



# OPEN Turnover of sex chromosomes in the Lake Tanganyika cichlid tribe Tropheini (Teleostei: *Cichlidae*)

Kristen A. Behrens<sup>1</sup>✉, Holger Zimmermann<sup>2,3</sup>, Radim Blažek<sup>2</sup>, Martin Reichard<sup>2,4</sup>, Stephan Koblmüller<sup>3</sup> & Thomas D. Kocher<sup>1</sup>

Sex chromosome replacement is frequent in many vertebrate clades, including fish, frogs, and lizards. In order to understand the mechanisms responsible for sex chromosome turnover and the early stages of sex chromosome divergence, it is necessary to study lineages with recently evolved sex chromosomes. Here we examine sex chromosome evolution in a group of African cichlid fishes (tribe Tropheini) which began to diverge from one another less than 4 MYA. We have evidence for a previously unknown sex chromosome system, and preliminary indications of several additional systems not previously reported in this group. We find a high frequency of sex chromosome turnover and estimate a minimum of 14 turnovers in this tribe. We date the origin of the most common sex determining system in this tribe (XY-LG5/19) near the base of one of two major sub-clades of this tribe, about 3.4 MY ago. Finally, we observe variation in the size of one sex-determining region that suggests independent evolution of evolutionary strata in species with a shared sex-determination system. Our results illuminate the rapid rate of sex chromosome turnover in the tribe Tropheini and set the stage for further studies of the dynamics of sex chromosome evolution in this group.

Early methods for characterizing sex chromosomes used cytological approaches that were only capable of identifying highly differentiated sex chromosomes<sup>1–4</sup>. Sex chromosome systems were therefore originally classified by the morphology of the sex chromosome pair. The availability of inexpensive genome sequencing techniques has allowed the identification of sex chromosome pairs that do not show overt morphological differences. Often these are young sex chromosomes that have not yet diverged enough to be considered heteromorphic. The same XY or ZW terminology can be used for these homomorphic sex chromosomes. Male heterogametic systems are referred to as XY, and female heterogametic systems are referred to as ZW.

New sex determining alleles arising in a population may initiate a transition to a new sex chromosome<sup>5–8</sup>. The rate of sex chromosome turnover is highly variable among lineages. With only a few exceptions, such as the spiny rat<sup>9</sup>, there have been no turnovers of sex chromosomes in therian mammals or birds in the last ~100–180 million years<sup>10,11</sup>. In stark contrast, groups such as fishes and squamate reptiles show a high frequency of such turnovers<sup>12–14</sup>. Why some organisms experience high rates of turnover while others do not remains an unanswered question<sup>15</sup>. Hypotheses about the evolutionary mechanisms promoting sex chromosome turnover include sexually antagonistic selection<sup>16,17</sup>, accumulation of deleterious mutations<sup>18,19</sup>, and genetic drift<sup>20</sup>. Distinguishing these mechanisms empirically has been challenging and will require the discovery and characterization of sex chromosomes in additional species.

In many species, the chromosome region immediately adjacent to the sex-determining locus shows a reduction in recombination. This sex-linked region may expand over time through further recombination suppression, often creating ‘evolutionary strata’ exhibiting various levels of differentiation between the sex chromosomes. A predominant theory for the development of evolutionary strata is that each expansion associates advantageous sexually antagonistic variants with the sex-determining gene<sup>8</sup>. However, this theory has been difficult to prove empirically<sup>21–23</sup> and has been challenged<sup>24</sup>. The mechanisms by which recombination becomes suppressed is also the subject of debate<sup>23</sup>. One recent model suggests that inversions expand in a stepwise manner from the sex-determining region because they shelter deleterious mutations<sup>25</sup>. Another theory proposes that co-evolution of cis and trans regulators of gene expression drives inversions on the Y<sup>26</sup>. All of these models should cause

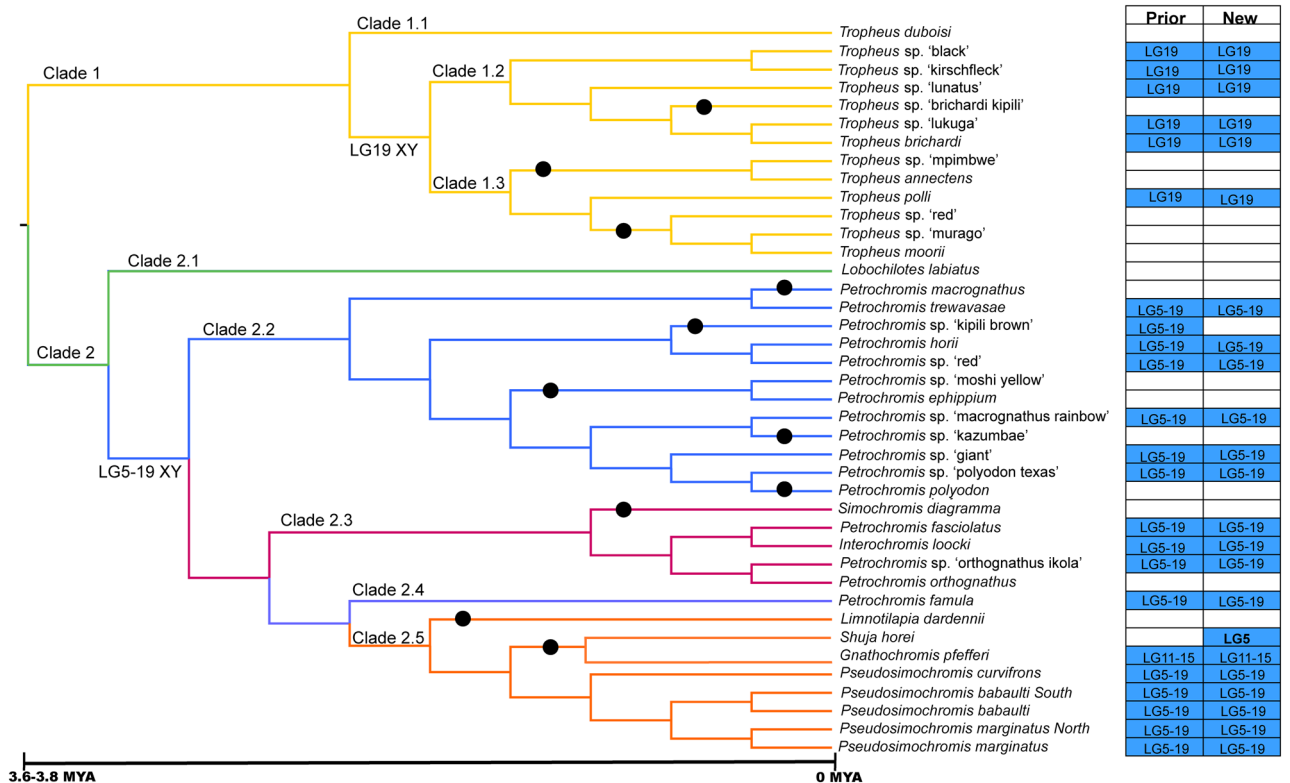
<sup>1</sup>Department of Biology, University of Maryland, College Park, MD 20742, USA. <sup>2</sup>Institute of Vertebrate Biology, Czech Academy of Sciences, Květná 8, 603 00 Brno, Czech Republic. <sup>3</sup>Institute of Biology, University of Graz, Universitätsplatz 2, 8010 Graz, Austria. <sup>4</sup>Department of Ecology and Vertebrate Zoology, University of Łódź, Łódź, Poland. ✉email: kbehen2@umd.edu

regions adjacent to the sex-determining region to cease recombination at different times, producing evolutionary strata<sup>27–31</sup>.

A large portion of the work on sex chromosome evolution has been conducted in mammals and birds. These two groups have ancient heteromorphic sex chromosomes with evolutionary strata that are readily apparent<sup>10,11</sup>. However, these systems are not ideal for studying the mechanisms of sex chromosome turnover or the early stages in the development of evolutionary strata. To understand the mechanisms of sex chromosome evolution we need to study lineages with a high rate of recent sex chromosome turnover. While highly divergent evolutionary strata have been identified in some fish, such as sticklebacks<sup>32–34</sup>, sex chromosomes in most fish species are homomorphic<sup>35–37</sup>. This homomorphy, and the high frequency at which sex chromosome turnover occurs, makes fishes ideal for studying sex chromosome evolution<sup>36,38</sup>. The direction of turnover, XY to ZW versus ZW to XY, is thought to be skewed in favor of the latter in fishes<sup>12</sup>. Instances of multiple sex chromosomes segregating within teleost species have been reviewed recently<sup>39</sup>, but the effects of polygenic systems on estimates of sex chromosome turnover have yet to be explored. Finally, a role for chromosome fusions in sex chromosome evolution has also been proposed in fishes<sup>40</sup>, and may contribute to reproductive isolation<sup>38</sup>.

