

XX/XY Sex Chromosomes in the South American Dwarf Gecko (*Gonatodes humeralis*)

Tony Gamble*; Department of Biological Sciences, Marquette University, Milwaukee, WI, USA; Bell Museum of Natural History, University of Minnesota, Saint Paul, MN USA; Milwaukee Public Museum, Milwaukee, WI USA; e-mail: tgamble@geckoevolution.org

Erin McKenna; Castleton University, Castleton, VT, USA; e-mail: EMM01161@castleton.edu

Wyatt Meyer; Department of Biological Sciences, Marquette University, Milwaukee, WI, USA; e-mail: whm.soccer9@gmail.com

Stuart V. Nielsen; Department of Biological Sciences, Marquette University, Milwaukee, WI, USA; e-mail: stuart.nielsen@marquette.edu

Brendan J. Pinto; Department of Biological Sciences, Marquette University, Milwaukee, WI, USA; e-mail: brendan.pinto@marquette.edu

Daniel P. Scantlebury; Washington, D.C., USA; e-mail: dscantle@gmail.com

Timothy E. Higham; Department of Evolution, Ecology, and Organismal Biology, University of California, Riverside, Riverside, CA, USA; e-mail: thigham@ucr.edu

*Corresponding author: Tony Gamble; Department of Biological Sciences, Marquette University, Milwaukee, WI, USA; Bell Museum of Natural History, University of Minnesota, Saint Paul, MN USA; Milwaukee Public Museum, Milwaukee, WI USA; e-mail: tgamble@geckoevolution.org

© The American Genetic Association 2017. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Abstract

Sex-specific genetic markers identified using restriction site-associated DNA sequencing, or RADseq, permits the recognition of a species' sex chromosome system in cases where standard cytogenetic methods fail. Thus, species with male-specific RAD markers have an XX/XY sex chromosome system (male heterogamety) while species with female-specific RAD markers have a ZZ/ZW sex chromosome (female heterogamety). Here we use RADseq data from five male and five female South American dwarf geckos (*Gonatodes humeralis*) to identify an XX/XY sex chromosome system. This is the first confidently known sex chromosome system in a *Gonatodes* species. We used a low-coverage *de novo* *G. humeralis* genome assembly to design PCR primers to validate the male-specificity of a subset of the sex-specific RADseq markers and describe how even modest genome assemblies can facilitate the design of sex-specific PCR primers in species with diverse sex chromosome systems.

Keywords: genome, lizard, Neotropics, PCR-RFLP, RADseq, Squamata

There are an incredible variety of sex-determining mechanisms among multicellular organisms (Bachtrog *et al.*, 2014). However, in order to understand the evolutionary mechanisms that have produced this diversity, we must first identify and catalog the sex-determining systems of major plant and animal lineages. Historically this work involved cytogenetics to karyotype males and females and identify morphologically distinct or heteromorphic sex chromosomes (Valenzuela, Adams and Janzen, 2003). Species where males are the heterogametic sex have an XX/XY sex chromosome system, while species where females are the heterogametic sex have a ZZ/ZW system (Bull, 1983; Graves, 2008). Unfortunately many species lack heteromorphic sex chromosomes and instead have morphologically identical or homomorphic sex chromosomes (Bachtrog *et al.*, 2014; Ezaz, *et al.*, 2009; Gamble and Zarkower, 2014). Sex chromosomes in these species cannot be identified using traditional cytogenetic methods but can only be recognized via breeding experiments involving sex-reversed animals (Wallace, Badawy and Wallace, 1999), advanced cytogenetic techniques such as comparative genomic hybridization or fluorescent in situ hybridization (FISH) of repetitive elements (Ezaz, *et al.*, 2005; Gamble, *et al.*, 2014; Pokorná, Kratochvíl and Kejnovsky, 2011), copy number variation of sex-linked markers (Gamble, *et al.*, 2014; Rovatsos, *et al.*, 2014), or the identification of sex-specific genetic markers (Berset-Brandli, *et al.*, 2006; Devlin, Biagi and Smailus, 2001; Gamble and Zarkower, 2014). Recently, restriction site-associated DNA sequencing, or RADseq, has been used to generate tens of thousands of molecular genetic markers from multiple confidently sexed males and females to identify genetic markers found in one sex but not the other (Baxter, *et al.*, 2011; Gamble, 2016; Gamble, *et al.*, 2015; Gamble and Zarkower, 2014; Palaiokostas, *et al.*, 2013; Pan, *et al.*, 2016). A species' sex chromosome system, either XX/XY or ZZ/ZW, can be inferred via the identification of these sex-specific genetic markers. Species with an excess of male-specific RAD markers have an XX/XY system, with markers

occurring on the Y chromosome, and conversely an excess of female-specific RAD markers indicates a ZZ/ZW system, those markers occurring on the W chromosome (Gamble, 2016; Gamble and Zarkower, 2014).

Once identified, sex-specific RAD markers can be further validated via PCR (Gamble, *et al.*, 2015; Gamble and Zarkower, 2014). Restricted recombination around the sex-determining locus and linked sexually antagonistic alleles results in increasing sequence divergence between the X and Y (or Z and W) chromosomes over time (Bull, 1983; Charlesworth and Charlesworth, 2000; Graves, 2008). Thus, PCR primers designed for most sex-specific RAD markers should amplify only in the heterogametic sex, males in XX/XY systems and females in ZZ/ZW systems. However, some Y (or W) linked RAD sequences, particularly in young or newly evolved sex chromosomes, can have high sequence similarity to homologous regions on the X (or Z) and PCR primers that are intended to be sex-specific may instead amplify in both males and females. Thus, PCR validation will be an overly conservative test of sex-specificity in these species (Gamble, 2016). There are, however, two ways to overcome this problem.

The first involves situations where the restriction site itself is sex specific, that is, the restriction site occurs on the Y (or W) allele but is lacking on the X (or Z) allele, and the regions flanking the restriction site are identical or nearly so. In these instances, the bioinformatic analyses will identify these RAD markers as sex-specific but subsequent PCR validation would fail because primers would produce amplicons in both males and females due to the similarity between the X and Y (or Z and W) regions flanking the restriction site to which the primers bind (Fowler and Buonaccorsi, 2016; Gamble, 2016). One solution is to

align sex-specific reads to the species genome (or to the genome of a closely related species), pairing adjacent RAD markers and then design PCR primers that flank either side of the restriction site. PCR amplicons can then be restriction digested, a technique known as PCR-RFLP. The X (or Z) amplicon will remain unchanged while the Y (or W) amplicon will be cut, resulting in different banding patterns on a gel between males and females (Fig 1). This strategy has been effective in validating sex-specific RAD markers in both fish and snakes (Fowler and Buonaccorsi, 2016; Gamble, *et al.*, 2017).

