9-2014

Identification of Sex-specific Molecular Markers Using Restriction Site-associated DNA Sequencing

Tony Gamble
Marquette University, anthony.gamble@marquette.edu

David Zarkower
University of Minnesota - Minneapolis

Identification of Sex-specific Molecular Markers Using Restriction Site-associated DNA Sequencing

Tony Gamble
Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, MN
Bell Museum of Natural History, University of Minnesota, Minneapolis, MN
David Zarkower
Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, MN

Abstract

A major barrier to evolutionary studies of sex determination and sex chromosomes has been a lack of information on the types of sex-determining mechanisms that occur among different species. This is particularly problematic in groups where most species lack visually heteromorphic sex chromosomes, such as fish, amphibians and reptiles, because cytogenetic analyses will fail to identify the sex chromosomes in these species. We describe the use of restriction site-associated DNA (RAD) sequencing, or RAD-seq, to identify sex-specific molecular markers and subsequently determine whether a species has male or female heterogamety. To test the accuracy of this technique, we examined the lizard *Anolis carolinensis*. We performed RAD-seq on seven male and ten female *A. carolinensis* and found one male-specific molecular marker. *Anolis carolinensis* has previously been shown to possess male heterogamety and the recently published *A. carolinensis* genome

facilitated the characterization of the sex-specific RAD-seq marker. We validated the male specificity of the new marker using PCR on additional individuals and also found that it is conserved in some other *Anolis* species. We discuss the utility of using RAD-seq to identify sex-determining mechanisms in other species with cryptic or homomorphic sex chromosomes and the implications for the evolution of male heterogamety in *Anolis*.

Introduction

The genetic mechanisms that determine sex in animals are highly varied, and species commonly have sex determination systems with either XX/XY (male heterogamety), or ZZ/ZW (female heterogamety) (Gamble & Zarkower **2012**). Transitions between these mechanisms have occurred repeatedly across the metazoan phylogeny (Mank *et al*. **2006**; Kaiser & Bachtrog **2010**; O'Meally *et al*. **2012**). Understanding the genetic basis of sex determination and how transitions occur between sex-determining mechanisms is central to understanding this fundamental biological process and requires the identification of sex chromosomes in a number of related species. However, a major barrier to evolutionary studies of sex determination has been the challenge of identifying sex chromosomes and a consequent lack of information on the types of mechanisms that occur across taxa.

Identifying a species' sex chromosome system is typically done using one of three techniques (Bull **1980**; Charlesworth & Mank **2010**): cytogenetic approaches that visualize heteromorphic sex chromosomes (Valenzuela *et al*. **2003**); breeding experiments involving sex-reversed animals (Wallace *et al*. **1999**); or the identification of sex-specific molecular markers (Devlin *et al*. **2001**; Felip *et al*. **2005**). Each of these methods has associated challenges. Cytogenetics can be problematic in vertebrate groups such as fish, amphibians and reptiles where most species lack visually heteromorphic sex chromosomes (Hillis & Green **1990**; Hayes **1998**; Devlin & Nagahama **2002**; Ezaz *et al*. **2009**). In these situations, cytogenetic analyses will fail to identify male or female heterogamety. Similarly, many species are not amenable to being reared or bred in captivity, making the production of sex-reversed individuals difficult or impossible (Bull **1983**). This leaves the identification of sex-linked or sex-specific markers as the method holding the most promise as a general approach to identify sex chromosome systems in the widest variety of species.

Sex-specific markers are found on the heterogametic sex chromosome, the Y in species with male heterogamety or the W in species with female heterogamety. Therefore sex-specific markers can be used to determine whether a species has genetic sex determination (GSD) with either male or female heterogamety (Charlesworth & Mank **2010**). The presence of a male-specific marker indicates an XX/XY system, while the presence of a female-specific marker indicates a ZZ/ZW system. Sex-specific DNA markers have been used for several decades in biology and agriculture to ascertain the sex of individual animals (Taberlet *et al*. **1993**; Ellegren **1996**). Molecular sexing has been especially valuable in species that lack readily distinguishable sexually dimorphic phenotypes or for individuals at a developmental stage lacking secondary sexual characteristics. Mammalian Y-specific markers such as *Zfy* and *Sry* are routinely used to sex mice and other mammals via PCR (Taberlet *et al*. **1993**; Hacker *et al*. **1995**). Similarly, avian W-specific loci have been identified that amplify in females, but not males (Griffiths & Orr **1999**). The most commonly used means of identifying new sex-linked markers involves AFLPs (Griffiths & Orr **1999**; Griffiths *et al*. **2000**; Chen *et al*. **2007**; Quinn *et al*. **2009a**,**b**; Stöck *et al*. **2011**) or microsatellites (Lee *et al*. **2003**; Berset-Brändli *et al*. **2006**). More recently, restriction site-associated DNA sequencing, or RAD-seq, has been used to discover sex-linked markers and sex-determining regions of the genome (Baxter *et al*. **2011**; Anderson *et al*. **2012**; Carmichael *et al*. **2013**; Palaiokostas *et al*. **2013a**,**b**).

Restriction site-associated DNA sequencing sequences the DNA-flanking specific restriction sites throughout the genome, thereby facilitating the discovery of tens of thousands of genetic markers (Baird *et al*. **2008**). RAD-seq is a particularly powerful tool for exploring genetic variation in 'nonmodel' species because it does not require a

fully sequenced genome. As mentioned, RAD-seq has been used previously to identify sex-specific markers, but nearly all of these studies have discovered these markers through the construction of linkage maps from test crosses (Baxter *et al*. **2011**; Anderson *et al*. **2012**; Palaiokostas *et al*. **2013a**,**b**). Unfortunately, in many species, generating test crosses is not feasible as they do not readily breed in captivity, have very long generation times or have small numbers of offspring (Amores *et al*. **2011**; Ritland **2011**).

Here, we demonstrate the utility of RAD-seq for identifying sex-specific markers without linkage maps using the green anole lizard (*Anolis carolinensis*). We chose *A. carolinensis* as a test species because it is known to have an XX/XY sex-determining mechanism, which allows us to assess the accuracy of the RAD-seq method. Additionally, *A. carolinensis* has sex chromosomes that are both small, that is microchromosomes, and morphologically indistinguishable (Alföldi *et al*. **2011**) as is typical of sex chromosomes in many fish, amphibian and reptile species. Therefore, a method that can identify sex-specific markers in *A. carolinensis* should be applicable to many other vertebrate species. Finally, *A. carolinensis* has a sequenced genome and associated genomic resources (Alföldi *et al*. **2011**; Eckalbar *et al*. **2013**), providing a framework to troubleshoot potential problems with the analyses and to characterize any sex-specific markers identified.

