens - ASH230
A. marmorata reticuloriens - ASH215
A. marmorata marmorata - DL423
A. marmorata reticuloriens - ASH183
A. marmorata marmorata - AMNH R127116
A. marmorata marmorata - AMNH R131081
A. marmorata reticuloriens - ASH193
A. marmorata reticuloriens - CER1028
A. marmorata marmorata - AMNH R127120
A. marmorata reticuloriens - ASH229
A. marmorata marmorata - AMNH R127121
A. marmorata reticuloriens - ASH231
A. marmorata reticuloriens - ASH219
A. marmorata marmorata - UNM7810
A. marmorata reticuloriens - ASH192
A. marmorata reticuloriens - ASH184
A. marmorata marmorata - KWF3106

Figure 2.6

Aspidoscelis gularis
Aspidoscelis tesselata
Aspidoscelis marmorata
Effective Migration Rates: m
Effective Diversity Rates: q
A
B
C
D
E
F
log(m)
2
1
0
−1
−2
log(q)
0.10
0.05
0.00
−0.05

Figure 2.7

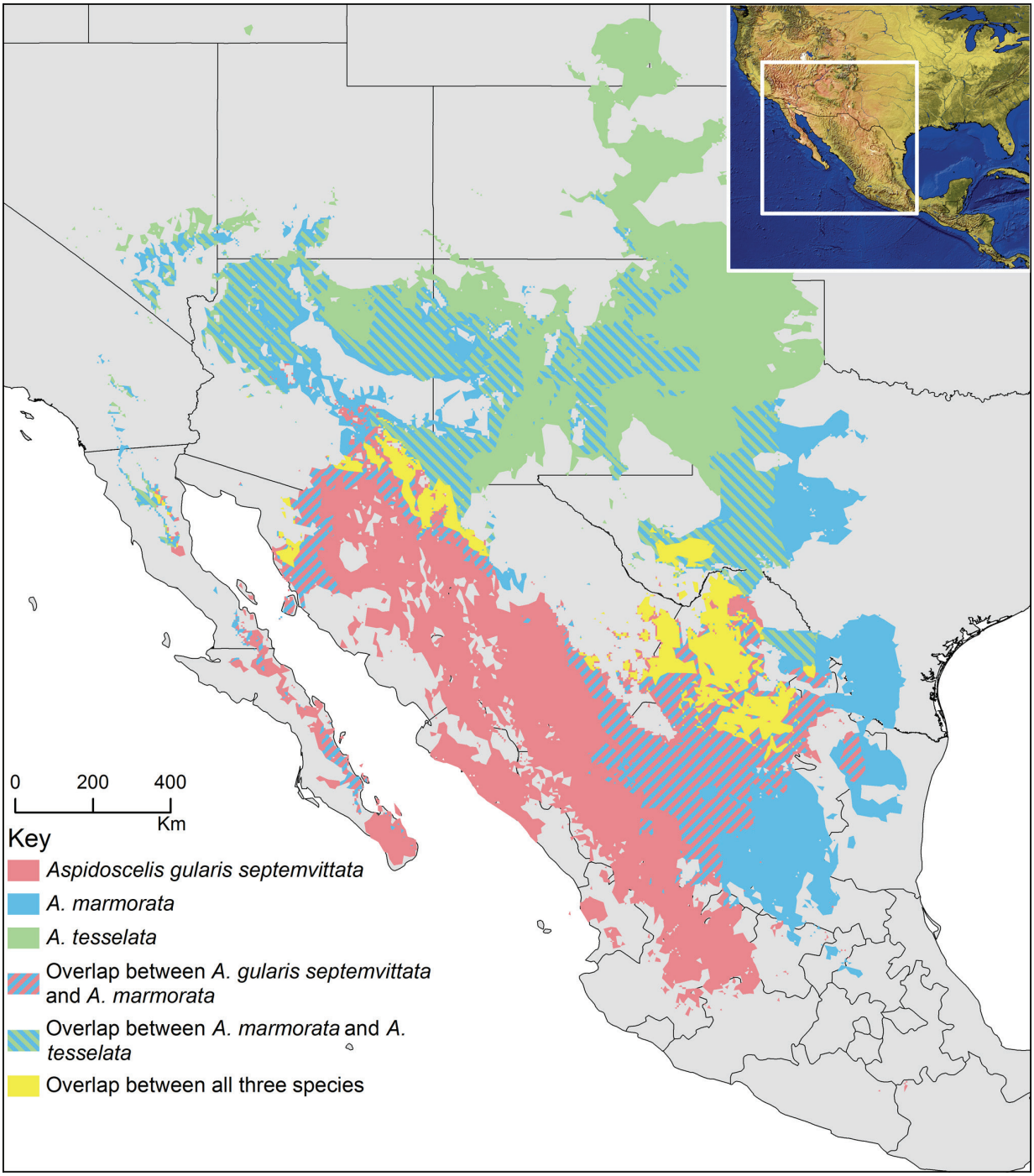
0
200
400
Km
Key
Aspidoscelis gularis septemvittata
A. marmorata
A. tesselata
Overlap between A. gularis septemvittata and A. marmorata
Overlap between A. marmorata and A. tesselata
Overlap between all three species

Figure 2.8

CHAPTER 3

DISCOVERY OF A NEW POPULATION OF *ASPIDOSCELIS DIXONI*, A VARIANT OF HYBRID AND PARTHENOGENETIC *ASPIDOSCELIS TESSELATA*

DISCOVERY OF A NEW POPULATION OF *ASPIDOSCELIS DIXONI*, A VARIANT OF HYBRID AND PARTHENOGENETIC *ASPIDOSCELIS TESSELATA*[1]

Alexander S. Hall, James E. Cordes, James M. Walker, and Matthew K. Fujita

[1]This article is in preparation for submission to a peer reviewed journal.

Abstract

Taxonomic assignments for whiptail lizards (*Aspidoscelis*; formerly *Cnemidophorus*) have previously been based on morphology, skin histocompatibility, karyotypes, and mitochondrial gene trees. However, hybridizations involving several whiptail species have produced parthenogenetic lineages and their derivatives that raise the question: what constitutes a parthenogenetic species? For example, hybridization of *A. marmorata* x *A. gularis septemvittata*, in either southern Texas or northern Mexico, produced the first individual of parthenogenetic *A. tesselata* which had one complete set of chromosomes from each parent. Nevertheless, many generations later *A. tesselata* now exhibits extensive postformational phenotypic variation in the form of distinctive pattern classes. One of these, pattern class F, was described as *A. dixoni* in 1973; recently shown to be histocompatible with *A. tesselata.* Thus, *A. dixoni* was derived from *A. tesselata* after its hybrid origin. Today, allopatric groups of *A. dixoni* occur in Hidalgo County, New Mexico, and Presidio County, Texas, where restrictions on collection have been imposed primarily because of small geographic areas of habitation. However, the most recent SSAR checklist treats *A. dixoni* as a mere synonym of *A. tesselata* (SSAR 2012). In this study, we sought to genetically characterize *A. dixoni* and *A. tesselata* in comparison to their maternal ancestor, *A. marmorata*. We sequenced entire mitochondrial genomes from over 50 animals that revealed low population structure in *A. marmorata*, one maternal origin of *A. tesselata*, and that *A. dixoni* mitochondria are derived from *A. tesselata*. This work also reports a previously unreported group *of A. dixoni* in northern Texas.

Introduction

Unisexual reproduction in vertebrates occurs nonrandomly and true all-female parthenogenesis only occurs in squamates (Kearney et al. 2009; Neaves and Baumann 2011). Existing parthenogenetic lizard species exhibit an array of color patterns, different skin histocompatibilities, and several mitochondrial haplotypes (Morán 2007). The phenotypic diversity seen in the most-studied parthenogenetic lizard, *Aspidoscelis tesselata* (formerly *Cnemidophorus* [Reeder et al. 2002]) led Zweifel (1965) to assign populations to six pattern classes A–F (Fig. 1.1). Further hybridization between diploid *A. tesselata* and *A. sexlineata* explains two of Zweifel's (1965) pattern classes A and B, now known to be triploid *A. neotesselata* (Fig. 3.1; Walker et al. 1997; Walker et al. 2012). The remaining four pattern classes C–F arose through hybridization between female *A. marmorata* and male *A. gularis* (Figs. 1.1, 2.1, 3.1). One pattern class, F, was described as *A. dixoni* by Scudday (1973), though the species is broadly recognized as a variant of *A. tesselata* (Cordes 1991; Walker et al. 1994; Cordes and Walker 2006). In this paper, we examine this lineage and describe a new population of *A. dixoni*.

Generally, previous investigations assert that all *A. tesselata* arose from one F_1 zygote (e.g., Maslin 1967; Cordes and Walker 2003, 2006; Taylor et al. 2003). In Chapter 2, we reported that *A. tesselata* may have arisen from at least two hybridization events (Figs. 2.1, 3.1). This may have involved one maternal population of *A. marmorata marmorata* having since gone extinct, thus omitting it from mitochondrial gene trees and study, including our own (Brown and Wright 1979; Wright et al. 1983; Densmore et al. 1989a, 1989b; Moritz et al. 1989; Chapter 2, this document). Nevertheless, postformational divergence best explains the diversity seen in most *A. tesselata*. This conclusion is reinforced by skin histocompatibility measured from

reciprocal skin transplants (Maslin 1967) which have been used in *Aspidoscelis* as a proxy for kinship (Cuellar 1977; Cordes and Walker 2006). Cordes and Walker (2006) found that reciprocal skin transplants were accepted between *Aspidoscelis tesselata* E and *A. dixoni* A–C. This occurs despite the substantial geographic disparity between the clades (Fig. 2.2).

It is currently unknown if *A. dixoni* as a phenotype has arisen multiple times from a background of *A. tesselata* (most likely pattern class E), or if it has evolved only once. The single origin hypothesis would seem highly unlikely given the hundreds of kilometers of fairly well-sampled area (at least north of the US border) between the combined habitat of *A. dixoni* A and B in the Chinati Mountains of west Texas and *A. dixoni* C in Animas Pass in the Peloncillo Mountains in southwestern New Mexico. What's even more surprising is our surprise discovery of a new population of *A. dixoni* (Fig. 3.3A, B). This new population occurs in the Texas panhandle in Palo Duro Canyon State Park. Very little is known about this population, as only one individual has thus been captured alive. Preliminary evidence from a skin histocompatibility test with a syntopic *A. tesselata* C (Fig. 3.3C) verify that this new *A. dixoni* population shares histocompatibility with *A. tesselata* pattern class C. At this locality, *A. gularis gularis* and *A. tesselata* C are abundant, but neither *A. marmorata* nor *A. g. septemvittata* are found. As these species are generally accepted to be the ancestral stock for diploid *A. tesselata* (Figs. 1.2, 2.1, 3.1; Neaves and Gerald 1968; Lowe and Wright 1966; Wright and Lowe 1967), a parsimonious explanation would be that this new *A. dixoni*, should it be recently evolved, is a postformational variant of either *A. tesselata* C or E.

In this study, we analyzed the genetic history of *A. dixoni* and *A. tesselata*. We use whole mitochondrial genomes and nuclear genomic SNP evidence for the first time in these lineages. We hypothesized that *A. dixoni* will nest within *A. tesselata* in phylogenetic analyses. This would

imply that *A. tesselata* is paraphyletic without *A. dixoni* and thus the taxonomy of this group bears revisiting.

Materials and Methods

Largely speaking, we followed a subset of the methods presented in Chapter 2 of this work. We added mitochondrial sequence data for 15 animals in addition to those presented in Figure 2.3 and Appendix A. Briefly, assembled and aligned mitochondrial genomes using a reference *Aspidoscelis gularis* mitochondrial genome. We then partitioned these data by codon and functional element as input for a MrBayes analysis (Ronquist et al. 2012). In MrBayes, we set up four runs with two chains each to run for 100,000,000 generations and sampled every 1,000 generations. The first 25% of these data were discarded as burn-in. Stationarity was confirmed in Tracer (ESS >200). For nuclear data, we collected SNPs from ddRADseq data using STACKS 1.37 as discussed in Chapter 2 (Catchen et al. 2011, 2013). We analyzed biallelic SNP data using the programs STRUCTURE (Pritchard et al. 2000) and discrete analysis of principle components (DAPC; Jombart et al. 2010) as implemented in *adegenet* (Jombart and Ahmed 2011). As we focused on *A. tesselata* and *A. dixoni* in this analysis, we used an r cutoff equal to 80 in *populations* (i.e., a SNP had to be present in ≥80% of the samples to be included). Additionally, we allowed a 50% missing data tolerance per animal. This resulted in a dataset consisting of 665 nuclear SNPs across 40 individuals (corresponds with dataset 4 in Chapter 2).

Results

Mitochondrial Origins of Aspidoscelis dixoni

In combination with the previously reported dataset (Chapter 2) we sequenced, assembled, and annotated 53 *Aspidoscelis* mitochondrial genomes from 300 bp paired-end reads with average coverage exceeding 100X per sample (Fig. 3.4). Assembly qualities were comparable

from the 37 mitochondrial genomes presented in Chapter 2. The MrBayes analysis recovered a mitochondrial phylogeny that agrees with previous findings by Reeder et al. (2002) and our work in Chapter 2. Additionally, the two sequence *A. dixoni* A specimens were nested within *A. tesselata* and are thus younger than the origin(s) of *A. tesselata*.

ddRADseq and Population Structure of Aspidoscelis dixoni

Since *Aspidoscelis dixoni* was recovered within *A. tesselata* based on mitochondrial data (Fig. 3.4), but is found in distant and restricted localities (Fig. 3.2), we also utilized ddRADseq data to investigate potential population structure. In a subset of the dataset presented in Chapter 2 (Fig. 2.4), we analyzed structure within *A. tesselata* and *A. dixoni* using 665 nuclear SNPs across 40 individuals using STRUCTURE. Based on the Evanno method (Evanno et al. 2005), this analysis suggested a K=2 was the most informative. See the results of Chapter 2 for a discussion of the two individuals assigned to a second population. Notably, no structure was inferred in these data between *A. dixoni* and *A. tesselata* or within *A. dixoni* pattern classes. Additionally, as previously reported in Chapter 2 (Fig. 2.5), DAPC from 1191 SNPs explained 77% of the variance in a data set that contained 104 individuals from *A. tesselata*, *A. dixoni*, and both parental species complexes (*A. marmorata* and *A. gularis*). All *A. dixoni* in our analysis were recovered in the conglomerate *A. tesselata* group (in green at the bottom of Fig. 2.5), with one exceptional *A. dixoni* A. This individual was recovered in the second-largest *A. tesselata* cluster which contained 5 *A. tesselata* E and one *A. tesselata* C.

Discussion

Post-formational clones of unisexual vertebrates are generally not favored for taxonomic consideration, though there is precedent for doing so (see Table 3 in Cordes and Walker 2006). Many parthenogenetic *Aspidoscelis* species were described before their clonal nature became

known to science (e.g., Lowe and Zweifel 1952; Lowe 1956). In tandem with further hybridization, ploidy elevation, and a 'surprising' degree of phenotypic variation within the clade, the hybrid parthenogenetic species complex *A. tesselata* has been a taxonomist's nightmare. Adding to this complexity, the pattern class described as *A. tesselata* F by Zweifel (1965) was later reclassified to *A. dixoni* by Scudday (1973). This maneuver has been critiqued in oblique and direct ways since its publication. As the authors and more recent articles have admitted (Cordes and Walker 2006; SSAR 2012), *A. dixoni* is most likely a postformational modification of *A. tesselata*. Nevertheless, at the time of writing the species *A. dixoni* is largely respected within herpetological nomenclature.

In this study, we provide for the first time a preliminary report of a newly discovered population of *A. dixoni* found in Palo Duro Canyon State Park (Figs. 3.2, 3.3). Similar to the nearly 500 km in between the *A. dixoni* A/B and *A. dixoni* C populations (Cordes 1991), the new population in north Texas is 600 km from the Chinati Mountains (*A. dixoni* A/B) and 740 km Antelope Pass (*A. dixoni* C). At least one major river and at mountain range lies between each direct route between the new population and either established population, making the discovery quite unexpected. Although we were not able to present new genetic evidence to confirm the placement of this new population within extant *A. dixoni*, we are convinced that it is based on phenotype and unpublished skin histocompatibility data by JEC. Previously, Cordes and Walker (2006) reported that reciprocal skin transplants between *A. dixoni* A, B, and C were compatible with *A. tesselata* E. Previous to the current study, this provides the strongest available evidence that *A. tesselata* E has a close evolutionary affinity with whiptail populations currently under the *A. dixoni* moniker, but these authors refrained from demoting *A. dixoni* to *A. tesselata* (as have others).

Agreeing with previous results, we found that *A. dixoni* A arises from within *A. tesselata* based on a full mitochondrial genome phylogeny (Fig. 3.4). Of course, since mitochondria are maternally inherited, *A. tesselata* was recovered from within *A. marmorata marmorata* (Fig. 3.4). We also failed to recover structure within *A. dixoni* pattern classes or between *A. dixoni* and *A. tesselata* C–E (Figs. 2.5, 3.5). In short, this implies that *A. tesselata*, as troubled a taxon as it is, is paraphyletic without the inclusion of *A. dixoni*. Given the disjunct distribution of *A. dixoni*, and its skin histocompatibility with *A. tesselata* E (Cordes and Walker 2006) - but skin histocompatibility between the new population and *A. tesselata* C – the most parsimonious explanation for *A. dixoni* populations is that they are separate postformational varieties of diploid *A. tesselata* E and possibly C.

It could be argued (as it was initially) that *A. neotesselata*, the triploid offspring of *A. tesselata* and *A. sexlineata viridis* (Fig. 3.1) should be considered within *A. tesselata*. However, the current state of the field in considering wild hybrids experiencing ploidy elevation is to consider these lineages to be different species. Essentially, it is argued, their genetic contribution has rather permanently departed from its diploid ancestors. Following this logic, Walker et al. (1997) formally elevated triploid *A. tesselata* A and B to *A. neotesselata* and assigned a neotype for *A. tesselata*, as the original holotype for the species was most likely a triploid animal (i.e., *A. neotesselata*).

As this dissertation is not peer-reviewed, we refrain from making a formal taxonomic recommendation. Nevertheless, although *A. dixoni* has been recognized as a species (Wright 1993, Walker et al. 1994), this was considered admissible because *A. dixoni* was considered to be the offspring of a separate hybridization event than *A. tesselata*. Evidence presented by Cordes and Walker (2006) and here refute this hypothesis. The findings in Cordes and Walker (2006) led

the most recent Society of the Study of Reptiles species listing (SSAR 2012) to include *A. dixoni* as a synonym for *A. tesselata* following Maslin and Secoy (1986). Until our results are formally published, we leave the taxonomic status of *A. dixoni* and *A. tesselata* unchanged.

Acknowledgements

We thank Chad Montgomery in assisting JEC with discovering the new population of *A. dixoni*. David Riskind ensured that JEC could collect within Texas state parks which allowed investigating the new population to be possible in the first place. See the acknowledgements from Chapter 2 regarding past work including grants and permits.

Literature Cited

Brown, W. M., and J. W. Wright. 1979. Mitochondrial DNA analyses and the origin and relative age of parthenogenetic lizards (genus *Cnemidophorus*). *Science*, 203:1247–1249.

Catchen, J., A. Amores, P. Hohenlohe, W. Cresko, and J. H. Postlethwait. 2011. *Stacks*: building and genotyping loci *de novo* from short-read sequences. *G3: Genes, Genomes, Genetics*, 1:171–182.

Catchen, J., P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. 2013. Stacks: an analysis tool set for population genomics. *Molecular Ecology*, 22:3124–3140.

Cordes, J. E. 1991. Biology of two sympatric clonal complexes of parthenogenetic whiptail lizards [TES-G (2n) and TES-H (2n)] in the *Cnemidophorus tesselatus* complex and identification of hybrid TES-G (2n) x *Cnemidophorus gularis septemvittatus* in the Chinati Mountains in trans-Pecos Texas. Thesis (Ph.D.). University of Arkansas.

Cordes, J. E., and J. M. Walker. 2003. Skin histocompatibility between syntopic pattern classes C and D of parthenogenetic *Cnemidophorus tesselatus* in New Mexico. *Journal of Herpetology*, 37:185–188.

Cordes, J. E., and J. M. Walker. 2006. Evolutionary and systematic implications of skin histocompatibility among parthenogenetic Teiid lizards: three color pattern classes of *Aspidoscelis dixoni* and one of *Aspidoscelis tesselata*. *Copeia*, 2006:14–26.