The second strategy is to design PCR primers in particular regions of the Y (or W) allele that are distinct from the X (or Z) allele. However, as with the PCR-RFLP example, identifying these regions is only possible if additional genomic resources are available. Here we utilized an approach that used paired-end Illumina sequence reads from an individual of the homogametic sex (XX or ZZ individual) to produce a *de novo* genome assembly. We subsequently aligned sex-specific RAD contigs from the heterogametic sex (XX/XY or ZZ/ZW individuals) to these genomic contigs and designed PCR primers in regions that should amplify in a sex-specific manner, such as over insertions or deletions (indels) (Fig 1). We implemented this approach using the South American dwarf gecko (*Gonatodes humeralis*), a species we identified as having an XX/XY sex chromosome system but were unable to PCR validate using standard methods.

Dwarf geckos of the genus *Gonatodes* (Gekkota: Sphaerodactylidae) consist of 31 described species distributed across Central and South America and some Caribbean islands (Gamble, *et al.*, 2008; Schargel, *et al.*, 2017). Nine species in the genus, including our focal species *G. humeralis*, have published karyotypes but, with one exception, no heteromorphic sex

chromosomes have been observed (Schmid, *et al.*, 2014). That exception is *Gonatodes ceciliae* from Trinidad. McBee, Bickham and Dixon (1987) karyotyped two males, which both had a pair of heteromorphic chromosomes that were interpreted as sex chromosomes. However, both individuals had different numbers of chromosomes, $2N = 26$ and $2N = 22$. The subsequent karyotyping of an additional male *G. ceciliae* exhibited a third, unique karyotype, $2N = 24$ (Schmid, *et al.*, 2014). This individual also had a pair of heteromorphic chromosomes although they did not form a sex bivalent in meiosis (Schmid, *et al.*, 2014). Multiple, unique male karyotypes and lack of published female karyotypes suggests more evidence is needed before the *G. ceciliae* sex chromosome system can be confirmed. Thus, there are no *Gonatodes* species with a confidently known sex-determining system. Furthermore, the Sphaerodactylidae overall are poorly known with regards to sex determining systems although at least one transition between XX/XY and ZZ/ZW systems is presumed to have occurred; two *Sphaerodactylus* species and *Euleptes europeae* have XX/XY sex chromosomes, while *Aristelliger expectatus* has a ZZ/ZW system (Gamble, 2010; Gamble, *et al.*, 2015; Gornung, *et al.*, 2013; Schmid, *et al.*, 2014). Thus, sex chromosome data from additional sphaerodactylid species are needed to better understand sex chromosome evolution in this group. It is this need for sex chromosome data from additional species that motivated the current study.

Methods

We extracted genomic DNA using the Qiagen® DNeasy Blood and Tissue extraction kit from tail clips of five adult male and five adult female *G. humeralis* collected on Trinidad (Supplementary Table 1). RADseq libraries were constructed following a modified protocol from Etter *et al.* (2011) as described in Gamble, *et al.* (2015). Briefly, we digested genomic

DNA using high-fidelity *SbfI* restriction enzyme (New England Biolabs). We ligated individually barcoded P1 adapters onto the *SbfI* cut site for each sample and pooled samples into multiple libraries, sonicated, and size selected into 200- to 500-bp fragments using magnetic beads in a PEG/NaCl buffer (Rohland and Reich, 2012). We blunt-end repaired libraries and added a dA tail before ligating a P2 adapter containing unique Illumina barcodes to each of these pooled libraries. We used 16 PCR cycles with Phusion high-fidelity DNA polymerase (New England Biolabs) and size-selected a second time into 250 to 600 bp library fragments using magnetic beads in PEG/NaCl buffer. Libraries were pooled and sequenced using paired-end 125 bp reads on an Illumina HiSeq2500 at the Institute for Integrative Genome Biology, University of California – Riverside (Supplementary Table 1).

We demultiplexed, trimmed, and filtered raw Illumina reads using the process_radtags function in STACKS [v1.23] (Catchen, *et al.*, 2011). We used RADtools 1.2.4 (Baxter, *et al.*, 2011) to generate candidate alleles for each individual and candidate loci across all individuals from the forward reads with parameters as previously described (Gamble, *et al.*, 2017; Gamble, *et al.*, 2015). We used a python script (Gamble, *et al.*, 2015) to identify putative sex-specific markers from the RADtools output. This script also produces a second list of 'confirmed' sex-specific RAD markers, which are a subset of the initial list of sex-specific RAD markers, but excludes from further consideration any sex-specific markers that also appear in the original reads files from the opposite sex. We assembled forward and reverse reads from the confirmed sex-specific RAD markers into sex-specific RAD contigs using Geneious R9 (Kearse, *et al.*, 2012).

Given some large number of confidently sexed male and female samples, the preceding methods will correctly identify sex-specific RAD markers. However, with small sample sizes, as we have here, we cannot rule out that some number of sex-specific RAD markers may be identified by chance. To address this we permuted the sex among sampled individuals to create an expected null distribution of the number of sex-specific RAD markers when none are actually present (Gamble, *et al.*, 2017). We then compared the observed number of sex-specific markers to this null distribution to test whether it is larger than expected by chance or contained within the lower 95% confidence interval of the null distribution. We performed 100 permutations using parameters from our original dataset including the total number of RAD markers and the same number of males and females. Note that permutations were evaluated using the total number of sex-specific markers, not the number of confirmed sex-specific markers.

We attempted to use PCR to validate the sex specificity of a subset of confirmed male-specific RAD markers, as done in previous studies (Fowler and Buonaccorsi, 2016; Gamble, *et al.*, 2017; Gamble, *et al.*, 2015; Gamble and Zarkower, 2014; Hayashi, *et al.*, 2017), but initial attempts were unsuccessful and every PCR resulted in amplification in both male and female samples (data not shown). This can result when PCR primers are designed on parts of the RAD marker that exhibit high sequence similarity between the X and Y (or Z and W) (Fowler and Buonaccorsi, 2016; Gamble, 2016; Gamble, *et al.*, 2015). One solution to this problem is to align sex-specific RAD markers to the species' genome, if available, and then: 1) determine whether the restriction site is sex-specific and design primers that flank the restriction site for subsequent PCR RFLP (Fowler and Buonaccorsi, 2016); or 2) design primers on regions of the sex-specific RAD marker that exhibit considerable divergence

from the reference genome, such as indels, that could amplify in a sex-specific fashion.

While there are two published gecko genomes (Liu, *et al.*, 2015; Xiong, *et al.*, 2016), they are too phylogenetically distant from *G. humeralis* to be useful for our purposes. Therefore, we sequenced and assembled the genome of a female *G. humeralis* to which we aligned male-specific RAD markers, and used these additional data to design new sex-specific PCR primers. We made an Illumina library from sheared DNA of a female *G. humeralis* (TG2241) from Trinidad using the NEBNext Ultra DNA Library Prep Kit with a 350bp insert size. We sequenced 634 million paired-end 150bp reads on an Illumina NextSeq at the Institute for Integrative Genome Biology, University of California – Riverside. Raw Illumina reads were trimmed and filtered using the process_shortreads function in STACKS [v1.23](Catchen, *et al.*, 2011). We *de novo* assembled reads into contigs using CLC Genomics Workbench [v10.1.1]. We estimated genome size and sequencing coverage from k-mer frequencies (Li, *et al.*, 2010). 31 bp k-mers were counted from cleaned, paired-end Illumina reads with Jellyfish [v2.2.6] (Marçais and Kingsford, 2011). We used perl scripts (https://github.com/josephryan/estimate_genome_size.pl) to generate a histogram of k-mer frequencies from the Jellyfish output, identify the peak k-mer frequency, and subsequently estimate genome size and sequence coverage.