We were able to develop a workflow, outlined below (Fig. **1**), to identify and validate sex-specific molecular markers by sampling adult individuals from natural populations without the need to perform experimental crosses. This involved using RAD-seq to identify putative sex-specific markers, followed by confirmation and validation of those markers. Finally, we characterized a newly discovered sex-specific marker from *A. carolinensis* to expand our understanding of *Anolis* sex chromosome evolution.
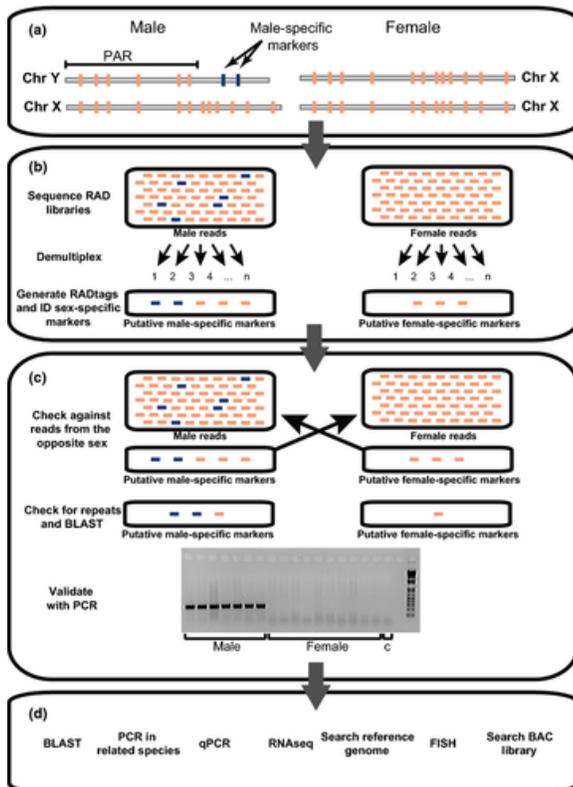
Figure 1 Outline of the proposed RAD-seq workflow. (a) Cartoon showing locations of restriction sites along the X and Y chromosomes in a male and female individuals. Restriction sites are similar between the X and Y chromosome on the pseudoautosomal region (PAR). Male-specific restriction sites on the Y chromosome are indicated and coloured blue. (b) Summary of the bioinformatic analysis of the RAD-seq libraries where Illumina reads are demultiplexed and loci and alleles identified. Putative sex-specific markers are identified although some, at this stage, may be false positives (illustrated here by the orange markers in the pools of putative sex-specific markers). (c) Confirmation and validation of putative sex-specific

markers through searches of the original read files, BLAST and PCR. (d) Characterizing sex-specific markers via a variety of methods.

Materials and methods

Restriction site-associated DNA sequencing

We constructed RAD-seq libraries following standard protocols (Etter *et al*. **2011**). Briefly, genomic DNA was extracted from tail tips of 7 male and 10 female field-collected adult *A. carolinensis* using the Qiagen DNeasy kit and digested with high-fidelity *SbfI* restriction enzyme (New England Biolabs). Individually barcoded P1 adapters (Table **1**, Table S1, Supporting Information) were ligated onto the *SbfI* cut site for each sample. One male was included twice, using two different barcodes, as an internal control for reproducibility, for a total of 18 samples (Table **1**). Samples were pooled by sex into separate male and female libraries and sonicated using a Fisher Scientific model 500 Ultrasonic Dismembrator. Libraries were manually size-selected into 200- to 500-bp fragments using gel electrophoresis. Libraries were blunt-end-repaired, and a 3' adenine overhang added to each fragment. We added a P2 adapter containing unique Illumina barcodes for separate male and female libraries. Libraries were amplified via 16 cycles of PCR and size-selected a second time into 250- to 550-bp fragments using gel electrophoresis. Complete adapter and barcode sequences are listed in Supplementary Table **1**. Samples were sequenced in 1/10th of a lane on an Illumina HiSeq2000 using 100-bp paired-end reads. Sequences are available at the NCBI Short Read Archive (PRJNA236043).

Table 1. Samples sequenced using RAD-seq and summary of RAD-seq analyses. MID sequence refers to multiplex identifier, the inline barcode sequence that identifies each individual in a RAD-seq library

Sample ID	Locality	Sex	MID sequence	Library ID	# of reads	# of RAD-tags	Mean # of reads/RAD-tag	Median # of reads/RAD-tag
tg1436	unknown	male	ATGTGTCGCCAA	I23	2 597 629	44 408	25.37	21
tg1609	Kona, Hawaii	male	CGACGATACTTG	I23	863 755	38 424	14.05	11
tg1423	unknown	male	CTAGATGCTGAC	I23	1 416 401	42 553	19.75	16
tg1537	Kona, Hawaii	male	GACACCGTATGT	I23	1 249 541	37 120	12.83	10
tg1578	Kona, Hawaii	male	AGAGT	I23	1 365 441	42 061	16.46	13
tg1578	Kona, Hawaii	male	GTACA	I23	1 086 041	41 353	16.21	13
tg1538	Kona, Hawaii	male	CAGTC	I23	1 061 778	41 211	14.82	12
tg1577	Kona, Hawaii	male	GTCAC	I23	843 086	37 503	12.6	10
tg1606	Kona, Hawaii	female	GATAGG	I27	488 277	28 533	10.51	8
tg1607	Kona, Hawaii	female	CAGAAG	I27	447 945	25 075	9.81	7
tg1608	Kona, Hawaii	female	TAATTG	I27	712 876	36 110	12.79	10
tg1424	unknown	female	ATGATG	I27	684 859	34 855	12.46	10
tg1437	unknown	female	AGAGT	I27	929 272	39 420	14.6	12
tg1536	Kona, Hawaii	female	CAGTC	I27	676 215	36 035	12.39	10
tg1540	Kona, Hawaii	female	GTCAC	I27	1 185 531	42 785	19.54	16

tg1541	Kona, Hawaii	female	TCTCT	I27	914 753	40 134	15.59	12
tg1579	Kona, Hawaii	female	ATGTGTCGCCAA	I27	495 010	23 524	9.74	7
tg1539	Kona, Hawaii	female	CGACGATACTTG	I27	703 947	37 549	13.74	11

Raw reads were demultiplexed using the process_radtags script from Stacks-0.9999 (Catchen *et al*. **2011**). We trimmed forward reads to 85 bp in length, removing low-quality bases at the 5' end of the reads and ensuring reads were all the same length for subsequent analyses. We generated candidate RAD-tags for each individual and candidate loci across all individuals from trimmed reads using RADtools version 1.2.4 (Baxter *et al*. **2011**). Settings for the RADtags script, which generates RAD-tags, included a cluster distance of ten, minimum quality score of 20, and read threshold, the minimum number of reads needed to form a RAD-tag, of five. Settings for the RADmarkers script, which uses the output from the RADtags script to generate candidate loci and alleles across all sampled individuals, included a tag count threshold of four and the maximum number of mismatches set at two.