We have focused on East African cichlid fishes (*Cichlidae*), a group known for its rapid radiation that generated species-rich lineages<sup>41</sup>. The relatively recent evolution of these species makes them ideal for studying evolutionary mechanisms. The high species density allows for a large number of comparisons among closely related species. Many sex chromosome turnovers have been detected in this clade, and at last count involved 12 of the 23 cichlid chromosomes<sup>15,42,43</sup>. The rate of sex chromosome turnover in this group has been calculated to be at least 0.186 turnovers per million years<sup>42</sup>.

The Tropheini, one of the thirteen cichlid tribes in Lake Tanganyika<sup>44</sup>, is particularly well-suited for such studies. The Tropheini consist of ~40 species grouped into eight genera<sup>45–47</sup> which shared a common ancestor ~3.6–3.8 million years ago<sup>44,48</sup>. Genera in the Tropheini are defined by their distinct jaw morphologies<sup>49</sup>, and previous phylogenetic analysis has identified several subclades within the tribe<sup>44,48</sup> (Fig. 1). Clade 1 consists of 13 species of *Tropheus* divided into 3 subclades. Clade 1.1 contains a single species, *T. duboisi*, which is sister to the rest of Clade 1. Clades 1.2 and 1.3 each contain 6 species. Clade 2 contains the 27 remaining species that can be grouped into five sub-clades. Clade 2.1 consists of the single species *Lobochilotes labiatus*, which is an outgroup to the other four clades. Clade 2.2 contains 12 species of *Petrochromis*. Clade 2.3 contains *Simochromis diagramma*, *Interochromis loocki* and 3 more species of *Petrochromis*. Clade 2.4 consists of a single species, *Petrochromis famula* and Clade 2.5 includes *Limnotilapia dardennii*, *Shuja horei*, *Gnathochromis pfefferi* and five species of *Pseudosimochromis*.



**Figure 1.** Phylogeny of Tropheini with previous sex chromosome calls and proposed sex chromosome systems. Some systems feature chromosomal fusions or multiple sex chromosomes. Cells in blue indicate an XY system, cells in orange indicate a ZW system. Bolded text indicates disagreement between previous call and current call. Black circles indicate loss of sex chromosome common to that clade. Topology adapted from Ronco et al.<sup>44</sup> and approximate time scale is from Singh et al.<sup>48</sup>.

Our previous study identified an XY system on LG19 in *Tropheus* sp. ‘black’<sup>43</sup>. A more recent survey of 40 taxa within the tribe identified three distinct sex chromosome systems, two of which are shared by multiple species<sup>42</sup>. Species of the genus *Tropheus* (Clade 1) were found to segregate the XY system on LG19, while species of *Petrochromis* and *Pseudosimochromis* (Clade 2) segregated an XY system involving an apparent fusion of LG5 and LG19. *Gnathochromis pfefferi* (Clade 2.4) carried an XY system involving an apparent fusion of LG11 and LG15<sup>42</sup> (Fig. 2).

Here we use new data from pooled sequencing of males and females in several species, together with a reanalysis of previously published sequences of single male and female individuals<sup>42</sup>, to identify new sex chromosome systems in this tribe. We quantify the high rate of sex chromosome turnover and also describe the independent development of evolutionary strata on LG19 among species. Our study provides new insights into the early stages of sex chromosome evolution in this group.

## Methods

Cichlid sex chromosomes are usually at an early stage of differentiation. This means the allele frequency differences between the sex chromosomes are small and often present at a relatively small number of sites. Furthermore, the characteristics of the signal changes with the age of the system. At first the main signal will be heterozygosity of a few sites in the heterogametic sex. However, with continued divergence, some sequences are lost from the Y (or W) chromosome, leading to apparently contradictory signals. For example, loss of a sequence from the Y chromosome makes males hemizygous at some sites, so that heterozygosity of these sites on the X falsely suggests a ZW system. Many signals typically quantified for older sex chromosome systems (e.g. differences in base composition or sequence coverage) are not present in the very recently evolved systems in cichlids.

Identification of these signals is further complicated by the small effective size of many cichlid populations. A small number of large haplotype blocks in these populations can lead to identification of spurious sex-associated blocks when small numbers of male and female individuals are sampled. Even when a larger number of individuals are sampled, the small  $N_e$  can lead to high background levels of allele frequency differences between the male and female samples.

Finally, there is a possibility that polygenic systems segregate in these species, arising from either the high rates of sex chromosome turnover, or the maintenance of balanced polymorphisms over hundreds of thousands of years. Sampling for sex chromosome surveys is generally not designed to identify multiple sex chromosome systems segregating in a population. With these challenges in mind, we used a variety of approaches to extract the maximum information from the limited data available.

## DNA samples and sequencing

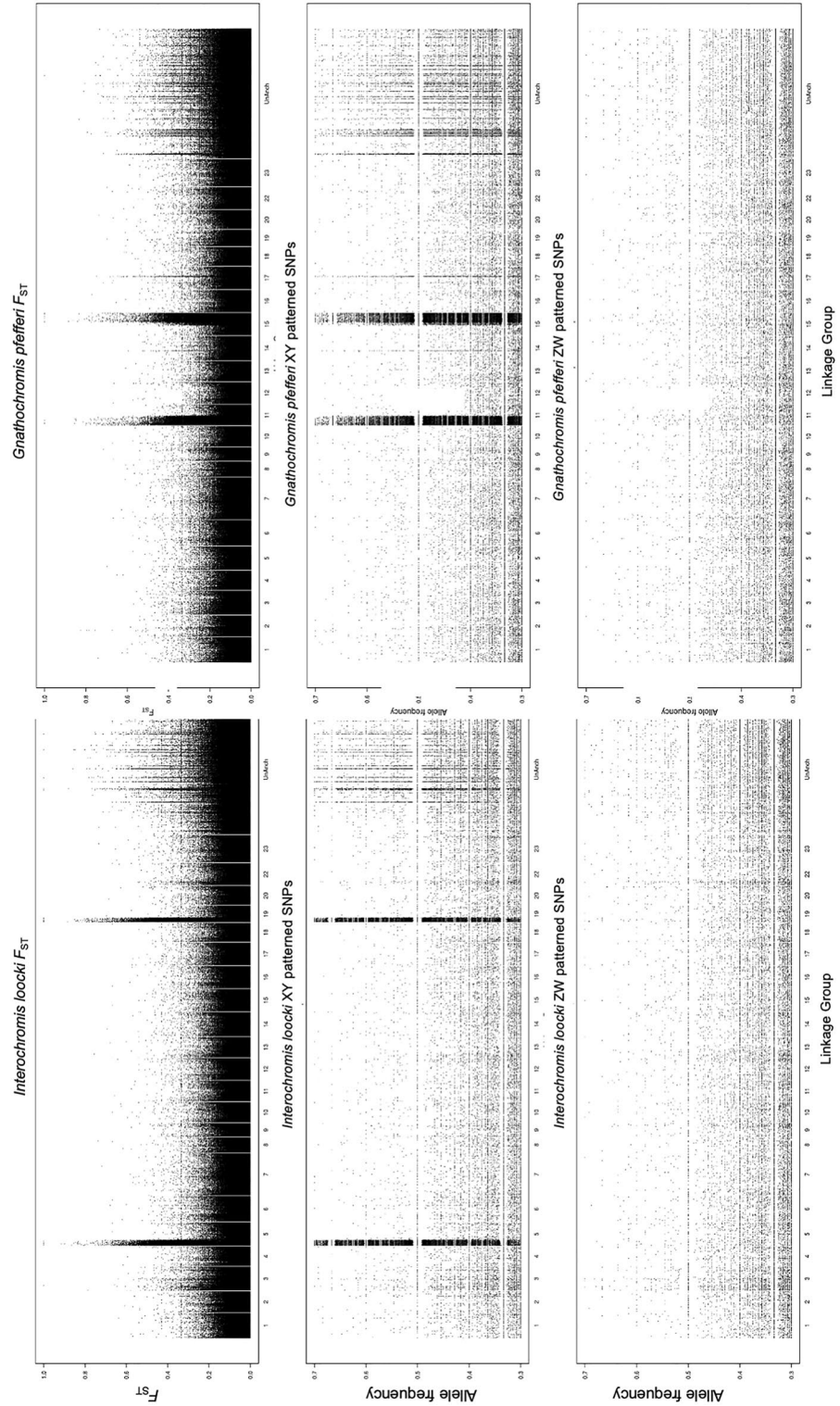
We obtained new sequencing data from pooled males and females from four species collected in 2019 near Kalambo Lodge, Zambia near the southern end of the lake (-8.623°N, 31.200°E): 25 males and 23 females of *Pseudosimochromis babaulti*, 23 males and 25 females of *Gnathochromis pfefferi*, 24 males and 21 females of *Interchromis loocki*, and 25 males and 25 females of *Shuja horei* (*Ctenochromis horei*)<sup>50</sup>. We also analyzed a full-sib family of 30 males and 28 females of *Simochromis diagramma* reared at U. Graz. Finally, we reanalyzed data from 30 males and 24 females of *Tropheus* sp. ‘black’ that were wild caught near Ikola, Tanzania<sup>43</sup>. Animal experiments were conducted in accordance with the Guide for Care and Use of Laboratory Animals and reported in accordance with ARRIVE guidelines<sup>51</sup>. All animal use was approved under Zambian research permit K-4335/18 KA/K.48/18 (S. Koblmüller) and animal care protocols R-OCT-19-48 (U. Maryland) and BMWFW-66.007/004-WF/V/3b/2016 (U. Graz). This study was carried out with the approval of the ethics committee of the University of Graz (permit number GZ. 39/115/63 ex 2022/23).