We used BLASTn to align male-specific RAD contigs to the female genome assembly and designed primers using Geneious R9 (Kearse, *et al.*, 2012). Primers were designed to amplify in males and not females by placing one primer onto an aligned region with a sex-specific indel (Fig 1). We used PCR of six males and five females (Supplementary Table 1) to validate sex-specificity and visualized the PCR amplicons using gel electrophoresis.

Results

We recovered 151,073 RAD markers, which included 156 male-specific and two female-specific markers. A subset of these RAD markers were excluded from further consideration because they occurred in the original reads files of the opposite sex. The RAD markers that remained are called “confirmed” sex-specific markers. We identified 25 confirmed male-specific RAD markers and two confirmed female-specific RAD markers. This excess of male-specific RAD markers indicates an XX/XY sex chromosome system. Permutations showed the observed number of male-specific markers (156) was greater than the lower 95% confidence interval of the null distribution (Fig. 2), whereas the observation of female-specific markers was contained within the lower 95% null distribution and considered an artifact of our small sample size.

The female genome was sequenced to about 36x from 629 million trimmed and cleaned Illumina reads. The *de novo* genome assembly resulted in 1.42 million contigs with an N50 of 2879 bp, 43.7% GC content, and estimated genome size of 1.3 Gb.

As mentioned previously, PCR amplicons from primers designed from male-specific RAD markers amplified in both sexes (not shown). Because PCR is likely to be an overly conservative means of validating sex-specificity (Gamble, 2016) we designed new PCR primers that could accurately validate male-specificity after aligning male-specific RAD contigs to the *de novo* female genome assembly using BLASTn. All 25 of the confirmed male-specific contigs had BLAST hits to the female genome with eight of the markers mapping in matched pairs to four separate genome contigs. The four female contigs that corresponded

to these eight RAD markers all had an *SbfI* restriction site and we subsequently found the corresponding RAD markers, which are presumably X-linked, in our RAD dataset. Thus, PCR-RFLP was not possible for these eight RAD markers. However, two of these alignments had indel differences between putative X and Y alleles that allowed us to design primers that produced male-specific PCR amplicons in the same individuals for both primer sets (Fig 3), confirming an XX/XY sex chromosome system. The following primer pairs amplified in a male-specific manner: Ghum_1and9-F 5'-GAGCAGATGATTGGGGCTGAT-3' and Ghum_1and9-R 5'-ATACCTTGGGTGAGACCAGGA-3'; Ghum_23and28-F1 5'-CGCAGCAAGGTTCCCTTGTACA-3' and Ghum_23and28-R1 5'-ACCGATAGTGTAAACCTTGCTTG-3'.

Discussion

These results provide the first robust evidence of sex chromosomes in the genus *Gonatodes* and increase our limited knowledge concerning the phylogenetic distribution of sex chromosomes in the gecko family Sphaerodactylidae (Fig 4). While it is clear that at least one transition between XX/XY and ZZ/ZW systems has occurred in Sphaerodactylidae, additional work is needed to assess whether the XX/XY systems in *Sphaerodactylus*, *Gonatodes*, and *Euleptes* are derived from the same ancestral XY system and thus homologous with each other. A shared XX/XY system among these genera would be quite ancient, as their most recent common ancestor is at least 95 mya (Gamble, *et al.*, 2015). Independently derived XX/XY systems would further increase the number of transitions in this clade making them an ideal model to study sex chromosome transitions and evolution. The strong sequence similarity between X and Y alleles coupled with the relatively few confirmed male-specific markers implies limited differentiation between the *G. humeralis* X

and Y and suggests that these sex chromosomes are relatively young (Charlesworth and Charlesworth, 2000; Rice, 1996). This scenario would support the independent origin of XX/XY sex chromosomes among the Sphaerodactylidae. However, poorly differentiated sex chromosomes are not necessarily newly evolved and a variety of processes may prevent Y degeneration in older sex chromosomes, e.g. occasional recombination between the X and Y; weak natural selection to restrict recombination between the X and Y; or the inability to evolve dosage compensating mechanisms (Adolfsson and Ellegren, 2013; Bachtrog, *et al.*, 2014; Perrin, 2009; Rice, 1987; Stöck, *et al.*, 2011). Determining whether XX/XY Sphaerodactylidae share a common sex chromosome system would resolve the age of sex chromosomes in the family, as could the identification of sex chromosomes in additional species and genera. Methods such as RADseq can be used to increase the number of sphaerodactylid species with known sex chromosome systems and targeting additional sphaerodactylid genera should be a high priority.

Squamate genomes range in size from 1.05 to 3.93 pg (1 pg of DNA = 978 Mb; (DeSmet, 1981; Doležel, *et al.*, 2003; Gregory, 2017). Thus, the estimated genome size of *G. humeralis* (1.3 Gb) is among the smallest of all squamate genomes and smaller than recent genome size measurements for other *Gonatodes* species (1.6 pg to an astonishing 7.8 pg; Schmid, *et al.*, 2014). Previously published *Gonatodes* genome size estimates should be interpreted with caution, however, as they were generated with flow cytometry using DAPI (4',6-diamidino-2-phenylindole) staining instead of the more commonly used propidium iodide (PI) or ethidium bromide (EB) staining (Schmid, *et al.*, 2014). DAPI preferentially binds AT-rich sites and can result in significantly different genome size estimates than estimates made from PI or EB, which lack such biases (Doležel, Sgorbati and Lucretti, 1992; Johnston,

et al., 1999; Kapuscinski, 1995). K-mer-based genome size estimates, as presented here, should also be interpreted cautiously as they can be biased by a variety of issues including sequencing errors, high heterozygosity, uneven sequence coverage, the overrepresentation of organelle sequences, e.g. mitochondrial sequence, and repetitive elements (Simpson, 2014; Sun, *et al.*, In press; Vurture, *et al.*, 2017; Williams, *et al.*, 2013). Improving the *G. humeralis* genome assembly through additional sequencing and scaffolding will facilitate more accurate genome size estimates.

The draft *G. humeralis* genome assembly, while modest, proved indispensable for designing PCR primers to validate the male-specificity of the RAD markers identified here. RADseq has been promoted as an efficient, cost-effective way to identify sex-specific genetic markers in species lacking genomic resources (Gamble, 2016; Gamble, *et al.*, 2015; Gamble and Zarkower, 2014). In a subset of cases though, PCR primers designed from the assembled RAD contigs will fail to amplify in a sex-specific manner (Fowler and Buonaccorsi, 2016; Gamble, *et al.*, 2015). Adding additional genomic resources can resolve these problems if initial attempts to perform PCR validation are unsuccessful. Similar efforts to design and validate sex-specific RAD markers in other taxa could see benefits from sequencing low coverage genomes to facilitate primer development. While additional uses for modest genome assemblies like this are limited, we are currently generating more data that will improve this initial assembly and elucidate which chromosomes are the *G. humeralis* sex chromosomes.

