Non-canonical sex chromosome evolution revealed by extreme heterogeneity in homomorphic Y chromosome differentiation in the common frog

Fantin Carpentier[1*], Paris Veltsos[1,2*†], Ezgi Unal[1], Kelsey L. Doucette[3], Ricard Fontseré[1], Alan Brelsford[4], Melissa A. Toups[2,3], Nicolas Perrin[2], Wen-Juan Ma[1,2*†]

1. Research group of Ecology, Evolution and Genetics, Biology Department, Vrije Universiteit Brussel, Brussels, Belgium
2. Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland
3. Department of Biology, University of Louisiana at Lafayette, Lafayette, LA, USA
4. Evolution Ecology and Organismal Biology Department, University of California, Riverside, CA, USA

[*]Equal contribution:
The two authors contribute to the work equally.

[†]Correspondence:
wen-juan.ma@vub.be
paris.veltsos@vub.be

Abstract

Sex chromosome evolution is traditionally viewed as a unidirectional process of recombination suppression followed by progressive Y degeneration, yet many vertebrates deviate from this canonical trajectory. Frogs represent a striking exception: they retain homomorphic sex chromosomes, exhibit extreme heterochiasmy, and undergo X-Y recombination in sex-reversed females, providing a powerful system to explore alternative evolutionary pathways. Here, we investigate a Swiss Alpine population of the European common frog (*Rana temporaria*) with homomorphic XY chromosomes and three coexisting male genotypes, using whole genome sequencing of pooled samples carrying distinct Y haplotypes, whole genome sequencing of doubled haploid YY individuals, and transcriptomic (RNA-seq) data. We uncover extreme heterogeneity in X–Y divergence: two fully differentiated Y haplotypes span nearly 90% of the sex chromosome (~618–625 Mb), forming the largest non-recombining region (large NRR) described in a vertebrate to date, whereas a semi-differentiated Y harbors only a NRR of 4.64 Mb (small NRR), and XX males display no detectable X-Y divergence. Despite extensive recombination suppression, Y degeneration is minimal even within the large NRR: we detect elevated transposable element insertions but no gene loss, Y-linked gene copy decay, faster-X effects, or enrichment of sex-biased genes. All Y haplotypes share a small NRR encompassing the candidate master sex-determining gene *Dmrt1*, defining a common sex-determining locus. Our data support a dynamic model in which extreme heterochiasmy and recurrent X–Y recombination via sex reversal repeatedly reset X-Y divergence, generating and potentially maintaining multiple Y haplotypes with distinct evolutionary histories. The long-term coexistence of these haplotypes across the species range is compatible with a dynamic equilibrium between drift-driven loss and mutation- and recombination-driven renewal of Y chromosome variation. Together, our findings reveal a non-canonical, reversible pathway of sex chromosome evolution shaped by sex-specific recombination patterns and sex reversal, challenging the universality of classical models and highlighting the value of non-model vertebrates for understanding sex chromosome diversity.

Keywords:

Sex chromosome evolution, non-canonical pathway, extreme heterochiasmy, sex-specific recombination, sex reversal, coexisting male genotypes, dynamic equilibrium, sex determination, *Dmrt1*

Introduction

Sex determination is one of the most fundamental developmental processes and is ubiquitous across eukaryotes [1–3]. Sex chromosomes have evolved in association with genetic sex determination and play pivotal roles in diverse evolutionary processes, including sexual selection, sex-biased gene flow and dispersal, genomic conflict, speciation and adaptation [4–7]. Sex chromosomes have independently evolved numerous times across the Tree of Life and demonstrate a remarkable diversity in their rate and level of genetic differentiation [1,8–10]. According to the canonical model of sex chromosome evolution, sex chromosomes are derived from a pair of homologous autosomes when a sex-determining locus arises by mutation on the sex-limited chromosome (e.g. the Y or W chromosome) (Fig. 1a). Various evolutionary forces result in recombination suppression between the sex-determining locus and the surrounding regions, causing the sex-limited Y/W chromosome to degenerate and progressively diverge in structure and function from the still-recombining X/Z chromosome. This can result in highly degenerated Y/W chromosomes which have lost much of their original gene content [11–13], as observed in most mammals [14–16], most birds [17], *Drosophila* [18] and many insect lineages [19,20] (Fig. 1a).

Various hallmarks of the canonical model are poorly supported by empirical evidence across the Tree of Life, particularly outside well-studied model lineages [21–25]. One hallmark is the initiation of recombination suppression by selection linking the sex-determining locus with sexually antagonistic loci (alleles beneficial for one sex but detrimental for the other) [26,27]. Another hallmark is the inevitable gene loss and degeneration of the Y/W chromosomes over evolutionary time [11], and the level of Y degeneration is positively correlated with evolutionary time [15,28]. Recent studies have found gene gains, in addition to gene losses on the differentiated sex chromosomes, for example in the expanded neo-Y chromosome of *Drosophila miranda* [29], and in the mating-type chromosomes in various anther smut *Microbotryum* fungi (Tatiana Giraud, personal comm.; Ma et al., in prep.). In addition, there is strong heterogeneity in the rate and differentiation of sex chromosomes both among closely related and within species [17,30–34]. It seems that the age of sex chromosomes does not correlate well with the degree of recombination suppression, as sex chromosomes can remain homomorphic or maintain a large recombining region (the pseudoautosomal region, PAR) despite long divergence times (e.g. 25-150 million years in ratites) [30,31,35–37].

It has recently been proposed that pre-existing recombination suppression can act as a genomic precondition that contributes to the tempo and dynamics of sex chromosome evolution [28,38–40] (Fig. 1b). In this scenario, the first recombination suppression event is not a consequence of the link between sex-determining loci and sexually antagonistic loci as the canonical model suggests, but rather a pre-existing feature of the genome. Such suppression can occur along a continuum ranging from local cold spots to extreme heterochiasmy, where recombination is strongly reduced in one sex (as in frogs and guppies) [41–43] and ultimately to achiasmy where recombination is entirely absent in one sex (as in *Drosophila* [44]). This pre-existing landscape of reduced recombination could fundamentally alter the trajectory of sex chromosome evolution. When a new sex-determining gene arises within a region already largely devoid of recombination, recombination arrest across the sex-limited Y/W chromosome occurs instantaneously, promoting rapid sequence divergence and degeneration (Fig. 1b). Accordingly, achiasmy in XY males of *Drosophila* flies and in ZW females of *Lepidoptera* (butterflies and moths) are all associated with highly heteromorphic and degenerated sex chromosomes [19,20,45]. A sex chromosome-autosome fusion can instantly stop recombination on the resulting neo-sex chromosome region, which accelerates the degeneration process[46]. This stands in contrast to the canonical view, where recombination suppression spreads through linkage with sexually antagonistic alleles, often in combination with inversions. Our understanding on how pre-existing

recombination suppression affects the evolution of sex chromosomes emerging *de novo* remains limited, as most data come from *Drosophila* neo-sex chromosomes originating by fusion between heteromorphic sex chromosomes and autosomes.

In sharp contrast to the highly degenerated sex chromosomes in most mammals and birds, most amphibians [47,48], fishes [33,49–53], many lizards [32,55,56], and flowering plants [9,54–56] show largely homomorphic sex chromosomes, with minimal degeneration. Theory predicts that X-linked genes may evolve faster than autosomal genes, a pattern known as Faster-X evolution, because recessive mutations are more efficiently exposed to selection in the hemizygous sex, and because effective population sizes differ between sex chromosomes and autosomes [57]. Empirical support for Faster-X evolution has been mixed and appears to depend on factors such as the degree of sex chromosome differentiation, recombination patterns, and the efficacy of selection. Whether Faster-X dynamics operate in young, homomorphic sex chromosome systems remains an open question. Possible explanations for maintenance of homomorphy include sex chromosome turnover [31,58–62], where closely related species use different chromosome pairs to determine sex. Sex-chromosome turnover results in the loss of one of the ancestral sex chromosomes (X or Y), through the loss of its sex linkage and the emergence of a new sex chromosome. This transition restores homomorphy to the ancestral sex chromosome pair (i.e. XX or YY), returning it to an autosomal state. Another non-exclusive mechanism is the 'fountain of youth' hypothesis, which posits that recombination is determined by phenotypic rather than genotypic sex. Under this scenario, sex-reversed (XY) females recombine, thereby purging deleterious mutations that would otherwise accumulate on the Y[63] (Fig. 1b). Sex reversals and the associated restoration of recombination along sex chromosomes have been documented in the European common frog (*Rana temporaria*). In this species, XY females appear to experience no detectable reproductive disadvantage [64,65] and even their rare occurrence, estimated to be as low as $1/10^5$ by simulations is sufficient to prevent Y-chromosome degeneration [66].

Frogs are excellent systems to investigate an alternative route of sex chromosome evolution, driven by extreme heterochiasmy and sex reversal evolutionary dynamics [40,47,64,67]. Over 78% of all studied anurans possess largely homomorphic sex chromosomes, suitable to address the initiation of recombination suppression between sex chromosomes which cannot be answered from studies of species with highly degenerated systems [47]. Extreme heterochiasmy has been reported in many anuran lineages: females recombine evenly along their chromosome length, while male recombination is largely restricted to chromosome ends [31,42,47,60]. Consequently, a newly arisen male-determining gene instantly defines a non-recombining haplotype spanning most of the neo-Y chromosome (Fig. 1b). Extreme heterochiasmy alone would result in sex chromosome differentiation over most of its length. Combined with deleterious mutation accumulation, it may explain the rapid sex chromosome turnover in many anurans [31]. Anurans are important for understanding sex chromosome evolution because several lineages, including XY and ZW lineages, have heteromorphic sex chromosomes, and closely related species transition between XY and ZW systems without preserving homomorphy (*Eleutherodactylus, Pristimantis, Leiopelma, Engystomops*) [47]. Either not all anurans have extreme heterochiasmy and reproduce sex-reversed XY females, or other evolutionary forces are at play.

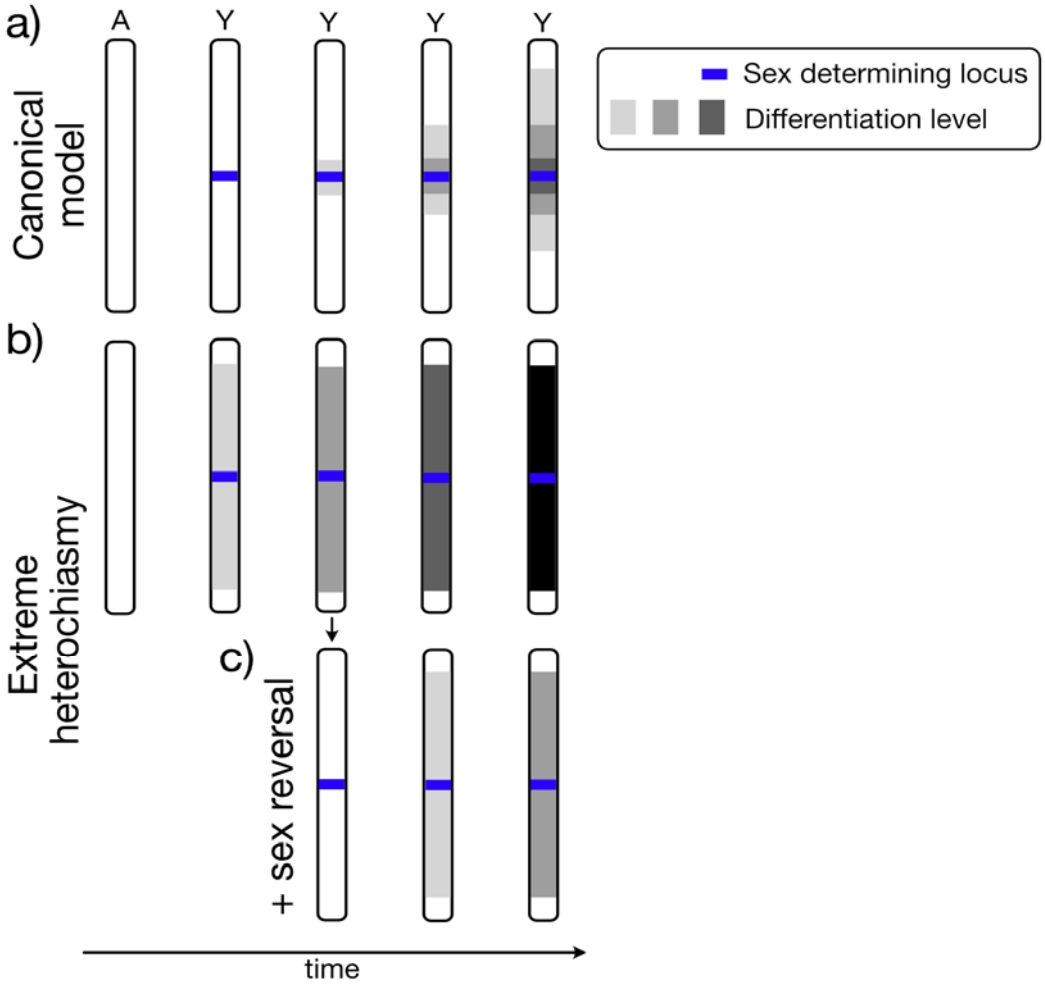

Fig. 1 | Divergence trajectories of the Y chromosome under the canonical model with chiasmata versus extreme heterochiasmy. a) In the canonical model, the origin of a sex-determining allele on the proto-Y initiates localized X–Y divergence through reduced recombination and the build-up of sexually antagonistic alleles, with recombination suppression subsequently expanding to form progressively differentiated strata. b) Under extreme heterochiasmy, where male recombination is confined to telomeric sites, the establishment of a sex-determining locus produces immediate, chromosome-wide X–Y differentiation except at the recombining ends. XY sex-reversal events transiently restore X–Y recombination and reset differentiation outside the sex-determining region; the resulting recombinant Y haplotype then accumulates independent mutations and follows a novel divergence trajectory.