Confirmation and validation of sex-specific markers

RADtools output includes the presence or absence of each locus and allele for every sampled individual, permitting the identification of putative sex-specific markers. The putative sex-specific markers identified with RADtools were screened against the raw read file of the opposite sex using a simple grep search looking for exact matches to the 85-bp trimmed RAD-tag. In other words, we grep-searched the female reads file for potentially male-specific markers and the male reads file for potentially female-specific markers. Any sex-specific markers with one or more matches in the raw reads of the opposite sex were excluded from subsequent analyses. This removed false positives that lacked a sufficient number of sequencing reads to create a RAD-tag in any individual from the library of the opposite sex but nonetheless had at least one sequence read in both sexes. All putative sex-specific markers that passed this confirmation step were assembled into contigs (RAD loci) with their paired-end reads using Sequencher5.0.1 (GenCodes). We further validated putative sex-linked RAD loci using three methods. First, we searched the female *A. carolinensis* genome (AnoCar2.0) for putative male-specific RAD loci using BLAT on the UCSC Genome Browser (Kent **2002**; Meyer *et al*. **2013**) and excluded any significant matches. Second, we used BLAST to search the NCBI nucleotide database for matches that might indicate whether the sequence was contamination or a fragment of a common repeat region. Third, we used PCR to validate sex-specific amplification of putative sex-linked RAD loci using primers (Table **2**) designed with Primer3 (Koressaar & Remm **2007**; Untergasser *et al*. **2012**). We used the following PCR profile: an initial 5-min denaturation at 94 °C followed by 32 cycles of denaturation (30 s at 94 °C), annealing (45 s at 55 °C) and extension (1 min at 72 °C), followed by a final extension of 5 min at 72 °C. PCR was performed using nine male samples and 14 female samples, which included a subset of the individuals used for RAD-seq as well as additional individuals (Supplementary Table **2**). This included additional individuals from a genetically distinct *A. carolinensis* population in west Florida (Campbell-Staton *et al*. **2012**; Tollis *et al*. **2012**). Including samples from this divergent population allowed us to infer that the most recent common ancestor of extant *A. carolinensis* populations possessed these male-specific markers.

Table 2. PCR primers used to amplify sex-specific markers in *Anolis*

Primer name	Primer sequence (5' to 3')
AcarB	
B-F1	GCAGGAAAGTATGGTGCCTTG
B-R1	GCTTCCACTCGATGACCTTGA
AcarC	
C-F1	AGGGAGAGAGATTGTGCCAAA

C-R1	ATGTCCCCTGGATCCTGATTT
AcarBC	
BC-F	CGTTCTGTGCCTATTACTGCTG
BC-R	CTCTCCTGGTCTTCCGTTAGTG
rtdr1y	
*rtdr1y*_F1	ACTAAAGGCACCCATTGCTTT
*rtdr1y*_R1	CATAAGCATGCTCTGGGTCAT

Characterization of sex-specific markers

Validated sex-specific RAD loci were subsequently examined with BLAST (Altschul *et al*. **1990**) against *Anolis* sequences in NCBI databases. It should be noted that a female *A. carolinensis* was sequenced, so no male markers exist in the assembled genome (Alföldi *et al*. **2011**). However, cDNAs from male tissues have been sequenced and deposited in GenBank (Eckalbar *et al*. **2013**). The genomic location of the top match from the wgs (whole-genome shotgun) database, a presumed X-linked gametolog to our male-specific marker, was uncertain, so we used qPCR of genomic DNA to determine whether the fragment occurs on an autosome, where we would expect an equal quantification between males and females, or the X chromosome, where males should have half the quantity as females. We designed two sets of primers from the contig identified by the BLAST searches of the validated sex-specific marker. The first marker, *rtdr1,* corresponded directly to the region of the BLAST match with the highest sequence similarity, while the other gene, *gnaz*, was approximately 7 kb upstream on the same contig. Other markers in the qPCR experiment included three autosomal genes: *kank1* (chromosome 2), *rag1* (chromosome 1) and *ngfb* (chromosome 4), and one X chromosome gene *pi4ka*. *ngfb* was used as the reference and *pi4ka* was a positive control as it is known to be X-linked in *A. carolinensis* (Gamble *et al*. **In Press**).

We conducted qPCR on genomic DNA from three males and three females (see Table S3, Supporting Information) using FastStart SYBR green (Roche) on an Eppendorf Realplex2 Mastercycler. Reactions were conducted in duplicate using 10 ng of genomic DNA in 12-µl reaction volumes. PCR primers are listed in Table S4 (Supporting Information). Cycle conditions involved an initial denaturation of 95 °C for 10 min followed by 40 cycles of: 95 °C for 20 s, 53 °C for 30 s, 72 °C for 40 s; and a final step to produce a melting curve going from 60 °C to 95 °C over 15 min. Data were analysed using the comparative quantification approach, which measures the relative amount of a gene in a group of male samples compared to a group of female samples, in rest 2009 software (Pfaffl *et al*. **2002**). Standard error and 95% confidence intervals of the normalized quantification values were calculated with 10,000 bootstrap replicates.

Genomic location of the BLAST match from *Anolis* sequences in the NCBI trace archive database was also uncertain, and we developed PCR primers (Table **2**) using Primer3 as described above. PCR was conducted in two male and two female *A. carolinensis* samples as well as one male and one female from nine additional *Anolis* species (Table S5, Supporting Information). We included *kank1* primers (Table S4, Supporting Information) in each reaction as a positive control. PCR using primers for the validated sex-specific RAD loci was also attempted in additional *Anolis* species to determine conservation and potential homology among *Anolis* Y chromosomes.

Genetic distances, measured as the number of base differences per site between sequences or p-distance, among aligned male-specific sequences and the related X chromosome gametolog were calculated in MEGA5 (Tamura *et al*. **2011**). We also compared homologous AcarB fragments from *Anolis sagrei* and *A. lineatopus*, which were Sanger-sequenced from PCR amplicons using primers B-F1 and B-R1. Ambiguous and missing positions were removed for each sequence pair.

The number of individuals needed to identify sex-specific markers

We explored the impact of sample size on the ability to identify sex-specific markers with RAD-seq by rerunning the analyses using a subset of individuals from the original analysis. We reran the RADtools analysis randomly choosing, in duplicate, four, five, six and seven males and females for each analysis (Table S6, Supporting Information). We subsequently screened putative sex-specific markers against a read file consisting of sampled individuals from the opposite sex using grep as described above.

Results

Restriction site-associated DNA sequencing

The male RAD-seq library had 11.6 million reads, and the female RAD-seq library had 8.3 million reads. An *in silico* restriction digest revealed 25,631 *SbfI* sites in the sequenced female *A. carolinensis* genome; thus, we would expect 51,262 RAD-tags (two RAD-tags flank each restriction site). We recovered slightly more RAD-tags than expected: 51,438 putative RAD-tags from the RADtools output (Supplementary File 1). The number of RAD-tags recovered in each individual varied from 23,524 to 44,408 and appeared strongly correlated with the number of reads recovered for each individual (Table **1**). Duplicate libraries from the male TG1578 shared 33,818 RAD-tags (81.6%). We identified 29 candidate RAD-tags with male-specific patterns, that is, RAD-tags occurring in at least seven of the eight male samples and no female samples. We identified two candidate RAD-tags with a female-specific pattern, occurring in at least nine of the ten female samples and no male samples. RADtools output is available via DRYAD (doi: **10.5061/dryad.pq608**).

