

OPEN

Identifying sex-linked markers in *Litoria aurea*: a novel approach to understanding sex chromosome evolution in an amphibian

Jarrod Sopniewski ^{1*}, Foyez Shams¹, Benjamin C. Scheele², Ben J. Kefford ¹ & Tariq Ezaz ^{1*}

Few taxa exhibit the variability of sex-determining modes as amphibians. However, due to the presence of homomorphic sex chromosomes in many species, this phenomenon has been difficult to study. The Australian frog, *Litoria aurea*, has been relatively well studied over the past 20 years due to widespread declines largely attributable to chytrid fungus. However, it has been subject to few molecular studies and its mode of sex determination remained unknown. We applied DArTseq™ to develop sex-linked single nucleotide polymorphisms (SNPs) and restriction fragment presence/absence (PA) markers in 44 phenotypically sexed *L. aurea* individuals from the Molonglo River in NSW, Australia. We conclusively identified a male heterogametic (XX-XY) sex determination mode in this species, identifying 11 perfectly sex-linked SNP and six strongly sex-linked PA markers. We identified a further 47 moderately sex-linked SNP loci, likely serving as evidence indicative of XY recombination. Furthermore, within these 47 loci, a group of nine males were found to have a feminised Y chromosome that significantly differed to all other males. We postulate ancestral sex-reversal as a means for the evolution of this now pseudoautosomal region on the Y chromosome. Our findings present new evidence for the 'fountain of youth' hypothesis for the retention of homomorphic sex chromosomes in amphibians and describe a novel approach for the study of sex chromosome evolution in amphibia.

The biological process of sex-determination amongst vertebrates is diverse and widely studied¹⁻⁶, and few taxa exhibit the variability of sex-determining modes as that of the amphibians^{1,5,7,8}. Specifically, within the anurans (frogs and toads), multiple variations of sex chromosome systems have been identified, with both male heterogamety (XX/XY)⁹ and female heterogamety (ZZ/ZW)⁶ being observed. Within anurans both XY and ZW sex chromosome systems have evolved multiple times. In addition, changes in the heterogametic sex have been observed involving the same chromosome pairs (i.e. XY chromosomes become ZW chromosomes) as reported in the Japanese frog *Glandirana rugosa*¹⁰, thereby suggesting frequent and easy transitions between heterogametic systems (Fig. 1). Despite all amphibian species studied thus far exhibiting genetic sex determination (GSD)¹, phenotypic sex has been shown to be influenced by both temperature^{5,11} and endocrine disrupting chemical pollutants^{7,12}. Adverse temperatures and pollution have resulted in sex reversal being observed in various anuran species, such as *Rana sylvatica*⁵ and *G. rugosa*¹³. Despite the identification of some amphibian sex determining modes¹⁴, in the context of the unique biosphere of Australia, knowledge of the sex determining mechanisms in amphibians is sparse. To our knowledge, no sex-linked markers have been reported for an anuran species in Australia, nor has sex reversal been noted in any Australian anurans.

The study of sex-determining systems in anurans has been difficult; many amphibians have homomorphic sex chromosomes¹⁵, thus sex chromosomes cannot be easily identified using conventional karyotyping. A variety of methods have been utilised to genotypically assign sex, such as the examination of inheritance patterns of allozymes¹⁶, identification of sex-linked microsatellite loci³ and more recently genotyping-by-sequencing (GBS) methods such as RAD-seq¹⁷ (Fig. 1). DArTseq™, developed by Diversity Arrays Technology Pty Ltd, has been proposed by Lambert *et al.*⁹ as an effective method to identify sex-linked markers in non-model species using single nucleotide polymorphic (SNP) loci, which also generate restriction site specific presence-absence (PA)

¹Institute for Applied Ecology, University of Canberra, Bruce 2617, Canberra, Australia. ²Fenner School of Environment and Society, The Australian National University, Canberra, ACT, 2601, Australia. *email: Jarrod.Sopniewski@canberra.edu.au; Tariq.Ezaz@canberra.edu.au

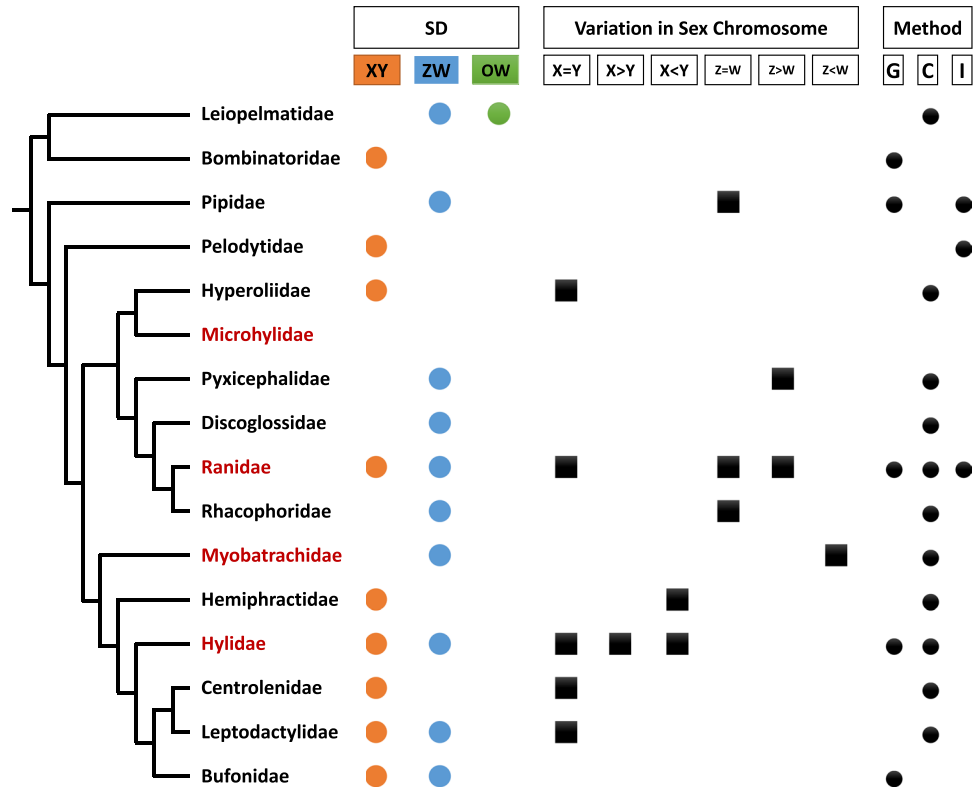


Figure 1. Truncated phylogeny (not according to scale) showing variations of sex determination modes and sex chromosome morphologies among different Anuran families. Australian frog species belong to red coloured families. G = genetic, C = cytological, I = immunologic. Lack of symbols indicate an absence of data. Summarized from Hillis and Green (1990)⁵⁶, Schmid and Steinlein (2001)⁵⁷, Sarre *et al.* (2011)⁵⁸, Miura (2017)⁵⁹ and Ito (2018)⁶⁰. Phylogeny adopted from Jetz and Pyron (2018)⁶¹.