Densmore, L. D., III, C. C. Moritz, J. W. Wright, and W. M. Brown. 1989a. Mitochondrial-DNA analyses and the origin and relative age of parthenogenetic lizards (genus *Cnemidophorus*). IV. Nine *sexlineatus*-group unisexuals. *Evolution*, 43:969–983.

Densmore, L. D., III, J. W. Wright, and W. M. Brown. 1989b. Mitochondrial-DNA analyses and the origin and relative age of parthenogenetic lizards (genus *Cnemidophorus*). II. *C. neomexicanus* and the *C. tesselatus* complex. *Evolution*, 43:943–957.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14:2611–2620.

Kearney, M., M. K. Fujita, and J. Ridenour. 2009. Lost sex in the reptiles: constraints and correlations. Pp 447–474 *in* Schön, I., K. Martens, and P. van Dijk, eds. Lost Sex. Springer, London, U.K.

Lowe, C. H., Jr. 1956. A new species and a new subspecies of whiptailed lizards (genus *Cnemidophorus*) of the inland Southwest. *Bulletin of the Chicago Academy of Sciences*, 10:137–150.

Lowe, C. H., and J. W. Wright. 1966. Chromosomes and karyotypes of cnemidophorine Teiid lizards. *Mammalian Chromosomes Newsletter*, 22:199–200.

Lowe, C. H., Jr., and R. G. Zweifel. 1952. A new species of whiptailed lizard (genus *Cnemidophorus*) from New Mexico. *Bulletin of the Chicago Academy of Sciences*, 9:229–247.

Maslin, T. P., and D. M. Secoy. 2012. A checklist of the lizard genus *Cnemidophorus* (Teiidae). *Contributions in Zoology*, 1:1–60.

Jombart, T., S. Devillard, and F. Balloux. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured population. *BMC Genetics*, 11:94.

Jombart, T., and I. Ahmed. 2011. *adegenet 1.3-1*: new tools for the analysis of genome-wide SNP data. *Bioinformatics* 27:3070–3071.

Maslin, T. P. 1967. Skin grafting in the bisexual Teiid lizard *Cnemidophorus sexlineatus* and in the unisexual *C. tesselatus*. *Journal of Experimental Zoolology*, 166:137–149.

Morán, N. M. 2007. Diversidad clonal en los lacertilios unisexuales del género *Aspidoscelis*. *Boletín de la Sociedad Herpetológica Mexicana*, 15:1–12.

Moritz, C. C., J. W. Wright, and W. M. Brown. 1989. Mitochondrial-DNA analyses and the origin and relative age of parthenogenetic lizards (genus *Cnemidophorus*). III. *C. velox* and *C. exsanguis*. *Evolution*, 43:958–968.

Neaves, W. B., and P. S. Gerald. 1968. Lactate dehydrogenase isozymes in parthenogenetic Teiid lizards (*Cnemidophorus*). *Science*, 160:1004–1005.

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

Parker, E. D., and R. K. Selander. 1976. The organization of genetic diversity in the parthenogenetic lizard *Cnemidophorus tesselatus*. *Genetics*, 84:791–805.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics*, 155:945–959.

Reeder, T. W., C. J. Cole, and H. C. Dessauer. 2002. Phylogenetic relationships of whiptail lizards of the genus *Cnemidophorus* (Squamata: Teiidae). *American Museum Novitates*, 3365:1–61.

Ronquist, F., M. Teslenko, P. van der Mark, D. L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M. A. Suchard, and J. P. Huelsenbeck. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61:539–542.

Scudday, J. F. 1973. A new species of lizard of the *Cnemidophorus tesselatus* group from Texas. *Journal of Herpetology*, 7:363–371.

Society for the Study of Amphibians and Reptiles (SSAR). 2012. Committee on Standard English and Scientific Names and B. I. Crother (committee chair). *Scientific and standard English names of amphibians and reptiles of North America north of Mexico, with comments regarding confidence in our understanding*, Seventh Edition. Herpetological Circular 39: Society for the Study of Amphibians and Reptiles.

Taylor, H. L., C. J. Cole, H. C. Dessauer, and E. D. Parker, Jr. 2003. Congruent patterns of genetic and morphological variation in the parthenogenetic lizard *Aspidoscelis tesselata* (Squamata: Teiidae) and the origins of color pattern classes and genotypic clones in eastern New Mexico. *American Museum Novitates*, 3424:1–40.

Walker, J. M., J. E. Cordes, C. C. Cohn, H. L. Taylor, R. V. Meyer, and L. Richard. 1994. Life history characteristics of three morphotypes in the parthenogenetic *Cnemidophorus dixoni* complex (Sauria: Teiidae) in Texas and New Mexico. *Texas Journal of Science*, 46:27–33.

Walker, J. M., J. E. Cordes, and H. L. Taylor. 1997. Parthenogenetic *Cnemidophorus tesselatus* complex (Sauria: Teiidae): a neotype for diploid *C. tesselatus* (Say, 1823), redescription of the taxon, and description of a new triploid species. *Herpetologica*, 53:233–259.

Walker, J. M., H. L. Taylor, G. J. Manning, J. E. Cordes, C. E. Montgomery, L. J. Livo, S. Keefer, and C. Loeffler. 2012. Michelle's lizard: identity, relationships, and ecological status of an array of parthenogenetic lizards (genus *Aspidoscelis*: Squamata: Teiidae) in Colorado, USA. *Herpetological Conservation and Biology*, 7:227–248.

Wright, J. W. 1993. Evolution of the lizards of the genus *Cnemidophorus*. *In* J. W. Wright and L. J. Vitt (eds.). *Biology of Whiptail Lizards [Genus* Cnemidophorus*]*: pp. 27–81, Oklahoma Museum of Natural History, Norman.

Wright, J. W., and C. H. Lowe. 1967. Evolution of the alloploid parthenospecies *Cnemidophorus tesselatus* (Say). *Mammalian Chromosome Newsletter*, 8:95–96.

Wright, J. W., C. Spolsky, and W. M. Brown. 1983. The origin of the parthenogenetic lizard *Cnemidophorus laredoensis* inferred from mitochondrial DNA analysis. *Herpetologica*, 39:410–416.

Zweifel, R. G. 1965. Variation in and distribution of the unisexual lizard, *Cnemidophorus tesselatus*. *American Museum Novitates*, 2235:1–49.

Figure Legends

Figure 3.1. Current hypotheses regarding the origin and evolution of the *Aspidoscelis tesselata* complex. Dashed lines indicate hybridization events. The *A. m. marmorata* population marked with the dagger† is thought to be extinct or otherwise not sampled.

Figure 3.2. *Aspidoscelis dixoni* populations are small and geographically distant. The dark green shaded area is a range map of *Aspidoscelis tesselata* and *A. neotesselata*. The *A. dixoni* nov. population is described in part in this text, but will be formally described in a forthcoming publication.

Figure 3.3. A) and B) Photos of a specimen from the newly discovered *Aspidoscelis dixoni* population from Palo Duro Canyon (UADZ 9409). C) Photo of a syntopic *A. tesselata* (pattern class C; UADZ 9408). Photos by James E. Cordes.

Figure 3.4. *Aspidoscelis dixoni* evolved from within *A. tesselata*. This phylogeny was recovered from Bayesian analysis of 53 Teiid mitochondrial genomes partitioned by gene and codon. Branches are colored by species: *A. gularis* is red, *A. marmorata* is blue, *A. tesselata* is green, and *A. dixoni* is pink. Nodal posterior probabilities below 0.95 are shown with connecting dotted lines. The scale bar represents substitutions per site. The *A. gularis* branch is truncated for aesthetics.

Figure 3.5. Population structure in *Aspidoscelis tesselata* and *A. dixoni*. Inferences were derived from 665 nuclear SNPs and 40 individuals.

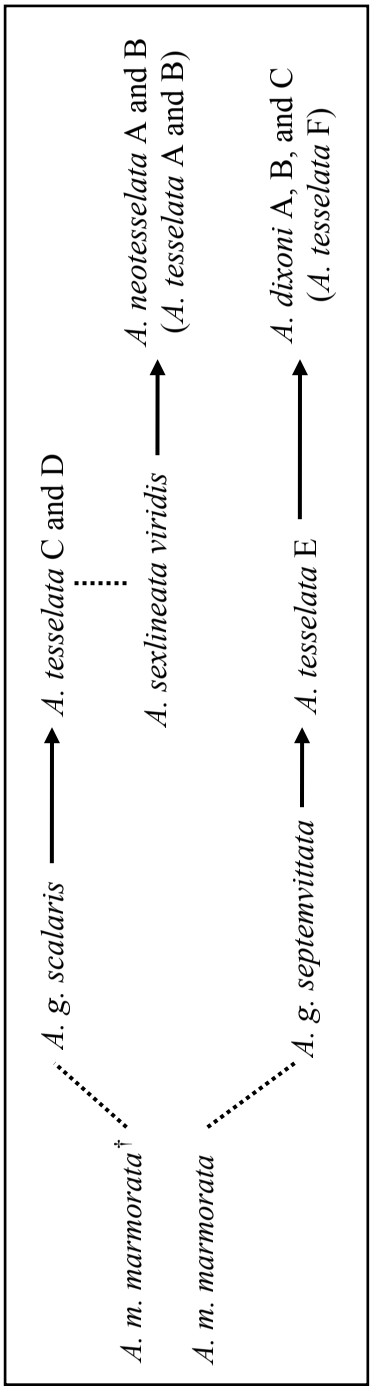
A. g. scalaris
A. tesselata C and D
A. m. marmorata†
A. sexlineata viridis
A. neotesselata A and B
(A. tesselata A and B)
A. m. marmorata
A. g. septemvittata
A. tesselata E
A. dixoni A, B, and C
(A. tesselata F)

Figure 3.1

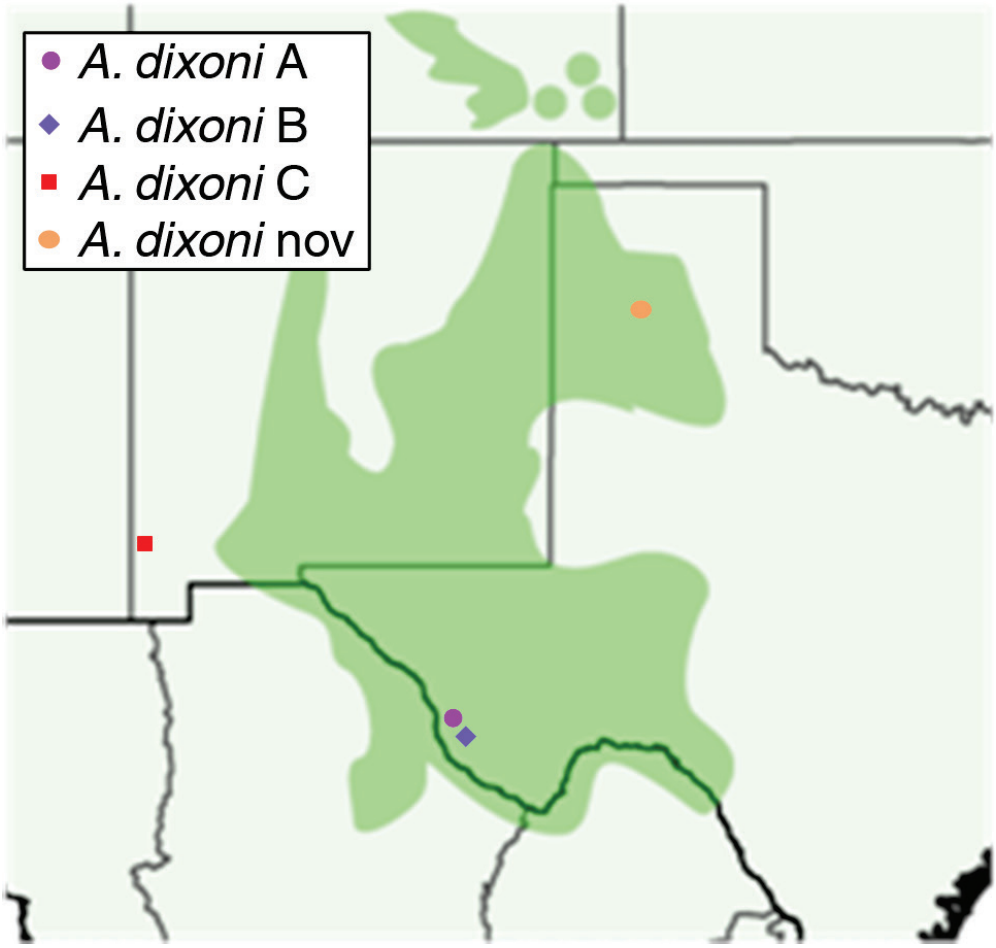

Figure 3.2

Figure 3.3

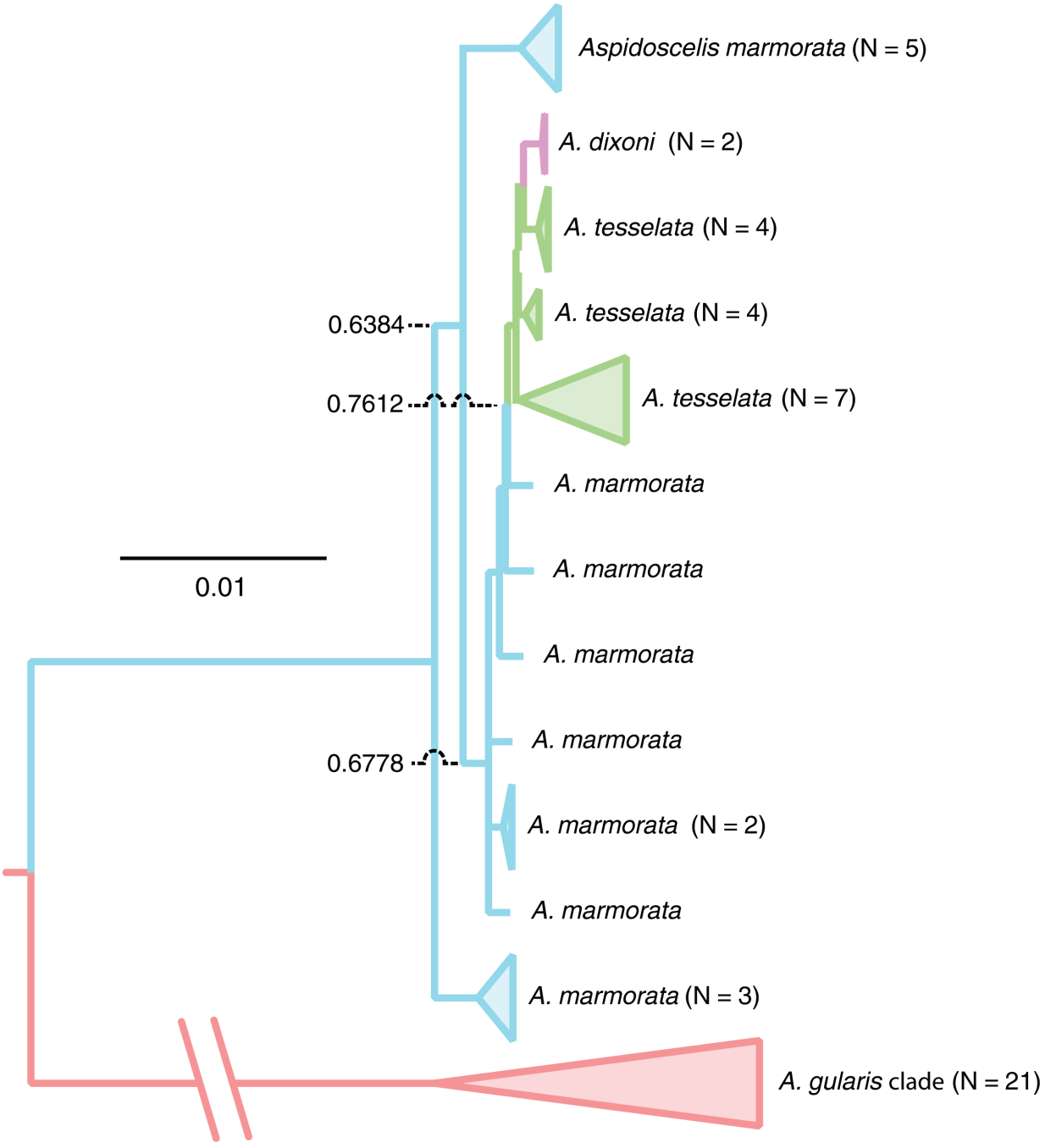
Aspidoscelis marmorata (N = 5)
A. dixoni (N = 2)
A. tesselata (N = 4)
A. tesselata (N = 4)
0.6384
0.7612
A. tesselata (N = 7)
A. marmorata
A. marmorata
0.01
A. marmorata
A. marmorata
0.6778
A. marmorata (N = 2)
A. marmorata
A. marmorata (N = 3)
A. gularis clade (N = 21)

Figure 3.4

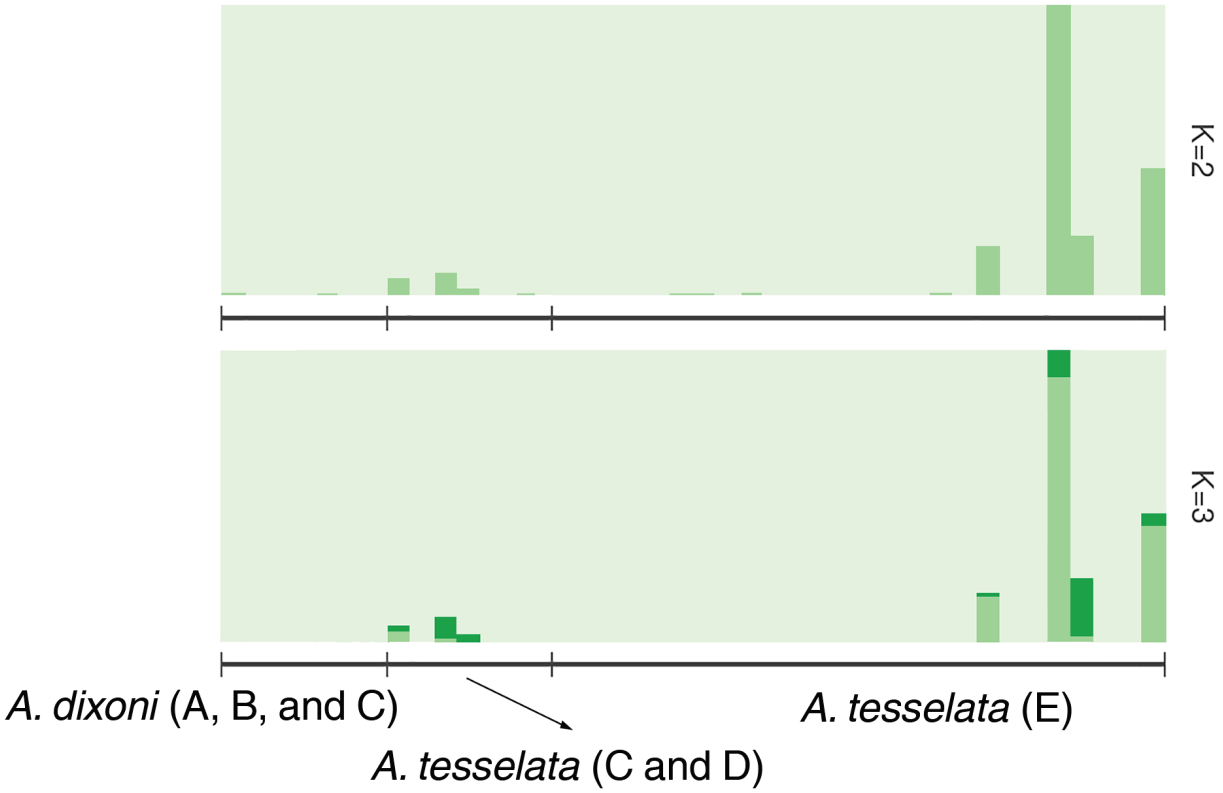

Figure 3.5

CHAPTER 4

ISOCHORE DIVERSITY ACROSS NEARLY 300 VERTEBRATE GENOMES

ISOCHORE DIVERSITY ACROSS NEARLY 300 VERTEBRATE GENOMES[1]

Alexander S. Hall, Peter E. Baumann, and Matthew K. Fujita

[1]This article is in preparation for submission to a peer reviewed journal.