Sex chromosomes, in addition, can have intraspecific heterogeneity in X-Y differentiation [31,68]. Of particular interest is the European common frog *Rana temporaria*. It has homomorphic sex chromosomes and intraspecific variation in Y differentiation based on 16 sex-linked markers spanning the sex chromosome genetic map: there are fully differentiated Y haplotypes (X-Y differentiation across all 16 markers along the sex linkage group), semi-differentiated Y haplotypes (differentiation restricted to a few markers spanning roughly 0.8cM) and undifferentiated Y chromosomes (no markers showed differentiation between males and females) [64,69–72]. Interestingly, two or three Y haplotypes can co-exist within a single pond in various localities across the species range [64,68,69,72–74]. Previous work showed no evidence of sexually antagonistic loci accumulating on these Y haplotypes: there was no difference in morphology, mating or siring success among males with different Y haplotypes from a pond (Meitreile) in the Swiss Alps [65]. There was little or no enrichment of genes with sex-biased expression on the sex chromosomes of any differentiation level in other populations [68,75,76]. Combined

with the rapid sex chromosome turnover that conserves male heterogamety (XY system) in Ranidae [31], the results provide little evidence for sexually antagonistic selection driving sex chromosome recombination suppression and differentiation in frogs [67]. However, key questions remain about how sex chromosomes evolve in this system. These include the genomic boundaries of X–Y divergence across various Y haplotypes, the timing and dynamics of Y-chromosome differentiation under extreme heterochiasmy and recurrent sex reversal, the identity and origin of the master sex-determining gene, as well as how multiple male genotypes coexist in a single pond.

In this study, we investigate a non-canonical model of sex chromosome evolution under extreme heterochiasmy (Fig. 1), by using next-generation sequencing to obtain the coexisting Y haplotypes showing varying degrees of differentiation within the pond of Meitreile, Switzerland, at high resolution. We integrate whole-genome pooled sequencing (Pool-seq), whole-genome sequencing of doubled haploid YY individuals, and transcriptomic (RNA-seq) data to investigate how pre-existing extreme heterochiasmy and occasional sex reversals ("the fountain of youth") shape the evolutionary history of sex chromosomes in the common frog. Consistent with earlier cytogenetic work, our genomic analyses reveal homomorphic sex chromosomes without gene loss, supporting recombination between X and Y via sex reversal. Remarkably, Y-chromosome differentiation exhibits extreme heterogeneity within a single pond, with non-recombining regions (NRRs), as inferred from markedly elevated male-female F_{ST}, ranging from extensive regions spanning 618-625 Mb (large NRR), to a much smaller region of 4.64 Mb (small NRR), and in some cases no detectable NRR (no NRR). We further detect a narrow peak of elevated F_{ST} containing 29 genes. Multiple independent lines of evidence including gene function, sex-specific gene expression, and population genetic signatures of SNP clustering inferred from homozygosity-heterozygosity patterns, consistently implicate *Dmrt1* as the master sex-determining gene. Although we detect no gene loss, we observe significantly elevated transposable element insertions in the most differentiated Y haplotypes, but without clear evidence of Y-linked gene decay nor enrichment of sex-biased genes on the sex chromosomes. Excitingly, our analyses date the shared small NRR across all Y haplotypes to approximately 3.47-5.70 Mya. The divergence patterns of fully differentiated Y haplotypes are consistent with ancestral XY sex reversal event occurring around 3.13 Mya or shortly afterwards (<2.94 Mya), whereas the semi-differentiated Y haplotypes appear to have emerged more recently, at approximately 0.98 Mya, from a fully differentiated Y. These results suggest a dynamic cycle between fully differentiated and semi-differentiated Y chromosomes. Our findings provide genomic evidence of a non-canonical route of sex chromosome evolution that complements and extends the canonical framework centred on sexually antagonistic selection and progressive recombination suppression.

Results

Whole genome sequencing of pooled individuals and doubled haploid YY individuals

To understand how extreme heterochiasmy and fountain of youth affect the differentiation and the evolutionary dynamics of sex chromosomes, we performed whole genome sequencing of pooled males and females of the European common frog *Rana temporaria*. Using 12-17 sex-linked markers, previous studies revealed three distinct levels of X-Y chromosome differentiation in males of the Swiss Meitreile population [65,72]. Utilizing the same markers, we sampled and sequenced 4 pools of males with either fully differentiated Y (with the most frequent haplotypes Y_{B1}^{a}, Y_{B2}^{a}), semi-differentiated Y (Y_{B2}^{0}), or undifferentiated Y (XX males), and 1 pool of XX females, most of which

have been documented previously [65,72] (Supplementary Table S1). These haplotypes follow the nomenclature of [65]. Note that Y_{B2}^{0} and Y_{B2}^{a} share the same Y haplotype at the 4 markers on 2 genes (*Dmrt1* and *Dmrt3*) but differ in the other markers. Less frequent Y chromosome haplotypes exist in the same population but were not included in our analysis. The number of individuals of each pool was 120, 120, 120, 50 and 48 respectively (Table 1). To obtain further insights from Y-specific features and avoid the challenges of phasing X and Y sequences of homomorphic sex chromosomes, we sequenced whole genomes from (double) haploid (Y)Y tadpoles which were generated with an androgenesis approach (Supplementary Table S2 and S3, see details in M&M). We further obtained individual genomic sequences from six such YY individuals with the same Y haplotype primarily from six clutches (Table S2, S3), 2 fully differentiated Y (Y_{B1}^{a}) and 4 semi-differentiated Y (Y_{B1}^{0}, Y_{B2}^{0}) (Table 1). To facilitate with Y genome assembly, we employed 3 mate-pair sequencing libraries of the fully differentiated YY male (Y_{B1}^{a}).

The sequencing of the 5 pools yielded between 548.8G and 744.1G bp (base pair) data (Supplementary Table S4). The reference genome of *R. temporaria* is based on one female collected in the National History Museum Wildlife Garden, London, UK [77]. The majority (99.5%) of the assembly was assigned to 13 chromosome-level scaffolds, and it has a 90.7% BUSCO completeness (tetrapoda_odb10 reference set)[78]. Given the reference genome size (approximately 4.11G bp), each pool is sequenced at 1.2 - 3.0-fold coverage per individual. For (doubled) haploid individual sequencing (Y_{B1}^{a}, Y_{B1}^{0} and Y_{B2}^{0}), there were 46.0G – 63.3G bp sequencing data (Table 1), and 3 additional mate-pair libraries with 1.50G – 21.1G bp data for the XY_{B1}^{a} double haploid individual (Supplementary Table S4).

We confirmed the order and physical position for most of the sex-linked markers of previous studies (Supplementary Table S5) using the reference genome assembly. Among the 15 sex-linked markers, 1 was missing on the X chromosome, 2 had a BLAST hit only for the reverse primers and 1 had multiple BLAST hits on the X chromosome with too large distances between the forward and reverse primers. One marker (Bfg191)'s order and position were sightly changed in the assembly (*Kank1, Dmrt1, Dmrt3, Bfg191,* compared to the published genetic map of *Kank1, Bfg191, Dmrt1* and *Dmrt3*). This could be due to population-specific local genomic rearrangement or genetic mapping error.

Table 1. Phenotypic and genotypic information on samples of pool-seq and doubled haploid YY individuals. Double haploid refers to the YY individuals generated through an androgenesis experiment. Family number refers to different individuals pooled from within the same clutch from Meitreile. * Frequency of each male genotype was calculated from the study of [65] , out of a total of 842 sampled males from Meitreile, Switzerland.

Sex	Y haplotype	Frequency of each male genotype*	X-Y differentiation (using sex-linked markers)	Pool-seq		(Double) haploid WGS	
				Pool size	Coverage	Family number	Coverage
Female	NA	NA	NA	120	197.10	NA	NA
Male	X	40.6%	Undifferentiated	120	147.81	NA	NA
Male	YB_1^{0}	6.5%	Semi-differentiated	NA	NA	Family 40	15.40
						Family 46	11.87
Male	YB_1^{a}	21.4%	Fully differentiated	120	146.88	Family 18	12.87
						Family 50	12.30
Male	YB_2^{0}	7.8%	Semi-differentiated	50	108.08	Family 33	11.20

						Family 4	12.76
Male	YB_2^a	11.6%	Fully differentiated	48	145.37	NA	NA

Extreme heterogeneity in homomorphic sex chromosome differentiation

To assess the level of differentiation between the X- and Y- chromosomes, we analysed the read coverage ratio, SNP density ratio and F_{ST} between each male pool and the female pool. The read coverage ratio between sexes is a useful metric to detect degeneration on the Y chromosome, particularly due to gene loss, and is most informative in systems with heteromorphic sex chromosomes. In such cases, Y-derived short reads align poorly to the X-chromosome reference, resulting in higher coverage in females than in males. The coverage differences can be used to detect the oldest strata of the sex chromosomes [79]. In contrast, in systems with largely homomorphic sex chromosomes and limited Y degeneration, the X- and Y-chromosomes differ mainly in base substitutions. Over time, SNPs can become fixed on the Y-chromosome, often leading to higher heterozygosity in males than in females. Thus, short reads from both the X- and Y-chromosome in the male pool align equally well to the X-chromosome reference, and the detection of differentiation relies on SNP density and allele frequency differences between the sexes. Regions with elevated genetic differentiation (F_{ST}) are indicative of sex linkage 78. In the absence of recombination, mutations on the Y accumulate due to reduced efficacy of purifying selection. Consequently, the Y is expected to show increasing levels of divergence from the X over time, including SNPs, indels, repetitive elements, and potential gene loss. Accordingly, the greater the X-Y differentiation, the lower the male-to-female coverage ratio is expected to be, approaching a $\log_2$ ratio of -1 (equivalent to a 0.5-fold coverage in males) when substantial Y-linked sequence loss has occurred [81].

First, we found no significant difference in the read coverage ratio between sexes across all autosomes in any pairwise comparison between male and female pools (Supplementary Fig. S1). The sex chromosomes showed no significant deviations from the 95% confidence interval (CI), calculated from 95% of the rolling mean of autosomal values, between the XX male, or the semi-differentiated Y_{B2}^0 and the female pool (Fig. 2: a1, a2), suggesting no detectable gene loss or heteromorphic regions. For the two fully differentiated male pools (Y_{B1}^a and Y_{B2}^a), the ratio was reduced but within the range of 0.01-0.07 and 0.01-0.05 respectively (Fig. 2: a3 and a4). It formed islands of low coverage along the sex chromosome, which slightly differed between the Y_{B1}^a and Y_{B2}^a pools. To further investigate gene loss, we mapped the short reads of (doubled) haploid individuals (with fully differentiated Y) to the female reference genome and retrieved genes with no coverage. Surprisingly, we found a significantly higher number of genes with zero coverage on autosomes compared to the sex chromosome (Fisher's exact tests: $P < 0.001$; Supplementary Table S6) and the gene losses on the sex chromosome were randomly distributed (Supplementary Fig. S2). These results suggest the genes with zero coverage might not be real gene losses, but false positives due to the low (on average 9.2-13x) and possibly uneven coverage distribution along the genome. Overall, we detect no dramatic gene loss on any male pool and mutation accumulation on the Y corresponded to its differentiation level.

