Genomics of *Aspidoscelis* whiptail lizards

Dissertation
Zur Erlangung des Grades
Doktor der Naturwissenschaften

Am Fachbereich Biologie
Der Johannes Gutenberg-Universität Mainz

David Vuong Ho

Mainz, 2024

Tag der mündlichen Prüfung: 14.11.2024

Table of Contents

List of Figures

List of Tables

List of Equations

Acknowledgements

1. Abstract

Whiptail lizards of the genus *Aspidoscelis* exhibit remarkable reproductive diversity, encompassing sexual reproduction, facultative parthenogenesis, and obligate parthenogenesis in a single genus. With close to 50 species, about which one-third are unisexual and reproduce through obligate parthenogenesis, whiptail lizards are a good model system for investigating how species can survive without the genetic variation typically introduced through sexual reproduction. Through sequencing and bioinformatic analysis, high-quality genome assemblies for five different whiptail lizards were produced. The first part of the thesis describes facultative parthenogenesis in whiptail lizards, shedding light on the mechanisms underlying this reproductive strategy using a high-quality genome assembly of *A. marmoratus*. The second part details the elevation of this assembly to chromosome-level and the genome assemblies for four additional gonochoristic species (*A. arizonae, A. tigris, A. gularis,* and *A. burti*). These assemblies not only are important to the understanding of diversity among gonochoristic species but also serve as a reference to explore the genomics of *A. neomexicanus*, a unisexual species of hybrid origin between *A. marmoratus* and *A. arizonae*. Genome sequencing reveals that rare occurrences of syntenic chromosome pairing leading to loss of heterozygosity in *A. neomexicanus*, leading one to question the long-term survival of the species. These genomes also provide a foundation for investigating whiptails of higher ploidy levels. The work here highlights the importance of genomics in studying whiptail lizards and will advance our understanding about the diversity within the genus and role of parthenogenesis in vertebrate evolution.

Die Rennechsen (oder „Whiptails“) der Gattung Aspidoscelis zeigen eine bemerkenswerte reproduktive Vielfalt, die sexuelle Fortpflanzung, fakultative Parthenogenese und obligate Parthenogenese in einer einzigen Gattung umfasst. Mit fast 50 Arten, von denen etwa ein Drittel unisexuell ist und sich durch obligate Parthenogenese fortpflanzt, sind Rennechsen ein gutes Modellsystem, um zu untersuchen, wie Arten ohne die genetische Variation überleben können, die normalerweise durch sexuelle Fortpflanzung eingeführt wird. Durch Sequenzierung und bioinformatische Analyse wurden hochwertige Genomassemblierung für fünf verschiedene Rennechsen erstellt. Der erste Teil der Arbeit beschreibt die fakultative Parthenogenese bei Rennechsen und beleuchtet die Mechanismen, die dieser Fortpflanzungsstrategie zugrunde liegen, anhand einer hochwertigen Genomassemblierung von A. marmoratus. Der zweite Teil beschreibt die Erhebung dieser Assembly auf Chromosomenebene und die Genomassemblierung für vier weitere gonochoristische Arten (A. arizonae, A. tigris, A. gularis und A. burti). Diese Genomassemblierung sind nicht nur wichtig für das Verständnis der Vielfalt unter gonochoristischen Arten, sondern dienen auch als Referenz, um die Genomik von A. neomexicanus, einer unisexuellen Art hybriden Ursprungs zwischen A. marmoratus und A. arizonae, zu erforschen. Die Genomsequenzierung zeigt, dass seltene Vorkommen von syntenischer Chromosomenpaarung zu einem Verlust der Heterozygotie in A. neomexicanus führen, was Fragen über das langfristige Überleben der Art aufwirft. Diese Genomassemblierungen bilden auch eine Grundlage für die Untersuchung von Rennechsen mit höherem Ploidiegrad. Die hier vorgestellte Arbeit unterstreicht die Bedeutung der Genomik bei der Untersuchung von Rennechsen und wird unser Verständnis der Vielfalt innerhalb der Gattung und der Rolle der Parthenogenese in der Evolution der Wirbeltiere erweitern.

2. Introduction

Parthenogenesis and unisexuality

Parthenogenesis, derived from the Greek words "parthenos" meaning "virgin" and "genesis" meaning "birth" or “creation,” is a form of asexual reproduction that has independently evolved across the tree of life, including plants, invertebrates, and vertebrates. A parthenogen produces offspring from an unfertilized egg. There are several forms of parthenogenesis, each with different implications for genetic diversity and evolutionary dynamics.

Facultative parthenogenesis is a form of reproduction where females of a species are able to reproduce both sexually and asexually. Depending on the sex determining mechanism, parthenogenetic offspring of ZW females are typically male (ZZ) and offspring of XX females are always female as well (Booth et al., 2010; Chapman et al., 2007; W. W. Olsen & Marsden, 1954; Watts et al., 2006). This reproductive strategy observed is observed various taxa and with many reports within the last two decades due to molecular techniques such as microsatellite analysis and sequencing that can confirm that the genetic inheritance of the offspring only comes from the mother. Taxa in which facultative parthenogenesis have been reported in include insects (Normark, 2003; Sperling et al., 2023), reptiles (Booth et al., 2010, 2023; Dubach et al., 1997; Kinney et al., 2013; Lenk et al., 2005; Mandáková et al., 2014), birds (M. W. Olsen, 1973; W. W. Olsen & Marsden, 1954; Ryder et al., 2021), and fish (Chapman et al., 2007, 2008; Dudgeon et al., 2017; Feldheim et al., 2010; Lampert, 2008).

Obligate parthenogenesis on the other hand occurs in species there are only females present, often referred to in the broader sense as “unisexual reproduction.” Among vertebrates, fish, amphibians, and reptiles all have representative species that are all female that reproduce more daughters (Neaves & Baumann, 2011). However, it is important to distinguish that even though these fish and amphibians species are all female, they are still dependent on the sperm from a related sexual species. The sperm in these cases do not contribute to the genetics of the progeny but stimulate embryogenesis like in the case of gynogenesis (Lampert & Schartl, 2008). In hybridogenesis, the paternal genome from the sperm contribute to the somatic development of the hybrid offspring, however the paternal genome is eliminated in early oogenesis resulting in haploid eggs with the maternal genome (R. C. Vrijenhoek et al., 1977). Only in reptiles is there true parthenogenesis where females are independent to reproduce and there is no contribution from any males (Figure 1; Fujita & Moritz, 2010).

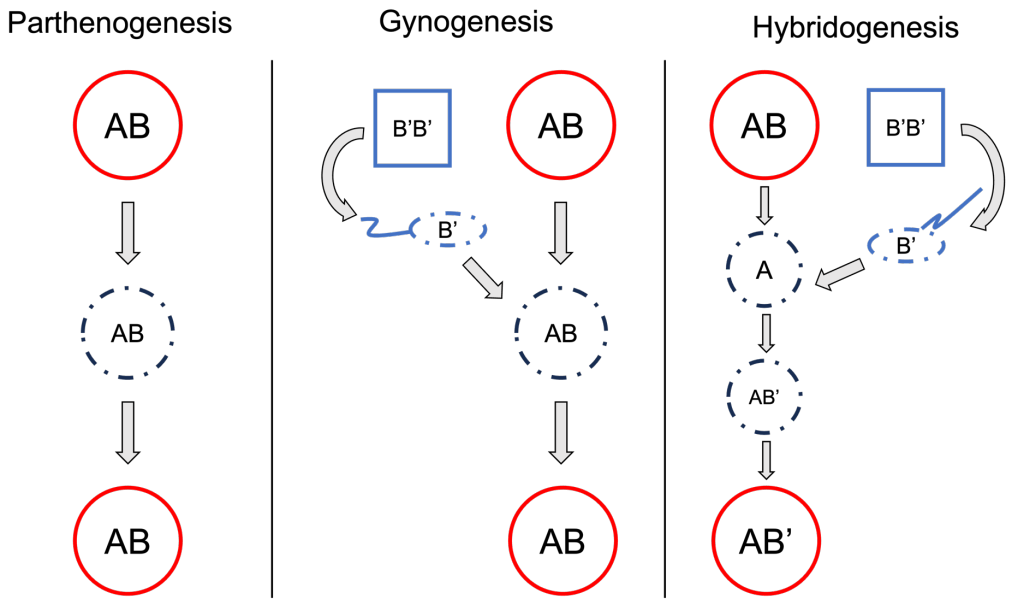

Figure 1. Comparison of different forms of parthenogenesis. In true parthenogenesis, there is no sperm requirement, and diploid oocytes develop directly into diploid offspring with the same genotype to the mother. Gynogenesis and hybridogenesis require sperm. In gynogenesis, a diploid oocyte is produced and the fertilization by a sperm from a male of a closely-related species is required for embryogenesis. However, the sperm does not contribute any genetic material and is subsequently eliminated. In hybridogenesis, only the genome A contributes to the oocyte, while the sperm from a closely-related species restore diploidy and allows for embryogenesis. In the next generation, the same genome A is retained.

One unifying aspect of unisexual reproduction in vertebrates is its hybrid origin. In all cases, with one exception (the Neotropical night lizard *Lepidophyma,* Sinclair et al., 2010), two closely related gonochoristic (or sometimes referred to as "bisexual") species breed and successfully produce viable offspring that can subsequently reproduce parthenogenetically (Neaves & Baumann, 2011). Without the need for males, true parthenogenesis would be postulated to be beneficial for population maintenance and expansion versus their sexual counterparts. Parthenogenesis allows every individual in the population, which are all female, to reproduce without the need or competition for mate-choice, which in theory can double its growth rate compared to a sexually reproducing population.

Despite these apparent advantages that parthenogens have, the prevalence of sexual reproduction is still widespread. Sexual reproduction offers the advantage of generating genetic diversity, which is crucial for adaptation to changing environments. Although there are additional costs such as mate-choice and the production of males that do not directly bear offspring, sexual reproduction generates new combinations of alleles through genetic recombination, allowing populations to adapt more effectively to environmental changes and evolving pathogens (Pomiankowski, 1987; Stapley et al., 2017). The success of sexually reproducing species in a variety of ecological niches suggests that

the benefits of generating genetic diversity often outweigh the costs associated with sexual reproduction (Hadany & Comeron, 2008; Lehtonen et al., 2012).

Unisexual species that undergo true parthenogenesis forgo these benefits, relying on mechanisms that bypass the need for a mate. One on hand, unisexuality can lead to short-term reproductive success, on the other hand, it imposes long-term evolutionary costs. In these populations, harmful mutations can accumulate over generations because there is no mechanism for recombination to produce offspring free of these mutations. Each time a new deleterious mutation arises, it is passed on and becomes fixed in the population over time, as defined by Muller's rachet. Over time, this accumulation can lead to a decline in fitness, making unisexual populations more vulnerable to extinction (Muller, 1964). The Red Queen hypothesis offers an additional disadvantage of unisexuals. Species must continually adapt and evolve to keep up with a changing environment and pathogens in a constant evolutionary "arms race" (Van Valen, 1973). In unisexual populations, the production of genetically identical offspring, for example, makes them easy targets for pathogens, which can evolve rapidly to eliminate genetically identical individuals.

The existence and persistence of unisexuality though challenges the notion that sexual reproduction is a universal necessity for the maintenance of species. There are several "asexual scandals" in nature where unisexual lineages have defied expectations and thrived for extended periods. The bdelloid rotifers, microscopic animals that have persisted 35-40 million years without sexual reproduction, are a classic example (Birky, 2004; Waggoner & Poinar, 1993). Among the close to 100 vertebrate species (Avise, 2008; R. Vrijenhoek et al., 1989), the salamander *Ambystoma* is estimated to have an origin of over five million years ago (Bi & Bogart, 2010).

Due to their hybrid origins, unisexual vertebrate species inherit two (or more) distinct genomes immediately when the initial hybridization occurs. The maintenance of these subgenomes in subsequent generations translates to high heterozygosity. Because distinct alleles are inherited, this genetic addition can lead to hybrid vigor and may allow it to compete effectively with its progenitors. This mirrors what is frequently observed in plants, where hybridization often leads to offspring that have phenotypes that are better than either parent (Birchler et al., 2006; Liu et al., 2020). The high heterozygosity observed in unisexual vertebrates is believed to also contribute significantly to their long-term persistence in nature.

While rarer than their sexual counterparts, studying and exploring these unisexual lineages provides valuable insights into the myriad ways in which species can evolve. These species challenge the conventional idea that sex is necessary and poses questions that are important for genetics, biodiversity, and conservation.

Lizards of the genus *Aspidoscelis*

Whiptail lizards were originally named and grouped with the South American genus *Cnemidophorus* (Burt, 1931). In 2002, there was a reclassification that led to the separation of the North American whiptail lizards into their own genus *Aspidoscelis* (“shield-leg”), supported by mitochondrial DNA, allozyme, and morphological characters (Reeder et al., 2002). To date, there are about 50 recognized whiptail species, with about one-third of them being unisexual species, derived from hybridizations between the gonochoristic species (Barley et al., 2022). Even though both gonochoristic and unisexual species have been described as early as the 19th century, it was not until the late 1950s and 1960s where there were publications to actually identify that some species were actually unisexual, owing to the fact that no males were ever found (Duellman & Zweifel, 1962; Maslin, 1962). These observations became more convincing in light that Darevsky (1958) had also described an all-female unisexual species found in Caucasus a few years earlier.

Within vertebrates, whiptail lizards have the highest count of unisexual lineages (http://www.reptile-database.org; v. 04 September 2024). Even before the recognition of unisexual lineages, whiptail species were already a problem for taxonomist to identify species boundaries, due to their diverse phenotypic variation within and between populations: “Most of the ‘species’ are so plastic, so variable, that they may well drive the systematist to despair. No two taxonomic authorities will, or can, possibly agree upon the. Number of admissible species” (Gadow, 1906). Adding now to the fact that unisexuals are derived from hybridizations of the gonochoristic species complicates that further. To this point about hybridization: not all whiptail hybridizations result in a unisexual lineage and sometimes the outcome is introgression (Barley et al., 2022). As such, to simplify our understanding of this, we will limit our discussion to just the outcome when a unisexual lineage or speciation event is produced.

Since the identification of unisexual whiptail species, considerable investigations using morphological traits, karyotypes, biochemical assays, and mitochondria comparisons have been done to identify the gonochoristic progenitors that hybridized to give rise to these unisexuals (Cole, 1985; Cole et al., 1988; Dessauer & Cole, 1984; C. H. Lowe & Wright, 1966; Neaves, 1969; Neaves & Gerald, 1969; Reeder et al., 2002). Strikingly, while gonochoristic species are all diploid, unisexuals species of diploidy, triploidy, and tetraploidy exist (Figure 2). In the case of tetraploids, single caught individuals were identified in the past (C. J. Lowe et al., 1970; Neaves & Neaves, 1971). But since, laboratory attempts to make tetraploids have either created sterile offspring or in three successful cases, tetraploid lineages that continue to reproduce in the lab (Charles J. Cole et al., 2023; Cole et al., 2017; Hardy & Cole, 1998; A. Lutes et al., 2011). It is not too alarming to assume that tetraploid

lineages exist in the wild as well, though cryptic to biologists without molecular tools to confidently identify them.

These increases in ploidy levels are not unique to just whiptail lizards but observed in other species as well (Lampert et al., 2007; Saccucci et al., 2016; Wynn et al., 1987). As discussed earlier, unisexuals of hybrid origin have high heterozygosity and therefore secondary and tertiary hybridizations that elevate ploidy levels would be hypothesized to increase heterozygosity levels to compensate for lack of recombination from sex. While unisexual whiptails are able to reproduce on their own without male contribution, in some instances, the female unisexual can incorporate the haploid genome from a gonochoristic male into her unreduced oocyte and consequential produces an offspring with an elevated ploidy level. One could argue that these ploidy elevation instances mirror the edge cases in gynogenesis where there is paternal leakage and incorporation of the paternal genome into the offspring (Lampert & Schartl, 2010). While these species are all still considered obligate parthenogens, there seemingly appears to be very much plasticity in reproduction that can introduce gene flow.

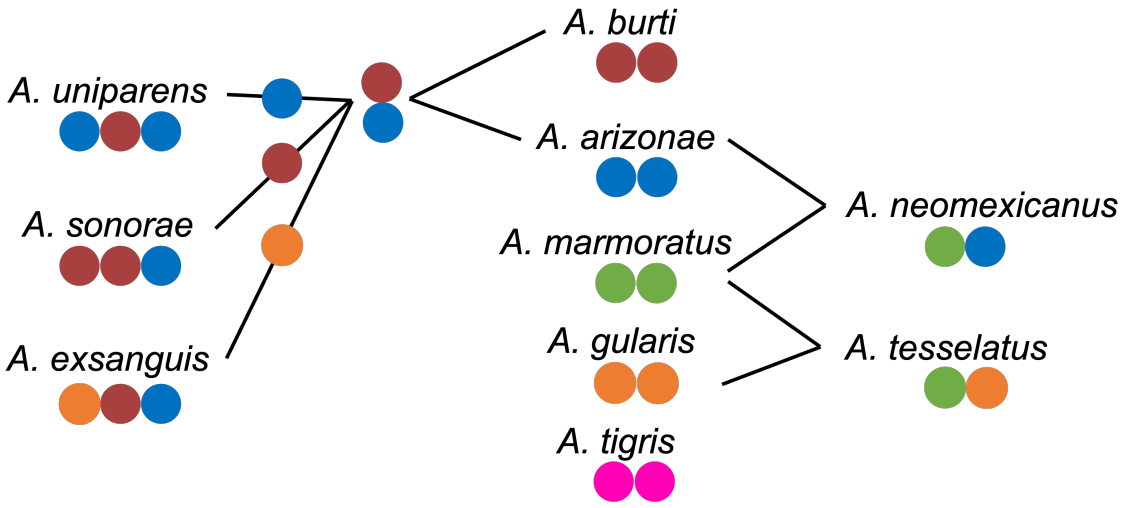

Figure 2. Schematic of hybridizations between gonochoristic *Aspidoscelis* species that resulted in diploid and triploid unisexual lineages. Not all hybridizations and unisexual lineages are represented. Each colored circle represents a haploid genome. Species with more than one color is indicative of a hybrid and the colors indicate the identity of the subgenomes that make up the hybrid.

Because all daughters are clones of their unisexual mothers, it is important to consider how the traditional meiotic process, which is reductive in nature, is different to meiosis in sexual species to pass on the complete subgenomes (and in turn maintain heterozygosity). One hypothesis would be that meiosis is suppressed and productive eggs are produced from a mitotic-like pathway to ensure the daughter egg is identical to the mother (Simon et al., 2003). What was shown by the Baumann lab though is that oocyte

production does not deviate significantly from the normal meiosis pathway (Figure 3). Unisexual lizards initiate meiosis with twice the number of chromosomes compared to their gonochoristic reproducing counterparts, resulting in an oocyte that still have the key number of chromosomes to initiate embryo development. Recombination still occurs, however since there are twice the number of chromosomes pairing partners are between identical chromosomes, rather than syntenic or homologous chromosomes between the subgenomes. By pairing and recombining identical chromosomes, this preserves heterozygosity (A. A. Lutes et al., 2010; Newton et al., 2016).

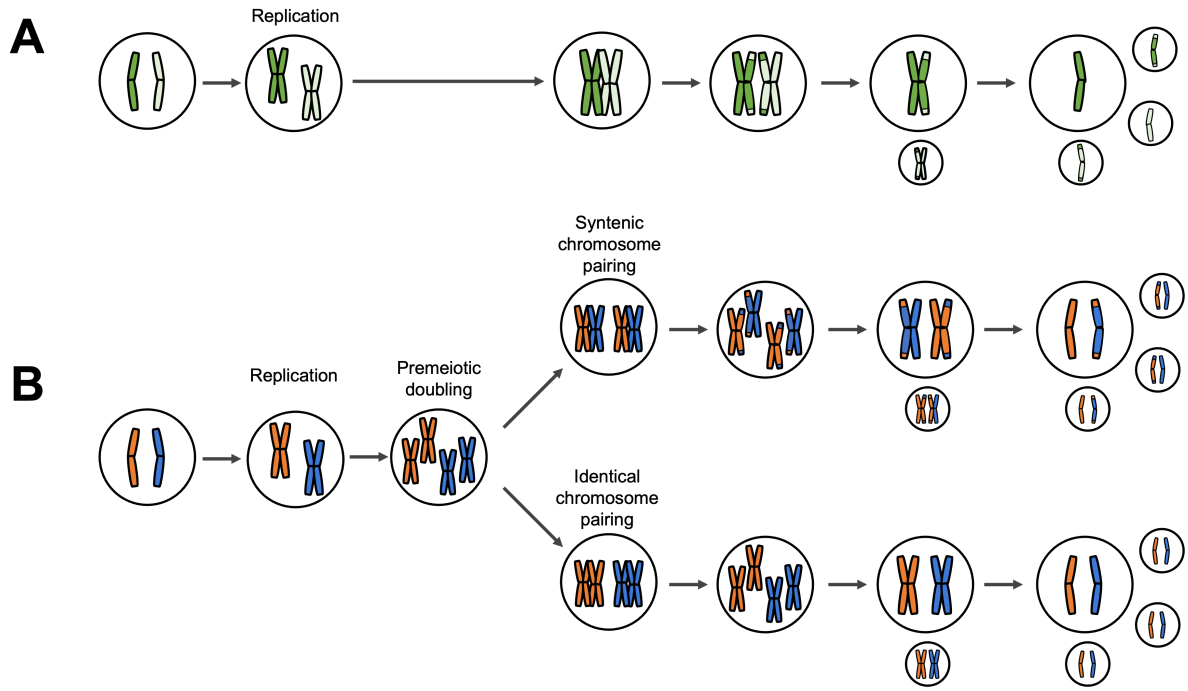

Figure 3. Comparison of canonical meiosis and two potential meiosis pathways that occurs in unisexual *Aspidoscelis* whiptails. (A) Canonical meiosis results in a haploid oocyte and three haploid polar bodies. (B) An additional premeiotic doubling ensures that canonical meiosis can still occur but the result is an oocyte with a complete set of chromosomes that can proceed to embryogenesis. Either syntenic chromosomes can pair resulting in oocytes that have loss of heterozygosity or identical chromosomes pair, ensuring maintenance of heterozygosity and the offspring being genetically identical to its mother. Adapted from Neaves & Baumann (2011).

Studying *Aspidoscelis*

From a theoretical point of view, parthenogenesis and unisexuality appear to be an evolutionary dead-end. However, recent studies showing maintenance of heterozygosity (A. A. Lutes et al., 2010) and further polyploidization to increase heterozygosity (Charles J. Cole et al., 2023) calls into question if these lineages are truly short-lived. Since hybridization is a

unifying theme within unisexual vertebrates, studying this in a genus that contains so many examples makes it a compelling model system for investigating unisexuality in other groups.

In addition to unisexual lineages, introgression between gonochoristic species have led to an overwhelming diversification of the group (Barley et al., 2022). While many species conform to Ernst Mayr's biological species concept, where species is defined as members of populations that actually or potentially interbreed, whiptail lizards appear to be an outlier. Constant gene flow between "species" reiterates the complexity within the genus and more data on their genomes need to be acquired to further comparative studies.

Genome assemblies are essential for advancing our understanding of how similar whiptail species are to each other and what are the genetic consequences of hybridization and unisexuality that cannot be answered solely with morphological and biochemical assays. In the past two decades, medicine and agriculture have benefitted from genome assemblies that facilitate the identification of genetic variants associated with diseases and assist in the selection of desirable traits in crop and livestock breeding. However, these early genome assemblies were not cheap and considerable resources were needed to even produce a usable assembly.

With the decreasing costs in sequencing, some have term the period we are in now as the "post-genomic era." Despite their importance, there is still a significant lack of high-quality genome assemblies, particularly for non-model organisms. In 2022, less than 1% of the submitted genome assemblies to the National Library of Medicine (NCBI) were from reptile taxa (Figure 4). Of the submitted reptile genomes, less than one-fifth were of chromosome-level scale.

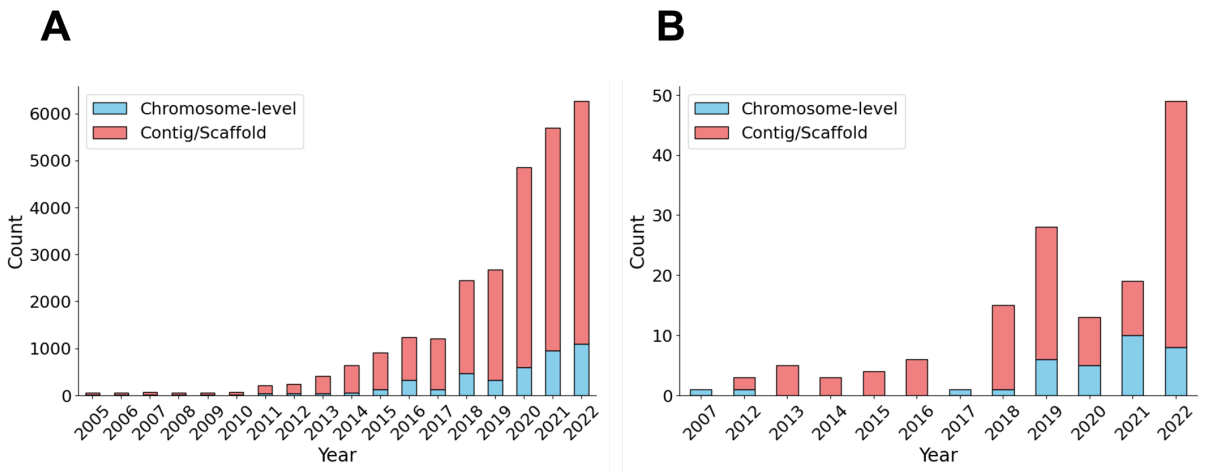

Figure 4. Genome assemblies submitted to NCBI. (A) Total and (B) reptile genome assemblies submitted per year to NCBI.

Some major factors contributing to this dearth is the technical challenges associated with genome sequencing and assembly. Many non-model organisms have large size genomes that are difficult to assemble due to the presence of repetitive sequences and polyploidy in some species presents an additional challenge.

In this thesis, I will present work that highlights use of sequencing and bioinformatics to further understand the genomics of whiptail lizards. The first part of a manuscript that describes for the first time facultative parthenogenesis in whiptail lizards and emphasizes how a high-quality genome assembly of *A. marmoratus* sheds light on the mechanisms underlying this reproductive strategy. The second part describes elevating the *A. marmoratus* genome assembly to chromosome-level and how we leverage Hi-C sequencing data to generating genome assemblies for four addition gonochoristic species (*A. arizonae, A. tigris, A. gularis,* and *A. burti*). Not only will these genome assemblies aid in whiptail researchers to understand the diversity among gonochoristic species, in the third part, I show how combining the *A. marmoratus* and *A. arizonae* genome assemblies together as a proxy for the unisexual *A. neomexicanus* provides insights into the genomics of unisexuals.

3. Facultative parthenogenesis in the genus *Aspidoscelis*

This chapter is the manuscript entitled “Post-meiotic mechanism of facultative parthenogenesis in gonochoristic whiptail lizard species” that has been the hard work of many contributing authors since the discovery of facultative parthenogenesis within the Baumann lab colony in two species: *A. marmoratus* and *A. arizonae*. This manuscript has gone through one round of revisions by the Refereed Preprint service *Review Commons* and after two rounds of revisions by the journal e*Life*. It was accepted and published in June 2024.

This body of work encompasses not only genomics and bioinformatic analysis, but also cytology, microscopy, and dedicated animal husbandry. As such, I thank my co-authors for laying the foundation for the data presented in this manuscript. My contributions include further analysis and combining the results together in a cohesive manner to be shared. The

figure numbers are self-contained to this chapter only. Figure supplements and supplementary files/tables come after the main text. The date presented in Figure 5-source data 1 can be assessed online at the journal’s website.

“Post-meiotic mechanism of facultative parthenogenesis in gonochoristic whiptail lizard species”

RESEARCH ARTICLE

Post-meiotic mechanism of facultative parthenogenesis in gonochoristic whiptail lizard species

David V Ho[1,2†], Duncan Tormey[3†‡], Aaron Odell[1], Aracely A Newton[3§], Robert R Schnittker[3], Diana P Baumann[3], William B Neaves[3], Morgan R Schroeder[3], Rutendo F Sigauke[3], Anthony J Barley[4], Peter Baumann[1,2,5*]

[1]Department of Biology, Johannes Gutenberg University, Mainz, Germany; [2]Institute of Quantitative and Computational Biosciences, Johannes Gutenberg University, Mainz, Germany; [3]Stowers Institute for Medical Research, Kansas City, United States; [4]School of Mathematical and Natural Sciences, Arizona State University–West Valley Campus, Glendale, United States; [5]Institute of Molecular Biology, Mainz, Germany

Abstract Facultative parthenogenesis (FP) has historically been regarded as rare in vertebrates, but in recent years incidences have been reported in a growing list of fish, reptile, and bird species. Despite the increasing interest in the phenomenon, the underlying mechanism and evolutionary implications have remained unclear. A common finding across many incidences of FP is either a high degree of homozygosity at microsatellite loci or low levels of heterozygosity detected in next-generation sequencing data. This has led to the proposal that second polar body fusion following the meiotic divisions restores diploidy and thereby mimics fertilization. Here, we show that FP occurring in the gonochoristic *Aspidoscelis* species *A. marmoratus* and *A. arizonae* results in genome-wide homozygosity, an observation inconsistent with polar body fusion as the underlying mechanism of restoration. Instead, a high-quality reference genome for *A. marmoratus* and analysis of whole-genome sequencing from multiple FP and control animals reveals that a post-meiotic mechanism gives rise to homozygous animals from haploid, unfertilized oocytes. Contrary to the widely held belief that females need to be isolated from males to undergo FP, females housed with conspecific and heterospecific males produced unfertilized eggs that underwent spontaneous development. In addition, offspring arising from both fertilized eggs and parthenogenetic development were observed to arise from a single clutch. Strikingly, our data support a mechanism for facultative parthenogenesis that removes all heterozygosity in a single generation. Complete homozygosity exposes the genetic load and explains the high rate of congenital malformations and embryonic mortality associated with FP in many species. Conversely, for animals that develop normally, FP could potentially exert strong purifying selection as all lethal recessive alleles are purged in a single generation.

Editor's evaluation

In this valuable paper, convincing evidence is provided for the production of facultatively parthenogenetic whiptail lizards through a gametic duplication. The audience for the work will be broad, given that parthenogenesis is such a fascinating topic.

***For correspondence:** peter@baumannlab.org

[†]These authors contributed equally to this work

Present address: [‡]Synthego Corporation, Redwood City, United States; [§]Missouri Western State University, Saint Joseph, United States

Competing interest: The authors declare that no competing interests exist.

Funding: See page 17

Preprinted: 22 September 2023
Received: 18 February 2024
Accepted: 17 May 2024
Published: 07 June 2024

Reviewing Editor: Detlef Weigel, Max Planck Institute for Biology Tübingen, Germany

Introduction

Incidences of facultative parthenogenesis (FP) have been reported to occur in diverse vertebrate clades including bony fish (***Lampert et al., 2007***), sharks (***Chapman et al., 2007***; ***Chapman et al., 2008***; ***Feldheim et al., 2010***; ***Dudgeon et al., 2017***), snakes (***Dubach et al., 1997***; ***Groot et al., 2003***; ***Booth and Schuett, 2011***; ***Germano and Smith, 2010***; ***Booth et al., 2012***; ***Kinney et al., 2013***; ***Booth et al., 2014***; ***Allen et al., 2018***; ***Card et al., 2021***), lizards (***Lenk et al., 2005***; ***Watts et al., 2006***; ***Kratochvíl et al., 2020***), crocodilians (***Booth et al., 2023***), and birds (***Bartelmez and Riddle, 1924***; ***Olsen and Marsden, 1954***; ***Sarvella, 1973***; ***Schut et al., 2008***; ***Ramachandran and McDaniel, 2018***; ***Parker et al., 2010***). The phenomenon was originally mistaken for long-term sperm storage occurring in zoo environments where females were housed without current or recent access to conspecific males. The most parsimonious explanation was, therefore, that the animal had previ- ously been in contact with a male and that stored sperm was responsible for delayed fertilization (***Booth and Schuett, 2011***; ***Holt and Lloyd, 2010***; ***Sever and Hamlett, 2002***). However, more recent studies involving microsatellite (MS) and/or amplified fragment length polymorphism (AFLP) analyses revealed no paternal contributions, as all alleles detected in the offspring were only found in the maternal ancestors (***Groot et al., 2003***; ***Booth and Schuett, 2011***; ***Shibata et al., 2017***; ***Ryder et al., 2021***; ***Levine et al., 2024***). Females with no access to males producing solely male (ZW systems) or female (XY systems) offspring that only harbor maternal genetic markers are now considered hallmarks of facultative parthenogenesis. Nevertheless, clear examples of long-term sperm storage have also been documented in the recent literature (***Levine et al., 2021***), underscoring the need for molecular methods such as MS analysis or sequencing data to elucidate the underlying mechanisms. Originally thought to only occur in captivity, more recent reports indicate that FP occurs in natural populations as well (***Booth et al., 2012***; ***Fields et al., 2015***). Serious concerns have been raised by conservation biologists, as species with dwindling population densities, including the endangered species Komodo dragon (***Watts et al., 2006***), small tooth sawfish (***Fields et al., 2015***), California condor (***Ryder et al., 2021***), and American crocodile (***Booth et al., 2023***) are overrepresented among reports of FP.

While overrepresentation could be a consequence of an increased likelihood of detection in species that are the subject of intense research and conservation efforts, the observations raise the question if FP is an adaptive trait aiding in the colonization of new areas and mitigating the effects of popula- tion bottlenecks or is simply a neutral trait (***Fields et al., 2015***). The adaptive trait hypothesis would of course require successful reproduction of FP animals either sexually or parthenogenetically, which to date has only been documented in a few cases (***Kratochvíl et al., 2020***; ***Straube et al., 2016***). At the same time, the association of FP with increased homozygosity constitutes a concern for conserva- tion biology, as an increase in FP within dwindling populations further accelerates the loss of genetic diversity, exposes deleterious alleles, and could compromise efforts to maintain the existing gene pool in selective breeding programs (***Groot et al., 2003***; ***Booth and Schuett, 2016***). FP is also being studied as a desirable outcome in the commercial production of poultry. However, examination of tens of thousands of unfertilized eggs from several different avian species and strains has not resulted in economically sustainable hatching rates thus far. One of the highest hatch rates for unfertilized eggs is seen in the Beltsville small white turkey with a rate of 0.88% (***Ramachandran and McDaniel, 2018***; ***Olsen, 1975***). A better understanding of the triggers and molecular mechanisms underlying FP and the fitness of the resulting offspring are, therefore, needed in a variety of contexts. These include: to understand a fundamental biological mechanism and its significance in vertebrate evolution, to aid in conservation efforts including captive breeding programs, and to possibly harness FP in an agricul- tural context (***Ryder et al., 2021***).

The high level of homozygosity observed in animals produced by FP has been interpreted as evidence for polar body fusion following meiosis II, also known as automixis, leading to the restoration of diploidy in unfertilized eggs (***Figure 1A***; ***Chapman et al., 2007***; ***Card et al., 2021***; ***Booth et al., 2023***; ***Booth and Schuett, 2016***; ***Reynolds et al., 2012***). If automixis involves the fusion of one of the meiotic products from the first polar body (central automixis), homozygosity will be concentrated near the chromosome ends and heterozygosity will be preferentially retained near the centromeres as premeiotic recombination strongly favors homologs over sister chromatids and homologs segre- gate during the first meiotic division. In contrast, a second polar body fusion (terminal automixis) would reunite sister chromatids, for which heterozygosity is preferentially seen near the chromosome termini (***Figure 1B***). In many cases, heterozygous and homozygous loci appeared to be inherited in FP

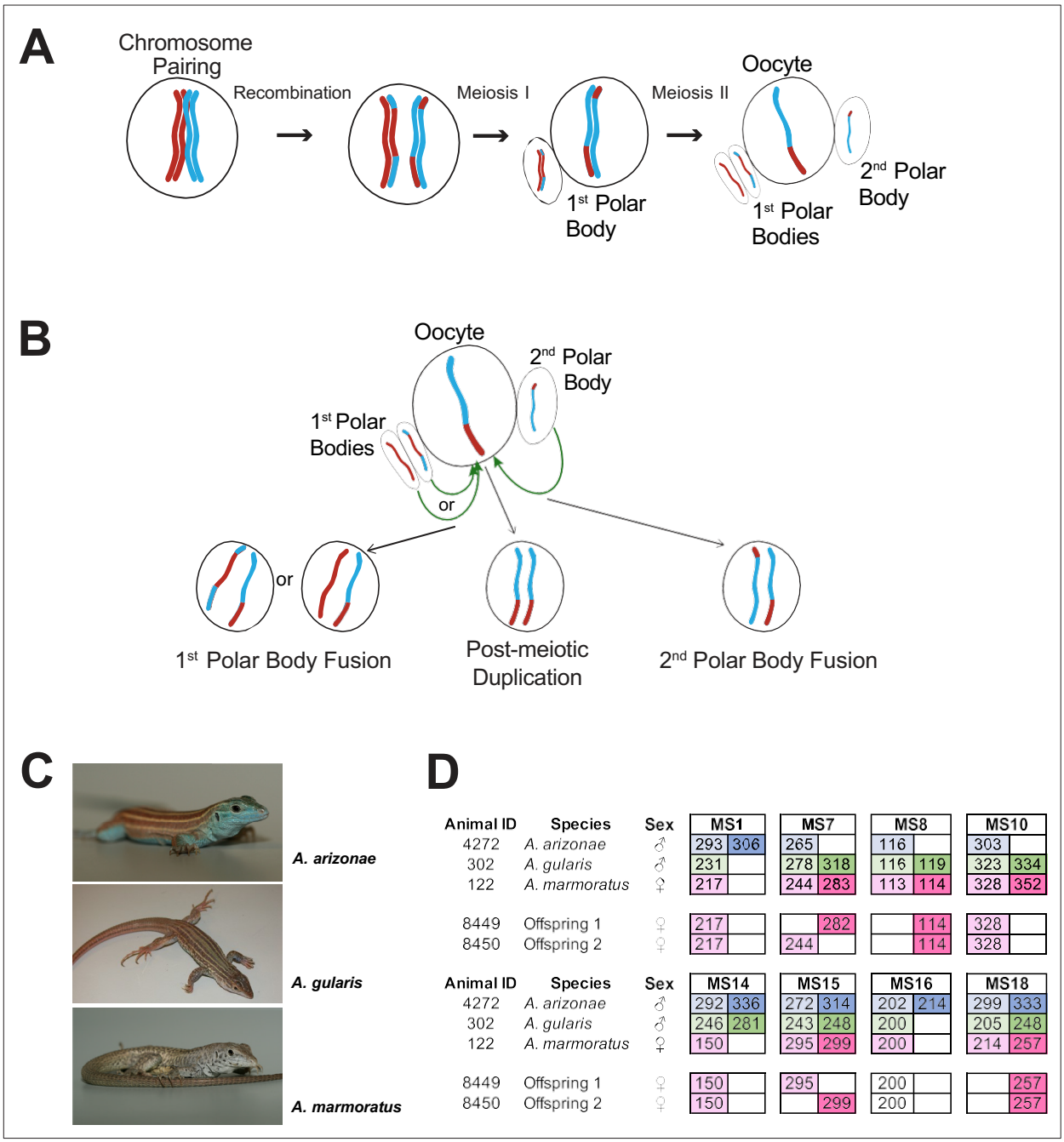

Figure 1. Overview. (**A**) Schematic of canonical meiosis. Only one pair of homologous chromosomes is shown using red and blue to distinguish homologs. (**B**) Schematic of main mechanisms by which a diploid oocyte may be produced in the context of facultative parthenogenesis. First polar body fusion, second polar body fusion, or post-meiotic duplication of chromosomes in the haploid gamete. (**C**) Photographs of *Aspidoscelis arizonae* with characteristic blue ventral coloration (top), *A. gularis* with light spots in dark fields that separate light stripes on dorsum (middle), and *A. marmoratus* with light and dark reticulated pattern on dorsum (bottom). (**D**) Microsatellite analysis for the three co-housed animals and two offspring (ID 8449 and 8450) produced in this enclosure. Alleles are color-coded as follows: *A. arizonae* male (blue), *A. gularis* male (green), and *A. marmoratus* female (red). Differences in shading highlight the two alleles at heterozygous loci. Both offspring are homozygous at all loci with most alleles matching only maternal alleles. For MS16, offspring alleles are not shaded because of this allele being shared between the mother and the *A. gularis* male. Single nucleotide differences in size are common binning artifacts and, therefore, are not scored as different alleles.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. Microsatellite analysis of eight loci for the three female *Aspidoscelis marmoratus* within the enclosure and the two offspring hatched in January 2009.

offspring (**Groot et al., 2003**; **Allen et al., 2018**; **Card et al., 2021**; **Kratochvíl et al., 2020**; **Booth et al., 2023**), but information as to the genomic location of these loci has been lacking. Central and terminal automixis are also distinguished by the extent of heterozygosity, with lower levels observed in next-generation sequencing data in snakes and crocodiles suggesting terminal automixis as the likely mechanism (**Allen et al., 2018**; **Card et al., 2021**; **Booth et al., 2023**).

While facultative parthenogenesis occurs in a wide range of vertebrate species, true obligate parthenogenesis is limited to a few taxa of squamate reptiles including the North American whiptail lizards of the genus *Aspidoscelis* (**Avise, 2015**). Historic hybridization events between distinct gonochoristic species in this clade has given rise to numerous hybrid individuals with the ability to reproduce clonally as all female lineages (**Vanzolini, 1993**; **Reeder et al., 2002**). In contrast to the increased homozygosity associated with FP, obligate parthenogenetic species are characterized by the long-term preservation of the high degree of heterozygosity that had its origin in the lineage-founding cross-species hybridization events.

Our laboratory has a longstanding interest in the mechanism of obligate parthenogenesis in whiptail lizards (**Lutes et al., 2010**; **Newton et al., 2016**). In this context, we are maintaining and propagating individuals of several obligate parthenogenetic as well as gonochoristic species. MS analysis revealed over 20 incidences of FP in the marbled whiptail lizard, *A. marmoratus* and the Arizona striped whiptail, *A. arizonae.* (The taxonomy of the genus *Aspidoscelis* has undergone frequent revisions and the maternal ancestor of the obligate parthenogenetic species *A. neomexicanus* was formerly known as the subspecies *A. tigris marmoratus* and the male ancestor as the subspecies *A. inornatus arizonae* (**Barley et al., 2022a**). For the purpose of this manuscript we follow the taxonomic conclusions by **Barley et al., 2021**). Whiptail lizards have an XX/XY sex determination system (**Cole et al., 1969**) and all FP offspring are consequently female. The identification of multiple incidences of FP provided us with the opportunity to investigate the mechanism of FP in whiptail lizards through next-generation sequencing. The generation of a genome assembly, in addition to whole-genome sequencing, allowed us to distinguish between different mechanisms for restoring diploidy in FP animals. To address the question of whether FP is limited to animals in captivity, we examined reduced-representation sequencing (RAD-seq) data of 321 whiptail lizards from 15 gonochoristic species sampled in nature. In aggregate, a combination of MS analysis, next-generation sequencing, and cytological analysis allows us to report on both the evidence and mechanism of FP in whiptail lizards and suggest that a baseline incidence of FP may coexist alongside sexual reproduction in some species.