DNA was extracted from fin clips by phenol–chloroform extraction using phase-lock gel tubes (5Prime, Gaithersburg, MD). DNA concentrations were quantified by fluorescence spectroscopy using a Quant-iT PicoGreen assay (ThermoFisher, Waltham, MA). Equimolar amounts of DNA from each individual were then pooled by sex for each species. Sequencing libraries were constructed, and 150 bp paired-end DNA sequencing was performed on a NovaSeq6000 S4 (Illumina, San Diego CA) by Novogene US (Davis, CA).

Sequence data from previous studies were retrieved from the NCBI Short Read Archive and GenBank. This included whole genome sequence data from single males and females for the species in the tribe *Tropheini*, analyzed by El Taher et al.<sup>42</sup>. We also examined transcriptome data for 9 species, which were also included in that study (Suppl Table S1).

## Sex-specific SNP analysis

The main basis of our analyses is the identification and analysis of sex-specific SNPs. These SNPs were identified following the methods described in Behrens et al.<sup>52</sup> using the pipeline developed by Gammerdinger et al.<sup>43</sup>. Previously reported code from that study is available (<https://github.com/Gammerdinger/sex-SNP-finder>). Briefly, the sequence reads were aligned with BWA version 0.7.12 using the default parameters along with read group labels<sup>53</sup>. We aligned all samples to the closest high-quality reference assembly, the Malawi zebra (*Maylandia zebra*—UMD2a, RefSeq GCF\_000238955.4)<sup>54</sup>. In some cases, some of the sex-specific signal mapped to unanchored contigs of the Malawi zebra assembly, so we remapped the reads to the more contiguous assembly of Nile tilapia (*Oreochromis niloticus*—UMD\_NMBU, RefSeq GCF\_001858055.2)<sup>55</sup>. At each variable nucleotide site we calculated the  $F_{ST}$  statistic between the populations of male and female sequence reads. The resulting  $F_{ST}$  plots provide a first indication of the differentiation between male and female genomes. We further identified XY- and ZW- patterned SNPs as SNPs that were fixed (frequency less than 0.1) in one sex and polymorphic (frequency between 0.3 and 0.7) in the other sex. Separate plots of the allele frequency of XY- and ZW- patterned SNPs often allowed determination of the type of heterogametic system segregating (XY or ZW).



**Figure 2.** Pairwise male x female  $F_{ST}$  and sex specific allele frequency in pool-seq data for *Interchromis loocki* and *Gnathochromis pfefferi*. The x-axis is position on the *Metriaclicma zebra* (UMD2a) reference assembly.

Next we used Bedtools<sup>56</sup> *make windows* and *coverage* to calculate the density of sex-patterned SNPs in 100kbp windows across the genome. We identified the top 1% of windows (~78 of 7,800 anchored windows) with the highest number of sex-patterned SNPs using the methodology described in Kocher et al.<sup>57</sup>. The  $\log_2(XY:ZW)$  ratio of SNP density was then calculated for each window<sup>58</sup>. A Kruskal–Wallis (KW) test on the ranked data was conducted in R (v.2023.03.0 + 386) using `kruskal.test` from the *stats* package to determine if the log ratio differed among chromosomes<sup>59</sup>. If the differences were statistically significant, the Dunn's test from the *rstatix* R package was conducted post-hoc to determine which chromosomes differed significantly from one another with Benjamini–Hochberg correction for multiple tests. Regions of elevated sex-specific SNP density were visualized in IGV<sup>60</sup> to identify candidate sex determining genes.

### k-mer analysis

To identify sex-linked variants shared by species with the same sex chromosome system we conducted a k-mer analysis. Following our previous work, we used using Jellyfish v. 2.2.7<sup>61</sup> to tabulate the frequency of k-mers of length 22 bases<sup>52</sup>. Code from that study is available (<https://github.com/KristenBehrens/K-mer-scripts>). Briefly, for each shared sex chromosome system, a core region was identified using a python script that compares k-mers from each species and generates a list of k-mers shared by all species with that system. These k-mers were then aligned to the *M. zebra* reference genome (M\_zebra\_UMD2a) using BLAST to determine the core region of differentiation. Samtools v. 1.10 was used for any necessary file format conversions, sorting, or indexing. This analysis was conducted for species with the XY-linkage group (LG) 5/19 system and to compare the XY-LG5 system between *Shuja horei* and *Cyprichromis pavo*.

## Results

### Identification of sex chromosomes from pooled sequence data

In the new pool-seq data for five species reported here, we confirmed three systems previously identified from individual sequence data<sup>42</sup>. The sex chromosomes confirmed were the XY-LG5/19 systems in *Interochromis loocki* and *Pseudosimochromis babaulti*, and the XY-LG11/15 system in *Gnathochromis pfefferi* (Fig. 2).