Confirmation and validation of sex-specific markers

Removing putative sex-specific RAD-tags identified via RAD-seq that occurred in the original reads files from the opposite sex left us with four male-specific RAD-tags and one female-specific RAD-tag. BLAST searches of the putative female-specific RAD-tag revealed a partial match to a fosmid cloning vector (NCBI ID: JX069764) so this tag was not considered further as it was likely the result of contamination during preparation of the female RAD-seq library. Two of the four putative male-specific RAD loci had multiple BLAT hits in the *A. carolinensis* genome with 97.7–100% identity, precluding their further use as putative male-specific markers because the published genome is of a female individual (Alföldi *et al.* **2011**). We designed PCR primers for the assembled paired-end contigs for the two remaining male-specific RAD loci, called here AcarB and AcarC (Table **2**). PCR for both RAD loci produced amplicons from nine male samples, but not from the 14 female samples (Fig. **2**, Table S2, Supporting Information). We used additional primers to verify that the two male-specific RAD loci were on opposite sides of the same restriction site and constitute a single locus, which we call AcarBC (Fig. **2**). Concatenating the two assembled RAD loci yielded a 572-bp fragment (Supplementary File 2).
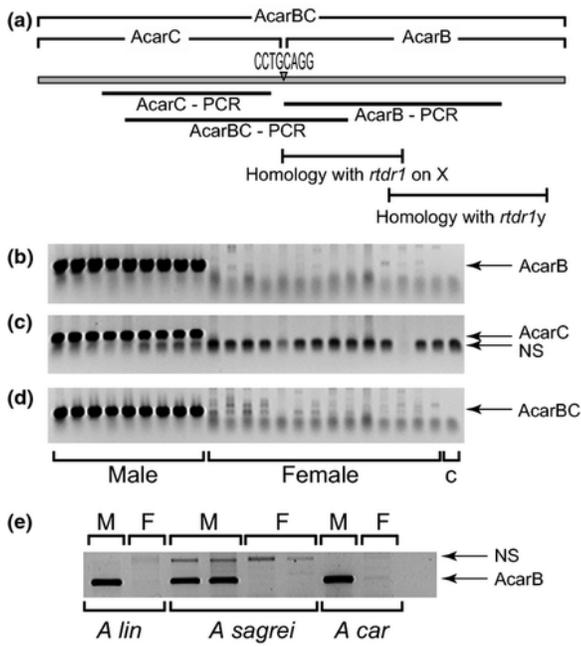
Figure 2 Validation of male-specific marker in *A. carolinensis*. (a) Physical map of the *A. carolinensis* male-specific marker AcarBC (572 bp) showing the locations of the *SbfI* restriction site and PCR amplicons for each male-specific marker. Regions that share sequence similarities with the X-linked *rtdr1* and *rtdr1y* are indicated. (b) Sex-specific amplification of marker AcarB-PCR (238 bp). (c) Sex-specific amplification of marker AcarC-PCR (184 bp). Bands labelled with 'NS' are nonspecific primer dimers. (d) Sex-specific amplification of marker AcarBC-PCR (243 bp) showing that markers AcarB and AcarC are on opposite sides of the same restriction site. Gel lane labelled with 'c' contained a negative control. (e) Sex-specific amplification of marker AcarB-PCR in *Anolis lineatopus* (*A lin*), *Anolis sagrei* and *Anolis carolinensis* (*A car*). 'M' and 'F' indicate male and female samples, respectively. Bands labelled with 'NS' are nonspecific PCR products.

Characterization of sex-specific markers

AcarBC had a partial BLAST match (BLASTN of *Anolis* sequences in the NCBI wgs database) to a gene, GAFZ01102862, transcribed from *A. carolinensis* liver. This fragment maps to the *rtdr1* gene on the contig chrUn_AAWZ02037698 in the *A. carolinensis* genome (Fig. **3**). This unmapped contig contains 3 genes: *gnaz*, *rtdr1* and *slc5a1*, all of which are on chromosome 15 in chicken, which is largely syntenic to the X chromosome in *A. carolinensis* (Alföldi *et al.* **2011**). qPCR of fragments of *rtdr1* showed differential abundance between males and females (Fig. **3**, Table S7, Supporting Information). *rtdr1* had quantification values in males that were half those of females, indicating that it is on the X chromosome, but not on the Y chromosome. The same pattern was seen for the X-linked control gene *pi4ka*. *gnaz*, by contrast, was present at similar levels in males and females, comparable to the two autosomal control genes *rag1* and *kank1*. Because *gnaz* and *rtdr1* are on the same contig in female *A. carolinensis, gnaz* must also be on the X chromosome. Unlike *rtdr1*, however, a gametolog of *gnaz* must also occur on the Y chromosome.
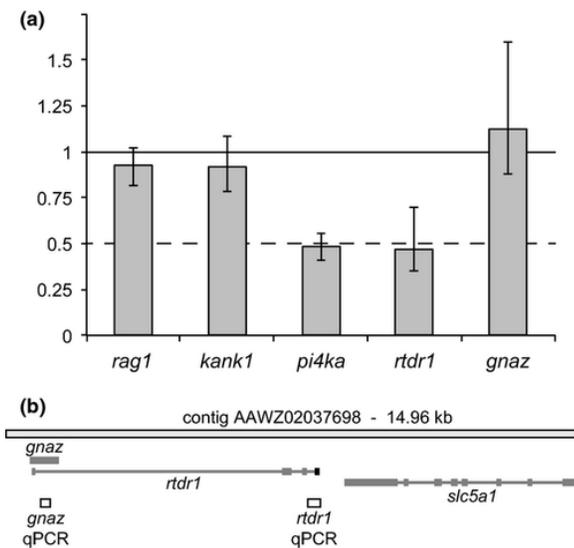
Figure 3 (a) Relative quantification and standard error of autosomal genes (*rag1*, *kank1*), a known X-linked gene (*pi4ka*) and two putative X-linked genes (*rtdr1*, *gnaz*) in male *Anolis* compared to females using qPCR of genomic DNA. A value of one (solid horizontal line) is equivalent to standard diploid copy number, for example single copy autosomal genes and X-linked genes in females, a value of 0.5 (dashed horizontal line) is expected of hemizygous loci, for example X-linked genes in males. The autosomal *ngfb* gene (not shown) was used as a standard. (b) Physical map of putative X-linked contig AAWZ02037698 showing the locations of three genes (grey boxes) on the contig and qPCR amplicons (white boxes) for *gnaz* and *rtdr1*. The *rtdr1* exon sharing homology with the *a. carolinensis* male-specific marker, AcarB, is in black.