Results

Identification of FP in *A. marmoratus*

In the context of studying interspecific hybridization among gonochoristic species of whiptail lizards, three female *A. marmoratus* (ID 122, 4238, 4239) were housed with a male *A. arizonae* (ID 4272) and a male *A. gularis* (ID 302) for close to three years. During this period, seven hybrid offspring between *A. marmoratus* 122 and *A. arizonae* 4272 were produced and confirmed by MS analysis. These animals will be described in more detail in due course. Surprisingly, two female hatchlings emerged that resembled *A. marmoratus* rather than the expected products of hybridization with either *A. arizonae* or *A. gularis*. Genotyping revealed only a single allele for each of eight MS markers in the two offspring (ID 8449 and 8450, **Figure 1D**) and identified *A. marmoratus* (ID 122) as the mother (**Figure 1—figure supplement 1**). The mother and the male *A. arizonae* were each heterozygous at five of the eight markers and the *A. gularis* male at six. Further supporting a uniparental origin of 8449 and 8450, all alleles found in the offspring were also present in the mother (**Figure 1D**). For seven of the eight markers, neither male shared the allele found in the hatchling lizards, providing strong evidence that neither male fathered the offspring. For the remaining marker MS16, the *A. marmoratus* mother and the *A. gularis* male were homozygous for the same allele found in the two offspring, therefore, not allowing a conclusion to be based on this locus. It is important to note that for two of the markers (MS7 and MS15), the two offspring inherited different alleles from the mother, indicating that they are not genetically identical to each other, but have randomly inherited one of the maternal alleles at each locus.

Two additional eggs (ID 8394 and 9070) were recovered from the same enclosure and found to contain developing embryos. MS analysis also revealed *A. marmoratus* 122 as the mother and

complete homozygosity at all loci. Given that all of these offspring are female, inherited only maternal alleles, and animal 122 had no history of being housed with a conspecific male during its lifetime, both interspecific hybridization and long-term sperm storage are all but ruled out and FP is strongly supported.

FP animals of *A. marmoratus* presented a unique opportunity to examine the underlying molecular mechanism. The observed homozygosity at all MS loci further promised to aid in the generation of a high-quality genome assembly, as homozygosity circumvents the challenge of collapsing haplotypes into a consensus sequence (***Kajitani et al., 2014***). To increase homozygosity, inbreeding for 15–20 generations is common practice prior to whole-genome sequencing and genome assembly (***Zhang et al., 2019***; ***Fang et al., 2012***). However, generation times of more than one year make this a costly and time-consuming strategy for many vertebrate species including *A. marmoratus*.

Genome sequencing and de novo assembly

The *A. marmoratus* genome is distributed over 23 chromosomes as previously demonstrated by metaphase spread analysis (***Lowe and Wright, 1966***). We used flow cytometry to compare nuclear DNA content of *A. marmoratus* erythrocytes from whole blood with cells from three species with well-characterized genome sizes. The nuclear DNA content of *A. marmoratus* was close to that of *Danio rerio* (1.4 Gb) and we calculated a haploid genome size for *A. marmoratus* of 1.55 Gb (***Figure 2—figure supplement 1A***).

Genomic DNA of FP animal 8450 was used to generate short insert paired-end, mate-pair (5 Kb, 8 Kb, 2–15 Kb, 40 Kb), and Chicago (***Putnam et al., 2016***) libraries for Illumina short-read sequencing. The paired-end and mate-pair reads were first assembled with Meraculous (***Chapman et al., 2011***) yielding an N50 of 1.6 Mb. The subsequent addition of Chicago reads and scaffolding with the HiRise pipeline by Dovetail Genomics produced an assembly of 1,639,530,780 bp distributed over 3826 scaffolds (***Supplementary file 1***) and raised the scaffold N50 to 32.22 Mb (***Figure 2—figure supplement 1B***). With a BUSCO completeness score of 96% the *A. marmoratus* genome assembly is comparable to other recently released reptilian genome assemblies (***Figure 2—figure supplement 1C***). Over 98% of the assembled sequences are contained within 90 scaffolds of more than 1 Mb in length, making this assembly highly contiguous.

Phylogenetic analysis of shared BUSCO genes with several other reference genomes (*Xenopus*, zebrafish, medaka, platyfish, tegu, green anole, chicken, mouse, rat, dog, cow, human) confirmed that *A. marmoratus* is most closely related to the tegu *Salvator merianae*, another representative of the family Teiidae (***Figure 2—figure supplement 2***). As transposable elements are a driving force in genome evolution, we examined the repeat content for the *A. marmoratus* genome. All classes of repeat elements combined amounted to 40.27% of the *A. marmoratus* assembly, only slightly below that found in other lizards *S. merianae* and *Anolis carolinensis* (***Figure 2—figure supplement 3A***). Strikingly, unclassified repeats make up the largest class of repeat elements in the *A. marmoratus* genome, an observation that parallels findings in *S. merianae*. However, a comparison between the unclassified repeats found in *A. marmoratus*, *S. merianae,* and *A. carolinensis* revealed few similar- ities with only around 10% of the unclassified repeats shared between *A. marmoratus* and *S. meri- anae*, and no significant overlap between these two species and *A. carolinensis* (***Figure 2—figure supplement 3B***). While further characterization of the unclassified repeat elements is needed, it is apparent that an impressive expansion of novel repeat element classes has occurred within this clade.

To annotate the *A. marmoratus* genome, we assembled a total of 119,728 transcripts from RNA-seq data generated from blood and embryo using Trinity (***Supplementary file 2***). These transcripts were subsequently used in the MAKER2 gene annotation pipeline, yielding 25,856 protein-coding genes and 44,461 protein isoforms (***Supplementary file 3***). To assign a putative function, we used BLASTp to query the UniProtKB/Swiss-Prot database (***UniProt Consortium, 2018***) and found significant hits for 76% of the putative protein-coding genes. Our assembly and annota- tion pipelines yielded 40 *HOX* genes and 2 *EVX* genes in four gene clusters (***Figure 2—figure supplement 4***). The *HOX* gene clusters are highly conserved among tetrapods and their complete presence and shared order within each cluster serve as a measure of assembly quality (***Kuraku and Meyer, 2009***).

Assessment of heterozygosity

The presence of only one allele for each of the examined MS markers already suggested widespread homozygosity in *A. marmoratus* produced by FP, consistent with similar observations in other vertebrate species. The highly contiguous genome assembly now afforded us the opportunity to probe the mechanism of FP by searching for regions of heterozygosity and mapping their relative genomic locations. Towards this aim, we performed whole-genome sequencing for an additional nine animals: four of them produced by FP (ID 12512, 12513, 6993, 9177), two mothers (ID 122, 9721), and three unrelated control animals (ID 003, 001, S30700; ***Supplementary file 4***). Each mother and the controls were heterozygous at several MS markers confirming their origin through sexual reproduction. Following alignment to the reference genome (***Supplementary file 5***), we defined heterozygous sites as those covered by an even number of reads with two alleles supported by the same number of reads. Sites covered by an odd number of reads were filtered out for this initial analysis. This stringent requirement was chosen to limit the search to apparent heterozygous sites with strong support, decreasing the chance of false positives.

For all ten individuals, the average sequencing coverage ranged between 15.91 and 20.08 (***Figure 2—figure supplement 5***). For FP animals, the number of heterozygous sites in a 10 Kb sliding window approaches zero for all sites with mean coverage (***Figure 2A***). For positions with coverage greater than the average, an increase in apparent heterozygosity was observed, due to the collapse of repetitive sequences during the assembly process. Based on this observation, we limited further analysis to positions in the genome where the coverage is equal to the mean sequencing depth (as defined by rounding the mean sequencing coverage value to the next even integer). For example, for animal 003, the average sequencing coverage is 18.31 (***Figure 2—figure supplement 5A***) and we only considered sites with a coverage of 20 (***Figure 2—figure supple- ment 6A***). This generated between 30,769 and 53,416 heterozygous sites for which two alleles were equally supported in the sexually produced mothers and control animals (***Figure 2—figure supple- ment 6A–E***). Far fewer heterozygous sites (between 649 and 928) were observed in the FP animals (***Figure 2—figure supplement 6F–J***). Plotting the heterozygous sites according to their position in the reference genome illustrates not only their sparsity in the genomes of FP animals, but also reveals their random distribution (***Figure 2B***). If FP genomes were the product of automixis, regions of homozygosity would be interspersed with regions of heterozygosity. The extent of heterozygosity within the latter would be the same as that observed in the respective mother. The few apparently heterozygous sites identified in FP animals are, therefore, not supporting either form of automixis but are most likely the result of over-assembly of repetitive regions (ie. collapsing paralogous loci into a single representative sequence) and a combination of biological (ie. somatic mutations) and technical errors (ie. PCR and sequencing errors).

When examining each mother-daughter group, the average number of heterozygous sites per 10 Kb window was greater in the sexually produced mothers across the entire assembly (***Figure 2C***). For most of the assembly, the average number of heterozygous sites remained close to zero for FP animals. The most notable exception was a region around genomic coordinate 1.0 Gb, but even there the extent of apparent heterozygosity remained below 50% of what is observed in this region for each of the mothers. Examination of the scaffold in question (Scaffold 45) revealed 167 genes annotated as homologous to *vomeronasal 2 receptor 26* (*Vmn2r26;* ***Figure 2—figure supplement 7***, ***Figure 2—figure supplement 8***). Members of this subfamily of receptors are found on the microvilli of the vomeronasal organ, where they are responsible for pheromone detection and play a signifi- cant role in social and environmental responses (***Ryba and Tirindelli, 1997***; ***Houck, 2009***; ***Su et al., 2009***). Given that this genomic region harbors a large cluster of highly similar genes, the most parsi- monious explanation for the elevated level of apparent heterozygosity is over-assembly. This conclu- sion is further supported by the increase in apparent heterozygosity in this region for the mothers and control animals. In aggregate, our analysis strongly supports genome-wide homozygosity for FP animals, inconsistent with either central or terminal automixis. Instead, the results favor a post-meiotic mechanism that restores diploidy by replicating the haploid genome residing in the oocyte following completion of the two meiotic divisions and thereby establishing genome-wide homozygosity in the offspring.

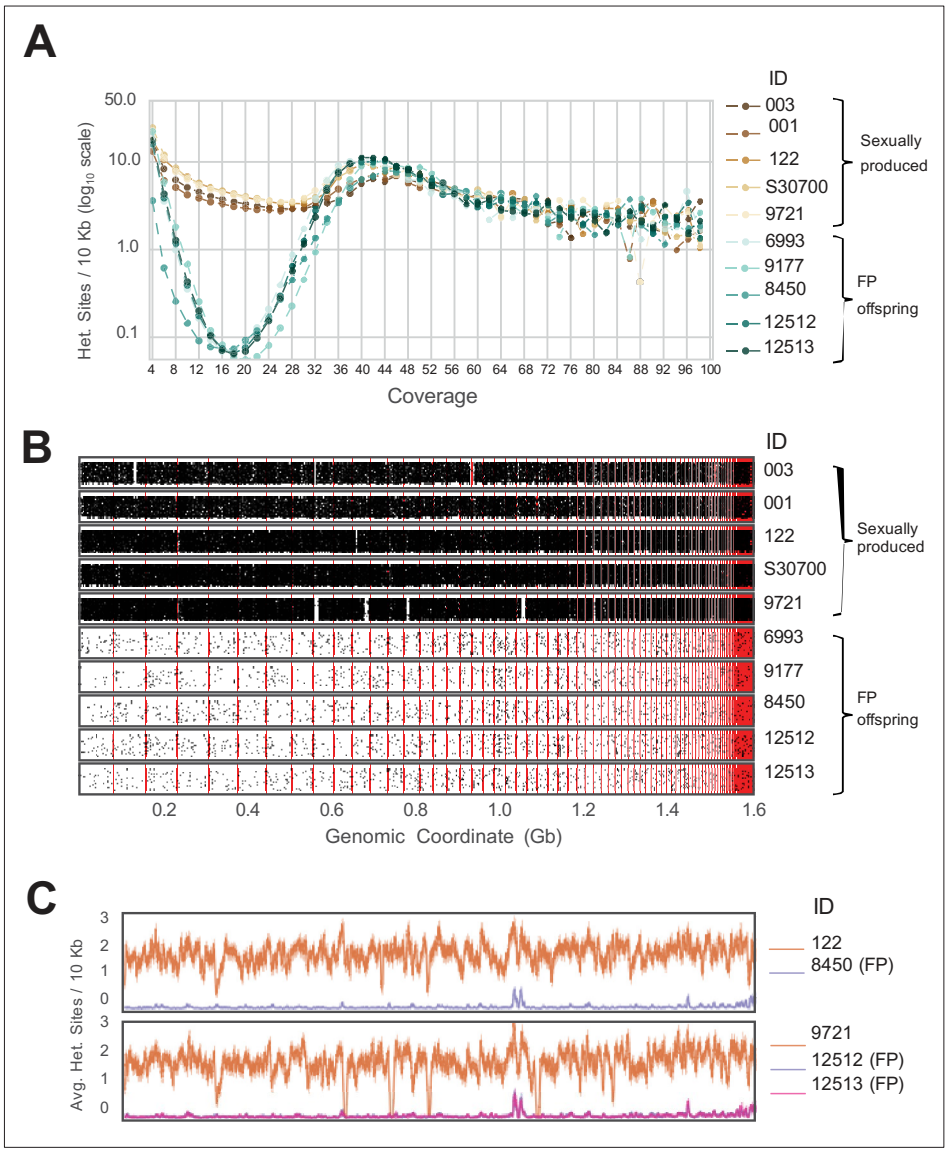

Figure 2. Genome-wide homozygosity in animals produced by facultative parthenogenesis. (**A**) Effect of coverage on the apparent rate of heterozygosity based on evenly split read counts supporting two alleles. Analysis of whole-genome sequencing data for five sexually produced animals (ID 003, 001, 122, S30700, 9721) and five individuals produced by facultative parthenogenetic (FP) animals (ID 6993, 9177, 8450, 12512, 12513) were aligned to the reference genome. In FP animals, the number of heterozygous sites approaches zero for sites with mean coverage ($\bar{x}$=18.37). (**B**) Scaffolds are ordered from largest to smallest on the x-axis. Red lines indicate borders between ordered scaffolds. Each black dot represents a heterozygous position in the genome defined by having a sequencing coverage equal to the average and equal support for only two alleles. The y-axis position of each data point is a random value between bounds of area shown to spread the data and better illustrate the density of heterozygous sites. (**C**) Average heterozygous sites, as defined in (**B**), per 10 Kb window for mothers (orange) and respective FP daughters (purple and pink).

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Genome assembly of *Aspidoscelis marmoratus*.

Figure supplement 2. Maximum likelihood tree for 13 vertebrate genomes, based on 1333 single-copy BUSCOs detected across all species analyzed.

Figure supplement 3. Identification of unclassified repeat elements in the *A. marmoratus* genome.

Figure supplement 4. *Aspidoscelis marmoratus* HOX gene clusters.

Figure supplement 5. Distribution of sequence coverage across the genome for each animal.

Figure 2 continued on next page

Figure 2 continued

Figure supplement 6. Analysis of all positions in the genome at average coverage, for which reads support exactly two alleles.

Figure supplement 7. A 1Mb sliding window for Scaffold 45 showing the annotated vomeronasal 2 receptors homologs.

Figure supplement 8. Phylogenetic tree showing the evolutionary distance between mouse V2Rs (red branches) and *Aspidoscelis marmoratus* homologs (black branches).

Cryptic FP in *A. arizonae*

Following the identification of several *A. marmoratus* generated by FP, we genotyped individuals from two other gonochoristic species housed in our laboratory. While no cases of FP were identified among 80 *A. gularis* produced eggs in captivity, we identified eight incidences of FP among 832 *A. arizonae* records between October 2007 and July 2018. During the same period, we recorded 15 incidences of FP among 286 *A. marmoratus* records (***Supplementary file 6***). Notably, in all cases, eggs undergoing FP development had been laid in enclosures where females were housed with conspecific males or males of a sister species known to mate with the heterospecific females. Isolation from mating partners was thus not a significant factor in triggering FP. In one enclosure, three *A. arizonae* females (ID 12850, 12851, 12852) were housed with a conspecific male (ID 12849; ***Figure 3A***). MS analysis of four hatchlings that originated from a single clutch laid in this enclosure identified animal 12852 as the mother of all four animals. Unexpectedly, two of her offspring were homozygous at all eight loci examined, whereas the two others were heterozygous at all loci, identifying animal 12849 as their father (***Figure 3B***). Therefore, both fertilized and unfertilized eggs developed alongside each other within the same clutch.

Mixoploid erythrocytes and developmental defects

Microscopic examination of blood from a newly hatched FP lizard revealed a striking bimodality in the sizes of erythrocyte nuclei when compared to a sexually produced animals (***Figure 4A, B***). Nuclear size correlates well with DNA content measurements (***Walker et al., 1991***), suggesting the presence of mixoploidy in the FP animal. Whereas most cells closely resembled those observed in the blood from the sexually produced animal, approximately 10% of red blood cells from the FP animal harbored smaller nuclei, consistent with half the amount of DNA (***Figure 4B***). In addition, 1.27% contained two small nuclei, indicating that the final cytokinesis during erythrocyte differentiation had failed for some haploid progenitor cells. DNA content analysis by flow cytometry confirmed the presence of haploid cells (***Figure 4C***). In the blood of sexually produced animals, no haploid or bi-nucleated cells were observed. These observations raise the possibility that the embryonic development of FP animals is initiated with consecutive divisions of a haploid, unfertilized oocyte. At a later stage in development, diploid cells most likely arise via failed cytokinesis. From that point forward, both haploid and diploid cells coexist, and the embryo develops in a mixoploid state. Indicative of a more widespread phenom- enon, mixoploidy was also observed in another FP *A. marmoratus* and in *A. arizonae* (***Figure 4— figure supplement 1***). The observed fraction of haploid cells was closer to 1% in these instances.

Genome-wide loss of heterozygosity exposes functionally compromised alleles that were previously covered by intact alleles on the homologous chromosomes. Depending on the extent of this genetic load, one would expect a substantial fraction of oocytes to not develop at all or for defects to manifest at various stages of embryonic and post-embryonic development. Indeed, of the 23 incidences of FP examined here, only 14 hatched, while the remaining lizards died in ovum (***Supplementary file 6***). For these nine unhatched eggs, we isolated developed embryos shortly after the expected hatch date and confirmed FP origin by MS analysis. The clutches that produced the 23 confirmed cases of FP contained an additional 24 eggs. For these, development did not initiate or terminated at an earlier stage of development precluding MS analysis. Based on the uncertainty regarding how many of these eggs underwent partial FP development, the incidence of FP may be even higher than reported here.

The observation of various malformations in several of the FP embryos and hatchlings further supports that genome-wide homozygosity unmasks deleterious alleles. Notable developmental defects included craniofacial abnormalities such as misaligned jaws, agenesis of eyes, missing limbs, and failed abdominal closure (***Figure 4—figure supplement 2***). Only six out of 16 FP animals (37.5%)

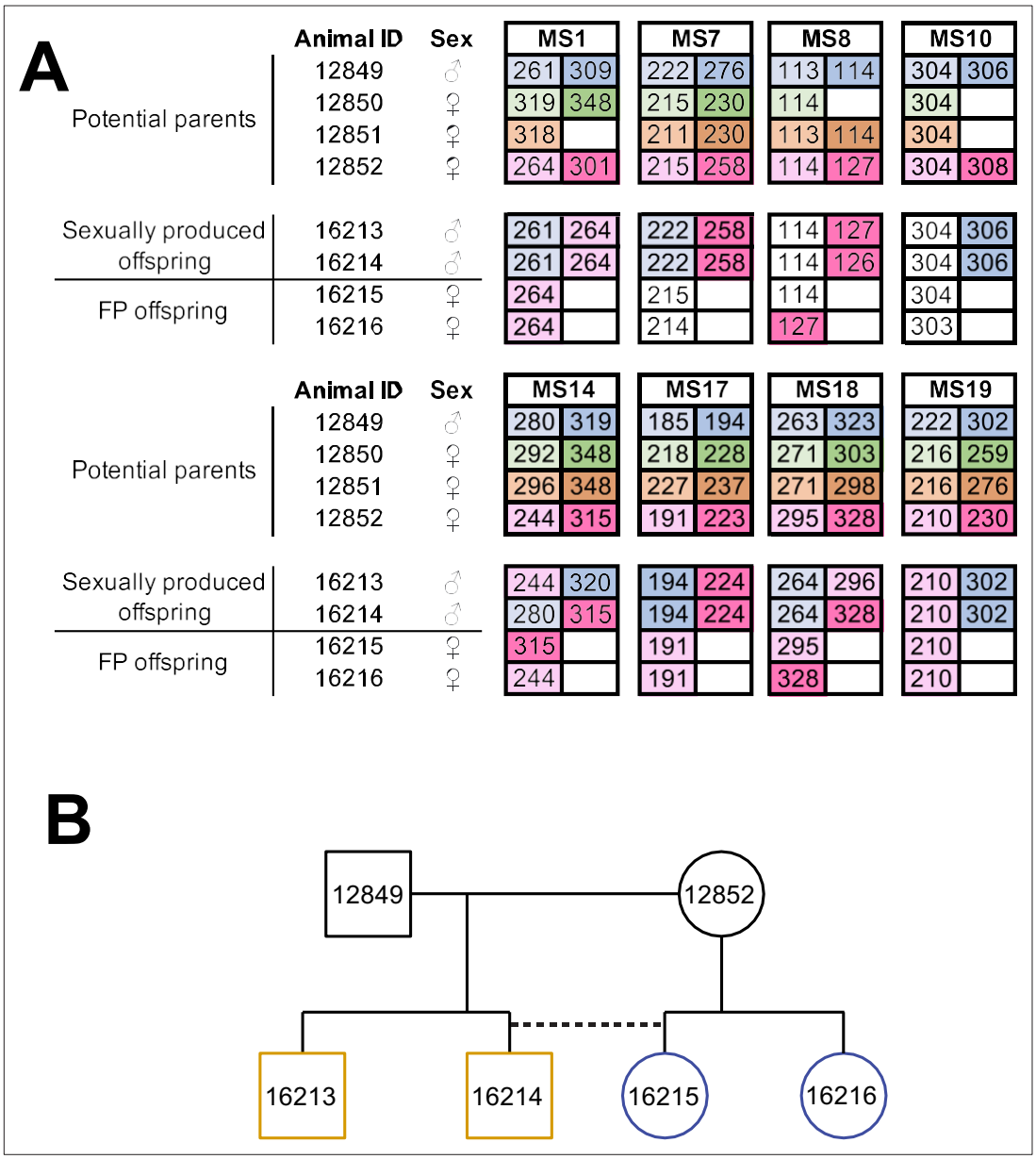

Figure 3. Facultative parthenogenesis is also found in *Aspidoscelis arizonae*. (**A**) Microsatellite analysis for the four co-housed adult animals (ID 12849, 12850, 12851, 12852) and the four hatchlings (ID 16213, 16214, 16215, 16216) produced in this enclosure. Alleles are color-coded for each potential parent: 12849 male (blue), 12850 female (green), 12851 female (orange), 12852 female (red). Differences in shading highlight the two alleles at heterozygous loci. Offspring 16213 and 16214 are heterozygous at all loci, with most loci having one allele matching 12849 and one allele matching 12852. Offspring 16215 and 16216 are homozygous at all loci, with most alleles matching only the 12852 female. Non-shaded offspring alleles indicate ambiguous inheritance as multiple adult animals share the same allele. Single nucleotide differences in size are common binning artifacts and, therefore, are not scored as different alleles. (**B**) Pedigree shows the relationship between the four offspring. The single clutch of four contains both sexually (yellow) and facultative parthenogenetically (blue) produced offspring.

hatched with no discernable developmental defects (***Figure 4—figure supplement 2*** *A-B*). This is in stark contrast to sexually produced animals, where over 98% of hatchlings (n=687) showed no abnormalities. Additionally, most of the defects noted in sexually produced animals were less severe than in FP animals including bulges in tails or truncated digits. While we have not recorded instances of FP animals producing offspring via FP, as described for the whitespotted bamboo shark (***Straube et al., 2016***), FP *A. marmoratus* 8450 did produce two eggs while housed in isolation, but these failed to hatch. Analysis of the ovaries of FP animal 8450 as well as germinal vesicles of its FP sister 8449

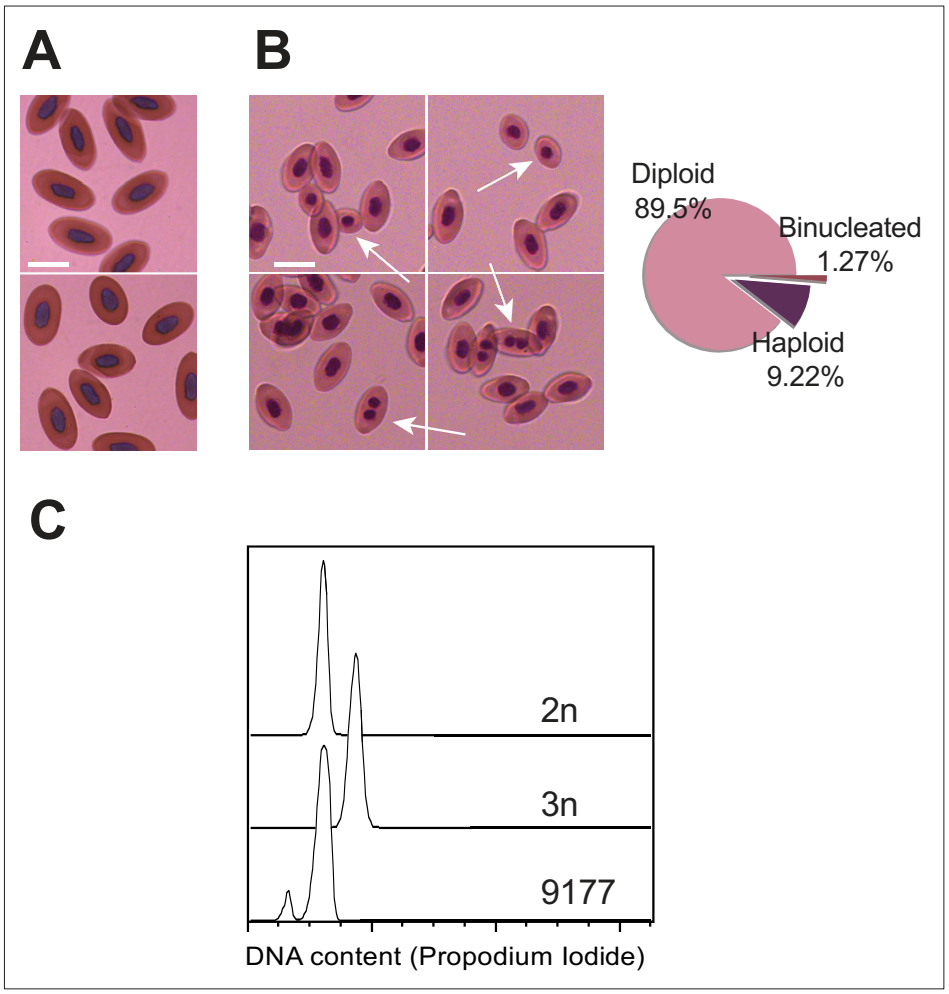

Figure 4. Detection of mixoploidy associated with facultative parthenogenesis. (**A**) Giemsa staining of erythrocytes from a sexually produced *Aspidoscelis marmoratus* (ID 14744). All cells are diploid (n=601). Scale bar corresponds to 10 µm. (**B**) Giemsa staining of erythrocytes from a newly hatched facultative parthenogenetic *A. marmoratus* (ID 9177). Diploid (n=844), smaller haploid (n=87), and binucleated (n=12) cells are evident. Scale bar corresponds to 10 µm. (**C**) DNA content from erythrocytes determined by propidium iodide staining and detection by flow cytometry. Samples are from a sexually produced *A. marmoratus* (2n, ID 5358), an obligate triploid parthenogen *A. exanguis* (3n, ID 4950), and facultative parthenogenesis (FP) *A. marmoratus* (ID 9177). Number of events scored by flow cytometry were 44,145 (2n), 44,043 (3n), and 44,060 (9177). The FP 9177 sample contained an additional peak to the left of the 2 n peak (90.04%), indicating the presence of haploid cells (9.62%).

The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. Mixoploidy detected in both *A. marmoratus* and *A. arizonae*.

Figure supplement 2. Animals produced by facultative parthenogenesis.

Figure supplement 3. Ovaries of *Aspidoscelis marmoratus* facultative parthenogenesis (FP) animal 8450 and germinal vesicles of FP sister 8449 revealed no differences in structure and anatomy compared to fertile sexually reproducing animals.

revealed no differences in structure and anatomy compared to fertile sexually reproducing animals (***Figure 4—figure supplement 3***).

Putative incidences of FP in wild populations of whiptail lizards

As FP has been associated with captivity in most species where it has been reported, we examined restriction site-associated DNA sequencing (RAD-seq) data for wild animals across 15 gonochoristic species (***Figure 5***). Because RAD-seq is a form of reduced-representation sequencing (***Rivera-Colón and Catchen, 2022***), we limited further analysis to the 321 individuals that had an average sequencing

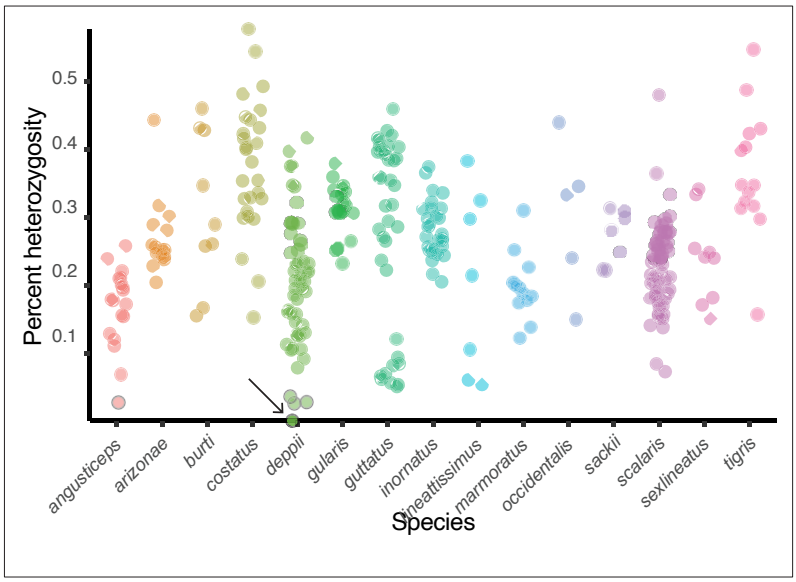

Figure 5. Heterozygosity estimates of whiptail lizards collected in nature. Percent heterozygosity estimates from reduced-representation sequencing (RAD-seq) for 321 whiptail lizards from 15 species. All individuals had an average coverage of at least 20. Each point is an individual, and percent heterozygosity was calculated only for sites where the coverage is equal to the average sequencing coverage. Five points with black borders indicate individuals (one *angusticeps* and four *deppii*) with low levels of heterozygosity. The heterozygosity of *Aspidoscelis deppii* ID LDOR30 (marked with arrow) is far less than that observed for individuals of the same species (Rosner's Test for Outliers within *deppii* individuals, p<0.001), having only one called heterozygous position. Species (sample size): *angusticeps* (**Booth et al., 2023**), *arizonae* (**Lenk et al., 2005**), *burti* (**Germano and Smith, 2010**), *costatus* (**Shibata et al., 2017**), *deppii* (**Ryba and Tirindelli, 1997**), *gularis* (**Olsen and Marsden, 1954**), *guttatus* (**Reeder et al., 2002**), *inornatus* (**Ryder et al., 2021**), *lineattissimus* (**Groot et al., 2003**), *marmoratus* (**Allen et al., 2018**), *occidentalis* (**Dudgeon et al., 2017**), *sackii* (**Groot et al., 2003**), *scalaris* (**Streisinger et al., 1981**), *sexlineatus* (**Germano and Smith, 2010**), *tigris* (**Card et al., 2021**).

The online version of this article includes the following source data for figure 5:

Source data 1. Source data of ***Figure 5***.

coverage of at least 20. Computational data analysis revealed five animals (one *A. angusticeps* and four *A. deppii*) that had very low levels of heterozygosity at positions of average coverage (<0.05% heterozygous positions). In contrast, the average level of heterozygosity was 0.261% across the dataset, with the highest heterozygosity value at 0.578%. Of the five low heterozygosity animals, *A. deppii* (ID LDOR30) showed the most striking level of homozygosity affecting all sites but one (Rosner's Test for Outliers within *A. deppii* individuals, log-transformation, R=5.127, λ=3.928, p<0.001). This is consis- tent with the pattern of homozygosity observed with the whole-genome sequencing from FP animals produced in the laboratory. Further fieldwork and analysis will be required to assess the level of FP in natural populations of gonochoristic *Aspidoscelis* species (and other factors that could influence the observed heterozygosity such as population size, levels of hybridization, and inbreeding). To assess whether all whiptail species can produce viable offspring through FP, larger and broader datasets will be required to compare the incidence of FP between species, especially because animals with developmental defects associated with FP would not have hatched or survived in the wild and would, therefore, not have shown up in the RAD-seq dataset.

Discussion

In this study, we report over 20 incidences of facultative parthenogenesis in marbled and Arizona striped whiptail lizards. By sequencing and assembling a highly contiguous *A. marmoratus* genome, we were able to refute automixis as the underlying mechanism in whiptail lizards. Instead, genome- wide homozygosity raises a possibility of a post-meiotic mechanism involving the activation of embryonic development in unfertilized haploid oocytes. Even though FP whiptail lizards are largely

comprised of homozygous diploid cells, a fraction of haploid cells persists through development and is readily detectable in young adults. Such mixoploidy and genome-wide homozygosity come at a price. In eight confirmed and 24 suspected cases of FP, development ceased prior to hatching and most FP animals that hatched showed congenital defects. Nevertheless, FP was observed at a rate of 1% and 5% in *A. arizonae* and *A. marmoratus*, respectively, and this occurred in the presence of mating partners. Interestingly, these rates are similar to what has been reported for wild populations of two North American pitviper species (**Booth et al., 2012**). Our findings indicate that FP is far more common in some vertebrate species than previously thought. The purifying selection associated with homozygosity may be an important force in generating additional resilience to counteract the effects of population bottlenecks and inbreeding depression. However, support for this hypothesis is predicated on the fitness and reproduction of FP offspring and, therefore, more long-term studies on seemingly healthy individuals of FP origin are needed.

The start of embryonic development is tightly coupled to fertilization in many vertebrates as the sperm entering the oocyte triggers a signaling cascade that is essential for the completion of female meiosis and initiation of cell division following karyogamy (**Sagata, 1996**). This process has been mimicked in the laboratory by piercing frog oocytes with a needle to trigger the signal to complete meiosis and initiate replication and division in the haploid oocyte (**Le Peuch et al., 1985**; **Wolf, 1974**). In zebrafish, homozygous embryos are routinely generated by fertilization of oocytes with UV-irradiated sperm, a treatment that destroys the paternal DNA (**Streisinger et al., 1981**). An oocyte treated in this manner will replicate the intact maternal genome in the absence of karyogamy. If the haploid oocyte is then subjected to heat shock treatment, cytokinesis is prevented resulting in a pseudodiploid oocyte that undergoes a second round of DNA replication followed by mitosis. Thus an entirely homozygous diploid embryo starts to develop (**Kroeger et al., 2014**). In contrast to human intervention forcing two consecutive rounds of DNA replication to occur without intervening mitosis at the start of embryonic development, our data indicate that one or multiple rounds of DNA replication and mitosis take place in some haploid oocytes of whiptail lizards prior to a skipped mitosis yielding a homozygous diploid cell followed by mixoploid development. Initiation of development in a haploid state may be conserved in avian species as unfertilized turkey eggs can yield embryos that contain 40% of haploid cells at blastoderm followed by a reduction to 1.3% within the blood of hatched birds (**Cassar et al., 1998**). Depending on the tissue and ploidy distribution, the presence of haploid cells may contribute to abnormal development of specific tissues reported here and elsewhere (**Ito et al., 1991**; **Tanaka et al., 2004**). Successful embryonic development from haploid cells that restore the diploid state by duplication has also been observed in a stick insect species (**Pijnacker, 1969**). In whiptail lizards, we have not been able to examine post-meiotic oocytes as locating the post-meiotic nucleus within a large yolked egg is inherently difficult. The difficulty is compounded by the unpredictability of which eggs will undergo FP development and the need to sacrifice animals to remove eggs.

In addition to mixoploidy, genome-wide homozygosity constitutes another obstacle to normal development as each recessive deleterious allele is exposed in either the hemizygous state (haploid cells) or homozygous state (diploid cells). Indeed, arrested development and abnormal phenotypes are observed in FP whiptails, as well as in FP animals across many other species (**Booth et al., 2023**; **Booth and Schuett, 2016**; **Adams et al., 2023**; **Olsen, 1973**). It is important to note though that some whiptails of FP origin developed normally, much like their sexually produced counterparts. At the population level, FP leads to a precipitous reduction in genetic diversity as only one set of alleles is inherited in the next generation. While FP could be an adaptive trait to bridge population bottlenecks when mate encounters are infrequent, small populations already rely heavily on inbreeding and FP further reduces the size of the gene pool (**Watts et al., 2006**; **Ryder et al., 2021**).

While FP can be considered the most extreme example of inbreeding, it is also the most powerful example of genetic purging as it eliminates most deleterious alleles in a single generation. FP in whiptail lizards and other species could, therefore, be considered a reproductive strategy, akin to mixed-mating systems in plants (**Goodwillie et al., 2005**). The ability to produce offspring via two different strategies provides a level of reproductive assurance in plants (**Busch and Delph, 2012**). Indeed, we see parallels to this in our own data in which female whiptails have produced offspring via both sexual reproduction and FP on separate occasions or simultaneously within a single clutch. Within plant species with mixed-mating, there are differences between the rates of selfing and outbreeding between populations, and hypotheses as to why these differences occur include limited pollinator

visitation and resource availability (***Whitehead et al., 2018***). To assess whether the co-occurrence of sexual and FP reproduction in vertebrates can indeed be considered a reproductive strategy rather than biological noise will require further studies to assess the reproductive competence and fecundity of offspring produced by either mode of reproduction. To gain a better understanding of the origin and outcomes of FP in whiptail lizards, it will also be important to identify the triggers. It has been proposed that lack of or limited mate encounters triggers FP, but our data in combination with many other reports (***Booth et al., 2012***; ***Booth et al., 2014***; ***Kratochvíl et al., 2020***; ***Ryder et al., 2021***; ***Booth et al., 2011***; ***Feldheim et al., 2023***; ***Larose et al., 2023***) rejects the idea that this is the key trigger. Recent work identifying key cell cycle genes inducing FP in two species of *Drosophila* (***Sper- ling et al., 2023***) and selection resulting in higher incidences of parthenogenesis in birds (***Parker et al., 2010***; ***Olsen, 1975***; ***Olsen et al., 1968***) suggest a genetic basis for the initiation of FP.

Our study adds two species of whiptail lizards to a growing list of vertebrates capable of FP and establishes that it occurs alongside sexual reproduction in the presence of males. Using whole- genome sequencing, we demonstrate that post-meiotic genome duplication is the underlying mecha- nism. One must now consider the possibility that FP is an adaptive trait and that low rates of successful FP could contribute significantly to genome purification. Sexually mature FP offspring will have a low genetic load and only pass on neutral or mildly deleterious alleles to the next generation. However, a role for FP as an adaptive trait hinges on further studies demonstrating the ability of parthenogens to reproduce themselves either through further FP or sexually. If successive reproduction occurs, FP may reduce the frequency of deleterious alleles within a population, as well as provide reproductive assur- ance when males are scarce. Additional whole-genome sequencing data for species with documented FP will be needed for a better understanding of the genetic basis, propensity, and evolutionary signif- icance of FP.

Materials and methods

Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Biological sample (*Aspidoscelis spp.*)	Erythrocytes, tail clippings, liver isolation	This paper, wild populations		See Ethics Statement
Chemical compound, drug	Giemsa stain	Sigma	GS500	0.40%
Chemical compound, drug	Schiff's reagent	Fisher Scientific	#SS32-500	
Commercial assay or kit	Roche gDNA Isolation Kit	Roche	#11814770001	
Commercial assay or kit	KAPA HTP kit	KAPA	KK8234	
Commercial assay or kit	Nextera Mate-Pair Library Prep Kit	Illumina	FC-132–1001	
Commercial assay or kit	TruSeq RNA Library Prep Kit v2	Illumina	RS-122–2001	
Sequence-based reagent	Oligos for MS analysis	This paper		See ***Supplementary file 7***
Software, algorithm	Meraculous 2.0	https://jgi.doe.gov/data-and-tools/software-tools/meraculous/	RRID:SCR_010700	
Software, algorithm	BUSCO 3.0.1	https://busco.ezlab.org/	RRID:SCR_015008	
Software, algorithm	RAxML 8.2.11	https://cme.h-its.org/exelixis/web/software/raxml/	RRID:SCR_006086	
Software, algorithm	RepeatModeler 1.0.11	https://www.repeatmasker.org/RepeatModeler/	RRID:SCR_015027	
Software, algorithm	RepeatMasker 4.0.9	https://www.repeatmasker.org/	RRID:SCR_012954	
Software, algorithm	BLAST 2.6.0 & 2.9.0+	https://blast.ncbi.nlm.nih.gov/Blast.cgi	RRID:SCR_004870	
Software, algorithm	BWA 0.7.15	https://bio-bwa.sourceforge.net/	RRID:SCR_010910	

Continued on next page

Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Software, algorithm	Picard 1.119	https://broadinstitute.github.io/picard/	RRID:SCR_006525	
Software, algorithm	GATK 3.5	https://gatk.broadinstitute.org/hc/en-us	RRID:SCR_001876	
Software, algorithm	seqtk 1.2-r94	https://github.com/lh3/seqtk ***Li, 2016***	1.2-r94	
Software, algorithm	pysam 0.12.0.1	https://github.com/pysam-developers/pysam ***pysam-developers, 2017***	0.12.0.1	
Software, algorithm	pysamstats 0.24.3	https://github.com/alimanfoo/pysamstats ***Miles, 2015***	0.24.3	
Software, algorithm	Trinity	https://github.com/trinityrnaseq/trinityrnaseq ***trinityrnaseq, 2015***	2.0.6	
Software, algorithm	seqclean	https://sourceforge.net/projects/seqclean		
Software, algorithm	MAKER2 2.31.8	https://www.yandell-lab.org/software/maker.html	RRID:SCR_005309	
Software, algorithm	Interproscan 5.13–52.0	https://interproscan-docs.readthedocs.io/en/latest/	RRID:SCR_005829	
Software, algorithm	Exonerate 2.4.0	https://www.ebi.ac.uk/about/vertebrate-genomics/software/exonerate	RRID:SCR_016088	
Software, algorithm	Geneious 10.1.3	https://www.geneious.com/	RRID:SCR_010519	
Software, algorithm	Stacks 2.62	https://catchenlab.life.illinois.edu/stacks/	RRID:SCR_003184	
Software, algorithm	Micromanager 1.4	https://micro-manager.org/	RRID:SCR_000415	
Software, algorithm	Flowjo treestar	https://www.flowjo.com/	RRID:SCR_008520	

Microsatellite analysis

DNA was extracted from tail samples for microsatellite genotyping as described in ***Lutes et al., 2011***. PCR products were analyzed by capillary electrophoresis on a 3730 DNA Analyzer and data was analyzed using GeneMapper (v. 4.0). Primer information can be found in ***Supplementary file 7***.