Abstract

Vertebrate genome evolution has been studied most extensively in a small number of species. From these species, mostly mammalian, the foundational work was laid for broad scale patterns of genome content evolution. Importantly, isochores were discovered in these genomes: long (often >300 kbp) stretches of the genome with substantially similar GC content. From decades of study in the human and bovine genomes, five isochoric families have been identified and correlated with histone binding affinity, gene density, and recombination rates. But not all vertebrate genomes possess all or even more than one or two isochore families. In this study, we investigated broad scale genomic heterogeneity in four groups: fish, mammals, reptiles, and birds. Agreeing with previous work, these groups exhibit quite different GC content landscapes. The reasons for this are currently unknown and would benefit from a phylogenetically aware analysis. We also investigate isochore diversity in a smaller subset of genomes including the new *Aspidoscelis marmorata* genome assembly. This work confirmed recently contested results from the *Anolis carolinensis* genome that suggested some lizards lack isochore family diversity.

Introduction

An ultimate goal of biology is to describe the diversity of life. Nonetheless, descriptions of diversity and studies of evolution in non-model organisms can be challenging due to their relative obscurity. With the universal language of nucleotides, comparative evolutionary study is made more accessible than relying on phenotype, and thus we are currently experiencing a surge in the use of non-model organisms for evolutionary study. Initiatives such as the 10K genome project and the avian genomics project are currently leading the way to provide genomic resources for vertebrate biologists. Despite this interest in discovery, model organisms and those of economic importance historically drive the field forward in a way that can hinder insight into how other organisms evolve.

For example, DNA derived from calf thymus has been used as a standard of eukaryotic DNA for nearly 65 years (e.g., Meselson et al. 1957). This DNA behaves consistently in digestion (Chargaff 1951) and ultracentrifugation (Meselson et al. 1957) experiments, but this also biased the literature for how the genomic architecture of eukaryotic genomes evolves. In the presence of sequence-specific ligands (such as Ag^{+}), ultracentrifugation of DNA in Cs_2SO_4 density gradients revealed that warm-blooded vertebrates contained patterns of compositional heterogeneity (Filipski et al. 1973; Macaya et al. 1976; Thiery et al. 1976). It was revealed that eukaryotic genomes consist, in large part, of stereotyped stretches (> 300 kb) of compositionally homogeneous nucleotide content. These stereotyped patterns were described as families of what are now called isochores (reviewed in Bernardi et al. 1985; Bernardi 1995, 2004, 2015). Indeed, many vertebrates exhibit the five isochoric families: L1, L2, H1, H2, and H3 (Table 4.1; Costantini et al. 2009), but it was noticed quite early that endothermic and ectothermic vertebrates exhibited – in aggregate – different patterns of isochoric family diversity (Bernardi

and Bernardi 1986; Bernardi 2004; Costantini et al. 2009). Some mammalian genomes have a full complement of isochore family diversity, but not all, such as the opossum (*Monodelphis domestica*) and platypus (*Ornithorhynchus anatinus*; Costantini et al. 2009).

The apparent discrepancy between relative isochore abundance across vertebrate genomes has led to speculation that isochores, and by extension genomic GC content, provide some functional importance for the organism in the form of thermodynamic stability (Bernardi and Bernardi 1986; Bernardi 2004, 2015). Though the purpose of isochores in eukaryotes is not fully resolved, they are imputed to be correlated with gene rich "core" portions of otherwise desert-like genomes loaded with repetitive content and pseudo-functionalized genetic remnants (reviewed in Bernardi 2015). These arguments are usually based on empirical data, and those data are biased towards model eukaryote systems: humans, mice, cows, dogs, chimpanzees, chicken, zebrafish, stickleback, *Xenopus* frogs, etc.

In this paper, we aim to elucidate the GC diversity across all publically available sequenced vertebrate genomes. We also focus on a newly sequenced lizard genome, the Teiid lizard *Aspidoscelis marmorata*, as it is an emerging model of vertebrate parthenogenesis and offers further contrast to the existing pool of well-characterized vertebrate genomes. We hypothesize that heterogeneity in GC content differs across four clades of vertebrates: fish, mammals, reptiles, and birds. We compare our results to previous work with a brief discussion of how these new results should affect broad interpretations of genome evolution.

Materials and Methods

GC Percentage Data Collection

Primarily, data were downloaded from GenBank in FASTA format (Fig. 4.1). Vertebrate genome data from GenBank included 64 fishes, 118 mammals, 15 reptiles, and 67 birds

(Appendix B). Additionally, we included a new species: the Teiid lizard *Aspidoscelis marmorata*. We excluded the Lissamphibia, as too few (N = 4) genomes were currently available for meaningful statistical comparisons within this clade. Generally speaking, the most up-to-date genomes were used for all analyses. In the case of multiple assemblies per species, we analyzed the genome with the largest contig N50 size. We ignored this rule if the largest N50 belonged to a WGS-only assembly when an assembly with scaffolds was available. Genomes labeled as partial were not considered. For easier downstream scripting, files were de-interleaved using a custom python script.

A custom python script assessed each file's GC content separately (Fig. 4.1) using the following window sizes (bp): 1000, 3000, 5000, 20000, 80000, and 320000. The script accepted non-overlapping window size as an integer, the species name, and class or group name. After checking that the input was valid and the supplied file was FASTA-formatted, the program read contigs into memory one at a time. This contig was then split into sub-contigs starting with the first base using the provided window size. If, after all splitting is finished, a contig or sub-contig was shorter than the window size; the names of the contig and sub-contig were added to a list of too-short DNA fragments. Sub-contigs passing this check were checked for the number of bases, including ambiguity codes. All hyphens, question marks, and N's were treated as missing data. All values of A, T, and W were treated as A or T. All values of C, G, and S were treated as G or C. All other ambiguity codes (M, R, Y, K, V, H, D, and B) were treated as ambiguous data. In calculating GC content, the script first checked if the amount of ambiguous and missing data of the contig was greater than or equal to 20%. If this value was met or exceeded, the name of the contig and sub-contig were added to a second output file for contigs with too much missing data.

All other sub-contigs were analyzed for GC content: the ratio of GC data (excluding gaps and ambiguous data) to the length of the sub-contig. These data were added to a third output file.

These data were summarized in two ways. First, a script summarized the results of a genome's GC content analysis at the provided window size. This produces a fourth output file that includes descriptive statistics: percent of sub-contigs recovered, mean GC content, GC content standard deviation, GC content range, median GC content, number of contigs, and contextual information about how the script was run. Helper scripts were then used to find all of these summary files and extract the data into a comma-separated spreadsheet of each analysis. This summary of summaries could be used as input for comparing summary statistics across taxa. The second summary method combined all of the GC content analyses done by sub-contig at a given window size. These summary files could be used to illustrate the distribution of GC content across genomes and between taxa.

GC Percentage Summary Data Analyses

Genomic standard deviation in GC content is used as a proxy for nucleic acid heterogeneity in sequenced genomes (Costantini et al. 2006, Fujita et al. 2011, Castoe et al. 2013). Thus, we compared standard deviations in GC content collected from our summary dataset. After removing outliers (Tukey 1977) and zeroes, we used Kruskal-Wallis rank sum tests to assess the presence of measurable differences in GC content standard deviation in each window size analysis across the four clades: fish, mammals, reptiles, and birds. We used Kruskal-Wallis tests since Shapiro-Wilk tests for normality (Royston 1982) indicated significant deviations from normality ($P < 0.005$) in all but the 80000 bp window analysis (where $P = 0.1333$). Pairwise comparisons for significant Kruskal-Wallis tests were analyzed using Dunn tests (Dunn 1964). With four groups and up to six window sizes to test, this results in up to 30

pairwise comparisons. To control for multiple testing while simultaneously circumventing issues with excessively conservative reductions in family-wise alpha (e.g., Bonferroni; see Nakagawa 2004), we implemented the Benjamini-Hochberg adjustment (Benjamini and Hochberg 1995). This step-down procedure controls the false discovery rate in a manner that more logically balances the chances of Type I (false positive) and Type II (false negative) errors than simpler and more conservative Bonferroni or Holm corrections (Benjamini and Hochberg 1995; Nakagavwa 2004; Dinno 2016).

GC Percentage across Genomes Analysis

To visually compare nucleic acid heterogeneity across the genomes of 265 vertebrates, each with different assembly sizes, we combined all GC percentage data analyzed by subcontig by window size and clade. This technique is useful for broad-level insight into the relative abundance of GC content throughout an entire genome (Pačes et al. 2004; Fujita et al. 2011). A tall peak would indicate a more homogeneous distribution of nucleotide content compared to a shorter, flatter peak. We also compared distributions split by species, rather than averaged across clade. In both of these comparisons, we overlaid the *Aspidoscelis marmorata* genome.

Isochore Data Analyses

Another approach to representing genomic heterogeneity is to is to consider isochore diversity by the amount of DNA assigned to bins of GC content at a given window size (e.g., Macaya et al. 1976, Costantini 2006). To do this, we analyzed the *Homo sapiens* (GRCh38.p7), *Anolis carolinensis* (AnoCar2.0), and *Aspidoscelis marmorata* genomes in IsoPlotter 2.4 (Elhaik et al. 2013). This program assesses GC content variability within a window size using a dynamic algorithm that considers allowing contiguous windows of similar GC content to be part of the same domain (Elhaik 2009; Elhaik et al. 2010). Domains are returned homogeneous (i.e.,

isochoric) or heterogeneous. Using default analysis options on each genome, we summed the length of heterogeneous or homogeneous DNA that fit within one-percent bins of mean GC content. We compared the distributions between the three taxa using a Kruskal-Wallis rank sum test followed by Dunn tests for each pairwise comparison.

We also calculated the number of domains attributed to each of five isochore families (Table 4.1) using the logic presented by Costantini et al. (2006, 2009) and references therein. Since genome size varied substantially between these three taxa, we also compared the proportion of isochores by family in each taxon. IsoPlotter allows you to exclude N runs (gaps) from analyses or include them. When we compared results, there was such substantial overlap we only report results where N runs were included. We performed the above analyses on all domains and separately on only those domains listed as homogeneous by IsoPlotter.

Results

As measured by standard deviation in genomic GC content, genomic heterogeneity varied between fish, mammals, birds, and reptiles (Fig. 4.1, Table 4.2). When corrected for multiple comparisons, 27 of 36 pairwise rank comparisons yielded significant differences at six window sizes and between the four clades (Fig. 4.1, Table 4.3). Fish generally had the most homogeneous genomic GC content and mammals the most heterogeneous. Birds and reptiles were somewhat intermediate in genomic heterogeneity. In half of the window sizes, bird genomes were on average more heterogeneous than reptiles, but the two groups substantially overlap. The pairwise differences between mammals, birds, and reptiles at the largest window sizes mostly did not differ (Fig. 4.1, Table 4.3). This was likely due to the homogenizing factor of increasing analysis window size (see Fujita et al. 2011; Castoe et al. 2013), but also because more data were omitted as window size was increased (Fig. 4.5). This is important because, before analyzing our data,

we removed instances when standard deviation equaled zero (points at the bottom of Fig. 4.1), and this occurred when all or nearly all the data were removed.

Relative GC content throughout vertebrate genomes varies by chosen window size (Figs. 4.3, 4.4). Generally, fish had the most homogeneous genomes as evidenced by tall, narrow peaks around a single mean value. This pattern is broken for 80000 and 320000 bp window sizes where a bimodal distribution emerges for fish and reptiles can be seen to have more homogeneous genomes. Nevertheless, with few exceptions, individual fish genomes exhibit the most "peaky" relative GC content, regardless of window size (Fig. 4.4). The spread of mean GC content increases for fish at 20000, 80000, and 320000 bp windows. This likely drives the bimodal pattern seen in the aggregate data (Fig. 4.3). A small number of genomes, mostly mammals, exhibit binomial relative GC content with a stereotyped pattern of a large peak centered around the genome's mean GC content and a much smaller, sometimes tighter, peak at a much higher GC content (Fig. 4.4). Interestingly, this pattern was reported in coarser-grained CsCl gradient data, mostly in mammals (Macaya et al. 1976).

In reporting the actual isochores, we found Aspidoscelis marmorata to have a somewhat intermediate spread of isochore family diversity between the isochore-homogeneous *Anolis carolinensis* and heterogeneous *Homo sapiens* (Figs. 4.6, 4.7). Our results for the relative family abundance of isochores in the human genome (Fig. 4.6B) vary slightly from previous work (Table 1 in Costantini et al. 2009). These authors reported a greater relative abundance of L2 and H1 isochores. We uncovered fairly equal amounts of L1, L2, H1, and H2 isochores (Fig. 4.6B). Note that Costantini et al. (2009) used hg17 and a default 100 kb window size for identifying and classifying isochores The distributions of GC content binned by one-percent windows differed between the three taxa (Fig. 4.7; $\chi^2 = 1616.5077$, df = 2, $P < 2.2 \times 10^{-16}$). Subsequent pairwise

comparisons all significantly differed (Table 4.4). Excluding N runs, mean GC content in *Aspidoscelis marmorata* was 42.666%. This is taken in comparison to *Homo sapiens* (40.9%; Costantini et al. 2009) and *Anolis carolinensis* (40.8238%; Aföldi et al. 2011).

Discussion

Genome Sampling and Window Size

Bias Isochore Research

Previous work on eukaryotic genome architecture has identified genomic heterogeneity (and by extension, isochores) are correlated with the arrangement and evolution of gene-encoding regions (Bernardi 2015). This argument hinges in large part on observations from only a small sampling of eukaryotic genomes (Costantini et al. 2009). In this study, we expanded past work by over an order of magnitude; thus, we were able to reconsider genomic architecture with less bias towards model organisms.

Overall, we found that GC heterogeneity varies between four vertebrate clades (Fig. 4.2; Table 4.2). The precision of these differences is influenced by the window size used to compute summary statistics (Table 4.3). In general, among vertebrate genomes analyzed, fish had the most homogeneous genomes and mammals the most heterogeneous. Bird and reptile genomes were intermediate between fish and mammals. The bird and reptile genomes studied exhibited somewhat equivalent GC heterogeneity in half of tested cases (Table 4.3). This challenges the previous view that homeothermic (i.e., "warm-blooded") animals should have the most features associated with genome-level adaptations to higher body temperatures (Bernardi 2007). This thermodynamic stability hypothesis (Bernardi and Bernardi 1986; Jabbari and Bernardi 2004) is based on the gene-centric (Bernardi and Bernardi 1986) and neoselectionist (Bernardi 2007) view that selection acts not only on the phenotype of the individual, but also at the genetic level.

In this scenario, genes that most faithfully copy are favored under selective pressures. This hypothesis largely lacks the nuance of a phylogenetic framework which would take phylogenetic correlation into consideration. Instead, a cladistic mindset dominates the field of isochore theory (and genomic architecture when presented by these same authors). A phylogenetically aware view of the pattern in reptiles and birds found in the present study (see also Fujita et al. 2011) may offer the best null model for isochore evolution and the evolution of genomic complexity.

Our conclusions pertaining to relative GC distribution and GC heterogeneity were strongly influenced by window size used for analysis (Figs. 4.2–4.5). We selected the six window sizes used in this study to cover the range utilized in most isochore identification studies (3–300 kb; Fujita et al. 2011; Costantini 2009). Of course, for larger window sizes less well assembled genomes yielded less data for analysis due to failing the < 20% missing data or minimum sub-contig size requirements imposed by our analytical approach (Figs. 4.1, 4.5). We do not interpret this result as a negative finding, as our choice of window size is largely arbitrary. Instead, we caution relying on a single window size for analyses, especially in comparative work. The patterns seen at one window size may not be recoverable at a different window size, even with the same data. 100 kb moving windows are used by Costantini et al. (2006, 2009, 2016) when draw_chromosome_gc.pl (Pačes et al. 2004) is invoked to (presumably) objectively identify GC content and isochoric regions of the genome. Our results draw into question the repeatability of their findings when other window sizes are considered. The 100 kb window size was initially selected by Costantini et al. (2006) since this is a value that appears to be a plateau in GC standard deviation as window size increases in the human genome. Importantly, this window size currently biases the literature towards a result that is sensible in the human genome, but may make little sense in other eukaryotic genomes (Fujita et al. 2011; Fig. 4.2).

The Aspidoscelis *Genome Exhibits Intermediate Isochore Patterns to* Anolis *and* Homo

In addition to our investigation of the GC content and heterogeneity in several eukaryotic genomes, we took a more focused view of a new genome assembly. The haploid *Aspidoscelis marmorata* genome consists of 1.63 Gb, and of that, 842.8 Mb were identified as isochoric/homogeneous by IsoPlotter (Fig. 4.7A). Unlike the *Anolis carolinensis* genome (Fig. 4.7B), the *Aspidoscelis marmorata* genome does exhibit isochoric diversity (Fig. 4.7A). We found that, proportional to the number of isochores in the *A. marmorata* genome, L2, H1, and H2 isochore families comprise about 80% of the genome (Fig. 4.6B). This does not, however, consider the length of these isochores (as in Fig. 4.7). Previously, Fujita et al. (2011) reported that the then-only lizard genome largely lacked compositionally heterogeneous isochores. In the present study, we recovered the same pattern using a different program (Fig. 4.7B).

We also recovered the characteristic wide spread of GC content in the human genome with a long tail extending towards high GC content (Fig. 4.7C; Zoubak et al. 1996; Bernardi 2001a, b). However, we found difficulty recovering the more clearly delineated peaks in GC content associated with isochore families reported by Costantini et al. (2006, 2009) and reviewed by Bernardi (2015) for hg17. Although we used an updated human genome assembly, the main differences likely occurred because of the different methodologies of extending isochores between the program we used (IsoPlotter; Elhaik and Graur 2013) and those used by Costantini et al. (draw_chromosome_gc.pl; Pačes et al. 2004). Indeed, an update by Cozzi et al. (2015) used hg38 and recovered the same distinguishable peaks in GC content. In the present study, we used IsoPlotter due to its ability to dynamically determine whether or not to extend isochoric domains. In contrast, the method described by Pačes et al. (2004) relies on a window size provided to the

program *a priori*. As previously discussed, it is our conclusion that window size choice matters in the identification of isochores and genomic heterogeneity. By allowing the program to essentially make the isochore expansion decision we side-step the window-size issue. In this way, it is unsurprising to find differences in the details between our results and previous findings. Indeed, Cozzi et al. (2015) compared extant isochore discovery methods and noted differences between these programs. Their solution was to create an automated program, isoSegmenter, but still this technique uses a 100 kb window size with the same justification (i.e., Costantini et al. 2006).

In conclusion, we extended previous efforts to compare genomic architecture across eukaryotic genomes by over an order of magnitude. In doing so, we confirm that vertebrate clades exhibit different GC heterogeneity, but not always in a way predicted by thermodynamic stability hypothesis (Bernardi and Bernardi 1986; Bernardi 2004). We would also like to emphasize that window size *does* matter when assessing comparative patterns of genomic architecture, at least for nucleotide composition. As more genomes are sequenced, and of higher quality, phylogenetically informed views of isochore evolution will become increasingly powerful and should offer more realistic predictive power.

Data Availability

The most up-to-date scripts used in our analyses can be downloaded from GitHub: https://github.com/allopatry/isochores. Appendix B lists genomes downloaded from GenBank. Supplementary data should be made available through a data repository in a future publication. Until then, contact ASH with any data requests.

Acknowledgements

We would like to thank Duncan Tormey, Morgan Schroeder, and Rutendo Sigauke for their irreplaceable assistance in assembling and annotating the first *Aspidoscelis* genome. Duncan, in particular, brought many ideas forward about the analysis of isochores in *A. marmorata*. This work was supported, in part, by the Stowers Institute for Medical Research and a UT Arlington Office of Graduate Studies Dissertation Fellowship to ASH. PEB is an Early Career Scientist with the Howard Hughes Medical Institute.

Tables

Table 4.1. Definitions of isochore families by mean GC percent. Final column is the expectation from *Homo sapiens* according to Costantini et al. (2006).

Isochore Family	Lower Bound	Upper Bound	Percent GC Range	Relative Abundance
L1	≥ 30	≤ 37	8	Moderate
L2	≥ 37	≤ 41	5	High
H1	≥ 41	≤ 46	6	Moderate
H2	≥ 46	≤ 53	8	Low
H3	≥ 53	≤ 65	13	Very Low

Table 4.2. Kruskal-Wallis rank sum tests of GC standard deviation between clades: fish, mammals, reptiles, and birds. Asterisks indicate statistical significance with alpha = 0.05.

Window Size	χ^2	*df*	*P*
1000 bp	134.8768	3	$< 2.2 \times 10^{-16}$ *
3000 bp	164.2148	3	$< 2.2 \times 10^{-16}$ *
5000 bp	138.8253	3	$< 2.2 \times 10^{-16}$ *
20000 bp	113.8441	3	$< 2.2 \times 10^{-16}$ *
80000 bp	87.2376	3	$< 2.2 \times 10^{-16}$ *
320000 bp	74.8951	3	3.816×10^{-16} *

Table 4.3. Post-hoc Dunn test comparisons of genomic GC standard deviation in non-overlapping windows. Asterisks indicate statistical significance after Dunn test false discovery rates were corrected by the Benjamini and Hochberg (1995) adjustment for multiple comparisons.