AcarBC had another partial BLAST match (BLASTN of *Anolis* sequences in the NCBI Trace archive database) to a gene, ti:2071506473 and its cognate cDNA ti:2071506472, sequenced from *A. carolinensis* testis cDNA. BLAT of this gene to the *A. carolinensis* genome using the UCSC genome browser revealed no significant matches, suggesting that it may be Y-linked. BLAST to the NCBI nucleotide database revealed it to be an ortholog of *rtdr1* in other vertebrates, with some limited sequence similarity to the X-linked *rtdr1* in the *A. carolinensis* genome. Male-specific PCR amplification of a fragment of this gene occurred in *A. carolinensis*, *A. sagrei, A. planiceps, A. lineatus, A. grahami* and *A. lineatopus*, suggesting that it is confined to the Y chromosome in these species (Fig. **4**). This gene failed to amplify in *A. richardii*, *A. aeneus*, *A. chlorocyanus* and *A. distichus*. We refer to this gene as *rtdr1y* to avoid confusion with the X-linked *Anolis rtdr1* that occurs on the chrUn_AAWZ02037698 contig. Primers for the AcarB RAD locus also amplified in a sex-specific manner in *A. sagrei* and *A. lineatopus* (Fig. **2**) but failed to amplify in any other *Anolis* species.
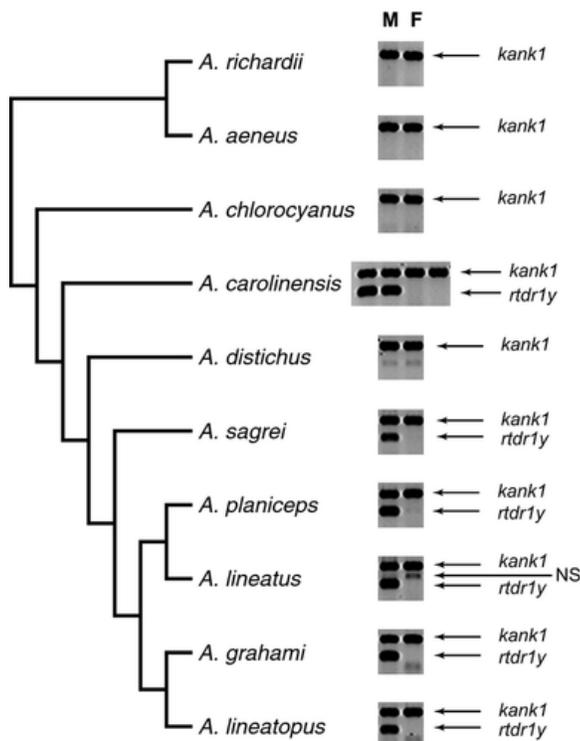
Figure 4 Phylogenetic relationships among sampled *Anolis* species (from Gamble *et al*. In Press) illustrating the sex-specific amplification of *rtdr1y*. The autosomal gene *kank1* was used as an internal positive control in all reactions. Bands labelled with 'NS' are nonspecific PCR products.

Genetic distances among AcarB (*rtdr1y*) sequences in *A. carolinensis*, *A. sagrei* and *A. lineatopus* ranged from 8.2–12.2% (Table 3). Genetic distances between AcarB (*rtdr1y*) sequences in *A. carolinensis*, *A. sagrei* and *A. lineatopus* and the X-linked *rtdr1* in *A. carolinensis* were substantially higher and ranged from 28.8–34.2%.

Table 3. Genetic distances (p-distance) among *A. carolinensis* X-linked *rtdr1* and male-specific gametologs in *A. carolinensis*, *A. sagrei* and *A. lineatopus*

	1	2	3
1. *A. carolinensis rtdr1* X chromosome	–	–	–
2. *A. carolinensis* Y chromosome BC	0.305	–	–
3. *A. sagrei* Y chromosome B	0.342	0.122	–
4. *A. lineaptopus* Y chromosome B	0.288	0.112	0.082

The number of individuals needed to identify sex-specific markers

Subsampling individuals from the original data set revealed the number of RAD-tags decreased as the number of sampled individuals decreased (Table 4, Supplementary File 3). Far more putative sex-specific RAD-tags were identified using the smaller data sets than with larger data sets. Similarly, the number of confirmed sex-specific RAD-tags, that is, sex-specific RAD-tags not found in the reads file of the opposite sex, was higher using fewer individuals than with larger data sets and the original data set (Fig. 5). However, despite the increase in false positives using smaller sample numbers, both of the validated male-specific RAD-tags from the complete data set, AcarB and AcarC, were identified as male specific in all of the subsampled analyses.

Table 4. Summary of RAD-seq analyses on the full data set and reanalyses of the data using a randomly sampled subset of individuals

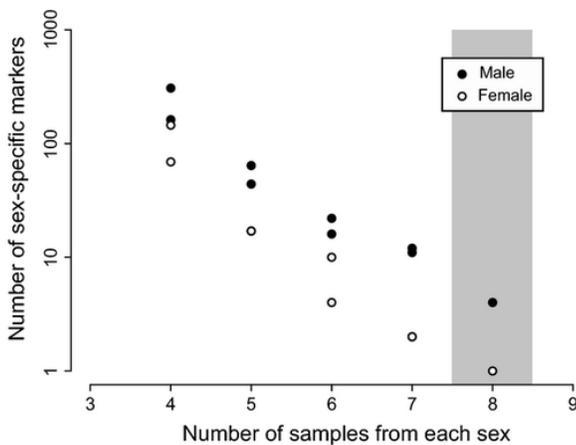| Analysis	# of male samples	# of female samples	Total # of RAD-tags	# of male only RAD-tags	# of female only RAD-tags	# of confirmed male only RAD-tags	# of confirmed female only RAD-tags	AcarB and AcarC identified?
Full data set	8	10	51 438	29	2	4	1	Yes
4a	4	4	48 525	1099	288	307	145	Yes
4b	4	4	47 751	1009	196	162	69	Yes
5a	5	5	49 074	439	45	64	17	Yes
5b	5	5	49 240	277	70	44	17	Yes
6a	6	6	49 352	154	14	22	4	Yes
6b	6	6	50 464	152	13	16	10	Yes
7a	7	7	50 676	109	3	12	2	Yes
7b	7	7	50 413	103	5	11	2	Yes

Figure 5 The number of confirmed sex-specific RAD-tags identified in replicate analyses using a randomly sampled subset of individuals from the original RAD-seq experiment. The number of individuals sampled of each sex is shown along the x-axis. Filled circles represent male samples, and open circles represent female samples. Data points for the original analyses are enclosed by the grey box, which used eight male samples and ten female samples. The y-axis is log-transformed for clarity.

Discussion

We used RAD-seq to identify a male-specific marker in the lizard *A. carolinensis* and subsequently validated the sex specificity of this marker using PCR. These results verify male heterogamety in *A. carolinensis*, provide a molecular 'beachhead' for further exploration of the Y chromosome in this species and underscore the utility of RAD-seq as a means of rapidly identifying sex chromosome systems in nonmodel species. This study is one of only a few to use RAD-seq to identify a sex-specific marker without generating linkage maps from test crosses. RAD-seq without linkage mapping was used to identify a sex-specific SNP in the salmon louse that was subsequently confirmed using PCR (Carmichael *et al*. **2013**). Another study looking at RAD-seq markers from guppies with male heterogamety found an excess of male-specific markers without generating linkage maps, although the authors did not validate the sex specificity of any of these markers (Willing *et al*. **2011**).