Genome size estimation

The genome size of *A. marmoratus* was estimated by fluorescence-activated cell sorting (FACs), in which a standard curve correlating fluorescence intensity of DNA-bound propidium iodide with known genome sizes was generated using cells from fruit flies, zebrafish, and mouse, and then comparing fluorescent intensity with that of erythrocytes from *A. marmoratus*. Samples were stained using the Sigma PI staining preparation and analyzed on the Influx cytometer. PI fluorescence was collected using the PI Texas red detector with linear amplification and data analysis was performed in FlowJo and Microsoft Excel.

DNA isolation, sequencing, and genome assembly for *A. marmoratus*

All genome sequencing libraries generated for the purpose of the *A. marmoratus* genome assembly were derived from the FP animal 8450. The liver tissue was first dissociated in a 10 mL Dounce

homogenizer using the tight-fitting pestle and then processed using the Roche gDNA Isolation Kit (#11814770001, MilliporeSigma, St. Louis, MO, USA).

A short insert, high-coverage library was generated using the KAPA HTP kit (KK8234), with 1 µg of gDNA. The resulting library was size selected for fragments between 500–850 bp on a Pippin Prep (Sage Science). Two 40 Kb mate-pair libraries were generated by Lucigen from 1 µg of gDNA using the CviQI and BfaI restriction enzymes, respectively. Each library was sequenced on the Illumina MiSeq using the MiSeq Reagent Kit v2 (500 cycles). An additional three mate-pair libraries were generated, spanning distances of 5 Kb, 8 Kb, and 2–15 Kb, using the Illumina Nextera Mate-Pair Library Prep Kit and 1 µg of gDNA for each. Size selection used the Gel-Plus protocol with Pippin for the 5 and 8 Kb libraries, and the Gel-Free protocol for the 2–15 Kb library. All three libraries were pooled and sequenced on three separate RapidSeq flow cells on an Illumina HiSeq 2500. Chicago libraries were prepared at Dovetail Genomics LLC, Santa Cruz, CA, USA from liver tissue to generate read pairs spanning distances up to 140 Kb and sequenced on an Illumina HiSeq 2500. The combined sequencing data was initially assembled at Dovetail Genomics using Meraculous and their in-house HiRise genome assembly algorithms to generate the *A. marmoratus* reference genome (AspMar1.0).

Assessing assembly completeness

In order to assess the completeness of the *A. marmoratus* reference genome, we used BUSCO (v. 3.0.1) (***Simão et al., 2015***) with the vertebrate_od9 dataset containing 2586 genes, with default parameters apart from changing the BLAST cutoff from 1e-3 to a more stringent value of 1e-6. We used BUSCO numbers generated in ***Gao et al., 2017*** for *Shinisaurus crocodilurus* and *Alligator mississippiensis* in ***Figure 2—figure supplement 1***.

To perform a phylogenetic analysis, 1333 shared 'complete' single-copy orthologs were identified in the genomes of green anole (*Anolis carolinensis*, anoCar2), cow (*Bos taurus*, ARS-UCD1.2), dog (*Canis lupus familiaris*, CanFam3.1), zebrafish (*Danio rerio*, danRer10), chicken (*Gallus gallus*, galGal5), human (*Homo sapiens*, GRCh38.p13), mouse (*Mus musculus*, GRCm38.p6), medaka (*Oryzias latipes*, oryLat2), rat (*Rattus norvegicus*, Rnor_6.0), Argentine black and white tegu (*Salvator merianae*, HLtupMer3), western clawed frog (*Xenopus tropicalis*, Xenopus_tropicalis_v9.1), platyfish (*Xiphophorus maculatus*, X_maculatus-5.0-male). For each amino acid sequence, a multiple sequence alignment was performed with MAFFT (v. 7.305) (***Katoh and Standley, 2013***). The alignments were concatenated into a supermatrix of 1,112,277 amino acids. Phylogenetic tree topology was estimated using the Maximum Likelihood inference method using the pthreads version of RAxML (v. 8.2.11) and the PROTOGAMMAAUTO model for sequence evolution with 100 bootstrap replicates (***Stamatakis, 2014***).

Repeat identification

We quantified and annotated the repetitive DNA content within the *A. marmoratus* genome assembly by using the RepeatMasker pipeline on *A. marmoratus* scaffolds greater than or equal to 10 Kb in length. We first generated a de novo list of *A. marmoratus* repetitive elements using RepeatMod- eler (v. 1.0.11) (***Smit and Hubley, 2008***). We then used these as input into RepeatMasker (v. 4.0.9) (***Smit et al., 2008***) using the NCBI/RMBLAST (v. 2.6.0+) search engine. Unclassified repeat element consensus sequences from the RepeatModeler output for each of the three lizards (*A. marmoratus, S. merianae,* and *A. carolinensis*) were compared to each other by identifying reciprocal best hits using BLAST (v. 2.9.0+).

Whole-genome sequencing, reference genome alignment, and heterozygosity determination

Genomic DNA isolated from either liver or tail was prepared for sequencing using the KAPA HTP Library Preparation Kit (KK8234). Stock adapters were used from the Nextflex kit and barcodes were from BioScientific. All libraries were sequenced on the Illumina HiSeq 2500 platform. Whole-genome sequencing data was aligned to the *A. marmoratus* reference genome with BWA (v. 0.7.15) (***Li and Durbin, 2010***) and marked for duplicates with Picard (RRID:SCR_006525; v. 1.119; https://broadinstitute.github.io/picard/). Because samples were sequenced over multiple lanes, the alignment files were merged subsequently, and another round of duplication marking was performed. The alignment files were realigned around small insertions and deletions with GATK (v. 3.5) (***DePristo et al., 2011***). Data

corresponding to lizard ID 122's bam file was down-sampled to 33% of its original size using seqtk (v. 1.2-r94) to match the expected average genome coverage of the other samples, as this animal was sequenced on one flow cell without multiplexing and, therefore, having much more sequenced reads.

The per position nucleotide profiles for each alignment were then generated using a combination of pysam (RRID:SCR_021017, v. 0.12.0.1) (https://github.com/pysam-developers/pysam) and pysamstats (***Miles, 2015***. v. 0.24.3) (https://github.com/alimanfoo/pysamstats) to determine the heterozygosity at any genomic position.

Transcriptome assembly and genome annotation

Two poly-A selected stranded RNA-sequencing libraries were generated with the TruSeq RNA Library Prep Kit v2 (RS-122–2001) and sequenced on an Illumina HiSeq 2500 for the purpose of an *A. marmoratus* transcriptome assembly. The first library was derived from a blood sample taken from a male animal, and the second library was derived from an embryo incubated at 28 °C and harvested 47–51 days post-egg deposition.

Trinity (v. 2.0.6) (***Grabherr et al., 2011***) was then used to generate an initial transcriptome assembly. The original reads were aligned to this transcriptome assembly using the Trinity companion script align_and_estimate_abundance.pl. Transcript isoforms with no read support were then filtered out and the remaining assembly was run through seqclean (https://sourceforge.net/projects/seqclean/). Evidence-based annotations for the transcriptome assembly were generated using the MAKER2 pipeline (v. 2.31.8) (***Holt and Yandell, 2011***). For MAKER2, the entire UniProtKB/Swiss-Prot database of proteins (***UniProt Consortium, 2018***) was used and the Repbase data base was used to mask repeats within the MAKER2 framework (***Bao et al., 2015***). Assigning putative functions to the gene models was performed using BLAST (v. 2.6.0) and Interproscan (v. 5.13–52.0) (***Jones et al., 2014***).

Copy number estimation for the Vomeronasal 2 receptor 26 (Vmn2r26) genes was based on aligning the mouse ortholog (http://www.orthodb.org), to the *A. marmoratus* reference assembly using Exonerate (v. 2.4.0) (***Slater and Birney, 2005***) with a maximum intron size set to 20 Kb. Genes annotated as Vmn2r26 in the MAKER2 annotations were concatenated and aligned using Geneious (v. 10.1.3) (***Geneious Prime, 2017***) with default settings. The FastTree plugin (v. 1.0) was used to generate the phylogenetic tree from the alignment with default parameters.

RAD-sequencing analysis

Double digest RAD-sequencing data was derived primarily from previous studies (***Barley et al., 2021***; ***Barley et al., 2022b***; ***Barley et al., 2019***; ***Barley et al., 2022a***). Fastq files were processed with Stacks (v. 2.62) using process_radtags and then samples from the same species were passed separately using denovo_map.pl. The script executes all the components of the standard Stacks pipeline. The gstacks program of denovo_map.pl calls variants for each position in each locus and assigns it either homozygous, heterozygous, or unknown. Output files were used to calculate the average coverage for each sample. Samples that have at least a coverage of 20 were considered for subsequent analysis. For each passing sample, the heterozygosity was calculated for sites at average coverage by adding up all heterozygous positions and dividing it by the total.

Giemsa staining of erythrocytes

Whole blood was collected aseptically from tails using acid citrate dextrose as an anticoagulant. All samples were prepared immediately after collection. A 5 µL aliquot of the diluted blood sample was used to prepare a blood smear on a 25 × 75 × 1 mm microscope slide. Once dry, the slide was placed in 95% ethanol for 5 min. Giemsa stain (0.4%; Sigma, GS500) was applied liberally to cover the slide every 4 min for a total incubation time of 16 min. The prepared slides were imaged on an Axioplan2 imaging microscope equipped with a plan-apochromat 100 x /1.40 Oil objective and an Axiocam HRc (color) camera (Zeiss). Micromanager (v. 1.4) software was used to acquire the images. The acquired images were then scored visually for the number of haploid, diploid, and binucleated erythrocytes.

Feulgen staining of erythrocytes

Whole blood was collected aseptically from tails using acid citrate dextrose as an anticoagulant. All samples were prepared immediately after collection. A 5 µL aliquot of the diluted blood sample was used to prepare a blood smear on a 25 × 75 × 1 mm microscope slide. Blood smears were treated

with 10% neutral buffered formalin for 5 min at RT, then rinsed twice in distilled water. The slides were immersed in 5 M HCl for 30 min at RT, and then rinsed 2 times in distilled water. Slides were then immersed in Schiff's reagent (Fisher Scientific #SS32-500) at RT for 15–30 min until nuclei were stained. The slides were transferred directly to bisulfite water that was prepared by dissolving 2.5 g of potassium metabisulfite in 500 mL of water and adjusting the pH to 4.0 by the addition of concentrated HCl. The bisulfite wash was repeated three times with 10–15 sec of agitation. The slide was then washed under running tap water for 2 min and dehydrated by incubating in 70% EtOH for 5 min and then 95% EtOH for 5 min. The preparations were cleared in xylene before mounting.

Flow cytometry of erythrocytes

Blood collected from animals was treated as previously described for flow cytometry with modifications: ethanol fixation was performed after RNase treatment and propidium iodide staining was performed overnight, followed by sonication to disrupt aggregate cells (***Lutes et al., 2011***). A minimum number of 50,000 events were collected for each sample. Flowjo (treestar) was used for data analysis.

Imaging of ovaries and germinal vesicle

Germinal vesicle isolation and acquisition of image stacks by confocal microscopy were performed as described (***Lutes et al., 2010***). Ovary images were acquired with a Leica M205FA dissection microscope with a planar 0.63 X objective using Micromanager (v. 1.4) software.

Acknowledgements

We are grateful for the husbandry staff at the Stowers Institute for Medical Research (SIMR) espe- cially Rick Kupronis and the team of dedicated reptile technicians, David Jewell, Alex Muensch, Jillian Schieszer, Christina Piraquive, and Kristy Winter, for outstanding husbandry and herpetocultural skills. We thank the members of the New Mexico Department of Game and Fish and the Arizona Game and Fish Department for their support. We thank Veronica Cloud for help with RNA extractions and the molecular biology, flow cytometry, and advanced microscopy core facilities at SIMR for their excellent technical support and the members of the Emergent AI Center at Johannes Gutenberg University Mainz (JGU) for outstanding IT support and helpful discussions. We acknowledge computing time granted on the supercomputer MOGON II at JGU as part of NHR South-West (nhrsw.de), as well as the Institute of Molecular Biology Bioinformatics Core for computing resources. This work was funded in part by the Howard Hughes Medical Institute, SIMR, and an Alexander von Humboldt Professor- ship awarded to P.B. at JGU. This work also received support from the National Science Foundation grant DEB-1754350 and the GenEvo RTG funded by the Deutsche Forschungsgemeinschaft (German Research Foundation) – GRK2526/1 – Projectnr. 40.7023052 and the Institute for Quantitative and Computational Biosciences (IQCB) in Mainz. We thank Charles (Jay) Cole, Uri García-Vázquez, Rick Kupronis, Daniel Lara-Tufiño, Norma Manríquez-Moran, Maximiliano Monroy, Priscilla Neaves, Adrián Nieto Montes de Oca, Harry Taylor, Robert Thomson, and Carol Townsend for assistance with field- work and specimen collection. We also thank the Natural History Museum of Los Angeles County, the UTEP Biodiversity Collections, Jay Cole and the American Museum of Natural History, the Museum of Vertebrate Zoology, Randy Klabacka and the Auburn University Museum of Natural History, and the LSU Museum of Natural Science for providing tissue loans in support of this research.

Additional information

Funding

Funder	Grant reference number	Author
Stowers Institute for Medical Research		Duncan Tormey
Howard Hughes Medical Institute		Peter Baumann

Funder	Grant reference number	Author
Alexander von Humboldt-Stiftung		Peter Baumann
Deutsche Forschungsgemeinschaft	GRK2526/1	David V Ho
National Science Foundation	DEB-1754350	Anthony J Barley

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

David V Ho, Data curation, Software, Formal analysis, Validation, Investigation, Visualization, Writing – original draft, Writing – review and editing; Duncan Tormey, Data curation, Software, Investigation, Methodology, Writing – original draft; Aaron Odell, Formal analysis, Validation, Investigation, Visualization, Methodology; Aracely A Newton, Robert R Schnittker, Formal analysis, Investigation, Methodology; Diana P Baumann, Resources, Data curation; William B Neaves, Investigation; Morgan R Schroeder, Rutendo F Sigauke, Software, Formal analysis, Investigation, Methodology; Anthony J Barley, Resources, Investigation, Writing – review and editing; Peter Baumann, Conceptualization, Resources, Supervision, Funding acquisition, Validation, Writing – original draft, Project administration, Writing – review and editing

Author ORCIDs

David V Ho http://orcid.org/0000-0002-4709-269X
Duncan Tormey https://orcid.org/0000-0002-0963-3424
Robert R Schnittker http://orcid.org/0000-0002-2950-8604
Anthony J Barley http://orcid.org/0000-0003-1675-6577
Peter Baumann http://orcid.org/0000-0003-4892-1485

Ethics

Animals used in this study were produced in the AAALAC International accredited Stowers Reptile and Aquatics Facility in compliance with protocols approved by the Institutional Animal Care and Use Committee. They descended from breeding stock collected in New Mexico under permit numbers 3199 and 3395 and Arizona under license number SP564133. Tissues used in the RAD- seq analyses were derived from samples collected under permits from the Arizona Game and Fish Department, California Department of Fish and Wildlife, New Mexico Department of Game and Fish, and the Secretariá de Medio Ambiente y Recursos Naturales, Dirección General de Fauna Silvestre of Mexico.

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.97035.sa1
Author response https://doi.org/10.7554/eLife.97035.sa2

Additional files

Supplementary files

- Supplementary file 1. *A. marmoratus* genome assembly statistics.
- Supplementary file 2. Trinity assembly statistics.
- Supplementary file 3. MAKER2 summary.
- Supplementary file 4. Information for all *A. marmoratus* animals sequenced.
- Supplementary file 5. Whole genome sequencing alignment statistics.
- Supplementary file 6. Animals confirmed by microsatellite analysis to be of facultative parthenogenesis (FP) origin.
- Supplementary file 7. Microsatellite primer information.
- MDAR checklist

Data availability

All raw sequencing data pertaining to the AspMar1.0 genome assembly are available at the National Center for Biotechnology Information under project accession number PRJNA360150. All other whole-genome and RNA sequencing data can be found under PRJNA980964. RAD-seq data can be located under PRJNA827355, PRJNA707030, PRJNA762930, and PRJNA1016487. Code and raw microsatellite data used for analysis are available at GitHub (copy archived at ***baumannlab, 2024***).

The following datasets were generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Baumann P	2020	Aspidoscelis marmoratus isolate:SIMRID8450 (Marbled whiptail)	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA360150	NCBI BioProject, PRJNA360150
Ho DV, Tormey D, Odell A, Newton AA, Schnittker RR, Baumann DP, Neaves WB, Schroeder MR, Sigauke RF, Barley AJ, Baumann P	2024	Sequencing of Aspidoscelis maramoratus	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA980964	NCBI BioProject, PRJNA980964
Ho DV, Tormey D, Odell A, Newton AA, Schnittker RR, Baumann DP, Neaves WB, Schroeder MR, Sigauke RF, Barley AJ, Baumann P	2024	Post-meiotic mechanism of facultative parthenogenesis in gonochoristic whiptail lizard species	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1016487	NCBI BioProject, PRJNA1016487

The following previously published datasets were used:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Barley AJ	2021	Evolutionary study of whiptail lizards (Aspidoscelis)	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA707030	NCBI BioProject, PRJNA707030
Barley AJ	2021	Genetic diversity and the origins of parthenogenesis in the teiid lizard Aspidoscelis laredoensis	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA762930	NCBI BioProject, PRJNA762930
Barley AJ, de Oca ANM, Normal L, Robert CT	2022	The evolutionary network of whiptail lizards reveals predictable outcomes of hybridization	https://www.ncbi.nlm.nih.gov/bioproject/PRJNA827355	NCBI BioProject, PRJNA827355

References

Adams L, Lyons K, Monday J, Larkin E, Wyffels J. 2023. Costs of parthenogenesis on growth and longevity in ex situ zebra sharks Stegostoma tigrinum. *Endangered Species Research* **50**:81–91. DOI: https://doi.org/10.3354/esr01224

Allen L, Sanders KL, Thomson VA. 2018. Molecular evidence for the first records of facultative parthenogenesis in elapid snakes. *Royal Society Open Science* **5**:171901. DOI: https://doi.org/10.1098/rsos.171901, PMID: 29515892

Avise JC. 2015. Evolutionary perspectives on clonal reproduction in vertebrate animals. *PNAS* **112**:8867–8873. DOI: https://doi.org/10.1073/pnas.1501820112, PMID: 26195735

Bao W, Kojima KK, Kohany O. 2015. Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mobile DNA* **6**:11. DOI: https://doi.org/10.1186/s13100-015-0041-9, PMID: 26045719

Barley AJ, Nieto-Montes de Oca A, Reeder TW, Manríquez-Morán NL, Arenas Monroy JC, Hernández-Gallegos O, Thomson RC. 2019. Complex patterns of hybridization and introgression across evolutionary timescales in Mexican whiptail lizards (Aspidoscelis). *Molecular Phylogenetics and Evolution* **132**:284–295. DOI: https://doi.org/10.1016/j.ympev.2018.12.016, PMID: 30562610

Barley AJ, Reeder TW, Nieto-Montes de Oca A, Cole CJ, Thomson RC. 2021. A new diploid parthenogenetic whiptail lizard from sonora, mexico, is the “missing link” in the evolutionary transition to polyploidy. *The American Naturalist* **198**:295–309. DOI: https://doi.org/10.1086/715056, PMID: 34260872

Barley AJ, Cordes JE, Walker JM, Thomson RC. 2022a. Genetic diversity and the origins of parthenogenesis in the teiid lizard Aspidoscelis laredoensis. *Molecular Ecology* **31**:266–278. DOI: https://doi.org/10.1111/mec.16213, PMID: 34614250

Barley AJ, Nieto-Montes de Oca A, Manríquez-Morán NL, Thomson RC. 2022b. The evolutionary network of whiptail lizards reveals predictable outcomes of hybridization. *Science* **377**:773–777. DOI: https://doi.org/10.1126/science.abn1593, PMID: 35951680

Bartelmez GW, Riddle O. 1924. On parthenogenetic cleavage and on the rôle of water absorption by the ovum in the formation of the subgerminal cavity in the pigeon's egg. *American Journal of Anatomy* **33**:57–66. DOI: https://doi.org/10.1002/aja.1000330104

baumannlab. 2024. Aspi_Fp. swh:1:rev:01a8d33519e700d701db67bb159fb5e693933327. Software Heritage. https://archive.softwareheritage.org/swh:1:dir:5cc3a8b9cefd8fb929893207512fb43194ee0a5f;origin=https://github.com/baumannlab/Aspi_FP;visit=swh:1:snp:25c60a968c6382793a701367d2e6f8e39d0dc415;anchor=swh:1:rev:01a8d33519e700d701db67bb159fb5e693933327

Booth W, Johnson DH, Moore S, Schal C, Vargo EL. 2011. Evidence for viable, non-clonal but fatherless Boa constrictors. *Biology Letters* **7**:253–256. DOI: https://doi.org/10.1098/rsbl.2010.0793, PMID: 21047849

Booth W, Schuett GW. 2011. Molecular genetic evidence for alternative reproductive strategies in North American pitvipers (Serpentes: Viperidae): long-term sperm storage and facultative parthenogenesis. *Biological Journal of the Linnean Society* **104**:934–942. DOI: https://doi.org/10.1111/j.1095-8312.2011.01782.x

Booth W, Smith CF, Eskridge PH, Hoss SK, Mendelson JR, Schuett GW. 2012. Facultative parthenogenesis discovered in wild vertebrates. *Biology Letters* **8**:983–985. DOI: https://doi.org/10.1098/rsbl.2012.0666, PMID: 22977071

Booth W, Schuett GW, Ridgway A, Buxton DW, Castoe TA, Bastone G, Bennett C, McMahan W. 2014. New insights on facultative parthenogenesis in pythons. *Biological Journal of the Linnean Society* **112**:461–468. DOI: https://doi.org/10.1111/bij.12286

Booth W, Schuett GW. 2016. The emerging phylogenetic pattern of parthenogenesis in snakes. *Biological Journal of the Linnean Society* **118**:172–186. DOI: https://doi.org/10.1111/bij.12744

Booth W, Levine BA, Corush JB, Davis MA, Dwyer Q, De Plecker R, Schuett GW. 2023. Discovery of facultative parthenogenesis in a new world crocodile. *Biology Letters* **19**:20230129. DOI: https://doi.org/10.1098/rsbl.2023.0129, PMID: 37282490

Brykczynska U, Tzika AC, Rodriguez I, Milinkovitch MC. 2013. Contrasted evolution of the vomeronasal receptor repertoires in mammals and squamate reptiles. *Genome Biology and Evolution* **5**:389–401. DOI: https://doi.org/10.1093/gbe/evt013, PMID: 23348039

Busch JW, Delph LF. 2012. The relative importance of reproductive assurance and automatic selection as hypotheses for the evolution of self-fertilization. *Annals of Botany* **109**:553–562. DOI: https://doi.org/10.1093/aob/mcr219, PMID: 21937484

Card DC, Vonk FJ, Smalbrugge S, Casewell NR, Wüster W, Castoe TA, Schuett GW, Booth W. 2021. Genome-wide data implicate terminal fusion automixis in king cobra facultative parthenogenesis. *Scientific Reports* **11**:7271. DOI: https://doi.org/10.1038/s41598-021-86373-1, PMID: 33790309

Cassar G, John TM, Etches RJ. 1998. Observations on ploidy of cells and on reproductive performance in parthenogenetic turkeys. *Poultry Science* **77**:1457–1462. DOI: https://doi.org/10.1093/ps/77.10.1457, PMID: 9776050

Chapman DD, Shivji MS, Louis E, Sommer J, Fletcher H, Prodöhl PA. 2007. Virgin birth in a hammerhead shark. *Biology Letters* **3**:425–427. DOI: https://doi.org/10.1098/rsbl.2007.0189, PMID: 17519185

Chapman DD, Firchau B, Shivji MS. 2008. Parthenogenesis in a large-bodied requiem shark, the blacktip *Carcharhinus limbatus*. *Journal of Fish Biology* **73**:1473–1477. DOI: https://doi.org/10.1111/j.1095-8649.2008.02018.x

Chapman JA, Ho I, Sunkara S, Luo S, Schroth GP, Rokhsar DS. 2011. Meraculous: de novo genome assembly with short paired-end reads. *PLOS ONE* **6**:e23501. DOI: https://doi.org/10.1371/journal.pone.0023501, PMID: 21876754

Cole CJ, Lowe CH, Wright JW. 1969. *Sex Chromosomes in Teiid Whiptail Lizards (Genus Cnemidophorus)*. American Museum of Natural History.

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics* **43**:491–498. DOI: https://doi.org/10.1038/ng.806

Dubach J, Sajewicz A, Pawley R. 1997. Parthenogenesis in the Arafuran filesnake (Acrochordus arafurae). *Herpetological Natural History* **5**:11–18.

Dudgeon CL, Coulton L, Bone R, Ovenden JR, Thomas S. 2017. Switch from sexual to parthenogenetic reproduction in a zebra shark. *Scientific Reports* **7**:40537. DOI: https://doi.org/10.1038/srep40537, PMID: 28091617

Fang X, Mou Y, Huang Z, Li Y, Han L, Zhang Y, Feng Y, Chen Y, Jiang X, Zhao W, Sun X, Xiong Z, Yang L, Liu H, Fan D, Mao L, Ren L, Liu C, Wang J, Li K, et al. 2012. The sequence and analysis of a Chinese pig genome. *GigaScience* **1**:16. DOI: https://doi.org/10.1186/2047-217X-1-16, PMID: 23587058

Feldheim KA, Chapman DD, Sweet D, Fitzpatrick S, Prodöhl PA, Shivji MS, Snowden B. 2010. Shark virgin birth produces multiple, viable offspring. *The Journal of Heredity* **101**:374–377. DOI: https://doi.org/10.1093/jhered/esp129, PMID: 20106913

Feldheim KA, Dubach J, Watson L. 2023. Parthenogenesis in an elasmobranch in the presence of conspecific males. *Journal of Fish Biology* **102**:525–527. DOI: https://doi.org/10.1111/jfb.15268, PMID: 36369968

Fields AT, Feldheim KA, Poulakis GR, Chapman DD. 2015. Facultative parthenogenesis in a critically endangered wild vertebrate. *Current Biology* **25**:R446–R447. DOI: https://doi.org/10.1016/j.cub.2015.04.018, PMID: 26035783

Gao J, Li Q, Wang Z, Zhou Y, Martelli P, Li F, Xiong Z, Wang J, Yang H, Zhang G. 2017. Sequencing, de novo assembling, and annotating the genome of the endangered Chinese crocodile lizard Shinisaurus crocodilurus. *GigaScience* **6**:gix041. DOI: https://doi.org/10.1093/gigascience/gix041

Geneious Prime. 2017. Geneious prime. 0.1. Graphstats. https://www.graphstats.net/geneious-prime

Germano DJ, Smith PT. 2010. Molecular evidence for parthenogenesis in the sierra garter snake, thamnophis couchii (Colubridae). *The Southwestern Naturalist* **55**:280–282. DOI: https://doi.org/10.1894/WL-29.1

Goodwillie C, Kalisz S, Eckert CG. 2005. The evolutionary enigma of mixed mating systems in plants: occurrence, theoretical explanations, and empirical evidence. *Annual Review of Ecology, Evolution, and Systematics* **36**:47–79. DOI: https://doi.org/10.1146/annurev.ecolsys.36.091704.175539

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* **29**:644–652. DOI: https://doi.org/10.1038/nbt.1883, PMID: 21572440

Groot TVM, Bruins E, Breeuwer JAJ. 2003. Molecular genetic evidence for parthenogenesis in the Burmese python, Python molurus bivittatus. *Heredity* **90**:130–135. DOI: https://doi.org/10.1038/sj.hdy.6800210, PMID: 12634818

Holt WV, Lloyd RE. 2010. Sperm storage in the vertebrate female reproductive tract: How does it work so well? *Theriogenology* **73**:713–722. DOI: https://doi.org/10.1016/j.theriogenology.2009.07.002

Holt C, Yandell M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* **12**:491. DOI: https://doi.org/10.1186/1471-2105-12-491, PMID: 22192575

Houck LD. 2009. Pheromone communication in amphibians and reptiles. *Annual Review of Physiology* **71**:161–176. DOI: https://doi.org/10.1146/annurev.physiol.010908.163134, PMID: 18847365

Ito M, Kaneko-Ishino T, Ishino F, Matsuhashi M, Yokoyama M, Katsuki M. 1991. Fate of haploid parthenogenetic cells in mouse chimeras during development. *The Journal of Experimental Zoology* **257**:178–183. DOI: https://doi.org/10.1002/jez.1402570206, PMID: 1990050

Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**:1236–1240. DOI: https://doi.org/10.1093/bioinformatics/btu031

Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, Okuno M, Yabana M, Harada M, Nagayasu E, Maruyama H, Kohara Y, Fujiyama A, Hayashi T, Itoh T. 2014. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Research* **24**:1384–1395. DOI: https://doi.org/10.1101/gr.170720.113, PMID: 24755901

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**:772–780. DOI: https://doi.org/10.1093/molbev/mst010

Kinney ME, Wack RF, Grahn RA, Lyons L. 2013. Parthenogenesis in a Brazilian rainbow boa (Epicrates cenchria cenchria). *Zoo Biology* **32**:172–176. DOI: https://doi.org/10.1002/zoo.21050, PMID: 23086743

Kratochvíl L, Vukić J, Červenka J, Kubička L, Johnson Pokorná M, Kukačková D, Rovatsos M, Piálek L. 2020. Mixed-sex offspring produced via cryptic parthenogenesis in a lizard. *Molecular Ecology* **29**:4118–4127. DOI: https://doi.org/10.1111/mec.15617, PMID: 32881125

Kroeger PT, Poureetezadi SJ, McKee R, Jou J, Miceli R, Wingert RA. 2014. Production of haploid zebrafish embryos by in vitro fertilization. *Journal of Visualized Experiments* **01**:51708. DOI: https://doi.org/10.3791/51708, PMID: 25046024

Kuraku S, Meyer A. 2009. The evolution and maintenance of Hox gene clusters in vertebrates and the teleost-specific genome duplication. *The International Journal of Developmental Biology* **53**:765–773. DOI: https://doi.org/10.1387/ijdb.072533km, PMID: 19557682

Lampert KP, Steinlein C, Schmid M, Fischer P, Schartl M. 2007. A haploid-diploid-triploid mosaic of the Amazon molly, Poecilia formosa. *Cytogenetic and Genome Research* **119**:131–134. DOI: https://doi.org/10.1159/000109629, PMID: 18160792

Larose C, Lavanchy G, Freitas S, Parker DJ, Schwander T. 2023. Facultative parthenogenesis: a transient state in transitions between sex and obligate asexuality in stick insects? *Peer Community Journal* **3**:283. DOI: https://doi.org/10.24072/pcjournal.283

Le Peuch C j, Picard A, Dorée M. 1985. Parthenogenetic activation decreases the polyphosphoinositide content of frog eggs. *FEBS Letters* **187**:61–64. DOI: https://doi.org/10.1016/0014-5793(85)81214-x, PMID: 2991011

Lenk P, Eidenmueller B, Staudter H, Wicker R, Wink M. 2005. A parthenogenetic Varanus. *Amphibia-Reptilia* **26**:507–514. DOI: https://doi.org/10.1163/156853805774806296

Levine BA, Schuett GW, Booth W. 2021. Exceptional long-term sperm storage by a female vertebrate. *PLOS ONE* **16**:e0252049. DOI: https://doi.org/10.1371/journal.pone.0252049, PMID: 34086677

Levine BA, Moresco A, Trout T, Schuett GW, Booth W. 2024. Female long-term sperm storage results in viable offspring in the Himalayan Mountain Pitviper, Ovophis monticola. *Zoo Biology* **43**:183–187. DOI: https://doi.org/10.1002/zoo.21820, PMID: 38234126

Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**:589–595. DOI: https://doi.org/10.1093/bioinformatics/btp698

Li H. 2016. seqtk. v1.2-r94. GitHub. https://github.com/lh3/seqtk

Lowe CH, Wright JW. 1966. Evolution of Parthenogenetic Species of Cnemidophorus (Whiptail Lizards) in Western North America. *Journal of the Arizona Academy of Science* **4**:81–87.

Lutes AA, Neaves WB, Baumann DP, Wiegraebe W, Baumann P. 2010. Sister chromosome pairing maintains heterozygosity in parthenogenetic lizards. *Nature* **464**:283–286. DOI: https://doi.org/10.1038/nature08818, PMID: 20173738

Lutes AA, Baumann DP, Neaves WB, Baumann P. 2011. Laboratory synthesis of an independently reproducing vertebrate species. PNAS **108**:9910–9915. DOI: https://doi.org/10.1073/pnas.1102811108, PMID: 21543715

Miles A. 2015. pysamstats. v0.24.3. GitHub. https://github.com/alimanfoo/pysamstats

Newton AA, Schnittker RR, Yu Z, Munday SS, Baumann DP, Neaves WB, Baumann P. 2016. Widespread failure to complete meiosis does not impair fecundity in parthenogenetic whiptail lizards. *Development* **143**:4486–4494. DOI: https://doi.org/10.1242/dev.141283, PMID: 27802173

Olsen WW, Marsden SJ. 1954. Natural parthenogenesis in turkey eggs. *Science* **120**:545–546. DOI: https://doi.org/10.1126/science.120.3118.545, PMID: 13195695

Olsen MW, Wilson SP, Marks HL. 1968. Genetic control of parthenogenesis in chickens. *The Journal of Heredity* **59**:41–42. DOI: https://doi.org/10.1093/oxfordjournals.jhered.a107639, PMID: 5656917

Olsen MW. 1973. Longevity and organ weights of Beltsville small white parthenogens and normal turkey males. *Poultry Science* **52**:666–670. DOI: https://doi.org/10.3382/ps.0520666, PMID: 4709417

MW Olsen, Avian parthenogensis. Agricultural Research Service USDA, ARS-NE 65, 1–82, (1975).

Parker HM, Kiess AS, Wells JB, Young KM, Rowe D, McDaniel CD. 2010. Genetic selection increases parthenogenesis in Chinese painted quail (Coturnix chinensis). *Poultry Science* **89**:1468–1472. DOI: https://doi.org/10.3382/ps.2009-00388

Pijnacker LP. 1969. Automictic parthenogenesis in the stick insectBacillus rossius Rossi (Cheleutoptera, phasmidae). *Genetica* **40**:393–399. DOI: https://doi.org/10.1007/BF01787364

Putnam NH, O'Connell BL, Stites JC, Rice BJ, Blanchette M, Calef R, Troll CJ, Fields A, Hartley PD, Sugnet CW, Haussler D, Rokhsar DS, Green RE. 2016. Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome Research* **26**:342–350. DOI: https://doi.org/10.1101/gr.193474.115, PMID: 26848124

pysam-developers. 2017. pysam. v0.12.0.1. GitHub. https://github.com/pysam-developers/pysam

Ramachandran R, McDaniel CD. 2018. Parthenogenesis in birds: a review. *Reproduction* **155**:R245–R257. DOI: https://doi.org/10.1530/REP-17-0728, PMID: 29559496

Reeder TW, Cole CJ, Dessauer HC. 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**:1–61. DOI: https://doi.org/10.1206/0003-0082(2002)365<0001:PROWLO>2.0.CO;2

Reynolds RG, Booth W, Schuett GW, Fitzpatrick BM, Burghardt GM. 2012. Successive virgin births of viable male progeny in the checkered gartersnake, *Thamnophis marcianus*. *Biological Journal of the Linnean Society* **107**:566–572. DOI: https://doi.org/10.1111/j.1095-8312.2012.01954.x

Rivera-Colón AG, Catchen J. 2022. Population genomics analysis with RAD, reprised: stacks 2. *Methods in Molecular Biology* **2498**:99–149. DOI: https://doi.org/10.1007/978-1-0716-2313-8_7, PMID: 35727543

Ryba NJ, Tirindelli R. 1997. A new multigene family of putative pheromone receptors. *Neuron* **19**:371–379. DOI: https://doi.org/10.1016/s0896-6273(00)80946-0, PMID: 9292726

Ryder OA, Thomas S, Judson JM, Romanov MN, Dandekar S, Papp JC, Sidak-Loftis LC, Walker K, Stalis IH, Mace M, Steiner CC, Chemnick LG. 2021. Facultative Parthenogenesis in California Condors. *The Journal of Heredity* **112**:569–574. DOI: https://doi.org/10.1093/jhered/esab052, PMID: 34718632

Sagata N. 1996. Meiotic metaphase arrest in animal oocytes: its mechanisms and biological significance. *Trends in Cell Biology* **6**:22–28. DOI: https://doi.org/10.1016/0962-8924(96)81034-8, PMID: 15157528

Sarvella P. 1973. Adult parthenogenetic chickens. *Nature* **243**:171. DOI: https://doi.org/10.1038/243171a0, PMID: 4706288

Schut E, Hemmings N, Birkhead TR. 2008. Parthenogenesis in a passerine bird, the Zebra Finch *Taeniopygia guttata*. *Ibis* **150**:197–199. DOI: https://doi.org/10.1111/j.1474-919X.2007.00755.x

Sever DM, Hamlett WC. 2002. Female sperm storage in reptiles. *The Journal of Experimental Zoology* **292**:187–199. DOI: https://doi.org/10.1002/jez.1154, PMID: 11754034

Shi P, Zhang J. 2007. Comparative genomic analysis identifies an evolutionary shift of vomeronasal receptor gene repertoires in the vertebrate transition from water to land. *Genome Research* **17**:166–174. DOI: https://doi.org/10.1101/gr.6040007, PMID: 17210926

Shibata H, Sakata S, Hirano Y, Nitasaka E, Sakabe A. 2017. Facultative parthenogenesis validated by DNA analyses in the green anaconda (Eunectes murinus). *PLOS ONE* **12**:e0189654. DOI: https://doi.org/10.1371/journal.pone.0189654, PMID: 29236745

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**:3210–3212. DOI: https://doi.org/10.1093/bioinformatics/btv351, PMID: 26059717

Slater GSC, Birney E. 2005. Automated generation of heuristics for biological sequence comparison. *BMC Bioinformatics* **6**:31. DOI: https://doi.org/10.1186/1471-2105-6-31, PMID: 15713233
Smit A, Hubley R. 2008. RepeatMasker Open-4.0. https://www.repeatmasker.org/ [Accessed September 18, 2019].
Smit A, Hubley R, Green R. 2008. Repeatmasker Open-4.0. 4.0.9. ISB. https://www.repeatmasker.org/
Sperling AL, Fabian DK, Garrison E, Glover DM. 2023. A genetic basis for facultative parthenogenesis in *Drosophila*. *Current Biology* **33**:3545–3560.. DOI: https://doi.org/10.1016/j.cub.2023.07.006, PMID: 37516115
Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**:1312–1313. DOI: https://doi.org/10.1093/bioinformatics/btu033
Straube N, Lampert KP, Geiger MF, Weiß JD, Kirchhauser JX. 2016. First record of second-generation facultative parthenogenesis in a vertebrate species, the whitespotted bambooshark Chiloscyllium plagiosum. *Journal of Fish Biology* **88**:668–675. DOI: https://doi.org/10.1111/jfb.12862, PMID: 26727105
Streisinger G, Walker C, Dower N, Knauber D, Singer F. 1981. Production of clones of homozygous diploid zebra fish (*Brachydanio rerio*). *Nature* **291**:293–296. DOI: https://doi.org/10.1038/291293a0, PMID: 7248006
Su CY, Menuz K, Carlson JR. 2009. Olfactory perception: receptors, cells, and circuits. *Cell* **139**:45–59. DOI: https://doi.org/10.1016/j.cell.2009.09.015, PMID: 19804753
Tanaka M, Yamaha E, Arai K. 2004. Survival capacity of haploid-diploid goldfish chimeras. *Journal of Experimental Zoology. Part A, Comparative Experimental Biology* **301**:491–501. DOI: https://doi.org/10.1002/jez.a.48, PMID: 15181643
trinityrnaseq. 2015. trinityrnaseq. v2.0.6. GitHub. https://github.com/trinityrnaseq/trinityrnaseq
UniProt Consortium T. 2018. UniProt: the universal protein knowledgebase. *Nucleic Acids Research* **46**:2699. DOI: https://doi.org/10.1093/nar/gky092
Vanzolini PE. 1993. Reptilians Too: *Biology of Whiptail Lizards*. (Genus *Cnemidophorus*.) John W. Wright and Laurie J. Vitt, Eds. University of Oklahoma and Oklahoma Museum of Natural History, Norman, 1993. xvi, 417 pp., illus. $29. Herpetologists' League Special Publication no. 3. Based on a symposium, Norman, OK, Aug. 1984. *Science* **262**:1282–1283. DOI: https://doi.org/10.1126/science.262.5137.1282.b
Walker JM, Abuhteba RM, Cordes JE. 1991. Morphological and experimental verification of hybridization between all-female cnemidophorus laredoensis b and gonochoristic cnemidophorus gularis (Squamata: Teiidae). *Herpetologica* **47**:152–164. DOI: https://doi.org/10.2307/3671928
Watts PC, Buley KR, Sanderson S, Boardman W, Ciofi C, Gibson R. 2006. Parthenogenesis in Komodo dragons. *Nature* **444**:1021–1022. DOI: https://doi.org/10.1038/4441021a, PMID: 17183308
Whitehead MR, Lanfear R, Mitchell RJ, Karron JD. 2018. Plant mating systems often vary widely among populations. *Frontiers in Ecology and Evolution* **6**:00038. DOI: https://doi.org/10.3389/fevo.2018.00038
Wolf DP. 1974. The cortical response in *Xenopus laevis* ova. *Developmental Biology* **40**:102–115. DOI: https://doi.org/10.1016/0012-1606(74)90112-2, PMID: 4213423
Zhang L, Huang Y, Wang M, Guo Y, Liang J, Yang X, Qi W, Wu Y, Si J, Zhu S, Li Z, Li R, Shi C, Wang S, Zhang Q, Tang Z, Wang L, Li K, Fei J-F, Lan G. 2019. Development and genome sequencing of a laboratory-inbred miniature pig facilitates study of human diabetic disease. *iScience* **19**:162–176. DOI: https://doi.org/10.1016/j.isci.2019.07.025

			MS1		MS2		MS6		MS7	
122	*A. marmoratus*	Female	217		259	272	174		244	282
4238	*A. marmoratus*	Female	217		269	275	158	175	267	298
4239	*A. marmoratus*	Female	217		259	269	158	175	267	298
8449	Offspring 1	Female	217		259		174		282	
8450	Offspring 2	Female	217		272		174		244	

			MS8		MS9		MS10		MS16	
122	*A. marmoratus*	Female	113		425	434	330	353	200	
4238	*A. marmoratus*	Female	114		415	421	346		200	
4239	*A. marmoratus*	Female	114		415	421	346		200	
8449	Offspring 1	Female	114		433		330		200	
8450	Offspring 2	Female	114		433		330		200	

Figure 1—figure supplement 1. Microsatellite analysis of eight loci for the three female *Aspidoscelis marmoratus* within the enclosure and the two offspring hatched in January 2009. Only alleles that are unique to a potential mother are colored. Single nucleotide differences in size are common binning artifacts and, therefore, are not scored as different alleles. The two offspring were homozygous for all eight markers. At three markers (MS7, MS9, MS10), both offspring inherited alleles only found in animal 122.