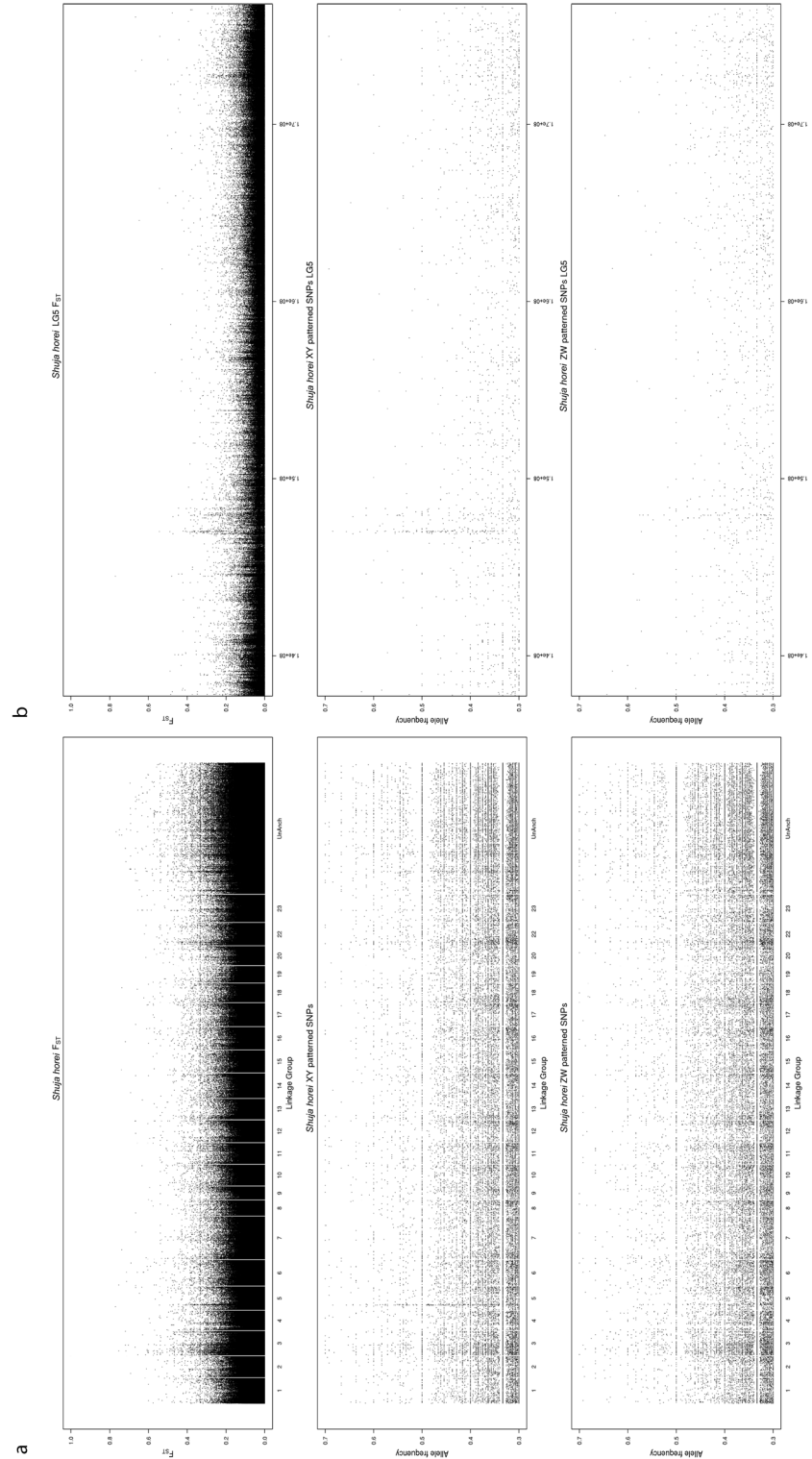
The  $F_{ST}$  plots of the pool-seq data for *S. horei* (Clade 2.5) show no evidence of the XY-LG5/19 system common in the rest of Clade 2, but do suggest an XY system on LG5 (Fig. 3). This narrow peak (~1.5 Mb wide) is located at 7.8–7.9 Mbp, which is the top window of differentiation containing 80 XY-patterned SNPs. However, this narrow peak did not yield a significant Kruskal–Wallis test for heterogeneity among chromosomes for either the single individual data or the pooled sequencing data (Suppl Table S2). The peak (7.8–7.9 Mbp) encompasses two genes, *ncoa3* and *id1*, both of which might be considered candidate genes for male sex determination. *ncoa3* (*src-3*, *ab1*) is a member of the p160 steroid receptor coactivator family that regulates transcription<sup>62,63</sup>. This gene has been found to control cell migration in the ovary in both *Drosophila* and humans<sup>64</sup>. *id1* is an inhibitor of differentiation protein that acts during embryogenesis to promote cell growth<sup>65</sup>. This gene is mediated upstream by the TGF-beta pathway<sup>66</sup> which has been implicated in sex determination in cichlids and other fishes<sup>67–69</sup>. However, the highest density of male-specific SNPs fell in the region between *ncoa3* and *id1*, which has no annotated gene in the *M. zebra* assembly. The differentiation on LG5 in *S. horei* is not related to the XY system on LG5/19 seen throughout Clade 2, as the peaks on LG5 are in different regions of the chromosome.

The  $F_{ST}$  plot for the pooled sample of *Simochromis diagramma* shows no evidence of the XY-LG5/19 system common in Clade 2. Instead, we observe strong differentiation on LG7 (Suppl Fig. S1). The top 1% windows include 45 ZW and 34 XY windows on LG7. Although the KW test for heterogeneity among chromosomes is significant, the Dunn's tests do not suggest unusual differentiation of LG7 relative to other chromosomes, but instead suggest differentiation of LG2, LG8 and LG20. In contrast, the top 1% windows in the comparison of a single male and female shows 22 XY windows on LG8 and 11 XY windows on LG16. The KW test for heterogeneity is also significant and the Dunn's test suggests unusual differentiation of LG8, LG16 and LG18. The multiple signals, and the differences between the individual samples and the pooled sequencing of the family suggest this species may be segregating multiple sex chromosome systems.

### Confirmation of sex chromosomes from individual sequence data

We also reanalyzed a published dataset consisting of whole genome sequence reads from a single male and single female for most species of *Tropheini* and were able to confirm many of the systems identified by El Taher<sup>42</sup>. These include an XY-LG19 system predominant in the *Tropheus* of Clade 1, which we confirmed in *Tropheus* sp. 'black', *T. sp.* 'kirschfleck', *T. sp.* 'lunatus', *T. sp.* 'lukuga', *T. brichardi* and *T. polli*. Each of these species showed strong statistical support for the XY-LG19 system. The KW tests for heterogeneity were significant, the Dunn's test for differences among chromosomes implicated LG19, and nearly all of the top 1% windows were located on LG19 (Suppl Table S2). Equally important, we confirmed the absence of the XY-LG19 system in many species of Clade 1. We found no evidence of XY-LG19 in *T. sp.* 'brichardi kipili', *T. sp.* 'mpimbwe', *T. annectens*, *T. sp.* 'red', *T. sp.* 'murago' or *T. moorii*. The strong XY-LG19 signals we observed in the sister taxa of each of these species allows the strong inference that the XY-LG19 system is not present in our samples of these species. We infer 3 independent losses of the XY-LG19 system in Clade 1 (Fig. 1).

We also confirmed an XY-LG5/19 system in many species from Clade 2, including *Petrochromis trewavasae*, *P. horii*, *P. sp.* 'red', *P. sp.* 'macrognathus rainbow', *P. sp.* 'giant', *Interochromis loocki*, *P. sp.* 'orthognathus Ikola', *P. famula*, *Pseudosimochromis curvifrons*, *P. babaulti* South, *P. babaulti*, *P. marginatus* North and *P. marginatus*. These species showed evidence for an XY-LG5/19 system across multiple tests (Suppl Table S2). We agree with the conclusion of El Taher et al.<sup>42</sup> that the LG19 system seen in the Clade 1 is different from the LG19 system in Clade 2. These two systems involve different regions of LG19, and their haplotypes do not group together phylogenetically<sup>42</sup>.



**Figure 3.** Pairwise male x female  $F_{ST}$  and sex specific allele frequency for *S. horei*, (a) whole genome, (b) LG5. The x-axis is position on the *Metriaclima zebra* (UMD2a) reference assembly.

We found less support for the XY-LG5/19 system in several other species (Suppl Table S2). *Petrochromis* sp. 'polyodon Texas' shows top 1% window support for XY-LG5/19, and KW heterogeneity among chromosomes. The Dunn's tests support differentiation of LG5 and LG19, but also LG8 and LG18. In the absence of top 1% window support for LG8 and LG18, we conclude this species is segregating the XY-LG5/19 system. *Petrochromis fasciolatus* shows KW heterogeneity and top 1% support for 5/19 but the Dunn's results are scattered. Again, we conclude this species segregates the XY-LG5/19, and the scattered Dunn's test results arise from artifacts of the small sample. Finally, we were unable to confirm the claim of an XY-LG5/19 system in *Petrochromis* sp. 'kipili brown'<sup>42</sup>. Our analysis shows KW heterogeneity, but the top 1% windows suggest XY signal for LG4, LG11, LG20, LG22. We are unable to identify a sex chromosome system in this species because there are many significant Dunn's tests, again suggesting artifacts of small sample size or identity by descent in a small population. For the remaining 9 species in Clade 2 we could find no support for an XY-LG5/19 system, possibly indicating that the XY-LG5/19 system has been lost and replaced by a new sex chromosome. Such replacements may have occurred as many as 8 times during the evolution of Clade 2 (Fig. 1).

We were also able to confirm the XY-LG11/15 system in *G. pfefferi*, which is the second instance of sex chromosome fusion in this tribe. The X and Y chromosomes are strongly differentiated, as seen in the  $F_{ST}$  plot (Fig. 2), and all of the top 1% windows are XY windows falling on LG11 and LG15 (Suppl Table S2). This high level of differentiation suggests that this sex chromosome is similar in age to the XY-LG5/19 system in Clade 2.

### Possible new systems identified in single individual data

We found provisional support for several additional systems in Clade 1. We mention them so that they may be examined with additional data in the future. None of these species showed evidence of the LG19 system found in other species of Clade 1. *Tropheus duboisi* has no obvious signal in the  $F_{ST}$  plot. The KW test for heterogeneity is significant, and Dunn's tests identify LG9 as distinct. There are 4 top 1% windows that suggest an XY system on LG9, but the overall ratio of XY to ZW SNPs on LG9 is strongly biased toward ZW. *Tropheus* sp. 'brichardi kipili' shows no obvious signals in the  $F_{ST}$  plot, and the KW test is not significant. We are unable to identify a sex chromosome for this species. *Tropheus* sp. 'mpimbwe' show no obvious signal in the  $F_{ST}$  plot. The KW test is moderately significant, and the Dunn's tests identify LG8 as distinct. However, there are no top 1% windows on LG8. *Tropheus annectens* shows a narrow peak of  $F_{ST}$  on LG8 suggestive of a ZW system. The KW test is significant, but the Dunn's tests suggest that LG14 and LG23 are distinct from the other chromosomes. The top 1% windows include one ZW window on LG8, two ZW windows on LG14, and one XY and six ZW windows on LG23. The  $F_{ST}$  plot for *Tropheus* sp. 'red' is unremarkable, and the KW test is not significant. The  $F_{ST}$  plot for *Tropheus* sp. 'murago' suggests a possible XY system on LG6, but the KW test is insignificant and there is no support in the top 1% windows. The  $F_{ST}$  plot for *Tropheus moorii* is unremarkable, the KW test is barely significant, and there is no obvious system indicated in the top 1% windows.