There are two advantages of identifying sex-specific markers via RAD-seq as compared to using microsatellites or AFLPs. One advantage is that the sequence data generated by RAD-seq allow for the rapid creation of PCR primers and subsequent validation of sex-specific markers. Another advantage is that if any one restriction enzyme fails to identify a sex-specific marker, it is a simple matter to switch to another enzyme that cuts more frequently in the genome. An infrequent cutter, like *SbfI* used here, may represent a good starting point in terms

of the number of markers generated and ability to obtain robust coverage for a large number of individuals analysed, but switching to a more frequent cutter may be necessary in some cases, for example in recently formed sex chromosome systems with relatively little sex chromosome divergence.

Using sex-specific sequences to identify sex chromosome systems has some advantages over cytogenetics. Cytogenetic methods will fail to identify sex chromosomes when a species lacks heteromorphic sex chromosomes, for example the majority of lizard, amphibian and fish species (Hillis & Green **1990**; Hayes **1998**; Devlin & Nagahama **2002**; Ezaz *et al*. **2009**). Advanced cytogenetic techniques, such as comparative genomic hybridization (CGH), can be useful in identifying homomorphic sex chromosomes in some cases (Traut *et al*. **2001**; Ezaz *et al*. **2005**). However, in contrast to RAD-seq, CGH does not reveal any specific DNA sequence information about the sex chromosomes and thus is limited as an entry point for further molecular analysis.

Reptiles, amphibians and fish are of particular interest to evolutionary biologists as transitions among sex-determining mechanisms have occurred repeatedly in these clades (Hillis & Green **1990**; Mank *et al*. **2006**; Takehana *et al*. **2007**; Volff *et al*. **2007**; Ezaz *et al*. **2009**; Ross *et al*. **2009**; Gamble **2010**; Miura *et al*. **2012**). RAD-seq has the potential to rapidly screen large numbers of species in these groups for their sex-determining mechanism, facilitating comparative analyses of sex chromosome evolution. In many cases, RAD-seq may also provide indications of conserved homology between species, particularly closely related species as we show here, and thus provide valuable information of sex chromosome homology.

Paired-end RAD-seq reads and RAD-seq generated linkage maps have been suggested as means to aid assembly of whole genomes (Amores *et al*. **2011**; Willing *et al*. **2011**). Similarly, sex-specific molecular markers identified via RAD-seq may prove useful in identifying Y or W chromosome contigs when assembling whole genomes. Sex-specific RAD markers can provide entry points to identify sex-specific contigs from whole-genome assemblies, thereby expanding our knowledge of sex chromosome gene content and evolution.

Our analyses of the RAD-seq data assumed a simple presence/absence model of sex-specific restriction sites. This assumption depends upon at least some divergence between the X and the nonrecombining region of the Y chromosomes that allows new restriction sites to evolve and spread through the population. Analyses that search for sex-specific SNPs associated with sex-nonspecific restriction sites, for example Bewick *et al*. (**2013**) and Carmichael *et al*. (**2013**), could yield additional sex-linked markers and are worth exploring. These SNPs do not depend on the evolution of male-specific restriction sites on the Y, simply mutations adjacent to existing restriction sites. This might be a particularly useful modification in species with newly evolved sex chromosome systems where the nonrecombining portion of the Y or W is presumed to be very small and has not had time to diverge significantly from the X (Charlesworth *et al*. **2005**; Volff *et al*. **2007**). SNP analyses might also prove useful to better understand sex chromosome degeneration and reduced recombination between the X and Y (Tripathi *et al*. **2009**; Bruneaux *et al*. **2013**).

Fully sequenced genomes and other genomic resources are not necessary for RAD-seq analysis but can be helpful for characterizing sex-linked markers that are identified. We were fortunate to identify a protein-coding gene via RAD-seq. Only a small proportion of RAD-tags will overlap with a coding region (Amores *et al*. **2011**; Chutimanitsakun *et al*. **2011**; Bruneaux *et al*. **2013**). Sequencing paired-end reads to generate larger RAD loci has been proposed as one means of increasing the probability of accurately mapping and identifying RAD-tags (Amores *et al*. **2011**). In our case, the subsequent characterization of the male-specific marker was only possible because we produced paired-end reads. We used resources associated with the published *A. carolinensis* genome (Alföldi *et al*. **2011**; Eckalbar *et al*. **2013**) to characterize the male-specific marker AcarBC. We found a gametolog of this marker by BLAST searching the NCBI databases and verified its X chromosome linkage using qPCR. We also identified the putative transcript from AcarBC, *rtdr1y*, from a testis cDNA library on

GenBank. Sequence similarity and shared male-specific amplification supports the hypothesis that AcarBC is a portion of the 3' UTR and an upstream intron from *rtdr1y*. Primers designed from a portion of *rtdr1y* upstream from the region sharing similarity with AcarBC amplified in a sex-specific manner in five additional *Anolis* species, indicating conservation of a portion of the Y chromosome in these taxa. This is consistent with the hypothesis that sex chromosomes in *Anolis* are derived from a single ancestral XY pair (Gamble *et al*. **In Press**). The failure to amplify *rtdr1y* in some other *Anolis* species is likely due to one of two things. First, sequence evolution of the primer sites could make the primers less efficient and cause PCR to fail. Second, an intron insertion in the amplified fragment could result in a gene product too large to amplify. Given the dynamic nature of the Y chromosome, either of these scenarios is plausible. Y chromosomes are known to degenerate at varying rates and undergo substantial reorganization over short periods of evolutionary time (Rice **1996**; Kirsch *et al*. **2008**; Hughes *et al*. **2012**), and further work will be needed to fully understand the history and phylogenetic distribution of *rtdr1y* on the *Anolis* Y chromosome.

The number of sex-specific markers identified using RAD-seq will likely vary depending on the size of the nonrecombining portion of the Y or W chromosome. Recovery of just a single male-specific marker with RAD-seq in *A. carolinensis* is not therefore unexpected. *A. carolinensis* has small, homomorphic sex chromosomes, and consequently, we expect few male-specific restriction sites to be present. Mice, on the other hand, like most mammals, have extremely heteromorphic sex chromosomes (Nesbitt & Francke **1973**; Graves **2006**) and should have more unique restriction sites and therefore more sex-specific molecular markers, than *A. carolinensis*. An *in silico* restriction digest of the sequenced portion of the mouse Y chromosome using *SbfI* (data not shown) confirmed this, revealing 24 unique Y-linked *SbfI* sites, far more than recovered here in *A. carolinensis*.

It can be difficult to determine a priori the number of individuals necessary to identify sex-specific markers with RAD-seq. Our original sample size was based on the number of available samples. Additional samples obtained subsequent to performing the RAD-seq experiment allowed us to validate sex-specific markers in animals beyond of our original sample set. Our reanalyses of smaller, subsampled data sets recovered the two male-specific RAD loci identified by the full analysis, suggesting that only a few individuals are needed to get the 'correct' answer. Even so, the smaller data sets identified an increasing number of false positives as sample size decreased. Validating the large number of putative sex-specific RAD loci would mean performing significantly more PCRs than necessary for the larger data sets. The trade-offs between doing RAD-seq on more individuals vs. increased validation via PCR will need to be weighed by each researcher but it appears that including more individuals should result in less effort validating putative sex-specific RAD-tags.