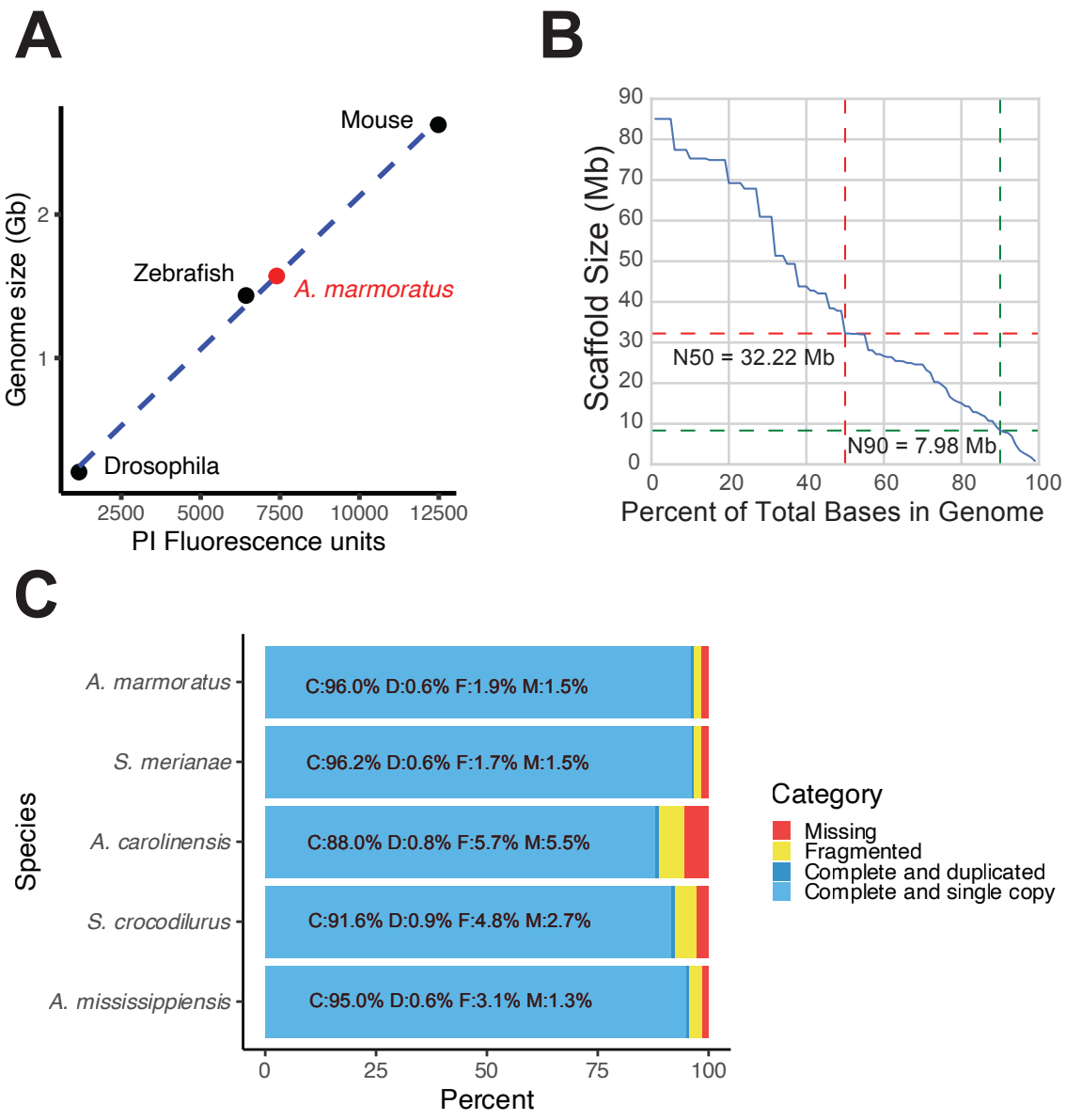

Figure 2—figure supplement 1. Genome assembly of *Aspidoscelis marmoratus*. (**A**) Genome size estimation using fluorescence-activated cell sorting (FACs). The standard curve was generated using known genome sizes from fruit flies, zebrafish, and mouse, and then comparing fluorescent intensity with that of erythrocytes from *A. marmoratus*. (**B**) N(x)% plot shows the ordered scaffold lengths plotted against the cumulative fraction of the genome. Dashed lines mark the N50 and N90 values. (**C**) BUSCO analysis of 2586 conserved vertebrate coding genes for *A. marmoratus*, *Salvator merianae*, *Anolis carolinensis*, *Shinisaurus crocodilurus*, and *Alligator mississippiensis*. Values for *S. crocodilurus* and *A. mississippiensis* were taken from ***Gao et al., 2017***.

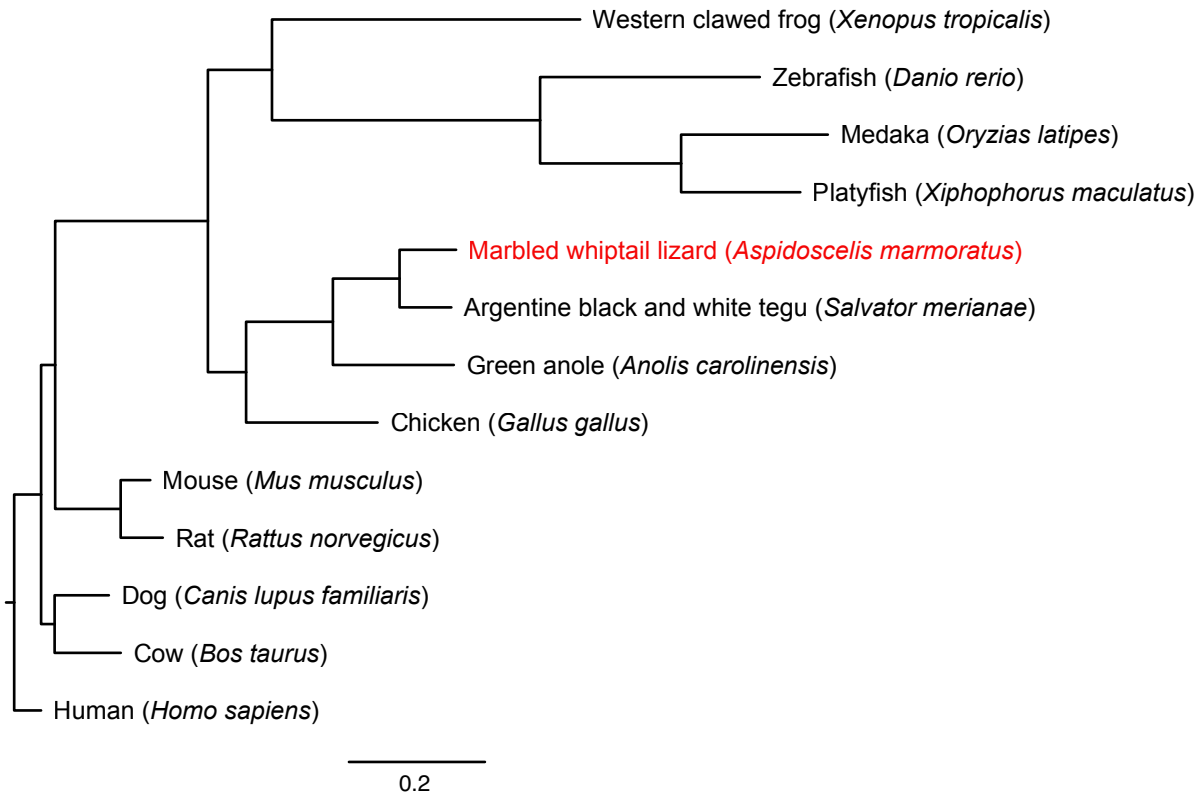

Figure 2—figure supplement 2. Maximum likelihood tree for 13 vertebrate genomes, based on 1333 single-copy BUSCOs detected across all species analyzed.

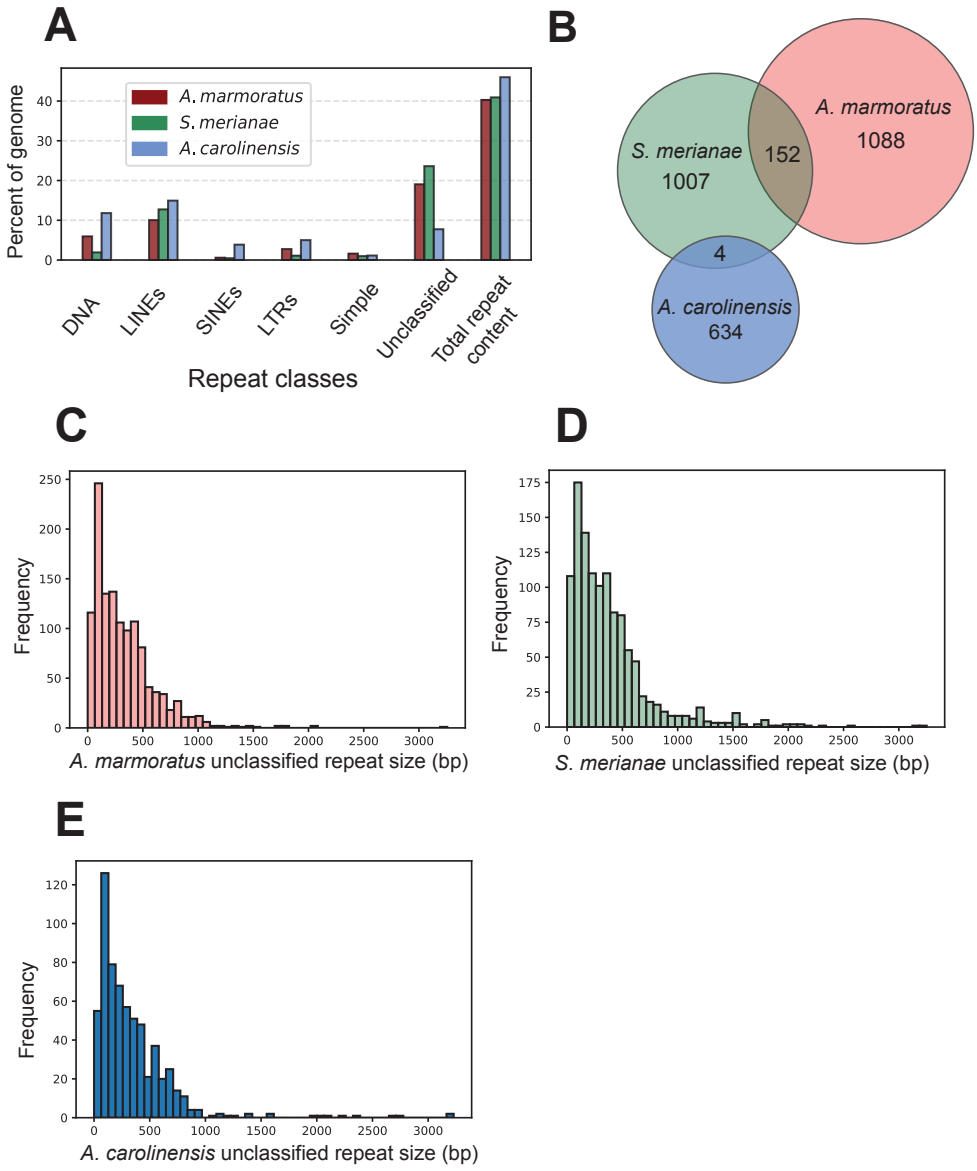

Figure 2—figure supplement 3. Identification of unclassified repeat elements in the *A. marmoratus* genome. (**A**) Repetitive DNA content in *Aspidoscelis marmoratus*, *Salvator merianae*, and *Anolis carolinensis* genomes separated by repeat classes as defined by RepeatMasker. (**B**) Overlap of unclassified repetitive elements between the three genomes as in (**A**). (**C–E**) Distribution of the sizes of unclassified repetitive elements, respectively, for *A. marmoratus*, *S. merianae*, and *A. carolinensis*.

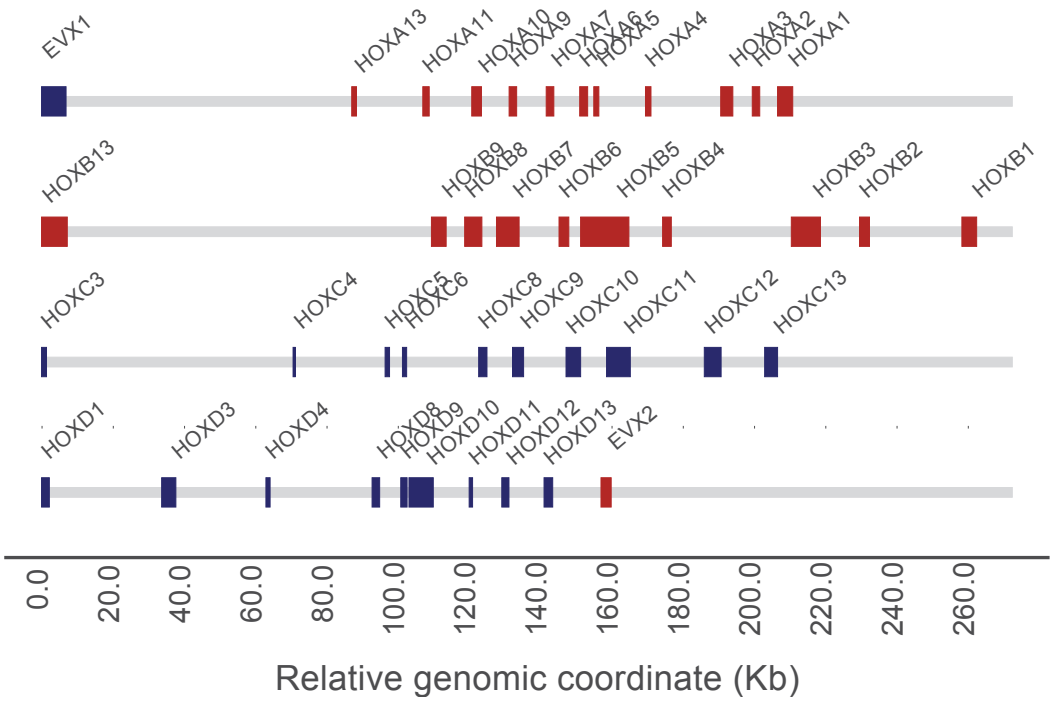

Figure 2—figure supplement 4. *Aspidoscelis marmoratus* HOX gene clusters. Red blocks indicate a region of homology was found on the sense strand, blue indicates the antisense strand. EVX1 and EVX2 are also shown, due to their relevance to HOX cluster evolution.

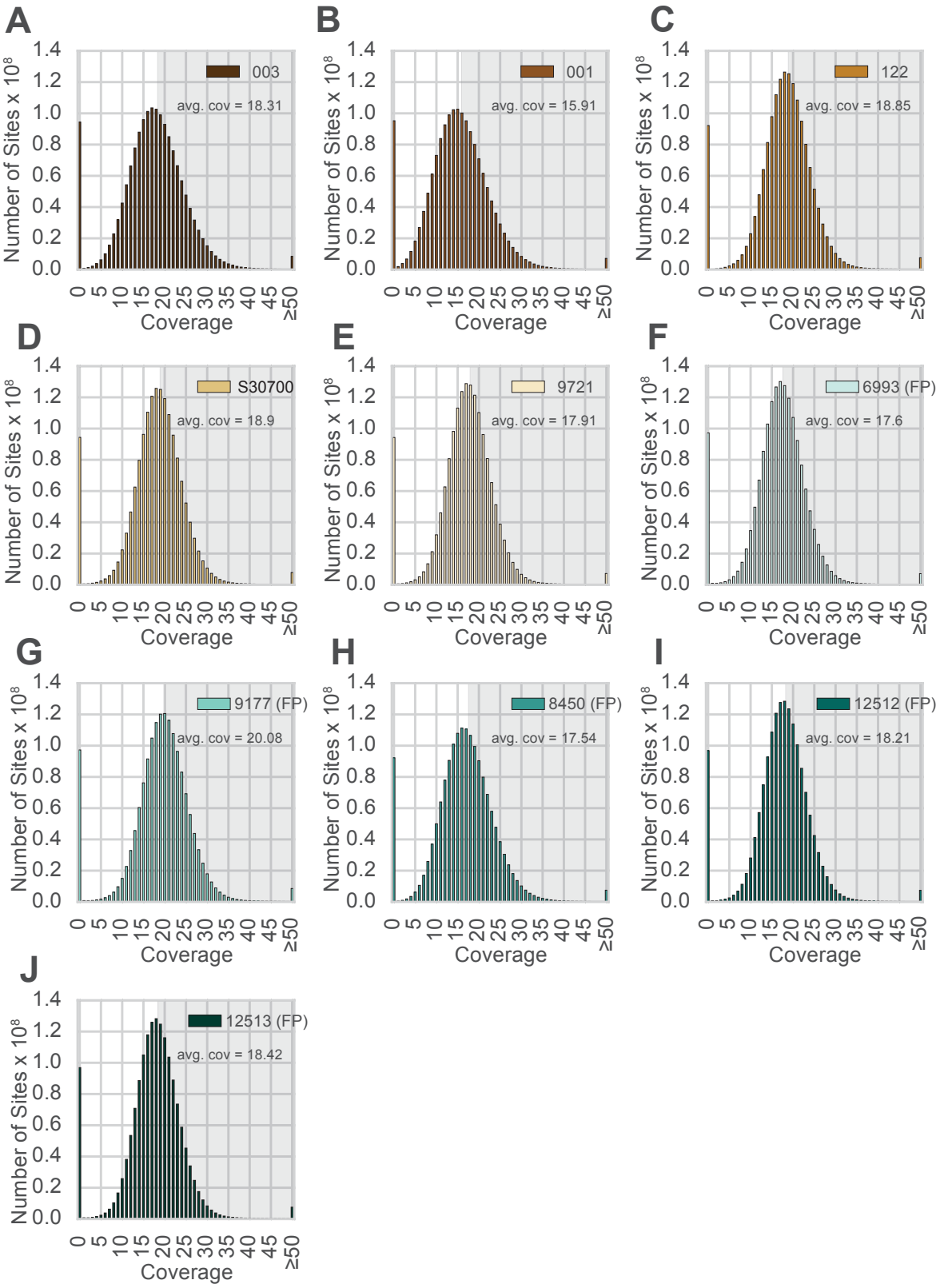

Figure 2—figure supplement 5. Distribution of sequence coverage across the genome for each animal. (**A–J**) Animal IDs are shown in the top right corner of each panel, with animals produced by facultative parthenogenesis denoted with 'facultative parthenogenesis (FP).' The distributions were generated by examining the coverage at every position in the reference assembly. Regions with coverage greater than the mean are shaded in gray in each panel.

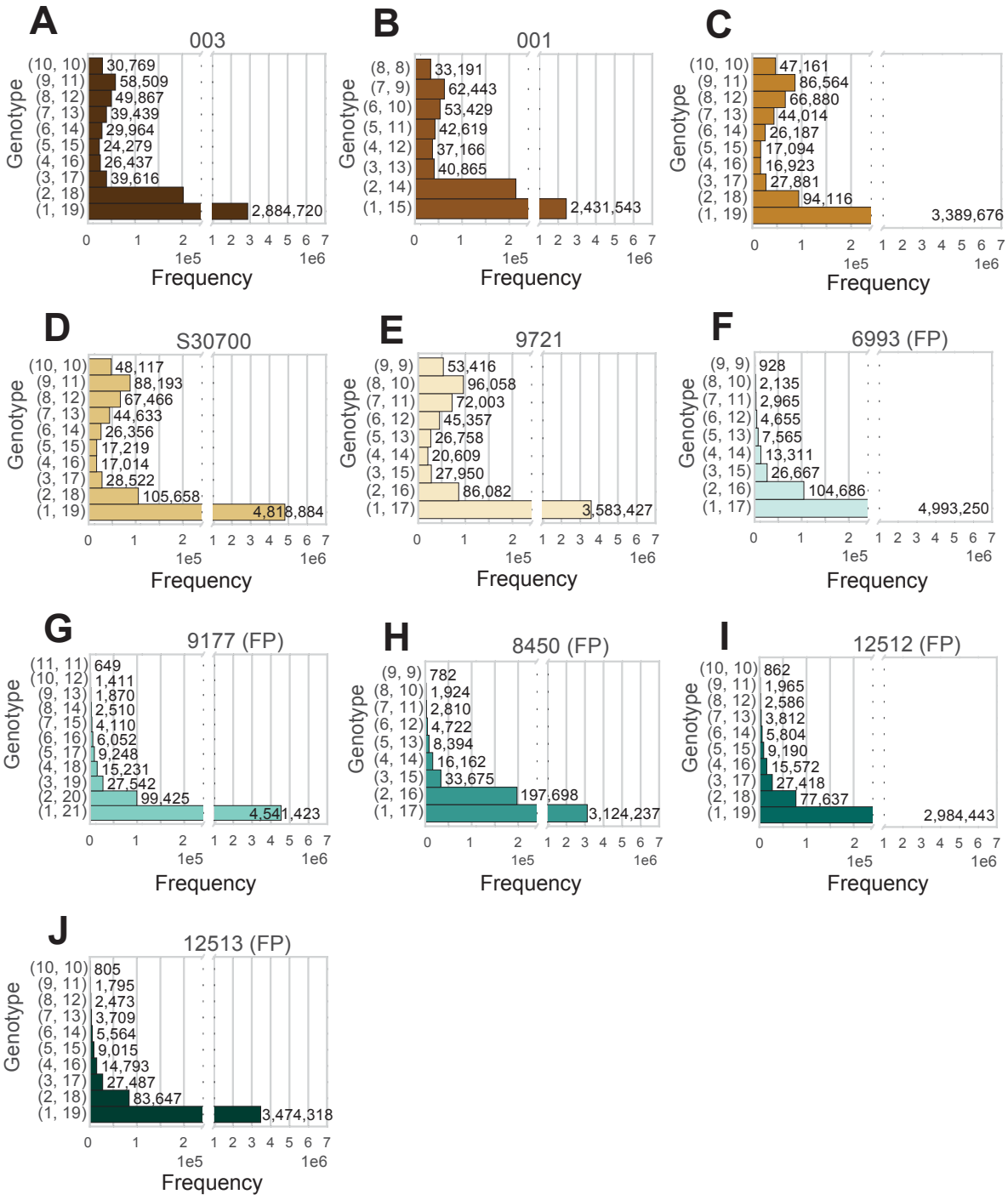

Figure 2—figure supplement 6. Analysis of all positions in the genome at average coverage, for which reads support exactly two alleles. **(A–J)** The two numbers to the left of each row indicate how many reads support each allele. Animal IDs are shown above each panel, with animals produced by facultative parthenogenesis denoted with 'facultative parthenogenesis (FP)'.

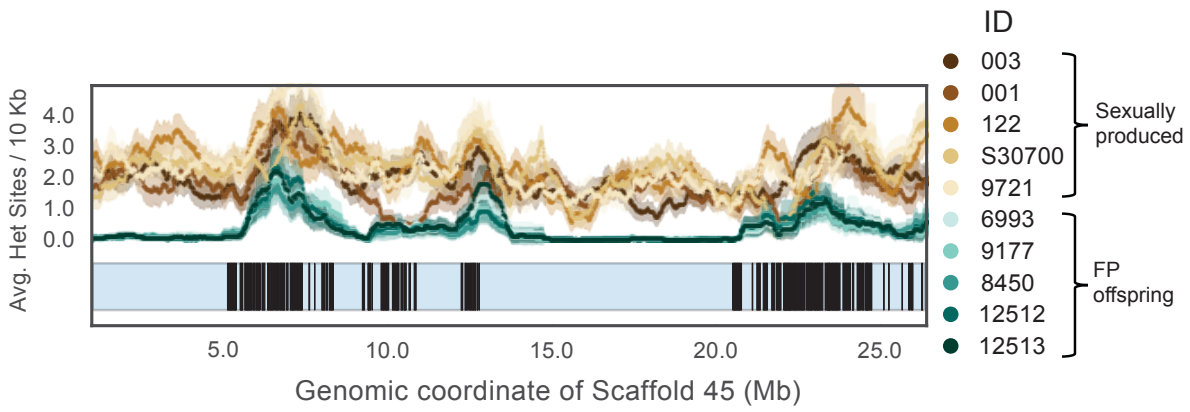

Figure 2—figure supplement 7. A 1Mb sliding window for Scaffold 45 showing the annotated vomeronasal 2 receptors homologs. A scaffold ideogram is shown in blue and coordinates of 167 annotated V2r26 homologs are shown in black.

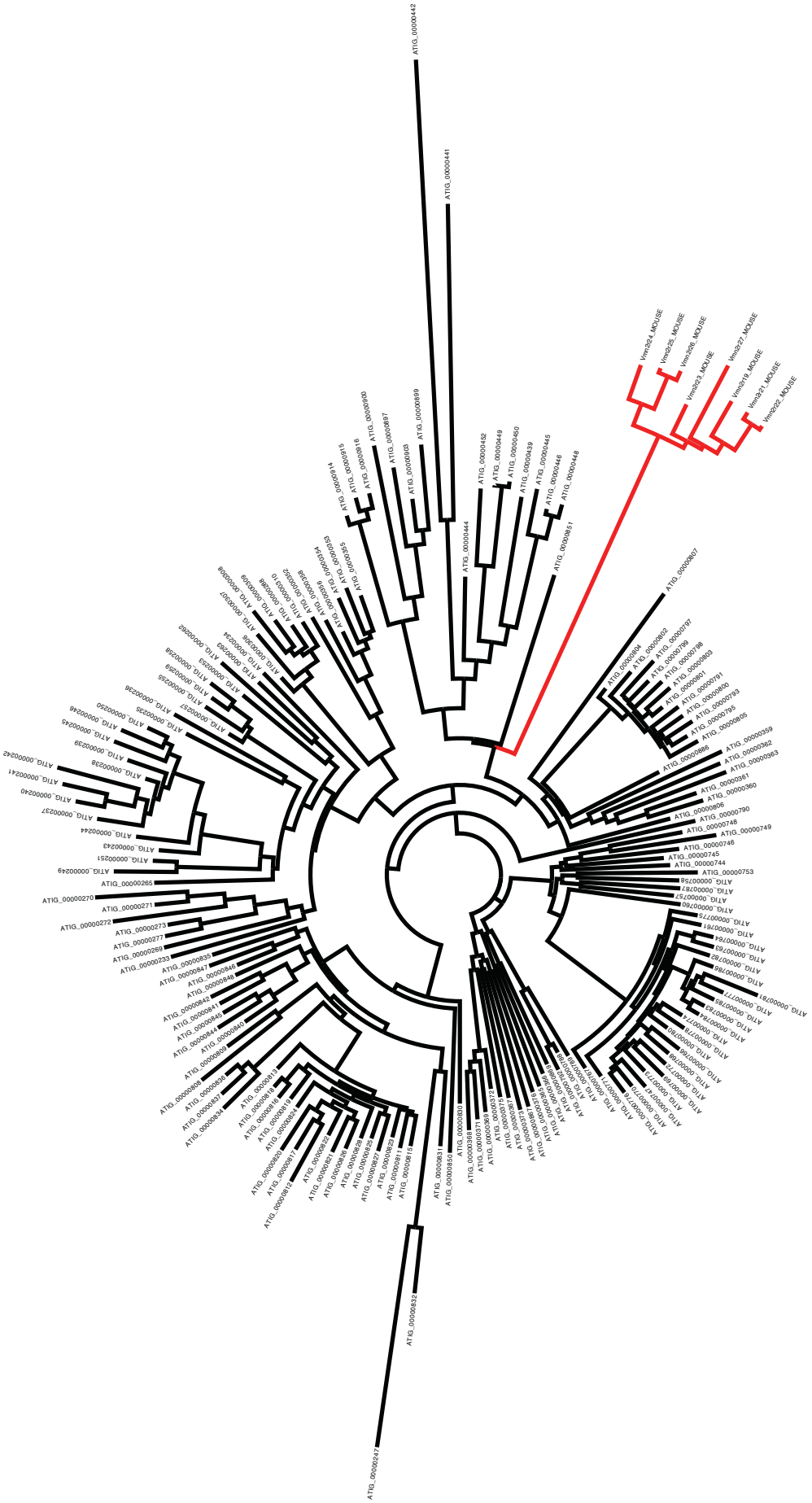
ATIG_00000442
ATIG_00000441
Vmn2r24_MOUSE
Vmn2r25_MOUSE
Vmn2r26_MOUSE
Vmn2r23_MOUSE
Vmn2r27_MOUSE
Vmn2r19_MOUSE
Vmn2r21_MOUSE
Vmn2r22_MOUSE
ATIG_00000452
ATIG_00000449
ATIG_00000460
ATIG_00000439
ATIG_00000445
ATIG_00000446
ATIG_00000448
ATIG_00000444
ATIG_00000899
ATIG_00000897
ATIG_00000900
ATIG_00000903
ATIG_00000916
ATIG_00000915
ATIG_00000914
ATIG_00000851
ATIG_00000807
ATIG_00000804
ATIG_00000805
ATIG_00000797
ATIG_00000799
ATIG_00000798
ATIG_00000803
ATIG_00000801
ATIG_00000791
ATIG_00000790
ATIG_00000793
ATIG_00000795
ATIG_00000805
ATIG_00000886
ATIG_00000359
ATIG_00000362
ATIG_00000363
ATIG_00000361
ATIG_00000360
ATIG_00000806
ATIG_00000790
ATIG_00000748
ATIG_00000749
ATIG_00000746
ATIG_00000745
ATIG_00000744
ATIG_00000753
ATIG_00000756
ATIG_00000757
ATIG_00000760
ATIG_00000775
ATIG_00000761
ATIG_00000764
ATIG_00000762
ATIG_00000763
ATIG_00000766
ATIG_00000768
ATIG_00000781
ATIG_00000777
ATIG_00000783
ATIG_00000784
ATIG_00000774
ATIG_00000780
ATIG_00000778
ATIG_00000769
ATIG_00000772
ATIG_00000771
ATIG_00000770
ATIG_00000773
ATIG_00000767
ATIG_00000765
ATIG_00000779
ATIG_00000782
ATIG_00000785
ATIG_00000786
ATIG_00000787
ATIG_00000788
ATIG_00000789
ATIG_00000792
ATIG_00000794
ATIG_00000796
ATIG_00000802
ATIG_00000800
ATIG_00000808
ATIG_00000809
ATIG_00000810
ATIG_00000811
ATIG_00000812
ATIG_00000813
ATIG_00000814
ATIG_00000815
ATIG_00000816
ATIG_00000817
ATIG_00000818
ATIG_00000819
ATIG_00000820
ATIG_00000821
ATIG_00000822
ATIG_00000823
ATIG_00000824
ATIG_00000825
ATIG_00000826
ATIG_00000827
ATIG_00000828
ATIG_00000829
ATIG_00000830
ATIG_00000831
ATIG_00000850
ATIG_00000849
ATIG_00000848
ATIG_00000847
ATIG_00000846
ATIG_00000845
ATIG_00000844
ATIG_00000843
ATIG_00000842
ATIG_00000841
ATIG_00000840
ATIG_00000839
ATIG_00000838
ATIG_00000837
ATIG_00000836
ATIG_00000835
ATIG_00000269
ATIG_00000253
ATIG_00000277
ATIG_00000273
ATIG_00000272
ATIG_00000271
ATIG_00000270
ATIG_00000265
ATIG_00000249
ATIG_00000251
ATIG_00000243
ATIG_00000244
ATIG_00000237
ATIG_00000240
ATIG_00000241
ATIG_00000242
ATIG_00000239
ATIG_00000245
ATIG_00000246
ATIG_00000250
ATIG_00000236
ATIG_00000235
ATIG_00000227
ATIG_00000258
ATIG_00000259
ATIG_00000254
ATIG_00000255
ATIG_00000256
ATIG_00000257
ATIG_00000260
ATIG_00000262
ATIG_00000261
ATIG_00000263
ATIG_00000264
ATIG_00000266
ATIG_00000267
ATIG_00000268
ATIG_00000252
ATIG_00000539
ATIG_00000538
ATIG_00000536
ATIG_00000535
ATIG_00000534
ATIG_00000533
ATIG_00000532
ATIG_00000530
ATIG_00000529
ATIG_00000528
ATIG_00000527
ATIG_00000526
ATIG_00000525
ATIG_00000524
ATIG_00000523
ATIG_00000522
ATIG_00000521
ATIG_00000520
ATIG_00000519
ATIG_00000518
ATIG_00000517
ATIG_00000516
ATIG_00000515
ATIG_00000514
ATIG_00000513
ATIG_00000512
ATIG_00000511
ATIG_00000510
ATIG_00000509
ATIG_00000508
ATIG_00000507
ATIG_00000506
ATIG_00000505
ATIG_00000504
ATIG_00000503
ATIG_00000502
ATIG_00000501
ATIG_00000500
ATIG_00000499
ATIG_00000498
ATIG_00000497
ATIG_00000496
ATIG_00000495
ATIG_00000494
ATIG_00000493
ATIG_00000492
ATIG_00000491
ATIG_00000490
ATIG_00000489
ATIG_00000488
ATIG_00000487
ATIG_00000486
ATIG_00000485
ATIG_00000484
ATIG_00000483
ATIG_00000482
ATIG_00000481
ATIG_00000480
ATIG_00000479
ATIG_00000478
ATIG_00000477
ATIG_00000476
ATIG_00000475
ATIG_00000474
ATIG_00000473
ATIG_00000472
ATIG_00000471
ATIG_00000470
ATIG_00000469
ATIG_00000468
ATIG_00000467
ATIG_00000466
ATIG_00000465
ATIG_00000464
ATIG_00000463
ATIG_00000462
ATIG_00000461
ATIG_00000460
ATIG_00000459
ATIG_00000458
ATIG_00000457
ATIG_00000456
ATIG_00000455
ATIG_00000454
ATIG_00000453
ATIG_00000452
ATIG_00000451
ATIG_00000450
ATIG_00000449
ATIG_00000448
ATIG_00000447
ATIG_00000446
ATIG_00000445
ATIG_00000444
ATIG_00000443
ATIG_00000442
ATIG_00000441
ATIG_00000440
ATIG_00000439
ATIG_00000438
ATIG_00000437
ATIG_00000436
ATIG_00000435
ATIG_00000434
ATIG_00000433
ATIG_00000432
ATIG_00000431
ATIG_00000430
ATIG_00000429
ATIG_00000428
ATIG_00000427
ATIG_00000426
ATIG_00000425
ATIG_00000424
ATIG_00000423
ATIG_00000422
ATIG_00000421
ATIG_00000420
ATIG_00000419
ATIG_00000418
ATIG_00000417
ATIG_00000416
ATIG_00000415
ATIG_00000414
ATIG_00000413
ATIG_00000412
ATIG_00000411
ATIG_00000410
ATIG_00000409
ATIG_00000408
ATIG_00000407
ATIG_00000406
ATIG_00000405
ATIG_00000404
ATIG_00000403
ATIG_00000402
ATIG_00000401
ATIG_00000400
ATIG_00000399
ATIG_00000398
ATIG_00000397
ATIG_00000396
ATIG_00000395
ATIG_00000394
ATIG_00000393
ATIG_00000392
ATIG_00000391
ATIG_00000390
ATIG_00000389
ATIG_00000388
ATIG_00000387
ATIG_00000386
ATIG_00000385
ATIG_00000384
ATIG_00000383
ATIG_00000382
ATIG_00000381
ATIG_00000380
ATIG_00000379
ATIG_00000378
ATIG_00000377
ATIG_00000376
ATIG_00000375
ATIG_00000374
ATIG_00000373
ATIG_00000372
ATIG_00000371
ATIG_00000370
ATIG_00000369
ATIG_00000368
ATIG_00000367
ATIG_00000366
ATIG_00000365
ATIG_00000364
ATIG_00000363
ATIG_00000362
ATIG_00000361
ATIG_00000360
ATIG_00000359
ATIG_00000358
ATIG_00000357
ATIG_00000356
ATIG_00000355
ATIG_00000354
ATIG_00000353
ATIG_00000352
ATIG_00000351
ATIG_00000350
ATIG_00000349
ATIG_00000348
ATIG_00000347
ATIG_00000346
ATIG_00000345
ATIG_00000344
ATIG_00000343
ATIG_00000342
ATIG_00000341
ATIG_00000340
ATIG_00000339
ATIG_00000338
ATIG_00000337
ATIG_00000336
ATIG_00000335
ATIG_00000334
ATIG_00000333
ATIG_00000332
ATIG_00000331
ATIG_00000330
ATIG_00000329
ATIG_00000328
ATIG_00000327
ATIG_00000326
ATIG_00000325
ATIG_00000324
ATIG_00000323
ATIG_00000322
ATIG_00000321
ATIG_00000320
ATIG_00000319
ATIG_00000318
ATIG_00000317
ATIG_00000316
ATIG_00000315
ATIG_00000314
ATIG_00000313
ATIG_00000312
ATIG_00000311
ATIG_00000310
ATIG_00000309
ATIG_00000308
ATIG_00000307
ATIG_00000306
ATIG_00000305
ATIG_00000304
ATIG_00000303
ATIG_00000302
ATIG_00000301
ATIG_00000300
ATIG_00000299
ATIG_00000298
ATIG_00000297
ATIG_00000296
ATIG_00000295
ATIG_00000294
ATIG_00000293
ATIG_00000292
ATIG_00000291
ATIG_00000290
ATIG_00000289
ATIG_00000288
ATIG_00000287
ATIG_00000286
ATIG_00000285
ATIG_00000284
ATIG_00000283
ATIG_00000282
ATIG_00000281
ATIG_00000280
ATIG_00000279
ATIG_00000278
ATIG_00000276
ATIG_00000275
ATIG_00000274
ATIG_00000254
ATIG_00000247

Figure 2—figure supplement 8. Phylogenetic tree showing the evolutionary distance between mouse V2Rs (red branches) and *Aspidoscelis marmoratus* homologs (black branches). There are 10 additional copies of Vmn2r26 homologs found on Scaffold 45 not identified by MAKER2 (177 total), while genome-wide, we annotated 323 copies. Phylogenetic analysis of the 323 copies with eight other mouse vomeronasal two receptors show that the mouse receptors form their own distinct clade from *A. marmoratus* and there is substantial sequence diversity among the *A. marmoratus* homologs. A manual search for additional copies of Vmn2r26 resulted in 478 further hits. Vmn2r26 belongs to a family of receptors known as V2Rs and this estimate on the number of Vmn2r26 homologs ranks *A. marmoratus* as one of the species with the largest expansion of V2Rs (***Shi and Zhang, 2007***; ***Brykczynska et al., 2013***).

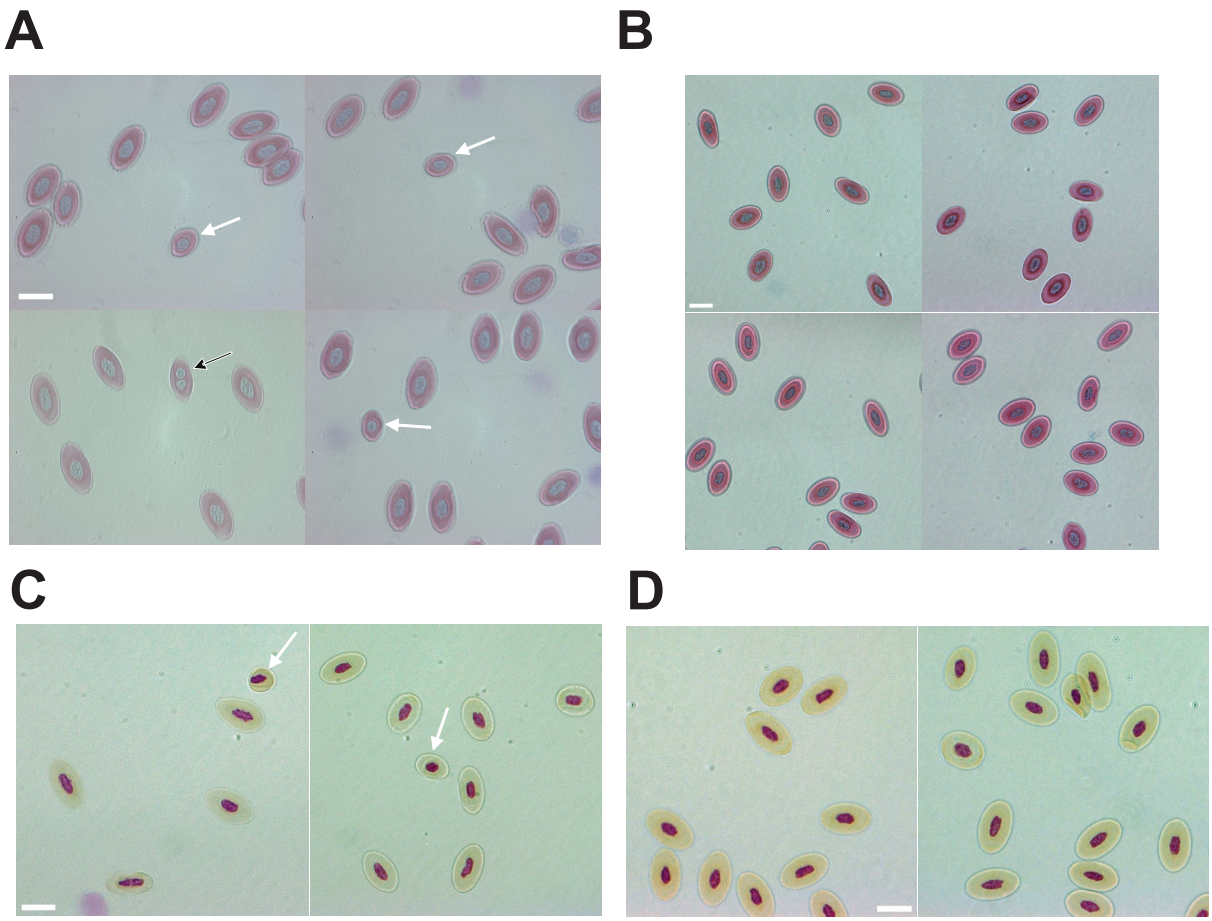

Figure 4—figure supplement 1. Mixoploidy detected in both *A. marmoratus* and *A. arizonae*. (**A**) Giemsa staining of erythrocytes from a one-year-old facultative parthenogenetic *Aspidoscelis marmoratus* (ID 25384). Diploid (n=1661), smaller haploid (white arrow, n=17), and binucleated (black arrow, n=1) cells are evident. Scale bar corresponds to 10 μm. (**B**) Giemsa staining of erythrocytes from a sexually produced *A. marmoratus* (ID 23880). All cells are diploid (n=924). Scale bar corresponds to 10 μm. (**C**) Feulgen staining of erythrocytes from a one-year-old facultative parthenogenetic *A. arizonae* (ID 23507). Diploid (n=130) and smaller haploid (white arrow, n=2) cells are evident. Scale bar corresponds to 10 μm. (**D**) Feulgen staining of erythrocytes from a sexually produced *A. arizonae* (ID 26714). All cells are diploid (n=1506). Scale bar corresponds to 10 μm.