Window Size		Fish		Mammals		Reptiles	
		z	*P*	*z*	*P*	*z*	*P*
1000 bp	Mammals	-9.177989	< 0.0001*	-	-	-	-
	Reptiles	-4.872381	< 0.0001*	0.639461	0.2613	-	-
	Birds	-11.07708	< 0.0001*	-3.315287	0.0007*	-2.574579	0.0060*
3000 bp	Mammals	-12.50793	< 0.0001*	-	-	-	-
	Reptiles	-2.232718	0.0128*	5.512479	< 0.0001*	-	-
	Birds	-6.869898	< 0.0001*	4.815147	< 0.0001*	-2.388626	0.0101*
5000 bp	Mammals	-11.07849	< 0.0001*	-	-	-	-
	Reptiles	-1.202888	0.1145	5.900486	< 0.0001*	-	-
	Birds	-4.225029	< 0.0001*	6.432921	< 0.0001*	-1.706016	0.0528
20000 bp	Mammals	-6.069892	< 0.0001*	-	-	-	-
	Reptiles	2.780831	0.0041*	6.986402	< 0.0001*	-	-
	Birds	2.527618	0.0069*	9.177011	< 0.0001*	-1.036436	0.1500
80000 bp	Mammals	-8.515756	< 0.0001*	-	-	-	-
	Reptiles	-6.916169	< 0.0001*	-1.809465	0.0422	-	-
	Birds	-6.145700	< 0.0001*	1.757114	0.394	2.768871	0.0042*
320000 bp	Mammals	-7.957674	< 0.0001*	-	-	-	-
	Reptiles	-5.099749	< 0.0001*	-0.316656	0.5636	-	-
	Birds	-7.186722	< 0.0001*	-0.004996	0.4980	0.299792	0.4586

Table 4.4. Post-hoc Dunn test comparisons of isochoric DNA content binned by 1% GC percentage. Asterisks indicates statistical significance with alpha = 0.05.

	Anolis		*Aspidoscelis*	
	z	*P*	*z*	*P*
Aspidoscelis	-29.78407	< 0.0001*	-	-
Homo	1.809094	< 0.0352*	39.77855	< 0.0001*

Literature Cited

Alföldi, J., F. Di Palma, M. Grabherr, C. Williams, L. Kong, E. Mauceli, P. Russell, C. B. Lowe, R. E. Glor, J. D. Jaffe, D. A. Ray, S. Boissinot, A.M. Shedlock, C. Botka, T. A. Castoe, J. K. Colbourne, M. K. Fujita, R. G. Moreno, B. F. ten Hallers, D. Haussler, A. Heger, D. Heiman, D. E. Janes, J. Johnson, P. J. de Jong, M. Y. Koriabine, M. Lara, P. A. Novick, C. L. Organ, S. E. Peach, S. Poe, D. D. Pollock, K. de Queiroz, T. Sanger, S. Searle, J. D. Smith, Z. Smith, R. Swofford, J. Turner-Maier, J. Wade, S. Young, A. Zadissa, S. V. Edwards, T. C. Glenn, C. J. Schneider, J. B. Losos, E. S. Lander, M. Breen, C. P. Ponting, and K. Lindblad-Toh. 2011. The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature*, 477:587–91.

Bernaola-Galván, P., P. Carpena, and J. L. Oliver. 2008. A standalone version of IsoFinder for the computational prediction of isochores in genome sequences. arXiv:0806.1292 [q-bio.GN]

Benjamini, Y., and Y. Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological*), 57:289–300.

Bernardi, G. 1995. The human genome: organization and evolutionary history. *Annual Review of Genetics*, 29:445–476.

Bernardi, G. 2001a. Isochores and the evolutionary genomics of vertebrates. *Gene*, 241:3–17.

Bernardi, G. 2001b. Misunderstandings about isochores. Part 1. *Gene*, 276:3–13.

Bernardi, G. 2004. *Structural and evolutionary genomics. Natural selection in genome evolution.* Elsevier, Amsterdam.

Bernardi, G. 2007. The neoselectionist theory of genome evolution. *Proceedings of the National Academy of Sciences*, 104:8385–8390.

Bernardi, G. 2015. Chromosome architecture and genome organization. *PLoS ONE*, 10:e0143739.

Bernardi, G. and G. Bernardi. 1986. Compositional constraints and genome evolution. *Journal of Molecular Evolution*, 24:1–11.

Castoe, T. A., A. P. J. de Koning, K. T. Hall, D. C. Card, D. R. Schield, M. K. Fujita, R. P. Ruggiero, J. F. Degner, J. M. Daza, W. Gu, J. Reyes-Velasco, K. J. Shaney, J. M. Castoe, S. E. Fox, A. W. Poole, D. Polanco, J. Dobry, M. W. Vandewege, Q. Li, R. K. Schott, A. Kapusta, P. Minx, C. Feschotte, P. Uetz, D. A. Ray, F. G. Hoffmann, R. Bogden, E. N. Smith, B. S. W. Chang, F. J. Vonk, N. R. Casewell, C. V. Henkel, M. K. Richardson, S. P. Mackessy, A. M. Bronikowski, M. Yandell, W. C. Warren, S. M. Secor, and D. D. Pollock. 2013. The Burmese python genome reveals the molecular basis for extreme adaptation in snakes. *Proceedings of the National Academy of Science*, 110:20645–20650.

Chargaff, E. 1951. Some recent studies on the composition and structure of nucleic acids. *Journal of Cellular Physiology*, 38:41–59.

Cosantini, M., O. Clay, F. Auletta, and G. Bernardi. 2006. An isochore map of human chromosomes. *Genome Research*, 16:536–541.

Costantini, M., G. Greif, F. Alvarez-Valin, and G. Bernardi. 2016. The *Anolis* lizard genome: an amniote genome without isochores? *Genome Biology and Evolution*, 8:1048–1055.

Costantini, M., R. Cammarano, and G. Bernardi. 2009. The evolution of isochore patterns in vertebrate genomes. *BMC Genomics*, 10:146.

Cozzi, P., L. Milanesi, and G. Bernardi. 2015. Segmenting the human genome into isochores. *Evolutionary Bioinformatics*, 11:253–261.

Dinno, A. 2016. dunn.test: Dunn's test of multiple comparisons using rank sums. R package version 1.3.2. http://CRAN.R-project.org/package=dunn.test

Dunn, O. J. 1964. Multiple comparisons using rank sums. *Technometrics*, 6:241–252.

Elhaik, E. 2009. The compositional organization of Mammalian genomes: characteristics and evolution. Thesis (Ph.D.). University of Houston.

Elhaik, E., D. Grauer, K. Josić, and G. Landan. 2010. Identifying compositionally homogeneous and nonhomogeneous domains within the human genome using a novel segmentation algorithm. *Nucleic Acids Research*, 38:e158.

Elhaik, E., and D. Graur. 2013. IsoPlotter^{+}: a tool for studying the compositional architecture of genomes. *ISRN Bioinformatics*, 2013.

Fearnhead, P., and D. Vasileiou. 2009. Bayesian analysis of isochores. *Journal of the American Statistical Association*, 104:132–141.

Filipski, J., J. P. Thiery, and G. Bernardi. 1973. An analysis of the bovine genome by Cs_2SO_4—Ag^{+} density gradient centrifugation. *Journal of Molecular Biology*, 80:177–197.

Fujita, M., S. V. Edwards, and C. P. Ponting. 2011. The *Anolis* lizard genome: an amniote genome without isochores. *Genome Biology and Evolution*, 3:971–984.

Jabbari, K., and G. Bernardi. 2004. Body temperature and evolutionary genomics of vertebrates: a lesson from the genomes of *Takifugu rubripes* and *Tetraodon nigroviridis*. *Gene*, 333:179–181.

Lercher, M. J., N. G. C. Smith, A. Eyre-Walker, and L. D. Hurst. 2002. The evolution of isochores: evidence from SNP frequency distributions. *Genetics*, 162:1805–1810.

Macaya, G., J.-P. Thiery, and G. Bernardi. 1976. An approach to the organization of Eukaryotic genomes at a macromolecular level. *Journal of Molecular Biology*, 108, 237–254.

Meselson, M., F. W. Stahl, and J. Vinograd. 1957. Equilibrium sedimentation of macromolecules in density gradients. *Proceedings of the National Academy of Sciences*, 47:581–588.

Nakagawa, S. 2004. A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioral Ecology*, 15:1044–1045.

Oliver, J. L., P. Carpena, M. Hackenberg, and P. Bernaola-Galván. 2004. IsoFinder: computational prediction of isochores in genome sequences. *Nucleic Acids Research*, 32: W287–W292.

Pačes, J., R. Zíka, V. Pačes, A. Pavlíček, O. Clay, and G. Bernardi. 2004. Representing GC variation along eukaryotic chromosomes. *Gene*, 333:135–141.

R Core Team (2015). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: http://www.R-project.org/ (accessed 6 November 2016).

Royston, P. 1982. Algorithm AS 181: the *W* test for normality. *Applied Statistics*, 31:176–180.

Thiery, J. P., G. Macaya, and G. Bernardi. 1976. An analysis of eukaryotic genomes by density gradient centrifugation. *Journal of Molecular Biology*, 108:219–235.

Tukey, J. W. 1977. Exploratory data analysis. Reading, PA: Addison-Wesley.

Zoubak, S., O. Clay, G. Bernardi. 1996. The gene distribution of the human genome. *Gene*, 174:95–102.

Figure Legends

Figure 4.1. Summary of methods used to compare GC content across vertebrate genomes. 1) FASTA-formatted genomes from GenBank were deinterleaved. 2) Using a window size provided by the user, a python script divided each FASTA contig starting at the first base. 3) Each divided sub-contig was analyzed for GC percent as described in the main text. Several example cases and script behavior are shown. 4) The python script concluded by producing four output files as described.

Figure 4.2. Standard deviation in GC content measures genomic heterogeneity. Across the six window sizes evaluated, mammals generally had the most heterogeneous genomes and fish had the most homogeneous. Birds and reptiles are somewhat intermediate between mammals and fish. As expected, the standard deviation in GC content decreases as window size increases. Box-plots show the median and interquartile range. Lines are the interquartile range*1.5. Points beyond these line are considered outliers (Tukey 1977). Zero standard deviation generally means that no contigs analyzed were long enough for analysis at that window size.

Figure 4.3. Relative GC content in 265 vertebrate genomes. Window sizes used: A) 1000bp, B) 3000bp, C) 5000bp, D) 20000bp, E) 80000bp, and F) 320000bp. The black and white line indicates the lizard *Aspidoscelis marmorata*.

Figure 4.4. Relative GC content summarized by clade across 265 vertebrate genomes. Window sizes used: A) 1000bp, B) 3000bp, C) 5000bp, D) 20000bp, E) 80000bp, and F) 320000bp. The black and white line indicates the lizard *Aspidoscelis marmorata*.

Figure 4.5. As assembly fragmentation increases the proportion of reads recovered decreases. Trend lines are LOESS curves, and the dark blue line is an average across all 265

samples with a localized standard error of the estimate in gray. Window sizes used: A) 1000bp, B) 3000bp, C) 5000bp, D) 20000bp, E) 80000bp, and F) 320000bp.

Figure 4.6. A) The number of isochoric (i.e., homogeneous) domains in *Aspidoscelis marmorata*, *Anolis carolinensis*, and *Homo sapiens*. B) Since genome size varies between these species, scaling by number of homogeneous domains offers a truer isochore family makeup comparison by proportion of number of isochores.

Figure 4.7. Scaling isochore abundance by isochore domain length, it becomes clear that the A) *Aspidoscelis marmorata* and B) *Anolis carolinensis* genomes lack heterogeneity in isochore families seen in C) *Homo sapiens*. Instead of clear breaks between isochoric families said to occur in *H. sapiens* (Costantini et al. 2006), DNA allotted to isochores distributes approximately normally in both lizard genomes.

1. Genome downloaded as FASTA and deinterleaved:

```
>Spp_A_Chr_1
ACGTAATGGAWGAGWSSTCGAATNNATGNNAYRCGBVCNGTAATGG
>Spp_A_Chr_2
NNAYRCAATGgacgtaatgtAAAATGGTSTCGAATNGTAGATGGAWGABV ...
...
```

2. Each FASTA contig divided by window size (10 bp in this example):

```
ACGTAATGGA | WGAGWSSTCG | AATNNATGNN | AYRCGBVCNG | TAATGG
NNAYRCAATG | gacgtaatgt | AAAATGGTST | CGAATNGTAG | ATGGAWGABV ...
```

3. Sub-contigs analyzed for GC content, if possible:

```
Spp_A_Chr_1_0: ACGTAATGGA >> GC% = 4/10 = 40%
Spp_A_Chr_1_1: WGAGWSSTCG >> GC% = 6/10 = 60%
Spp_A_Chr_1_2: AATNNATGNN >> >20% missing data; moved to missing list
Spp_A_Chr_1_3: AYRCGBVCNG >> >20% missing data; moved to missing list
Spp_A_Chr_1_4: TAATGG     >> Contig too short; moved to short list
Spp_A_Chr_2_0: NNAYRCAATG >> >20% missing data; moved to missing list
Spp_A_Chr_2_1: gacgtaatgt >> GC% = 4/10 = 40%
...
```

4. Outputs summarized into four files:
   1) Summary file for each genome at that window size. Lists GC%, GC% SD, number of contigs, etc.
   2) tsv of each sub-contig name that could be analyzed and its GC%
   3) tsv of headers for sub-contigs too short to be analyzed for GC%
   4) tsv of headers for sub-contigs with too much missing or ambiguous data to be analyzed for GC%

Figure 4.1

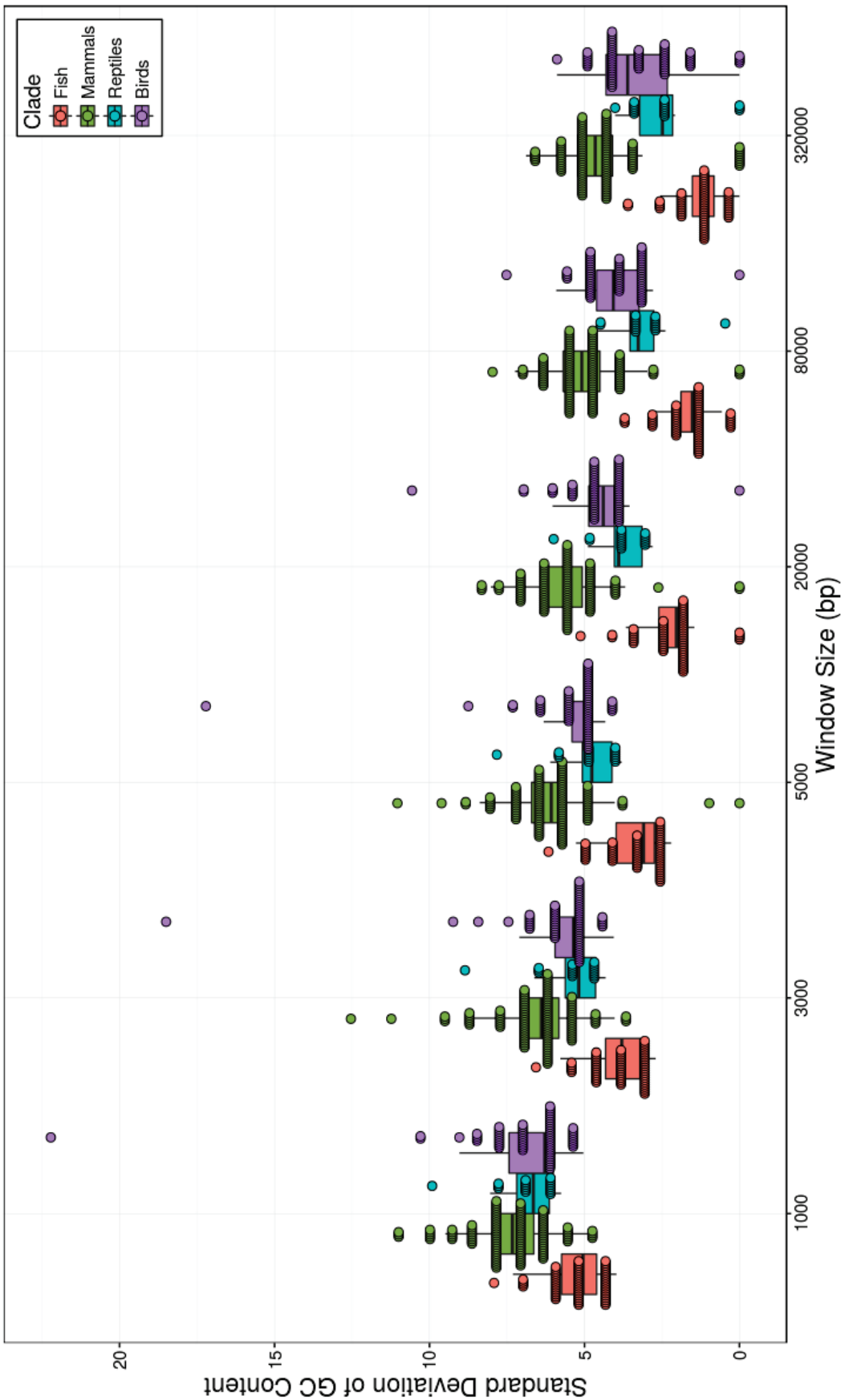
Clade
Fish
Mammals
Reptiles
Birds
Standard Deviation of GC Content
0
5
10
15
20
Window Size (bp)
1000
3000
5000
20000
80000
320000

Figure 4.2

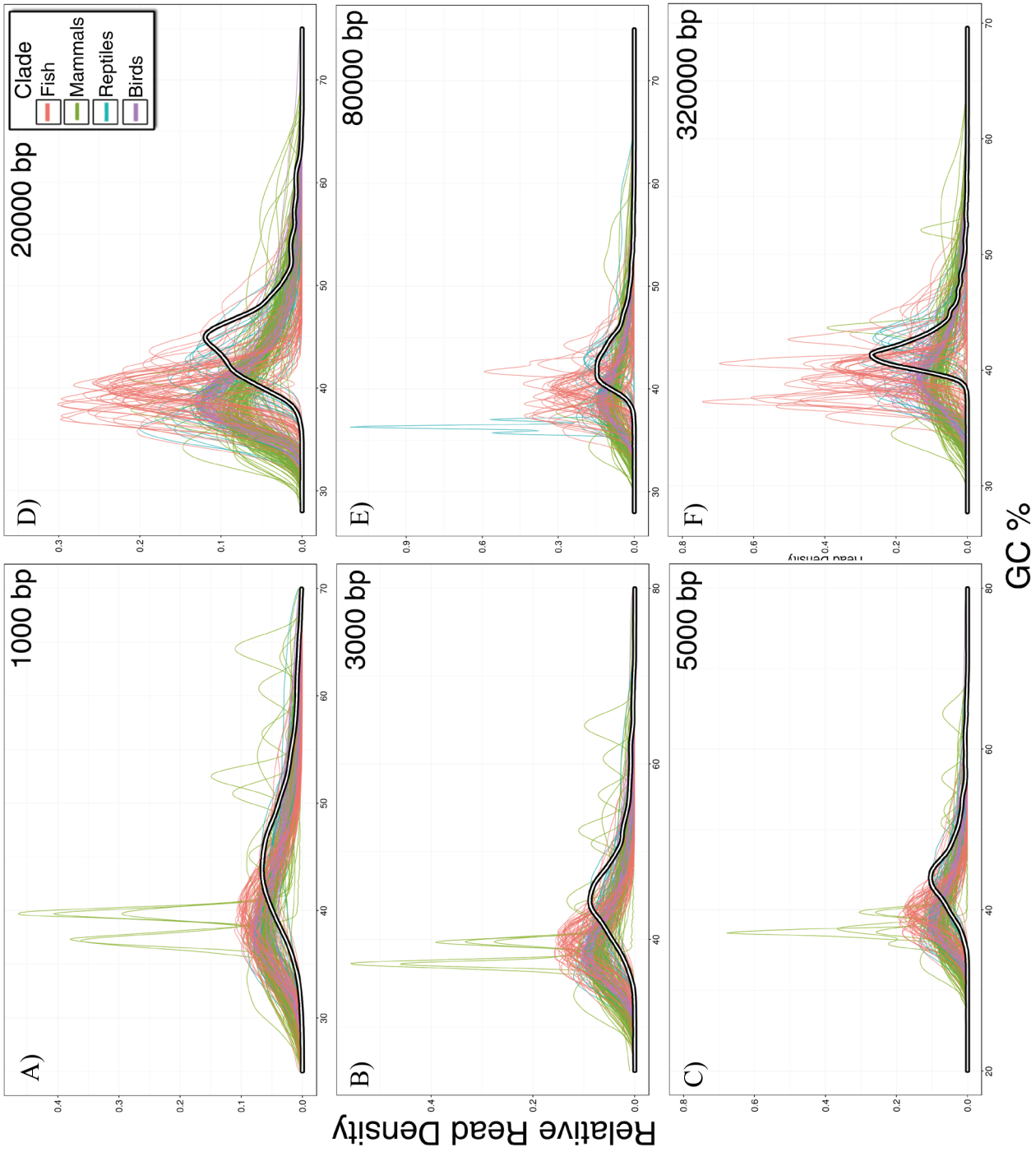
A)
1000 bp
B)
3000 bp
C)
5000 bp
D)
20000 bp
E)
80000 bp
F)
320000 bp
Clade
Fish
Mammals
Reptiles
Birds
Relative Read Density
GC %