In addition to providing information regarding the sex-determining mechanism of a species, sex-specific markers can still perform their traditional role as a simple means of identifying males and females via PCR. In this regard, primers for AcarB and *rtdr1y* can be used to sex embryonic tissue from *A. carolinensis* and additional *Anolis* species in the Norops clade, in our samples the clade consisting of *A. sagrei*, *A. grahami*, *A. lineatopus*, *A. lineatus* and *A. planiceps*. This will prove useful for studies of the development of sexually dimorphic phenotypes, sex ratio evolution and more.

Conclusions

Sex-specific markers have proven useful in identifying the sex chromosome systems of many animal and plant species (Charlesworth & Mank **2010**) and will be especially helpful in species lacking heteromorphic sex chromosomes. We provide a workflow that identifies and validates sex-specific markers using RAD-seq data from multiple male and female individuals. We illustrate the utility of this workflow using the lizard *A. carolinensis* and identify a male-specific marker, AcarBC. Utilizing the resources associated with the recently published *Anolis* genome, we discover the putative Y chromosome gene *rtdr1y*, which corresponds to AcarBC, as

well as an X chromosome gametolog, *rtdr1*. These results highlight the potential utility of RAD-seq as a tool to uncover the sex chromosome systems of large numbers of nonmodel species in a rapid, cost-effective manner.

Acknowledgements

Supporting Information

men12237-sup-0001-TableS1-7.pdf

Please note: The publisher is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing content) should be directed to the corresponding author for the article.

References

Alföldi J, Di Palma F, Grabherr M *et al*. (2011) The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature*, **477**, 587– 591.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403– 410.

Amores A, Catchen J, Ferrara A, Fontenot Q, Postlethwait JH (2011) Genome evolution and meiotic maps by massively parallel DNA sequencing: Spotted gar, an outgroup for the teleost genome duplication. *Genetics*, **188**, 799– 808.

Anderson JL, Marí AR, Braasch I *et al*. (2012) Multiple sex-associated regions and a putative sex chromosome in zebrafish revealed by RAD mapping and population genomics. *PLoS ONE*, **7**, e40701.

Baird NA, Etter PD, Atwood TS, *et al*. (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE*, **3**, e3376.

Baxter SW, Davey JW, Johnston JS *et al*. (2011) Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. *PLoS ONE*, **6**, e19315.

Berset-Brändli L, Jaquiery J, Dubey S, Perrin N (2006) A sex-specific marker reveals male heterogamety in European tree frogs. *Molecular Biology and Evolution*, **23**, 1104– 1106.

Bewick AJ, Chain FJJ, Zimmerman LB *et al*. (2013) A large pseudoautosomal region on the sex chromosomes of the frog *Silurana tropicalis*. *Genome Biology and Evolution*, **5**, 1087– 1098.

Bruneaux M, Johnston SE, Herczeg G *et al*. (2013) Molecular evolutionary and population genomic analysis of the nine-spined stickleback using a modified restriction-site-associated DNA tag approach. *Molecular Ecology*, **22**, 565– 582.

Bull JJ (1980) Sex determination in reptiles. *Quarterly Review of Biology*, **55**, 3– 21.

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

Campbell-Staton SC, Goodman RM, Backström N *et al*. (2012) Out of Florida: mtDNA reveals patterns of migration and Pleistocene range expansion of the Green Anole lizard (*Anolis carolinensis*). *Ecology and Evolution*, **2**, 2274– 2284.

Carmichael SN, Bekaert M, Taggart JB *et al*. (2013) Identification of a sex-linked SNP marker in the salmon louse (*Lepeophtheirus salmonis*) using RAD sequencing. *PLoS ONE*, **8**, e77832.

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**, 171– 182.

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**, 9– 31.

Charlesworth D, Charlesworth B, Marais G (2005) Steps in the evolution of heteromorphic sex chromosomes. *Heredity*, **95**, 118– 128.

Chen S-L, Li J, Deng S-P *et al*. (2007) Isolation of female-specific AFLP markers and molecular identification of genetic sex in half-smooth tongue sole (*Cynoglossus semilaevis*). *Marine Biotechnology*, **9**, 273– 280.

Chutimanitsakun Y, Nipper RW, Cuesta-Marcos A *et al*. (2011) Construction and application for QTL analysis of a Restriction Site Associated DNA (RAD) linkage map in barley. *BMC Genomics*, **12**, 4.

Devlin RH, Nagahama Y (2002) Sex determination and sex differentiation in fish: an overview of genetic, physiological, and environmental influences. *Aquaculture*, **208**, 191– 364.

Devlin R, Biagi C, Smailus D (2001) Genetic mapping of Y-chromosomal DNA markers in Pacific salmon. *Genetica*, **111**, 43– 58.

Eckalbar WL, Hutchins ED, Markov GJ *et al*. (2013) Genome reannotation of the lizard *Anolis carolinensis* based on 14 adult and embryonic deep transcriptomes. *BMC Genomics*, **14**, 49.

Ellegren H (1996) First gene on the avian W chromosome (CHD) provides a tag for universal sexing of non-ratite birds. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **263**, 1635– 1641.

Etter PD, Bassham S, Hohenlohe PA, Johnson EA, Cresko WA (2011) SNP discovery and genotyping for evolutionary genetics using RAD sequencing. In: *Molecular Methods for Evolutionary Genetics* (eds V Orgogozo, MV Rockman), pp. 157– 178. Springer, New York.

Ezaz T, Quinn AE, Miura I *et al*. (2005) The dragon lizard *Pogona vitticeps* has ZZ/ZW micro-sex chromosomes. *Chromosome Research*, **13**, 763– 776.

Ezaz T, Sarre S, O'Meally D, Graves J, Georges A (2009) Sex chromosome evolution in lizards: independent origins and rapid transitions. *Cytogenetic and Genome Research*, **127**, 249– 260.

Felip A, Young WP, Wheeler PA, Thorgaard GH (2005) An AFLP-based approach for the identification of sex-linked markers in rainbow trout (*Oncorhynchus mykiss*). *Aquaculture*, **247**, 35– 43.

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

Gamble T, Zarkower D (2012) Sex determination. *Current Biology*, **22**, R257– R262.

Gamble T, Geneva AJ, Glor RE, Zarkower D (In Press) *Anolis* sex chromosomes are derived from a single ancestral pair. *Evolution*, doi: 10.1111/evo.12328, in press.

Graves JAM (2006) Sex chromosome specialization and degeneration in mammals. *Cell*, **124**, 901– 914.

Griffiths R, Orr K (1999) The use of amplified fragment length polymorphism (AFLP) in the isolation of sex-specific markers. *Molecular Ecology*, **8**, 671– 674.