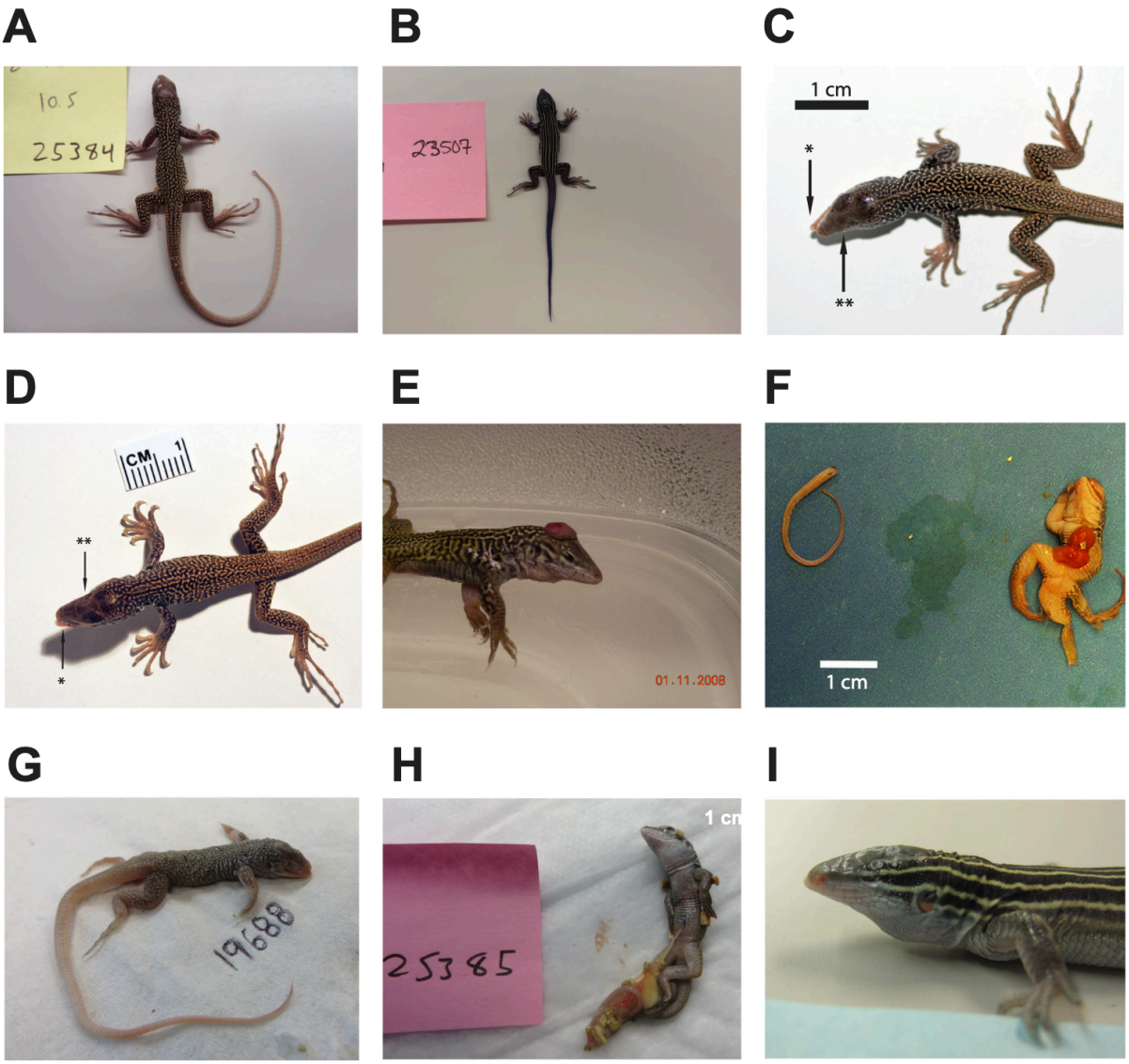

Figure 4—figure supplement 2. Animals produced by facultative parthenogenesis. (**A**) *Aspidoscelis marmoratus* 25384. No abnormal phenotypes noted. (**B**) *Aspidoscelis arizonae* 23507. No abnormal phenotypes noted. (**C**) *A. marmoratus* 12512. Misalignment of jaws (*). Agenesis of left eye (**). (**D**) *A. marmoratus* 12513. Misalignment of jaws (*). Agenesis of right eye (**). (**E**) *A. marmoratus* 6993. Animal was cut from egg with exposed brain. (**F**) *A. marmoratus* 8394. Did not hatch. Missing a leg, failure of abdomen closure leading to exposed organs, face abnormalities, and hunched back. (**G**) *A. marmoratus* 19688. Did not hatch. Multiple craniofacial deformities. (**H**) *A. marmoratus* 25385. Partially emerged from egg with egg yolk still attached. (**I**) *A. arizonae* 16216. Agenesis of left eye.

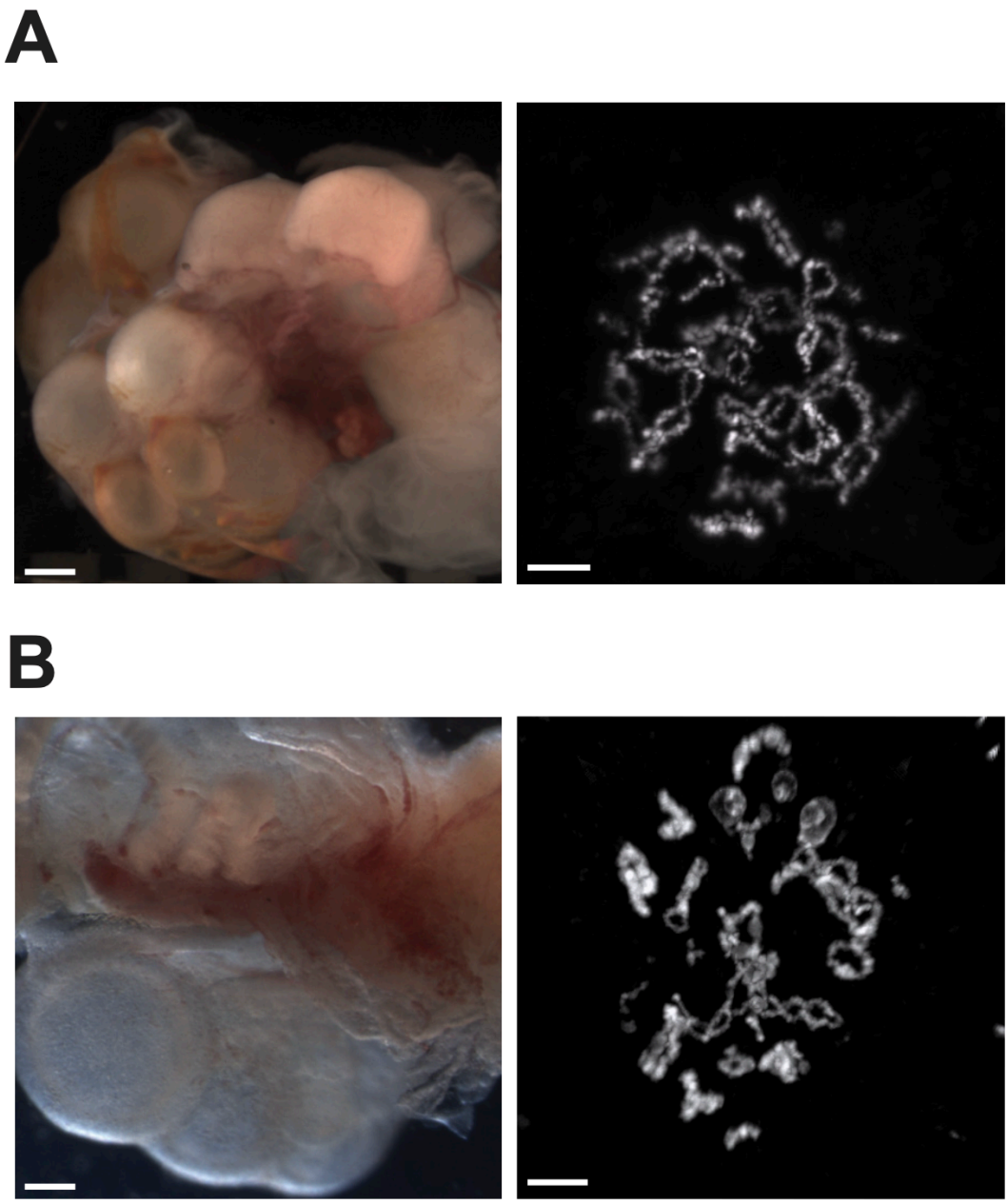

Figure 4—figure supplement 3. Ovaries of *Aspidoscelis marmoratus* facultative parthenogenesis (FP) animal 8450 and germinal vesicles of FP sister 8449 revealed no differences in structure and anatomy compared to fertile sexually reproducing animals. (**A**) Ovary (left; ID 8450) and diplotene stage germinal vesicle (right; ID 8449). The lizard ovary image is taken with a Leica M205FA dissection microscope. Scale bar corresponds to 1 mm. The germinal vesicle is stained with DAPI. Scale bar corresponds to 10 μm. (**B**) Ovary (left; ID 13103) and diplotene stage germinal vesicle (right ID 5359) from sexually produced *A. marmoratus*. The lizard ovary image is taken with a Leica as described previously. Scale bar corresponds to 1 mm. The germinal vesicle is stained with Acridine Orange (0.01%). Scale bar corresponds to 10 μm.

Supplementary file 1. *A. marmoratus* genome assembly statistics

Statistic	Meraculous	HiRise	Units
N50	1600688	32220929	bp
N90	413554	7979424	bp
max_scaffold	11507399	85027298	bp
mean_scaffold	188287.18	428523.466	bp
median_scaffold	1703	1450	bp
min_scaffold	1000	927	bp
number of scaffolds	8705	3826	bp
perc_A	27.06	27.04	%
perc_C	20.14	20.12	%
perc_G	20.13	20.13	%
perc_N	5.64	5.67	%
perc_T	27.04	27.04	%
perc_genome_covered_by_scaffolds_greater_than_100kb	97.99	99.46	%
perc_genome_covered_by_scaffolds_greater_than_10kb	99.28	99.63	%
perc_genome_covered_by_scaffolds_greater_than_1mb	69.83	98.47	%
scaffolds_greater_than_100Mb	0	0	
scaffolds_greater_than_100kb	1588	133	
scaffolds_greater_than_10Mb	2	45	
scaffolds_greater_than_10kb	2076	223	
scaffolds_greater_than_1Mb	545	90	
total bases	1639039918	1639530780	bp
non_n_GC%	42.67	42.67	%
non_n_AT%	57.33	57.33	%

Supplementary file 2. Trinity assembly statistics

	Trinity initial assembly	Final filtered Trinity assembly
Mean contig (bp)	1146.41188	1332.98712
Median contig (bp)	505	686
Min contig (bp)	224	174
Max contig (bp)	24454	24454
Number of contigs	513013	119728
N50 (bp)	2415	2490
N90 (bp)	393	512
Total bases (bp)	588124198	159595882

Supplementary file 3. MAKER2 summary

Feature	Source	Count
CDS	MAKER2	222450
exon	MAKER2	164669
five_prime_UTR	MAKER2	22402
gene	MAKER2	25856
intron	MAKER2	189732
mRNA	MAKER2	44461
three_prime_UTR	MAKER2	21729

Supplementary file 4. Information for all *A. marmoratus* animals sequenced

ID	Analysis type	Library type	Sex	Tissue	Sequencing	Number of read pairs	Read length
8450	Genome assembly	40 kb mate pair BfaI	Female	Liver	Illumina MiSeq	7325967	251
8450	Genome assembly	40 kb mate pair CVIQI	Female	Liver	Illumina MiSeq	7578268	251
8450	Genome assembly	5 kb mate pair	Female	Liver	Illumina HiSeq 2500	68124235	151
8450	Genome assembly	8 kb mate pair	Female	Liver	Illumina HiSeq 2500	62138640	151
8450	Genome assembly	2-15 kb mate pair	Female	Liver	Illumina HiSeq 2500	67949265	151
8450	Genome assembly	Chicago library	Female	Liver	Illumina HiSeq 2500	52923610	101
8450	Genome assembly	Chicago library	Female	Liver	Illumina HiSeq 2500	207139609	101
8450	Genome assembly	Paired-end	Female	Liver	Illumina HiSeq 2500	376374707	251
S30700	Full genome sequencing	Paired-end	Female	Tail	Illumina HiSeq 2500	106787117	151
9177	Full genome sequencing	Paired-end	Female	Tail	Illumina HiSeq 2500	113372074	151
6993	Full genome sequencing	Paired-end	Female	Tail	Illumina HiSeq 2500	100116690	151
12512	Full genome sequencing	Paired-end	Female	Tail	Illumina HiSeq 2500	102760203	151
12513	Full genome sequencing	Paired-end	Female	Tail	Illumina HiSeq 2500	103885909	151
9721	Full genome sequencing	Paired-end	Female	Tail	Illumina HiSeq 2500	101265101	151
122	Full genome sequencing	Paired-end	Female	Tail	Illumina HiSeq 2500	321903175	151
001	Full genome sequencing	Paired-end	Male	Liver	Illumina HiSeq 2500	96536963	151
003	Full genome sequencing	Paired-end	Male	Liver	Illumina HiSeq 2500	108407900	151
8450	Full genome sequencing	Paired-end	Female	Liver	Illumina HiSeq 2500	103851051	151
11225	Transcriptome assembly	Poly A stranded RNA-seq	Male	Blood	Illumina HiSeq 2500	260779309	101
S237-2	Transcriptome assembly	Poly A stranded RNA-seq	Unkown	Embryo	Illumina HiSeq 2500	253832665	101

Supplementary file 5. Whole genome sequencing alignment statistics

Animal ID	12512	12513	S30700	6993	9177	9721	001	003	122	8450
reads total	2.1E+08	2.12E+08	2.18E+08	2.04E+08	2.32E+08	2.07E+08	1.97E+08	2.21E+08	6.56E+08	2.12E+08
secondary	4371764	4293355	4427220	4120929	4761643	4246124	3932037	4507406	12408444	4081669
supplementary	0	0	0	0	0	0	0	0	0	0
duplicates	1829281	1841233	1848550	2314935	2169035	2151550	1523198	1283723	14760637	1368644
mapped	2.07E+08	2.09E+08	2.15E+08	2.01E+08	2.29E+08	2.04E+08	1.93E+08	2.18E+08	6.5E+08	2.08E+08
paired in sequencing	2.06E+08	2.08E+08	2.14E+08	2E+08	2.27E+08	2.03E+08	1.93E+08	2.17E+08	6.44E+08	2.08E+08
read1	1.03E+08	1.04E+08	1.07E+08	1E+08	1.13E+08	1.01E+08	96536963	1.08E+08	3.22E+08	1.04E+08
read2	1.03E+08	1.04E+08	1.07E+08	1E+08	1.13E+08	1.01E+08	96536963	1.08E+08	3.22E+08	1.04E+08
properly paired	1.88E+08	1.9E+08	1.95E+08	1.82E+08	2.07E+08	1.85E+08	1.84E+08	2.06E+08	5.97E+08	1.96E+08
with itself and mate mapped	2.02E+08	2.04E+08	2.09E+08	1.95E+08	2.22E+08	1.99E+08	1.88E+08	2.13E+08	6.36E+08	2.02E+08
singletons	1212491	1504914	1467678	1679077	1523933	1408129	1072182	1099052	1912936	993280
with mate mapped to a different chr	12512434	12380462	12772150	11827242	13772076	12316058	3751076	5431708	34786180	5413508
with mate mapped to a different chr (mapQ>=5)	7450173	7412852	7583394	7060453	8164436	7358325	2149235	3193570	21253145	3196901

Supplementary file 6. Animals confirmed by microsatellite analysis to be of FP origin

Species	Animal ID	Clutch ID	Laid-Begin	Laid-End	Hatch	Notes
A. marmoratus	6993	X48	21-Oct-07	22-Oct-07		Animal was cut from egg, alive, with brain exposed from skull
A. marmoratus	8377	J68	21-Oct-08	25-Oct-08		
A. marmoratus	8394	P68	29-Oct-08	2-Nov-08		Missing a leg, organs exposed at abdomen, face abnormalities, and hunched back
A. marmoratus	8449	M69	18-Nov-08	22-Nov-08	25-Jan-09	Appeared to have balance issues; hatched with substance still attached to abdomen
A. marmoratus	8450	M69	18-Nov-08	22-Nov-08	25-Jan-09	Appeared to have balance issues
A. arizonae	8677	K72	7-Feb-09	10-Feb-09		
A. marmoratus	9070	V78	12-Jun-09	16-Jun-09		
A. marmoratus	9177	F80	30-Jul-09	3-Aug-09	10-Oct-09	
A. marmoratus	12512	L131	26-Apr-12	30-Apr-12	8-Jul-12	Deformed jaw; missing left eye
A. marmoratus	12513	L131	26-Apr-12	30-Apr-12	8-Jul-12	Deformed jaw, missing right eye
A. arizonae	16215	B187	7-Mar-14	11-Mar-14	2-May-14	Right side of skull seems to have stunted growth compared to left side, lower mandible protrudes out slightly, right eye seems slightly smaller than left
A. arizonae	16216	B187	7-Mar-14	11-Mar-14	2-May-14	Missing left eye, malformed jaw, shortened torso
A. marmoratus	19687	T237	19-May-15	23-May-15	1-Aug-15	
A. marmoratus	19688	T237	19-May-15	23-May-15		Multiple craniofacial deformities
A. arizonae	23304	Q288	11-Sep-16	15-Sep-16	15-Nov-16	
A. arizonae	23507	D291	11-Oct-16	15-Oct-16	15-Dec-16	
A. marmoratus	24514	L303	12-Jan-17	16-Jan-17		
A. arizonae	25339	I313	3-Mar-17	7-Mar-17		Deformed snout, missing left eye
A. marmoratus	25384	R313	5-Mar-17	9-Mar-17	6-Jun-17	
A. marmoratus	25385	R313	5-Mar-17	9-Mar-17	6-Jun-17	Found partially emerged from egg, barely moving, had large mass of egg yolk still attached
A. arizonae	29606	E364	27-Apr-18	1-May-18	3-Jul-18	
A. arizonae	29607	E364	27-Apr-18	1-May-18	3-Jul-18	Congenital shortened torso; hatched alive and died immediately
A. marmoratus	30050	A369	29-Jun-18	3-Jul-18		No eyes

Supplementary file 7. Microsatellite primer information

Microsatellite	Oligo 1 sequence (5' → 3')	Oligo 2 sequence (5' → 3')
MS1	TGCATGATGGAGGAATCTTC	CTAGTGGTGATAGAAACATGG
MS2	GATCCTGTTGCCGTTTGGAC	ATGCTTCTAGATGAACCCAC
MS3	CTCTCCTTGCAAACAGCAGG	ACATGGGTGAGTTGAGGGTG
MS5	TATCTCCTGGCTATCAAGAC	TTAGTTAGCCCTTAGCACAG
MS6	CACACCCATATTATAAGTGG	CATTCAGATGAAACCTAACC
MS7	AACTAAGTGCTAAGTGTGAC	ACAGTCTTAGAGATCACAAG
MS8	ACACCCAAAGTCCTCAACAG	CTAGTACATGTGTAAGGGTG
MS9	GCAGACTCATAGTTGAAGTG	TCTGGGAATACCAGGTTCAC
MS10	GACCAATAATGTGGAAGCTG	ACATGGCTGAGTAATTGGTG
MS12	TACCCACCTGGAGATGTTTAG	AGGACGCCTTAAAATAGGAAG
MS14	TGGAGGCAGTCTTGGTATC	GAACATTGACCGCATCAC
MS15	TTAAAGCAGAGGTCAGGTTATC	GATGGAAGAATAGGATGATGAA
MS16	TTTAATGCATCCACTGAGTC	GGAATATAGTGGCATATCAG
MS17	AATCCTGAACCTACGGTAAGC	TGCCAGAAAATAGAGGGAAG
MS18	AATTAATGTGCAGCACTAT	GGCAGTTTTTCAGCTAAG
MS19	AAAAAGAAAAGGAAGAACTAA	TGAGACAAGTTGGGTAGA

4. Chromosome-level assemblies of five whiptail lizard species

In this chapter, I detail the primary data that went into the generation of five different whiptail lizard genome assemblies.

Aspidoscelis marmoratus (Marbled whiptail)

In our manuscript “Post-meiotic mechanism of facultative parthenogenesis in gonochoristic whiptail lizard species,” we showed genome-wide homozygosity in *A. marmoratus* (coming from the Larin word for “marbled”) by assembling a highly contiguous genome (AspMar1.0). While there are positions in the genome that indicates heterozygous calls (Chapter 3, Figure 2B), the expectation is that if polar body fusion was occurring, we would see blocks of heterozygosity interspersed throughout the genome. The lack of heterozygous regions indicate that any heterozygous signals in animals produced through facultative parthenogenesis (FP) should be considered noise attributed to assembly, sequencing, and PCR errors. Another possibility would be somatic mutations.

While the AspMar1.0 genome is highly contiguous with an N50 length of 32.22 Mb, our goal is to be able to get to chromosome-sized scaffolds because this would be a closer digital representation of the genome *in vivo*. To this end, we further sequenced the *A. marmoratus* reference animal (ID 8450) with Hi-C. Double stranded DNA is compact into chromatin in an organized way, guided by DNA-associated proteins. The seminal work by Dekker et al. (2002) showed that the interactions between genomic loci can be quantified in a high-throughput manner by cross-linking segments of the genome that are physically near each other called chromosome conformation capture (3C). Since then, various editions of 3C have been developed including Hi-C, where crosslinking and Illumina paired-end sequencing is used to capture all possible pairwise interactions between different regions of the genome (Lajoie et al., 2015; Lieberman-Aiden et al., 2009).

The use of Hi-C in genome research allows researchers to study chromatin structure, chromatin interactions, and topologically associating domains (TADs) that play a role in gene regulation (Sefer, 2022). In addition, Hi-C can be taken advantage of as well to generate more contiguous genome assemblies by using the frequency of contacts between genomic loci to not only assign smaller contigs to scaffolds but also order and orient these contigs as well. To this aim, we used Dovetail Genomics (San Diego, USA) to generate Hi-C libraries from the FP animal 8450. These libraries were subsequently sequenced at Stowers Institute

for Medical Research (SIMR). Dovetail Genomics used their in-house HiRise assembly program (Putnam et al., 2016) to scaffold the AspMar1.0 genome assembly to a newer version: AspMarm2.0 (Figure 5). This version is the most contiguous genome assembly for *A. marmoratus* with the majority of the sequences found on 23 scaffolds, representing the 23 chromosomes in the species (C. H. Lowe & Wright, 1966). The inclusion of Chicago and Hi-C reads were essential to get a chromosome-level assembly (Figure 6).

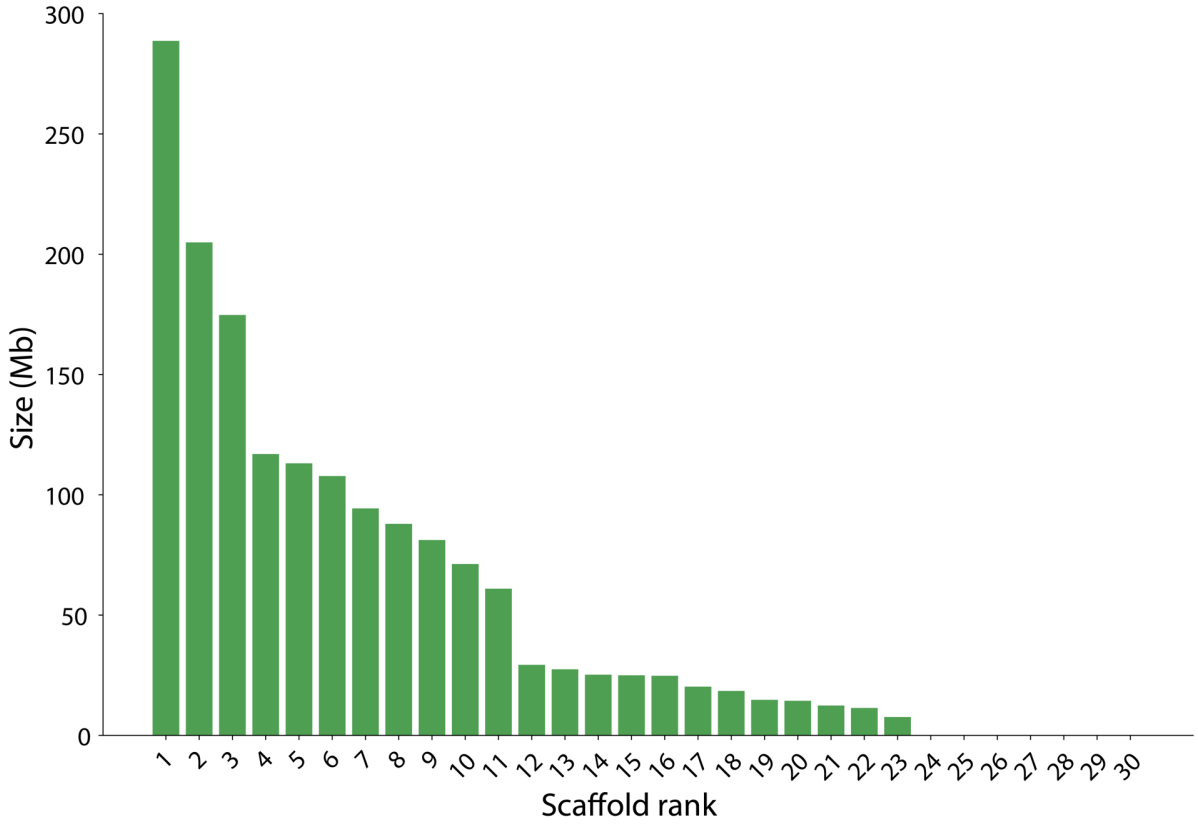

Figure 5. Sizes of the 30 largest scaffolds in the *A. marmoratus* assembly (AspMarm2.0) produced by Dovetail Genomics' HiRise. Over 99.58% of the sequences are anchored on 23 scaffolds.

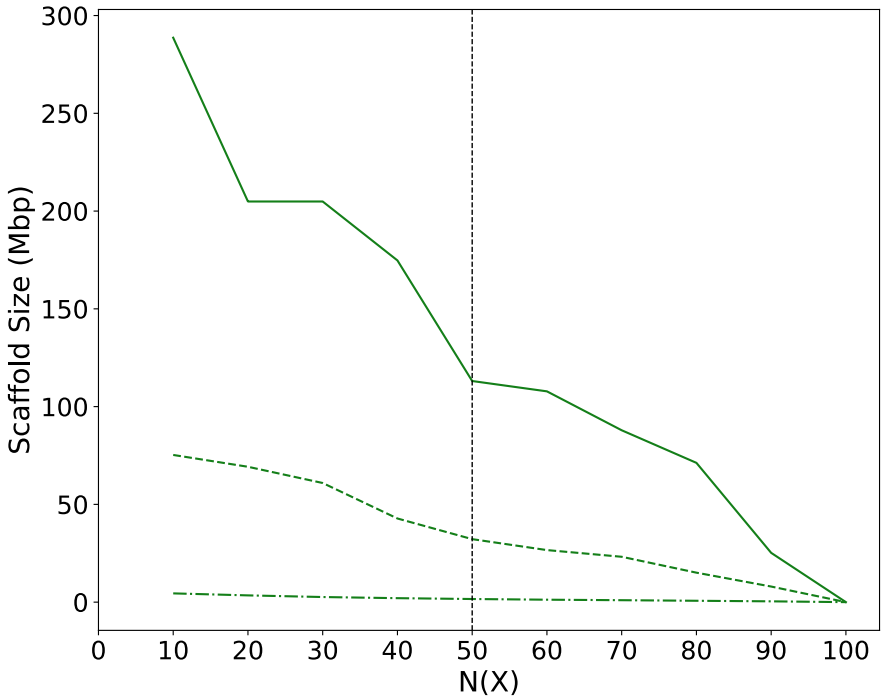

Figure 6. N(X) plot showing the improvement in the genome assembly of *A. marmoratus*. Vertical line is N50, where 50% of the genome sequences are on scaffolds at the particular size on the y-axis or greater. Bottom line = assembly with just Meraculous; N50 length is 1.6 Mb. Middle line = assembly including Chicago-reads with HiRise; N50 length is 32.2 Mb. Top line = assembly including Hi-C reads; N50 is chromosome 5 with a length of 113 Mb.

Genome Assembly Booster (GAB)

Among the handful of scaffolding tools available the leverages Hi-C data to make more contiguous genome assemblies (and hopefully to chromosome-level), SALSA (Ghurye et al., 2017) and LACHESIS (Burton et al., 2013) are two well-known programs. As Dovetail Genomics' HiRise program is proprietary and therefore we need to pay for their services, we wanted to test SALSA and LACHESIS and utilize either of them for genome assemblies. Unfortunately, we were not able to install LACHESIS and the use of the depreciated python2 in SALSA means these were not programs that would work for our purposes. While one can argue that many programmers still use python2, the development of bioinformatic software should aim to keep the user in mind and therefore Aaron Odell, a bioinformatician in our group, developed a new scaffolding program: Genome Assembly Booster (GAB).

The program is dependent on the output of HiC-Pro, a well-established bioinformatics software package, that first processes the Hi-C data (Servant et al., 2015). Hi-C data is first aligned to a primary genome assembly, artifacts and noise due to small contigs are removed

(scaffolds less than 10 Kb are recommended to be removed post alignment of Hi-C data), and the resulting contact matrix between genomic loci are normalized. HiC-Pro normalizes via a matrix balancing method known as ICE (iterative correction and eigenvector decomposition) normalization (reference in HiCPro as ICED). This approach to Hi-C data processing breaks up the primary genome assembly into non-overlapping equal sized pieces. The size of these pieces is commonly referred to as the “resolution size.” Read pairs are then counted between any two pieces of the genome, and the resulting square matrix is then normalized to generate the final contact map (Imakaev et al., 2012).

As defined by the contact map, genomic loci are clustered into discrete scaffold/chromosome groups by contact values alone. An advantage of the algorithm is it is non-heuristic and the number of chromosome groups are not pre-defined by the user. By not incorporating scaffold information and only looking at contact values between non-overlapping equal sized loci, our process allows for an unbiased approach in clustering genomic loci into chromosome sized groups. Scaffold information can then be leveraged to assess the accuracy of the algorithm as well as identify potential errors in the primary genome assembly itself. We found that UPGMA (average) clustering of this distance matrix using the python SciPy package followed by the dendrogram “`count_sort ='ascending'`” function and option, works remarkably well in clustering loci into chromosome groups.

Average clustering ultimately results in a dendrogram in which groups (made up with genomic loci) within a dendrogram can be separated and called chromosome groups. However, separating these groups are a challenge for any scaffolding program because chromosome sizes within a species are not all the same. Cytological studies of chromosomes in eukaryotes reveal diverse karyotypes, not only in the total size of the genomes but also in number of chromosomes (Mayrose & Lysak, 2020). At one extreme end, the ciliate *Oxytricha trifallax* has a haploid genome size of 50 Mb spread over 16,000 chromosomes (Swart et al., 2013), while on the other extreme, telomere fusion and centromere inactivation led to the Jack jumper ant (*Myrmecia pulosula*) having all of its genome located on one pair of chromosome (Crosland & Crozier, 1986; Imai & Taylor, 1989). Furthermore, within a single species, the size of each chromosome can be vastly different, in addition to differences in chromosome architecture, such as the location of centromeres and banding (Schalch & Steiner, 2017). Within vertebrates, birds and reptiles share similar karyotypes (Waters et al., 2021) in that their nuclei contain chromosomes of vastly different sizes: macrochromosomes (larger than 150 Mb) and microchromosomes (smaller than 50 Mb). The origin of these different sized chromosomes in these species is most likely through a combination of inter- and intrachromosomal rearrangements, transposable elements, differences in recombination, and genome duplication to just name a few reasons (Muffato et

al., 2023; Sacerdot et al., 2018). Not only does this diversity in karyotypes make these genomes harder to reconstruct from an evolutionary history viewpoint, but it also makes genome assemblies difficult to separate out the chromosome groups because separation needs to be done dynamically.

From a dendrogram, large sized chromosome groups are easy to pick out, even by eye, but smaller chromosomes are harder to separate (Figure 7). As observed in other species, smaller chromosomes tend to interact more with each other and therefore lead to a higher noise to signal ratio (Duan et al., 2010). To address this issue, a Hidden Markov Model (HMM) based method was developed to separate out the larger chromosomes. An additional second method based on a series of hypergeometric tests also resulted in comparable results but has the advantage of requiring less parameter inputs. As both methods perform well, they are both available options within GAB (with the default being the hypergeometric method).

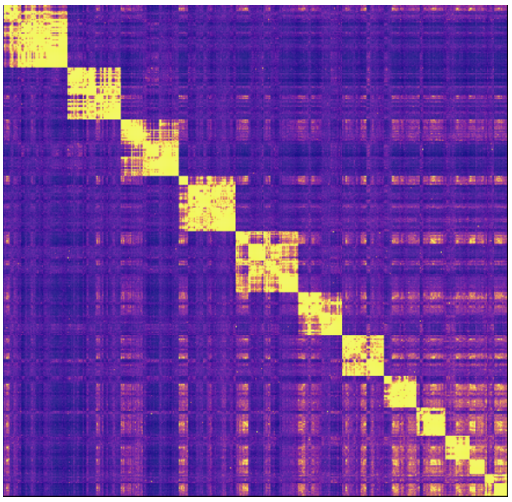

Figure 7. Example of an average clustering plot from the Genome Assembly Booster (GAB). Loci are ordered on the left and top-most margins and interactions between loci are colored by yellow in the resulting matrix. Larger chromosome groups are clearly identifiable (top-left), while smaller chromosome groups appear to interact with each other more and therefore make separating them harder (lower-right).

Once the HMM or hypergeometric test has defined all chromosome groups within the allotted fraction of the data (the large chromosome groups), we then switch over to a concept known as modularity which has origins in computer science to detect community structures in networks. Modularity has also been adopted in evolutionary biology to help identify and explain gene networks and compartmentalizing complex systems into distinct parts (Hatleberg & Hinman, 2021; Melo et al., 2016). To partition the remaining genomic loci into discrete communities (ie. chromosome groups), the Louvain method for community detection is used (Blondel et al., 2008). In general, this method treats each locus as a node

and measures the density of interactions between each node, as compared to interactions to other communities. Different ordering initiations are used to get to a convergent result to avoid any potential local optima.

After establishing chromosome groups, the algorithm continues with ordering and orienting the scaffolds. For a chromosome group, the number of possible permutations of different orderings and orientations is given by the following equation:

$$\frac{(N!)2^N}{2}$$

Equation 1. Number of possible permutations of different orderings and orientations for a given number of scaffolds assigned to a chromosome group.

1. *N* represents the number of scaffolds on a chromosome.
2. *N*! represents the number of permutations of arranging *N* scaffolds.
3. 2^N represents the number of ways each scaffold can be oriented. Since each scaffold can be oriented in two ways (forward or reverse), there are 2 choices for each scaffold.
4. The division by 2 is there because each ordering and orientation is counted twice due to the orientation being reversible.

Among all the different possibilities, we define the best permutation as the maximum value of a cost function that weights contact values that are close in linear space more heavily than values that are further away. For an example chromosome group with 10 scaffolds assigned, the number of possible order and orientation to test is already 1,857,946,600 which easily becomes an unfeasible computation problem with more scaffolds assigned to a chromosome group. To reduce the search space, we first generate a submatrix containing only the largest scaffolds (default is 6) within that chromosome and brute force calculate all possible permutations within this submatrix and choose the one with the best score. The reasoning for this is that the largest scaffolds offer a much higher signal to noise ratio than do the smaller scaffolds. Therefore, even with missing scaffolds in between, the correct permutation of the largest scaffolds with respect to each other can be determined. After this step, the remaining scaffolds are added one at a time. With each new addition, we keep the original order and orientation of the scaffolds currently contained within the submatrix. Again, the best scoring permutation found and repeat until all scaffolds have been added. This is repeated for all chromosome groups.

With the establishment of GAB by Aaron Odell (Figure 8), we were excited to use this algorithm for other whiptail species as described later in this chapter. We tested GAB with the same data input for the *A. marmoratus* genome assembly performed by Dovetail Genomcs with HiRise and had similar results, confirming GAB is as suitable for scaffolding and getting to a chromosome-level genome assembly as HiRise. The finer details of the program, such as the cost function, and the quality checks that were perform will be detailed in a manuscript in preparation.

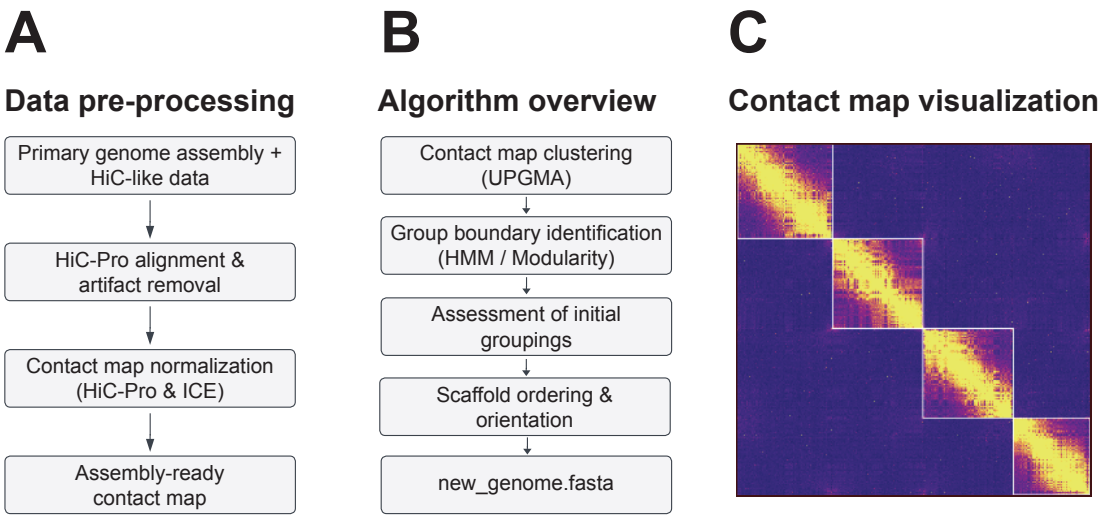

Figure 8. Outline of the Genome Assembly Booster (GAB). The program is written in python3 and utilizes a primary assembly and Hi-C data that has been processed with HiC-Pro. Grouping scaffolds together, the algorithm takes advantage of the Hi-C contact map to order and orient the scaffolds. The output includes contact map visualization to see the chromosomes assembled and a final fasta file.

Aspidoscelis arizonae (Arizona striped whiptail)

In Chapter 3, we also have documented cases of *A. arizonae* reproducing through FP. Dovetail Genomics was provided with liver tissue from a wild-caught male (ID 17121; Alamogordo, New Mexico, USA) to make Chicago libraries that were subsequently sequenced at SIMR (Table 1). We also performed whole-genome sequencing (266 M paired-end reads, 2 x 250) on the animal as well to make a primary assembly using Meraculous (N50 length of 30.4 Kb). This was provided to Dovetail Genomics to use in their HiRise program along with the Chicago reads (256 M paired-end reads, 2 x 101). Similar to the Chicago-level *A. marmoratus* assembly provided by Dovetail Genomics, the N50 length for the *A. arizonae* genome was elevated to 35.8 Mb with a final genome size of 1.528 Gb.

Phase Genomics was hired to perform the Hi-C library preparation and sequencing from a tail tissue of the same animal (134 M paired-end reads, 2 x 80). Using the GAB, we were able to group the majority of the scaffolds into 23 chromosomes (Figure 9).

Table 1. Statistics on the three versions of the *A. arizonae* assembly.

Assembly version	Initial primary assembly	Dovetail	GAB
Assembly software	Meraculous	HiRise	GAB
Reads used	Illumina paired-end	Illumina paired-end, Chicago	Illumina paired-end, Chicago, Hi-C
Longest scaffold	420,575 Kb	108.89 Mb	211.80 Mb
Number of scaffolds	61880	3277	3150
Contig N50 length	30.4 Kb	34.1 Kb	106.147 Mb

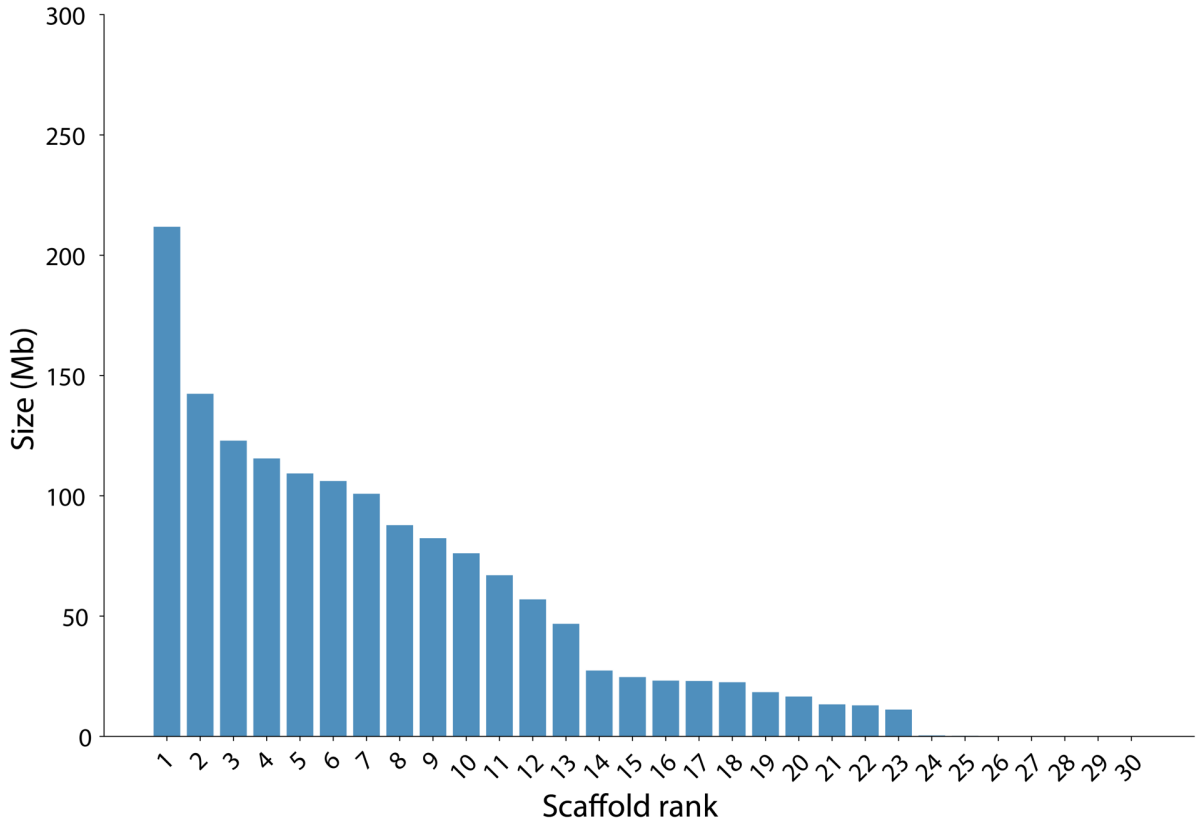

Figure 9. Sizes of the 30 largest scaffolds in the *A. arizonae* assembly produced by GAB. Over 99.4% of the sequences are anchored on 23 scaffolds.

Aspidoscelis tigris (Western whiptail)

As part of the California Conservation Genomics Project (CCGP; https://www.ccgproject.org/), *A. tigris* was chosen as one of the 100+ species of plants and animals to have a genome assembly generated and perform whole genome sequencing on a number of individuals to “manage regional biodiversity” in California. A reference genome for a wild male *A. tigris* individual was generated by the CCGP and uploaded to NCBI with

two versions: primary (JALMGA000000000, GCA_023333525.1) and alternate (JALMGB000000000, GCA_023333555.1). The primary assembly is haploid representation of the genome ("principal haplotype of diploid"), while the alternate assembly contains the haplotigs in heterozygous regions.