markers. In comparison to the aforementioned methods, DARtseq™ provides a relatively inexpensive, simple and reliable source of genetic data, which can also be utilised further for additional population analysis^{9,18,19}. DARtseq™ combines the genome complexity reduction method DArT, first reported in 2001²⁰, with modern next generation sequencing (NGS) platforms. The technology integrates high throughput processing of DNA to generate complexity reduced genome representations using combinations of restriction enzymes selecting the optimal fraction of the genome for sequencing with dedicated analytical processes¹⁹. This method has been successfully deployed in many plant¹⁸ and animal²¹ species. DARtseq™ allows for a lower density of high-quality markers, providing reasonable coverage of the genome with a low level of excluded data¹⁸. The markers generated by DARtseq™ reveal loci assorted to a particular sex, thereby accurately identifying loci tightly linked to the sex-determining region of the sex chromosomes, as such proving to be a useful molecular tool for the deciphering of sex-determining modes in non-model species with cryptic sex chromosomes. Furthermore, through the observation of discordance between genotypic and phenotypic sex, these sex-linked markers allow sex-reversal to be identified in non-model species⁹.

In amphibians, sex reversal has been observed frequently both in the laboratory²² and in natural populations^{12,23–25}. As previously mentioned, the sex chromosomes of amphibians are largely unique among taxa, both due to the common occurrence of homomorphy as well as transitions between sex-determining modes. The retention of homomorphy between sex chromosomes among amphibians has been shown to be somewhat attributable to occasional recombination between XY sex chromosomes as reported in the hylids *H. arborea*, *H. intermedia* and *H. molleri*¹⁵. In these male heterogametic species, XY recombination has been shown to significantly differ from 0, though be less than between the X chromosomes of homogametic females. Furthermore, the ‘fountain of youth’ hypothesis²⁶ suggests that sex chromosome recombination is phenotypically moderated, rather than genotypically; in a male heterogametic species, sex chromosome recombination is suppressed by ‘maleness’, rather than differing sex chromosomes^{26–28}. As such, sex-reversal has great ramifications for the evolution of the Y chromosome in amphibians; Y recombination may traditionally be suppressed by ‘maleness’, however in an XY female, recombination is instigated, and the Y chromosome is reinvigorated²⁷. The identification and analysis of sex-linked genetic markers provides a genotypic tool for the practical sexing of individuals and allows for the detection of sex-reversal in populations and insight into the evolutionary history of the Y chromosome (or W chromosome in a female heterogametic species).

One non-model species from which the identification of sex-linked markers will prove useful to the study of sex chromosome evolution is the Australian frog, *Litoria aurea* (Lesson, 1827 - Anura: Hylidae), or the ‘Green and Golden Bell Frog’. *Litoria aurea*, listed in 2004, is deemed ‘vulnerable’ on the IUCN Red List²⁹, and is experiencing

ongoing declines³⁰ and was thought to have completely disappeared from areas above 250 m in elevation³¹. As such, this species has been relatively well studied over 20 years³². The species' range once extended from the East Gippsland region in Victoria along the eastern coastline to northern New South Wales (NSW), though has since suffered disappearances from over 90% of its range³³. There is robust evidence that the demise of this species is principally due to the introduction of chytrid fungus^{34,35}, although pollution³⁶ and the predation of eggs and tadpoles from the introduced fish *Gambusia holbrooki*³² may have also contributed. Despite the numerous studies undertaken on this frog, only one has used molecular data, this being a population analysis of several microsatellite loci in the early 2000s^{33,37}. Thus far, *L. aurea*'s sex-determining mode is unknown and sex reversal in this species is undocumented. Inclusive of our study population, wild populations that have been studied often exhibit a male bias sex ratio^{38,39}. Detection of males is far higher than females due to the loudness of their call, thus this sex bias is likely attributed to sampling bias⁴⁰. However, new sex-linked molecular markers, such as those obtainable from DArTseq™, will allow for exploration as to whether sex reversal is contributing to the reported male bias sex ratio in this species. As a member of the hylids, a family that has already been studied with regard to sex chromosome evolution^{15,41}, *L. aurea* acts as a novel model species to study this phenomenon. Furthermore, the markers we have developed not only contribute to the evolutionary understanding of the sex chromosomes of the anuran, they also have potential for future use as a conservation resource, by enabling identification of whether sex-reversal is occurring in wild populations.

In this study, we aimed to address the knowledge gap that exists with respect to the sex-determining modes in an Australian frog, *L. aurea*, as well as identify any sex-linked markers in this taxon. From these markers, we also aimed to postulate the evolutionary history of the sex chromosomes present in this species. Our findings elucidate the previously unknown sex-determining mode, whilst also demonstrating a novel approach to understanding the fascinating interactions between alternate sex chromosomes found in amphibians.

Results

Filtration of DArTseq™ data and population analysis. We obtained 40,152 SNP markers and 26,119 DArT silico markers (presence-absence/PA markers) from initial DArT sequencing data. We retained a total of 20,111 SNP markers after filtration following our cut off filtering criteria (call rate = 1, reproducibility = 1). Additionally, we retained 19,121 PA markers after filtering (call rate > 0.90, reproducibility = 1). Initial population analysis utilising these markers suggested the individuals across the three sampling sites to be members of a single, interbreeding population, with little to no population substructure evident between the sites. Pairwise F_{ST} values reported between sampling sites were no higher than 0.026 and the overall population had an F_{ST} value of 0.016, a value generally indicative of little population substructure⁴². Our principal coordinates analysis plot (Supplementary Fig. S1) supports this conclusion, revealing no clustering of individuals by sampling site. Therefore, we considered samples from the three sites to be a single population comprising of 31 males, 13 females and 8 ambiguous individuals whose phenotypic sex could not be conclusively determined in the field. Other studies have shown the degree of a potential marker's sex-linkage to differ between subpopulations^{8,18}; thus being a single population, we could assume that any markers correlated with sex at a particular sampling site, though not others, are highly unlikely to be sex-linked.

Identification and validation of sex-linked markers. Under the parameters outlined in the methods, a total of 11 SNP markers were found to be perfectly linked to maleness, i.e. all males were found to be heterozygous and all females were found to be homozygous for the reference allele at 11 distinct loci (Table 1). In addition, six PA markers were identified as present in over 90% of males and absent in over 90% of females (Table 2). For these selected PA markers, none were found to be present in females or absent in males; however, some results returned 'null alleles.' No SNP loci were found to assort to females, and no PA markers were noted to be present in over 90% of females and absent in over 90% of males. As such, a male heterogametic sex determination system (XX/XY) has been conclusively determined in *L. aurea*. Furthermore, perfect sex linkage amongst the SNP markers implies that these loci are likely to be tightly linked to the sex determining locus of the sex chromosomes.