We then characterised the level of genetic differentiation between sex chromosomes by analysing the SNP density ratio between each male pool and the female pool. Once recombination between X and Y chromosome stops, Y-specific mutation accumulation is expected to result in higher heterozygosity on the sex chromosomes of XY compared to XX individuals [81]. This expectation, however, should be adjusted for pool-seq data: if there is only one Y haplotype, even though the Y may have more SNPs than any X haplotype, a pool of such XY haplotypes may have fewer SNPs than a pool of XX haplotypes, because the latter have a larger genetic diversity. The SNP density ratio is thus expected to be skewed toward XX females in regions with higher Y differentiation. In line with

the modified prediction, we found no significant difference of SNP density ratio across 12 autosomes between any of the male and female pool comparisons (Supplementary Fig. S3), as well as along the sex chromosome when comparing the XX male and XX female pools (Fig. 2: b1). For the semi-differentiated pool, the SNP density ratio (XY_{B2}^{0}/XX) showed no significant deviation and was within the 95% CI (Fig. 2: b2). The fully differentiated pools had SNP density ratios (XY_{B1}^{a} /XX and XY_{B2}^{a} /XX) that were significantly biased towards females along most of the sex chromosome, except the chromosome ends (Fig. 2: b3 and b4).

Next, we computed the average F_{ST} between each male and the female pool, in sliding windows (window size = 10,000bp; stride size = 5,000bp). As expected, F_{ST} was near zero (median: 0.00104 – 0.00167) and homogeneous across all 12 autosomes (Supplementary Fig. S4) in all comparisons. F_{ST} along the sex chromosome was not significantly different from the F_{ST} of autosomes in the XX male and XX female pool comparison (permutation test P = 0.054, Fig. 2: c1, d1). Excitingly, we detected a very narrow elevated F_{ST} peak (median = 0.0663, light blue, 4.64 Mb, ~ 0.67% of the chromosome) when comparing the semi-differentiated male (Y_{B2}^{0}) pool to the female pool, which we define as the sex-linked small non-recombining region (henceforth small NRR) (permutation test P = 0.001). F_{ST} was near zero and homogenous for the rest of the sex chromosome (median = 0.00656) but slightly higher than the autosomes (median = 0.00153; permutation test P = 0.001; Fig. 2: c2, d2; Supplementary Fig. S4), consistent with it being the pseudo-autosomal region (PAR). When comparing the two male pools of fully differentiated Y haplotypes Y_{B1}^{a} or Y_{B2}^{a} to the female pool, F_{ST} along the entire sex chromosome was significantly elevated (median = 0.0511, 0.0558) except its ends, which was still significantly elevated compared to the autosomes (permutation test P = 0.001 respectively; Fig. 2: c3, d3, c4 and d4; Supplementary Table S7). We identified the boundaries of the elevated F_{ST} region with changepoint analysis, and found both boundaries were at a similar position in the two Y haplotypes (Y_{B1}^{a}:39,910,000, Y_{B2}^{a}:36,385,000; Y_{B1}^{a}: 664,610,000, Y_{B2}^{a}:654,095,000). Again, we refer to the elevated F_{ST} region as large NRR, and the end of chromosome as the PAR. As with the semi-differentiated Y pool, F_{ST} of the PAR of the fully differentiated haplotypes was slightly higher than that of the autosomes (median = 0.00241 and 0.00645; permutation test P = 0.001 respectively). Taken together, we found three distinct levels of variation in F_{ST} between male pools and the female pool, ranging from almost the entire sex chromosome (618 – 625 Mb, except both ends in the fully differentiated Y) to a very small region (4.64 Mb in the semi-differentiated Y), and no differentiation (in XX males). Along with the results of read coverage ratio and SNP density ratio between the two sexes, this is good evidence of extreme heterogeneity in homomorphic X-Y chromosome differentiation in a single pond.

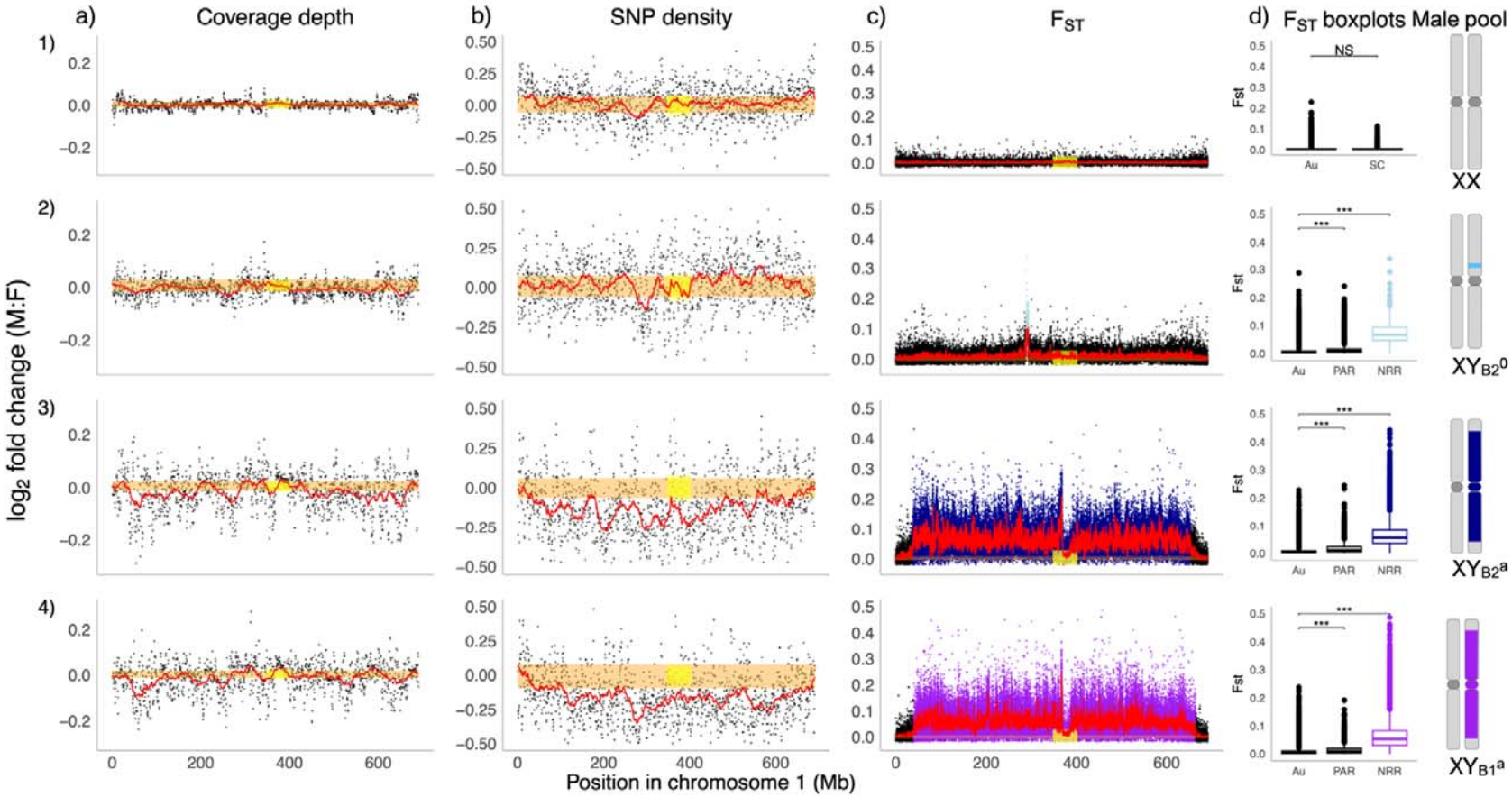

Fig. 2 | Genomic signatures of sex-chromosome differentiation across male pools. Read-coverage ratio (a), SNP-density ratio (b), F_{ST} across the sex chromosome (c), and F_{ST} boxplots across genomic compartment (autosome, PAR, NRR) (d) are shown for four male pools relative to the female pool: an undifferentiated XX male pool (1), a semi-differentiated Y_{B2}^0 pool (2), and two fully differentiated Y_{B2}^a (3) and Y_{B1}^a (4) pools. The red horizontal line denotes the rolling mean (window size = 50) along the sex chromosome, and the orange band marks the central 95% of autosomal rolling-mean values. Yellow shaded regions indicate putative centromeres. For panels (a) and (b), each point represents the male: female ratio of mean read coverage or SNP density calculated in 1-Mb windows with 500-kb step size. For panel (c), F_{ST} values are averaged within 10-kb windows with 5-kb step size. Au: autosome, PAR: pseudoautosome region, NRR: non-recombining region.

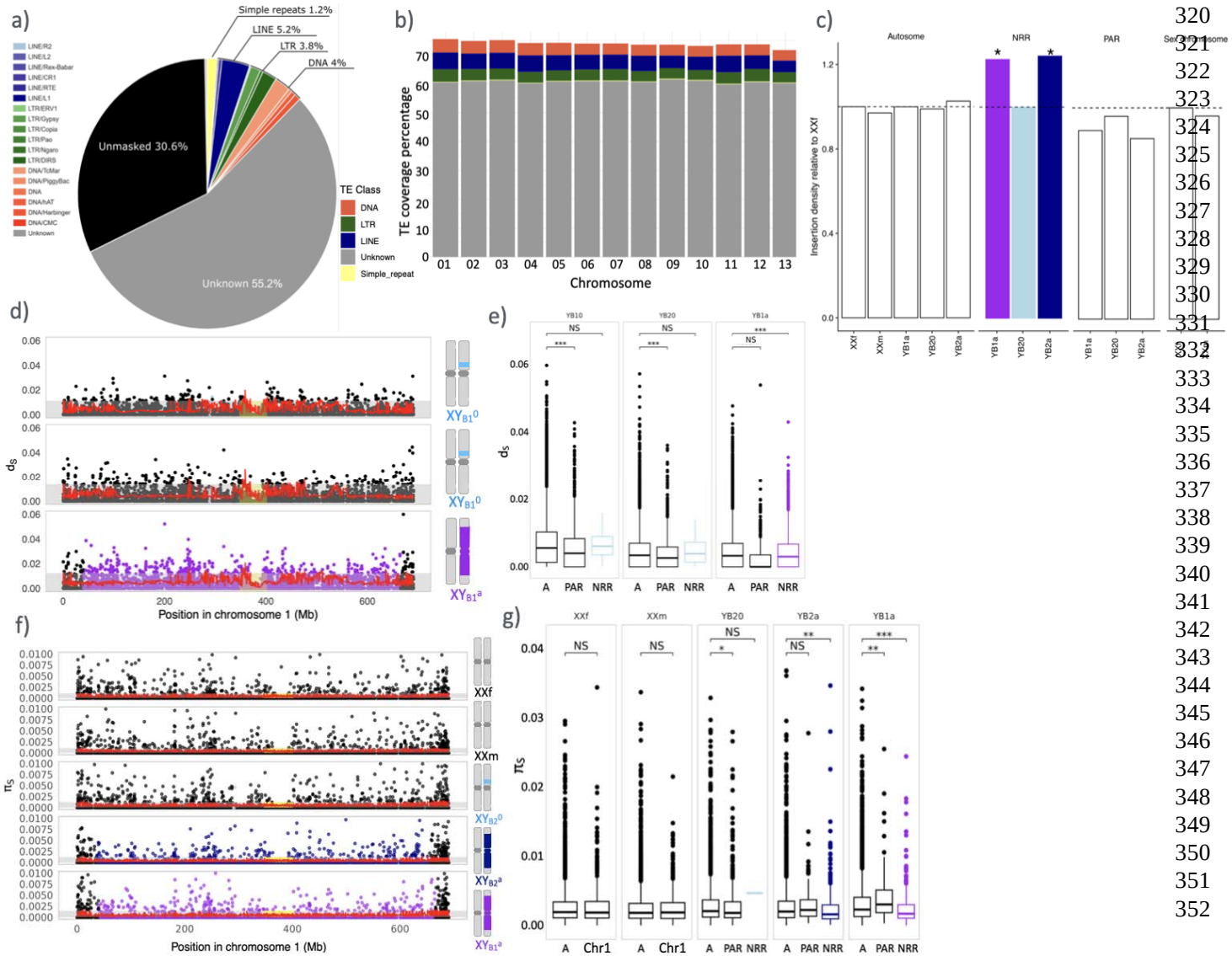

Fig. 3 | Genomic signatures of degeneration on the sex chromosome. (a) Proportion of major transposable element (TE) classes and simple repeats in the reference genome, based on a *de novo* species-specific TE library. (b) Proportion of major TE classes across the 13 chromosomes. (c) Relative TE insertion rates for the 4 male pools comparing to the female pool. (d) Synonymous divergence (d_S) along the sex chromosome for two semi-differentiated Y haplotypes and one fully differentiated Y haplotype. (e) d_S comparison across different genomic compartments (A: autosome, PAR, and NRR) within each Y haplotypes. (f) Synonymous nucleotide diversity (π_S) along the sex chromosome for the female pool and four male pools. (g) π_S comparison across different genomic compartments within each Y haplotypes.