Figure 4.3

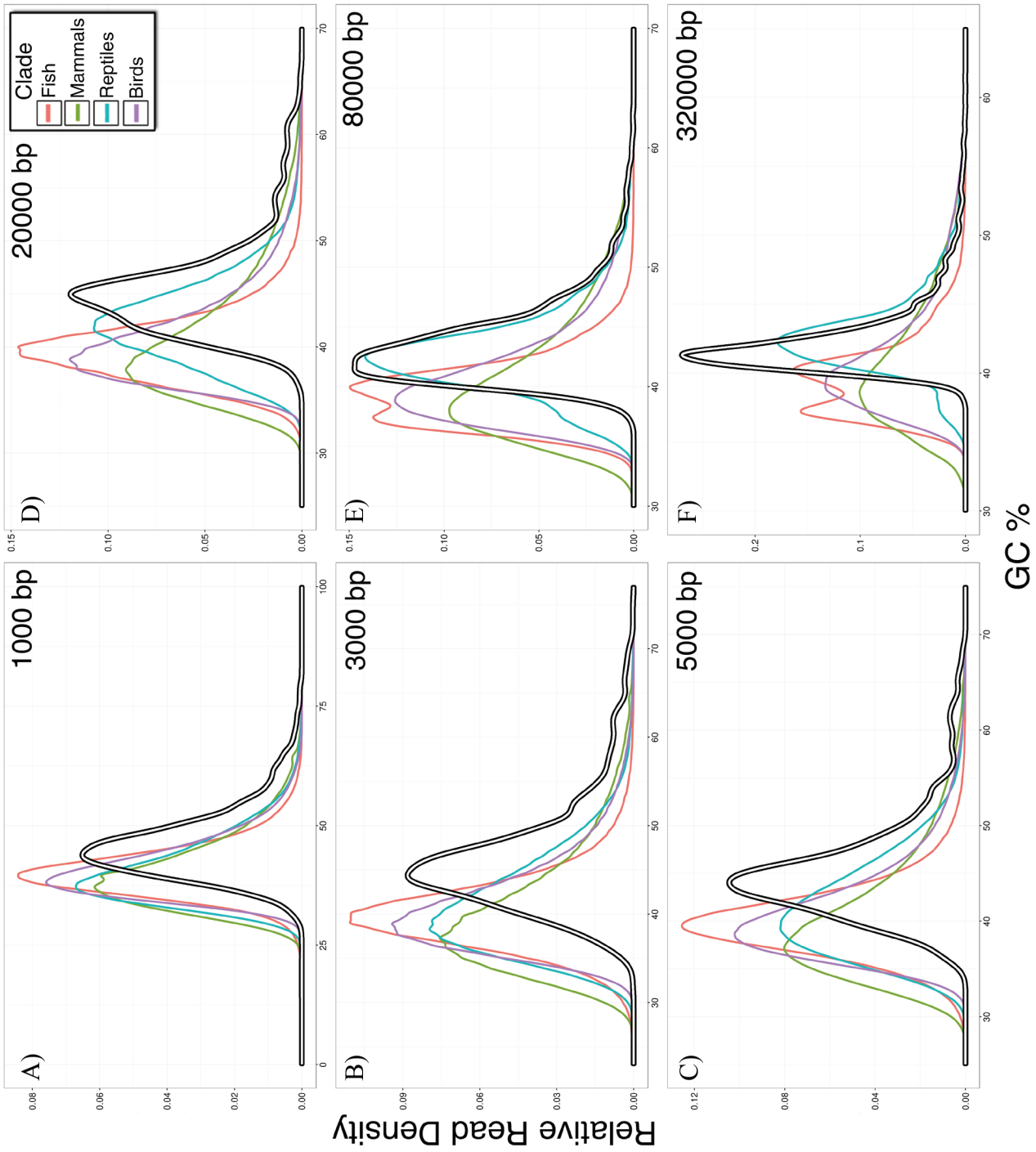
A)
1000 bp
B)
3000 bp
C)
5000 bp
D)
20000 bp
E)
80000 bp
F)
320000 bp
Clade
Fish
Mammals
Reptiles
Birds
Relative Read Density
GC %

Figure 4.4

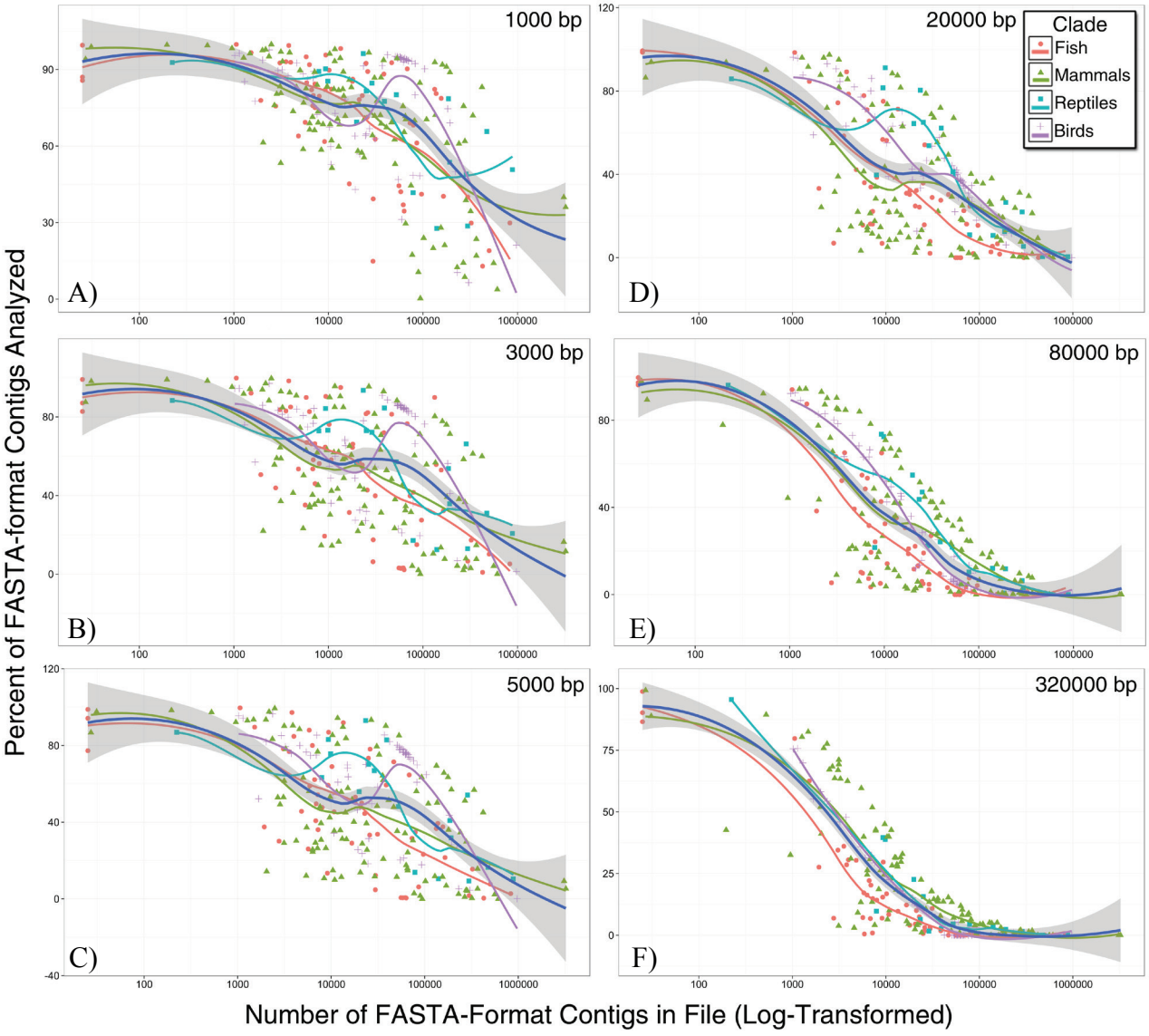
1000 bp
20000 bp
3000 bp
80000 bp
5000 bp
320000 bp
A)
B)
C)
D)
E)
F)
Clade
Fish
Mammals
Reptiles
Birds
Percent of FASTA-format Contigs Analyzed
Number of FASTA-Format Contigs in File (Log-Transformed)

Figure 4.5

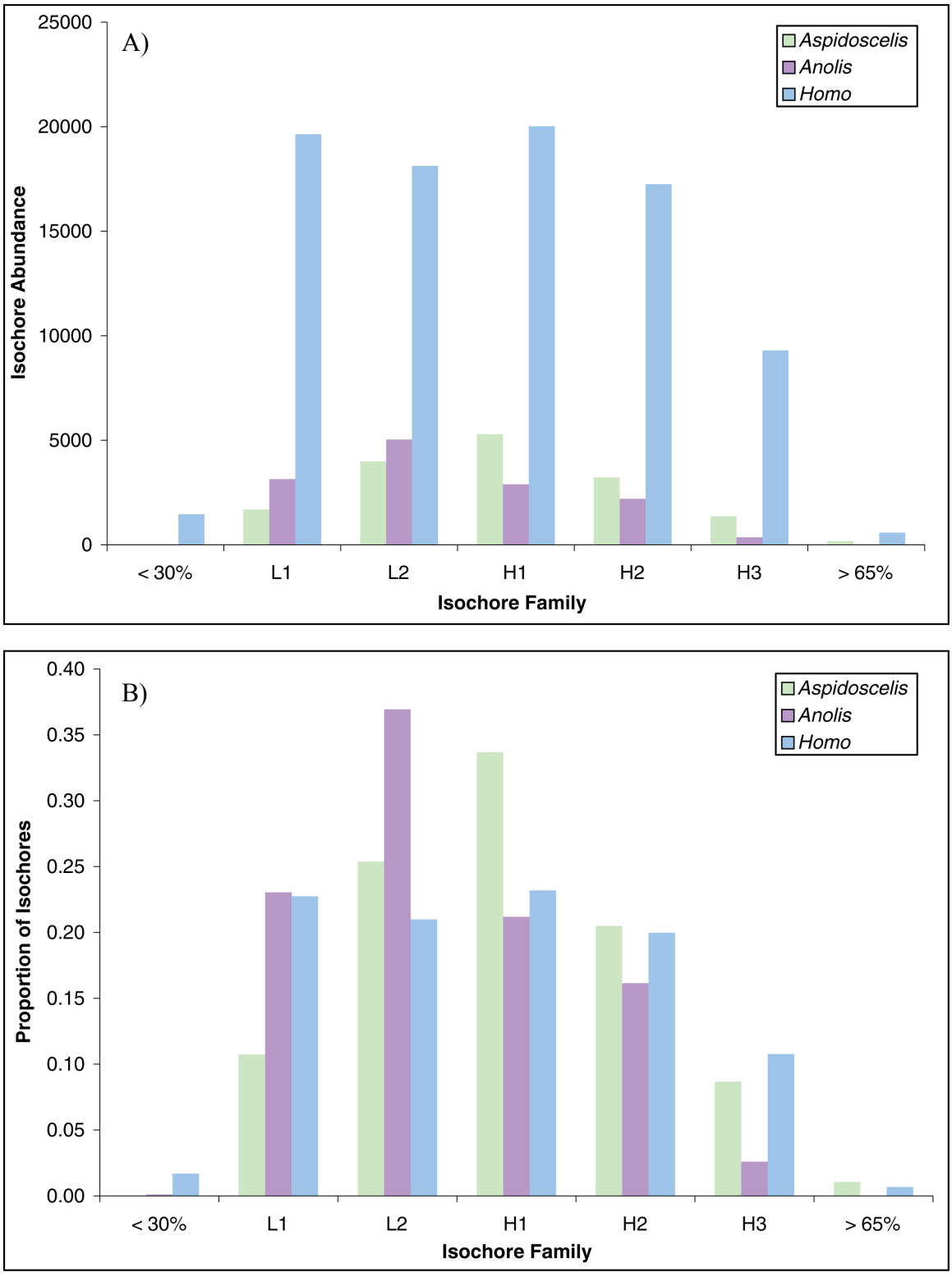
A)
Aspidoscelis
Anolis
Homo
Isochore Abundance
25000
20000
15000
10000
5000
0
< 30%
L1
L2
H1
H2
H3
> 65%
Isochore Family
B)
Aspidoscelis
Anolis
Homo
Proportion of Isochores
0.40
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
< 30%
L1
L2
H1
H2
H3
> 65%
Isochore Family

Figure 4.6

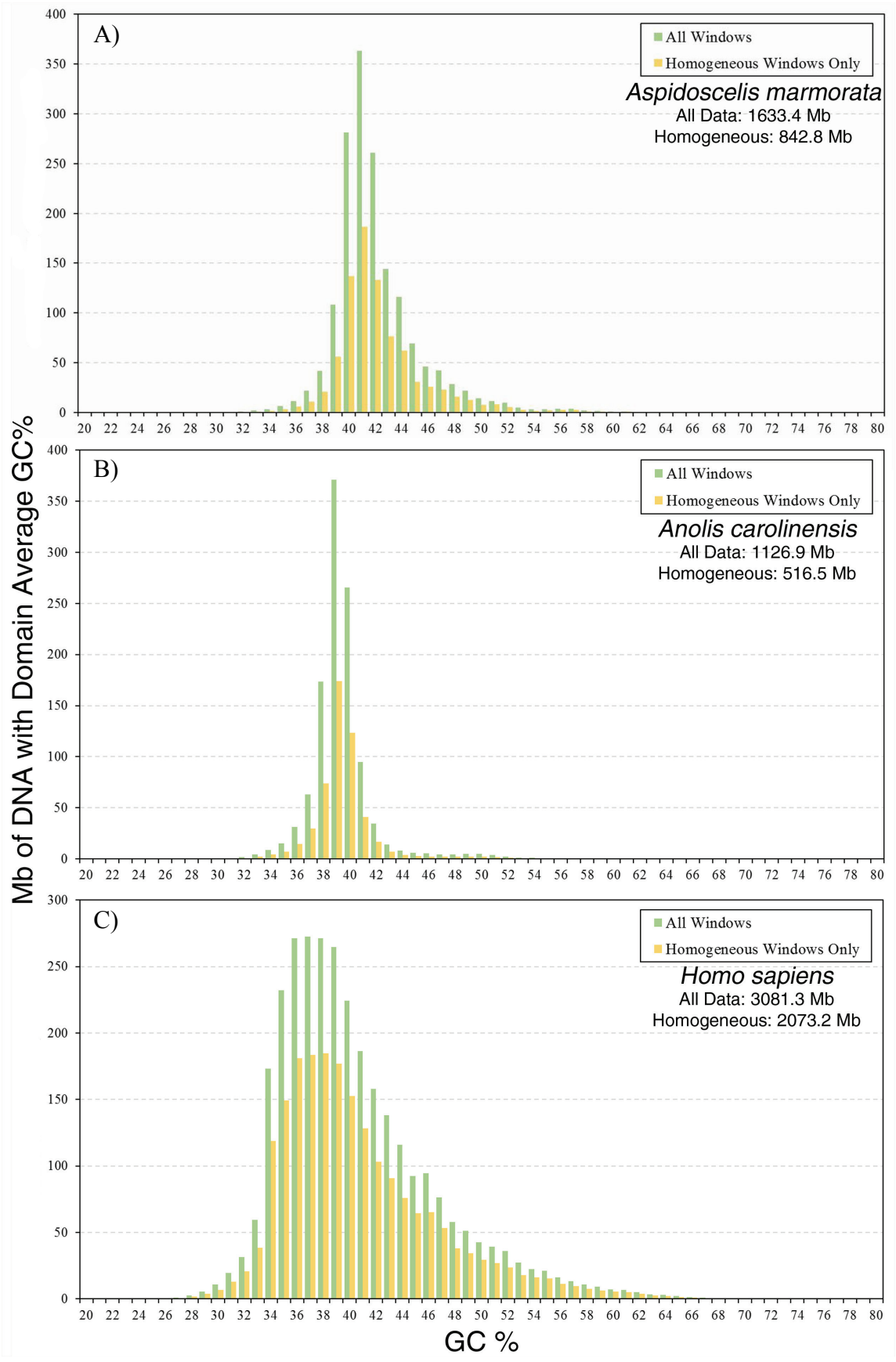
A)
All Windows
Homogeneous Windows Only
Aspidoscelis marmorata
All Data: 1633.4 Mb
Homogeneous: 842.8 Mb
B)
All Windows
Homogeneous Windows Only
Anolis carolinensis
All Data: 1126.9 Mb
Homogeneous: 516.5 Mb
C)
All Windows
Homogeneous Windows Only
Homo sapiens
All Data: 3081.3 Mb
Homogeneous: 2073.2 Mb
Mb of DNA with Domain Average GC%
GC %

Figure 4.7

APPENDIX A – Specimens used in Chapters 2 and 3

Institution	Institution Number	Field Number	Species	Latitude	Longitude	Country	State	County	mtDNA	ddRAD
UTA	R-62282	ASH022	*Aspidoscelis gularis gularis*	30.99736	-101.055	USA	Texas	Crockett		x
UTA	R-62283	ASH023	*Aspidoscelis gularis gularis*	30.99736	-101.055	USA	Texas	Crockett		x
UTA	R-62284	ASH024	*Aspidoscelis gularis gularis*	30.99736	-101.055	USA	Texas	Crockett		x
UTA	R-62285	ASH025	*Aspidoscelis gularis gularis*	31.26679	-100.525	USA	Texas	Tom Green	x	x
UTA	R-62335	ASH031	*Aspidoscelis gularis gularis*	32.98793	-98.7631	USA	Texas	Young		x
UTA	R-62286	ASH052	*Aspidoscelis gularis gularis*	32.2373	-97.8305	USA	Texas	Somervell		x
UTA	NA	ASH062	*Aspidoscelis tesselata*	34.44275	-101.078	USA	Texas	Briscoe		x
UTA	R-62334	ASH063	*Aspidoscelis gularis gularis*	34.44275	-101.078	USA	Texas	Briscoe	x	x
UTA	R-62276	ASH070	*Aspidoscelis tesselata*	34.44275	-101.078	USA	Texas	Briscoe	x	x
UTA	R-62369	ASH072	*Aspidoscelis tesselata*	33.33212	-104.33	USA	New Mexico	Chaves		x
UTA	R-62300	ASH076	*Aspidoscelis exsanguis*	34.6263	-104.366	USA	New Mexico	De Baca		x
UTA	R-62287	ASH078	*Aspidoscelis neomexicana*	34.6263	-104.366	USA	New Mexico	De Baca		x
UTA	R-62288	ASH079	*Aspidoscelis exsanguis*	34.61827	-104.373	USA	New Mexico	De Baca		x
UTA	R-62289	ASH080	*Aspidoscelis tesselata*	34.61827	-104.373	USA	New Mexico	De Baca	x	x
UTA	R-62290	ASH097	*Aspidoscelis tesselata*	34.76442	-104.535	USA	New Mexico	Guadalupe	x	
UTA	R-62370	ASH098	*Aspidoscelis tesselata*	34.76442	-104.535	USA	New Mexico	Guadalupe	x	x
UTA	R-62371	ASH099	*Aspidoscelis tesselata*	34.76442	-104.535	USA	New Mexico	Guadalupe	x	x
UTA	R-62291	ASH102	*Aspidoscelis tesselata*	34.76442	-104.535	USA	New Mexico	Guadalupe	x	x