For the PA markers, no loci were found to be absent in males, and no loci present in females; for each potentially sex-linked PA marker, some individuals returned a null allele. A null allele appears as a dash in the DArTseq™ file, indicative of a non-zero count recorded, though with too low counts to be considered '1' (present). As such, this often is indicative of heterozygotes. Although it is likely that all males possess the PA marker restriction fragment for each locus, and all females do not, this cannot be confirmed from our data set. As such, though these loci are likely tightly linked to the sex determining region, as with all SNP alleles identified, this cannot be validated without further investigation. Additionally, each of the eight individuals whose sex was not confidently determined in field sampling perfectly conformed to one sex based on each of these selected markers; five perfectly conformed to the female genotype, whereas three aligned perfectly with the male genotype (Fig. 2). This provides validation of our sex-linked markers and overall, the sex ratio of our sample was 65.4% male. This also provides further support to our claim that these markers are indicative of the sex determining region of the sex chromosomes.

Furthermore, we identified an additional 47 SNP markers for which each female was homozygous for the reference allele and 13 of the males (herein termed 'Males 1') were heterozygous (expressing the genotype for both reference and alternate alleles), when considering the initially non-sex assigned frogs with their genotypic sex. Interestingly, the remaining 23 males that were not perfectly heterozygous assorted into two distinct groups: one group (termed here as 'Males 2'), comprising of 12 males, were heterozygous for an average of 92.2% (s.d. = 5.90%) of the 47 moderately sex-linked loci. The remaining group of individuals, comprising of 9 males (termed here as 'Males 3'), were heterozygous for an average of 16.3% (s.d. = 2.38%) of the 47 moderately sex-linked loci. A two-sample t-Test assuming unequal variances, considering the percent heterozygosity of each individual, shows these two groups to be significantly different to each other ($t(15) = 40.38, p = 1.017^{-17}$, two-tailed). A

Locus	Allele	Sequence	Proportion Homozygous (Reference Allele)		Proportion Heterozygous	
			Male	Female	Male	Female
LiauCT01	REF ALT	TGCAGAAGAAATAGATGAGGTGGAGAGGTGTAATG <u>A</u> ATGTGTGCAGCTATCTTCGTTGTAGCTCATGG TGCAGAAGAAATAGATGAGGTGGAGAGGTGTAATG <u>C</u> ATGTGTGCAGCTATCTTCGTTGTAGCTCATGG	0	1	1	0
LiauCT02	REF ALT	TGCAGAGGAAATGTGTAACAGACCAAAATCCGAAATTACATACAGTAGTCACTAACAATTTAAACATGTG TGCAGAGGAAATGTGTAACAGAC <u>A</u> AAATCCGAAATTACATACAGTAGTCACTAACAATTTAAACATGTG	0	1	1	0
LiauCT03	REF ALT	TGCAGATATGTACAGAGTGAAGAGGAGGGTGA <u>A</u> AGGGGAGAAGTCTGAGAGCTGCCAGGATGGATTG TGCAG <u>G</u> TATGTACAGAGTGAAGAGGAGGGTGG <u>G</u> AAGGGGAGAAGTCTGAGAGCTGCCAGGATGGATTG	0	1	1	0
LiauCT04	REF ALT	TGCAGATCA <u>G</u> CAGTGTGTAGAATTGTATGTTGAGCATGGTTTCAAATATGCAAAGTCCACACCCTG TGCAGATCA <u>A</u> CAGTGTGTAGAATTGTATGTTGAGCATGGTTTCAAATATGCAAAGTCCACACCCTG	0	1	1	0
LiauCT05	REF ALT	TGCAGTGGAGCTTACACCAACAAGCAGAGGAGAGCTGTAGCATCAGATAAGCAATAGGAAGTCCAGAG TGCAGTGGAGCTT <u>T</u> ACACCAACAAGCAGAGGAGAGCTGTAGCATCAGATAAGCAATAGGAAGTCCAGAG	0	1	1	0
LiauCT06	REF ALT	TGCAGGTGCCTCATACAGATGC <u>A</u> ACAGGCAGATTTAGGCACTGTTACACTTGTCTTCTTGTCTCTG TGCAGGTGCCTCATACAGATGC <u>A</u> ACAGGCAGATTTAGGCACTGTTACACTTGTCTTCTTGTCTCTG	0	1	1	0
LiauCT07	REF ALT	TGCAGTAGAAGGTGGCTATTGGC <u>A</u> TCCTGTTCTTGTCTGTGTAACCCAGGGACCTTCAAATGGGCCACTG TGCAGTAGAAGGTGGCTATTGGC <u>T</u> TCCTGTTCTTGTCTGTGTAACCCAGGGACCTTCAAATGGGCCACTG	0	1	1	0
LiauCT08	REF ALT	TGCAGTCTGTAACATGGTTTAAATTTGTTATTTGGTAAAAGGAGGCTAAGTATACCCAGGGTAGTCTTA TGCAGTCTGTAACATGGTTTAAATTTGTTATTTGGTAAAAGGAGGCTAAGTATACCCAGGGTAGTCTTA	0	1	1	0
LiauCT09	REF ALT	TGCAGTGCAGTCTGATATACAGTATACTCCCTTATGCAGTTCAGTGTCTCAAATTGACCATTATCTTT TGCAGTGCAGTCTGATATACAGTATACTCCCTTATGCAGTTCAGTGTCTCAAAGTGCACCATATCTTT	0	1	1	0
LiauCT10	REF ALT	TGCAGTAAAAGTGTGTCTCAACAGCCAACCTGTACAGATATTGGTCACTGCATGAGATCGGAAGAGC TGCAGTAAAAGTGTGTCTCAACAGCCAACCTGTACAGATATTGGTCACTGCATGAGATCGGAAGAGC	0	1	1	0
LiauCT11	REF ALT	TGCAGAACATCTTTTCTATGAAGTGGTGCAGAACTAAACTGCATATCCAGATGGGGTTCGCATCAA TGCAGAACATCTTTTCTATGAAGTGG <u>A</u> GTC <u>A</u> AACTAAACTGCATATCCAGATGGGGTTCGCATCAA	0	1	1	0

Table 1. Sequences of both reference (REF) and alternate SNP-containing (ALT) alleles for 11 sex linked SNP loci in *Litoria aurea*, with the position of the polymorphism/s underlined. All loci perfectly assort to males, indicating a male heterogametic (XX-XY) sex determining system.

Locus	Call Rate	Sequence	Male		Female	
			Proportion Present	Proportion Null	Proportion Absent	Proportion Null
LiauCT12	0.980769	TGCAGTAGAAGGTGGCTATTGGCTTCTGTTCTGCTGTGTAACCCAGGGACCTTCAAATGGGCCACTG	0.98	0.02	1	0
LiauCT13	0.942308	TGCAGAAGTGGGGTGTATCTATTTTCGGTAATATTATGATAAATGCCAAGTCAAAGGGGCAGGTAG	0.90	0.10	1	0
LiauCT14	0.942308	TGCAGTGAAGGATGCCTTACAGGTGCTGTGGCCAGATGCAGAGAAAGAGGATAAGTTACCGTTGTATGT	0.94	0.06	1	0
LiauCT15	0.923077	TGCAGTTGCACAATCTCTTTCGAGAAGCTCCTTGCAGAGTATTGCAAAACCACCATTTGCAAAGTCT	0.94	0.06	1	0
LiauCT16	0.942308	TGCAGAGCATTGGGTTACAGTGAAGTCCCCTTTATCTAGGTAGGTGACAAGTCACTGGAATGCTTAT	0.90	0.10	1	0
LiauCT17	0.942308	TGCAGCTTATGAAACTCCTTACACAGTTTTCTCATCTCTTTGAAAGCAAAGTCTACACTACCGTT	0.90	0.10	1	0

Table 2. Six sex-linked PA markers. Call rate refers to the proportion of individuals that returned a result (present or absent). For all loci, where the proportion present or absent for a marker was found to be < 1 , this was solely due to the presence of a null allele (i.e. a valid result was not returned during the sequencing process). No markers that were present in males were found present in any females, and vice versa.

principal coordinates analysis plot (Supplementary Fig. 2) and a Nei's genetic distance heat map (Fig. 3) show how the group 'Males 3' is more similar to females than their male counterparts when considering the perfectly and moderately sex-linked loci.