We also found provisional support for additional systems in Clade 2. None of these species showed evidence of the XY-LG5/19 system found in other species of Clade 2. The  $F_{ST}$  plots for *Petrochromis macrognathus* do not show obvious sex signals, and the KW and Dunn's tests are only moderately significant. *Petrochromis* sp. 'moshi yellow' shows some possible  $F_{ST}$  peaks, and the KW test is significant. However, the Dunn's results are weak and scattered, and the top 1% windows do not support any particular chromosome. The  $F_{ST}$  plot for *Petrochromis ephippium* is unremarkable, but the KW test is significant and the Dunn's tests highlight LG3 and LG4. The top 1% windows suggest a ZW system on LG4. The  $F_{ST}$  plot for *Petrochromis* sp. 'kazumbae' is also unremarkable, but the KW test is significant, and the Dunn's tests highlight LG23. *Petrochromis polyodon* shows no obvious  $F_{ST}$  peaks, and the KW test is moderately significant. The Dunn's tests do not highlight any chromosome. The evidence for multiple sex systems segregating in *Simochromis diagramma* was discussed in detail in the section on pooled sequencing results above. The  $F_{ST}$  plot for *Petrochromis orthognathus* is unremarkable, but the KW is significant, and the Dunn's tests suggest a system on LG13. There is a single top 1% window on LG13 that contains 18 ZW-patterned SNPs. The  $F_{ST}$  plot for *Limnotilapia dardennii* is unremarkable, the KW test is moderately significant but only a handful of the Dunn's test comparisons show weak significance. The evidence for an XY-LG5 system in *Shuja horei* was discussed in detail in the section on pooled sequencing results above.

Discovery of sex chromosomes using the single individual data was impeded by high levels of background noise in the sex-specific SNP analyses (e.g. *Petrochromis* sp. 'giant' in Suppl Fig. S2). Single individual data is not ideal for identifying sex chromosomes regardless of analytical methodology. We hope that calling attention to these possible sex chromosome signals will motivate further research.

### Origins of the sex chromosomes in Clades 1 and 2

Clades 1 and 2 both include species with an XY system that involves LG19. In order to determine whether these species share elements of an ancestral sex determination system, we compared male-specific k-mers between Clade 1 and Clade 2.2. We found only 10 shared k-mers that aligned to the reference assembly. Of these, eight k-mers representing 2 SNPs aligned to LG19, one aligned to LG2, and the last aligned to LG6. Of the k-mers associated with LG19, all fell in the same region (~10.1 Mbp) in the gene *kiaa1109*, well within the sex-linked region in Clade 1, but just outside the right-most boundary of the sex-linked region in Clade 2.2. Thus, it is likely not a candidate sex determining gene, and it is unlikely that this handful of shared k-mers, representing just 2 SNPs, are evidence for a shared origin of LG19.

The species in Clade 1 with an XY-LG19 system share 320,209 k-mers. Of these, 245,240 (76%) are Y-specific k-mers on LG19 representing approximately 10,000 SNPs. The large number of shared SNPs on LG19 are indicative of an old sex chromosome system with a relatively long period of shared ancestry before the current species diverged. The SNPs are distributed relatively evenly over a large portion of LG19, and so do not help to localize a candidate sex determination gene.

We also quantified shared k-mers for species with the XY-LG5/19 system in Clade 2. The four subclades (2.2–2.5) of Clade 2 share only 20 Y-specific k-mers, of which 18 aligned to LG5. The majority of these k-mers (14) correspond to a single SNP that maps to ~5.127 Mbp within an intron of the gene *bin2b*, which has no obvious function in sex determination. The other four k-mers on LG5 represent a single SNP mapped to ~17.57 Mbp in a region that contained no annotations. The absence of a large number of SNPs shared across clades suggests that the radiation of species occurred very quickly after the emergence of the new sex determiner. Therefore, most of the sequence differentiation in the sex-determining region has occurred independently in each lineage.

The distantly related tribe Cyprichromini also has an XY system on LG5<sup>52</sup>. We counted k-mers shared between *Cyprichromis pavo* and *P. trewavasae* (XY-LG5/19). These two species shared slightly more k-mers on LG5 and LG19 than on the other chromosomes (Suppl Table S3). However, given the phylogenetic distance between these clades, these similarities are unlikely to represent shared origin of the sex chromosome systems in these two tribes. A comparison between *Cyprichromis pavo* and the recently evolved XY-LG5 system in *S. horei* did not show an excess of shared k-mers on LG5.

### Evolutionary strata

Sex chromosome systems shared by multiple species offer the opportunity to study the early divergence of sex-linked regions into evolutionary strata. We characterized male-specific SNP density per 100 kb window and found interspecific variation in the extent and levels of sequence differentiation of the sex determining regions that suggests the presence of multiple evolutionary strata.

Six species in Clade 1 share the XY-LG19 system. Although the number of XY-patterned SNPs varies among species, there are no obvious differences in the pattern or chromosomal extent of the differentiation (Suppl Fig. S3). The variation in the number of sex-patterned SNPs likely reflect differences in  $N_e$  among species.

Sixteen species in Clade 2 share the XY-LG5/19 system. The first 9 Mb of LG5 shows differentiation in all species (Suppl Fig. S4). The species in Clades 2.2, 2.3 and 2.4 show no further expansion from the core 9 Mb. In Clade 2.5, *P. babaulti South* does not appear to have expanded beyond the 9 Mb core, but the remaining 3 species share a stratum that extends and additional ~4 Mb. *Pseudosimochromis babaulti* appears to have an additional stratum that extends and additional ~4 Mb beyond the core.

The expansion of evolutionary strata on LG19 appears to have occurred independently in each clade (Suppl Fig. S5). The species in Clade 2.2 share a stratum from 0 to 5 Mb. The species in Clade 2.3 share a stratum from 0 to 6 Mb, while in Clade 2.4 (*P. famula*) it extends approximately 0–8 Mb. In Clade 2.5 the differentiated region in *P. curvifrons* extends 0–5 Mb, but in *P. babaulti South*, *P. marginatus* and *P. marginatus North* it extends 0–7 Mb, and in *P. babaulti* it extends all the way from 0 to 9 Mb. The expansion of these two regions in each species is summarized in Fig. 4.

## Discussion

### Rates and patterns of sex chromosome turnover

Cichlids are an ideal system in which to examine the rates and patterns of sex chromosome turnover due to their high species diversity and relatively recent divergence. Our analysis of the tribe Tropheini further contributes to the narrative that East African cichlid fishes have some of the highest rates of sex chromosome turnover in vertebrates.

The ancestral sex chromosome state for the tribe Tropheini is unknown. We were unable to identify sex chromosome systems in the deepest lineage of Clade 1 (*T. duboisi*) or Clade 2 (*L. labiatus*). Given the high rate of sex chromosome turnover in the group, and the large phylogenetic distance to neighboring tribes, we are unable to reconstruct the ancestral state for the tribe with any confidence. A previous study found significant probabilities for 5 different ancestral states for the common ancestor of the tribe<sup>42</sup>.