When we looked at the statistics of the assembly though, what stood out immediately is that the primary assembly was larger than the alternate assembly, suggesting that perhaps the two assemblies were switched and identified incorrectly (as confirmed through personal communication with the CCGP staff; Table 2). The assemblies created by CCGP uses a combination of HiFiasm, purge_dups, and SALSA2, with the input data being long-read PacBio HiFi data and Dovetail Genomics' Omni-C reads. Omni-C reads are like Hi-C reads in that it is a derivative of 3C but the difference lies in the enzyme of choice to digest the crosslinked DNA. In Hi-C, either one or multiple restriction enzymes are used and the resulting reads therefore have specific sequences to look for when filtering the resulting paired-end reads. In Omni-C, Dovetail Genomics states that it uses an "endonuclease" that is non-specific in order to digest the crosslinked DNA (Dovetail Genomics Technical Note Omni-C^{TM}). One can hypothesize that this endonuclease is likely or similar to DNase I, which is an endonuclease that can act on single- and double-stranded DNA.

Because the alternate assembly is larger and closer to the expected genome size of around 1.6 Gb, a quick check to see the distribution of scaffold sizes revealed that the majority of contigs were not anchored to just 23 scaffolds representing the 23 chromosomes, and therefore this is not a chromosome-level assembly (Figure 10). Since *A. tigris* is sister to *A. marmoratus*, the expectation is that the chromosomes should be largely one-to-one syntenic with each other. We performed a synteny analysis using Satsuma (Grabherr et al., 2010), taking the largest 50 scaffolds in the alternate *A. tigris* assembly and finding the syntenic regions in the *A. marmoratus* genome that was previously described. This analysis revealed that while 15 of the chromosomes in *A. marmoratus* were assembled and found in *A. tigris,* the rest of the *A. marmoratus* chromosomes were still syntenic to a large handful of *A. tigris* scaffolds (Figure 11).

Table 2. Genome assembly primary data and statistics for *A. tigris stejnegeri* as uploaded to NCBI by CCGP.

PacBio HiFi reads	SRX15576191	
Omni-C Illumina reads	SRX15576192, SRX15576193	
Assembly	Primary	Alternate
NCBI Genome accessions	JALMGA000000000	JALMGB000000000
Assembly accession	GCA_023333525.1	GCA_023333555.1
Number of contigs	113	272
Number of scaffolds	74	228
Scaffold N50 (bp)	93,690,953	98,364,735
Largest scaffold (bp)	272,933,357	285,531,925
Size of final assembly (bp)	1,335,711,111	1,675,921,800

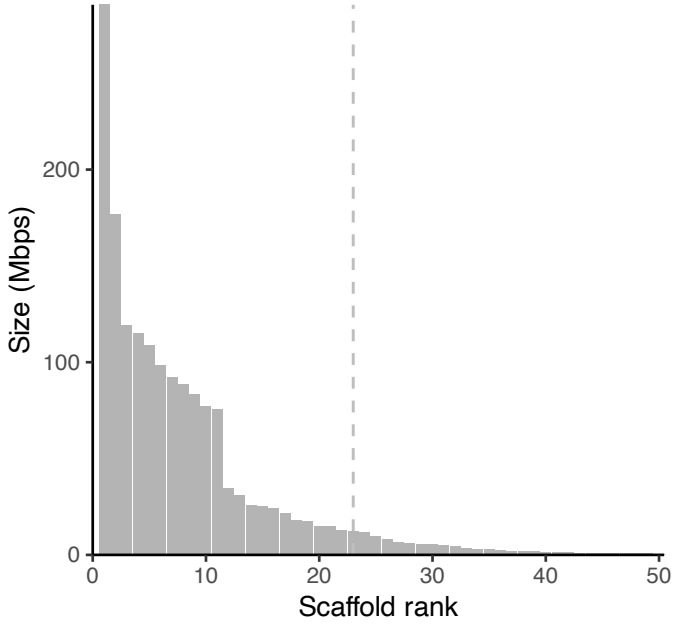

Figure 10. Sizes of the 50 largest scaffolds in the *A. tigris* alternate assembly (GCA_023333555.1). The dashed line is at x = 23, indicating the number of chromosomes in *A. tigris*.

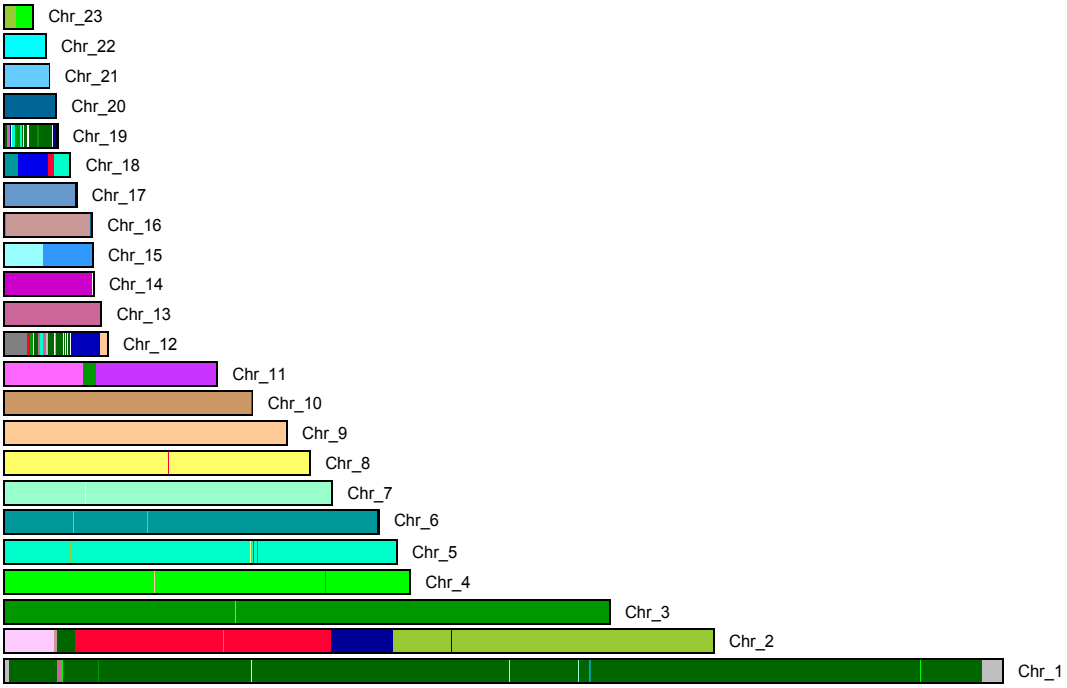

Figure 11. Synteny analysis between the largest 50 scaffolds in *A. tigris* (alternate assembly) and the chromosome-level assembly of *A. marmoratus*. Each large bar represents one of the 23 *A. marmoratus* chromosomes. Color blocks within each bar represent one of the *A. tigris* scaffolds that is syntenic to a particular region in the *A. marmoratus* genome. Colors are repeated because the ChromosomePaint function of Satsuma only allows for 28 unique colors. Fifteen of the chromosomes in *A. marmoratus* were assembled and found fully assembled in *A. tigris* (as indicated by a *A. marmoratus* chromosome only containing one color block*).*

Since the data was publicly available, the PacBio HiFi and Omni-C reads were downloaded to hopefully use the GAB to generate a higher quality genome assembly. The HiFi reads (77.1 Gb) were first passed through an adapter trimmer (HiFiAdapterFilt v. 2.0.0) and assembled using hifiasm (v. 0.16.1-r375) with the following options: `-l 3 -s 0.25`. The resulting genome assembly was then passed through the HiC-Pro pipeline along with the Omni-C reads (120.9 M paired-end reads, 2 x 150), like with the *A. arizonae* genome assembly. Since Omni-C uses a non-specific endonuclease, the HiC-Pro pipeline had to be modified to exclude the step that identifies the presence of certain sequences as a result of the choice of restriction enzymes. Furthermore, because non-specific digestion has a higher change of short-range interactions being included in the paired-end sequencing data, we further filter out short range contacts below 1000 bp to reduce any noise.

The final assembly produced by GAB has over 99% of the genome anchored on 23 chromosomes (Table 3, Figure 12). The assembly size is 1.7 Gb in total and there are a total of 61 scaffolds. To further assess the quality of this genome, BUSCO (Simão et al., 2015) was used to identify almost-universal single-copy genes. A BUSCO single-copy, completeness score of 96.40% was identified when looking at 5,310 tetrapoda genes (an improvement over the 94.80% in the CCPG alternate assembly). As like before, we

compared this assembly to the sister *A. marmoratus* assembly and in this case, all 23 chromosomes of each species were syntenic one-to-one with each other (Figure 13, Figure 14). Interestingly, there are regions in the *A. tigris* assembly not found in the *A. marmoratus* assembly. Due to the location of these regions, they likely represent centromeres and telomeres that were readily assembled due to the use of PacBio HiFi long reads in *A. tigris* assembly, thus showing the advantage of long-read technology in genome assembly. Centromeres and telomeres are highly repetitive regions which are easily collapsed down in assemblies when short-reads are used.

Table 3. Comparison between the CCGP assembly and the assembly using GAB. All assemblies utilized the same primary data.

PacBio HiFi reads	SRX15576191		
Omni-C Illumina reads	SRX15576192, SRX15576193		
Assembly	Primary	Alternate	GAB
NCBI Genome accessions	JALMGA000000000	JALMGB000000000	
Assembly accession	GCA_023333525.1	GCA_023333555.1	
Number of contigs	113	272	79
Number of scaffolds	74	228	61
Scaffold N50 (bp)	93,690,953	98,364,735	115,585,196
Largest scaffold (bp)	272,933,357	285,531,925	298,763,426
Size of final assembly (bp)	1,335,711,111	1,675,921,800	1,714,001,480
Assembly software	HiFiasm, purge_dups, and SALSA2	HiFiasm, purge_dups, and SALSA2	HiFiasm, GAB

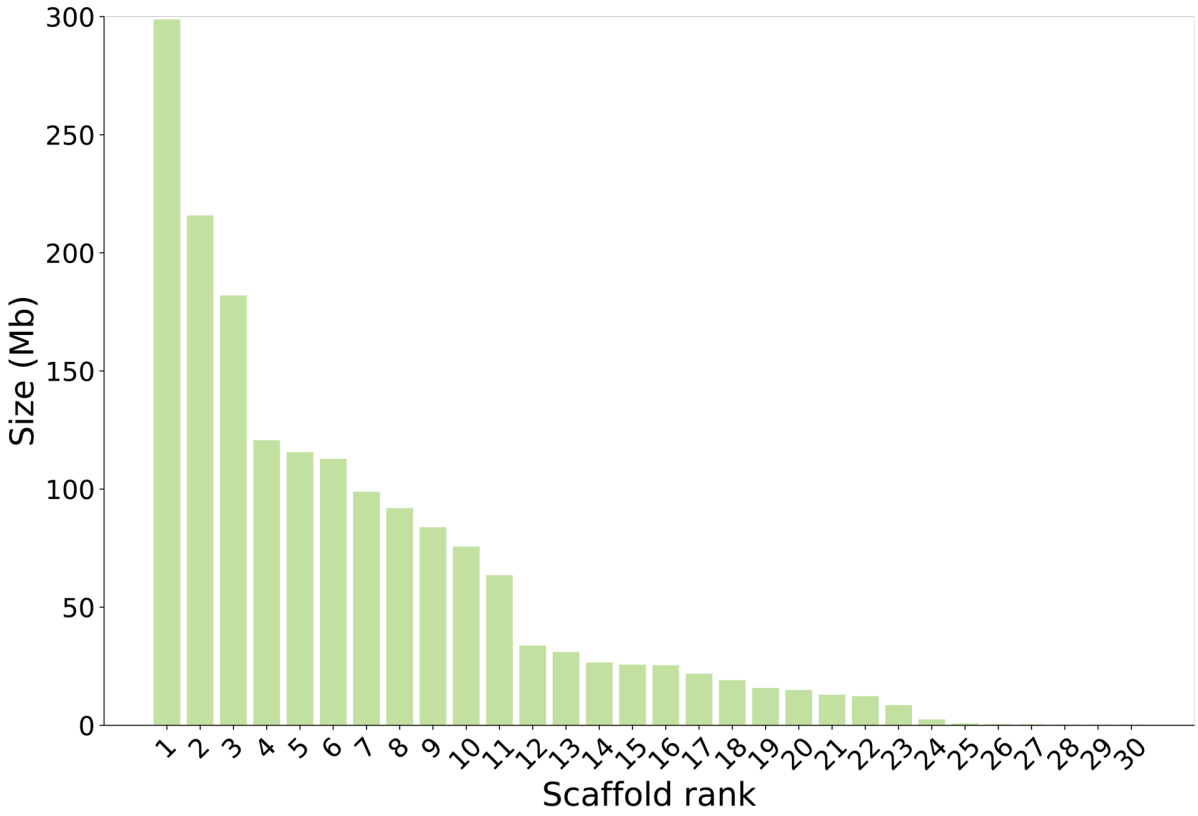

Figure 12. Sizes of the 30 largest scaffolds in the *A. tigris* produced by GAB. Over 99.58% of the sequences are anchored on the first 23 scaffolds.

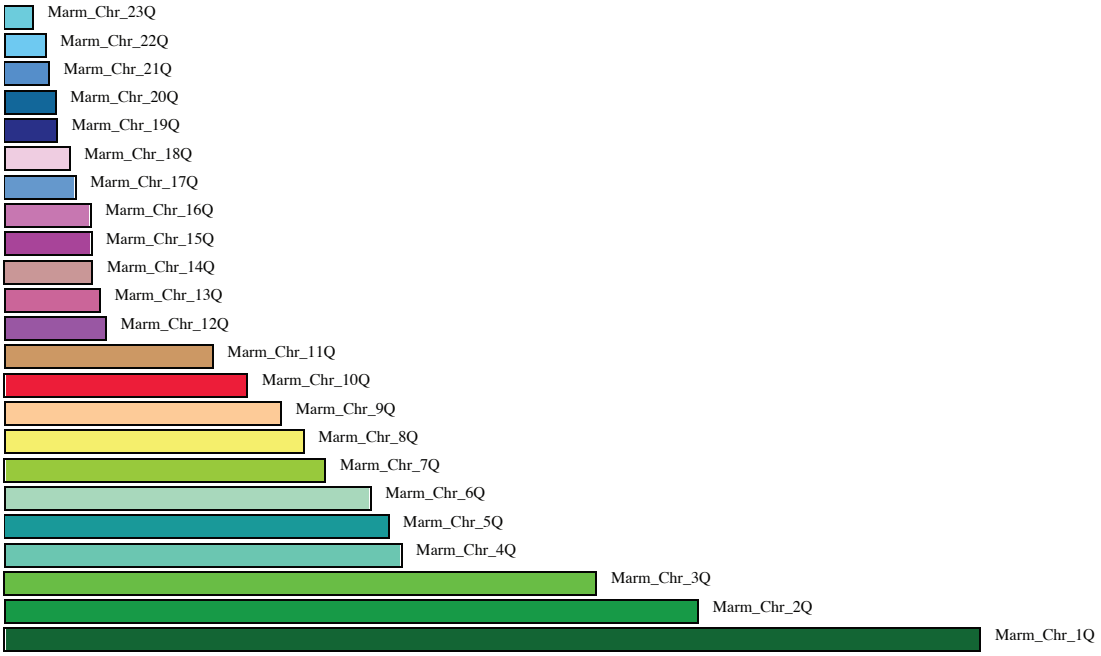

Figure 13. Synteny analysis between the 23 chromosomes in the *A. marmoratus* assembly and the 23 chromosomes in the *A. tigris* assembly. Each large bar represents one of the 23 *A. marmoratus* chromosomes. Color blocks within each bar represent one of the *A. tigris* scaffolds that is syntenic to a particular region in the *A. marmoratus* genome. Each of the 23 *A. tigris* chromosomes is syntenic to the corresponding chromosome in *A. marmoratus*.

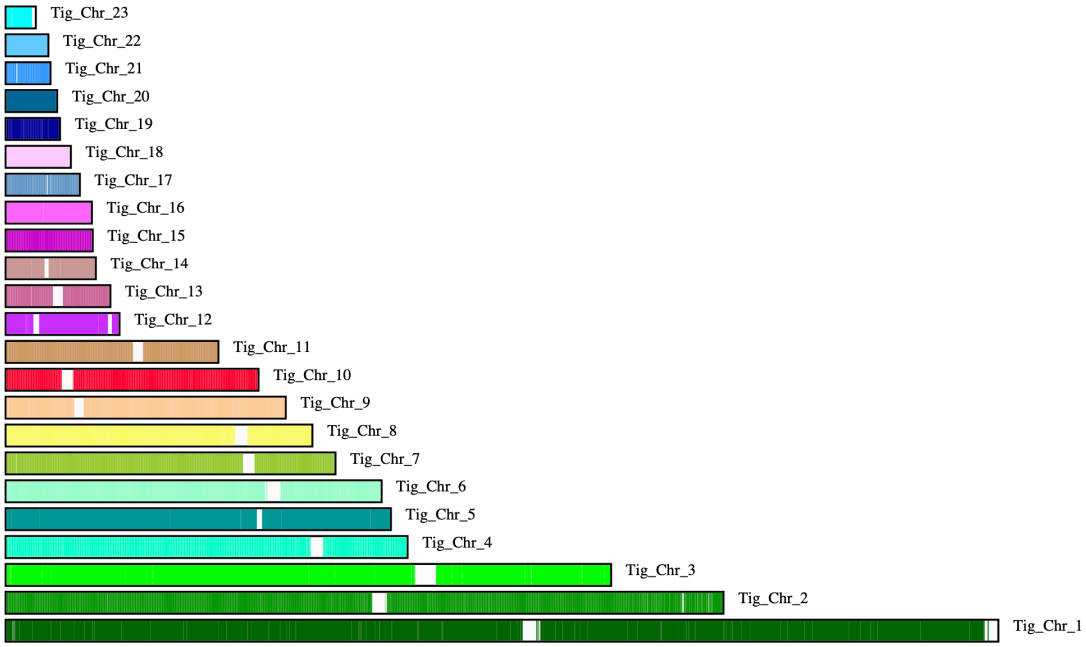

Figure 14. Synteny analysis between the 23 chromosomes in the *A. tigris* assembly and the 23 chromosomes in the *A. marmoratus* assembly. Each large bar represents one of the 23 *A. tigris* chromosomes. Color blocks within each bar represent one of the *A. marmoratus* scaffolds that is syntenic to a particular region in the *A. tigris* genome. Each of the 23 *A. marmoratus* chromosomes is syntenic to the corresponding chromosome in *A. tigris*. Blocks of white represent regions assembled on the *A. tigris* chromosomes that were not found in the *A. marmoratus* chromosomes. These regions represent centromeres and telomeres that were readily assembled due to the use of PacBio HiFi long reads in *A. tigris* assembly.

Aspidoscelis gularis (Common spotted whiptail)

Similar to the *A. marmoratus* and *A. arizonae* genomes, tissue sample from a wild-caught male *A. gularis* (ID 9993; Spur, Texas, USA) was also sent to Dovetail Genomics. The species name *gularis* comes from the Latin *gula* meaning throat, referecing to coloration on the throats of the males. Dovetail Genomics created a Chicago library that was sequenced at SIMR. The data was returned to them to use their HiRise program to provide us a genome assembly. Since this genome assembly was still not at chromosome-level (Table 4), we decided to make Hi-C libraries and sequence in-house. Hi-C libraries were constructed using the Arima HiC+ Kit (A51008-ARI). This kit was chosen over the Arima-HiC High Coverage Kit (A101030-ARI) because it uses a two enzyme digestion rather than four enzymes. As observed from the genome assembly with the Omni-C reads for *A. tigris*, the more digested the genome is, the more sequencing noise is generated that needs to be filtered out.

We used 142.7 mg of flash frozen tail tissue from *A. gularis* ID 9993. The tissue was pulverized first in liquid nitrogen in a mortar placed on a bed of dry ice. The pulverized tissue

was added to 5 mL of 1X PBS. TC Buffer for crosslinking (100 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 50 mM HEPES pH 8.0, 22% Formaldehyde) was added for a final formaldehyde concentration of 2% and incubated at room temperature (RT) for 20 min. Next, 289 µL Stop Solution 1 from the kit was added, mixed, and incubated at RT. The sample was then pelleted by centrifugation at 2000 x g at RT for 15 min. The pellet was resuspended in 1 mL of 1X PBS and pelleted again at 2000 x g at RT for 10 min. The resulting pellet was kept on dry ice. Half of the crosslinked pellet was used in the subsequent steps for the proximity ligation according to the manufacture's protocol. Briefly, lysis buffer, conditioning buffer, a stop solution, and the digestion enzyme mix was used to generate proximity ligated products that was incubated overnight at 68 °C because there were still tissue particulates in the solution due to the tail sample used. The following day, following the manufacture's protocol, DNA Purification beads were used, along with two 80% ethanol washes to purify the proximity ligated DNA. The DNA was then sheared with the Covaris E220 to a target size of 400 bp (Duty Factor: 10%, Peak Incident power (W): 140, Cycles per Burst: 200, Time: 55 sec, Water level: 6). Successful shearing was determined using a Tapestation with the D5000 Screentape. Because the proximity ligation results in biotinylated fragments, a biotin enrichment was performed with the kit's T1 beads. Illumina library preparation was completed with the included reagents in the kit, using primer pair #6 (P7 Index: AGAAGCAA and P5 Index: TCAGCATC) and the resulting library was amplified with 10 cycles during the last PCR step. The final library was sequenced on a high output flow cell on an Illumina NextSeq 500 (2 x 150 bp).

The Hi-C sequencing resulted in 193.34 million paired-end reads that was passed through HiC-Pro. The resulting ICED matrix and the Chicago-level HiRise assembly provided by Dovetail Genomics were used in GAB resulting in a genome assembly with over 99.56% of the sequences assembled on 23 scaffolds (Figure 15).

Table 4. Comparison between the input *A. gularis* assembly and chromosome-level assembly. The input genome assembly was provided by Dovetail Genomics using Chicago-reads and the resulting assembly incorporating Hi-C reads and GAB as the scaffolding program.

	Chicago-level assembly (Dovetail Genomics)	**Chicago + Hi-C (GAB)**
Main genome scaffold total:	2275	2173
Main genome contig total:	82846	82846
Main genome scaffold sequence total:	1503.826 MB	1503.836 MB
Main genome contig sequence total:	1489.262 MB	1489.262 MB
Main genome scaffold N/L50:	11/40.157 MB	6/104.19 MB
Main genome contig N/L50:	11693/36.149 KB	11693/36.149 KB
Main genome scaffold N/L90:	51/6.104 MB	15/24.718 MB
Main genome contig N/L90:	43003/9.13 KB	43003/9.13 KB
Max scaffold length:	111.48 MB	209.339 MB
Max contig length:	425.99 KB	425.99 KB
Number of scaffolds > 50 KB:	126	26
% main genome in scaffolds > 50 KB:	99.60%	99.61%

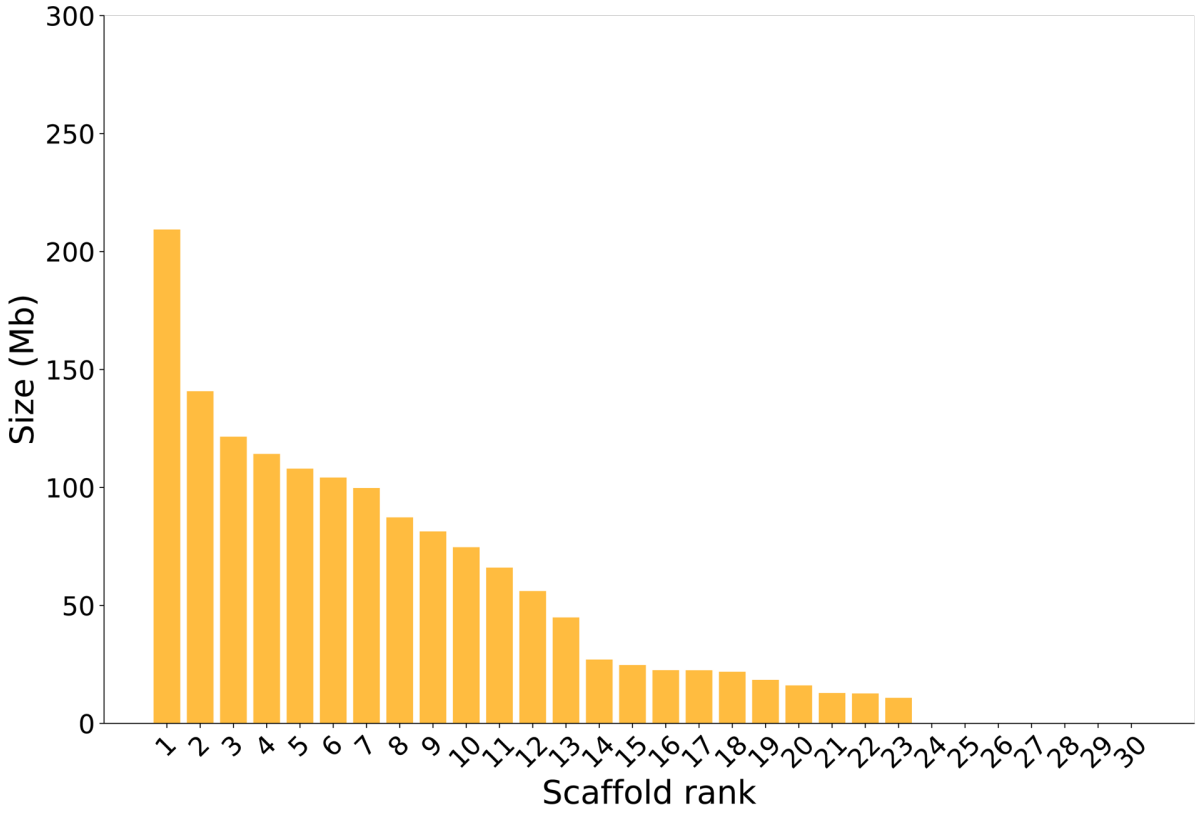

Figure 15. Sizes of the 30 largest scaffolds in the *A. gularis* produced by GAB. Over 99.59% of the sequences are anchored on 23 scaffolds.

Aspidoscelis burti (Canyon spotted whiptail)

The Canyon spotted whiptail, *A. burti* (named after American herpetologist Dr. Charles Earle Burt), is a species of interest for genome assembly because many naturally occurring triploid species are derived from at least one hybridization event with an individual from the *A. burti* species complex: *A. opatae*, *A. sonorae*, *A. uniparens*, and *A. velox*. While speculations that a diploid hybrid with the genome combination from both *A. arizonae* and *A. burti* was the intermediate that led to the establishment of the aforementioned triploids, despite reports in the 1960s (by herpetologist John Wright), no species with this combination had been formally found and described until recently. Barley et al. (2021) described a species with this exact combination found in Sonora, Mexico, along the Rio Bavispe. Named *A. preopatae*, the discovery of this species is important because it was most likely the progenitor for the triploid species.

High-molecular weight gDNA isolated from the liver of a male *A. burti stictogrammus* was sent to Novogene to prepare PacBio HiFi libraries for sequencing. Novogene created two sequencing libraries after shearing the gDNA to fragment sizes centered around 18,000 bp. A total of 29.52 Gb of data was produced over two PacBio flow cells with N50 read lengths of 17,440 and 16,953 bp for the two libraries. Similar to the preparation of the Hi-C library for *A. gularis*, a tail sample from *A. burti* was also used. The resulting Hi-C library was sequenced (along with the *A. gularis* Hi-C library) to generate 230.66 million paired-end reads. The HiFi reads were assembled to a primary genome assembly using HiFiasm as with the *A. tigris* assembly. As compared to the amount of *A. tigris* input HiFi reads (77.1 Gb), we had less than half the data. This resulted in the assembly produced from HiFiasm to have not be as contiguous. A round of purge haplotigs (Roach et al., 2018) was performed to reduce duplications arising from low coverage HiFi reads. After using the Hi-C data with GAB and manual curation, we were able to produce a chromosome-level assembly for *A. burti* with 98.83% of the sequences being anchored on the expected 23 chromosomes. Even at lower sequencing coverage than what was used in *A. tigris*, we were still able to produce this contiguous assembly because of the use of long-read sequencing (Table 5, Figure 16). Since *A. burti* is sister to *A. arizonae*, we performed a synteny analysis between the chromosomes and saw one-to-one synteny between the 23 chromosome pairs (Figure 17).

Table 5. Comparison of the statistics for the *A. burti* genome assembly at each step of the assembly process.

Steps in the assembly process	Hifiasm	Purge haplotigs	GAB	Final curated genome
Total number of scaffolds	1016	545	102	109
Total genome size (Gb)	1.732	1.688	1.688	1.688
N50 scaffold and value (Mb)	52 / 7.687	49 / 7.858	6 / 117.559	6 / 117.559
N90 scaffold and value (Mb)	277 / 1.186	246 / 1.375	15 / 26.337	15 / 26.057
Largest scaffold / Chromosome 1 (Mb)	42.633	42.633	239.589	239.589
Cumulative percentage of genome for top 23 scaffolds	32.9	33.76	99.13	98.83

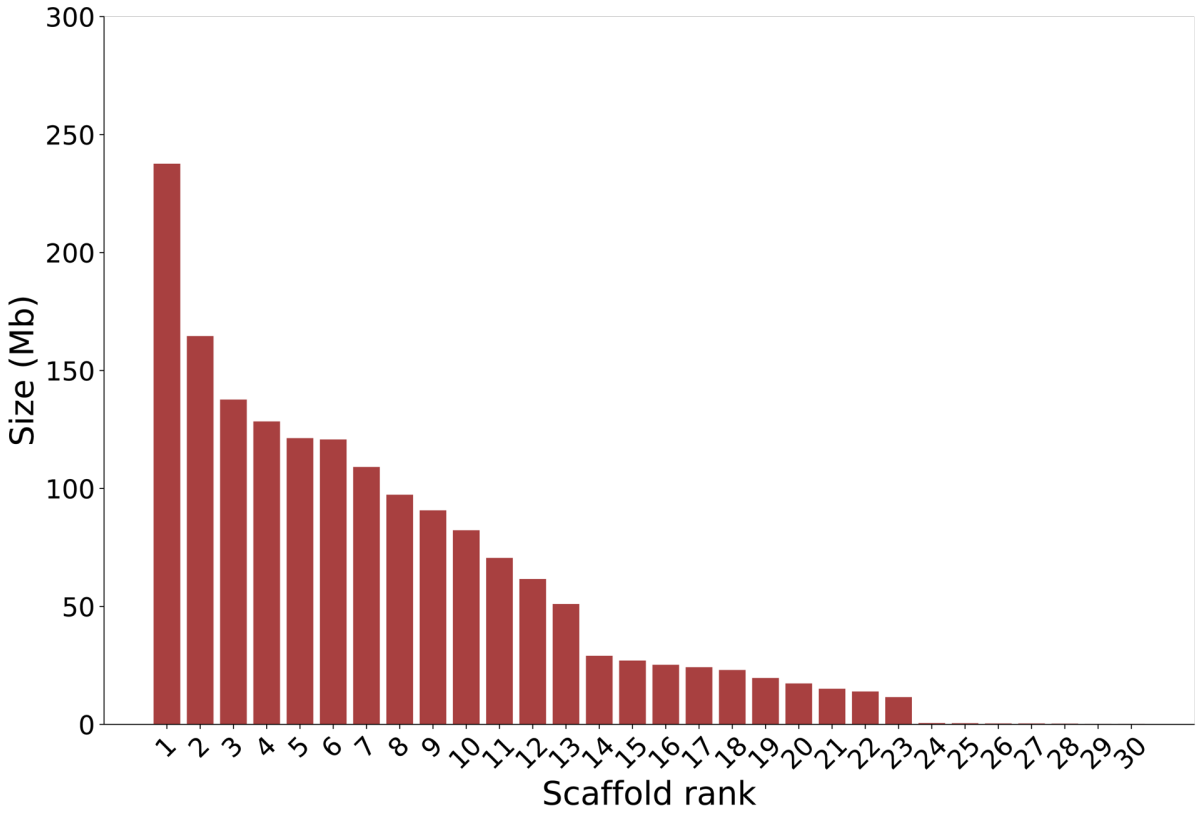

Figure 16. Sizes of the 30 largest scaffolds in the *A. burti* produced by GAB. Over 98.83% of the sequences are anchored on 23 scaffolds.

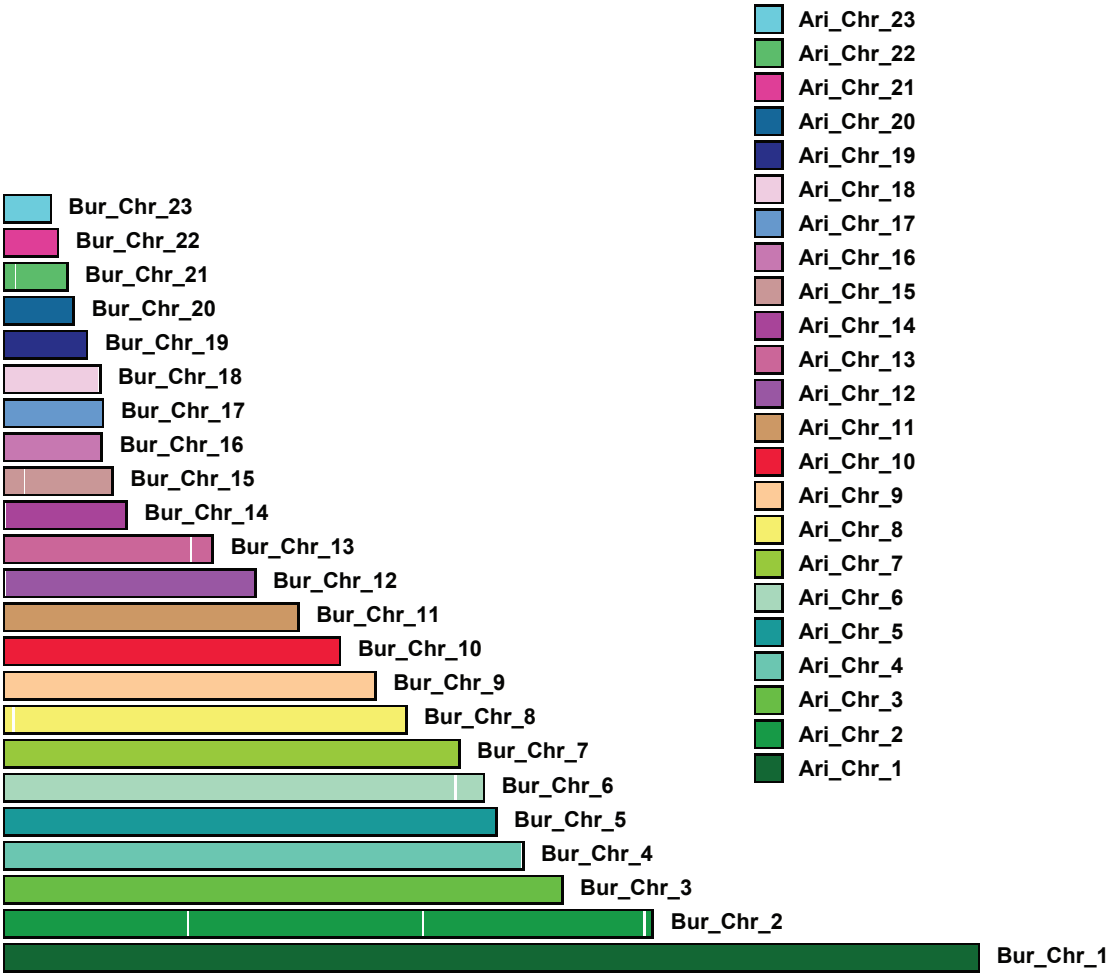

Figure 17. Synteny analysis between the 23 chromosomes in the *A. burti* assembly and the 23 chromosomes in the *A. arizonae* assembly. Each large bar represents one of the 23 *A. burti* chromosomes. Color blocks within each bar represent one of the *A. arizonae* scaffolds that is syntenic to a particular region in the *A. burti* genome. Each of the 23 *A. burti* chromosomes is syntenic to the corresponding chromosome in *A. arizonae*.

Mitochondria assemblies

Mitochondrial genomes (mitogenome) are the most abundantly sequenced genomes. Encoding genes that produce the protein subunits for the electron transport chain, ribosomal RNAs (rRNAs), and transfer RNAs (tRNAs), the mitochondrial genome is largely conserved in metazoans, being around 16 Kb in size . However, in plants and fungi, the size of the mitogenomes can be as big as 750 Kb in angiosperms (Fonseca et al., 2021; Gualberto et al., 2014; Kubo & Newton, 2008). Inherited exclusively from the mother (with the exception of male inheritance observed in Plymouth Rock chickens), the mitochondria is not only appreciated for its production of ATP and signaling apoptosis, its genome can also be used in animals as a form of barcoding (Alexander et al., 2015; Gray, 2012). Mitochondrial genes

tend to have a higher rate of mutations as compared to the nuclear genome and therefore genes such as the cytohrome *c* oxidase subunit 1 (COX1, CO1, MT-CO1) have used to discriminate between species (Guo et al., 2022; Mueller, 2006; Paul D.N. Hubert et al., 2003). In addition to the assembly of the much larger nuclear genome, the assembly of the mitogenome is important as well.

The assembly of the mitogenomes for *A. marmoratus*, *A. arizonae*, *A. gularis*, and *A. burti* was done with NOVOPlasty (v. 4.3.3) (Dierckxsens et al., 2017). The *A. tigris* mitogenome was already assembled by the CCGP with MitoHiFi (GenBank: CM041363.1). In brief, the Illumina paired-end short-reads used for the initial assemblies of the four species were first trimmed for adapters and low quality bases with trimmomatic (Bolger et al., 2014) and then passed through NOVOPlasty with a COX1 gene sequence (*A. inornatus,* NCBI: NC_067819, ASH120 isolate; *A.* gularis, NCBI: NC_067815, ASH022 isolate) as the seed sequence and k-mer setting set to 65 or 71. NOVOPlasty takes reads that map to the seed sequence and continues to prime off of the match. With enough coverage, the algorithm is able to circularize the genome. Any ambiguities are converted to Ns manually afterwards. Annotation of the genes in these mitogenomes were annotated using Geneious (*Geneious Prime 2023.2.1*). The average size of the mitogenomes for the five species is 17279.8 bp (± 313.5 s.d.; Table 6).

Table 6. Mito-genome assembly statistics.

Species	ID	MT genome size (bp)	Number of WGS reads used in NOVOPlasty	MT coverage
A. marmoratus	8450	17,237	157,758,540	3,036
A. arizonae	17121	17,261	495,690,594	18,792
A. tigris	HBS 135688	17,692	N/A	N/A
A. gularis	9993	17,387	470,734,448	15,372
A. burti	12836	16,822	471,193,250	11,794

Summarizing and comparing the genomes

Taken together, I have shown the assemblies of five high quality *Aspidoscelis* genomes. The scaffold N50 is on chromosome 5 for the *A. marmoratus* and *A. tigris* genomes, while chromosome 6 represents the scaffold N50 for the *A. arizonae*, *A. gularis*, and *A. burti* genomes. The range of the total assembled genome sizes is from 1.50 to 1.71 Gb. Given that the *A. tigris* and *A. burti* genomes were assembled with long-read PacBio HiFi data, it is not surprising that the assemblies are a bit bigger because the chances of

repetitive parts of the genome being collapsed is less likely. This is also mirrored in the total number of scaffolds for each assembly. Due to the short Illumina reads being used as the first round of assembly, there are much more small scaffolds still present, even though over 99% of the genome can be found on 23 scaffolds (Table 7). All five genome assemblies also have a high BUSCO score as well, which looks for the presence of single-copy orthologs found within vertebrates (Figure 18).

Table 7. Summary statistics for the five assembled *Aspidoscelis* genomes.

	marmoratus	*tigris*	*arizonae*	*gularis*	*burti*
Total number of scaffolds	3519	61	3150	2173	109
Total genome size (Gb)	1.64	1.71	1.53	1.50	1.69
N50 scaffold and value (Mb)	5 / 113.087	5 / 115.585	6 / 106.147	6 / 104.19	6 / 117.559
N90 scaffold and value (Mb)	14 / 25.213	14 / 26.587	15 / 24.636	15 / 24.718	15 / 26.057
Largest scaffold / Chromosome 1 (Mb)	288.666	298.763	211.803	209.339	239.589
Cumulative percentage of genome for top 23 scaffolds	99.59	99.58	99.4	99.59	98.83

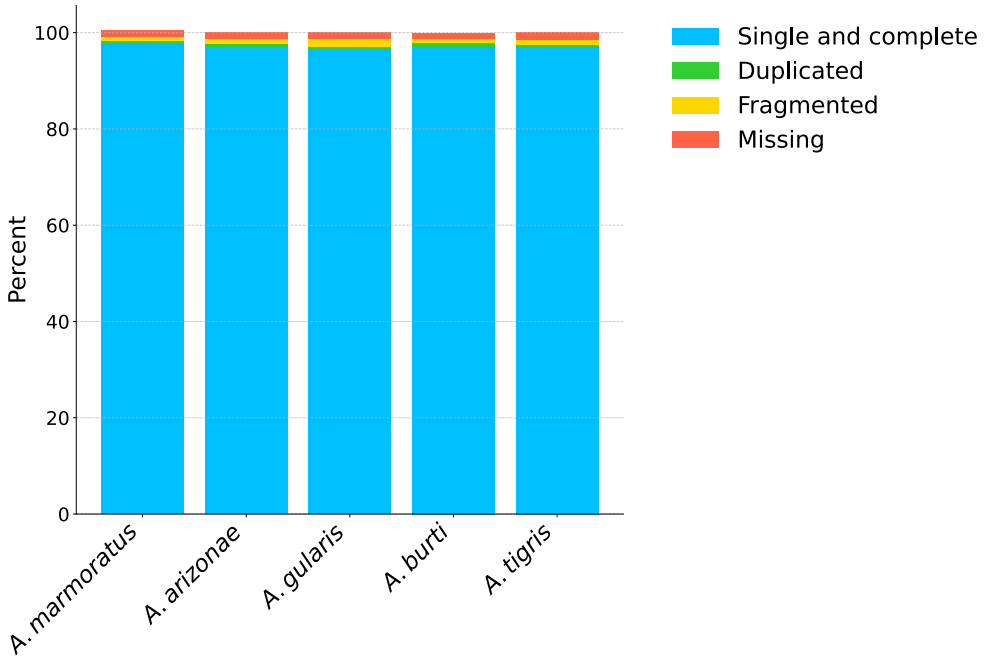

Figure 18. BUSCO values for the five assembled Aspidoscelis genomes. The vertebrata_odb10 database was used as a query (n = 3354).

Additional comparisons between these chromosome-level assemblies can be made by looking synteny between all five gnomes at the same time (Figure 19). In this view, chromosome fusions and fissions certainly distinguish the *A. marmoratus* and *A. tigris* group from the rest of the genomes. Structural variation between the species also seems to be common. For example, despite being closely related, *A. arizonae* and *A. burti* appear to

have inversions between syntenic chromosomes. Whether or not these inversions are seen across whole populations and if they impact the successful outcomes of hybridization between related species still needs to be investigated.