Spurious sex-linkage estimation. Following the model described in Lambert *et al.*⁹, our filtered dataset of 20,111 SNP markers, for a sample of 44 phenotypically sexed individuals, revealed that 1.14×10^{-9} markers are likely to be spuriously sex-linked. Similarly, for the PA markers (19,121), 1.08×10^{-9} markers would show sex-linkage by chance. As such, it can be stated with confidence that all markers identified are tightly linked with the sex determining locus of this species. As shown in Supplementary Fig. 3, for this number of SNP and PA markers, 14 individuals would be sufficient to be confident that less than one marker would be spuriously sex-linked (adopted from Lambert *et al.*⁹).

Discussion

Our study successfully identified 11 SNP markers to be perfectly assorted to maleness across all known sex individuals, thus providing decisive evidence that *L. aurea* utilises a male heterogametic sex determining system (XX/XY). In addition, we identified a further six PA markers that are strongly sex-linked; excluding null allele results, each identified marker was present in all males and absent in all females. These results demonstrate the usefulness

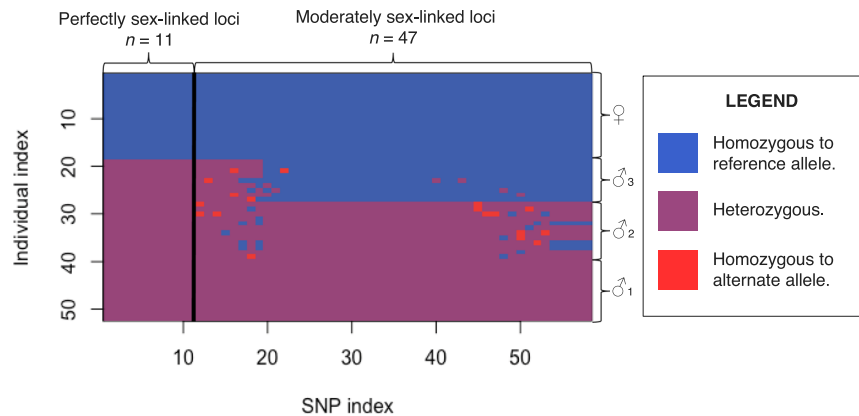


Figure 2. An index of the 11 perfectly sex aligned sequences and the further 47 sex-related sequences created with the ‘glPlot’ function in the R package ‘dartR’⁵³. Blue indicates homozygous to the reference allele, purple is indicative of heterozygosity, and red indicates homozygosity to the alternate SNP-containing allele. Sequences 1–11 (left of the figure) show the perfectly sex related sequences, with all females showing blue and all males showing purple. From this point, four clusters emerge: the identical homozygous females, males group 1 (100% heterozygous), males group 2 (high heterozygosity) and males group 3 (low heterozygosity).

of the DARtseq™ methodology for identifying sex-linked markers⁹. Furthermore, the remaining non-sex-linked loci obtained from DARtseq™ can be used to conduct a number of population analyses^{43–45}.

A further 47 SNP loci were also identified as moderately sex-linked, with all females expressing homozygosity to the reference allele and a total of 13 males recorded as heterozygous across all loci. Of these, eight markers (markers 12–19 inclusive) largely assorted to males (i.e. males were generally heterozygous), as presented in Fig. 2. Here, in the groups ‘Males 2’ and ‘Males 3’, discordance with the heterozygous genotype could reflect potential recombination events between the X and Y sex chromosomes, as has been shown in other hylids⁴¹. Stöck *et al.* concluded that the evolutionary young sex chromosomes of hylids are the result of occasional recombination between the X and Y chromosomes, rather than turnover events with autosomes²⁸. It has been shown on a separate occasion that recombination between the X and Y chromosomes in hylids is significantly greater than 0, though does occur at a rate 10^5 times slower than recombination between X chromosomes¹⁵. Partial assortment of these loci to males suggest that these markers were perhaps restricted to the Y chromosome in the past, however have undergone recombination with the X chromosome, hence no longer exhibit perfect male concordance; these potential recombination events are indicated by the alternate coloured rectangles in Fig. 2. This subset of moderately sex-linked markers provide evidence for the occasional recombination between the X and Y sex chromosomes as reported in other hylid species⁴¹.

When considering the 47 non-perfectly sex-linked markers in their entirety, as previously discussed, the males fall into three distinct groups: ‘Males 1’ (heterozygous for all markers, $n = 13/34$), ‘Males 2’ (mostly heterozygous, $n = 12/34$) and ‘Males 3’ (predominately homozygous to the reference allele, $n = 9/34$). These findings may be indicative of alternate origins of the Y chromosome between these groups. As observed in Fig. 3, ‘Males 1’ appears to be a subset of ‘Males 2’, and we postulate that the minor variations between these two groups may represent further evidence of occasional recombination between the sex chromosomes⁴¹ as previously discussed; the Y chromosome in these individuals is likely of similar origin. The rate of evolution of sex chromosomes in amphibians is far higher than that observed in other taxa⁴⁶, contributed to by occasional XY recombination¹⁵. Interestingly, however, our finding of 11 perfectly sex-correlated markers suggest the presence of a distinct male-specific region of the Y chromosome. In order for this formation of distinct sex chromosomal regions to occur and this degree of sex-specific genotypes to be observed, recombination between the X and Y chromosomes in *L. aurea* have significantly slowed⁴⁷. The vast differentiation between the groups ‘Males 3’ and the other males identified is therefore difficult to interpret.

One possible explanation of this phenomenon is the occurrence of a historical sex reversal event in this population that led to the origin of a differentiated Y chromosome, the lineage of males now represented as ‘Males 3’. The ‘fountain of youth’ hypothesis for the retention of homomorphic sex chromosomes in amphibians postulates that recombination between sex chromosomes is suppressed by phenotypic, not genotypic, sex^{26,27}. That is, recombination between sex chromosomes is encouraged by ‘femaleness’, and suppressed by ‘maleness’, rather than being influenced by the types of sex chromosomes present²⁷. As such, recombination between X and Y chromosomes would be invigorated by sex-reversed XY females; the male phenotype suppresses recombination, thus the female phenotype would allow reinvigoration of the Y chromosome⁴⁸. We postulate that the ‘feminised’ Y chromosome we observe in the ‘Males 3’ group may be resultant from rapid recombination events in an ancestral sex-reversed XY female in a recent sex-reversal event.