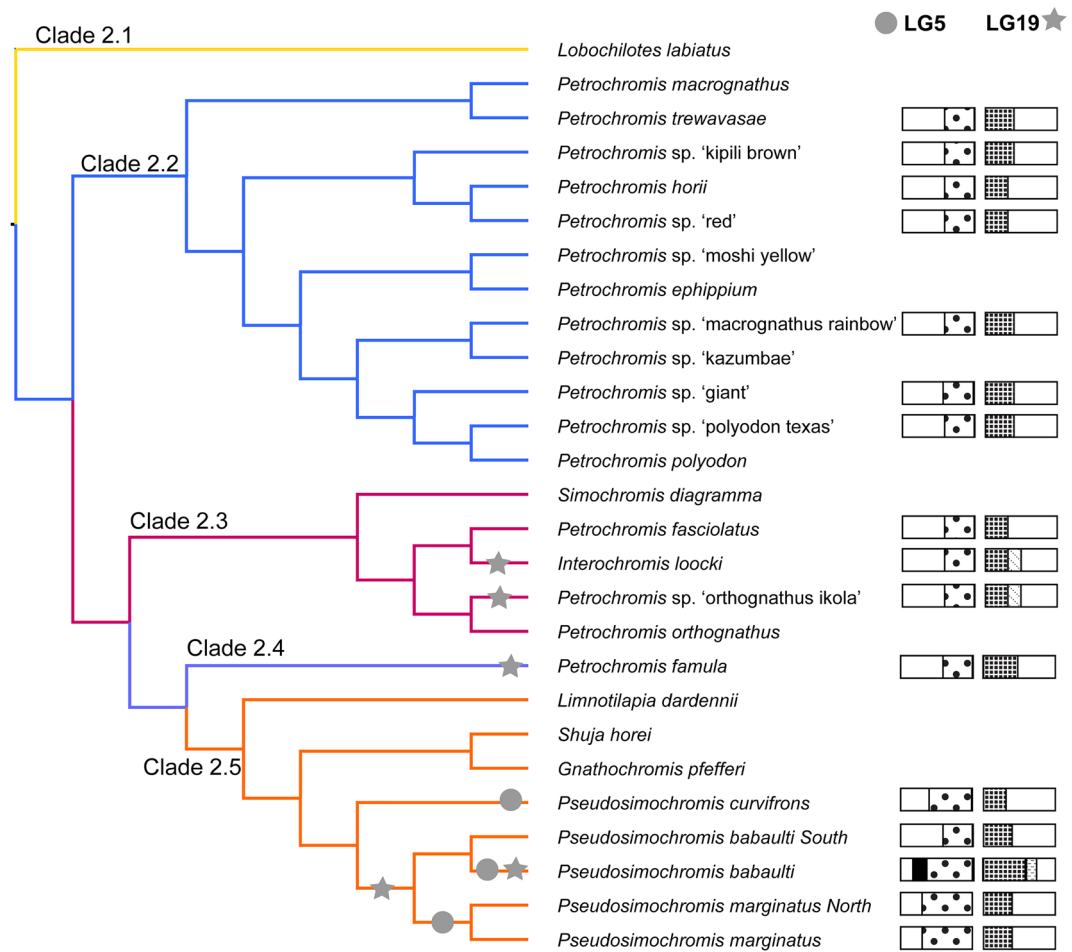
Although we were unable to identify a sex chromosome system in the deepest lineage of Clade 1 (*T. duboisi*), the ancestral state for the rest of Clade 1 appears to have been XY-LG19. The signal for LG19 in several species in Clade 1.2 and in *T. polli* (Clade 1.3) is strong and encompasses at least 60% of the chromosome. This signal should be detectable in other species in Clade 1 with our current methodology even in samples of single males and females. The absence of such signals in six species of Clades 1.2 and 1.3 is thus evidence for at least 3 additional sex chromosome turnovers.

We were also unable to identify the sex chromosome system in the deepest lineage of Clade 2 (*L. labiatus*). The ancestral state for the remainder of Clade 2 appears to have been XY-LG5/19, which is found in each of the major subclades (2.2, 2.3, and 2.5). This system also shows high levels of differentiation between the X and Y chromosomes over roughly 30% of both LG5 and LG19. This signal should also have been detectable in samples of single individuals. The absence of this signal in 12 species of Clade 2 suggests at least 10 additional sex chromosome turnovers have occurred.

A turnover to XY-LG11/15 in *G. pfefferi* is strongly supported by the pool-seq data. The sex-differentiated regions in *G. pfefferi* encompass roughly half of each chromosome and show differentiation similar in magnitude to the XY-LG5/19 system and is thus relatively old. Our data also suggest a recent turnover to an XY-LG5 system in *S. horei*. The region of differentiation is only about 100 kb, but already shows 82 XY-patterned SNPs.

The other 9 species lacking the XY-LG5/19 system represent at least 7 independent losses, presumably accompanied by the rise of different system of genetic sex determination. In total, from the root of the tribe Tropheini there have been at least 12 sex chromosome turnovers, leaving aside the unknown sex systems in *T. duboisi* and *L. labiatus* (Fig. 3). Assuming 12 turnovers and the branch lengths of the time-calibrated phylogeny from Singh et al.<sup>48</sup> which suggests that the tribe Tropheini is 3.6–3.8 million years old, we estimate the rate of sex chromosome turnover for this tribe is at least 0.259 turnovers per million years.





**Figure 4.** Expansion of the XY-LG5/19 sex chromosome on the phylogeny of Clade 2 as estimated from sex-specific SNP density data. Circles indicate expansion of LG5, stars indicate expansion of LG19. For LG5 the dotted pattern indicates what is presumably the original region, as does the plaid pattern for LG19, with any further additions in different patterns representing potential evolutionary strata. Phylogeny adapted from Ronco et al.<sup>44</sup>.

### Polygenic sex determination

There is disagreement on the stability of polygenic sex chromosome systems, as some argue they should be unstable intermediates or the result of recent hybridization<sup>70,71</sup>. However evidence in house flies (*Musca domestica*)<sup>72,73</sup> and in cichlids, where multiple sex chromosome systems can persist within species for periods of up to a million years, suggests otherwise. This is evident from work on Lake Malawi cichlids, where sex loci associated with both the orange-blotch (OB) inversion<sup>74,75</sup> and B chromosomes<sup>57,76,77</sup> are broadly shared among species. Further support comes from the polymorphism of sex chromosomes in *Astatotilapia burtoni*<sup>57,74,75,77,78</sup>. Finally, long-term sex chromosome polymorphism has been described in the Lake Tanganyika tribe *Cyprichromini*<sup>52</sup>. Similar polymorphisms may have been overlooked in less species-rich clades, especially because most sampling strategies are not designed to detect multiple sex chromosome systems segregating in a population.

We discovered some evidence for sex chromosome polymorphism in *S. diagramma*. The pool-seq analysis revealed strong differentiation on LG7, but the analysis of single individuals from another sample suggested XY systems on LG8 and 16. The lack of sex chromosome signal in the other species which do not show either the XY-LG19 or XY-LG5/19 system can most likely be attributed to the recent invasion of a new, relatively undifferentiated sex chromosome that is not detectable using our GWAS approaches. But it might also indicate the segregation of multiple sex chromosome systems that would tend to obscure the signal in pool-seq analyses.

### Chromosome fusions

The relationship between the evolution of karyotypes, sex chromosomes and speciation has been the subject of considerable research<sup>79,80</sup>. Sex chromosomes can drive overall karyotype evolution via inversions, deletions, and fusions<sup>40</sup>. Chromosome fusions have been proposed to be involved in sex chromosome turnover<sup>38</sup>, which may further result in speciation as has been proposed in stickleback<sup>34</sup>. Supernumerary B chromosomes have become

sex-determining in some Lake Malawi cichlids<sup>81</sup>, and we have suggested that a fusion between a B chromosome and a sex chromosome generated the giant LG3 chromosome found only in the tilapia lineage<sup>76</sup>.

The typical cichlid diploid chromosomal number is ( $2n = 44$ )<sup>82</sup>. Karyotypic data is not available for most species of the Tropheini, but we predict at least two fusions in this clade. The first is predicted by the strong sex signals of an XY-LG5/19 system in Clade 2. We predict that karyotypes of these species will display  $2n = 42$ . The second is the XY-LG11/15 system in *G. pfefferi*, which we predict will display  $2n = 40$ . In both cases, the fusions likely preceded the development of sex chromosome characteristics. The differentiated regions are found at the end of these chromosomes, where cichlid recombination rates are typically low<sup>54</sup>. These regions of reduced recombination may have facilitated the development of sex chromosomes in these species.

### Evolutionary strata

The multiple versions of the Y chromosomes in these clades provide an opportunity to study the dynamics of the formation of evolutionary strata on the sex chromosomes. We summarize our findings in Fig. 4. *Pseudosimochromis babaulti* in Clade 2.5 appears to have the largest region of differentiation on both LG5 and LG19. The other species in Clades 2.4 and 2.5 have experienced some expansion as well, but not to the same extent as *P. babaulti*. The overall pattern suggests that the initial stratum was established over the ~5 Mb terminal regions of each chromosome, which were telomeric regions prior to chromosome fusion. Low recombination rates seem to persist in these formerly telomeric regions for millions of years after chromosome fusion (e.g. LG23)<sup>54</sup>. These regions of inherently low recombination may represent a favorable environment for the evolution of new sex chromosomes.