Griffiths R, Orr K, Adam A, Barber I (2000) DNA sex identification in the three-spined stickleback. *Journal of Fish Biology*, **57**, 1331– 1334.

Hacker A, Capel B, Goodfellow P, Lovell-Badge R (1995) Expression of *Sry*, the mouse sex determining gene. *Development*, **121**, 1603– 1614.

Hayes TB (1998) Sex determination and primary sex differentiation in amphibians: Genetic and developmental mechanisms. *Journal of Experimental Zoology*, **281**, 373– 399.

Hillis DM, Green DM (1990) Evolutionary changes of heterogametic sex in the phylogenetic history of amphibians. *Journal of Evolutionary Biology*, **3**, 49– 64.

Hughes JF, Skaletsky H, Brown LG *et al*. (2012) Strict evolutionary conservation followed rapid gene loss on human and rhesus Y chromosomes. *Nature*, **483**, 82– 86.

Kaiser VB, Bachtrog D (2010) Evolution of sex chromosomes in insects. *Annual Review of Genetics*, **44**, 91– 112.

Kent WJ (2002) BLAT – The BLAST-like alignment tool. *Genome Research*, **12**, 656– 664.

Kirsch S, Münch C, Jiang Z *et al*. (2008) Evolutionary dynamics of segmental duplications from human Y-chromosomal euchromatin/heterochromatin transition regions. *Genome Research*, **18**, 1030– 1042.

Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. *Bioinformatics*, **23**, 1289– 1291.

Lee B, Penman D, Kocher T (2003) Identification of a sex-determining region in Nile tilapia (*Oreochromis niloticus*) using bulked segregant analysis. *Animal Genetics*, **34**, 379– 383.

Mank JE, Promislow DEL, Avise JC (2006) Evolution of alternative sex-determining mechanisms in teleost fishes. *Biological Journal of the Linnean Society*, **87**, 83– 93.

Meyer LR, Zweig AS, Hinrichs AS *et al*. (2013) The UCSC genome browser database: extensions and updates 2013. *Nucleic Acids Research*, **41**, D64– D69.

Miura I, Ohtani H, Ogata M (2012) Independent degeneration of W and Y sex chromosomes in frog *Rana rugosa*. *Chromosome Research*, **20**, 47– 55.

Nesbitt MN, Francke U (1973) A system of nomenclature for band patterns of mouse chromosomes. *Chromosoma*, **41**, 145– 158.

O'Meally D, Ezaz T, Georges A, Sarre SD, Graves JAM (2012) Are some chromosomes particularly good at sex? Insights from amniotes. *Chromosome Research*, **20**, 7– 19.

Palaiokostas C, Bekaert M, Davie A *et al*. (2013a) Mapping the sex determination locus in the Atlantic halibut (*Hippoglossus hippoglossus*) using RAD sequencing. *BMC Genomics*, **14**, 566.

Palaiokostas C, Bekaert M, Khan MG *et al*. (2013b) Mapping and validation of the major sex-determining region in Nile tilapia (*Oreochromis niloticus* L.) using RAD sequencing. *PLoS ONE*, **8**, e68389.

Pfaffl MW, Horgan GW, Dempfle L (2002) Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Research*, **30**, e36.

Quinn A, Ezaz T, Sarre S, Graves JM, Georges A (2009a) Extension, single-locus conversion and physical mapping of sex chromosome sequences identify the Z microchromosome and pseudo-autosomal region in a dragon lizard, *Pogona vitticeps*. *Heredity*, **104**, 410– 417.

Quinn AE, Radder RS, Sarre SD *et al*. (2009b) Isolation and development of a molecular sex marker for *Bassiana duperreyi*, a lizard with XX/XY sex chromosomes and temperature-induced sex reversal. *Molecular Genetics and Genomics*, **281**, 665– 672.

Rice WR (1996) Evolution of the Y sex chromosome in animals. *BioScience*, **46**, 331– 343.

Ritland K (2011) Evolutionary potential in the wild: more than meets the eye. *Molecular Ecology*, **20**, 3494– 3495.

Ross JA, Urton JR, Boland JE, Shapiro MD, Peichel CL (2009) Turnover of sex chromosomes in the stickleback fishes (Gasterosteidae). *PLoS Genetics*, **5**, 1– 12.

Stöck M, Croll D, Dumas Z *et al*. (2011) A cryptic heterogametic transition revealed by sex-linked DNA markers in Palearctic green toads. *Journal of Evolutionary Biology*, **24**, 1064– 1070.

Taberlet P, Mattock H, Dubois-Paganon C, Bouvet J (1993) Sexing free-ranging brown bears *Ursus arctos* using hairs found in the field. *Molecular Ecology*, **2**, 399– 403.

Takehana Y, Demiyah D, Naruse K, Hamaguchi S, Sakaizumi M (2007) Evolution of different Y chromosomes in two medaka species, *Oryzias dancena* and *O. latipes*. *Genetics*, **175**, 1335– 1340.

Tamura K, Peterson D, Peterson N *et al*. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, **28**, 2731– 2739.

Tollis M, Ausubel G, Ghimire D, Boissinot S (2012) Multi-locus phylogeographic and population genetic analysis of *Anolis carolinensis*: Historical demography of a genomic model species. *PLoS ONE*, **7**, e38474.

Traut W, Eickhoff U, Schorch JC (2001) Identification and analysis of sex chromosomes by comparative genomic hybridization (CGH). *Methods in Cell Science*, **23**, 155– 161.

Tripathi N, Hoffmann M, Weigel D, Dreyer C (2009) Linkage analysis reveals the independent origin of poeciliid sex chromosomes and a case of atypical sex inheritance in the guppy (*Poecilia reticulata*). *Genetics*, **182**, 365– 374.

Untergasser A, Cutcutache I, Koressaar T *et al*. (2012) Primer3 - New capabilities and interfaces. *Nucleic Acids Research*, **40**, e115.

Valenzuela N, Adams DC, Janzen FJ (2003) Pattern does not equal process: exactly when is sex environmentally determined? *American Naturalist*, **161**, 676– 683.

Volff JN, Nanda I, Schmid M, Schartl M (2007) Governing sex determination in fish: regulatory putsches and ephemeral dictators. *Sexual Development*, **1**, 85– 99.

Wallace H, Badawy G, Wallace B (1999) Amphibian sex determination and sex reversal. *Cellular and Molecular Life Sciences*, **55**, 901– 909.

Willing E-M, Hoffmann M, Klein JD, Weigel D, Dreyer C (2011) Paired-end RAD-seq for de novo assembly and marker design without available reference. *Bioinformatics*, **27**, 2187– 2193.

Data Accessibility

Sequencing reads have been deposited in the NCBI Sequence Read Archive (Accession no. PRJNA236043)

Supplementary Files are accessible via DRYAD: doi:**10.5061/dryad.pq608**

Supplementary Files

Supplementary File 1. RADtools output – full data set.

Supplementary File 2. Fasta file with male-specific RAD loci.

Supplementary File 3. RADtools output – subsampled data sets.