Although we posit the occurrence of historic sex-reversal in this population, we did not find evidence of this phenomena currently occurring. During sampling, one individual (Frog31 of Site 2) was tentatively sexed as female in the field, though we were not confident with this morphological sexing, thus for analytical purposes we deemed this individual as ‘unknown sex’. Despite the original tentative female identification, both SNP and PA

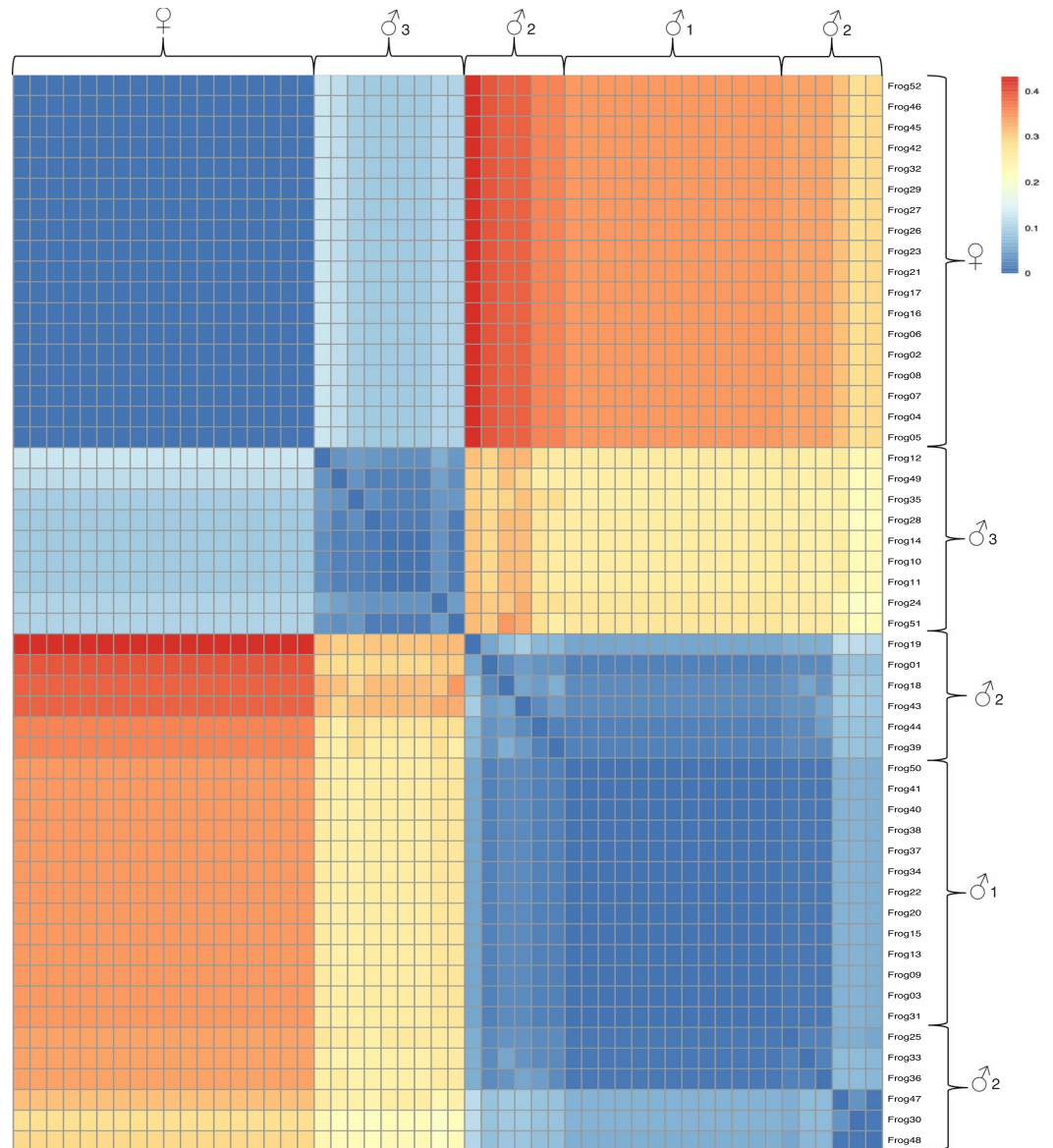


Figure 3. A heat map plotted using the R package ‘pheatmap’⁵⁵ from a genetic distance matrix using the 11 perfectly sex-linked loci and 47 moderately sex-linked loci. The genetic distance matrix was generated using the function ‘stampNeisD’ from the R package StAMPP⁵⁴. This function calculates Nei’s genetic distance between individuals. Blue indicates identical genotypes, whereas high genetic dissimilarity is displayed in red. Females are genetically identical to each other (homozygous to the reference allele at all 58 loci). Similarly, Males 1, heterozygous at all 58 loci, are identical to each other. Males 2, predominately heterozygous, share high similarity with and are clustered around Males 1, although Males 3 possess a more feminine genotype, hence are more similar to females than their male counterparts.

markers perfectly genotyped this individual as a male. This discordance highlights the greater accuracy of our genetic sexing method over morphological sexing, though also signifies the potential these markers possess for the detection of sex reversed individuals. Pollution in the form of endocrine disruptors, particularly increased levels of estrogens in the environment, has been shown to cause sex reversal in other amphibians^{12,24}. Our study population of *L. aurea* resides in an area polluted with metals from a disused mine⁴⁹, and many other populations of *L. aurea* frequent polluted habitats, providing the potential for the occurrence of sex-reversal.

In conclusion, the genotypic markers identified by this study are the first sex-linked molecular markers developed for *L. aurea*, and for an Australian anuran. We were able to identify a notable number of perfectly sex-linked markers given the small portion of the genome represented through DArTseq analysis and the general homomorphy between sex chromosomes often observed in hylids and other amphibians. Furthermore, the 11 perfectly sex-linked markers are indicative of a male specific region of the Y chromosome, present despite major differences observed elsewhere on the chromosome, implying that these markers will likely be sex-linked in other populations. Our results not only elucidated the modes of sex determination in *L. aurea*, though also provide

molecular support for both higher rates of recombination between sex chromosomes in amphibians^{15,28,41} and the 'fountain of youth' hypothesis for the retention of homomorphic sex chromosomes^{26,27}.

Methods

Sample collection. In November and December 2016, 52 *Litoria aurea* adults were sampled from the study population on the Molonglo River and associated floodplains in NSW (an exact location cannot be revealed due to a confidentiality agreement with the private landowner). At the time of sampling, individual sex (male or females) was determined by the presence or absence of male nuptial pads, male throat discolouration and/or female gravidness; where morphological evidence was inconclusive, the individual was designated 'unknown sex'. A single digit from each individual was removed at the base of the third phalange and stored in 70% ethanol. Each individual was handled with a new pair of sterile gloves. Field work was conducted under the NSW government scientific licence SL101797 and approved by The Australian National University animal ethics committee (protocol: A2016/25).