Institution	Institution Number	Field Number	Species	Latitude	Longitude	Country	State	County	mtDNA	ddRAD
UTA	R-62292	ASH103	*Aspidoscelis tesselata*	34.76442	-104.535	USA	New Mexico	Guadalupe		x
UTA	R-62372	ASH104	*Aspidoscelis tesselata*	34.76983	-104.522	USA	New Mexico	Guadalupe	x	x
UTA	R-62293	ASH110	*Aspidoscelis tesselata*	34.60423	-104.383	USA	New Mexico	De Baca	x	x
UTA	R-62294	ASH111	*Aspidoscelis tesselata*	34.60487	-104.373	USA	New Mexico	De Baca	x	x
UTA	R-62295	ASH112	*Aspidoscelis tesselata*	34.38285	-104.264	USA	New Mexico	De Baca	x	
UTA	R-62301	ASH120	*Aspidoscelis inornata*	32.56632	-104.382	USA	New Mexico	Eddy	x	x
UTA	R-62342	ASH131	*Aspidoscelis marmorata reticuloriens*	32.5454	-104.367	USA	New Mexico	Eddy	x	x
UTA	R-62373	ASH133	*Aspidoscelis exsanguis*	32.5698	-104.355	USA	New Mexico	Eddy		x
UTA	R-62374	ASH134	*Aspidoscelis tesselata*	32.5698	-104.355	USA	New Mexico	Eddy		x
UTA	R-63270	ASH135	*Aspidoscelis tesselata*	32.56632	-104.382	USA	New Mexico	Eddy	x	x
UTA	R-62376	ASH140	*Aspidoscelis tesselata*	32.56535	-104.38	USA	New Mexico	Eddy		x
UTA	R-62296	ASH141	*Aspidoscelis tesselata*	32.56535	-104.38	USA	New Mexico	Eddy	x	x
UTA	R-62297	ASH142	*Aspidoscelis inornata heptagramma*	32.56535	-104.38	USA	New Mexico	Eddy		x
UTA	R-62298	ASH143	*Aspidoscelis inornata heptagramma*	32.56535	-104.38	USA	New Mexico	Eddy		x
UTA	R-62277	ASH144	*Aspidoscelis marmorata reticuloriens*	32.1458	-104.382	USA	New Mexico	Eddy		x
UTA	R-62278	ASH145	*Aspidoscelis marmorata reticuloriens*	32.1458	-104.382	USA	New Mexico	Eddy	x	x
UTA	R-62279	ASH146	*Aspidoscelis marmorata reticuloriens*	32.1458	-104.382	USA	New Mexico	Eddy	x	x
UTA	R-62280	ASH147	*Aspidoscelis marmorata reticuloriens*	32.1458	-104.382	USA	New Mexico	Eddy		x

Institution	Institution Number	Field Number	Species	Latitude	Longitude	Country	State	County	mtDNA	ddRAD
UTA	R-62281	ASH148	*Aspidoscelis marmorata reticuloriens*	32.1458	-104.382	USA	New Mexico	Eddy	x	x
UTA	R-62340	ASH150	*Aspidoscelis gularis gularis*	30.59303	-103.94	USA	Texas	Jeff Davis		x
UTA	R-62378	ASH153	*Aspidoscelis exsanguis*	30.59303	-103.94	USA	Texas	Jeff Davis		x
UTA	R-62343	ASH159	*Aspidoscelis gularis gularis*	30.29727	-103.599	USA	Texas	Brewster		x
UTA	R-62384	ASH160	*Aspidoscelis exsanguis*	30.29727	-103.599	USA	Texas	Brewster		x
UTA	R-62341	ASH162	*Aspidoscelis gularis gularis*	30.54213	-103.838	USA	Texas	Jeff Davis		x
UTA	R-63271	ASH164	*Aspidoscelis exsanguis*	30.54213	-103.838	USA	Texas	Jeff Davis		x
UTA	R-63272	ASH165	*Aspidoscelis marmorata marmorata*	**30.9537**	**-104.899**	USA	Texas	Hudspeth	x	x
UTA	R-62386	ASH167	*Aspidoscelis tesselata*	30.9346	-104.912	USA	Texas	Hudspeth		x
UTA	R-63273	ASH168	*Aspidoscelis tesselata*	30.93195	-104.919	USA	Texas	Hudspeth		x
UTA	R-62388	ASH173	*Aspidoscelis tesselata*	30.22233	-104.062	USA	Texas	Presidio	x	x
UTA	R-62338	ASH174	*Aspidoscelis inornata heptagramma*	30.22233	-104.062	USA	Texas	Presidio	x	
UTA	R-62343	ASH175	*Aspidoscelis gularis septemvittata*	30.06368	-104.177	USA	Texas	Presidio		x
UTA	R-62339	ASH176	*Aspidoscelis gularis septemvittata*	30.06368	-104.177	USA	Texas	Presidio	x	x
UTA	R-62389	ASH178	*Aspidoscelis tesselata*	29.47505	-104.942	USA	Texas	Presidio	x	x
UTA	R-63274	ASH179	*Aspidoscelis tesselata*	29.47505	-104.942	USA	Texas	Presidio	x	x
UTA	R-62390	ASH180	*Aspidoscelis tesselata*	29.47642	-103.972	USA	Texas	Presidio	x	x
UTA	R-63275	ASH182	*Aspidoscelis marmorata reticuloriens*	29.47408	-103.937	USA	Texas	Presidio		x
UTA	R-62391	ASH183	*Aspidoscelis marmorata marmorata*	**29.37968**	**-104.115**	USA	Texas	Presidio		x
UTA	R-62392	ASH184	*Aspidoscelis marmorata marmorata*	**29.37968**	**-104.115**	USA	Texas	Presidio	x	x
UTA	R-62394	ASH190	*Aspidoscelis inornata heptagramma*	29.52935	-102.921	USA	Texas	Brewster		
UTA	R-62395	ASH192	*Aspidoscelis marmorata reticuloriens*	29.55232	-102.933	USA	Texas	Brewster	x	x

Institution	Institution Number	Field Number	Species	Latitude	Longitude	Country	State	County	mtDNA	ddRAD
UTA	R-63276	ASH193	*Aspidoscelis marmorata reticuloriens*	29.48113	-102.856	USA	Texas	Brewster	x	x
UTA	R-62396	ASH205	*Aspidoscelis inornata heptagramma*	29.62142	-103.033	USA	Texas	Brewster	x	
UTA	R-62397	ASH206	*Aspidoscelis marmorata reticuloriens*	29.62683	-103.021	USA	Texas	Brewster	x	x
UTA	R-62398	ASH207	*Aspidoscelis marmorata reticuloriens*	29.62683	-103.021	USA	Texas	Brewster		x
UTA	R-62400	ASH214	*Aspidoscelis marmorata reticuloriens*	29.5523	-102.933	USA	Texas	Brewster		x
UTA	R-62401	ASH215	*Aspidoscelis marmorata reticuloriens*	29.59765	-102.988	USA	Texas	Brewster		x
UTA	R-62402	ASH216	*Aspidoscelis marmorata reticuloriens*	29.55227	-102.933	USA	Texas	Brewster		x
UTA	R-62344	ASH217	*Aspidoscelis gularis septemvittata*	29.52557	-103.049	USA	Texas	Brewster		x
UTA	R-62404	ASH219	*Aspidoscelis marmorata reticuloriens*	29.31013	-103.164	USA	Texas	Brewster	x	x
UTA	R-62405	ASH220	*Aspidoscelis marmorata reticuloriens*	29.31013	-103.164	USA	Texas	Brewster	x	x
UTA	R-62406	ASH221	*Aspidoscelis marmorata reticuloriens*	29.33978	-103.345	USA	Texas	Brewster	x	x
UTA	R-62345	ASH223	*Aspidoscelis gularis septemvittata*	29.21462	-103.377	USA	Texas	Brewster		x
UTA	R-62407	ASH229	*Aspidoscelis marmorata reticuloriens*	29.43233	-103.397	USA	Texas	Brewster	x	x
UTA	R-62408	ASH230	*Aspidoscelis marmorata reticuloriens*	30.0314	-103.561	USA	Texas	Brewster	x	x
UTA	R-62409	ASH231	*Aspidoscelis marmorata reticuloriens*	30.0314	-103.561	USA	Texas	Brewster		x
UTA	R-62346	ASH232	*Aspidoscelis gularis gularis*	29.69577	-101.321	USA	Texas	Val Verde		x
UTA	R-62348	ASH234	*Aspidoscelis gularis gularis*	29.69577	-101.321	USA	Texas	Val Verde		x
UTA		MKF811	*Aspidoscelis sexlineata viridis*			USA			x	x
	NC_011607		*Podarcis muralis*			Austria			x	

Institution	Institution Number	Field Number	Species	Latitude	Longitude	Country	State	County	mtDNA	ddRAD
AMNH	R128249	JC4850	*Aspidoscelis gularis*	30.90357	-103.791	USA	Texas	Reeves		x
AMNH	R129176	JC5124	*Aspidoscelis scalaris*	26.38914	-105.355	Mexico	Durango			x
AMNH	R127001	JC4391	*Aspidoscelis tesselata*	31.93045	-106.513	USA	Texas	El Paso		x
AMNH	R127002	JC4412	*Aspidoscelis tesselata*	35.44664	-106.439	USA	New Mexico	Sandoval		x
AMNH	R129216	JC5178	*Aspidoscelis tesselata*	30.90357	-103.791	USA	Texas	Reeves		x
AMNH	R129217	JC5179	*Aspidoscelis tesselata*	30.90357	-103.791	USA	Texas	Reeves		x
AMNH	R134864	NMGF421	*Aspidoscelis dixoni C*	31.94285	-108.927	USA	New Mexico	Hidalgo		x
AMNH	R136875	JC5756	*Aspidoscelis tesselata*	35.37569	-104.199	USA	New Mexico	San Miguel		x
AMNH	R136876	JC5757	*Aspidoscelis tesselata*	35.37569	-104.199	USA	New Mexico	San Miguel		x
AMNH	R148361	UADZ4005	*Aspidoscelis dixoni* A or B	29.9092	-104.489	USA	Texas	Presidio		x
AMNH	R148426	UADZ6434	*Aspidoscelis dixoni* A	29.87371	-104.463	USA	Texas	Presidio		x
AMNH	R148427	UADZ6438	*Aspidoscelis dixoni* A	29.87371	-104.463	USA	Texas	Presidio		x
AMNH	R154170	JC8187	*Aspidoscelis tesselata*	33.3556	-106.594	USA	New Mexico	Sierra		x
AMNH	R154171	JC8188	*Aspidoscelis tesselata*	33.3556	-106.594	USA	New Mexico	Sierra		x
AMNH	R127116	JC4260	*Aspidoscelis marmorata*	32.65212	-107.185	USA	New Mexico	Dona Ana		x
AMNH	R127120	JC4596	*Aspidoscelis marmorata*	32.2062	-108.085	USA	New Mexico	Luna		x
AMNH	R127121	JC4597	*Aspidoscelis marmorata*	32.2062	-108.085	USA	New Mexico	Luna		x
AMNH	R131080	JC5339	*Aspidoscelis marmorata*	33.91699	-106.894	USA	New Mexico	Socorro		x
AMNH	R131081	JC5340	*Aspidoscelis marmorata*	33.91699	-106.894	USA	New Mexico	Socorro		x
Smithsonian	USNM315523	KdQ551	*Aspidoscelis gularis gularis*	28.2	-98.333	USA	Texas	McMullen		x

Institution	Institution Number	Field Number	Species	Latitude	Longitude	Country	State	County	mtDNA	ddRAD
Smithsonian		KdQ194	*Aspidoscelis gularis gularis*		30.48741	-97.8076	Texas	Travis		x
BRTC	91365	DL907	*Aspidoscelis marmorata*	33.0206	-103.92	USA	New Mexico	Chaves		x
BRTC	92137	MTH370	*Aspidoscelis gularis gularis*	30.79169	-98.3509	USA	Texas	Burnet		x
BRTC	92211	DL1134	*Aspidoscelis marmorata*	31.49589	-102.659	USA	Texas	Crane		x
BRTC	94085	TJH2108	*Aspidoscelis gularis gularis*	30.9444	-100.624	USA	Texas	Schleicher		x
BRTC	94089	TJH2112	*Aspidoscelis gularis gularis*	30.76873	-100.288	USA	Texas	Schleicher		x
BRTC	95195	TJH2818	*Aspidoscelis gularis gularis*	29.96305	-98.0852	USA	Texas	Hays		x
BRTC	95229	TJH2853	*Aspidoscelis marmorata*	31.63433	-102.871	USA	Texas	Ward		x
BRTC	95278	TJH2903	*Aspidoscelis marmorata*	31.58394	-102.708	USA	Texas	Crane		x
BRTC	95291	TJH2916	*Aspidoscelis marmorata*	31.65069	-102.775	USA	Texas	Ward		x
BRTC	95347	TJH2972	*Aspidoscelis marmorata*	32.19819	-102.687	USA	Texas	Andrews		x
BRTC	95383	TJH3008	*Aspidoscelis marmorata*	32.1371	-102.803	USA	Texas	Andrews		x
BRTC	95410	TJH3035	*Aspidoscelis gularis gularis*	32.12197	-102.785	USA	Texas	Andrews		x
BRTC	95434	TJH3059	*Aspidoscelis marmorata*	31.99728	-103.065	USA	Texas	Winkler		x
BRTC		ESP9319	*Aspidoscelis marmorata*	33.5	-104.5	USA	New Mexico	Chaves		x
BRTC		ESP9320	*Aspidoscelis marmorata*	33.5	-104.5	USA	New Mexico	Chaves		x
BRTC		ESP9618	*Aspidoscelis marmorata*	32.7	-103.2	USA	New Mexico	Lea		x
BRTC		ESP9619	*Aspidoscelis marmorata*	32.7	-103.2	USA	New Mexico	Lea		x
BRTC		ESP9629	*Aspidoscelis marmorata*	32.2	-104	USA	New Mexico	Eddy		x
BRTC		ESP9630	*Aspidoscelis marmorata*	32.2	-104	USA	New Mexico	Eddy		x
BRTC		TJH3454	*Aspidoscelis gularis gularis*	26.71971	-98.5225	USA	Texas	Starr		x

Institution	Institution Number	Field Number	Species	Latitude	Longitude	Country	State	County	mtDNA	ddRAD
BRTC		TJH3385	*Aspidoscelis gularis gularis*	26.76105	-98.7766	USA	Texas	Starr		x
BRTC		TJH3382	*Aspidoscelis gularis gularis*	27.2258	-97.8516	USA	Texas	Kenedy		x
BRTC		H7725	*Aspidoscelis gularis gularis*	30.78961	-98.355	USA	Texas	Burnet		x
BRTC		H4647	*Aspidoscelis dixoni*	29.8755	-104.507	USA	Texas	Presidio		x
BRTC		H4648	*Aspidoscelis dixoni*	29.8755	-104.507	USA	Texas	Presidio		x
BRTC		H4649	*Aspidoscelis dixoni*	29.8755	-104.507	USA	Texas	Presidio		x
BRTC		H4650	*Aspidoscelis dixoni*	29.8755	-104.507	USA	Texas	Presidio		x
BRTC		H4651	*Aspidoscelis gularis septemvittata*	29.8755	-104.507	USA	Texas	Presidio		x
BRTC		H4652	*Aspidoscelis gularis septemvittata*	29.8755	-104.507	USA	Texas	Presidio		x
TNHC	60571	TJL305	*Aspidoscelis gularis gularis*	30.08726	-97.1735	USA	Texas	Bastrop		x
TNHC	67456	GBP789	*Aspidoscelis gularis gularis*	30.46877	-99.7851	USA	Texas	Kimble		x
TNHC	67457	GBP810	*Aspidoscelis gularis gularis*	33.834	-100.516	USA	Texas	King		x
TNHC	52029	AHP3383	*Aspidoscelis gularis gularis*	29.12139	-100.477	USA	Texas	Kinney		x
TNHC	53225	DCC3019	*Aspidoscelis gularis gularis*	29.95	-99.9639	USA	Texas	Real		x
TNHC	65456	TJL1320	*Aspidoscelis gularis gularis*	30.2836	-97.8993	USA	Texas	Travis		x
TNHC	68824	TJL1574	*Aspidoscelis gularis gularis*	30.235	-97.6453	USA	Texas	Travis		x
TNHC	68825	TJL1677	*Aspidoscelis gularis gularis*	30.42917	-97.6536	USA	Texas	Travis		x
TNHC	66886	TJL1289	*Aspidoscelis tesselata*	30.54944	-104.662	USA	Texas	Presido		x
TNHC	60372	AHP3665	*Aspidoscelis gularis septemvittata*	29.89038	-104.521	USA	Texas	Presido		x
TNHC	60373	AHP3666	*Aspidoscelis gularis septemvittata*	29.89038	-104.521	USA	Texas	Presido		x
TNHC	60370	AHP3664	*Aspidoscelis dixoni*	29.89038	-104.521	USA	Texas	Presido		x

Institution	Institution Number	Field Number	Species	Latitude	Longitude	Country	State	County	mtDNA	ddRAD
TNHC	60369	AHP3663	*Aspidoscelis dixoni*	29.89038	-104.521	USA	Texas	Presido		x
TNHC		TJD984	*Aspidoscelis gularis gularis*	30.32288	-103.743	USA	Texas	Brewster		x
TNHC		TJL2434	*Aspidoscelis tesselata*	30.44806	-104.635	USA	Texas	Jeff Davis		x
TNHC		TJL2445	*Aspidoscelis tesselata*	30.63873	-104.63	USA	Texas	Jeff Davis		x
UTA	R-42096	MSM13	*Aspidoscelis motaguae*			Mexico				x
UTA	R-52259	ENS9235	*Cnemidophorus lemniscatus*							x
UNAM		ANMO878	*Aspidoscelis deppii*			Mexico				x
UTA	R-39895	JWS061	*Aspidoscelis gularis ssp*	32.61068	-98.5725	USA	Texas	Palo Pinto		x
UTA	R-40426	JWS425	*Aspidoscelis sp*	30.7069	-104.104	USA	Texas	Jeff Davis		x
UNAM	JAL23303	MX13_27	*Aspidoscelis sp*			Mexico				x
UTA		JWS666	*Aspidoscelis sp*	30.59396	-103.939	USA	Texas	Jeff Davis		x
OCGR	43070	4943	*Aspidoscelis gularis gularis*	35.07459	-99.8928	USA	Oklahoma	Beckham		x
OCGR	43071	4944	*Aspidoscelis gularis gularis*	35.07459	-99.8928	USA	Oklahoma	Beckham		x
UNM	UNM74810		*Aspidoscelis marmorata*	32.1997	-104.252	USA	New Mexico	Eddy		x
UNM	UNM79169		*Aspidoscelis tesselata*	33.61666	-103.858	USA	New Mexico	Chaves		x
UNM	UNM74685		*Aspidoscelis tesselata*	32.09501	-108.975	USA	New Mexico	Hidalgo		x
UNM	UNM73545		*Aspidoscelis gularis gularis*	34.443	-107.012	USA	New Mexico	Socorro		x
LACM	178804	TC923	*Aspidoscelis tesselata*	29.06	-103.104	USA	Texas	Brewster		x
LACM	130598	TC136	*Aspidoscelis gularis ssp*	21.40472	-99.6369	Mexico	Queretaro			x
LACM	130599	TC137	*Aspidoscelis gularis ssp*	21.40472	-99.6369	Mexico	Queretaro			x
LACM	134314	TC372	*Aspidoscelis gularis gularis*	29.6631	-98.5577	USA	Texas	Bexar		x
UTEP	20329	CSL9299	*Aspidoscelis gularis gularis*	31.66947	-103.626	USA	Texas	Loving		x
UTEP	20330	CSL9300	*Aspidoscelis gularis gularis*	31.66995	-103.626	USA	Texas	Loving		x
UTEP	20725	CSL9464	*Aspidoscelis gularis gularis*	30.58618	-104.287	USA	Texas	Jeff Davis		x