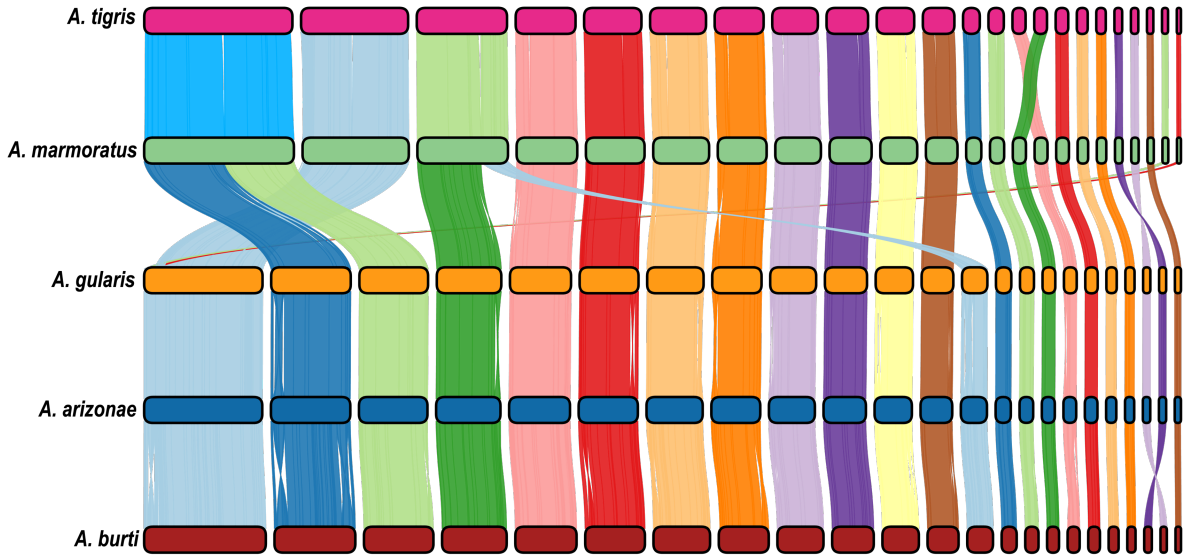

Figure 19. Full genome synteny comparisons between five *Aspidoscelis* species. All five species each have 23 chromosomes. Like other reptiles, whiptail genomes are characterized by macro- and micro-chromosomes. Ribbons represent blocks of synteny between species. Many chromosome pairs are 1:1 syntenic however evidence of fusion and fission of chromosomes are apparent. For example, *A. tigris* and *A. marmoratus* chromosome 1 are fusion between chromosomes 2 and 3 of *A. gularis, A. arizonae*, and *A. burti*. Large scale inversions between syntenic chromosomes can also be observed.

Within the genus *Aspidoscelis*, historic work using mitochondrial genes, morphological data, and allozyme analysis revealed the phylogenetic relationships between the many species within the genus and within the sister genus *Cnemidophorus* (Reeder et al., 2002). More recently, the evolutionary network history of whiptail lizards was constructed, showing similar patterns and relationships but incorporating information about how hybridization between whiptail species can lead to two outcomes or lineages: introgression or parthenogenesis (Barley et al., 2022). Using the genome assemblies, we wanted to reconstruct the relationship between the species using genomic data. Similar to the methods in Chapter 3, shared single-copy BUSCO genes (amino acids) were extracted from all five genomes and the *Anolis carolinensis* genome. A multiple sequence alignment of the supermatrix containing 1,799,948 residues was performed with MAFFT (v. 7.520). The resulting alignment was then passed through to RAxML-NG (v. 1.2.0) with 10 randomized parsimony starting trees, fixed empirical substitution matrix (LG), 8 discrete GAMMA categories, *A. carolinensis* as the outgroup, and 100 bootstrap replicates (Kozlov et al., 2019). The resulting tree showed full 100% support for all branches and the relationship between the five species is similar to the former mentioned studies, with *A. marmoratus* and *A. tigris* forming their own clade, *A. arizonae*, *A. gularis*, and *A. burti* in another clade (Figure

20A). MAFFT alignment (v. 7.490) of the mito-genomes followed by tree generation with FastTree (v. 2.1.11) also revealed the same relationships between the species (Figure 20B).

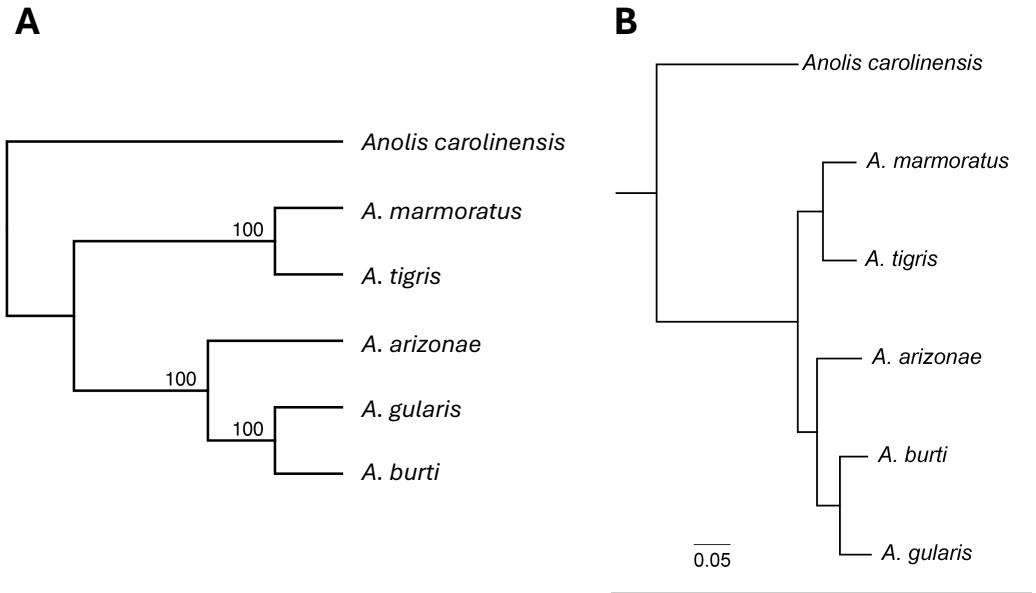

Figure 20. Phylogenetic tree of five *Aspidoscelis* species.The tree was derived from multiple sequence alignment of (A) shared BUSCOs and (B) the mitogenomes. The *Anolis carolinensis* mitogenome was used as an outgroup in both cases. Values next to the nodes represent bootstrap support in (A). Scale bar represents substitutions per site in (B).

The multiple sequence alignment of the BUSCO peptide sequences and the mito-genomes not only provide information about the phylogenetic relationships between the species, but we can also quantify it by looking at the pairwise identities between sequences. For the five genomes, there is a 94.05% average pairwise identity, while the mito-genomes were on average 84.39% different from each other (Figure 21). This 10% difference between the genes located within the nucleus and the ones located in the mitochondria is consistent with other species where higher mutation rates are consistently observed in the mito-genomes (Boore, 1999; Montooth & Rand, 2008).

A

	Anolis carolinensis	*A. marmoratus*	*A. tigris*	*A. arizonae*	*A. gularis*	*A. burti*
Anolis carolinensis						
A. marmoratus	71.42					
A. tigris	71.44	96.75				
A. arizonae	71.09	93.34	93.25			
A. gularis	71.06	93.25	93.47	93.76		
A. burti	71.14	93.66	93.82	94.04	95.17	

B

	Anolis carolinensis	*A. marmoratus*	*A. tigris*	*A. arizonae*	*A. gularis*	*A. burti*
Anolis carolinensis						
A. marmoratus	52.28					
A. tigris	51.87	88.6				
A. arizonae	51.84	82.76	81.65			
A. gularis	51.12	82.9	83.12	84.53		
A. burti	49.94	83.63	80.53	87.39	88.78	

Figure 21. Pairwise percent identities between the *Anolis carolinensis* and the five *Aspidoscelis* genomes.Comparisons were made between (A) shared BUSCOs and (B) mito-genomes. Percent identity is defined as number of residues or bases, respectively, that are identical over the total alignment length. An alignment length of 1,799,948 residues derived from BUSCO genes (2747 / 3354) were compared. For the mito-genomes, the total alignment length was 23,173 bp. The average pairwise percent identity for BUSCO proteins is 94.05% and the average pairwise percent identity for the mito-genomes is 84.39%.

5. *Aspidoscelis neomexicanus* (New Mexico whiptail)

The New Mexican whiptail *A. neomexicanus* has the origin as a hybrid between *A. marmoratus* (♀) and *A. arizonae* (♂) and therefore harbor one set of 23 chromosomes from each species, resulting in 46 unique chromosomes. The formation of *A. neomexicanus* by hybridization is estimated to have occurred at least 200 thousand years ago (Barley et al., 2022). We can confirm the maternal progenitor is *A. marmoratus* by comparing our own assembled mito-genomes with publicly available data. The resulting phylogeny shows that the mito-genome within *A. neomexicanus* groups with mito-genomes from different *A. marmoratus* individuals (Figure 22).

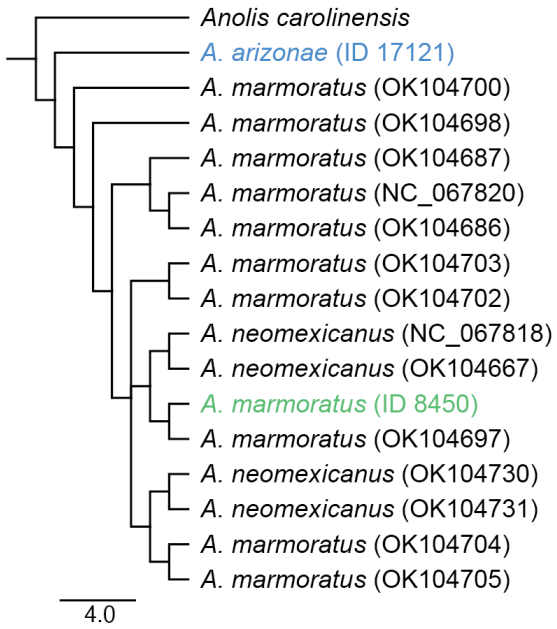

Figure 22. Phylogenetic tree showing the relationship *A. neomexicanus* and *A. marmoratus* mito-genomes.Black branches represent of publicly available assemblies while our own assembled mito-genomes for *A. arizonae* are in blue and for *A. marmoratus* in green. The *A. neomexicanus* mito-genomes cluster with the *A. marmoratus* and not *A. arizonae*. Scale bar represents substitutions per site.

By generating chromosome-level assemblies for *A. arizonae* and *A. marmoratus*, not only can we study the bisexual species separately, but the combination of the two representative-haploid genome assemblies provides a representation of the *A. neomexicanus* genome. Analysis using Satsuma2 reveal even though the two species each have 23 chromosomes, 1:1 synteny is observed only between 18 of the chromosomes in *A. marmoratus* and *A. arizonae*. Evidence of fusion and fission of chromosomes from the most

recent common ancestor is apparent. For example, *A. marmoratus* chromosome 3 is syntenic to *A. arizonae* chromosome 4 and 13 (Figure 23).

Chromosomal fusion and fission are important evolutionary processes that shape genomes and are not unique to just whiptail lizards (de Vos et al., 2020; Leaché et al., 2016; Pennell et al., 2015). The consequences of these structural changes can influence how genes are organized, their expression patterns, and linkage in inheritance. In reptiles, chromosomal fusions and fissions are documented in many taxa with squamates having the largest variability in karyotypes (Olmo, 2005). Chromosomal changes don't just affect gene expression but in some cases, they can lead to reproductive isolation and speciation as documented in other taxa (Mackintosh et al., 2023; Potter et al., 2017). Even though the rate of fusion and fission vary between groups, they nevertheless are important in genome evolution and contribute to the diversity seen across vertebrates. The fissions and fusions observed through microscopy and confirmed with genomics within whiptail lizards add another layer to the hybridization success between different whiptail species (Barley et al., 2022).

Within this *A. neomexicanus* genome, the divergence between syntenic chromosomes can be measured by aligning the chromosomes pairs to each other and calculating the gap-compressed percent identity. The gap-compressed percent identity puts less weight on differences between sequences and compresses consecutive gaps in alignment as one difference only. As the both assemblies are not without errors, gaps, and various other problems associated with genome assembly, the gap-compressed percent identity is a much better estimation of the similarity between the genomes. By using wfmash, an accelerated alignment algorithm, we calculated the average gap-compressed percent identity between syntenic *A. arizonae* and *A. marmoratus* to be 94.67% (s.d. = 0.73%; Table 8). This is remarkably similar to the average percent identity between shared BUSCO genes of the two species: 93.34% (s.d. = 3.91). The percent identity was calculated by performing a BUSCO (v5.4.2) search of the genome against the vertebrate gene database. There were 2907 shared genes between the two genomes and pairwise alignments of the orthologs were performed to derive the average BUSCO gene percent identity.

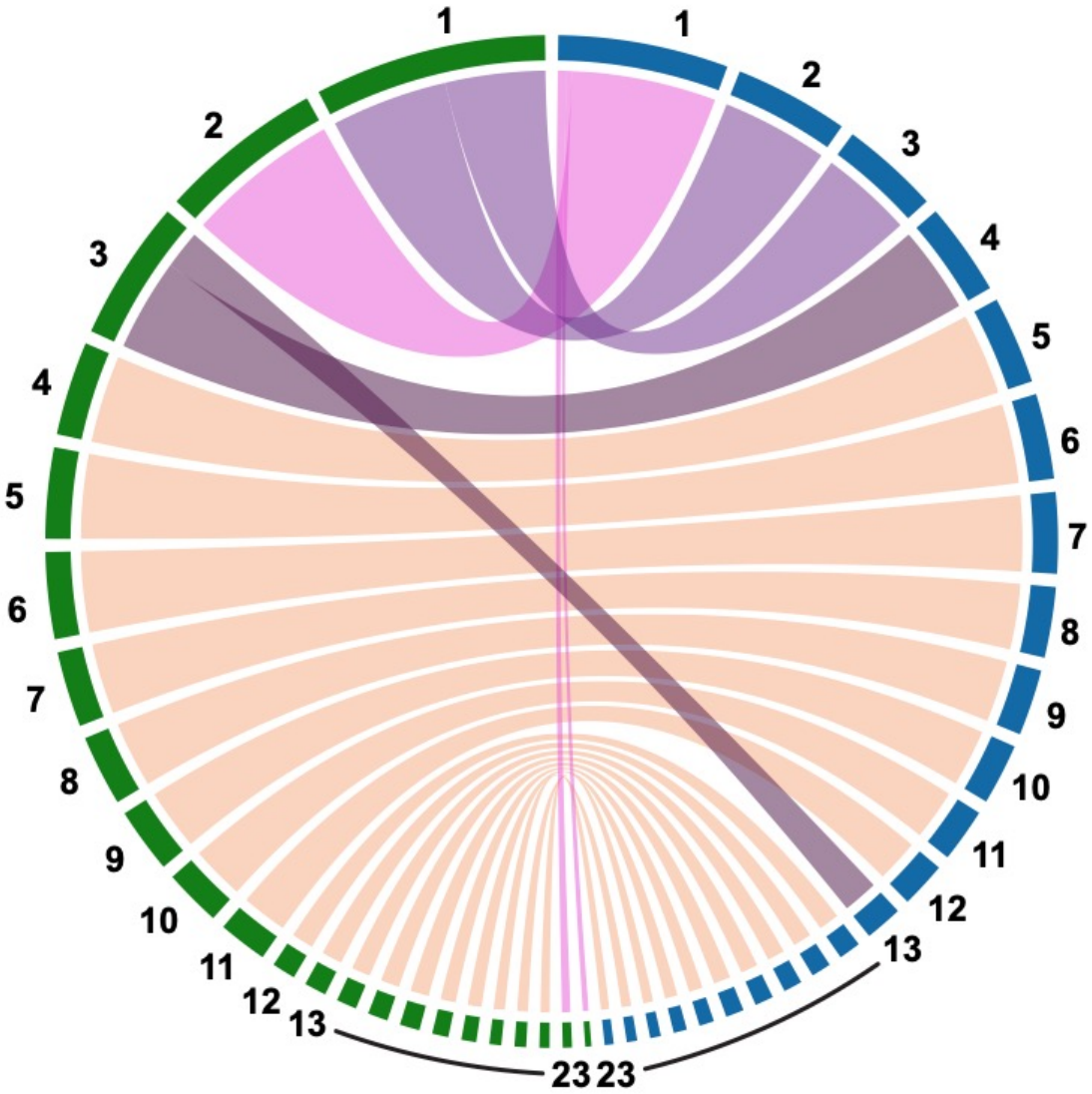

Figure 23. Circos plot showing the synteny between *A. marmoratus* and *A. arizonae* chromosomes.*A. marmoratus* chromosomes are green (left hemisphere) and *A. arizonae* chromosomes are blue (right hemisphere). Tracks between the chromosomes are colored to indicate 1:1 synteny (orange) between chromosomes or fusions of chromosomes (shades of purple and pink). Each species has 23 chromosomes and the concatenation of the two genomes is a representation of the *A. neomexicanus* genome.

Table 8. Gap-compressed percent identity between syntenic *A. marmoratus* and *A. arizonae* chromosomes. Percent identities were calculated by alignment using wfmash. The average percent identity between syntenic chromosomes is 94.67%.

A. marmoratus **chromosomes**	***A. arizonae*** **chromosomes**	**Gap-compressed Percent Identity**
1	3	94.96
1	2	95.02
2	1	93.12
3	13	95.26
3	4	95.52
4	5	95.11
5	6	95.66
6	7	95.10
7	8	95.01
8	9	95.21
9	10	95.26
10	11	95.33
11	12	95.15
12	14	95.08
13	15	94.78
14	16	94.85
15	17	93.47
16	18	94.94
17	19	94.61
18	20	94.47
19	22	93.83
20	21	94.17
21	23	93.83
22	1	93.91
23	1	93.12

Microsatellite differences between clonal individuals

Hybridization leading to obligate parthenogenesis are thought to be successful because of the high level of heterozygosity in the offspring that is maintained from generation to generation (A. A. Lutes et al., 2010). However, it is still important to note that individuals might differ due to accumulated mutations that are maintained into the subsequent generations. One approach to understanding genetic diversity in seemingly clonal individuals is to genotype individuals through microsatellite (MS) analysis. MS analysis involves looking at repeating sequences of DNA that are generally two to six base pairs long. These sequences are highly polymorphic because of slippage of the DNA polymerase enzyme, lead to an increase or decrease in the number of repeat units. Overtime, variability in the size of the locus makes them useful for genetic studies to distinguish individuals from each other (Ellegren, 2004). By comparing the specific markers in offspring to potential parents, paternity or maternity can also be inferred.

Analysis looking at 12 loci from 377 *A. neomexicanus* individuals genotyped at the SIMR reveal that not all individuals share the exact same MS at all loci. For five loci, all individuals were heterozygous each for the same two alleles. For two loci, all individuals were homozygous for the same allele. And for the remaining five loci, there were various allele combinations (Table 9). Taken together, there are 14 unique combinations or groups (of the last five alleles) represented in the 377 animals. The genotypes can be converted to numerical values (R package adegenet) and through a linear combination we can visualize the groups in a PCA (Figure 24).

Table 9. Microsatellites used to distinguish *A. neomexicanus* individuals.

Microsatellite	Type of repeat	Number of allele combinations	Alleles
MS1	AAAG	1	AB
MS6	TG/TC interspersed	1	AB
MS7	GGAA	1	AB
MS8	CA	1	AB
MS18	AGGT, ATAG	1	AB
MS17	AAGA, CAA	1	AA
MS12	AGAT, GATT, AGGT	1	AA
MS10	CA	3	AC, BD, BC
MS16	CA	2	AA, AB
MS19	AGAA	2	AB, BB
MS15	TATC, TGTC	3	AA, AB, BB
MS14	CAGG, AGAT, GGCA, ACAG	5	AA, BB, CC, DD, BC

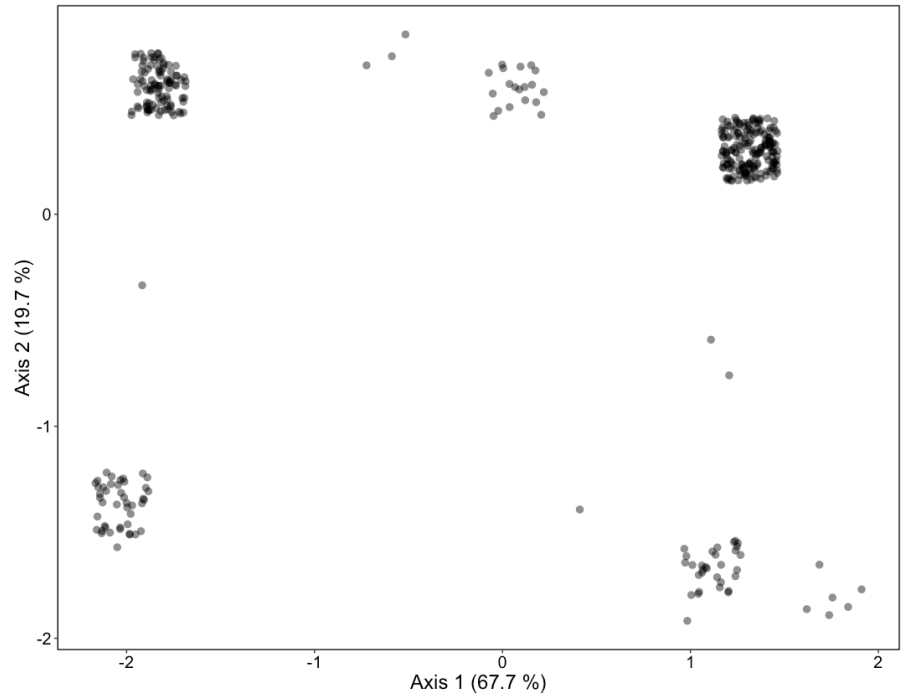

Figure 24. PCA clustering *A. neomexicanus* individuals. Each point represents an individual (n = 377) with the same 12 microsatellites (MS) analyzed. Genotypes of each MS was converted to numerical values to generate the similarity matrix. Clear groups with the exact same set of unique MS are observed in the plot. Points are jittered to reflect the density of each cluster.

The fact that we can cluster these seemingly clonal individuals into distinct groups and detect variation through microsatellite analysis provides an important first hint of potential genetic differences between individuals. In this case, it is important to consider the origin of the diversity we are observing. Since only six founding *A. neomexicanus* individuals were originally taken from the wild (in 2003 and 2004, Socorro County) and maintained and reproduced in captivity, some of the genetic variation we see could be a reflection of the diversity already present at the time of collection. However, it's equally possible that some of this variation has been generated within the lab over time. This raises interesting questions about the balance between pre-existing wild diversity, lab variation, and the natural mutation rate within the species.

Loss of heterozygosity

While microsatellites are useful for identifying these initial patterns, they offer only a partial view of the entire genome. Sequencing of the entire genome is the first step in looking at potential diversity genome-wide.

The choice of a reference genome though is not trivial. Assembling the genomes of a hybrid species like *A. neomexicanus* is inherently more difficult than assembling those of its progenitor *A. arizonae* and *A. marmoratus*. This is due to the presence of two distinct sets of genetic material that are still similar enough that standard assembly tools often struggle to distinguish between the homologous sequences. Mis-assemblies, large gaps, collapsed regions, and tandem duplicated regions are some of the few errors that can occur. In plants, hybrid species often arise frequently and naturally, and the resulting genomes can be quite large. For example, wheat (*Triticum aestivum*) species have genomes composed of three different genomes and despite its agricultural significance, only within the last six years has there been chromosome-level assemblies published (Aury et al., 2022; Walkowiak et al., 2020). To achieve these polyploid genome assemblies, a variety of sequencing technologies need to be utilized, in addition to improved computation algorithms, not to mention the economic feasibility.

For *A. neomexicanus*, like the other whiptail genome assemblies, we sought to create a genome assembly utilizing short read, Chicago, and Hi-C data. All of these data were combined together by Dovetail Genomics using their HiRise assembly (Table 10). While the expected genome size of around 3.2 Gb was produced and many of the scaffolds were big enough to be labeled as chromosome sized, the issue arose when we tried to do a BUSCO analysis. The expectation from a good genome assembly of a diploid hybrid is every gene should actually yield two copies and fall under the “Complete and duplicated” category. In the analysis of the HiRise *A. neomexicanus* genome, we find that only 86.3% of the vertebrate genes were found in duplicate and 11.1% of the genes were single copy. This is highly indicative that many homologous regions between the progenitor species were collapses in the genome assembly. Furthermore, of the duplicated genes, 10.95% of these genes were found on the same scaffold, next to each other. All of this highlights the need for long-read and other types of sequencing to be generated and used in order to assemble the 46 chromosomes harbored within *A. neomexicanus*.

Table 10. Summary statistics of the *A. neomexicanus* genome assembly provided by Dovetail Genomics.The comparison is made to the genome assemblies of the two progenitor species.

	A. arizonae	***A. marmoratus***	***A. neomexicanus***
Sequencing	Paired-end, mate-pair, Chicago, HiC	Paired-end, mate-pair, Chicago, HiC	Paired-end, Chicago, HiC
Assembly algorithm	Genome Assembly Booster	HiRise	HiRise
Scaffold total	3150	3519	6977
Contig total	88819	170667	173004
Size (Scaffold)	1527.901 MB	1639.567 MB	3165.703 MB
Scaffold # / N50	6 / 106.147 MB	5 / 113.087 MB	11 / 107.969 MB
Scaffold # / N90	15 / 24.636 MB	14 / 25.213 MB	17 / 78.745 MB
Number of scaffolds > 50 Kb	31	28	58
% main genome in scaffolds > 50 Kb	99.48	99.61	99.59

One potential workaround for assembling hybrid genomes is to use the genomes of the progenitor species as proxies. By concatenating the genomes, we can use this composite genome as a reference for further analysis. This approach is simpler because it circumvents the need to disentangle the two genomes at every locus within the *A. neomexicanus*. However, the caveat to note is since the hybridization that led to *A. neomexicanus* was over 200 thousand years ago, the original *A. arizonae* and *A. marmoratus* that copulated therefore would be different from the modern-day individuals used in sequencing.

Whole-genome sequencing using Illumina short-reads of several *A. neomexicanus* individuals was performed. After alignment to the concatenated *A. arizonae* and *A. marmoratus* genomes, we looked at sequencing coverage across both genomes. While the majority of chromosomes showed the expected coverage across the entirety of their lengths, what was surprising was there were regions where there was essentially no coverage (i.e. no reads mapping) and regions where there was elevated coverage, twice the expected. The most notable example that was observed in all animals sequenced so far is on the 3'-end of *A. arizonae* Chromosome 17, where there is a region of elevated coverage followed by a region of no coverage. In contrast to the 3' end of *A. marmoratus* Chromosome 15, where there is a region where no coverage was observed followed by elevated coverage (Figure 25). As seen in Figure 25, these two chromosomes are syntenic to each other and therefore one hypothesis is that there was recombination between the two chromosomes that led to missing and duplicated regions. Alternatively, other genomics rearrangements could also explain the results seen with the whole-genome sequencing.

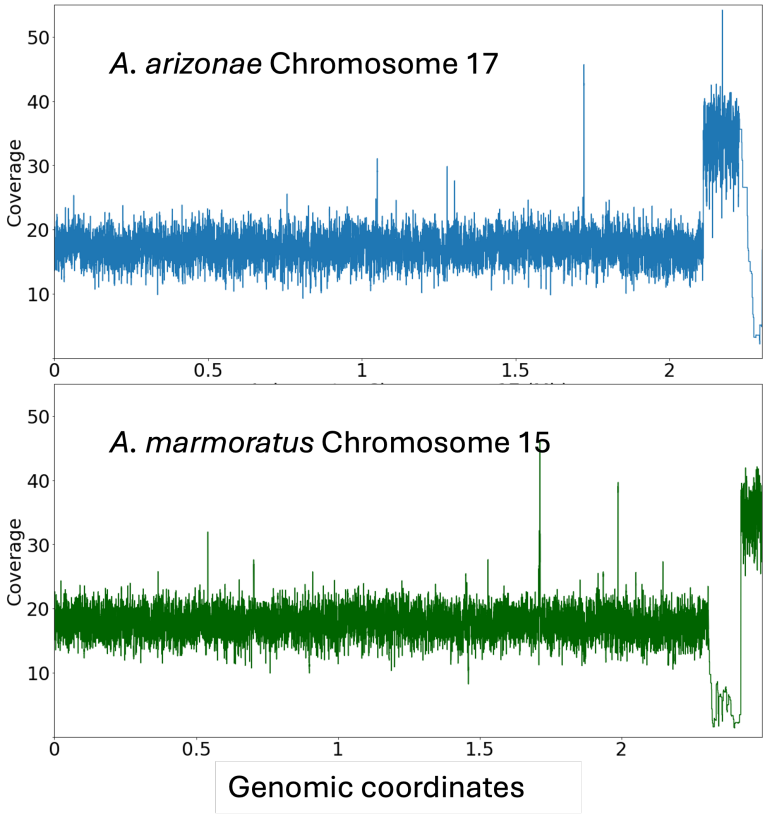

Figure 25. Whole-genome sequencing of *A. neomexicanus* reveals missing and duplicated on *A. arizonae* Chromosome 17 and *A. marmoratus* Chromosome 15.

From our understanding that identical chromosomes pair following a duplication of the nuclear content in oocytes, the expectation is that entire chromosomes are inherited in the offspring in its entirety (Figure 26A), maintaining heterozygosity at all loci. To test that there is indeed recombination between these two syntenic chromosomes, primers were designed to be specific to each chromosome that surround the junctions (Table 11). Junctions were defined by the regions just upstream or downstream of the elevated or missing regions. Different combinations of the primers were used to amplify the proposed fusions between the chromosomes (Figure 26B), with various conditions and polymerases tested (Table 12).

The successful amplification showing fusions between the two syntenic chromosomes (Figure 26C) is significant because for the first time in a unisexual whiptail, we show that heterozygosity is not maintained in full fidelity. For the individuals in our colony, not all 23 *A. arizonae* and all 23 *A. marmoratus* chromosomes are inherited in subsequent generations. While the mechanism of endoduplication prior to meiosis and identical chromosome pairing holds true to generate successful offspring, these results suggest that identical chromosome pairing does not occur every single time and in rare instances, pairing

and crossover of syntenic chromosomes of the two different species within *A. neomexicanus* can occur. The consequence is a loss of heterozygosity (LOH).

Table 11. Primers used for the amplification of junctions between *A. arizonae* Chromosome 17 and *A. marmoratus* Chromosome 15.

Oligo number	Sequence (5' --> 3')
8427	GTGTATCCAAACTCAGCACCT
8429	GAGCCAAACATCTTTCCTTAGAG
8459	CAGAGAGGATGGCATTCCGA
8460	GGCAGCAAATTCCACAGACC
8461	CTCCTGGCAGCAAATTCCGT
8464	ACCTGAGCATCCATGTACCT

Table 12. Primer pairs and PCR conditions for amplification of junctions between *A. arizonae* Chromosome 17 and *A. marmoratus* Chromosome 17.

				PCR conditions		
Fragment amplified (Figure 26)	**Primer pair**	**Polymerase**	**Annealing temperature (°C)**	**Initial denaturation**	**Cycles (30 X)**	**Final extension**
A-D	8427/ 8429	Taq	58	95 °C (30 sec)	95 °C (30 sec), 58 °C (30 sec), 68 °C (30 sec)	68 °C (5 min)
I-A	8461/ 8459	Taq	58	95 °C (30 sec)	95 °C (30 sec), 58 °C (30 sec), 68 °C (30 sec)	68 °C (5 min)
M-A	8460/ 8459	Q5	68	98 °C (30 sec)	98 °C (30 sec), 68 °C (30 sec), 72 °C (30 sec)	72 °C (2 min)
M-C	8460/ 8464	Taq	58	95 °C (30 sec)	95 °C (30 sec), 58 °C (30 sec), 68 °C (30 sec)	68 °C (5 min)

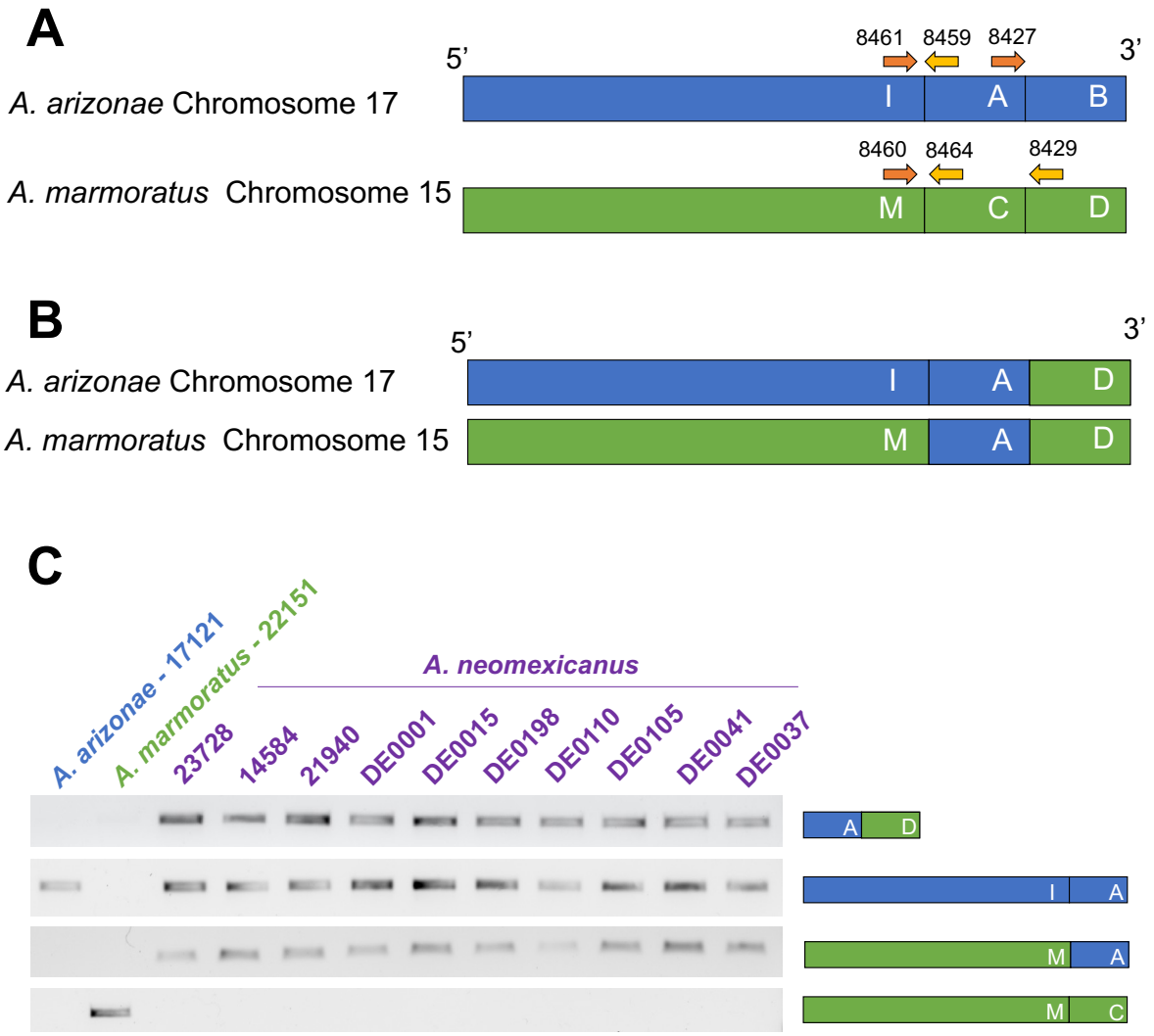

Figure 26. Recombination between syntenic chromosomes in *A. neomexicanus* leads to loss of heterozygosity. (A) *A. arizonae* chromosome 17 ("I") and *A. marmoratus* chromosome 15 ("M") are syntenic chromosomes within *A. neomexicanus*. A-D are approximately 1 Mb regions on the 3' end of the chromosomes. The null hypothesis is that in *A. neomexicanus* the entirety of both chromosomes are inherited. Arrows with numbers denote primer names that bind in the relative parts of the chromosomes used for PCR. (B) Proposed chromosome structure due to recombination between the syntenic chromosomes. Segments C and B are no longer present leading to a loss of heterozygosity. (C). Agarose gel showing bands from PCR products amplifying the junctions where recombination has occurred. Ten *A. neomexicanus* animals were tested and all individuals showed the same pattern.

Another approach that was attempted was to use Nanopore sequencing to sequence the junction on *A. arizonae* chromosome 17 where the *A. marmoratus* chromosome 15 is fused ("A-D" in Figure XXXB). The CRISPR-Cas9 protocol paired with Nanopore sequencing utilizes a guide RNA to direct the Cas9 endonuclease to a targeted region of the DNA. A guide RNA (gRNA) is created using a combination of CRISPR RNA (crRNA) that matches the target sequence in the genome and a trans-activating CRISPR RNA (tracRNA). This

gRNA forms a complex with the addition of the Cas9 protein, which creates a double-stranded break at the target site. After the break, Nanopore adapters are added to create a sequencing library that is loaded onto a Nanopore flow cell. These adapters include motor proteins that help guide the DNA molecules through the nanopores in the flow cell. One of the theoretical advantages of this approach is the ability to directly sequence the native DNA without the need for amplification or extensive library preparation. This is particularly useful when analyzing regions of the genome that are difficult to amplify, such as repetitive regions.

In this case, four different crRNA was designed 1000 to 2000 bp upstream of the A-D junction in Figure 26B. The crRNAs were designed as such to be specific to the *A. arizonae* chromosome 17 and *A. marmoratus* chromosome 15 (Table 13). By having more than one crRNA in a reaction, the probability of enriching for the region of interest is higher through this "tiling" approach. Once the gRNA was assembled, 5000 ng of gDNA from *A. neomexicanus* (ID D37) was used as the input. The library preparation was done with the Cas9 enrichment kit from ONT (SQK-CS9109) in conjunction with the LSK-109 kit (SQK-LSK109).

The resulting library was loaded on an R9.4.1 minION flow cell and sequenced on a Linux platform (MinKNOW v. 20.10.3, MinKNOW Core v. 4.1.2) and base called with Guppy (v. 4.2.2) on high-accuracy mode. A total of 866 reads were sequenced with an N50 read length of 26,943 bp. The low number of reads could have been a result of two main reasons: the low abundance of targeted regions in the input gDNA (due to shearing, poor quality gDNA) and the less-than-optimal library preparation from a first-time user. Despite the less-than-disappointing output in terms of number of reads, which the majority which were off-target sequences, there were indeed four reads (0.46% on-target) that aligned to the A-D junction on *A. arizonae* Chromosome 17 (Figure 27). Despite the low output, 0.46% is not outside the range that was published at the time for this Cas9-target sequencing, which ranged from 0.44% - 4.61% (Gilpatrick et al., 2020).

Table 13. crRNA used for the Nanopore Cas9-enrichment protocol.

Name	Sequence	Product type	Orientation	Chromosome	Start position in ref. genome
BLgRNA1	TTTCACGCCTGACGTGCCAT	Alt-R® CRISPR-Cas9 crRNA	Forward	*A. arizonae* Chr 17	22255049
BLgRNA2	CCCACTACCATCCACCACGG	Alt-R® CRISPR-Cas9 crRNA	Forward	*A. arizonae* Chr 17	22256311
BLgRNA3	CGTCCGGAGGAGCCACGCAC	Alt-R® CRISPR-Cas9 crRNA	Reverse	*A. marmoratus* Chr 15	24221998
BLgRNA4	CTCAAAGGCATATGATCGCA	Alt-R® CRISPR-Cas9 crRNA	Reverse	*A. marmoratus* Chr 15	24220894

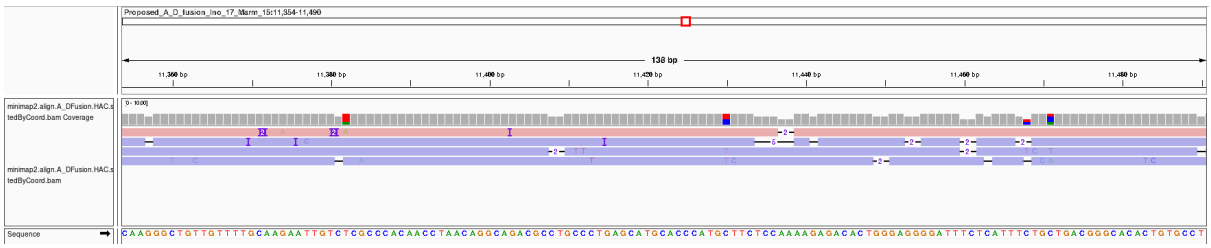

Figure 27. Integrative Genomics Viewer (IGV) screenshot showing the four reads from the Nanopore Cas9-enrichment experiment.These reads align to the A-D junction on *A. arizonae* Chromosome 17. The pink read indicates it came from the gRNA targeting the A region, upstream from the junction (BLgRNA1/2) Purple indicates the reads came from the gRNA targeting the D region, downstream from the junction (BLgRNA3/4). The insertions, deletions, and SNPs observed are typical of sequencing data generated from R9.4.1 Nanopore flow cells.

Despite obtaining reads showing that recombination occurs between syntenic *A. arizonae* chromosome 17 and *A. marmoratus* chromosome 15, using Cas9-enrichment paired with Nanopore sequencing did not appear to be the most ideal use of the technology. Designing guide RNAs for non-model organisms was not an easy task. In human cell lines and mice, databases and tools are already developed for the user to choose a specific locus in mind, and it automatically outputs the best crRNAs to use. In non-model organisms like whiptails, we do not have information about the probability of off-target effects and the efficiency of cutting by Cas9.

In comparison, WGS using Nanopore is a more promising use of the technology as many researchers use it for this purpose and therefore there is more support from both the company and the research community. The library preparation for Nanopore-based WGS begins with high-molecular weight (HMW) gDNA extraction to ensure long reads. The gDNA is then subjected to an optional fragmentation step (for example by sonication with the Covaris machine), followed by the attachment of sequencing adapters. The amount of output highly depends on the quality of the input gDNA, quality of the sequencing library made, and quality of the flow cell used. Despite the potential for lower coverage due to many factors that are either in the control or not of the user, these long reads can have the potential to provide valuable information on structural variations that might be missed by short-read sequencing. For example, large insertions, deletions, inversions, or translocations can be more accurately detected with long-read data.

Performing Nanopore WGS on an *A. neomexicanus* individual, we can see that even though the coverage is lower than an Illumina run, there is even coverage across all 46 chromosomes (Figure 27A). Like in the Illumina data, there are regions of elevated and missing coverage between syntenic chromosomes indicating pairing between these chromosomes have occurred (Figure 27B-C).