DNA extraction and genotyping-by-sequencing (GBS). DNA extraction and sequencing were conducted by Diversity Arrays Technology Pty Ltd using DArTseq™, a methodology combining DArT genome complexity reduction methods and next generation sequencing technologies⁵⁰. The methodology used was adapted from similar process utilised by Lambert *et al.*⁹ and Hill *et al.*²¹. The DNA samples were digested using *Pst*I and *Sph*I before ligation reaction were performed using two adaptors: a *Pst*I compatible adaptor (consisting of an Illumina flow cell attachment sequence, sequencing primer sequence and a unique barcode sequence), as well as a *Sph*I compatible adaptor (consisting of an Illumina flow-cell attachment region). The ligated fragments then underwent 30 rounds of PCR (94 °C for 20 seconds, 58 °C for 30 seconds and 74 °C for 45 seconds), followed by an extension of seven minutes at 72 °C. Following PCR, equimolar amounts of amplification products derived from each individual were bulked and applied to Illumina's proprietary cBot bridge PCR, which was followed by sequencing on an Illumina HiSeq2000. The single read sequencing was run for 77 cycles.

Proprietary DArTseq™ analytical pipelines were used to process the sequences generated. Initially, the fastq files generated from HiSeq2000 were processed to filter for poor quality sequences. The parameters included the discarding of any results with reproducibility <90% and read depth <3.5 for SNPs and <5 for PA markers, as well as the application of more stringent selection criteria to the barcode region (compared to the remainder of the sequence). Identical sequences were then collapsed into 'fastqcoll' files. These were subjected to a secondary pipeline (DArTsoft14), further DArT proprietary pipeline that is differentiated for single nucleotide polymorphism (SNP) and SilicoDArT (presence/absence of restriction fragments in representation). DArTsoft14 utilises a reference-free algorithm whereby each unique sequence from the fastqcoll file is identified and clustered by sequence similarity (three base pair variation is used as the distance threshold) using an optimised clustering algorithm. The resulting SNP and SilicoDArT markers are labelled with a number of metadata parameters derived from the quantity and distribution of each sequence across all samples analysed. The genotyping process utilised by DArTseq™ includes high levels of technical replication, thereby allowing the parameter of 'reproducibility' to be calculated, an integral metric.

Marker filtering and population analysis. The data obtained from DArTseq™ analysis (provided as Supplementary Data) initially contained 40,152 SNP markers and 26,119 presence-absence markers from the SilicoDArT file after the filtering processes undertaken by DArT. The SNP file scored homozygotes to the reference allele as '0', homozygotes to the alternate SNP-containing allele as '1' and heterozygotes as '2'. In the SNP file, a result of '-', known as a null allele, indicates an absence of the fragment containing the SNP in the genomic representation of that individual. SilicoDArT determines genetically 'dominant' markers, scored in a binary fashion. Results of "1" indicate presence and "0" indicate absence of a restriction fragment with the marker sequence in the genomic representation of the sample being studied. A null allele result in a SilicoDArT file represents non-zero counts of the restriction fragment, though with too low a calling to confidently score as '1'. Call rate refers to the proportion of samples that return a confident result non-null allele results, whereas reproducibility indicates the proportion of technical replicate assay pairs for which the marker score is consistent. As such, we applied additional filtering criteria to the data we received from DArT. For all SNP data, loci with call rate <1 and reproducibility <1 were excluded from analysis. Similarly, for the SilicoDArT file, all loci with a call rate <0.90 and reproducibility <1 were removed from the data set. The call rate threshold was lowered for the SilicoDArT data due to the slightly poorer quality of data received. These stringent filtering methods ensured that only the most reliable sequences were interrogated. These data subsets were used for all subsequent analysis.

Samples were collected from three locations separated by 1–2 kms, and we used the statistical program R⁵¹ to perform various population analyses. The R package 'PopGenReport'⁵² was used to calculate the pairwise F_{ST} values between each sampling location using the function 'pairwise.fstb'. A principal coordinates analysis was also conducted using the functions 'gl.pcoa' and 'gl.pcoa.plot' of the R package 'dartR'⁵³. The results of these tests were then interrogated to confirm status of our study population as a single population. This was important to ascertain, as other studies have found some sex-linked markers to correlate with geography and differ between subpopulations⁸.

Selection of sex-linked markers and identification of sex-determining mode. The markers produced by DArTseq™ comprise of two alleles for each SNP locus: a reference allele (the allele most frequently sequenced in the data subset) and an alternative allele (the allele containing a single nucleotide polymorphism with respect to the reference allele). In a male heterogametic (XX-XY) system, the reference allele, sequenced most frequently, would appear on the X chromosome, and the alternate allele would appear on the Y-chromosome. SNP

markers that are tightly linked with the sex determining region of the Y chromosome should therefore express 100% heterozygosity in the heterogametic sex (males) and 100% homozygosity to the reference allele in the homogametic sex (females)¹². For a female heterogametic (ZZ-ZW) system, the opposite would be true. For PA markers, an XX-XY system would be characterised by all males being marked 'present' for the marker-containing restriction fragment and all females being marked 'absent', and vice versa for a ZZ-ZW system.

We used Microsoft Excel (2015) to test for the sex determining mode used by *L. aurea* and identify and sex-linked markers. Initially, we tested for a male heterogametic system (XX-XY). We manipulated the data such that the columns of results for each individual were arranged in blocks of females, males and unknown sex. For each row, representing each individual for a particular allele, we used the 'countif' function to count all heterozygous results (2) for males, divided by 31 (the number of confidently sexed males). For loci where all males are heterozygous, this would give a result of '1'; i.e. a proportion of 1 (100%) males are heterozygous. A similar process was used to evaluate the level of homozygosity to the reference allele in females. These two results were then added together, meaning a result of '2' would mean that all males are heterozygous, and all females are homozygous to the reference allele; these would be sex-linked loci. The data was then custom sorted so that each allele that gave a result of '2' appeared at the top of the file, and all alleles were sorted into descending order. The opposite procedure was then conducted to test for a female heterogametic system. A similar approach was used for the SilicoDART PA subset; however, no loci were found to have 100% concordance with sex. This was largely due to the lower average call rate present in the SilicoDART file, so we rather selected markers with >90% presence (a result of <0.9) in the heterogametic sex and >90% absence in the homogametic sex and determined these to be sex-linked loci. For all selected PA loci, none returned an 'absence' result in males or a 'presence' result in females.

Further analysis. Upon identification of sex-linked loci, we noticed a trend in our SNP data whereby all females continued to be homozygous to the reference allele and most males appeared to be largely heterozygous, though a group of 9 males were homozygous to the reference allele, similarly to females, consistently at the same loci. We identified a further 47 loci for which 13 males were 100% homozygous (Males 1), 12 males were highly heterozygous (>85% heterozygous; Males 2) and the 'female-like' group of males expressed low heterozygosity (<20% heterozygosity; Males 3). We deemed these to be moderately sex-linked loci. Using Microsoft Excel, we used the 'descriptive statistics' function for Males 2 and Males 3 respectively to determine the mean heterozygosity and standard deviation for each. We then performed a two-sample t-test assuming unequal variances (based off results from the descriptive statistics) to determine if these two groups were significantly different from each other.