Significantly higher TE insertions in the fully differentiated Y haplotypes

To assess the level of Y chromosome degeneration, we further quantify the accumulation of transposable elements (TEs) and repeats, non-synonymous and synonymous substitution rate (d_N/π_N, d_S/π_S) across the genome and among the identified genomic compartments of different X-Y differentiation within the sex chromosome. TEs and repeats would accumulate more on the NRR due to less efficient purifying selection. We generated a *de novo* TE library of the XX reference genome of *R. temporaria*, which makes up 69.4% of the genome and the vast majority (80.9%) is unknown (Fig. 3a). Divergence analysis, estimated from Kimura substitution level, suggests a recent burst of TE activity (Supplementary Fig. S5). However, there was no enrichment in any TE class on the X chromosome compared to the autosomes (Fig. 3b). We then computed TE insertions (i.e. insertion rate as the number of detected TEs per Mb) for all four male pool-seq datasets relative to the XX female pool-seq data using PopoolationTE2. We found no significant differences in TE insertions among the male pools for all autosomes (Glm, P = 0.524, Fig. 3c). For the fully differentiated male pools (Y_{B1}^a, Y_{B2}^a), TE insertion was significantly higher in the NRR compared to the autosomal regions (Glm, P < 0.003, Fig. 3c). This was, however, not the case for the semi-differentiated male pool (P = 0.504, Y_{B2}^0).

Synonymous divergence (d_S) and polymorphism (π_S) are commonly used to infer the timing of recombination suppression and divergence between sex chromosomes, whereas ratios of nonsynonymous to synonymous substitutions (d_N/d_S, π_N/π_S) are used to infer the mode of selection. We computed d_S of the X-linked genes between the reference XX genome and Y-linked genes from the doubled haploid individuals (Y_{B1}^0, Y_{B2}^0 and Y_{B1}^a). d_S among autosomal genes were uniformly low across all pairwise comparisons between each male and the reference genome (median = 0.0034-0.0055, Supplementary Table S8a, Fig. S6). Similarly, d_S remained consistently low along the sex chromosome NRR (median = 0.0026-0.0061) and fell within the 95% CI for all males examined (Y_{B1}^0, Y_{B2}^0 or Y_{B1}^a; Fig. 3d). Consistent with this pattern, d_S did not differ significantly between genes located in the NRR and those on the autosomes (permutation test, P = 0.79, 0.79, 0.10 respectively). Changepoint analysis based on d_S identified one compartment for the entire sex chromosome. Regarding the d_N/d_S, overall, there was no significant difference between sex chromosome compartments (NRR, PAR) and autosomes (permutation test, all P > 0.82; Supplementary Fig. S7, S8, Table S8b), except significant higher ratio on the PAR of fully differentiated YB_1^a (permutation test, P = 0.0006). We thus find no strong support for Faster-X evolution on this homomorphic sex chromosome system.

Regarding π_S, no significant difference was detected between sex chromosomes (compartments) and autosomes for XX male and XY_{B2}^0 male pools (permutation test, all P > 0.23, Fig. 3f & 3g, Supplementary Table S9a, Fig. S9), but π_S was significantly lower for the NRR for both fully differentiated XY_{B1}^a and XY_{B2}^a male pools (permutation test P < 0.0035, Supplementary Table S9a). Interestingly, π_S, was slightly but significantly higher for the PAR than for autosomes (permutation test, P = 0.0035). However, none of the compartment on sex chromosomes was significantly different from autosomes for π_N/π_S for all four male pools (permutation test, all P > 0.29, Supplementary Table S9b, Fig. S10-S12). Taken together, comparing to autosomes, we find evidence of significantly lower π_S on NRR of two fully differentiated male pools (Y_{B1}^a, Y_{B2}^a) but with no effect on π_N/π_S ratio. Furthermore, no NRR showed significant difference in synonymous substitution accumulation (d_S) or in the d_N/d_S ratio for both fully differentiated and semi-differentiated males. These results suggest that mutations have not accumulated substantially on the NRR coding regions of Y_{B1}^a, Y_{B2}^a, and Y_{B2}^0, and therefore have not markedly influenced X-Y divergence.

No clear sign of Y-linked gene expression decay and no enrichment of sex-biased genes across the various Y haplotypes

To investigate the extent of Y gene activity decay across various Y haplotypes, we estimated major allele ratio (MAR) patterns from RNA-seq data in gonad and brain tissues from Meitreile. If the X and Y chromosomes are transcriptionally active at the same level, we expect a probability of around 0.5 of recovering mRNA from either chromosome. As expected, we found that females have a similar MAR of autosomal and sex-linked genes (0.56) in both gonad and brain tissues (permutation test, all P = 0.15 and 0.13 respectively, Fig. 4a, 4b). Similarly, no significant deviations of MAR patterns for sex-linked genes were detected for males of all Y differentiation types in gonad (permutation test, all P > 0.20, 0.58-0.59), and in brain tissue (permutation test, all P > 0.21, 0.56-0.57) (Fig. 4a, 4b). Very similar patterns of rolling means of MAR were detected along the sex chromosome among females and males with various Y haplotypes (Supplementary Fig. S13). Next, we looked for differential gene expression in pairwise comparisons of all male genotypes representing different Y differentiation levels. In both brain and testis tissues, no differentially expressed (DE) genes were detected across comparisons ($\log_2$FC(male/male) >= 1, FDR < 0.05), with only a single gene identified in each tissue as an exception (Supplementary Fig. S14-15, Table S10). As all individuals were phenotypic males with comparable traits, the near absence of significant differential expression suggests that Y-linked gene expression does not differ detectably among the various Y haplotypes. Taken together, our MAR and differential gene expression results provide no evidence of Y gene activity decay.

To investigate whether sex-biased gene enrichment has evolved on the sex chromosomes, we compared gene expression between males of each Y differentiation level and females, using brain and gonad tissues. Only one DE gene (on Chr13) was detected for brain tissue so chromosomal enrichment analysis was not performed. For gonads we detected 6598, 6449 and 6057 female-biased genes, as well as 8106, 8032, and 7554 male-biased genes when comparing respectively XY, XY^0 and XX males to females (|$\log_2$FC(male/female)| >= 1, FDR < 0.05, Supplementary Fig. S16). We then conducted permutation tests for both gene expression magnitude ($\log_2$ fold change), and the number of sex-biased genes, between the sex chromosome and autosomes. With one exception (see below) there were no significant differences between the sex chromosome and autosomes in female- and male-biased genes in any of the contrasts ($\log_2$ fold change permutation test, all P > 0.31, Supplementary Fig. S17, Table S11-12; gene number permutation test, all P > 0.33, Fig. 4c, Supplementary Table S13-14). We also separately compared the NRR and PAR, defined from F_{ST} (Fig. 2), to autosomes for the semi-differentiated (Y_B^0) and fully differentiated (Y_B^a) males. Again, there were no differences in gene expression magnitude (Fig. 4d, permutation test, all P > 0.15; Supplementary Table S15) or DE gene number (permutation test, all P > 0.12; Supplementary Fig. S18, Table S16) for all contrasts.

Only the magnitude of female-biased gene expression, but not the number of DE genes (permutation test, P = 0.120), was higher on the PAR than on autosomes in XY male testes (permutation test, P = 0.00045). Because the PAR is largely located in telomeric regions with high gene density, and because we observed significantly elevated synonymous polymorphisms π_S (Fig. 3g), likely reflecting enhanced crossover rates in males relative to females, we asked whether this pattern could be a general property of telomeric regions. To test this, we performed analogous permutation tests on each autosome by partitioning them into a “bulk” region and a distal telomeric region (approximately 30 Mb). Three (Chr3, Chr6, Chr7) of the 12 autosomes showed a similar enrichment of female-biased expression magnitude in telomeric regions in both fully-differentiated XY_B^a males and semi-differentiated XY_B^0 males (permutation test, all P < 0.005). Finally, applying the PAR boundaries defined from the Y_B^a uniformly across all three male genotypes (XY_B^a, XY_B^0, XX), we consistently detected a significantly higher magnitude of female-biased gene expression in this

region in all male types (permutation test, all $P < 0.005$; Supplementary Table S17). Collectively, these patterns are consistent with telomeric regions being gene-rich and have higher crossover rate in males than females, which may contribute to the observed enrichment of female-biased expression.

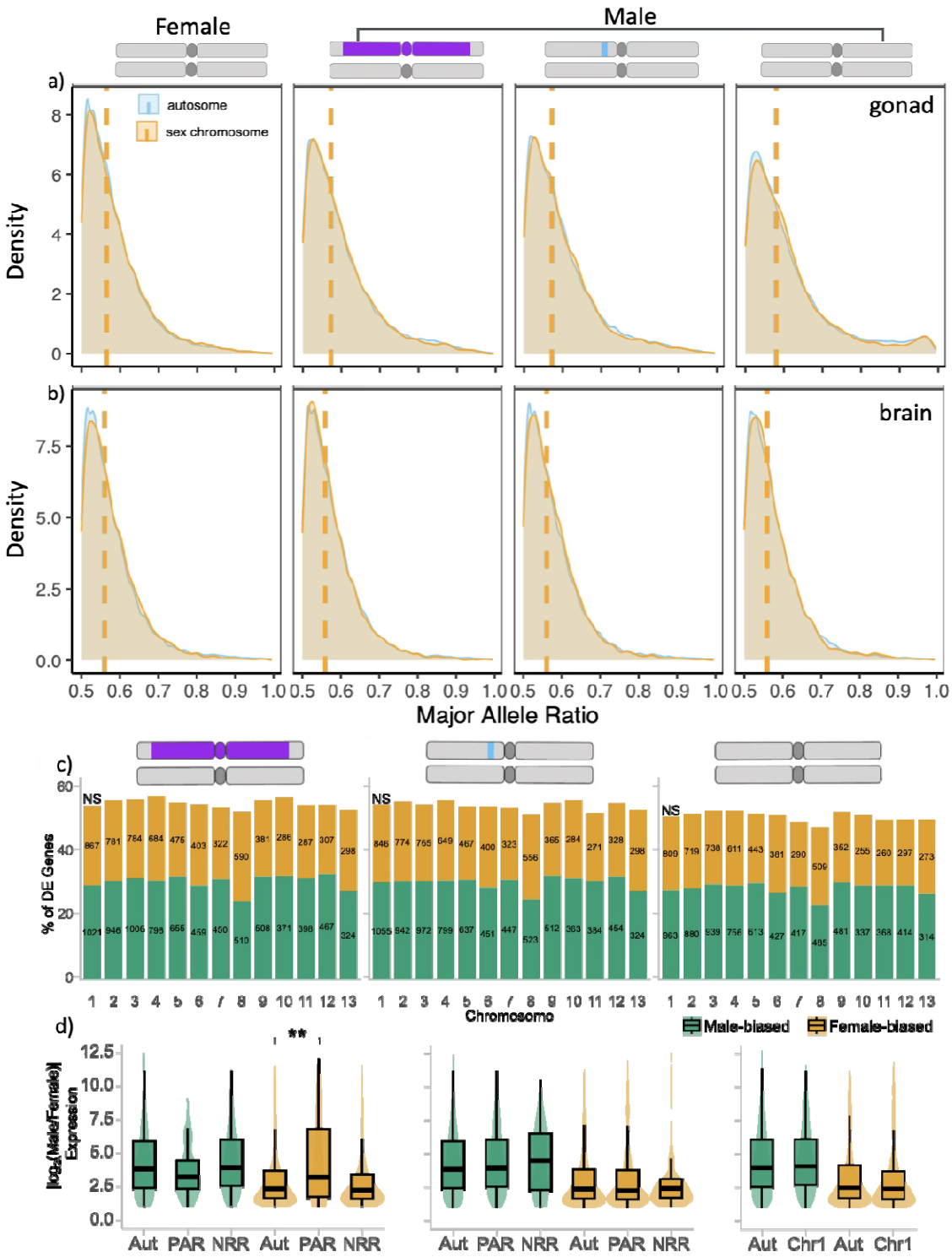

Fig. 4 | Distributions of major allele ratios and sex-biased gene expression across males with three distinct levels of Y-chromosome differentiation, illustrated by the cartoon where purple indicates a large NRR and blue a small

NRR, ordered from fully-differentiated Y to semi-differentiated Y, to undifferentiated Y. Density plots show the distribution of major allele ratio (MAR) for autosomal (blue) and sex-chromosomal (orange) genes in gonad (a) and brain (b) tissues of females and males. Vertical dotted lines denote median MAR values. (c) Proportion of female-, and male-biased genes per chromosome among three types of males. (d) Boxplot of differential gene expression (in log2(male/female)) among different genomic compartments among three types of males (Aut: autosome, NRR: non-recombining region and Chr1: sex chromosome).