Sex-specific genetic markers have proven useful in identifying the sex chromosome systems in many plant and animal species (Charlesworth and Mank, 2010; Gamble, 2016; Gamble

and Zarkower, 2014). However, these markers have many other uses, including identifying an individual's sex in species lacking recognizable secondary sexual traits (Griffiths, *et al.*, 1998; Rovatsos, *et al.*, 2015), sexing embryonic material in developmental studies (Hacker, *et al.*, 1995; Smith, Smith and Sinclair, 1999), identifying sex-reversed individuals (Hayes, *et al.*, 2010; Holleley, *et al.*, 2015; Quinn, *et al.*, 2007), and the breeding and management of endangered species (Griffiths and Tiwari, 1995; Literman, *et al.*, 2017; Ortega, *et al.*, 2004; Robertson, *et al.*, 2006). Therefore, the identification of sex-specific markers and subsequent development of accurate PCR-based sex assays has important practical uses on top of their importance in studying sex chromosome evolution. The utilization of low-coverage whole genome sequencing to facilitate sex-specific PCR primer design should be considered in cases where the development and validation of PCR-based sex assays from RADseq data is problematic. Not only do the additional genomic resources enable the development of PCR primers in challenging cases, like *G. humeralis*, but the alignment of sex-specific RAD markers to a genome could enable primer design at many more independent loci than would otherwise be discovered. Having multiple PCR-based assays for a single species is useful as it can reduce sexing inaccuracies due to technical errors and other mistakes (Robertson and Gemmell, 2006). The combination of RADseq with whole genome sequencing appears to offer the greatest flexibility in identifying sex-specific markers and sex chromosomes for a wide range of uses.

Data Availability

Demultiplexed *G. humeralis* RADseq reads were deposited to NCBI SRA: SAMN06827899-06827908; cleaned paired-end Illumina reads used in whole genome sequencing were submitted to NCBI SRA: SAMN07634392 (Supplementary Table 1).

Funding

This work was supported by the U.S. National Science Foundation NSF-IOS1146820 [to D. Zarkower]; NSF-DBI1659595 [to A. Abbott and E. Blumenthal]; Marquette University startup funds [to TG]; and UC Riverside startup funds [to TEH].

Acknowledgements

The authors thank E. Charles and D. Zarkower for help in the field and A. Abbott, E. Blumenthal, and other Marquette University Summer REU participants for help, kindness, and inspiration. Two anonymous reviewers provided useful comments that improved the quality of the manuscript. All experiments were carried out in accordance with animal use protocols at University of Minnesota: 0810A50001 and 1108A03545. Samples were collected and exported under Trinidad and Tobago permit FW: 4/10/1.

References

ADOLFSSON S, ELLEGREN H. 2013 Lack of dosage compensation accompanies the arrested stage of sex chromosome evolution in ostriches. *Molecular Biology and Evolution*. 30(4):806–810.

BACHTROG D, MANK JE, PEICHEL CL, KIRKPATRICK M, OTTO SP, ASHMAN T-L, HAHN MW, KITANO J, MAYROSE I, MING R, PERRIN N, ROSS L, VALENZUELA N, VAMOSI JC, CONSORTIUM TTOS. 2014 Sex determination: Why so many ways of doing it? *PLoS Biology*. 12(7):e1001899.

BAXTER SW, DAVEY JW, JOHNSTON JS, SHELTON AM, HECKEL DG, JIGGINS CD, BLAXTER ML. 2011 Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. *PLoS ONE*. 6(4):e19315.

BERSET-BRANDLI L, JAQUIERY J, DUBEY S, PERRIN N. 2006 A sex-specific marker reveals male heterogamety in European tree frogs. *Molecular Biology and Evolution*. 23(6):1104-1106.

BULL JJ. 1983 *Evolution of Sex Determining Mechanisms*. Benjamin Cummings Publishing Company, Inc., Menlo Park, California.

CATCHEN JM, AMORES A, HOHENLOHE P, CRESKO W, POSTLETHWAIT JH. 2011 Stacks: Building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*. 1(3):171-182.

CHARLESWORTH D, CHARLESWORTH B. 2000 The degeneration of Y chromosomes. *Philosophical Transactions of the Royal Society of London B Biological Sciences*. 355:1563-1572.

CHARLESWORTH D, MANK JE. 2010 The birds and the bees and the flowers and the trees: Lessons from genetic mapping of sex determination in plants and animals. *Genetics*. 186(1):9-31.

DESMET W. 1981 The nuclear Feulgen-DNA content of the vertebrates (especially reptiles), as measured by fluorescence cytophotometry, with notes on the cell and the chromosome size. *Acta Zoologica et Pathologica Antverpiensia*. (76):119-167.

DEVLIN RH, BIAGI CA, SMAILUS DE. 2001 Genetic mapping of Y-chromosomal DNA markers in Pacific salmon. *Genetica*. 111(1-3):43-58.

DOLEŽEL J, BARTOS J, VOGLMAYR H, GREILHUBER J. 2003 Nuclear DNA content and genome size of trout and human. *Cytometry. Part A: The Journal of the International Society for Analytical Cytology*. 51(2):127–128.

DOLEŽEL J, SGORBATI S, LUCRETTI S. 1992 Comparison of three DNA fluorochromes for flow cytometric estimation of nuclear DNA content in plants. *Physiologia Plantarum*. 85(4):625–631.

EZAZ T, QUINN AE, MIURA I, SARRE SD, GEORGES A, GRAVES JAM. 2005 The dragon lizard *Pogona vitticeps* has ZZ/ZW micro-sex chromosomes. *Chromosome Research*. 13(8):763–776.

EZAZ T, SARRE SD, O'MEALLY D, GRAVES JAM, GEORGES A. 2009 Sex chromosome evolution in lizards: Independent origins and rapid transitions. *Cytogenetic and Genome Research*. 127(2-4):249–260.

FOWLER BL, BUONACCORSI VP. 2016 Genomic characterization of sex-identification markers in *Sebastes carnatus* and *S. chrysomelas* rockfishes. *Molecular Ecology*. 25:2165–2175.

GAMBLE T. 2010 A review of sex determining mechanisms in geckos (Gekkota: Squamata). *Sexual Development*. 4:88–103.

GAMBLE T. 2016 Using RAD-seq to recognize sex-specific markers and sex chromosome systems. *Molecular Ecology*. 25:2114–2116.

GAMBLE T, CASTOE TA, NIELSEN SV, BANKS JL, CARD DC, SCHIELD DR, SCHUETT GW, BOOTH W. 2017 The discovery of XY sex chromosomes in a boa and python. *Current Biology*. 27(14):2148–2153.

GAMBLE T, CORYELL J, EZAZ T, LYNCH J, SCANTLEBURY DP, ZARKOWER D. 2015

Restriction site-associated DNA sequencing (RAD-seq) reveals an extraordinary number of transitions among gecko sex-determining mechanisms. *Molecular Biology and Evolution*. 32:1296–1309.