We plotted each of these 58 markers against each individual using the 'glPlot' function in the R package 'dartR'⁵³. Following this, using the 'dartR' functions 'gl.pcoa' and 'gl.pcoa.plot' respectively, we created a principal coordinates analysis plot to visualise the relatedness between females and the three groups of males identified based on these 58 perfectly and moderately sex-linked loci. Using the same loci, using the R package 'StAMPP'⁵⁴, we used the function 'stampNeisD' to create a distance matrix based upon the Nei genetic calculated between each individual based upon the 58 isolated loci. We represented this matrix visually through the creation of a heat map, using the function 'pheatmap' in the R package 'pheatmap'⁵⁵.

Random sex-linkage estimation. Due to the large number of loci provided by DARTseqTM, spurious sex association may be possible if analysing a small sample¹⁷. To ensure the sample size used ($n = 44$ – only confidently sexed samples used) was adequate to minimise the probability of random sex-linked markers being uncovered, the formula presented in Lambert *et al.*⁹ was utilised:

$$P_i = 0.5^n$$

Where P is the probability that any locus, i , is sex-linked by chance, and n is the sample size. Following this calculation, the result was multiplied by the number of quality SNP markers analysed to give an estimation of the number of random sex-linked loci produced through SNP analysis, with a similar process being undertaken for the PA markers.

Received: 14 June 2019; Accepted: 25 October 2019;

Published online: 12 November 2019

References

- Eggert, C. Sex determination: the amphibian models. *Reprod. Nutr. Dev.* **44**, 539–549 (2004).
- Maruo, K., Suda, M., Yokoyama, S., Oshima, Y. & Nakamura, M. Steroidogenic gene expression during sex determination in the frog *Rana rugosa*. *General and Comparative Endocrinology* **158**, 87–94 (2008).
- Matsuba, C., Miura, I. & Merilä, J. Disentangling genetic vs. environmental causes of sex determination in the common frog, *Rana temporaria*. *BMC genetics* **9**, 3 (2008).
- Valenzuela, N. & Lance, V. *Temperature-dependent sex determination in vertebrates*. (Smithsonian Books Washington, D. C., 2004).
- Witschi, E. Studies on sex differentiation and sex determination in amphibians. II. Sex reversal in female tadpoles of *Rana sylvatica* following the application of high temperature. *Journal of Experimental Zoology* **52**, 267–291 (1929).
- Yoshimoto, S. & Ito, M. A ZZ/ZW-type sex determination in *Xenopus laevis*. *The FEBS Journal* **278**, 1020–1026 (2011).
- Hayes, T. B. *et al.* Hermaphroditic, demasculinized frogs after exposure to the herbicide atrazine at low ecologically relevant doses. *Proceedings of the National Academy of Sciences* **99**, 5476–5480 (2002).
- Schmid, M., Steinlein, C. & Haaf, T. Chromosome banding in Amphibia: XXVII. DNA replication banding patterns in three anuran species with greatly differing genome sizes. *Cytogenetic and Genome Research* **101**, 54–61 (2003).
- Lambert, M. R., Skelly, D. K. & Ezaz, T. Sex-linked markers in the North American green frog (*Rana clamitans*) developed using DARTseq provide early insight into sex chromosome evolution. *BMC genomics* **17**, 844 (2016).