Institution	Institution Number	Field Number	Species	Latitude	Longitude	Country	State	County	mtDNA	ddRAD
UTEP	20726	CSL9471	*Aspidoscelis gularis gularis*	30.53463	-104.064	USA	Texas	Jeff Davis		x
UTEP	20724	CSL9473	*Aspidoscelis gularis gularis*	30.90688	-104.147	USA	Texas	Jeff Davis		x
UTEP	1267	DIL298	*Aspidoscelis gularis gularis*	30.34114	-97.8443	USA	Texas	Travis		x
UTEP	12308	12308	*Aspidoscelis marmorata*	31.82362	-104.112	USA	Texas	Culberson		x
UTEP	18536	DIL423	*Aspidoscelis marmorata*	32.40727	-106.084	USA	New Mexico	Otero		x
UTEP	20543	CSL9383	*Aspidoscelis marmorata*	31.80402	**-106.867**	USA	New Mexico	Dona Ana		x
UTEP		KWF3106	*Aspidoscelis marmorata*	32.23919	-107.124	USA	New Mexico	Dona Ana		x
UTEP		KWF3107	*Aspidoscelis marmorata*	32.23919	-107.124	USA	New Mexico	Dona Ana		x
UTEP	18926	CSL8739	*Aspidoscelis tesselata*	32.08045	-105.092	USA	New Mexico	Otero		x
UTEP	18971	CSL8784	*Aspidoscelis tesselata*	30.78593	-104.961	USA	Texas	Hudspeth		x
UTEP	18972	CSL8785	*Aspidoscelis tesselata*	30.78547	-104.979	USA	Texas	Hudspeth		x
UTEP	18971	CSL9434	*Aspidoscelis tesselata*	30.78592	-104.961	USA	Texas	Hudspeth		x
UTEP	20759	CSL9497	*Aspidoscelis tesselata*	31.21395	-104.851	USA	Texas	Culberson		x
UTEP	20758	CSL9501	*Aspidoscelis tesselata*	31.03883	-104.897	USA	Texas	Culberson		x
UTEP	20737	CSL9480	*Aspidoscelis neomexicana*	31.7936	-106.039	USA	Texas	El Paso		x
UTEP	18998	JJ10	*Aspidoscelis tigris*	33.81097	-111.644	USA	Arizona	Maricopa		x
UTEP	20720	CSL9444	*Aspidoscelis uniparens*	31.94337	-108.889	USA	New Mexico	Hidalgo		x
UTEP	19164	DIL706	*Aspidoscelis velox*	**34.58117**	-105.111	USA	New Mexico	Guadalupe		x
LSU-MNS		H3229	*Aspidoscelis gularis gularis*	29.97	-98.108	USA	Texas	Travis		x
UTA	R-63474	CER1028	*Aspidoscelis marmorata*	29.63518	-103.333	USA	Texas	Brewster		x
UTA	R-63476	CER1030	*Aspidoscelis gularis gularis*	30.53135	-104.359	USA	Texas	Jeff Davis		x
UTA	R-63477	CER1031	*Aspidoscelis gularis gularis*	30.53135	-104.359	USA	Texas	Jeff Davis		x

Institution	Institution Number	Field Number	Species	Latitude	Longitude	Country	State	County	mtDNA	ddRAD
UTA	R-63478	CER1032	*Aspidoscelis gularis gularis*	30.53135	-104.359	USA	Texas	Jeff Davis		x
UTA	R-63479	CER1033	*Aspidoscelis gularis gularis*	30.53135	-104.359	USA	Texas	Jeff Davis		x
UTA	R-63470	EW2	*Aspidoscelis gularis gularis*	33.3098	-97.6079	USA	Texas	Wise		x
UTA		MKF810	*Aspidoscelis gularis gularis*			USA	Texas			x
UTA	R-59335	CLC752	*Aspidoscelis gularis gularis*	31.03063	-100.997	USA	Texas	Crockett		x
TNHC	60973	DCC3558	*Aspidoscelis gularis gularis*	29.9401	-100.971	USA	Texas	Val Verde		x
TNHC	53225	DCC3019	*Aspidoscelis gularis gularis*	29.95	-99.9639	USA	Texas	Real		x
TNHC	52029	AHP3383	*Aspidoscelis gularis gularis*	29.12139	-100.477	USA	Texas	Kinney		x
TNHC	60571	TJL305	*Aspidoscelis gularis gularis*	30.08726	-97.1735	USA	Texas	Bastrop		x
TNHC	66795	WLH1220	*Aspidoscelis gularis gularis*	30.23762	-97.2789	USA	Texas	Bastrop		x
TNHC	68824	TJL1574	*Aspidoscelis gularis gularis*	30.235	-97.6453	USA	Texas	Travis		x
TNHC	65456	TJL1320	*Aspidoscelis gularis gularis*	30.2836	-97.8993	USA	Texas	Travis		x
TNHC	68827	TJL1738	*Aspidoscelis gularis gularis*	30.35889	-97.9492	USA	Texas	Travis		x
TNHC	68825	TJL1677	*Aspidoscelis gularis gularis*	30.42917	-97.6536	USA	Texas	Travis		x
TNHC	84905		*Aspidoscelis gularis gularis*	31.4985	-100.467	USA	Texas	Tom Green		x
TNHC	67456	GBP 789	*Aspidoscelis gularis gularis*	30.46877	-99.7851	USA	Texas	Kimble		x
TNHC	67457	GBP 810	*Aspidoscelis gularis gularis*	33.834	-100.516	USA	Texas	King		x
SIMS			*Aspidoscelis marmorata*			USA				x
SIMS			*Aspidoscelis tesselata*			USA				x

APPENDIX B – Genomes downloaded from GenBank

#Organism/Name	Assembly	SubGroup	Size (Mb)	GC%	Scaffolds	GenBank FTP
Anolis carolinensis	GCA_000090745.2	Reptiles	1799.14	40.8238	6646	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000090745.2_AnoCar2.0
Apalone spinifera	GCA_000385615.1	Reptiles	1931.08	42.8	286620	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000385615.1_ASM38561v1
Chelonia mydas	GCA_000344595.1	Reptiles	2208.41	43.7	140023	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000344595.1_CheMyd_1.0
Chrysemys picta bellii	GCA_000241765.2	Reptiles	2365.77	44.564	78631	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000241765.2_Chrysemys_picta_bellii-3.0.3
Crocodylus porosus	GCA_000768395.1	Reptiles	2120.57	44.3	23365	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000768395.1_Cpor_2.0
Crotalus horridus	GCA_001625485.1	Reptiles	1520.33	34.3	186068	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001625485.1_ASM162548v1
Crotalus mitchellii pyrrhus	GCA_000737285.1	Reptiles	1126.79	38.6	473380	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000737285.1_CrotMitch1.0
Gekko japonicus	GCA_001447785.1	Reptiles	2490.27	45.5	191500	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001447785.1_Gekko_japonicus_V1.1
Ophiophagus hannah	GCA_000516915.1	Reptiles	1594.07	40.6	296399	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000516915.1_OphHan1.0
Pantherophis guttatus	GCA_001185365.1	Reptiles	1404.22	38.3	883920	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001185365.1_PanGut1.0
Pelodiscus sinensis	GCA_000230535.1	Reptiles	2202.48	44.4999	19904	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000230535.1_PelSin_1.0
Protobothrops mucrosquamatus	GCA_001527695.2	Reptiles	1673.88	40.6	52281	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001527695.2_P.Mucros_1.0
Python bivittatus	GCA_000186305.2	Reptiles	1435.05	39.7	39113	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000186305.2_Python_molurus_bivittatus-5.0.2
Thamnophis sirtalis	GCA_001077635.2	Reptiles	1424.9	41.8	7930	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001077635.2_Thamnophis_sirtalis-6.0
Vipera berus berus	GCA_000800605.1	Reptiles	1532.39	41.3	28883	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000800605.1_Vber.be_1.0
Acinonyx jubatus	GCA_001443585.1	Mammals	2372.55	41.4	14383	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001443585.1_aciJub1
Ailuropoda melanoleuca	GCA_000004335.1	Mammals	2299.51	41.7	81467	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000004335.1_AilMel_1.0
Aotus nancymaae	GCA_000952055.1	Mammals	2926.58	41.7	29223	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000952055.1_Anan_1.0
Apodemus sylvaticus	GCA_001305905.1	Mammals	3758.14	22.6	559629	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001305905.1_ASM130590v1
Balaenoptera acutorostrata scammoni	GCA_000493695.1	Mammals	2431.69	41.4	10776	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000493695.1_BalAcu1.0
Balaenoptera bonaerensis	GCA_000978805.1	Mammals	2234.64	40.7	421444	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000978805.1_ASM97880v1
Bison bison bison	GCA_000754665.1	Mammals	2828.03	42.2	128431	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000754665.1_Bison_UMD1.0

#Organism/Name	Assembly	SubGroup	Size (Mb)	GC%	Scaffolds	GenBank FTP
Bos indicus	GCA_000247795.2	Mammals	2673.95	42.3028	31	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000247795.2_Bos_indicus_1.0
Bos mutus	GCA_000298355.1	Mammals	2645.16	42	41192	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000298355.1_BosGru_v2.0
Bos taurus	GCA_000003205.6	Mammals	2724.98	41.8685	5998	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000003205.6_Btau_5.0.1
Bubalus bubalis	GCA_000471725.1	Mammals	2836.17	42.2	366983	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000471725.1_UMD_CASPUR_WB_2.0
Callithrix jacchus	GCA_000004665.1	Mammals	2914.96	41.3414	16399	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000004665.1_Callithrix_jacchus-3.2
Camelus bactrianus	GCA_000767855.1	Mammals	1992.66	41.4	35455	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000767855.1_Ca_bactrianus_MBC_1.0
Camelus dromedarius	GCA_000767585.1	Mammals	2004.06	41.3	32573	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000767585.1_PRJNA234474_Ca_dromedarius_V1.0
Camelus ferus	GCA_000311805.2	Mammals	2009.19	41.3	13334	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000311805.2_CB1
Canis lupus familiaris	GCA_000002285.2	Mammals	2410.98	47.2	3310	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000002285.2_CanFam3.1
Capra aegagrus	GCA_000765075.1	Mammals	2583.32	42.3	6616	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000765075.1_Caeg1
Capra hircus	GCA_000317765.1	Mammals	2635.85	42.1792	77432	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000317765.1_CHIR_1.0
Capreolus capreolus	GCA_000751575.1	Mammals	2785.38	41.8	3088511	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000751575.1_kmer631
Cavia aperea	GCA_000688575.1	Mammals	2716.4	42.2	3131	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000688575.1_CavAp1.0
Cavia porcellus	GCA_000151735.1	Mammals	2723.22	40.1	3144	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000151735.1_Cavpor3.0
Cebus capucinus imitator	GCA_001604975.1	Mammals	2717.7	41	7156	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001604975.1_Cebus_imitator-1.0
Ceratotherium simum simum	GCA_000283155.1	Mammals	2464.37	41.2	3087	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000283155.1_CerSimSim1.0
Cercocebus atys	GCA_000955945.1	Mammals	2848.25	41.1	11433	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000955945.1_Caty_1.0
Chinchilla lanigera	GCA_000276665.1	Mammals	2390.87	41.4	2839	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000276665.1_ChiLan1.0
Chlorocebus sabaeus	GCA_000409795.2	Mammals	2789.66	40.9303	2022	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000409795.2_Chlorocebus_sabeus_1.1
Choloepus hoffmanni	GCA_000164785.2	Mammals	3286.01	40	269084	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000164785.2_C_hoffmanni-2.0.1
Chrysochloris asiatica	GCA_000296735.1	Mammals	4210.11	41.8	20500	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000296735.1_ChrAsi1.0
Colobus angolensis palliatus	GCA_000951035.1	Mammals	2970.12	41.6	13124	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000951035.1_Cang.pa_1.0
Condylura cristata	GCA_000260355.1	Mammals	1769.66	41.9	2040	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000260355.1_ConCri1.0

#Organism/Name	Assembly	SubGroup	Size (Mb)	GC%	Scaffolds	GenBank FTP
Cricetulus griseus	GCA_000223135.1	Mammals	2399.79	41.6	109152	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000223135.1_CriGri_1.0
Dasypus novemcinctus	GCA_000208655.2	Mammals	3631.52	41.5	46559	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000208655.2_Dasnov3.0
Daubentonia madagascariensis	GCA_000241425.1	Mammals	2855.37	39.6	3231305	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000241425.1_DauMad_1.0
Dipodomys ordii	GCA_000151885.2	Mammals	2236.37	42.6	65193	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000151885.2_Dord_2.0
Echinops telfairi	GCA_000313985.1	Mammals	2947.02	43.6	8402	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000313985.1_EchTel2.0
Eidolon helvum	GCA_000465285.1	Mammals	1837.75	39.2	133538	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000465285.1_ASM46528v1
Elephantulus edwardii	GCA_000299155.1	Mammals	3843.98	41.5	8768	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000299155.1_EleEdw1.0
Eptesicus fuscus	GCA_000308155.1	Mammals	2026.63	43.5	6789	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000308155.1_EptFus1.0
Equus asinus	GCA_001305755.1	Mammals	2391.05	41.4	2167	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001305755.1_ASM130575v1
Equus caballus	GCA_000002305.1	Mammals	2474.93	41.6532	9688	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000002305.1_EquCab2.0
Equus przewalskii	GCA_000696695.1	Mammals	2395.95	41.3	53097	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000696695.1_Burgud
Erinaceus europaeus	GCA_000296755.1	Mammals	2715.72	42.4999	5803	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000296755.1_EriEur2.0
Eulemur flavifrons	GCA_001262665.1	Mammals	2115.57	40.5	38367	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001262665.1_Eflavifronsk33QCA
Eulemur macaco	GCA_001262655.1	Mammals	2119.88	38.3	26772	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001262655.1_Emacaco_refEf_BWA_oneround
Felis catus	GCA_000181335.3	Mammals	2641.34	42.0311	267928	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000181335.3_Felis_catus_8.0
Fukomys damarensis	GCA_000743615.1	Mammals	2333.89	40.5	74730	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000743615.1_DMR_v1.0
Galeopterus variegatus	GCA_000696425.1	Mammals	3187.66	41.2	179514	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000696425.1_G_variegatus-3.0.2
Gorilla gorilla gorilla	GCA_000151905.1	Mammals	3035.66	41.1641	57197	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000151905.1_gorGor3.1
Heterocephalus glaber	GCA_000247695.1	Mammals	2618.2	41.2	4229	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000247695.1_HetGla_female_1.0
Homo sapiens	GCA_000001405.22	Mammals	3232.55	41.4522	805	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000001405.22_GRCh38.p7
Ictidomys tridecemlineatus	GCA_000236235.1	Mammals	2478.39	40.5	12483	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000236235.1_SpeTri2.0
Jaculus jaculus	GCA_000280705.1	Mammals	2835.25	42.7	10898	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000280705.1_JacJac1.0
Leptonychotes weddellii	GCA_000349705.1	Mammals	3156.9	43.8	16711	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000349705.1_LepWed1.0

#Organism/Name	Assembly	SubGroup	Size (Mb)	GC%	Scaffolds	GenBank FTP
Lipotes vexillifer	GCA_000442215.1	Mammals	2429.21	41.4	30713	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000442215.1_Lipotes_vexillifer_v1
Loxodonta africana	GCA_000001905.1	Mammals	3196.74	40.9	2352	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000001905.1_Loxafr3.0
Macaca fascicularis	GCA_000364345.1	Mammals	2946.84	41.3398	7625	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000364345.1_Macaca_fascicularis_5.0
Macaca mulatta	GCA_000772875.3	Mammals	3236.22	41.2015	286263	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000772875.3_Mmul_8.0.1
Macaca nemestrina	GCA_000956065.1	Mammals	2948.7	41.3	9733	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000956065.1_Mnem_1.0
Macropus eugenii	GCA_000004035.1	Mammals	3075.18	40.4	277711	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000004035.1_Meug_1.1
Mandrillus leucophaeus	GCA_000951045.1	Mammals	3061.99	41.6	12821	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000951045.1_Mleu.le_1.0
Manis pentadactyla	GCA_000738955.1	Mammals	2204.73	41.6	92772	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000738955.1_M_pentadactyla-1.1.1
Marmota marmota marmota	GCA_001458135.1	Mammals	2510.59	40.2	14543	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001458135.1_marMar2.1
Megaderma lyra	GCA_000465345.1	Mammals	1735.93	40.3	192872	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000465345.1_ASM46534v1
Mesocricetus auratus	GCA_000349665.1	Mammals	2504.93	43.2	21484	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000349665.1_MesAur1.0
Microcebus murinus	GCA_000165445.2	Mammals	2438.8	41.2	10311	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000165445.2_Mmur_2.0
Microtus agrestis	GCA_001305995.1	Mammals	3124.14	26.4	230202	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001305995.1_ASM130599v1
Microtus ochrogaster	GCA_000317375.1	Mammals	2287.34	42.8312	6450	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000317375.1_MicOch1.0
Miniopterus natalensis	GCA_001595765.1	Mammals	1803.1	42.4	1269	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001595765.1_Mnat.v1
Monodelphis domestica	GCF_000002295.2	Mammals	3598.44	38.1446	5223	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF_000002295.2_MonDom5
Mus musculus	GCA_000001635.6	Mammals	2803.57	41.9419	293	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000001635.6_GRCm38.p4
Mus spretus	GCA_001624865.1	Mammals	2625.59	42.4818	5404	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001624865.1_SPRET_EiJ_v1
Mustela putorius furo	GCA_000239315.1	Mammals	2400.18	41.7	4245	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000239315.1_MusPutFurMale1.0
Myodes glareolus	GCA_001305785.1	Mammals	3443.07	25.8	367242	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001305785.1_ASM130578v1
Myotis brandtii	GCA_000412655.1	Mammals	2107.24	42.9	169750	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000412655.1_ASM41265v1
Myotis davidii	GCA_000327345.1	Mammals	2059.8	43.1	101769	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000327345.1_ASM32734v1
Myotis lucifugus	GCA_000147115.1	Mammals	2034.58	42.7	11654	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000147115.1_Myoluc2.0

#Organism/Name	Assembly	SubGroup	Size (Mb)	GC%	Scaffolds	GenBank FTP
Nannospalax galili	GCA_000622305.1	Mammals	3061.42	41.6	154976	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000622305.1_S.galili_v1.0
Nasalis larvatus	GCA_000772465.1	Mammals	3011.97	42.2536	319549	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000772465.1_Charlie1.0
Nomascus leucogenys	GCA_000146795.3	Mammals	2962.06	41.3959	17524	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000146795.3_Nleu_3.0
Ochotona princeps	GCA_000292845.1	Mammals	2229.84	44	10421	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000292845.1_OchPri3.0
Octodon degus	GCA_000260255.1	Mammals	2995.89	42.5	7135	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000260255.1_OctDeg1.0
Odobenus rosmarus divergens	GCA_000321225.1	Mammals	2400.15	41.7	3893	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000321225.1_Oros_1.0
Odocoileus virginianus	GCA_000191625.1	Mammals	37.6562	45.2	-	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000191625.1_RRL-RSL_1.0
Orcinus orca	GCA_000331955.2	Mammals	2372.92	41.7	1668	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000331955.2_Oorc_1.1
Ornithorhynchus anatinus	GCF_000002275.2	Mammals	1995.61	45.6584	200283	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF_000002275.2_Ornithorhynchus_anatinus_5.0.1
Orycteropus afer afer	GCA_000298275.1	Mammals	4444.08	42.1	22509	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000298275.1_OryAfe1.0
Oryctolagus cuniculus	GCA_000003625.1	Mammals	2737.46	44.0526	3318	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000003625.1_OryCun2.0
Otolemur garnettii	GCA_000181295.3	Mammals	2519.72	41.5	7793	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000181295.3_OtoGar3
Ovis aries	GCA_000298735.2	Mammals	2615.52	41.9473	5466	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000298735.2_Oar_v4.0
Pan paniscus	GCA_000258655.2	Mammals	3286.64	42.3185	10984	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000258655.2_panpan1.1
Pan troglodytes	GCA_000001515.4	Mammals	3309.58	41.8745	27005	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000001515.4_Pan_troglodytes-2.1.4
Panthera tigris altaica	GCA_000464555.1	Mammals	2391.08	41.5	1479	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000464555.1_PanTig1.0
Pantholops hodgsonii	GCA_000400835.1	Mammals	2696.89	42.4	15059	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000400835.1_PHO1.0
Papio anubis	GCA_000264685.1	Mammals	2948.4	41.1229	72501	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000264685.1_Panu_2.0
Peromyscus maniculatus bairdii	GCA_000500345.1	Mammals	2630.54	42.7	30921	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000500345.1_Pman_1.0
Physeter catodon	GCA_000472045.1	Mammals	2280.73	41.3	11711	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000472045.1_Physeter_macrocephalus-2.0.2
Pongo abelii	GCF_000001545.4	Mammals	3441.24	41.5894	79342	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF_000001545.4_P_pygmaeus_2.0.2
Procavia capensis	GCA_000152225.2	Mammals	3602.18	41.8	65694	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000152225.2_Pcap_2.0
Propithecus coquereli	GCA_000956105.1	Mammals	2798.15	43.2	22539	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000956105.1_Pcoq_1.0