GAMBLE T, GENEVA AJ, GLOR RE, ZARKOWER D. 2014 *Anolis* sex chromosomes are derived from a single ancestral pair. *Evolution*. 68:1027–1041.

GAMBLE T, GREENBAUM E, JACKMAN TR, BAUER AM. 2015 Into the light: Diurnality has evolved multiple times in geckos. *Biological Journal of the Linnean Society*. 115:896–910.

GAMBLE T, SIMONS AM, COLLI GR, VITT LJ. 2008 Tertiary climate change and the diversification of the Amazonian gecko genus *Gonatodes* (Sphaerodactylidae, Squamata). *Molecular Phylogenetics and Evolution*. 46(1):269–277.

GAMBLE T, ZARKOWER D. 2014 Identification of sex-specific molecular markers using restriction site associated DNA sequencing. *Molecular Ecology Resources*. 14:902–913.

GORNUNG E, MOSCONI F, ANNESI F, CASTIGLIA R. 2013 The first cytogenetic description of *Euleptes europaea* (Gené, 1839) from Northern Sardinia reveals the highest diploid chromosome number among sphaerodactylid geckos (Sphaerodactylidae, Squamata). *Comparative Cytogenetics*. 7(2):153–161.

GRAVES JAM. 2008 Weird animal genomes and the evolution of vertebrate sex and sex chromosomes. *Annual Review of Genetics*. 42:565–586.

GREGORY TR. 2017. Animal Genome Size Database [Internet]. Available from: URL <http://www.genomesize.com>.

GRAFFITHS R, DOUBLE MC, ORR K, DAWSON RJ. 1998 A DNA test to sex most birds. *Molecular Ecology*. 7(8):1071–1075.

GRAFFITHS R, TIWARI B. 1995 Sex of the last wild Spix's macaw. *Nature*. 375:454.

HACKER A, CAPEL B, GOODFELLOW P, LOVELL-BADGE R. 1995 Expression of *Sry*, the mouse sex determining gene. *Development*. 121(6):1603–1614.

HAYASHI Y, OGUCHI K, YAMAGUCHI K, KITADE O, MAEKAWA K, MIURA T, SHIGENOBU S. 2017 Male-specific molecular genetic markers in the Japanese subterranean termite *Reticulitermes speratus*. *Insectes Sociaux*. 64(3):1–8.

HAYES TB, KHOURY V, NARAYAN A, NAZIR M, PARK A, BROWN T, ADAME L, CHAN E, BUCHHOLZ D, STUEVE T, GALLIPEAU S. 2010 Atrazine induces complete feminization and chemical castration in male African clawed frogs (*Xenopus laevis*). *Proceedings of the National Academy of Sciences of the United States of America*. 107(10):4612–4617.

HOLLELEY CE, O'MEALLY D, SARRE SD, GRAVES JAM, EZAZ T, MATSUBARA K, AZAD B, ZHANG X, GEORGES A. 2015 Sex reversal triggers the rapid transition from genetic to temperature-dependent sex. *Nature*. 523(7558):79–82.

JOHNSTON JS, BENNETT MD, RAYBURN AL, GALBRAITH DW, PRICE HJ. 1999 Reference standards for determination of DNA content of plant nuclei. *American Journal of Botany*. 86(5):609–613.

KAPUSCINSKI J. 1995 DAPI: A DNA-specific fluorescent probe. *Biotechnic & Histochemistry*. 70(5):220–233.

KEARSE M, MOIR R, WILSON A, STONES-HAVAS S, CHEUNG M, STURROCK S, BUXTON S, COOPER A, MARKOWITZ S, DURAN C, THIERER T, ASHTON B, MEINTJES P, DRUMMOND AJ.

2012 Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 28(12):1647–1649.

LI R, FAN W, TIAN G, ZHU H, HE L, CAI J, HUANG Q, CAI Q, LI B, BAI Y, ZHANG Z, ZHANG Y, WANG W, LI J, WEI F, LI H, JIAN M, LI J, ZHANG Z, NIELSEN R, LI D, GU W, YANG Z, XUAN Z, RYDER OA, LEUNG FC-C, ZHOU Y, CAO J, SUN X, FU Y, FANG X, GUO X, WANG B, HOU R, SHEN F, MU B, NI P, LIN R, QIAN W, WANG G, YU C, NIE W, WANG J, WU Z, LIANG H, MIN J, WU Q, CHENG S, RUAN J, WANG M, SHI Z, WEN M, LIU B, REN X, ZHENG H, DONG D, COOK K, SHAN G, ZHANG H, KOSIOL C, XIE X, LU Z, ZHENG H, LI Y, STEINER CC, LAM TT-Y, LIN S, ZHANG Q, LI G, TIAN J, GONG T, LIU H, ZHANG D, FANG L, YE C, ZHANG J, HU W, XU A, REN Y, ZHANG G, BRUFORD MW, LI Q, MA L, GUO Y, AN N, HU Y, ZHENG Y, SHI Y, LI Z, LIU Q, CHEN Y, ZHAO J, QU N, ZHAO S, TIAN F, WANG X, WANG H, XU L, LIU X, VINAR T, WANG Y, LAM T-W, YIU S-M, LIU S, ZHANG H, LI D, HUANG Y, WANG X, YANG G, JIANG Z, WANG J, QIN N, LI L, LI J, BOLUND L, KRISTIANSEN K, WONG GK-S, OLSON M, ZHANG X, LI S, YANG H, WANG J, WANG J. 2010 The sequence and de novo assembly of the giant panda genome. *Nature*. 463(7279):311–317.

LITERMAN R, RADHAKRISHNAN S, TAMPLIN J, BURKE R, DRESSER C, VALENZUELA N. 2017 Development of sexing primers in *Glyptemys insculpta* and *Apalone spinifera* turtles uncovers an XX/XY sex-determining system in the critically-endangered bog turtle *Glyptemys muhlenbergii*. *Conservation Genetics Resources*. 9(4):1–8.

LIU Y, ZHOU Q, WANG Y, LUO L, YANG J, YANG L, LIU M, LI Y, QIAN T, ZHENG Y, LI M, LI J, GU Y, HAN Z, XU M, WANG Y, ZHU C, YU B, YANG Y, DING F, JIANG J, YANG H, GU X. 2015 *Gekko japonicus* genome reveals evolution of adhesive toe pads and tail regeneration. *Nature Communications*. 6(10033).

MARÇAIS G, KINGSFORD C. 2011 A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics*. 27(6):764–770.

MCBEE K, BICKHAM JW, DIXON JR. 1987 Male heterogamety and chromosomal variation in Caribbean geckos. *Journal of Herpetology*. 21(1):68–71.

ORTEGA J, FRANCO MDR, ADAMS BA, RALLS K, MALDONADO JE. 2004 A reliable, non-invasive method for sex determination in the endangered San Joaquin kit fox (*Vulpes macrotis mutica*) and other canids. *Conservation Genetics*. 5(5):715–718.

PALAIOKOSTAS C, BEKAERT M, DAVIE A, COWAN ME, ORAL M, TAGGART JB, GHARBI K, MCANDREW BJ, PENMAN DJ, MIGAUD H. 2013 Mapping the sex determination locus in the Atlantic halibut (*Hippoglossus hippoglossus*) using RAD sequencing. *BMC Genomics*. 14(1):566.