The definition of very young evolutionary strata remains a challenging problem. Because the levels of differentiation are low, and the differences between strata even smaller, it is difficult to delineate the boundaries of individual strata from measures of sequence differentiation alone. Stringent definition of these strata may require phased assemblies of the sex chromosomes for each species.

### Conclusion

With each new study it becomes evident that the variability of the sex chromosomes in East African cichlids is greater than previously anticipated. This is an important insight because when we underestimate the number of sex chromosome systems, we undercount the number of turnovers and may infer a different pattern of sex chromosome replacement. We also undercount the number of autosomes that have become sex chromosomes, which skews our understanding of which chromosomes and genes can become involved in sex determination. Most of what we understand about sex chromosome evolution comes from distantly related organisms with old and very diverged sex chromosomes. Only in groups like cichlids, with recent, species-rich clades, can we begin to understand the early stages of sex chromosome evolution.

### Data availability

The new sequence data for *Pseudosimochromis babaulti*, *Shuja horei* (previously called *Ctenochromis horei*), *Gnathochromis pfefferi*, *Simochromis diagramma*, and *Interochromis loocki* are available on NCBI under BioProject: PRJNA802233. Previously published data are available under their respective BioProjects on NCBI: Gamberdinger et al. 2018 BioProject: PRJNA400462 and El Taher et al. 2021 BioProject: PRJNA552202.

Received: 14 July 2023; Accepted: 25 January 2024

Published online: 30 January 2024

### References

1. Ezaz, T. *et al.* The dragon lizard *Pogona vitticeps* has ZZ/ZW micro-sex chromosomes. *Chromosome Res.* **13**, 763–776 (2005).
2. Valenzuela, N., Adams, D. C. & Janzen, F. J. Pattern does not equal process: Exactly when is sex environmentally determined?. *Am. Nat.* **161**, 676–683 (2003).
3. Iannucci, A. *et al.* Conserved sex chromosomes and karyotype evolution in monitor lizards (Varanidae). *Heredity* **123**, 215–227 (2019).
4. Kawai, A. *et al.* Different origins of bird and reptile sex chromosomes inferred from comparative mapping of chicken Z-linked genes. *Cytogenet. Genome Res.* **117**, 92–102 (2007).
5. Bull, J. J. *Evolution of sex determining mechanisms.* (Benjamin/Cummings Publishing Company, 1983).
6. Charlesworth, D., Charlesworth, B. & Marais, G. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* **95**, 118–128 (2005).
7. Charlesworth, D. & Charlesworth, B. Sex differences in fitness and selection for centric fusions between sex-chromosomes and autosomes. *Genet. Res.* **35**, 205–214 (1980).
8. Rice, W. R. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution* **41**, 911–914 (1987).
9. Terao, M. *et al.* Turnover of mammal sex chromosomes in the Sry-deficient Amami spiny rat is due to male-specific upregulation of Sox9. *Proc. Natl. Acad. Sci. U. S. A.* **119**, e2211574119 (2022).
10. Zhou, Q. *et al.* Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science* **346**, 1246338 (2014).
11. Cortez, D. *et al.* Origins and functional evolution of y chromosomes across mammals. *Nature* **508**, 488–493 (2014).
12. Pennell, M. W., Mank, J. E. & Peichel, C. L. Transitions in sex determination and sex chromosomes across vertebrate species. *Mol. Ecol.* **27**, 3950–3963 (2018).
13. Mank, J. E. & Avise, J. C. Evolutionary diversity and turn-over of sex determination in teleost fishes. *Sex. Dev.* **3**, 60–67 (2009).
14. Jeffries, D. L. *et al.* A rapid rate of sex-chromosome turnover and non-random transitions in true frogs. *Nat. Commun.* **9**, 4088 (2018).
15. Tao, W. *et al.* High-quality chromosome-level genomes of two tilapia species reveal their evolution of repeat sequences and sex chromosomes. *Mol. Ecol. Resour.* **21**, 543–560 (2021).
16. Van Doorn, G. S. & Kirkpatrick, M. Turnover of sex chromosomes induced by sexual conflict. *Nature* **449**, 909–912 (2007).

17. Van Doorn, G. S. & Kirkpatrick, M. Transitions between male and female heterogamety caused by sex-antagonistic selection. *Genetics* **186**, 629–645 (2010).
18. Blaser, O., Neuenschwander, S. & Perrin, N. Sex-chromosome turnovers: The hot-potato model. *Am. Nat.* **183**, 140–146 (2014).
19. Blaser, O., Grossen, C., Neuenschwander, S. & Perrin, N. Sex-chromosome turnovers induced by deleterious mutation load. *Evolution* **67**, 635–645 (2013).
20. Bull, J. J. & Charnov, E. L. Changes in the heterogametic mechanism of sex determination. *Heredity* **39**, 1–14 (1977).
21. Cavoto, E., Neuenschwander, S., Goudet, J. & Perrin, N. Sex-antagonistic genes, XY recombination and feminized Y chromosomes. *J. Evol. Biol.* **31**, 416–427 (2018).
22. Charlesworth, D. Evolution of recombination rates between sex chromosomes. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **372**, 20160456 (2017).
23. Ponnikas, S., Sigeman, H., Abbott, J. K. & Hansson, B. Why do sex chromosomes stop recombining?. *Trends Genet.* **34**, 492–503 (2018).
24. Ironside, J. E. No amicable divorce? Challenging the notion that sexual antagonism drives sex chromosome evolution. *BioEssays* **32**, 718–726 (2010).
25. Jay, P., Tezenas, E., Véber, A. & Giraud, T. Sheltering of deleterious mutations explains the stepwise extension of recombination suppression on sex chromosomes and other supergenes. *PLoS Biol.* **20**, e3001698 (2022).
26. Lenormand, T. & Roze, D. Y recombination arrest and degeneration in the absence of sexual dimorphism. *Science* **375**, 663–666 (2022).
27. Charlesworth, D. When and how do sex-linked regions become sex chromosomes?. *Evolution* **75**, 569–581 (2021).
28. Lahn, B. T. & Page, D. C. Four evolutionary strata on the human X chromosome. *Science* **286**, 964–967 (1999).
29. Olito, C., Ponnikas, S., Hansson, B. & Abbott, J. K. Consequences of partially recessive deleterious genetic variation for the evolution of inversions suppressing recombination between sex chromosomes. *Evol. Int. J. Org. Evol.* **76**, 1320–1330 (2022).
30. Rice, W. R. Evolution of the Y sex chromosome in animals. *BioScience* **46**, (1996).
31. Skaletsky, H. *et al.* The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **423**, 825–837 (2003).
32. Dagilis, A. J. *et al.* Searching for signatures of sexually antagonistic selection on stickleback sex chromosomes. *Philos. Trans. R. Soc. B Biol. Sci.* **377**, 1–9 (2022).
33. Sardell, J. M., Josephson, M. P., Dalziel, A. C., Peichel, C. L. & Kirkpatrick, M. Heterogeneous histories of recombination suppression on stickleback sex chromosomes. *Mol. Biol. Evol.* **38**, 4403–4418 (2021).
34. Yoshida, K. *et al.* Sex chromosome turnover contributes to genomic divergence between incipient stickleback species. *PLOS Genet.* **10**, e1004223 (2014).
35. Charlesworth, D. & Mank, J. E. 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 (2010).
36. Devlin, R. H. & Nagahama, Y. Sex determination and sex differentiation in fish: An overview of genetic, physiological, and environmental influences. *Aquaculture* **208**, 191–364 (2002).
37. Gamble, T. Using RAD-seq to recognize sex-specific markers and sex chromosome systems. *Mol. Ecol.* **25**, 2114–2116 (2016).
38. Kitano, J. & Peichel, C. L. Turnover of sex chromosomes and speciation in fishes. *Environ. Biol. Fishes* **94**, 549–558 (2012).
39. Sember, A. *et al.* Multiple sex chromosomes in teleost fishes from a cytogenetic perspective: state of the art and future challenges. *Philos. Trans. R. Soc. B Biol. Sci.* **376**, 20200098 (2021).
40. Pennell, M. W. *et al.* Y Fuse? Sex chromosome fusions in fishes and reptiles. *PLOS Genet.* **11**, e1005237 (2015).
41. Kocher, T. D. Adaptive evolution and explosive speciation: The cichlid fish model. *Nat. Rev. Genet.* **5**, 288–298 (2004).
42. El Taher, A., Ronco, F., Matschiner, M., Salzburger, W. & Böhne, A. Dynamics of sex chromosome evolution in a rapid radiation of cichlid fishes. *Sci. Adv.* **7**, eabe8215 (2021).
43. Gammerdinger, W. J. *et al.* Novel sex chromosomes in 3 cichlid fishes from Lake Tanganyika. *J. Hered.* **109**, 489–500 (2018).
44. Ronco, F. *et al.* Drivers and dynamics of a massive adaptive radiation in cichlid fishes. *Nature* **589**, 76–81 (2021).
45. Koblmüller, S., Egger, B., Sturmbauer, C. & Sefc, K. M. Rapid radiation, ancient incomplete lineage sorting and ancient hybridization in the endemic Lake Tanganyika cichlid tribe Tropheini. *Mol. Phylogenet. Evol.* **55**, 318–334 (2010).
46. Koblmüller, S., Sefc, K. M. & Sturmbauer, C. The Lake Tanganyika cichlid species assemblage: recent advances in molecular phylogenetics. *Hydrobiologia* **615**, 5–20 (2008).
47. Ad Konigs. *Tanganyika cichlids in their natural habitat.* (Cichlid Press, 2019).
48. Singh, P. *et al.* Phylogenomics of trophically diverse cichlids disentangles processes driving adaptive radiation and repeated trophic transitions. *Ecol. Evol.* **12**, 1–15 (2022).
49. Poll, M. Classification des Cichlidae du lac Tanganika. Tribus, genres et especes. *Mem Acad. R Belg. Cl Sci.* **45**, 1–163 (1986).
50. Genner, M. J. *et al.* Revision of the African cichlid fish genus *Ctenochromis* (Teleostei, Cichliformes), including a description of the new genus *Shuja* from Lake Tanganyika and the new species *Ctenochromis scatebra* from northern Tanzania. *Eur. J. Taxon.* **819**, 23–54 (2022).
51. Percie du Sert, N. *et al.* The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. *BMC Vet. Res.* **16**, 242 (2020).
52. Behrens, K. A., Koblmüller, S. & Kocher, T. D. Sex chromosomes in the tribe Cyprichromini (Teleostei: Cichlidae) of Lake Tanganyika. *Sci. Rep.* **12**, 1–13 (2022).
53. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
54. Conte, M. A. *et al.* Chromosome-scale assemblies reveal the structural evolution of African cichlid genomes. *GigaScience* **8**, giz030 (2019).
55. Conte, M. A. & Kocher, T. D. An improved genome reference for the African cichlid *Metriacrima zebra*. *BMC Genomics* **16**, 724 (2015).
56. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
57. Kocher, T. D. *et al.* New sex chromosomes in lake victoria cichlid fishes (Cichlidae: Haplochromini). *Genes Basel* **13**, 804 (2022).
58. Darolti, I., Almeida, P., Wright, A. E. & Mank, J. E. Comparison of methodological approaches to the study of young sex chromosomes: A case study in *Poecilia*. *J. Evol. Biol.* **35**, 1646–1658 (2022).
59. R Core Development Team. R: A language and environment for statistical computing. *Vienna Austria* (2019).
60. Thorvaldsdóttir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Brief. Bioinform.* **14**, 178–192 (2013).
61. Marçais, G. & Kingsford, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. *Bioinformatics* **27**, 764–770 (2011).
62. Szwarc, M. M., Kommagani, R., Lessey, B. A. & Lydon, J. P. The p160/steroid receptor coactivator family: Potent arbiters of uterine physiology and dysfunction. *Biol. Reprod.* **91**(122), 1–11 (2014).
63. Villavicencio, A. *et al.* Androgen and estrogen receptors and co-regulators levels in endometria from patients with polycystic ovarian syndrome with and without endometrial hyperplasia. *Gynecol. Oncol.* **103**, 307–314 (2006).
64. Yoshida, H. *et al.* Steroid receptor coactivator-3, a homolog of Taiman that controls cell migration in the *Drosophila* ovary, regulates migration of human ovarian cancer cells. *Mol. Cell. Endocrinol.* **245**, 77–85 (2005).