#Organism/Name	Assembly	SubGroup	Size (Mb)	GC%	Scaffolds	GenBank FTP
Pteronotus parnellii	GCA_000465405.1	Mammals	1960.32	40.8	177401	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000465405.1_ASM46540v1
Pteropus alecto	GCA_000325575.1	Mammals	1985.98	39.9	65598	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000325575.1_ASM32557v1
Pteropus vampyrus	GCA_000151845.2	Mammals	2198.28	40.5	36094	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000151845.2_Pvam_2.0
Rattus norvegicus	GCA_000001895.4	Mammals	2870.18	42.3282	1395	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000001895.4_Rnor_6.0
Rhinolophus ferrumequinum	GCA_000465495.1	Mammals	1926.44	40.6	160500	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000465495.1_ASM46549v1
Rhinopithecus roxellana	GCA_000769185.1	Mammals	2899.55	41	135512	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000769185.1_Rrox_v1
Rousettus aegyptiacus	GCA_001466805.2	Mammals	1910.25	40	2490	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001466805.2_Raegyp2.0
Saimiri boliviensis boliviensis	GCA_000235385.1	Mammals	2608.59	41.1	2686	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000235385.1_SaiBol1.0
Sarcophilus harrisii	GCA_000189315.1	Mammals	3174.69	37	35974	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000189315.1_Devil_ref_v7.0
Sorex araneus	GCA_000181275.2	Mammals	2423.16	43.4	12845	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000181275.2_SorAra2.0
Sus scrofa	GCA_000003025.4	Mammals	2808.53	42.1	9906	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000003025.4_Sscrofa10.2
Tarsius syrichta	GCA_000164805.2	Mammals	3453.86	41	337189	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000164805.2_Tarsius_syrichta-2.0.1
Trichechus manatus latirostris	GCA_000243295.1	Mammals	3103.81	41.6	6323	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000243295.1_TriManLat1.0
Tupaia belangeri	GCA_000181375.1	Mammals	2137.23	41.4	-	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000181375.1_ASM18137v1
Tupaia chinensis	GCA_000334495.1	Mammals	2846.58	42	50750	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000334495.1_TupChi_1.0
Tursiops truncatus	GCA_000151865.3	Mammals	2551.42	42.1	240558	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000151865.3_Ttru_1.4
Ursus maritimus	GCA_000687225.1	Mammals	2301.38	41.7	23819	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000687225.1_UrsMar_1.0
Vicugna pacos	GCA_000164845.3	Mammals	2172.21	41.7	276725	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000164845.3_Vicugna_pacos-2.0.2
Vicugna pacos huacaya	GCA_000767525.1	Mammals	2013.7	41.5	52275	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000767525.1_Vi_pacos_V1.0
Acanthisitta chloris	GCA_000695815.1	Birds	1035.88	41.6	53875	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000695815.1_ASM69581v1
Amazona aestiva	GCA_001420675.1	Birds	1129.54	42.2	3232	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001420675.1_ASM142067v1
Amazona vittata	GCA_000332375.1	Birds	1175.4	41.7	182974	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000332375.1_AV1
Anas platyrhynchos	GCA_000355885.1	Birds	1105.05	41.2001	78488	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000355885.1_BGI_duck_1.0

#Organism/Name	Assembly	SubGroup	Size (Mb)	GC%	Scaffolds	GenBank FTP
Anser cygnoides domesticus	GCA_000971095.1	Birds	1119.15	41.5001	7593	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000971095.1_AnsCyg_PRJNA183603_v1.0
Apaloderma vittatum	GCA_000703405.1	Birds	1070.84	41.4	54728	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000703405.1_ASM70340v1
Aptenodytes forsteri	GCA_000699145.1	Birds	1254.35	42	10672	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000699145.1_ASM69914v1
Apteryx australis mantelli	GCA_001039765.1	Birds	1523.97	42.5	24720	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001039765.1_AptMant0
Aquila chrysaetos canadensis	GCA_000766835.1	Birds	1192.74	41.9001	1142	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000766835.1_Aquila_chrysaetos-1.0.2
Ara macao	GCA_000400695.1	Birds	1204.7	42.5001	192790	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000400695.1_SMACv1.1
Balearica regulorum gibbericeps	GCA_000709895.1	Birds	1127.62	41.2001	53491	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000709895.1_ASM70989v1
Buceros rhinoceros silvestris	GCA_000710305.1	Birds	1065.78	42.6	62257	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000710305.1_ASM71030v1
Calidris pugnax	GCA_001431845.1	Birds	1229.09	42.7	3753	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001431845.1_ASM143184v1
Calypte anna	GCA_000699085.1	Birds	1105.68	41.3	54736	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000699085.1_ASM69908v1
Caprimulgus carolinensis	GCA_000700745.1	Birds	1119.68	40.8	70122	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000700745.1_ASM70074v1
Cariama cristata	GCA_000690535.1	Birds	1132.25	41.2	53474	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000690535.1_ASM69053v1
Cathartes aura	GCA_000699945.1	Birds	1152.57	41.1	104141	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000699945.1_ASM69994v1
Chaetura pelagica	GCA_000747805.1	Birds	1119.19	41.6	19072	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000747805.1_ChaPel_1.0
Charadrius vociferus	GCA_000708025.2	Birds	1219.86	42	15167	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000708025.2_ASM70802v2
Chlamydotis macqueenii	GCA_000695195.1	Birds	1086.57	41.1	59693	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000695195.1_ASM69519v1
Colinus virginianus	GCA_000599465.1	Birds	1171.86	43	220307	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000599465.1_NB1.1
Colius striatus	GCA_000690715.1	Birds	1075.93	40.9	70188	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000690715.1_ASM69071v1
Columba livia	GCA_000337935.1	Birds	1107.99	41.6001	14923	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000337935.1_Cliv_1.0
Corvus brachyrhynchos	GCA_000691975.1	Birds	1091.31	42	10547	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000691975.1_ASM69197v1
Corvus cornix cornix	GCA_000738735.1	Birds	1049.96	41.7	1299	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000738735.1_Hooded_Crow_genome
Coturnix japonica	GCA_001577835.1	Birds	927.657	41.3684	2531	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001577835.1_Coturnix_japonica_2.0
Cuculus canorus	GCA_000709325.1	Birds	1153.89	41.7	14930	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000709325.1_ASM70932v1

#Organism/Name	Assembly	SubGroup	Size (Mb)	GC%	Scaffolds	GenBank FTP
Egretta garzetta	GCA_000687185.1	Birds	1206.5	42.5	11791	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000687185.1_ASM68718v1
Eurypyga helias	GCA_000690775.1	Birds	1088.02	42.3	62699	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000690775.1_ASM69077v1
Falco cherrug	GCA_000337975.1	Birds	1174.81	41.8	5863	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000337975.1_F_cherrug_v1.0
Falco peregrinus	GCA_000337955.1	Birds	1171.97	41.8	7021	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000337955.1_F_peregrinus_v1.0
Ficedula albicollis	GCA_000247815.2	Birds	1118.34	44.2999	21836	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000247815.2_FicAlb1.5
Fulmarus glacialis	GCA_000690835.1	Birds	1141.4	41.2	57389	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000690835.1_ASM69083v1
Gallus gallus	GCA_000002315.3	Birds	1230.26	42.9197	23870	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000002315.3_Gallus_gallus-5.0
Gavia stellata	GCA_000690875.1	Birds	1129.69	41.1001	61831	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000690875.1_ASM69087v1
Geospiza fortis	GCA_000277835.1	Birds	1065.29	41.7	27239	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000277835.1_GeoFor_1.0
Haliaeetus albicilla	GCA_000691405.1	Birds	1133.55	40.9	50905	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000691405.1_ASM69140v1
Haliaeetus leucocephalus	GCA_000737465.1	Birds	1178.41	41.8	1023	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000737465.1_Haliaeetus_leucocephalus-4.0
Lepidothrix coronata	GCA_001604755.1	Birds	1079.58	41.9	4612	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001604755.1_Lepidothrix_coronata-1.0
Leptosomus discolor	GCA_000691785.1	Birds	1136.24	41.8	57160	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000691785.1_ASM69178v1
Lyrurus tetrix tetrix	GCA_000586395.1	Birds	657.025	40.9	-	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000586395.1_tetTet1
Manacus vitellinus	GCA_000692015.2	Birds	1145.85	41.2	92755	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000692015.2_ASM69201v2
Meleagris gallopavo	GCA_000146605.3	Birds	1128.34	41.7219	233806	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000146605.3_Turkey_5.0
Melopsittacus undulatus	GCA_000238935.1	Birds	1117.37	41.4001	25212	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000238935.1_Melopsittacus_undulatus_6.3
Merops nubicus	GCA_000691845.1	Birds	1062.96	41.7	53499	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000691845.1_ASM69184v1
Mesitornis unicolor	GCA_000695765.1	Birds	1087.29	41.3	67520	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000695765.1_ASM69576v1
Nestor notabilis	GCA_000696875.1	Birds	1053.56	41.1	42180	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000696875.1_ASM69687v1
Nipponia nippon	GCA_000708225.1	Birds	1223.86	42.0001	59555	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000708225.1_ASM70822v1
Opisthocomus hoazin	GCA_000692075.1	Birds	1203.71	42.7	10256	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000692075.1_ASM69207v1
Parus major	GCA_001522545.1	Birds	1020.31	41.4159	1676	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001522545.1_Parus_major1.0.3

#Organism/Name	Assembly	SubGroup	Size (Mb)	GC%	Scaffolds	GenBank FTP
Pelecanus crispus	GCA_000687375.1	Birds	1160.92	41.4	63982	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000687375.1_ASM68737v1
Phaethon lepturus	GCA_000687285.1	Birds	1152.96	41.5	66785	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000687285.1_ASM68728v1
Phalacrocorax carbo	GCA_000708925.1	Birds	1138.97	41.3	64312	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000708925.1_ASM70892v1
Phoenicopterus ruber ruber	GCA_000687265.1	Birds	1132.18	41.9	76189	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000687265.1_ASM68726v1
Picoides pubescens	GCA_000699005.1	Birds	1167.32	44.6	31254	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000699005.1_ASM69900v1
Podiceps cristatus	GCA_000699545.1	Birds	1134.92	41.5	82923	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000699545.1_ASM69954v1
Pseudopodoces humilis	GCA_000331425.1	Birds	1043	41.8001	5406	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000331425.1_PseHum1.0
Pterocles gutturalis	GCA_000699245.1	Birds	1069.32	41.4	58607	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000699245.1_ASM69924v1
Pygoscelis adeliae	GCA_000699105.1	Birds	1216.62	41.8001	19265	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000699105.1_ASM69910v1
Serinus canaria	GCA_000534875.1	Birds	1152.1	42.6	304400	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000534875.1_SCA1
Struthio camelus australis	GCA_000698965.1	Birds	1225.04	41.3	6915	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000698965.1_ASM69896v1
Sturnus vulgaris	GCA_001447265.1	Birds	1036.76	41.7	2361	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001447265.1_Sturnus_vulgaris-1.0
Taeniopygia guttata	GCA_000151805.2	Birds	1232.14	41.4526	37422	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000151805.2_Taeniopygia_guttata-3.2.4
Tauraco erythrolophus	GCA_000709365.1	Birds	1155.54	41.6	59587	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000709365.1_ASM70936v1
Tinamus guttatus	GCA_000705375.2	Birds	1047.06	41.5	82514	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000705375.2_ASM70537v2
Tyto alba	GCA_000687205.1	Birds	1120.14	40.2	62122	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000687205.1_ASM68720v1
Zonotrichia albicollis	GCA_000385455.1	Birds	1052.6	41.8	6018	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000385455.1_Zonotrichia_albicollis-1.0.1
Amphilophus citrinellus	GCA_000751415.1	Fishes	844.903	41.4	6637	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000751415.1_Midas_v5
Anguilla anguilla	GCA_000695075.1	Fishes	1018.7	42.9	501148	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000695075.1_Anguilla_anguilla_v1_09_nov_10
Anguilla japonica	GCA_000470695.1	Fishes	1151.14	43.6	323740	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000470695.1_japanese_eel_genome_v1_25_oct_2011_japonica_c401b400k25m200_sspacepremiumk3a02n24_extra.final.scaffolds
Anguilla rostrata	GCA_001606085.1	Fishes	1413.05	41	79210	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001606085.1_ASM160608v1
Astyanax mexicanus	GCA_000372685.1	Fishes	1191.24	40	10735	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000372685.1_Astyanax_mexicanus-1.0.2
Austrofundulus limnaeus	GCA_001266775.1	Fishes	866.963	41.1	29785	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001266775.1_Austrofundulus_limnaeus-1.0

#Organism/Name	Assembly	SubGroup	Size (Mb)	GC%	Scaffolds	GenBank FTP
Boleophthalmus pectinirostris	GCA_000788275.1	Fishes	955.735	40.1	16619	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000788275.1_BP.fa
Callorhinchus milii	GCA_000165045.2	Fishes	974.499	42.5998	21204	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000165045.2_Callorhinchus_milii-6.1.3
Clupea harengus	GCA_000966335.1	Fishes	807.712	44.5001	6915	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000966335.1_ASM96633v1
Cottus rhenanus	GCA_001455555.1	Fishes	563.609	36.8	164693	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001455555.1_ASM145555v1
Cynoglossus semilaevis	GCA_000523025.1	Fishes	470.199	41.2788	31181	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000523025.1_Cse_v1.0
Cyprinodon nevadensis pectoralis	GCA_000776015.1	Fishes	1011.85	39	96516	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000776015.1_ASM77601v1
Cyprinodon variegatus	GCA_000732505.1	Fishes	1035.18	39.5001	9259	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000732505.1_C_variegatus-1.0
Cyprinus carpio	GCA_000951615.1	Fishes	1713.64	37.3	9377	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000951615.1_common_carp_genome
Danio rerio	GCA_000002035.3	Fishes	1371.72	36.6591	3399	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000002035.3_GRCz10
Dicentrarchus labrax	GCA_000689215.1	Fishes	675.917	40.4	25	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000689215.1_seabass_V1.0
Esox lucius	GCA_000721915.2	Fishes	904.453	42.1896	1709	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000721915.2_ASM72191v2
Fundulus heteroclitus	GCA_000826765.1	Fishes	1021.9	41.2	10180	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000826765.1_Fundulus_heteroclitus-3.0.2
Gadus morhua	GCA_000231765.1	Fishes	824.311	46.3	427427	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000231765.1_GadMor_May2010
Haplochromis burtoni	GCA_000239415.1	Fishes	831.412	41.9	8001	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000239415.1_AstBur1.0
Kryptolebias marmoratus	GCA_001649575.1	Fishes	680.349	37.8	3072	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001649575.1_ASM164957v1
Labeotropheus fuelleborni	GCA_000150875.1	Fishes	70.8584	42.1	58245	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000150875.1_ASM15087v1
Labrus bergylta	GCA_900080235.1	Fishes	805.481	40.9	13466	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_900080235.1_BallGen_V1
Larimichthys crocea	GCA_000972845.1	Fishes	678.922	41.4	6013	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000972845.1_L_crocea_1.0
Lates calcarifer	GCA_001640805.1	Fishes	668.465	40.8	3807	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001640805.1_ASM164080v1
Latimeria chalumnae	GCA_000225785.1	Fishes	2860.59	43	22819	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000225785.1_LatCha1
Lepisosteus oculatus	GCA_000242695.1	Fishes	945.878	40.4083	2106	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000242695.1_LepOcu1
Lethenteron camtschaticum	GCA_000466285.1	Fishes	1030.66	48.1	86125	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000466285.1_LetJap1.0
Maylandia zebra	GCA_000238955.3	Fishes	859.842	41.4002	3555	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000238955.3_M_zebra_UMD1

#Organism/Name	Assembly	SubGroup	Size (Mb)	GC%	Scaffolds	GenBank FTP
Mchenga conophoros	GCA_000150855.1	Fishes	73.4256	41.8	61923	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000150855.1_ASM15085v1
Melanochromis auratus	GCA_000150895.1	Fishes	68.2386	41.5	63297	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000150895.1_ASM15089v1
Miichthys miiuy	GCA_001593715.1	Fishes	619.301	39.3	6294	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001593715.1_ASM159371v1
Neolamprologus brichardi	GCA_000239395.1	Fishes	847.91	42.0001	9099	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000239395.1_NeoBri1.0
Nothobranchius furzeri	GCA_001465895.2	Fishes	1242.52	44.8104	6013	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001465895.2_Nfu_20140520
Notothenia coriiceps	GCA_000735185.1	Fishes	636.614	40.8002	38657	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000735185.1_NC01
Oreochromis niloticus	GCA_000188235.2	Fishes	927.696	41.5012	5910	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000188235.2_Orenil1.1
Oryzias latipes	GCA_000313675.1	Fishes	869.818	42.2741	7307	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000313675.1_ASM31367v1
Pampus argenteus	GCA_000697985.1	Fishes	350.449	38.7	298139	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000697985.1_PamArg1.0
Periophthalmodon schlosseri	GCA_000787095.1	Fishes	679.761	40.2	46662	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000787095.1_PS.fa
Periophthalmus magnuspinnatus	GCA_000787105.1	Fishes	701.697	40	26060	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000787105.1_PM.fa
Petromyzon marinus	GCA_000148955.1	Fishes	885.535	46.8	25005	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000148955.1_Petromyzon_marinus-7.0
Pimephales promelas	GCA_000700825.1	Fishes	1219.33	41.7	73057	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000700825.1_FHM_SOAPdenovo
Poecilia formosa	GCA_000485575.1	Fishes	748.923	39.6	3985	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000485575.1_Poecilia_formosa-5.1.2
Poecilia latipinna	GCA_001443285.1	Fishes	815.145	40.8	17988	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001443285.1_P_latipinna-1.0
Poecilia mexicana	GCA_001443325.1	Fishes	801.711	40.7	18105	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001443325.1_P_mexicana-1.0
Poecilia reticulata	GCA_000633615.2	Fishes	731.622	40.2816	3029	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000633615.2_Guppy_female_1.0_MT
Pundamilia nyererei	GCA_000239375.1	Fishes	830.133	41.9	7236	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000239375.1_PunNye1.0
Rhamphochromis esox	GCA_000150935.1	Fishes	71.2951	42.3	55751	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000150935.1_ASM15093v1
Salmo salar	GCA_000233375.4	Fishes	2966.89	43.8912	241573	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000233375.4_ICSASG_v2
Scartelaos histophorus	GCA_000787155.1	Fishes	695.009	39.1	156044	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000787155.1_SH.fa
Scleropages formosus	GCA_001624265.1	Fishes	777.343	44	4818	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001624265.1_ASM162426v1
Sebastes nigrocinctus	GCA_000475235.1	Fishes	687.55	40.6	138020	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000475235.1_Snig1.0
Sebastes rubrivinctus	GCA_000475215.1	Fishes	756.297	40.7	68206	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000475215.1_SRub1.0

#Organism/Name	Assembly	SubGroup	Size (Mb)	GC%	Scaffolds	GenBank FTP
Sinocyclocheilus anshuiensis	GCA_001515605.1	Fishes	1632.72	38.0001	85682	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001515605.1_SAMN03320099.WGS_v1.1
Sinocyclocheilus grahami	GCA_001515645.1	Fishes	1750.29	38.7	31277	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001515645.1_SAMN03320097.WGS_v1.1
Sinocyclocheilus rhinocerous	GCA_001515625.1	Fishes	1655.79	38.1	164173	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001515625.1_SAMN03320098_v1.1
Stegastes partitus	GCA_000690725.1	Fishes	800.492	42.1	5818	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000690725.1_Stegastes_partitus-1.0.2
Takifugu flavidus	GCA_000400755.1	Fishes	378.032	45.6	34332	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000400755.1_version_1_of_Takifugu_flavidus_genome
Takifugu rubripes	GCA_000180615.2	Fishes	391.485	45.8414	7091	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000180615.2_FUGU5
Tetraodon nigroviridis	GCA_000180735.1	Fishes	342.403	46.6	25773	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000180735.1_ASM18073v1
Thunnus orientalis	GCA_000418415.1	Fishes	684.497	39.7	-	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000418415.1_Thunnus_orientalis_ver_Ba_1.0
Xiphophorus couchianus	GCA_001444195.1	Fishes	708.396	40.9	25	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001444195.1_Xiphophorus_couchianus-4.0.1
Xiphophorus hellerii	GCA_001443345.1	Fishes	733.802	41.2	25	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_001443345.1_Xiphophorus_hellerii-3.0.1
Xiphophorus maculatus	GCA_000241075.1	Fishes	729.664	39.8002	20632	ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA_000241075.1_Xiphophorus_maculatus-4.4.2