Support for *Dmrt1* as the candidate master sex-determining gene

We detected an elevated F_{ST} peak, over ~4.64 Mb, between the semi-differentiated male (Y_{B2}^{0}) and the female pools (Fig. 2c2, Fig. 5a). As discussed in the introduction, in species with extreme heterochiasmy, a newly originating sex-determining gene would result in immediate recombination arrest between the X and that Y chromosome haplotype and a fully differentiated Y haplotype with elevated F_{ST} except the sex chromosome ends. The only way to evolve a small, differentiated Y region is explained by the 'fountain of youth' model where recombination is controlled by phenotypic (not genotypic) sex and sex-reversed XY females restore X-Y recombination. Changepoint analysis further divided the high F_{ST} 4.64 Mb region into two strata with distinctive mean F_{ST} (Fig. 5a, Wilcoxon test, $P < 8.4e-11$): the high differentiation (blue) stratum is ~3.25 Mb and the intermediate differentiation (dark blue) stratum is ~1.39 Mb.

The high differentiation stratum contains 19 coding genes (including *Dmrt1*) while the low differentiation stratum contains 10 coding genes, based on the XX genome annotation (Fig. 5a, Supplementary Table S18). Functionally, 2 of those 29 genes stand out: *Dmrt1* and *Spata6l* (Supplementary Table S18) which have sex determination and spermatogenesis functions respectively. When looking at sex-specific expression in gonads from Meitreile samples (Fig. 5a, 5b), 2 genes stand out. One is again *Dmrt1*, which has testis-specific expression. The other is *LOC120943612*, which has ovary-specific expression and is an ortholog of anuran class-A G protein-coupled receptors (GPCRs), which respond to small ligands and regulate intracellular signalling. The gene seems to be involved in steroidogenesis and ovarian signalling. *Dmrt1* is a transcriptional factor and its function in male cell development and testis development is highly conserved across all major vertebrate groups in fish, amphibians, non-avian reptiles and birds [82]. Master sex-determining genes, often show sexually dimorphic gene expression during early embryogenesis prior to gonad sexual differentiation. Using published developmental RNA-seq data from other populations (Ammarnäs with fully differentiated XY, Tvedöra with semi-differentiated XY^{0}) [75,76], *LOC120943612* shows low expression in early larval stages but strong female-specific expression at the froglet stage (Supplementary Fig. S19), while *Dmrt1* shows near-zero expression for all developmental stages including metamorph and froglet stages (Supplementary Fig. S20). This could be due to tissue allometric scaling issues that affected the whole-tadpole RNA-seq used [83]. Future progress can be made using gonadal-only RNA-seq of samples at early developmental stages.

The sex-determining gene can be identified by population genetic patterns in its genetic variation. It is expected to be heterozygous in all males, which should form a single genetic cluster. We used SNPs obtained from RNA-seq data from Meitreile (gonads) to identify genes showing these patterns. Dendrograms were generated by hierarchical clustering of samples using Euclidean distance calculated from genotypic data. Clustering analyses showed male-specific clusters, separate from the XX genotype only for the *Dmrt1*, *Dmrt3* and *Dmrt2* genes from the entire sex chromosome (Fig. 5c). Note that since those genes were largely expressed in testes only, XX genotype SNPs were obtained from XX males. Furthermore, we analysed heterozygosity in three ways. First, there were 55 SNPs showing homozygosity in XX females and heterozygosity in XY_{B}^{a}, XY_{B}^{0}, XX males, which were

randomly distributed across the genome (Supplementary Table S19). Second, we identified 25 SNPs, with homozygosity in XX males but heterozygosity in XY_B^0 and XY_B^a males, and most (21 out of 25) were on *Dmrt1* or between *Dmrt1* and *Dmrt3* (Supplementary Table S20). Finally, only 1 SNP, with homozygosity in XX individuals (females and males), but heterozygosity in XY_B^0 and XY_B^a males, was identified, which was located between *Kank1* and *Dmrt1* (Supplementary Table S21). The results point to *Dmrt1* being the master sex-determining gene but functional studies like RNAi or CRISPR-Cas9 are needed to confirm this.

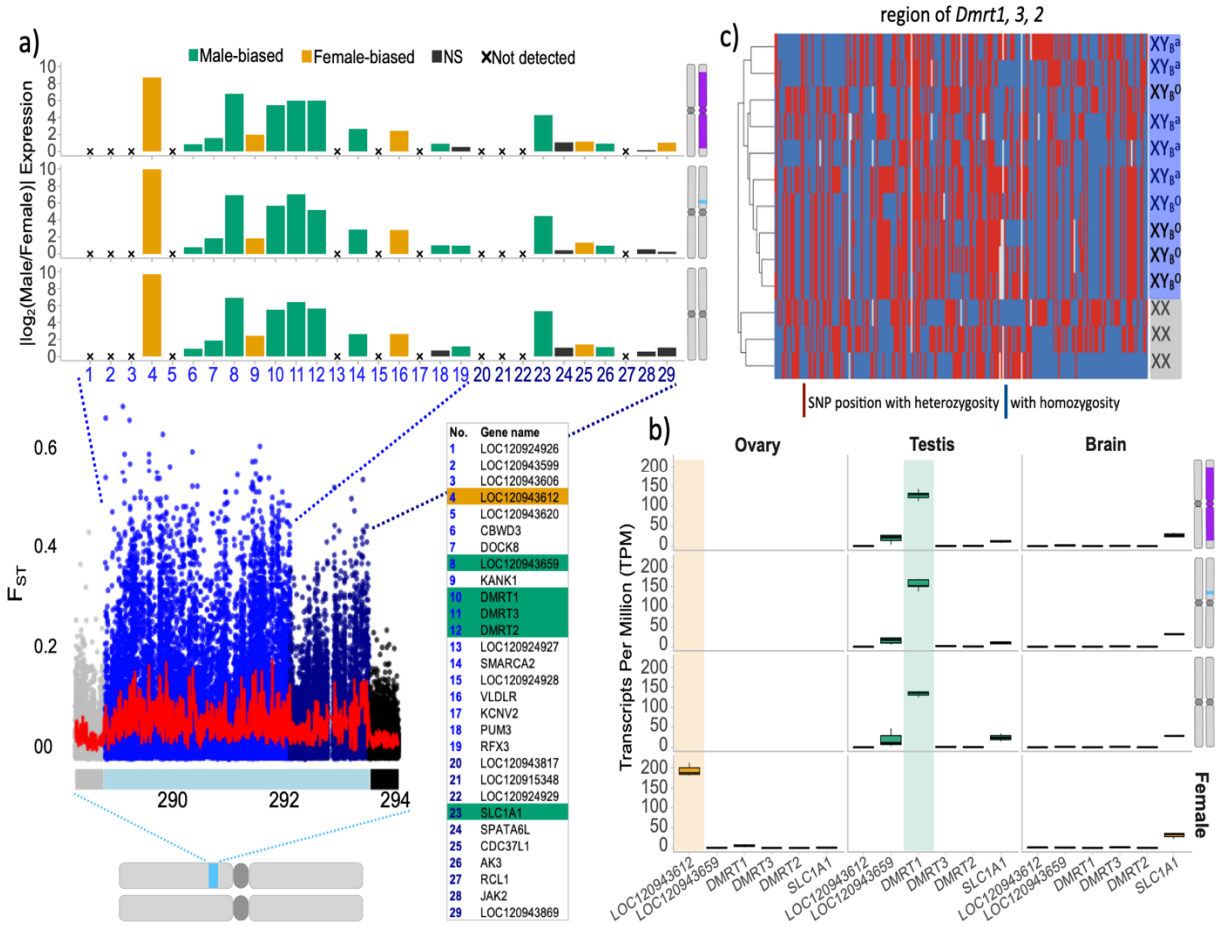

Fig. 5 | The elevated F_{ST} peak between males carrying the semi-differentiated Y_{B2}^0 haplotype and females is subdivided into two strata, a light-blue and a dark-blue region. (a) Differential expression analyses identify male- and female-specific or biased genes in adult testis versus ovary across three male types representing distinct stages of Y-chromosome differentiation. (b) Expression profiles of the top six candidate genes, including *Dmrt1*, display pronounced sex-specific or sex-biased expression. (c) A clustering dendrogram based on homozygosity in XX individuals and heterozygosity in XY individuals including both fully and semi-differentiated Y haplotypes was constructed from SNPs derived from the RNA-seq datasets.

Table 2. Estimated age of small and large NRR regions reflected by the highly elevated F_{ST} region from Fig. 2. The large NRR was subdivided into small NRR and the rest of large NRR. The age estimation and the confidence interval was based on the d_S between the YY doubled haploid males with Y_{B2}^0 or Y_{B1}^a haplotypes and the reference X chromosome sequence. Mutation rate per synonymous site per year (0.776×10^{-9}) was based on the Tibetan frog *Nanorana parkeri* [84].

Y haplotype	Small NRR	d_S (mean ±se)	Age (Myr)	Age with 95%CI	Rest of large NRR	d_S (mean ±se)	Age (Myr)	Age with 95%CI
Y_{B2}^0	100%	$(26.9\pm7.6) \times 10^{-4}$	3.47	1.55-5.39	100%	$(7.4\pm0.5) \times 10^{-4}$	0.95	0.82-1.08
Y_{B1}^a	100%	$(44.3\pm8.2) \times 10^{-4}$	5.70	3.62-7.78	100%	$(26.9\pm7.6) \times 10^{-4}$	3.13	2.94-3.32

Single origin and evolutionary history of Y haplotypes dynamics

The evolutionary divergence among the X, Y_{B2}^0, Y_{B1}^a and Y_{B2}^a haplotypes was examined using complementary approaches. First, we analysed the evolutionary history of the various Y haplotypes using pairwise male-male comparisons among the male pools. The average F_{ST} per sliding window (window size = 10,000 bp; stride size = 5,000 bp) across the genome showed distinct patterns of differentiation on the sex chromosomes, but no significant difference for the near zero F_{ST} across 12 autosomes (i.e. comparison between XX male and XY_{B1}^a male pools, Supplementary Fig. S21). Comparing to the XX male pool, F_{ST} was significantly and consistently elevated along the large NRR for Y_{B1}^a and Y_{B2}^a haplotypes, including the small NRR (Wilcoxon test, FDR corrected $P < 0.0001$, Fig. 6a, 6b). Interestingly, when comparing the Y_{B2}^0 male to the XX male pools, elevated F_{ST} formed a sharp peak only at the small NRR along the sex chromosome (Wilcoxon test, $P < 0.0001$, Fig. 6c). When male pools with Y haplotypes (Y_{B2}^0, Y_{B1}^a and Y_{B2}^a) were compared, F_{ST} was lowest (near zero) at the small NRR (Fig. 6d, 6e, 6f), and no significant differences were detected among them (Wilcoxon test, FDR corrected, all $P > 0.29$; Supplementary Fig. S22). The boundaries of this low F_{ST} region coincided with those of the small NRR across all pairwise male pool comparisons, and encompassed both the high (blue) and the intermediate (dark blue) F_{ST} strata identified by changepoint analysis (Fig. 6g; Fig. 2c2). The result is compatible with a very recent and shared origin of the small NRR (3.47-5.70 Mya, Table 2). In contrast, within the large NRR (excluding the small NRR), genetic differentiation F_{ST} among male pools was highest between Y_{B1}^a and Y_{B2}^a (0.116 ± 0.089), intermediate between Y_{B1}^a and Y_{B2}^0 (0.061 ± 0.050), and lowest between Y_{B2}^a and Y_{B2}^0 (0.045 ± 0.034) (Wilcoxon test, FDR corrected $P < 0.0001$ for all comparisons; Fig 6d-f, Supplementary Fig. S22). This gradient of differentiation is consistent with Y_{B1}^a being the oldest NRR, followed by Y_{B2}^a and Y_{B2}^0 (i.e. Table 2).