65. Massari, M. E. & Murru, C. Helix-loop-helix proteins: Regulators of transcription in eucaryotic organisms. *Mol. Cell. Biol.* **20**, 429–440 (2000).
66. Korchynskiy, O. & ten Dijke, P. Identification and functional characterization of distinct critically important bone morphogenetic protein-specific response elements in the Id1 promoter. *J. Biol. Chem.* **277**, 4883–4891 (2002).
67. Kamiya, T. *et al.* A trans-species missense SNP in Amhr2 is associated with sex determination in the tiger pufferfish, *Takifugu rubripes* (Fugu). *PLoS Genet.* **8**, e1002798 (2012).
68. Li, M. *et al.* A tandem duplicate of anti-Müllerian hormone with a missense SNP on the Y chromosome is essential for male sex determination in Nile Tilapia, *Oreochromis niloticus*. *PLoS Genet.* **11**, e1005678 (2015).
69. Nacif, C. L. *et al.* Molecular parallelism in the evolution of a master sex-determining role for the anti-Müllerian hormone receptor 2 gene (amhr2) in Midas cichlids. *Mol. Ecol.* **32**, 1–13 (2023).
70. Rice, W. R. On the instability of polygenic sex determination: the effect of sex-specific selection. *Evolution* **40**, 633–639 (1986).
71. Scharlt, M., Georges, A. & Graves, J. A. M. Polygenic sex determination in vertebrates—is there any such thing?. *Trends Genet.* **39**, 242–250 (2023).
72. Meisel, R. P. *et al.* Is multifactorial sex determination in the house fly, *Musca domestica* (L.), stable over time? *J. Hered.* **107**, 615–625 (2016).
73. Meisel, R. P. The maintenance of polygenic sex determination depends on the dominance of fitness effects which are predictive of the role of sexual antagonism. *G3 Bethesda* **11**, jkab149 (2021).
74. Böhne, A., Wilson, C. A., Postlethwait, J. H. & Salzburger, W. Variations on a theme: Genomics of sex determination in the cichlid fish *Astatotilapia burtoni*. *BMC Genomics* **17**, 883 (2016).
75. Heule, C., Göppert, C., Salzburger, W. & Böhne, A. Genetics and timing of sex determination in the East African cichlid fish *Astatotilapia burtoni*. *BMC Genet.* **15**, 140 (2014).
76. Conte, M. A. *et al.* Origin of a giant sex chromosome. *Mol. Biol. Evol.* **38**, 1554–1569 (2021).
77. Lichilín, N., Salzburger, W. & Böhne, A. No evidence for sex chromosomes in natural populations of the cichlid fish *Astatotilapia burtoni*. *G3 Bethesda* **13**, jkad011 (2023).
78. Roberts, N. B. *et al.* Polygenic sex determination in the cichlid fish *Astatotilapia burtoni*. *BMC Genomics* **17**, 835 (2016).
79. Payseur, B. A., Presgraves, D. C. & Filatov, D. A. Introduction: Sex chromosomes and speciation. *Mol. Ecol.* **27**, 3745–3748 (2018).
80. de Vos, J. M., Augustijnen, H., Batscher, L. & Lucek, K. Speciation through chromosomal fusion and fission in Lepidoptera. *Philos. Trans. R. Soc. B Biol. Sci.* **375**, 20190539 (2020).
81. Clark, F. E. & Kocher, T. D. Changing sex for selfish gain: B chromosomes of Lake Malawi cichlid fish. *Sci. Rep.* **9**, 20213 (2019).
82. Majtánová, Z., Indermaur, A., Bitja Nyom, A. R., Ráb, P. & Musilova, Z. Adaptive radiation from a chromosomal perspective: Evidence of chromosome set stability in cichlid fishes (Cichlidae: Teleostei) from the Barombi Mbo Lake. *Cameroon. Int. J. Mol. Sci.* **20**, 4994 (2019).

## Acknowledgements

We would like to thank the administrators of the high-performance computing clusters at University of Maryland for their continued assistance.

## Author contributions

The study was conceived by K.A.B. and T.D.K. Samples for pooled sequencing were collected by S.K., H.Z., R.B., and M.R.; DNA extraction and quantification was performed by T.D.K. Sequence analyses were performed by K.A.B. and T.D.K. The manuscript was prepared by K.A.B. and edited by T.D.K., S.K., H.Z., R.B., and M.R. All authors have read and agreed to the published version of the manuscript.

## Funding

This work was supported by an award from the National Science Foundation (DEB-1830753) to TDK.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-53021-3>.

**Correspondence** and requests for materials should be addressed to K.A.B.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024