10. Miura, I. An Evolutionary Witness: the Frog *Rana rugosa* Underwent Change of Heterogametic Sex from XY Male to ZW Female. *Sexual Development* **1**, 323–331 (2007).
11. Wallace, H. & Wallace, B. M. N. Sex reversal of the newt *Triturus cristatus* reared at extreme temperatures. *International Journal of Developmental Biology* **44**, 807–810 (2000).
12. Pettersson, I., Arukwe, A., Lundstedt-Enkel, K., Mortensen, A. S. & Berg, C. Persistent sex-reversal and oviducal agenesis in adult *Xenopus (Silurana) tropicalis* frogs following larval exposure to the environmental pollutant ethynylestradiol. *Aquatic toxicology* **79**, 356–365 (2006).
13. Ohtani, H., Miura, I. & Ichikawa, Y. Effects of dibutyl phthalate as an environmental endocrine disruptor on gonadal sex differentiation of genetic males of the frog *Rana rugosa*. *Environmental Health Perspectives* **108**, 1189–1193 (2000).
14. Mahony, M. J. Heteromorphic sex chromosomes in the Australian frog *Crinia bilineata* (Anura: Myobatrachidae). *Genome* **34**, 334–337 (1991).
15. Guerrero, R. F., Kirkpatrick, M. & Perrin, N. Cryptic recombination in the ever-young sex chromosomes of Hyliid frogs. *Journal of Evolutionary Biology* **25**, 1947–1954 (2012).
16. Elinson, R. P. Inheritance and expression of a sex-linked enzyme in the frog, *Rana clamitans*. *Biochemical Genetics* **21**, 435–442 (1983).
17. Breltsford, A., Dufresnes, C. & Perrin, N. High-density sex-specific linkage maps of a European tree frog (*Hyla arborea*) identify the sex chromosome without information on offspring sex. *Heredity* **116**, 177 (2016).
18. Baloch, F. S. *et al.* A Whole Genome DArTseq and SNP Analysis for Genetic Diversity Assessment in Durum Wheat from Central Fertile Crescent. *PLOS ONE* **12**, 1–18 (2017).
19. Sansaloni, C. *et al.* Diversity Arrays Technology (DArT) and next-generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. *BMC Proceedings* **5**, P54 (2011).
20. Jaccoud, D., Peng, K., Feinstein, D. & Kilian, A. Diversity arrays: a solid state technology for sequence information independent genotyping. *Nucleic acids research* **29**, e25–e25 (2001).
21. Hill, P. L., Burrige, C. P., Ezaz, T. & Wapstra, E. Conservation of Sex-Linked Markers among Conspecific Populations of a Viviparous Skink, *Niveoscincus ocellatus*, Exhibiting Genetic and Temperature-Dependent Sex Determination. *Genome biology and evolution* **10**, 1079–1087 (2018).
22. Saidapur, S. K., Gramapurohit, N. P. & Shanbhag, B. A. Effect of sex steroids on gonadal differentiation and sex reversal in the frog, *Rana curtipipes*. *General and Comparative Endocrinology* **124**, 115–123 (2001).
23. Alho, J. S., Matsuba, C. & Merilä, J. Sex reversal and primary sex ratios in the common frog (*Rana temporaria*). *Molecular Ecology* **19**, 1763–1773 (2010).
24. Cotton, S. & Wedekind, C. Population Consequences of Environmental Sex Reversal. *Conservation Biology* **23**, 196–206 (2009).
25. Hayes, T. B. *et al.* Atrazine induces complete feminization and chemical castration in male African clawed frogs (*Xenopus laevis*). *Proceedings of the National Academy of Sciences* **107**, 4612–4617 (2010).
26. Perrin, N. Sex reversal: a fountain of youth for sex chromosomes? *Evolution: International Journal of Organic Evolution* **63**, 3043–3049 (2009).
27. Rodrigues, N., Studer, T., Dufresnes, C. & Perrin, N. Sex-chromosome recombination in common frogs brings water to the fountain-of-youth. *Molecular biology and evolution* **35**, 942–948 (2018).
28. Stöck, M. *et al.* Ever-Young Sex Chromosomes in European Tree Frogs. *PLOS Biology* **9**, 1–9 (2011).
29. Hero, J.-M., Gillespie, G., Cogger, H., Lemckert, F. & Robertson, P. *Litoria aurea*. *The IUCN Red List of Threatened Species* (2004).
30. Threlfall, C., Jolley, D., Evershed, N., Goldingay, R. & Buttemer, W. Do green and golden bell frogs *Litoria aurea* occupy habitats with fungicidal properties? *Australian Zoologist* **34**, 350–360 (2008).
31. Pyke, G. & White, A. Habitat requirements for the green and golden bell frog *Litoria aurea* (Anura: Hylidae). *Australian Zoologist* **30**, 224–232 (1996).
32. Burns, E. L. *Phylogeography, population history and conservation genetics of the endangered Green and Golden Bell Frog (Litoria aurea)*. (University of New South Wales, 2004).
33. Pyke, G. & White, A. A review of the biology of the green and golden bell frog *Litoria aurea*. *Australian Zoologist* **31**, 563–598 (2001).
34. Mahony, M. J. *et al.* Identifying conservation and research priorities in the face of uncertainty: a review of the threatened bell frog complex in eastern Australia. *Herpetological Conservation and Biology* **8**, 519–538 (2013).
35. Scheele, B. C. *et al.* After the epidemic: ongoing declines, stabilizations and recoveries in amphibians afflicted by chytridiomycosis. *Biological Conservation* **206**, 37–46 (2017).
36. Goldingay, R. The Green and Golden Bell Frog *Litoria aurea*—from riches to ruins: conservation of a formerly common species. *Australian Zoologist* **30**, 248–256 (1996).
37. Burns, E. L., Eldridge, M. D. B. & Houlden, B. A. Microsatellite variation and population structure in a declining Australian Hyliid *Litoria aurea*. *Molecular Ecology* **13**, 1745–1757 (2004).
38. Goldingay, R. & Newell, D. Population estimation of the green and golden bell frog *Litoria aurea* at Port Kembla. *Australian Zoologist* **33**, 210–216 (2005).
39. Greer, A. & Byrne, M. Sex ratio and frequency of osteological abnormalities in the Australian hyliid frog *Litoria aurea* from two apparently unpolluted localities in Sydney, New South Wales. *Australian Zoologist* **30**, 43–47 (1995).
40. Pickett, E. J. *et al.* Estimates of sex ratio require the incorporation of unequal catchability between sexes. *Wildlife Research* **39**, 350–354 (2012).
41. Horn, A. *et al.* Sex-Chromosome Homomorphy in Palearctic Tree Frogs Results from Both Turnovers and X–Y Recombination. *Molecular Biology and Evolution* **32**, 2328–2337 (2015).
42. Hendricks, S. *et al.* Conservation implications of limited genetic diversity and population structure in Tasmanian devils (*Sarcophilus harrisii*). *Conservation Genetics* **18**, 977–982 (2017).
43. Ndjiondjop, M.-N. *et al.* Genetic variation and population structure of *Oryza glaberrima* and development of a mini-core collection using DArTseq. *Frontiers in plant science* **8**, 1748 (2017).
44. Ren, R. *et al.* Construction of a high-density DArTseq SNP-based genetic map and identification of genomic regions with segregation distortion in a genetic population derived from a cross between feral and cultivated-type watermelon. *Molecular Genetics and Genomics* **290**, 1457–1470 (2015).
45. Steane, D. A. *et al.* Genome-wide scans reveal cryptic population structure in a dry-adapted eucalypt. *Tree Genetics & Genomes* **11**, 33 (2015).
46. Ezaz, T., Stiglec, R., Veyrunes, F. & Graves, J. A. M. Relationships between Vertebrate ZW and XY Sex Chromosome Systems. *Current Biology* **16**, R736–R743 (2006).
47. Wright, A. E., Dean, R., Zimmer, F. & Mank, J. E. How to make a sex chromosome. *Nature Communications* **7**, 12087 (2016).
48. Matsuba, C., Alho, J. S. & Merilä, J. Recombination rate between sex chromosomes depends on phenotypic sex in the common frog. *Evolution: International Journal of Organic Evolution* **64**, 3634–3637 (2010).
49. Reich, J. K. A., Nichols, S. J., Maher, W. A. & Kefford, B. J. Is metal flocculation from mining activities a previously overlooked mechanism for impairing freshwater ecosystems? *Science of The Total Environment* (2019).
50. Kilian, A. *et al.* Diversity arrays technology: a generic genome profiling technology on open platforms. In *Data production and analysis in population genomics* 67–89 (Springer, 2012).
51. RCore Team. R: A Language and Environment for Statistical Computing. (2013).

52. Adamack, A. T. & Gruber, B. PopGenReport: simplifying basic population genetic analyses in R. *Methods in Ecology and Evolution* **5**, 384–387 (2014).
53. Gruber, B. & Georges, A. dartR: Importing and Analysing SNP and Silicodart Data Generated by Genome-Wide Restriction Fragment Analysis. (2019).
54. Pembleton, L. W., Cogan, N. O. I. & Forster, J. W. StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources* **13**, 946–952 (2013).
55. Kolde, R. pheatmap: Pretty Heatmaps (2019).
56. Hillis, D. M. & Green, D. M. Evolutionary changes of heterogametic sex in the phylogenetic history of amphibians. *Journal of evolutionary Biology* **3**, 49–64 (1990).
57. Schmid, M. & Steinlein, C. Sex chromosomes, sex-linked genes, and sex determination in the vertebrate class Amphibia. In *Genes and Mechanisms in Vertebrate Sex Determination* 143–176 (Springer, 2001).
58. Sarre, S. D., Ezaz, T. & Georges, A. Transitions between sex-determining systems in reptiles and amphibians. *Annual review of genomics and human genetics* **12**, 391–406 (2011).
59. Miura, I. Sex determination and sex chromosomes in amphibia. *Sexual Development* **11**, 298–306 (2017).
60. Ito, M. Sex Determination and Differentiation in Frogs. in *Reproductive and Developmental Strategies* 349–366 (Springer, 2018).
61. Jetz, W. & Pyron, R. A. The interplay of past diversification and evolutionary isolation with present imperilment across the amphibian tree of life. *Nature ecology & evolution* **2**, 850 (2018).

Acknowledgements

JS is an undergraduate student at the University of Canberra and performed this project as part of his 3rd Year Research Project in Applied Science unit.

Author contributions

Study conception and design: T.E., B.J.K. and B.C.S.; sample collection: B.C.S.; data analysis: J.S., F.S. & T.E.; writing first draft: J.S.; all authors contributed in writing, reviewing & editing the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-019-52970-4>.

Correspondence and requests for materials should be addressed to J.S. or T.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019