PAN Q, ANDERSON J, BERTHO S, HERPIN A, WILSON C, POSTLETHWAIT JH, SCHARTL M, GUIGUEN Y. 2016 Vertebrate sex-determining genes play musical chairs. *Comptes Rendus Biologies*. 339(7):258–262.

PERRIN N. 2009 Sex reversal: A fountain of youth for sex chromosomes? *Evolution*. 63(12):3043–3049.

POKORNÁ M, KRATOCHVÍL L, KEJNOVSKY E. 2011 Microsatellite distribution on sex chromosomes at different stages of heteromorphism and heterochromatinization in two lizard species (Squamata: Eublepharidae: *Coleonyx elegans* and Lacertidae: *Eremias velox*). *BMC Genetics*. 12(1):90–97.

QUINN AE, GEORGES A, SARRE SD, GUARINO F, EZAZ T, GRAVES JAM. 2007 Temperature sex reversal implies sex gene dosage in a reptile. *Science*. 316(5823):411.

RICE WR. 1987 The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution*. 41(4):911–914.

RICE WR. 1996 Evolution of the Y sex chromosome in animals. *Bioscience*. 46(5):331–343.

ROBERTSON BC, ELLIOTT GP, EASON DK, CLOUT MN, GEMMELL NJ. 2006 Sex allocation theory aids species conservation. *Biology Letters*. 2(2):229–231.

ROBERTSON BC, GEMMELL NJ. 2006 PCR-based sexing in conservation biology: Wrong answers from an accurate methodology? *Conservation Genetics*. 7(2):267–271.

ROHLAND N, REICH D. 2012 Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome research*. 22(5):939–946.

ROVATSOS M, ALTMANOVÁ M, POKORNÁ MJ, KRATOCHVÍL L. 2014 Novel X-linked genes revealed by quantitative polymerase chain reaction in the green anole, *Anolis carolinensis*. *G3: Genes, Genomes, Genetics*. 4(11):2107–2113.

ROVATSOS M, POKORNÁ MJ, ALTMANOVÁ M, KRATOCHVÍL L, VELENSKÝ P, VODIČKA R, REHÁK I. 2015 Sexing of Komodo Dragons, *Varanus komodoensis*. *Gazella*. 42:93–107.

SCHARGEL WE, RIVAS GA, GARCIA-PEREZ JE, RIVERO-BLANCO C, CHIPPINDALE PT, FUJITA MK. 2017 A new species of *Gonatodes* (Squamata: Sphaerodactylidae) from the western versant of the Cordillera de Merida, Venezuela. *Zootaxa*. 4291(3):549–562.

SCHMID M, STEINLEIN C, FEICHTINGER W, HAAF T, MIJARES-URRUTIA A, SCHARGEL WE, HEDGES SB. 2014 Cytogenetic studies on *Gonatodes* (Reptilia, Squamata, Sphaerodactylidae). *Cytogenetic and Genome Research*. 144(1):47–61.

SIMPSON JT. 2014 Exploring genome characteristics and sequence quality without a reference. *Bioinformatics*. 30(9):1228–1235.

SMITH CA, SMITH MJ, SINCLAIR AH. 1999 Gene expression during gonadogenesis in the chicken embryo. *Gene*. 234(2):395–402.

STÖCK M, HORN A, GROSSEN C, LINDTKE D, SERMIER R, BETTO-COLLIARD C, DUFRESNES C, BONJOUR E, DUMAS Z, LUQUET E, MADDALENA T, SOUSA HC, MARTINEZ-SOLANO I, PERRIN N. 2011 Ever-young sex chromosomes in European tree frogs. *PLoS Biology*. 9(5):e1001062.

SUN H, DING J, PIEDNOËL M, SCHNEEBERGER K. In press findGSE: estimating genome size variation within human and *Arabidopsis* using k-mer frequencies. *Bioinformatics*.

VALENZUELA N, ADAMS DC, JANZEN FJ. 2003 Pattern does not equal process: Exactly when is sex environmentally determined? *American Naturalist*. 161(4):676–683.

VURTURE GW, SEDLAZECK FJ, NATTESTAD M, UNDERWOOD CJ, FANG H, GURTOWSKI J, SCHATZ MC. 2017 GenomeScope: Fast reference-free genome profiling from short reads. *Bioinformatics*. 33:2202–2204.

WALLACE H, BADAWY G, WALLACE B. 1999 Amphibian sex determination and sex reversal. *Cellular and Molecular Life Sciences*. 55(6-7):901–909.

WILLIAMS D, TRIMBLE WL, SHILTS M, MEYER F, OCHMAN H. 2013 Rapid quantification of sequence repeats to resolve the size, structure and contents of bacterial genomes. *BMC Genomics*. 14(1):537.

XIONG Z, LI F, LI Q, ZHOU L, GAMBLE T, ZHENG J, KUI L, LI C, LI S, YANG H, ZHANG G. 2016 Draft genome of the leopard gecko, *Eublepharis macularius*. *GigaScience*. 5(1):47.

Accepted Manuscript

Figure Legends

Figure 1. Conceptual examples using genomic resources to validate sex-specific RAD markers via PCR and PCR-RFLP. Examples illustrate an XX/XY sex chromosome system but results are similar in species with a ZZ/ZW system. Male RAD contigs are presumed to be Y alleles while the genomic contig is presumed to be the X allele. Black vertical segments indicate Y-specific SNPs that do not occur on the X. Thin horizontal segments in the contigs represent insertions or deletions (indels) that differ between X and Y alleles. **(A)** Alignment of two adjacent, male-specific RAD contigs to the homologous region of the female genome. The restriction site is male-specific (as indicated by SNPs in the restriction site on both male RAD contigs). With no prior genomic resources, primer pair B would be designed. However, the sequence similarity between the X and Y alleles would result in PCR amplification in both sexes. In light of the genomic alignment, primer pair C can be developed and used for PCR-RFLP with cartoon scissors indicating where the Y allele would be cut via restriction digest. This is the approach developed by Fowler and Buonaccorsi (2016). **(B)** Gel electrophoresis of primer pair B showing PCR amplification in both sexes. **(C)** Gel electrophoresis of PCR amplicons using primer pair C after restriction digest (PCR-RFLP) showing different banding patterns between male and female samples. **(D)** Alignment of two adjacent, male-specific RAD contigs to the homologous region of the female genome. The restriction site is not male-specific but instead found in both sexes. With no further genomic resources, primer pair E would be designed. However, the sequence similarity between the X and Y alleles would lead to PCR amplification in both sexes. In light of the genomic alignment, primer pair F can be designed for the Y allele, which places one of the primers on top of a male-specific indel. This primer placement should restrict PCR

amplification to just the Y allele. This is the approach taken in this study. **(E)** Gel electrophoresis of primer pair E showing PCR amplification in both sexes. **(F)** Gel electrophoresis of primer pair F showing male-specific PCR amplification. See online version for full colors.