To independently infer the evolutionary relationships among various Y haplotypes, we performed dendrogram-based clustering analyses in sliding windows along the sex chromosome, using SNPs derived from RNA-seq data that follow the expected pattern of XX-specific homozygosity and XY heterozygosity (Supplementary Fig. S23-S27). Because the clustering analyses are based on patterns of SNP heterozygosity and homozygosity inferred from transcriptomic data, they capture divergence among Y-linked variants but do not resolve specific SNP states or individual Y haplotypes. Clustering analyses were conducted in sliding windows across PAR1 (Supplementary Fig. S23a-d), large NRR part1 (Supplementary Fig. S24a-d), small NRR (4.64 Mb, Supplementary Fig. S25a-c), large NRR part2 (Supplementary Fig. S26a-g) and PAR2 separately (Supplementary Fig. S27ffa-c,

Fig. 8). In PAR1 and PAR2, clustering based on genotype-specific heterozygosity was random with respect to sex, consistent with ongoing recombination. In contrast, within the large NRR part1 and part2, males with fully-differentiated Y largely clustered together, whereas XX males grouped with females. Notably, one fully-differentiated male clustered outside the main Y clade near the PAR1 boundary, suggesting occasional recombination that may locally extend the PAR1 boundary. Only in the *Dmrt1*, *Dmrt3 and Dmrt2* region, all male samples with Y-haplotypes clustered (fully- and semi-differentiated Ys) together. This independent confirmation of male clustering in the *Dmrt* region (*Dmrt1*, *Dmrt3*, *Dmrt2*) strengthens *Dmrt1* as the primary candidate master sex-determining gene in this species (Fig. 7, Fig. 5c) and additionally indicates a shared origin of the sex-determining locus in both fully-differentiated and semi-differentiated Y haplotypes.

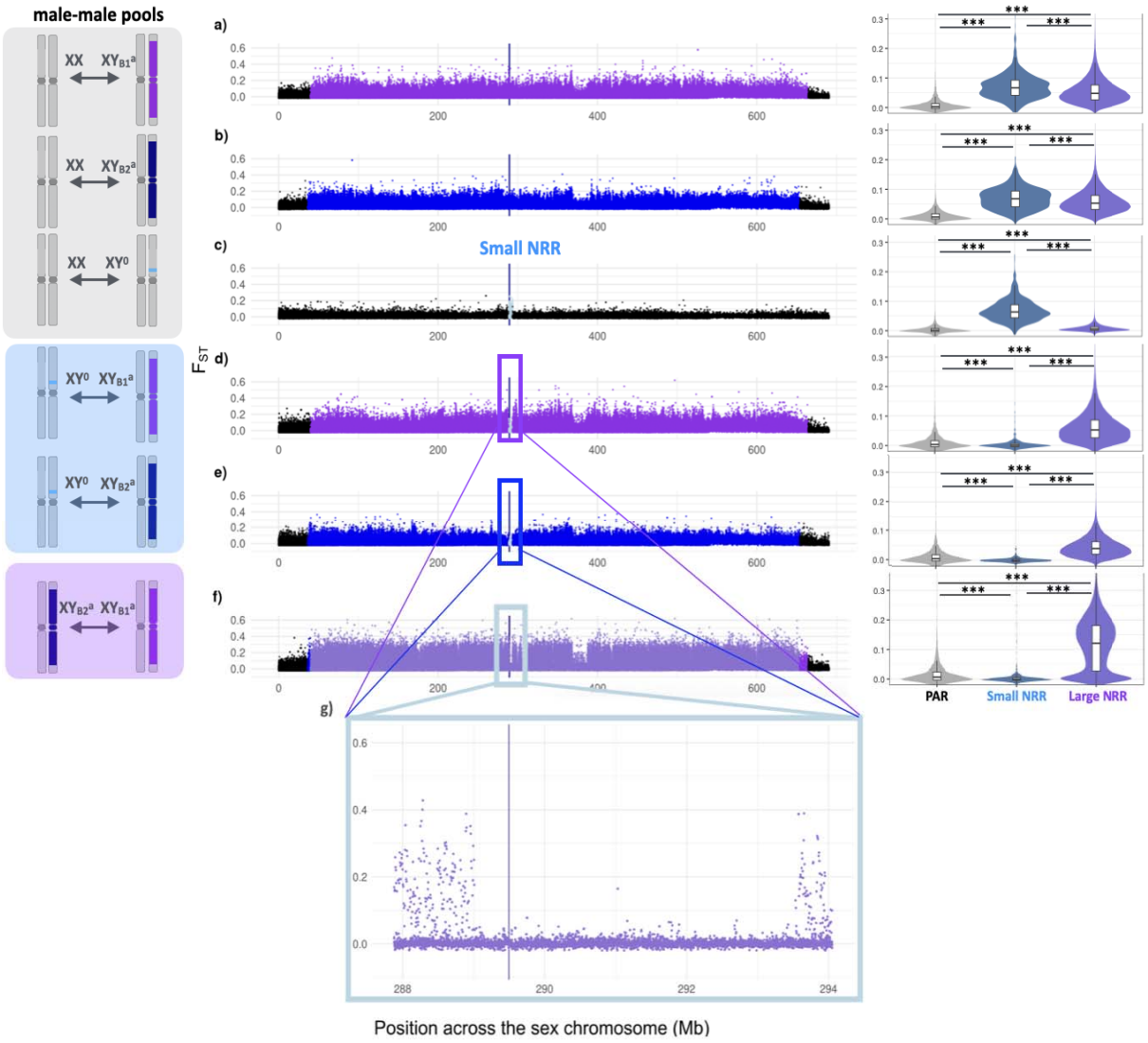

Fig. 6 | Pairwise F_{ST} comparisons among male pools across the sex chromosome. a) XY_{B1}^{a} males compared with XX male pools in sliding window and boxplots among genomic compartments (PAR, small NRR and large NRR subtracting small NRR), b) XY_{B2}^{a} males compared to XX male pools, c) XY_{B2}^{0} males compared to XX male pools, d) XY_{B1}^{a} males compared to XY_{B2}^{0} male pools, e) XY_{B2}^{a} males compared to XY_{B2}^{0} male pools, and f) XY_{B1}^{a} males compared to XY_{B2}^{a} male pools. g) A magnified view of the small NRR (4.64 Mb) identified from

the early F_{ST} peak in the XY_{B2}^{0} – XX male pools comparison reveals near-zero F_{ST} among the semi-differentiated and fully-differentiated Y haplotypes (corresponding to comparisons in panels d, e. and f).

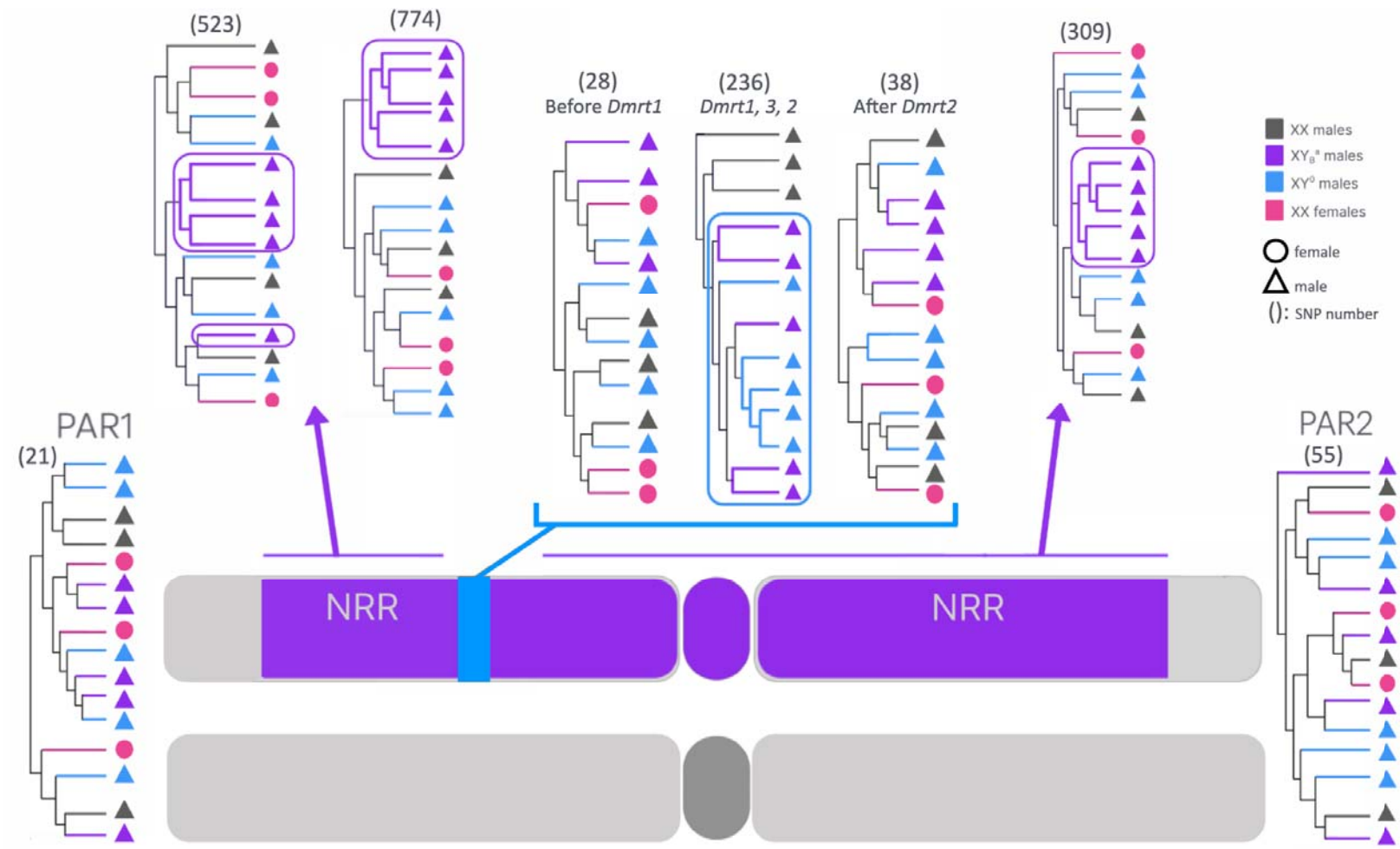

Fig. 7 | Representative clustering dendrograms based on SNPs derived from RNA-seq read mappings of male and female gonad and brain tissues to the reference genome, illustrating patterns of homozygosity and heterozygosity across discrete compartments of the sex chromosome (PAR1, large NRR part1, small NRR of 4.64 Mb, large NRR part2 and PAR2). Each dendrogram was generated from SNP sets extracted using a 10-Mb sliding window across PAR1 and PAR2, a 50-Mb sliding window across the NRR, and three subdivided windows spanning small NRR. Number of SNPs used in each clustering analysis is indicated in parentheses.

Finally, we estimated the age of the NRR (4.64 Mb for semi-differentiated Y and 618-625 Mb for fully-differentiated Y) using the method described by [85], based on synonymous substitution d_S values across each region divided by the annual mutation rate. For the NRR of the fully-differentiated Y, we analysed the shared small NRR (4.64 Mb including *Dmrt1*) separately from the remaining large NRR. The shared small NRR (4.64 Mb) was dated to approximately 3.47-5.70 Mya based on d_S values, between doubled haplotype Y-linked sequences and reference X sequences, of all related coding genes of Y_{B2}^{0} and Y_{B1}^{a}, after correcting for within-XX male polymorphism (Table 2). In contrast, divergence estimates for the remaining NRR varied among haplotypes, with older ages inferred for Y_{B1}^{a} (2.94–3.32 Mya) and substantially younger estimates for Y_{B2}□ (0.82–1.08 Mya). Assuming a generation time of approximately five years, the most recent inferred XY sex-reversal event occurred roughly 200,000 generations ago and was not associated with detectable genetic differentiation for the large NRR.

Based on these results, we propose a model for the evolutionary history of Y-chromosome haplotypes that explains the relative age and their dynamics in the single Meitreile pond (Fig. 8). Under male-specific chromosomal end-restricted recombination, the emergence of a master sex-determining gene, likely *Dmrt1*, could have promoted independent differentiation of various Y haplotypes (e.g., Y_{B1}^{a}, Y_{B2}^{a}) across most of the sex chromosome, except the chromosomal end. The shared small NRR likely evolved early (3.47-5.70 Mya), immediately followed by independent mutation accumulation on the large NRR for each Y haplotype. The near-zero F_{ST} observed in the small NRR between male pools carrying semi- and fully differentiated Y haplotypes is striking, indicating that Y haplotypes within this region are nearly identical despite an estimated absence of recombination for approximately 3.47–5.70 Myr. The non-overlapping divergence time estimates are consistent with temporarily distinct XY sex-reversal events following the establishment of the small NRR, including an older event associated with Y_{B1}^{a} (~3.13 Mya) and then with Y_{B2}^{a} (likely <2.94 Mya), and a more recent event giving rise to the semi-differentiated Y_{B2}^{0} haplotypes (0.98-1.08 Mya, Fig. 8, Table 2). The Y haplotypes subsequently accumulated haplotype-specific divergence from the X chromosome, whereas occasional XY sex reversals are inferred to have reintroduced recombination, eroding differentiation and generating semi-differentiated Y haplotypes. This process is consistent with a dynamic cycle between fully differentiated and semi-differentiated Y chromosomes. In contrast, XX males likely arise continuously through non-genetic sex determination, potentially driven by stochastic processes or epigenetic regulation [86], and remain undifferentiated due to ongoing recombination in XX females. Together, ancient and ongoing XY sex-reversal events are inferred to recurrently reset X–Y divergence, thereby promoting diversity in Y chromosome haplotype length and enabling the long-term coexistence of genetic (XY, XY□) and non-genetic (XX) male phenotypes within the population.