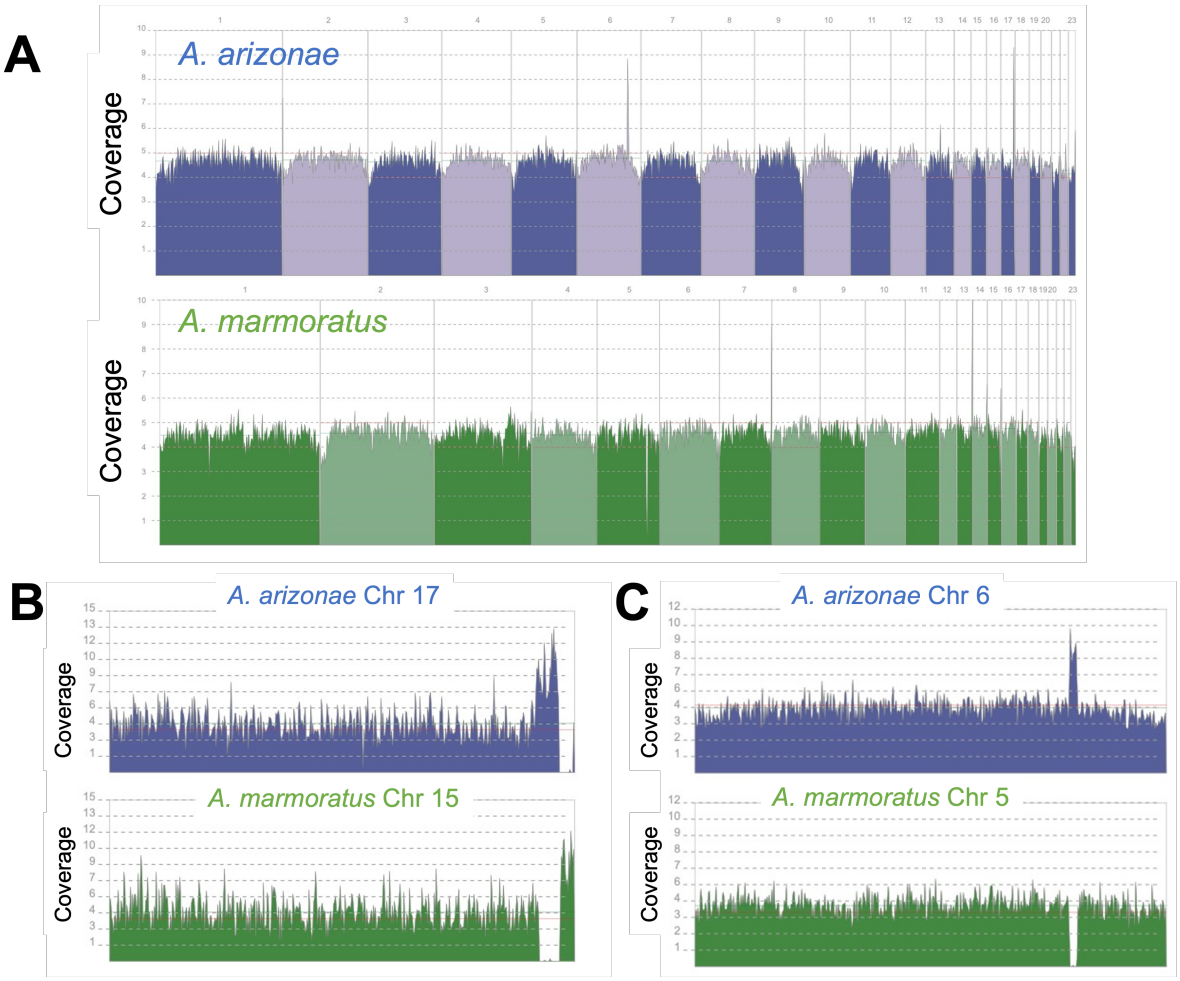

Figure 28. Nanopore whole-genome sequencing of *A. neomexicanus*. (A) Coverage across all 46 chromosomes from *A. arizonae* (blue) and *A. marmoratus* (green). Ordered from largest to smallest, different shades of blue and green delineate the chromosomes. (B) Syntenic pair *A. arizonae* chromosome 17 and *A. marmoratus* chromosome 15. Elevated and missing regions at the 3' end of the chromosomes are evident, reflecting the recombination that has occurred between the two chromosomes. (C) Syntenic pair *A. arizonae* chromosome 6 and *A. marmoratus* chromosome 5. Elevated and missing regions at the 3' end of the chromosomes are evident, reflecting the recombination that has occurred between the two chromosomes.

All of the six *A. neomexicanus* animals with whole-genome sequencing (IDs 16923, 23563, D15, D37, 14584, 11194) show a LOH between *A. arizonae* Chromosome 17/*A. marmoratus* Chromosome 15 and *A. arizonae* Chromosome 6/*A. marmoratus* Chromosome 5. Interestingly, two of these animals (IDs 14584 and 11194) show a LOH between *A. arizonae* Chromosome 2/*A. marmoratus* Chromosome 1 as well. The differences between these individuals reflect similar patterns seen in the MS data: there is genetic diversity within *A. neomexicanus*. In this case, the genetic diversity in question is the different patterns of LOH.

Recombination between syntenic chromosomes in polyploid hybrids are often called homeologous exchanges as opposed to homolog exchanges that are associated with meiosis (Deb et al., 2023). While homeologous exchange is highly studied in plants,

especially crops, there are examples within the animal kingdom as well. Whether the mechanism is due to identical or homolog chromosome recognition failure or homeolog pairing favoritism, it is clear that polyploid fish (Allendorf et al., 2015; Li et al., 2021) and amphibians (Bi & Bogart, 2006) also show this phenomenon. With the results presented here, we add whiptails to the list and is the first known example in reptiles.

In evolutionary biology, these syntenic or homeologous exchanges are hypothesized to be significant because they potentially facilitate hybrid vigor. Examples of the advantages of reshuffling chromosomes from the subgenomes are particularly limited to the plant literature where polyploid crop plants can develop pathogen resistance and are less sensitive to salinity conditions (Wang et al., 2022). The benefits from homeologous exchange in plants can be explained for example by beneficial alleles gaining copy numbers and therefore promote stronger phenotypes (Zhang et al., 2020). In addition, plants of higher ploidy-level could have exchanges that are reciprocal and therefore no LOH is observed in these cases (Deb et al., 2023). The shuffling of the subgenomes though could still result in structural variations that promote or disrupt gene expression.

However, while homeologous exchanges contribute to genetic diversity and phenotypic innovation, they can also pose significant evolutionary risks due to LOH. Through the lens of Muller's ratchet, LOH is irreversible and therefore these homeologous exchanges can lead to a deterioration of the genome through the accumulation of deleterious mutations. Additionally, homeologous exchange could be hypothesized to place deleterious mutations into regions of the genome where they cannot easily be purged through selection, potentially leading to an increased genetic load.

Another potential disadvantage of homeologous exchange is tied to the Red Queen hypothesis. In the absence of LOH, homeologous recombination can introduce novel genetic combinations through for example gene fusions, however, it may also disrupt co-adapted gene networks. Depending on the rate of exchange, homeologous recombination may be another barrier to adaptation because these cells now have to additionally compensate for continuing disruption of cellular pathways. Under these considerations, homeologous exchanges represent a double-edged sword and understanding the mechanisms and consequences is essential for both evolutionary biology in general and agricultural programs, where successful manipulation of these exchanges can help improve crop traits while minimizing negative outcomes.

In addition to LOH, there are patterns observed in crop plants that mirror our results in *A. neomexicanus*. For example, syntenic chromosomes are the most frequently observed exchanges. *Brassica napus* (rapeseed) is an allotetraploid hybrid between *B. rapa* and *B. oleracea* and the major subgenome exchanges are observed between the syntenic chromosomes pairs 1, 2, and 9 (Higgins et al., 2018; Lloyd et al., 2018). Despite high levels

of synteny, not all chromosome pairs in hybrids show evidence of exchanges, suggesting there is more sophisticated control during meiosis of which chromosomes pair together (Deb et al., 2023). Alternatively, smaller regions of exchange can occur but not detected because of conserved sequence identity between the subgenomes. In *A. neomexicanus*, we have found only three pairs of chromosomes that have regions of exchanges, despite the percent identity across all pairs to be very similar (Table 8). What is clear though is that these regions of exchange across taxa are not random and appears to be either exclusively closer to the telomeres (Nicolas et al., 2012; Wu et al., 2021) or centromeres (Mandáková et al., 2014). In contrast, the regions of exchange in *A. neomexicanus* appear to be positioned not only near the telomeres (*A. arizonae* Chromosome 17/*A. marmoratus* Chromosome 15) but also near the centromeres (*A. arizonae* Chromosome 6/*A. marmoratus* Chromosome 5 and *A. arizonae* Chromosome 2/*A. marmoratus* Chromosome 1).

Other observations near the sites of these exchanges include: high gene density, transposable element enrichment, reduced methylation levels, and low nucleosome density (Deb et al., 2023). Taken together, these patterns suggest a non-randomness to homeologous exchanges and more examples and bigger sample sizes will be needed to determine if those mentioned factors are important in the rate of exchange in *A. neomexicanus*. Importantly, these homeologous exchanges lead to LOH in *A. neomexicanus* but in its own way generates genetic diversity among seemingly clonal individuals. Whether there are benefits is predicated on the evolutionary history of the species and how often these exchanges occur and maintained.

Here we have shown that both short-read and low-coverage long-read sequencing can show the presence of recombination or homeologous exchanges between syntenic chromosomes in *A. neomexicanus*. While identical (homologous) chromosome pairing does occur to maintain heterozygosity in the majority of times, there are instances where syntenic chromosomes recombine instead leading to a LOH. While the animals we have sequences show that only 1% of the genome is affected by this, the mechanisms and consequences of this still need to be investigated.

6. Conclusions and future prospects

Within a single genus *Aspidoscelis*, sexual reproduction, facultative parthenogenesis, and obligate parthenogenesis can all occur. Interestingly, this demonstrates the reproductive plasticity within these whiptail lizards and highlights why studying them is important in the context of not only reproductive biology but also evolutionary biology.

We have shown through that genomics through genome assemblies is essential in propelling research into whiptail biology further. The genome assembly of *A. marmoratus* helped us understand the mechanism of facultative parthenogenesis in this genus and how an unfertilized oocyte of a gonochoristic whiptail can develop into an embryo by a duplication of its haploid genome. Chromosome-level assemblies of *A. marmoratus* and *A. arizonae* used together serve well as a proxy for the genome reference of *A. neomexicanus* and helped reveal that even though identical chromosomes pair most of the time, there are instances of syntenic chromosome pairing that leads to LOH. The other chromosome-level assemblies shown here can be applied to look at other unisexual species with different hybrid combinations. These chromosome-level assemblies provide us the foundation to start investigating whiptails of higher ploidy levels through the lens of sequencing, to be complemented by cytology and microscopy.

The continuing use of next-generation sequencing by other research groups to understand patterns of hybridization and inter- and intra-specific diversity within whiptails will benefit from these genome assemblies to use as a reference. Still, many questions concerning whiptail lizards still are left for investigation that can benefit from these assemblies: (1) Are there patterns of LOH in other diploid unisexuals? (2) Do triploid and tetraploid unisexuals exhibit LOH as well? (3) What are the repeat element landscapes in each species and how does this change in unisexual lineages?

In our results here concerning *A. neomexicanus*, that these syntenic exchanges span at least 1 Mb long and therefore the signal here is clear from the sequencing data but what is not clear is if there are smaller-sized exchanges. With triploids such as *A. sonarae* where the subgenomes can be described as AAB, are there exchanges between the A genomes and B genome or are they exclusively between the A genomes, leading to harlequin chromosomes (Rachel et al., 1991)? At the transcriptional level, while there are already gene models in progress for some of these species, comprehensive annotations of all the genomes are still needed to investigate questions regarding allele-specific expression, gene dosage compensation, and the interaction between subgenomes within unisexuals.

As we continue to sequence and investigate whiptail lizards, these genome assemblies equip the community to better address these outstanding questions in a frustrating group of lizards that can “drive the systematist to despair.”

7. References

Alexander, M., Ho, S. Y. W., Molak, M., Barnett, R., Carlborg, Ö., Dorshorst, B., Honaker, C., Besnier, F., Wahlberg, P., Dobney, K., Siegel, P., Andersson, L., & Larson, G. (2015). Mitogenomic analysis of a 50-generation chicken pedigree reveals a rapid rate of mitochondrial evolution and evidence for paternal mtDNA inheritance. *Biology Letters*, *11*(10), 20150561. https://doi.org/10.1098/rsbl.2015.0561

Allendorf, F. W., Bassham, S., Cresko, W. A., Limborg, M. T., Seeb, L. W., & Seeb, J. E. (2015). Effects of Crossovers Between Homeologs on Inheritance and Population Genomics in Polyploid-Derived Salmonid Fishes. *Journal of Heredity*, *106*(3), 217–227. https://doi.org/10.1093/jhered/esv015

Aury, J.-M., Engelen, S., Istace, B., Monat, C., Lasserre-Zuber, P., Belser, C., Cruaud, C., Rimbert, H., Leroy, P., Arribat, S., Dufau, I., Bellec, A., Grimbichler, D., Papon, N., Paux, E., Ranoux, M., Alberti, A., Wincker, P., & Choulet, F. (2022). Long-read and chromosome-scale assembly of the hexaploid wheat genome achieves high resolution for research and breeding. *GigaScience*, *11*, giac034. https://doi.org/10.1093/gigascience/giac034

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

Barley, A. J., Nieto-Montes de Oca, A., Manríquez-Morán, N. L., & Thomson, R. C. (2022). The evolutionary network of whiptail lizards reveals predictable outcomes of hybridization. *Science*, *377*(6607), 773–777. https://doi.org/10.1126/science.abn1593

Barley, A. J., Reeder, T. W., Nieto-Montes de Oca, A., Cole, C. J., & Thomson, R. C. (2021). A New Diploid Parthenogenetic Whiptail Lizard from Sonora, Mexico, Is the “Missing Link” in the Evolutionary Transition to Polyploidy. *The American Naturalist*, *198*(2), 295–309. https://doi.org/10.1086/715056

Bi, K., & Bogart, J. P. (2006). Identification of intergenomic recombinations in unisexual salamanders of the genus Ambystoma by genomic in situ hybridization (GISH). *Cytogenetic and Genome Research*, *112*(3–4), 307–312. https://doi.org/10.1159/000089885

Bi, K., & Bogart, J. P. (2010). Time and time again: Unisexual salamanders (genus Ambystoma) are the oldest unisexual vertebrates. *BMC Evolutionary Biology*, *10*(1), 238. https://doi.org/10.1186/1471-2148-10-238

Birchler, J. A., Yao, H., & Chudalayandi, S. (2006). Unraveling the genetic basis of hybrid vigor. *Proceedings of the National Academy of Sciences*, *103*(35), 12957–12958. https://doi.org/10.1073/pnas.0605627103

Birky, C. W. (2004). Bdelloid rotifers revisited. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(9), 2651–2652. https://doi.org/10.1073/pnas.0308453101

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boore, J. L. (1999). Animal mitochondrial genomes. *Nucleic Acids Research*, *27*(8), 1767–1780.

Booth, W., Johnson, D. H., Moore, S., Schal, C., & Vargo, E. L. (2010). Evidence for viable, non-clonal but fatherless Boa constrictors. *Biology Letters*, *7*(2), 253–256. https://doi.org/10.1098/rsbl.2010.0793

Booth, W., Levine, B. A., Corush, J. B., Davis, M. A., Dwyer, Q., De Plecker, R., & Schuett, G. W. (2023). Discovery of facultative parthenogenesis in a new world crocodile. *Biology Letters*, *19*(6), 20230129. https://doi.org/10.1098/rsbl.2023.0129

Burt, C. E. (1931). *A Study of the Teiid Lizards of the Genus Cnemidophorus with Special Reference to Their Phylogenetic Relationships*. U.S. Government Printing Office.

Burton, J. N., Adey, A., Patwardhan, R. P., Qiu, R., Kitzman, J. O., & Shendure, J. (2013). Chromosome-scale scaffolding of de novo genome assemblies based on chromatin interactions. *Nature Biotechnology*, *31*(12), 1119–1125. https://doi.org/10.1038/nbt.2727

Chapman, D. D., Firchau, B., & Shivji, M. S. (2008). Parthenogenesis in a large-bodied requiem shark, the blacktip Carcharhinus limbatus. *Journal of Fish Biology*, *73*(6), 1473–1477. https://doi.org/10.1111/j.1095-8649.2008.02018.x

Chapman, D. D., Shivji, M. S., Louis, E., Sommer, J., Fletcher, H., & Prodöhl, P. A. (2007). Virgin birth in a hammerhead shark. *Biology Letters*, *3*(4), 425–427. https://doi.org/10.1098/rsbl.2007.0189

Charles J. Cole, Diana P. Baumann, Harry L. Taylor, Nadine Bobon, David V. Ho, William B. Neaves, & Peter Baumann. (2023). RETICULATE PHYLOGENY: A NEW TETRAPLOID PARTHENOGENETIC WHIPTAIL LIZARD DERIVED FROM HYBRIDIZATION AMONG FOUR BISEXUAL ANCESTRAL SPECIES OF ASPIDOSCELIS (REPTILIA: SQUAMATA: TEIIDAE). *Bulletin of the Museum of Comparative Zoology*, *163*(7), 247–279. https://doi.org/10.3099/MCZ76

Cole, C. J. (1985). Taxonomy of Parthenogenetic Species of Hybrid Origin. *Systematic Zoology*, *34*(3), 359–363. https://doi.org/10.2307/2413153

Cole, C. J., Cole, C. J., Dessauer, H. C., & Barrowclough, G. F. (1988). *Hybrid origin of a unisexual species of whiptail lizard, Cnemidophorus neomexicanus, in western North America: New evidence and a review: Vol. no. 2905* (pp. 1–40). American Museum of Natural History. https://www.biodiversitylibrary.org/item/317104

Cole, C. J., Taylor, H. L., Neaves, W. B., Baumann, D. P., Newton, A., Schnittker, R., & Baumann, P. (2017). The Second Known Tetraploid Species of Parthenogenetic Tetrapod (Reptilia: Squamata: Teiidae): Description, Reproduction, Comparisons With Ancestral Taxa, And Origins Of Multiple Clones. *Bulletin of the Museum of Comparative Zoology*, *161*(8), 285–321. https://doi.org/10.3099/MCZ37.1

Crosland, M. W. J., & Crozier, R. H. (1986). Myrmecia pilosula, an Ant with Only One Pair of Chromosomes. *Science*, *231*(4743), 1278–1278. https://doi.org/10.1126/science.231.4743.1278

Darevsky, I. S. (1958). Natural parthenogenesis in certain subspecies of rocky lizard, Lacerta saxicola Eversmann. *Dokl Nauk SSSR Biol. Sci*, 877–879.

de Vos, J. M., Augustijnen, H., Bätscher, L., & Lucek, K. (2020). Speciation through chromosomal fusion and fission in Lepidoptera. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1806), 20190539. https://doi.org/10.1098/rstb.2019.0539

Deb, S. K., Edger, P. P., Pires, J. C., & McKain, M. R. (2023). Patterns, mechanisms, and consequences of homoeologous exchange in allopolyploid angiosperms: A genomic and

epigenomic perspective. *New Phytologist*, *238*(6), 2284–2304. https://doi.org/10.1111/nph.18927

Dekker, J., Rippe, K., Dekker, M., & Kleckner, N. (2002). Capturing Chromosome Conformation. *Science*, *295*(5558), 1306–1311. https://doi.org/10.1126/science.1067799

Dessauer, H. C., & Cole, C. J. (1984). Influence of gene dosage on electrophoretic phenotypes of proteins from lizards of the genus *Cnemidophorus*. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, *77*(1), 181–189. https://doi.org/10.1016/0305-0491(84)90241-4

Dierckxsens, N., Mardulyn, P., & Smits, G. (2017). NOVOPlasty: De novo assembly of organelle genomes from whole genome data. *Nucleic Acids Research*, *45*(4), e18. https://doi.org/10.1093/nar/gkw955

Duan, Z., Andronescu, M., Schutz, K., McIlwain, S., Kim, Y. J., Lee, C., Shendure, J., Fields, S., Blau, C. A., & Noble, W. S. (2010). A three-dimensional model of the yeast genome. *Nature*, *465*(7296), 363–367. https://doi.org/10.1038/nature08973

Dubach, J., Sajewicz, A., & Pawley, R. (1997). Parthenogenesis in the Arafuran filesnake (Acrochordus arafurae). *Herpetological Natural History*, *5*(1), 11–18.

Dudgeon, C. L., Coulton, L., Bone, R., Ovenden, J. R., & Thomas, S. (2017). Switch from sexual to parthenogenetic reproduction in a zebra shark. *Scientific Reports*, *7*, 40537. https://doi.org/10.1038/srep40537

Duellman, W. E., & Zweifel, R. G. (1962). *A synopsis of the lizards of the sexlineatus group (genus Cnemidophorus)*. http://archive.org/details/bulletin-american-museum-natural-history-123-159-210

Ellegren, H. (2004). Microsatellites: Simple sequences with complex evolution. *Nature Reviews Genetics*, *5*(6), 435–445. https://doi.org/10.1038/nrg1348

Feldheim, K. A., Chapman, D. D., Sweet, D., Fitzpatrick, S., Prodöhl, P. A., Shivji, M. S., & Snowden, B. (2010). Shark Virgin Birth Produces Multiple, Viable Offspring. *Journal of Heredity*, *101*(3), 374–377. https://doi.org/10.1093/jhered/esp129

Fonseca, P. L. C., De-Paula, R. B., Araújo, D. S., Tomé, L. M. R., Mendes-Pereira, T., Rodrigues, W. F. C., Del-Bem, L.-E., Aguiar, E. R. G. R., & Góes-Neto, A. (2021). Global Characterization of Fungal Mitogenomes: New Insights on Genomic Diversity and Dynamism of Coding Genes and Accessory Elements. *Frontiers in Microbiology*, *12*, 787283. https://doi.org/10.3389/fmicb.2021.787283

Fujita, M. K., & Moritz, C. (2010). Origin and Evolution of Parthenogenetic Genomes in Lizards: Current State and Future Directions. *Cytogenetic and Genome Research*, *127*(2–4), 261–272. https://doi.org/10.1159/000295177

Gadow, H. (1906). A Contribution to the Study of Evolution based upon the Mexican Species of Cnemidophorus. *Proceedings of the Zoological Society of London*, *76*(1–2), 277–375. https://doi.org/10.1111/j.1469-7998.1906.tb08438.x

Geneious Prime 2017.10.1.3. (n.d.). [Computer software]. https://www.geneious.com

Ghurye, J., Pop, M., Koren, S., Bickhart, D., & Chin, C.-S. (2017). Scaffolding of long read assemblies using long range contact information. *BMC Genomics*, *18*(1), 527. https://doi.org/10.1186/s12864-017-3879-z

Gilpatrick, T., Lee, I., Graham, J. E., Raimondeau, E., Bowen, R., Heron, A., Downs, B., Sukmar, S., Sedlazeck, F. J., & Timp, W. (2020). Targeted nanopore sequencing with Cas9-guided adaptor ligation. *Nature Biotechnology*, *38*(4), 433–438. https://doi.org/10.1038/s41587-020-0407-5

Grabherr, M. G., Russell, P., Meyer, M., Mauceli, E., Alföldi, J., Di Palma, F., & Lindblad-Toh, K. (2010). Genome-wide synteny through highly sensitive sequence alignment: Satsuma. *Bioinformatics*, *26*(9), 1145–1151. https://doi.org/10.1093/bioinformatics/btq102

Gray, M. W. (2012). Mitochondrial Evolution. *Cold Spring Harbor Perspectives in Biology*, *4*(9), a011403. https://doi.org/10.1101/cshperspect.a011403

Gualberto, J. M., Mileshina, D., Wallet, C., Niazi, A. K., Weber-Lotfi, F., & Dietrich, A. (2014). The plant mitochondrial genome: Dynamics and maintenance. *Biochimie*, *100*, 107–120. https://doi.org/10.1016/j.biochi.2013.09.016

Guo, M., Yuan, C., Tao, L., Cai, Y., & Zhang, W. (2022). Life barcoded by DNA barcodes. *Conservation Genetics Resources*, *14*(4), 351–365. https://doi.org/10.1007/s12686-022-01291-2

Hadany, L., & Comeron, J. M. (2008). Why are sex and recombination so common? *Annals of the New York Academy of Sciences*, *1133*, 26–43. https://doi.org/10.1196/annals.1438.011

Hardy, L. M., & Cole, C. J. (1998). *Morphology of a sterile, tetraploid, hybrid whiptail lizard (Squamata, Teiidae, Cnemidophorus). American Museum novitates ; no. 3228*. http://hdl.handle.net/2246/3369

Hatleberg, W. L., & Hinman, V. F. (2021). Chapter Two - Modularity and hierarchy in biological systems: Using gene regulatory networks to understand evolutionary change. In S. F. Gilbert (Ed.), *Current Topics in Developmental Biology* (Vol. 141, pp. 39–73). Academic Press. https://doi.org/10.1016/bs.ctdb.2020.11.004

Higgins, E. E., Clarke, W. E., Howell, E. C., Armstrong, S. J., & Parkin, I. A. P. (2018). Detecting de Novo Homoeologous Recombination Events in Cultivated Brassica napus Using a Genome-Wide SNP Array. *G3 Genes|Genomes|Genetics*, *8*(8), 2673–2683. https://doi.org/10.1534/g3.118.200118

Imai, H. T., & Taylor, R. W. (1989). Chromosomal polymorphisms involving telomere fusion, centromeric inactivation and centromere shift in the ant Myrmecia (pilosula) n=1. *Chromosoma*, *98*(6), 456–460. https://doi.org/10.1007/BF00292792

Imakaev, M., Fudenberg, G., McCord, R. P., Naumova, N., Goloborodko, A., Lajoie, B. R., Dekker, J., & Mirny, L. A. (2012). Iterative correction of Hi-C data reveals hallmarks of chromosome organization. *Nature Methods*, *9*(10), 999–1003. https://doi.org/10.1038/nmeth.2148

Kinney, M. E., Wack, R. F., Grahn, R. A., & Lyons, L. (2013). Parthenogenesis in a Brazilian Rainbow Boa (Epicrates cenchria cenchria). *Zoo Biology*, *32*(2), 172–176. https://doi.org/10.1002/zoo.21050

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: A fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, *35*(21), 4453–4455. https://doi.org/10.1093/bioinformatics/btz305

Kubo, T., & Newton, K. J. (2008). Angiosperm mitochondrial genomes and mutations. *Mitochondrion*, *8*(1), 5–14. https://doi.org/10.1016/j.mito.2007.10.006

Lajoie, B. R., Dekker, J., & Kaplan, N. (2015). The Hitchhiker's Guide to Hi-C Analysis: Practical guidelines. *Methods (San Diego, Calif.)*, *72*, 65–75. https://doi.org/10.1016/j.ymeth.2014.10.031

Lampert, K. P. (2008). Facultative Parthenogenesis in Vertebrates: Reproductive Error or Chance? *Sexual Development*, *2*(6), 290–301. https://doi.org/10.1159/000195678

Lampert, K. P., & Schartl, M. (2008). The origin and evolution of a unisexual hybrid: Poecilia formosa. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1505), 2901–2909. https://doi.org/10.1098/rstb.2008.0040

Lampert, K. P., & Schartl, M. (2010). A little bit is better than nothing: The incomplete parthenogenesis of salamanders, frogs and fish. *BMC Biology*, *8*, 78. https://doi.org/10.1186/1741-7007-8-78

Lampert, K. P., Steinlein, C., Schmid, M., Fischer, P., & Schartl, M. (2007). A haploid-diploid-triploid mosaic of the Amazon molly, Poecilia formosa. *Cytogenetic and Genome Research*, *119*(1–2), 131–134. https://doi.org/10.1159/000109629

Leaché, A. D., Banbury, B. L., Linkem, C. W., & de Oca, A. N.-M. (2016). Phylogenomics of a rapid radiation: Is chromosomal evolution linked to increased diversification in north american spiny lizards (Genus Sceloporus)? *BMC Evolutionary Biology*, *16*(1), 63. https://doi.org/10.1186/s12862-016-0628-x

Lehtonen, J., Jennions, M. D., & Kokko, H. (2012). The many costs of sex. *Trends in Ecology & Evolution*, *27*(3), 172–178. https://doi.org/10.1016/j.tree.2011.09.016

Lenk, P., Eidenmueller, B., Staudter, H., Wicker, R., & Wink, M. (2005). A parthenogenetic Varanus. *Amphibia-Reptilia*, *26*(4), 507–514. https://doi.org/10.1163/156853805774806296

Li, J.-T., Wang, Q., Huang Yang, M.-D., Li, Q.-S., Cui, M.-S., Dong, Z.-J., Wang, H.-W., Yu, J.-H., Zhao, Y.-J., Yang, C.-R., Wang, Y.-X., Sun, X.-Q., Zhang, Y., Zhao, R., Jia, Z.-Y., & Wang, X.-Y. (2021). Parallel subgenome structure and divergent expression evolution of allo-tetraploid common carp and goldfish. *Nature Genetics*, *53*(10), 1493–1503. https://doi.org/10.1038/s41588-021-00933-9

Lieberman-Aiden, E., van Berkum, N. L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B. R., Sabo, P. J., Dorschner, M. O., Sandstrom, R., Bernstein, B., Bender, M. A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L. A., Lander, E. S., & Dekker, J. (2009). Comprehensive mapping of long range interactions reveals folding principles of the human genome. *Science (New York, N.Y.)*, *326*(5950), 289–293. https://doi.org/10.1126/science.1181369

Liu, J., Li, M., Zhang, Q., Wei, X., & Huang, X. (2020). Exploring the molecular basis of heterosis for plant breeding. *Journal of Integrative Plant Biology*, *62*(3), 287–298. https://doi.org/10.1111/jipb.12804

Lloyd, A., Blary, A., Charif, D., Charpentier, C., Tran, J., Balzergue, S., Delannoy, E., Rigaill, G., & Jenczewski, E. (2018). Homoeologous exchanges cause extensive dosage-dependent gene expression changes in an allopolyploid crop. *The New Phytologist*, *217*(1), 367–377. https://doi.org/10.1111/nph.14836

Lowe, C. H., & Wright, J. W. (1966). Evolution of Parthenogenetic Species of Cnemidophorus (Whiptail Lizards) in Western North America. *Journal of the Arizona Academy of Science*, *4*(2), 81–87.

Lowe, C. J., Wright, J. W., Cole, C. J., & Bezy, R. L. (1970). Natural Hybridization Between the Teiid Lizards Cnemidophorus Sonorae (Parthenogenetic) and Cnemidophorus Tigris (Bisexual). *Systematic Biology*, *19*(2), 114–127. https://doi.org/10.2307/2412449

Lutes, A. A., Neaves, W. B., Baumann, D. P., Wiegraebe, W., & Baumann, P. (2010). Sister chromosome pairing maintains heterozygosity in parthenogenetic lizards. *Nature*, *464*(7286), 283–286. https://doi.org/10.1038/nature08818

Lutes, A., Baumann, D., Neaves, W., & Baumann, P. (2011). Laboratory synthesis of an independently reproducing vertebrate species. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 9910–9915. https://doi.org/10.1073/pnas.1102811108

Mackintosh, A., Vila, R., Laetsch, D. R., Hayward, A., Martin, S. H., & Lohse, K. (2023). Chromosome Fissions and Fusions Act as Barriers to Gene Flow between Brenthis Fritillary Butterflies. *Molecular Biology and Evolution*, *40*(3), msad043. https://doi.org/10.1093/molbev/msad043

Mandáková, T., Marhold, K., & Lysak, M. A. (2014). The widespread crucifer species Cardamine flexuosa is an allotetraploid with a conserved subgenomic structure. *The New Phytologist*, *201*(3), 982–992. https://doi.org/10.1111/nph.12567

Maslin, T. P. (1962). All-Female Species of the Lizard Genus Cnemidophorus, Teiidae. *Science*, *135*(3499), 212–213. https://doi.org/10.1126/science.135.3499.212

Mayrose, I., & Lysak, M. A. (2020). The Evolution of Chromosome Numbers: Mechanistic Models and Experimental Approaches. *Genome Biology and Evolution*, *13*(2), evaa220. https://doi.org/10.1093/gbe/evaa220

Melo, D., Porto, A., Cheverud, J. M., & Marroig, G. (2016). Modularity: Genes, development and evolution. *Annual Review of Ecology, Evolution, and Systematics*, *47*, 463. https://doi.org/10.1146/annurev-ecolsys-121415-032409

Montooth, K. L., & Rand, D. M. (2008). The Spectrum of Mitochondrial Mutation Differs across Species. *PLoS Biology*, *6*(8), e213. https://doi.org/10.1371/journal.pbio.0060213

Mueller, R. L. (2006). Evolutionary rates, divergence dates, and the performance of mitochondrial genes in Bayesian phylogenetic analysis. *Systematic Biology*, *55*(2), 289–300. https://doi.org/10.1080/10635150500541672

Muffato, M., Louis, A., Nguyen, N. T. T., Lucas, J., Berthelot, C., & Roest Crollius, H. (2023). Reconstruction of hundreds of reference ancestral genomes across the eukaryotic kingdom. *Nature Ecology & Evolution*, *7*(3), 355–366. https://doi.org/10.1038/s41559-022-01956-z

Muller, H. J. (1964). The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, *1*(1), 2–9. https://doi.org/10.1016/0027-5107(64)90047-8

Neaves, W. B. (1969). Adenosine deaminase phenotypes among sexual and parthenogenetic lizards in the genus Cnemidophorus (teiidae). *Journal of Experimental Zoology*, *171*(2), 175–183. https://doi.org/10.1002/jez.1401710205

Neaves, W. B., & Baumann, P. (2011). Unisexual reproduction among vertebrates. *Trends in Genetics*, *27*(3), 81–88. https://doi.org/10.1016/j.tig.2010.12.002

Neaves, W. B., & Gerald, P. S. (1969). Gene dosage at the lactate dehydrogenase b locus in triploid and diploid teiid lizards. *Science (New York, N.Y.)*, *164*(3879), 557–559. https://doi.org/10.1126/science.164.3879.557

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

Newton, A. A., Schnittker, R. R., Yu, Z., Munday, S. S., Baumann, D. P., Neaves, W. B., & Baumann, P. (2016). Widespread failure to complete meiosis does not impair fecundity in parthenogenetic whiptail lizards. *Development*, *143*(23), 4486–4494. https://doi.org/10.1242/dev.141283

Nicolas, S. D., Monod, H., Eber, F., Chèvre, A.-M., & Jenczewski, E. (2012). Non-random distribution of extensive chromosome rearrangements in Brassica napus depends on genome organization. *The Plant Journal*, *70*(4), 691–703. https://doi.org/10.1111/j.1365-313X.2012.04914.x

Normark, B. B. (2003). The evolution of alternative genetic systems in insects. *Annual Review of Entomology*, *48*, 397–423. https://doi.org/10.1146/annurev.ento.48.091801.112703

Olmo, E. (2005). Rate of Chromosome changes and Speciation in Reptiles. *Genetica*, *125*(2), 185–203. https://doi.org/10.1007/s10709-005-8008-2

Olsen, M. W. (1973). Longevity and Organ Weights of Beltsville Small White Parthenogens and Normal Turkey Males. *Poultry Science*, *52*(2), 666–670. https://doi.org/10.3382/ps.0520666

Olsen, W. W., & Marsden, S. J. (1954). Natural parthenogenesis in turkey eggs. *Science (New York, N.Y.)*, *120*(3118), 545–546. https://doi.org/10.1126/science.120.3118.545

Paul D.N. Hubert, Alina Cywinska, Shelley L. Ball, & Jeremy R. deWaard. (2003). *Biological identifications through DNA barcodes*. *270*(1512). https://royalsocietypublishing.org/doi/10.1098/rspb.2002.2218

Pennell, M. W., Kirkpatrick, M., Otto, S. P., Vamosi, J. C., Peichel, C. L., Valenzuela, N., & Kitano, J. (2015). Y Fuse? Sex Chromosome Fusions in Fishes and Reptiles. *PLoS Genetics*, *11*(5), e1005237. https://doi.org/10.1371/journal.pgen.1005237

Pomiankowski, A. (1987). The costs of choice in sexual selection. *Journal of Theoretical Biology*, *128*(2), 195–218. https://doi.org/10.1016/s0022-5193(87)80169-8

Potter, S., Bragg, J. G., Blom, M. P. K., Deakin, J. E., Kirkpatrick, M., Eldridge, M. D. B., & Moritz, C. (2017). Chromosomal Speciation in the Genomics Era: Disentangling Phylogenetic Evolution of Rock-wallabies. *Frontiers in Genetics*, *8*. https://doi.org/10.3389/fgene.2017.00010

Putnam, N. H., O'Connell, B. L., Stites, J. C., Rice, B. J., Blanchette, M., Calef, R., Troll, C. J., Fields, A., Hartley, P. D., Sugnet, C. W., Haussler, D., Rokhsar, D. S., & Green, R. E. (2016). Chromosome-scale shotgun assembly using an in vitro method for long-range linkage. *Genome Research*, *26*(3), 342–350. https://doi.org/10.1101/gr.193474.115

Rachel, A. J., Sharma, T., & Menon, V. V. (1991). Harlequin banding and localisation of sister-chromatid exchanges. *Mutation Research Letters*, *264*(2), 71–80. https://doi.org/10.1016/0165-7992(91)90048-9

Reeder, T., Dessauer, H. C., & Cole, C. J. (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 ; no. 3365*. https://digitallibrary.amnh.org/handle/2246/2854

Roach, M. J., Schmidt, S. A., & Borneman, A. R. (2018). Purge Haplotigs: Allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics*, *19*(1), 460. https://doi.org/10.1186/s12859-018-2485-7

Ryder, O. A., Thomas, S., Judson, J. M., Romanov, M. N., Dandekar, S., Papp, J. C., Sidak-Loftis, L. C., Walker, K., Stalis, I. H., Mace, M., Steiner, C. C., & Chemnick, L. G. (2021). Facultative Parthenogenesis in California Condors. *The Journal of Heredity*, *112*(7), 569–574. https://doi.org/10.1093/jhered/esab052

Saccucci, M. J., Denton, R. D., Holding, M. L., & Gibbs, H. L. (2016). Polyploid unisexual salamanders have higher tissue regeneration rates than diploid sexual relatives. *Journal of Zoology*, *300*(2), 77–81. https://doi.org/10.1111/jzo.12339

Sacerdot, C., Louis, A., Bon, C., Berthelot, C., & Roest Crollius, H. (2018). Chromosome evolution at the origin of the ancestral vertebrate genome. *Genome Biology*, *19*(1), 166. https://doi.org/10.1186/s13059-018-1559-1

Schalch, T., & Steiner, F. A. (2017). Structure of centromere chromatin: From nucleosome to chromosomal architecture. *Chromosoma*, *126*(4), 443–455. https://doi.org/10.1007/s00412-016-0620-7

Sefer, E. (2022). A comparison of topologically associating domain callers over mammals at high resolution. *BMC Bioinformatics*, *23*(1), 127. https://doi.org/10.1186/s12859-022-04674-2

Servant, N., Varoquaux, N., Lajoie, B. R., Viara, E., Chen, C.-J., Vert, J.-P., Heard, E., Dekker, J., & Barillot, E. (2015). HiC-Pro: An optimized and flexible pipeline for Hi-C data processing. *Genome Biology*, *16*(1), 259. https://doi.org/10.1186/s13059-015-0831-x

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Simon, J.-C., Delmotte, F., Rispe, C., & Crease, T. (2003). Phylogenetic relationships between parthenogens and their sexual relatives: The possible routes to parthenogenesis in animals. *Biological Journal of the Linnean Society*, *79*(1), 151–163. https://doi.org/10.1046/j.1095-8312.2003.00175.x

Sinclair, E. A., Pramuk, J. B., Bezy, R. L., Crandall, K. A., & Sites, J. W. (2010). DNA evidence for nonhybrid origins of parthenogenesis in natural populations of vertebrates. *Evolution; International Journal of Organic Evolution*, *64*(5), 1346–1357. https://doi.org/10.1111/j.1558-5646.2009.00893.x

Sperling, A. L., Fabian, D. K., Garrison, E., & Glover, D. M. (2023). A genetic basis for facultative parthenogenesis in Drosophila. *Current Biology*, *33*(17), 3545-3560.e13. https://doi.org/10.1016/j.cub.2023.07.006

Stapley, J., Feulner, P. G. D., Johnston, S. E., Santure, A. W., & Smadja, C. M. (2017). Recombination: The good, the bad and the variable. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1736), 20170279. https://doi.org/10.1098/rstb.2017.0279

Swart, E. C., Bracht, J. R., Magrini, V., Minx, P., Chen, X., Zhou, Y., Khurana, J. S., Goldman, A. D., Nowacki, M., Schotanus, K., Jung, S., Fulton, R. S., Ly, A., McGrath, S., Haub, K., Wiggins, J. L., Storton, D., Matese, J. C., Parsons, L., … Landweber, L. F. (2013). The Oxytricha trifallax Macronuclear Genome: A Complex Eukaryotic Genome with 16,000 Tiny Chromosomes. *PLoS Biology*, *11*(1), e1001473. https://doi.org/10.1371/journal.pbio.1001473

Van Valen, L. (1973). A new evolutionary law. *Evol Theory*, *1*, 1–30.

Vrijenhoek, R. C., Angus, R. A., & Schultz, R. J. (1977). Variation and Heterozygosity in Sexually Vs. Clonally Reproducing Populations of Poeciliopsis. *Evolution*, *31*(4), 767–781. https://doi.org/10.1111/j.1558-5646.1977.tb01069.x

Vrijenhoek, R., Dawley, R., Cole, C. J., & Bogart, J. (1989). A list of known unisexual vertebrates. In *Evolution and Ecology of Unisexual Vertebrates* (pp. 19–23).

Waggoner, B. M., & Poinar, G. O. (1993). Fossil habrotrochid rotifers in Dominican amber. *Experientia*, *49*(4), 354–357. https://doi.org/10.1007/BF01923421

Walkowiak, S., Gao, L., Monat, C., Haberer, G., Kassa, M. T., Brinton, J., Ramirez-Gonzalez, R. H., Kolodziej, M. C., Delorean, E., Thambugala, D., Klymiuk, V., Byrns, B., Gundlach, H., Bandi, V., Siri, J. N., Nilsen, K., Aquino, C., Himmelbach, A., Copetti, D., … Pozniak, C. J. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature*, *588*(7837), 277–283. https://doi.org/10.1038/s41586-020-2961-x

Wang, B., Lv, R., Zhang, Z., Yang, C., Xun, H., Liu, B., & Gong, L. (2022). Homoeologous exchange enables rapid evolution of tolerance to salinity and hyper-osmotic stresses in a synthetic allotetraploid wheat. *Journal of Experimental Botany*, *73*(22), 7488–7502. https://doi.org/10.1093/jxb/erac355

Waters, P. D., Patel, H. R., Ruiz-Herrera, A., Álvarez-González, L., Lister, N. C., Simakov, O., Ezaz, T., Kaur, P., Frere, C., Grützner, F., Georges, A., & Graves, J. A. M. (2021). Microchromosomes are building blocks of bird, reptile, and mammal chromosomes. *Proceedings of the National Academy of Sciences*, *118*(45), e2112494118. https://doi.org/10.1073/pnas.2112494118

Watts, P. C., Buley, K. R., Sanderson, S., Boardman, W., Ciofi, C., & Gibson, R. (2006). Parthenogenesis in Komodo dragons. *Nature*, *444*(7122), Article 7122. https://doi.org/10.1038/4441021a

Wu, Y., Lin, F., Zhou, Y., Wang, J., Sun, S., Wang, B., Zhang, Z., Li, G., Lin, X., Wang, X., Sun, Y., Dong, Q., Xu, C., Gong, L., Wendel, J. F., Zhang, Z., & Liu, B. (2021). Genomic mosaicism due to homoeologous exchange generates extensive phenotypic diversity in nascent allopolyploids. *National Science Review*, *8*(5), nwaa277. https://doi.org/10.1093/nsr/nwaa277

Wynn, A. H., Cole, C. J., & Gardner, A. L. (1987). *Apparent triploidy in the unisexual brahminy blind snake, Ramphotyphlops braminus: Vol. no. 2868*. American Museum of Natural History. https://www.biodiversitylibrary.org/item/316647

Zhang, Z., Gou, X., Xun, H., Bian, Y., Ma, X., Li, J., Li, N., Gong, L., Feldman, M., Liu, B., & Levy, A. A. (2020). Homoeologous exchanges occur through intragenic recombination generating novel transcripts and proteins in wheat and other polyploids. *Proceedings of the National Academy of Sciences*, *117*(25), 14561–14571. https://doi.org/10.1073/pnas.2003505117

Curriculum vitae

David Vuong Ho