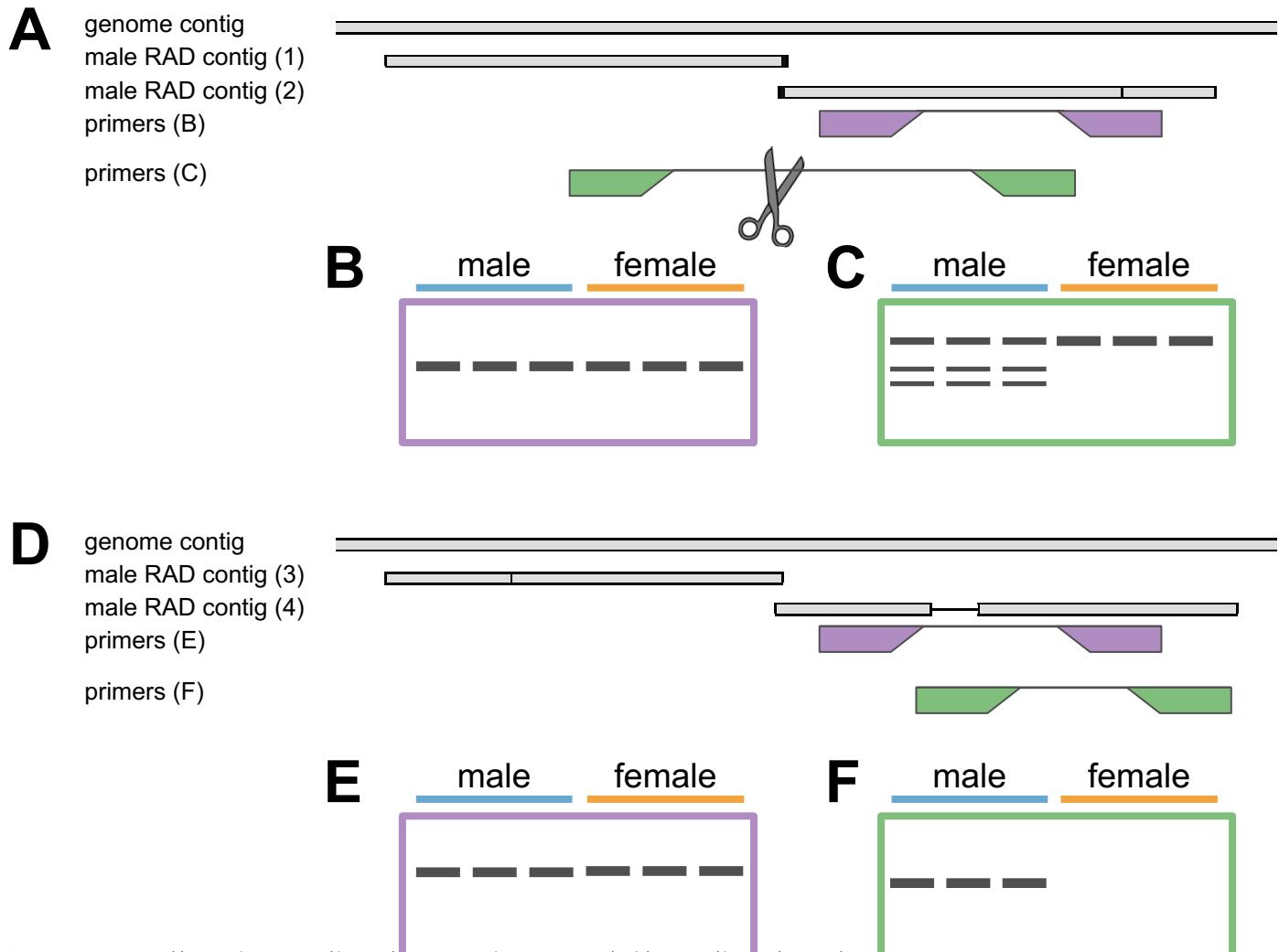
Figure 2. Permutations of the number of sex-specific markers expected by chance for the gecko *Gonatodes humeralis*. Blue and orange vertical lines denote the observed number of male- and female-specific RAD markers, respectively. The observed number of male-specific markers, 156, is greater than the upper 95% confidence interval of the null distribution (dashed vertical line) and is considered significantly different than expected by chance. See online version for full colors.

Figure 3. Primer design and PCR validation of male-specific RAD markers in the gecko *Gonatodes humeralis*. Both (A) and (B) show the alignment of two adjacent, male-specific RAD contigs to the homologous region of the female genome on the X chromosome. In light of this genomic information, one primer at each locus was designed on top of a male-specific indel. This primer placement should restrict PCR amplification to just the male-specific Y allele. Sequence alignments and PCR primers were implemented in Geneious R9 (Kearse, *et al.*, 2012). **(A)** Alignment of male-specific RAD markers 1 and 9, to design primers Ghum_1and9-F/Ghum_1and9-R. **(B)** Alignment of male-specific RAD markers 23 and 8, to design primers Ghum_23and8-F/Ghum_23and8-R. **(C)** Male-specific PCR amplification of primers Ghum_1and9-F/Ghum_1and9-R in six male and five female *G. humeralis*. **(D)** Male-specific PCR amplification of primers Ghum_23and8-F/Ghum_23and8-R in six male and five

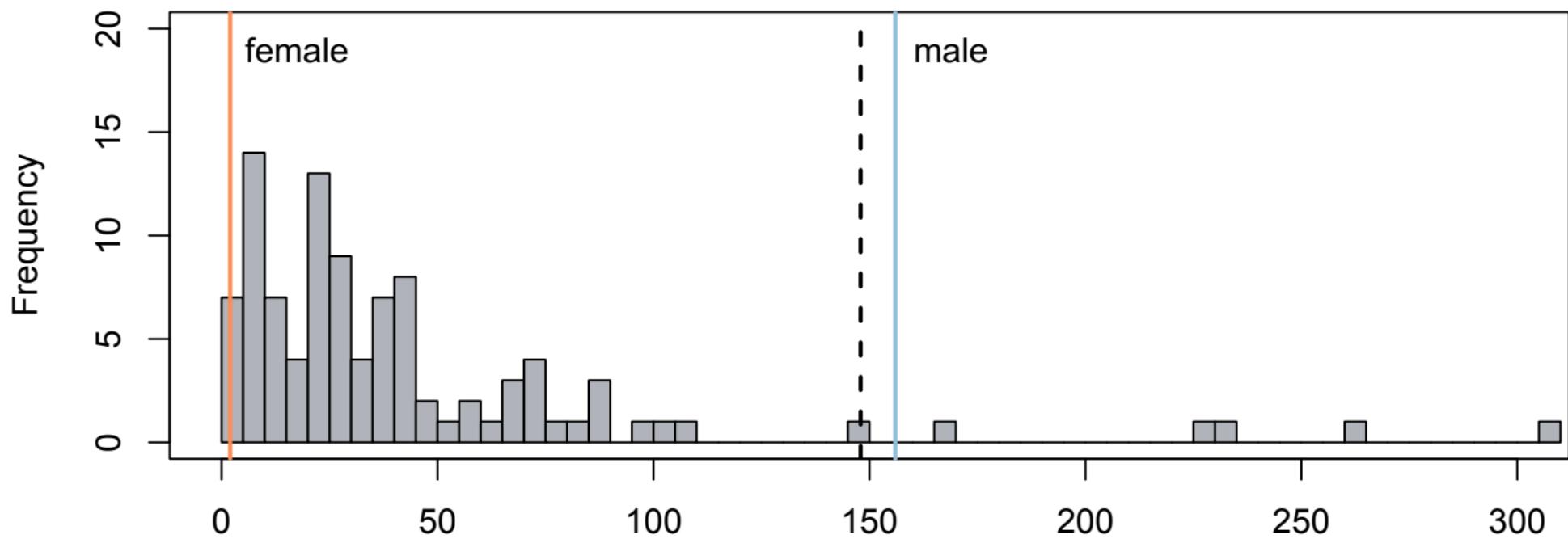
female *G. humeralis*. Specimen ID numbers are listed below every well for both gel images.