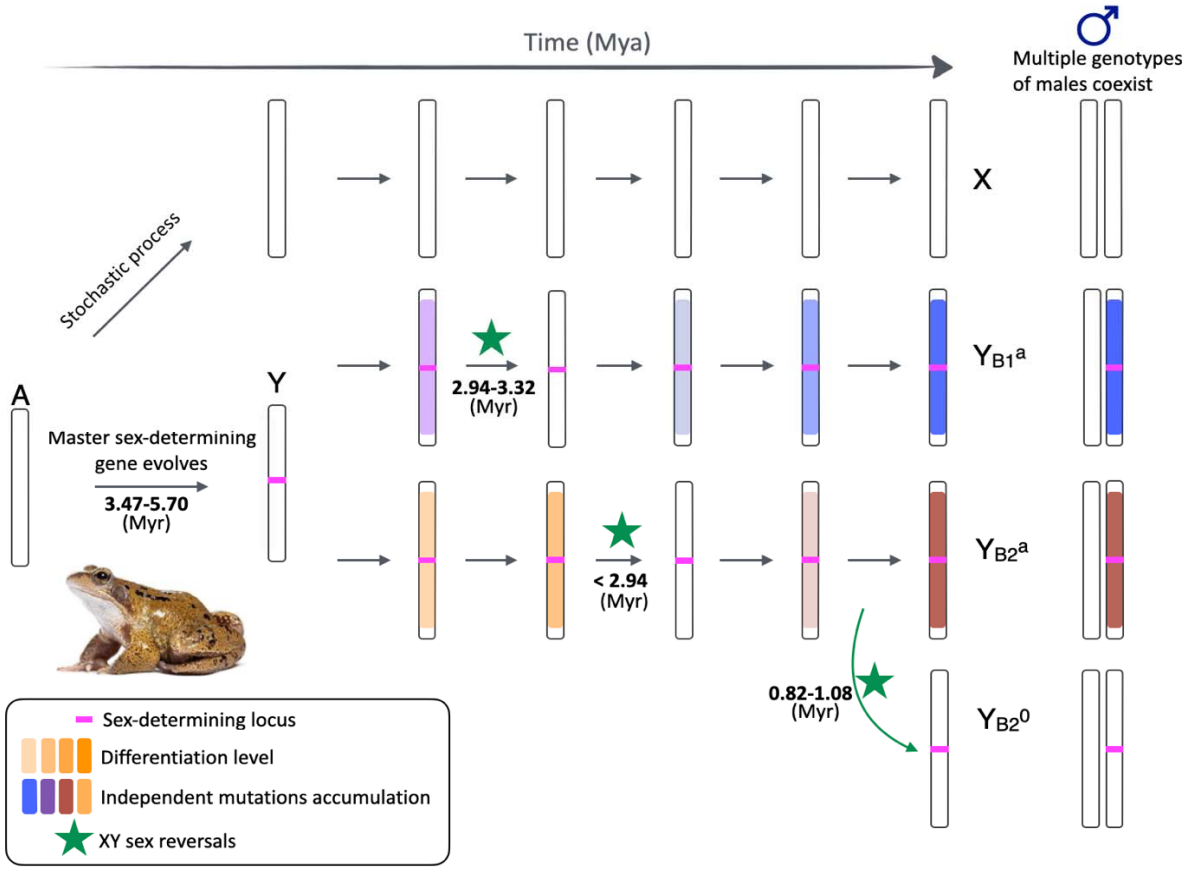

Fig. 8 | Proposed evolutionary history of Y-haplotype lineages in the Meitreile population. XX males may arise stochastically through random sex determination, during which no detectable sex-linked differentiation is present. A shared, *Dmrt1*-based small non-recombining region (NRR; 4.64 Mb) subsequently evolved approximately 3.47-5.07 Mya. Following the establishment of the small NRR, further suppression of recombination led to the expansion of a large NRR and the emergence of fully differentiated Y_{B1}^a haplotypes under extreme heterochiasmy. Divergence time estimates for the large NRR are consistent with an ancient XY sex-reversal event occurring approximately 2.94–3.32 Mya. Differences in genetic differentiation among Y haplotypes further suggest a subsequent XY sex-reversal event associated with Y_{B2}^a occurring after 2.94 Mya, followed by a more recent event around 0.82–1.08 Mya, giving rise to the Y_{B2}□ haplotype. Chromosomes are coloured according to the degree of differentiation, with darker shading indicating higher differentiation. Different colours indicate independent mutation accumulation among Y haplotypes during the indicated evolutionary period. Green stars denote inferred XY sex-reversal events.

Discussion

In contrast to the canonical model, where NRRs gradually expand via inversions or other cis-acting recombination mechanisms (Fig. 1a), we examine sex chromosomes in a system with extreme heterochiasmy, where arrested recombination can be restored through sex reversals (XY females), offering alternative trajectories of sex chromosome evolution (Fig. 1b). We identified extreme heterogeneity in X-Y divergence (NRR is 618-625 Mb, 4.64 Mb or 0 bp) in a single Swiss Alpine population of *R. temporaria* with homomorphic sex chromosomes. To our knowledge, the two fully differentiated Y_{B1}^a, Y_{B2}^a chromosomes likely harbour the largest non-recombining region (~618 Mb and ~ 625Mb, ~90% of Chr1) known in vertebrates to date, which is nearly double the size of the previous record (~349Mb of NRR ~ 60% of Chr1) in the Emei moustache toad with also XY sex chromosome systems [87]. The presence of a large NRR arising from male-specific, chromosomal end-restricted recombination is consistent with theoretical predictions. In contrast, NRRs in ZW sex

chromosome systems are not expected to be affected in the same way, implying that additional forces or mechanisms must contribute to their evolution. Consistent with this expectation, anurans with well-characterized ZW system generally exhibit comparatively smaller NRRs. For example, NRR sizes range from 10.7 Mb (7.7% of Chr7) in *Xenopus tropicalis* [88] to 54.1 Mb (44.2% of Chr8L) in *X. borealis*[89]. Extending investigations of extreme heterochiasmy to taxa beyond anurans, and across both XY and ZW systems, will be critical for evaluating the generality of the NRR being larger in XY compared to ZW systems.

Beyond the large NRR region detected along the sex chromosome, the PAR1 and PAR2 of *R. temporaria* are encompassing only the chromosome ends, as expected from the pattern of male-specific recombination. This study thus provides the first direct fine-scale genomic evidence of the consequences of males recombining primarily at their chromosome ends, which is documented in most studied anurans including *R. temporaria*[31,42,43,47,88,90,91]. Our data support that new sex-determining genes immediately define a near-zero recombination region in organisms with extreme heterochiasmy, so that recombination arrest does not require sexually antagonistic selection as in the canonical model of sex chromosome evolution[40,92]. We defined the NRR of the previously discovered[69,72] semi-differentiated Y haplotype at high resolution to be only 4.64 Mb. These haplotypes can be explained by historical recombination between the X and Y in sex-reversed XY females, which has been observed by linkage mapping in the offspring of a rare XY mother from the same population[64].

It remains unclear whether male chromosomal end-restricted recombination adaptively evolves to capture sex-beneficial alleles close to the male-determining gene, or whether it is a pre-existing characteristic of male meiosis [44]. Chromosomal end-restricted recombination has been detected in XY males in all studied frog lineages in Ranidea, Hylidae, Bufonidae, and Megophryidae [31,42,93,94], with cytogenetic or sex-specific linkage maps, and also in ZZ males in a few species including *Hyla sarda*, *Hyla savignyi* and Pipidae [43,89], suggesting it is ancestral across anurans. A systematic study of additional ZZ/ZW species across diverse anuran lineages could help further establish this [47]. Extending such analyses to other highly divergent groups with extreme heterochiasmy and their sex-chromosome systems, such as in certain fishes (i.e. guppies, sea horses) and flowering plants (i.e. *Rumex hastatulus*), could further clarify whether extreme heterochiasmy evolves adaptively or as a consequence of molecular or cellular constraints during meiosis, which has broader implications for sex chromosome evolution.

Despite harboring the largest known sex-linked NRR in vertebrates (618–625 Mb), the Y chromosome shows remarkably limited degeneration. Fully differentiated Y haplotypes exhibit only a modest increase in TE insertions, whereas semi-differentiated Y haplotypes and XX males show no difference to XX females. We likewise found no enrichment of sex-biased gene expression within the NRR, no evidence of gene loss, reduced Y-allele expression, or faster-X effect. These results indicate that the Y chromosome remains at an exceptionally early stage of divergence and that the earliest detectable hallmark of differentiation in this system is the accumulation of TEs and other repetitive elements, consistent with patterns reported for young sex chromosomes in flowering plants such as *Silene latifolia* and multiple fish lineages, where repeat proliferation typically precedes substantial gene loss, dosage-compensation evolution or strong sexualization of gene expression [95,96].

Across highly differentiated sex chromosome systems, X chromosomes often become enriched for female-biased genes and Z chromosomes for male-biased genes, reflecting their asymmetric occurrence in the two sexes (reviewed in [97]). In contrast, our data reveal no sexualization of the NRR. We detected an enrichment of female-biased gene expression in the PAR across all Y haplotypes, as well as on several autosomes, indicating that this pattern does not conform to expectations of

feminization driven by sex-linked inheritance and evolution. Instead, the PAR, particularly regions adjacent to the NRR boundary, is theoretically predicted to accumulate sexually antagonistic polymorphisms due to reduced local recombination [98]. Confirming the presence of sexually antagonistic polymorphism on the PAR will require functional characterization of candidate genes and fitness assays. The minimal degeneration observed in the NRR (estimated age 3.47–5.70 Mya) is comparable to similarly young sex chromosome systems in fishes, including madaka (*Oryzias latipes*, ~10 Mya), guppy (*Poecilia rebiculata*, ~20 Mya), tongue sole (*Cynoglossus semilaevis*, ~30 Mya), stickleback (*Gasterosteus aculeatus*, ~10 Mya) and fugu (*Takifugu rubripes*, ~5 Mya), where extensive gene loss or strong sex-biased expression is also not consistently observed[41,81,95,99,100].

The absence of Y-linked gene decay, the lack of enrichment of sex-biased genes, and the limited evidence for faster-X effects, together with the presence of a single evolutionary stratum in fully-differentiated haplotypes, point to a recent origin and/or ongoing recombination between the X and Y chromosome in this sex-determining system. This interpretation does not preclude the presence of large inversions or other structural variations along the sex chromosomes, provided that such rearrangements arose recently. Notably, in the Emei moustache toad, which harbors an extensive NRR of ~ 394 Mb, phased chromosome-level X and Y chromosome assemblies were largely colinear and no substantial inversions were detected, and XY sex-reserved females were documented[87]. Taken together, we propose that the limited sex chromosome divergence in *R. temporaria* reflects the combined effects of a recent origin of sex determination, extreme heterochiasmy that rapidly establishes non-recombining regions, and recurrent X–Y recombination via sex reversal, which repeatedly erodes divergence across much of the sex chromosome. Collectively, these processes counteract the long-term accumulation of deleterious mutations typically associated with non-recombining chromosomes[63].