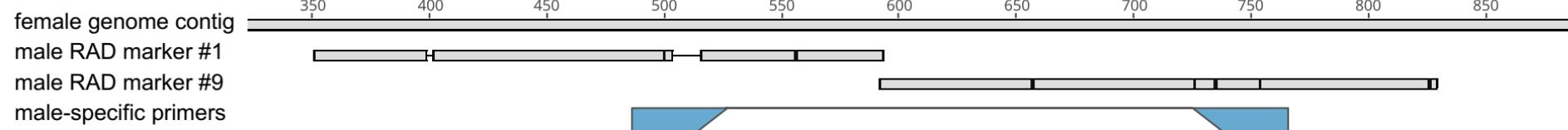
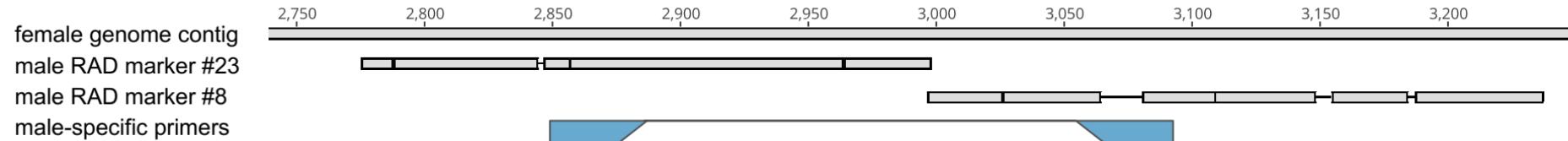
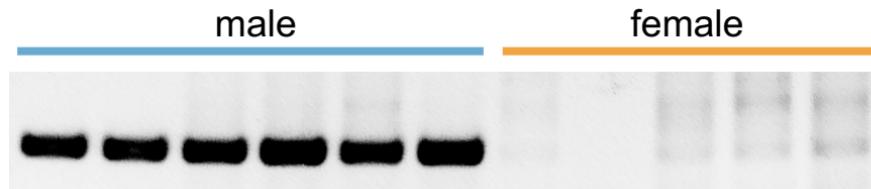
See online version for full colors.

Figure 4. Sex chromosome evolution in the gecko family Sphaerodactylidae. Time-calibrated phylogeny (Gamble, *et al.*, 2015) shows relationships among sphaerodactylid genera and sex chromosome systems, if known, are indicated by colored boxes to the left of generic names. An XX/XY sex chromosome system in the genus *Gonatodes* (in bold) is confirmed here for the first time. Image: Female *G. humeralis* from Trinidad (T. Gamble). Phylogenetic time, along the x-axis, is measured in millions of years. See online version for full colors.

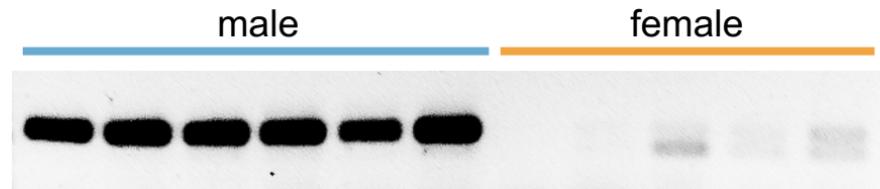


Gonatodes humeralis (XY) 5M,5F

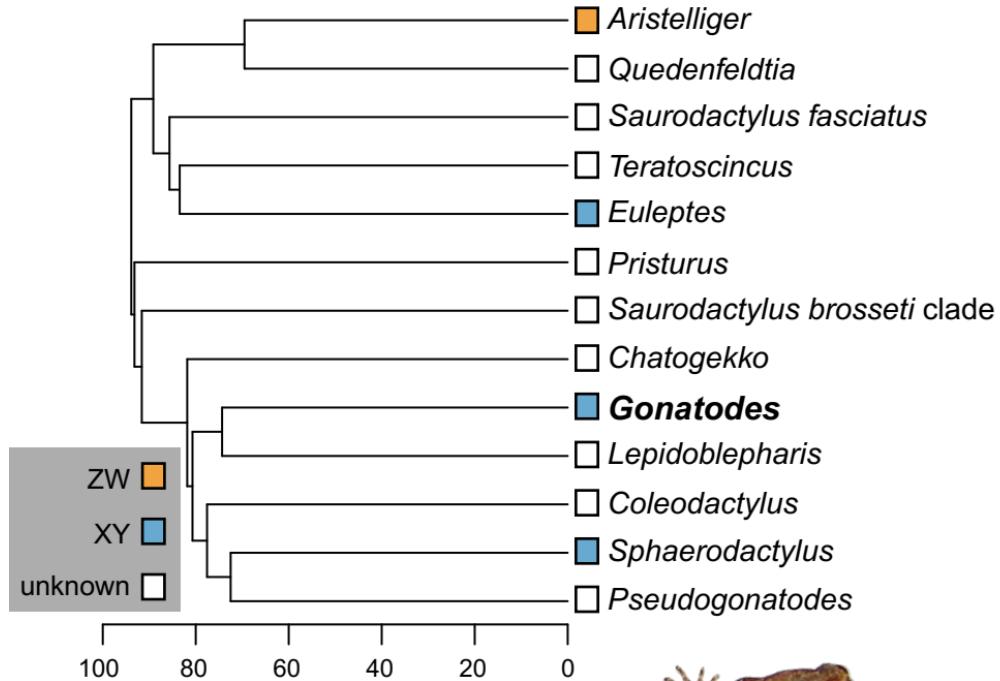


A**B****C**

T	G	T	G	T	G	T	G	T	G	T	G
1	1	1	1	1	1	1	1	1	1	1	1
8	8	8	8	6	7	8	8	7	8	8	8
9	6	4	4	8	3	9	6	2	5	5	5
7	1	2	3	1	0	6	2	9	0	1	1

D

T	G	T	G	T	G	T	G	T	G	T	G
1	1	1	1	1	1	1	1	1	1	1	1
8	8	8	8	8	8	8	8	6	7	8	8
9	6	4	4	4	4	4	4	8	3	9	6
7	1	2	3	1	0	6	2	3	1	0	6



100 80 60 40 20 0