Interestingly, the fully-differentiated two Y haplotypes and semi-differentiated Y haplotype share the same boundary of small NRR (4.64 Mb), identified by pairwise F_{ST} analyses among male pools, and our analyses propose it evolved once approximately 3.47-5.70 Mya (Fig. 8). The 95% CI of the age estimates for the small NRR of Y_{B2}^{0} (1.55-5.39 Myr) and Y_{B1}^{a} (3.62-7.78 Myr) overlap substantially, and pairwise F_{ST} comparisons among male pools reveal near-zero differentiation in this region, contrasted by significantly elevated F_{ST} across the remaining portions of the large NRR (Fig. 6). This raises the question of why distinct Y haplotypes retain nearly identical sequences within the small NRR despite an estimated suppression of recombination for approximately 3.47-5.70 Myr. Ongoing gene conversion provides a parsimonious explanation for this pattern, although additional processes may also contribute, including strong purifying selection acting on functionally constrained loci and rare or cryptic recombination events likewise cannot be entirely excluded. Once a master sex-determining gene evolves, the instant lack of recombination of the sex chromosome would initiate independent and parallel evolution of different Y haplotypes that diverge over time. Each would accumulate deleterious mutations independently, until secondary X-Y recombination via sex reversals generates semi-differentiated Y haplotypes, which begin to differentiate anew (Fig. 8). Under this scenario, one would expect multiple NRR boundaries and greater heterogeneity in X-Y differentiation, rather than the three discrete levels observed here. In contrast, the shared small NRR encompassing *Dmrt1*, which exhibits identical boundaries in both semi- and fully-differentiated Y haplotypes, suggests that a Y-specific inversion may suppress X-Y recombination cross this 4.64 Mb region. Alternatively, this pattern could reflect the presence of a supergene involved in sex determination, which as a whole is under selection to preserve the same boundary. One mechanism that could support this hypothesis is the accumulation of sexually antagonistic alleles at the small NRR boundary, which may reinforce linkage and further suppress recombination. Consistent with this idea, we identified two genes located at the small NRR boundaries (*LOC120943612*, *CDC37L1*) that display persistent

female-biased expression across all comparisons between female and three male types with three distinctive Y differentiation levels. This is related to the Y-specific inversion hypothesis, but it is possible to have a non-supergene inversion or a non-inversion supergene. Given that male recombination in this species is largely restricted to chromosomal ends, the latter scenario appears less likely, although it cannot be ruled out. Resolving these alternative mechanisms will require future analyses using long-read sequencing to directly characterize structural variation in this region.

Our evidence points to *Dmrt1* as a likely candidate master sex-determining gene. *Dmrt1* has testis-specific expression in adult tissues, but extremely low (non-detectable) expression during early embryonic sex-determining stages from our whole-tadpole RNA-seq data. It is possible that *Dmrt1* expression at early development is not detectable by RNA-seq because we used whole tadpole tissue of which the gonads are a very small proportion. *Dmrt1* is shown to be recruited as master sex determining gene across various vertebrate groups[101], including the medaka fish [102], and domestic chicken [103]. The homolog *DmW* gene in the tropical clawed frog *Xenopus laevis* controls ovary determination [104], and is a key upstream sex-determining gene in the red-eared sliding turtle [104–106]. In addition, it has been identified as a candidate master sex-determining gene in 27 species across many vertebrate groups [101]. Multiple known genetic mechanisms by which *Dmrt1* acts as the master sex-determining gene are as the follows: i). Sex-biased/sex-specific expression during early embryonic developmental stage (as in medaka and *Xenopus laevis*) achieved by Y- or W- specific expression; ii). Sex-specific dosage, achieved for example by having two copies in males and one copy in females as in chicken with ZW system; iii). Allelic diversification, in cases when there is no coding sequence difference between the X and Y copy but transposable element insertion in the promoter or intronic region to enhance Y-copy expression or reduce X-copy expression (i.e. in various fish species) (see review in [101]). In the common frog, we detected no fixed differences between the coding sequence of X and Y copies of *Dmrt1*[69]. We therefore hypothesize that allelic diversification through sequence variation in promoter, intron, or intragenic regions underlies allele-specific *Dmrt1* expression prior to gonadal sexual differentiation, thereby initiating the male sex-determining pathway. Testing this hypothesis will require long-read sequencing to resolve regulatory differences between X and Y copies, together with gonad-specific RNA-seq experiments across early embryonic stages and functional validation through targeted gene knockout experiments to confirm *Dmrt1* as the master sex determining gene.

The coexistence of multiple Y haplotypes within a single population is unexpected. Previous work found no detectable fitness differences among the three male types[107], and XX genotypes in lowland Swiss populations already produce an even 1:1 sex ratio despite lacking differentiated Y chromosomes. Under such conditions, the introduction of any Y haplotype is predicted to skew the sex ratio toward males, generating selection against Y chromosomes. In the absence of strong selective advantages, classical expectations predict the eventual loss of Y-linked variation through genetic drift, potentially leading to fixation of a single Y haplotype along a counter-evolved X haplotype that ensures an even sex ratio, or Y chromosome loss. The persistence of three distinct Y haplotypes within this pond, and across the species range, therefore, requires an alternative explanation. The most parsimonious interpretation of the data is the Y haplotype variation is a consequence of a dynamic balance between processes that erode variation and those that regenerate it. Specifically, genetic drift continuously acts to eliminate Y haplotypes, whereas extreme heterochiasmy, mutation accumulation and recurrent X–Y exchange mediated by sex reversal generate new Y-chromosome variants, thereby maintaining neutral genetic diversity over time. Such a balance between stochastic loss and recurrent renewal is a well-established mechanism for the maintenance of neutral genetic variation within populations and provides a natural explanation for the long-term persistence of multiple Y haplotypes. Additional processes may further contribute to this pattern, especially since sex ratio selection would

remove Y haplotypes if the XX genotypes already produce a roughly 50% sex ratio, as measured in Meitreile[65]. Gene flow from neighboring populations, each potentially harboring different Y haplotypes, might be a major source of Y haplotype diversity. In addition, selection on Y-linked traits not captured by current fitness measures cannot be excluded. For instance, different Y haplotypes may confer advantages under distinct environmental conditions, producing temporally fluctuating or balancing selection mediated by inter-annual variation in climate. Alternatively, differences in male–male competition or mating tactics, such as the sneaking behavior reported in a Spanish *R. temporaria* population [108], could generate sexual selection favoring alternative Y haplotypes. Such selective mechanisms remain challenging to test, yet they are plausible given broader geographic patterns. Across Switzerland, two male types with differing levels of Y differentiation frequently co-occur [72,73]. Similarly, a latitudinal cline in Sweden reveals stable mixtures of fully differentiated Y, semi-differentiated Y, and XX males, with frequencies shifting from predominantly full-Y in northern Ammarnäs to mainly semi-Y and XX males in southern Tvedöra [69,74]. Taken together, these patterns suggest that the coexistence of multiple male types, previously considered transient, may instead represent a long-term and recurrent feature of *R. temporaria* populations. Disentangling the relative contributions of neutral processes, metapopulation dynamics and selection to the persistence and coexistence of multiple Y haplotypes will require further investigation.

Conclusion

Our study reveals that sex chromosomes can follow an evolutionary trajectory under extreme heterochiasmy and recurrent sex reversal which is profoundly different than predicted by canonical models of gradual recombination suppression. Within a single natural population of *R. temporaria*, we document the coexistence of three Y haplotypes spanning the full continuum of X–Y divergence, from completely homomorphic to carrying the largest non-recombining region yet described in vertebrates. Remarkably, even the most highly differentiated Y haplotypes show little evidence of degeneration: transposable element accumulation is limited, sex-biased expression is weak, and gene loss or Y-copy decay is undetectable. These patterns, together with shared boundaries of the small non-recombining region encompassing the master sex-determining gene *Dmrt1* and evidence for episodic X–Y recombination via sex-reversed females, indicate that recombination is repeatedly reintroduced and divergence recurrently reset. As a result, extensive recombination suppression does not translate into irreversible Y degeneration but instead gives rise to a dynamic equilibrium between drift-driven loss and mutation- and recombination-driven renewal of Y chromosome variation.

We therefore uncover an alternative pathway of sex-chromosome evolution in which recombination architecture, rather than sexually antagonistic selection or progressive structural rearrangements, governs both the origin and persistence of non-recombining regions. Extreme heterochiasmy can instantaneously generate vast Y-linked non-recombining regions, while recurrent sex reversal periodically restores recombination, maintaining multiple Y haplotypes within a single population and preventing long-term degeneration. This cyclical process provides a mechanistic explanation for the widespread homomorphy of frog sex chromosomes and challenges the canonical view of sex chromosomes as evolutionarily irreversible end points.

More broadly, our results suggest that sex chromosomes are best understood as dynamic genomic states rather than fixed evolutionary outcomes, with their tempo and trajectory shaped by the interplay between meiotic constraints, sex reversal, and episodic recombination. We propose that similar dynamics may operate widely in taxa experiencing strong heterochiasmy and labile sex determination, a hypothesis now directly testable with comparative genomics, long-read phasing, and explicit modelling of sex-reversal–mediated recombination.

Materials & Methods

Field sampling

All sampling was performed over three consecutive years (2014-2016) in Meitreile (46º22'4.9"N, 7º9'53.1"E), from a breeding pond located at 1,798m at lower subalpine area in the Western Swiss Alps. The exact sampling method has described previously 69,75. Briefly, all adult samples were sampled during the short breeding season (late March to late April). Females and males were distinguished by the amplexus status and/or the presence of nuptial pad on the thumbs of the males. Toeclips or buccal cells were sampled from each individual with sterile cotton swabs[42], which were taken back to the lab for DNA extractions. The sampled frogs were immediately released after swabbing in the field[107].

Approximately 10 amplexus mating pairs were caught during 2015 breeding season and brought back to the lab for egg spawning at the University of Lausanne for androgenesis experiment. Each amplexus was put in a separate tank (11 L plastic tank) in a climatic room at condition to best mimic the natural environment they were collected from (16ºC, with 12:12 light: dark cycle and approximately 50% humidity). A small amount of natural pond water was brought back to the lab, together with natural vegetation from the pond, in order to minimise animal stress.

Androgenesis experiment

Experiment protocols were approved by the ethics committee for research on animals at the University of Lausanne. The care and treatment of animals used in this research was conducted in accordance with policies on animal care provided by regulations of the Swiss government. Frogs or tadpoles were killed by immersion in 2 g/L or 0.2 g/L bicarbonate-buffered tricaine methanesulfonate (MS-222, Sigma-Aldrich) respectively.

We performed an androgenesis experiment to generate doubled haploid individuals with Y chromosomes (YY individuals) to facilitate genome sequencing of Y chromosomes. Viable embryos derived solely from the paternal genome were generated using protocol described previously [109]. Briefly, females (wild-caught in amplexus) were injected with human chorionic gonadotropin (Sigma-Aldrich) to ovulate with the protocol described earlier [109,110]. Eggs were collected manually in individual petri dishes to inhibit premature activation. Testes were dissected from freshly killed males (with known Y chromosome haplotype) and were crushed in 1x PBS. The collected eggs were irradiated with UV (50-70,000 mJ) for 3 min and incubated with sperm for 40 min by slowly pouring the sperm suspension. These eggs then were cold shocked in an ice bath for 7.5 min and returned to 16ºC for further development. We then periodically checked embryonic and tadpole development and collected individuals for DNA extraction and genotyping using sex-linked markers to detect YY individuals for further genome sequencing (Supplementary Table S2, S3).

DNA extraction and genotyping

DNA was extracted from adult swabs or toeclips or tails of tadpoles, with an overnight treatment of 10% proteinase K (Qiagen) at 56ºC. A Qiagen DNeasy kit and BioSprint 96 workstation (Qiagen) were used to obtain 100-200 µl DNA elution in buffer AE (Qiagen). For identifying the different haplotypes of Y chromosomes, sex-linked markers at the candidate sex-determining gene *Dmrt1* (*Dmrt1_1, Dmrt1_2, Dmrt1_5 and Dmrt3*) [69], and nine to 13 diagnostic sex-linked

microsatellite loci (*Bfg092, Bfg131, Bfg172, Bfg053, Kank1, Bfg191, Bfg093, RtuB, Bfg266, Btemp5, Bfg021, Bfg072* and *Bfg147*) over the entire length of sex chromosome (Chr01) were used to perform genotyping analysis and protocol was described previously [69,70,72,111] (Supplementary Table S1). Briefly, DNA fragments were amplified with the primers of these sex-linked markers. Primers and PCR protocol are available in these publications respectively. Afterwards, genotyping was performed with four-colour fluorescent capillary electrophoresis using an Applied Bosystems Prism 3100 sequencer (Applied Biosystems), and the generated alleles were scored using GENEMAPPER v4.0.

Haplotypes were characterized based both on the presence of Y-specific *Dmrt* alleles and the rest of sex-linked markers[69]. Three categories of X-Y differentiation were previously detected earlier: i) XY males with Y-specific alleles fixed at all *Dmrt* region markers and 13 sex-linked markers, and different Y haplotypes were described as Y_{B1}^{a}, Y_{B2}^{a} and so on; ii) XY^{0} (semi-differentiated Y) with Y-specific alleles fixed only at all *Dmrt* region markers but not for other five sex-linked markers; iii) XX with no Y-specific alleles fixed at any of the 13 sex-linked markers. Finally, we considered individuals with nine Y-linked markers being homozygous as (likely doubled) haploid (Y)Y individuals (Supplementary Table S2, S3).

Genome-wide sequencing of pooled samples (pool-seq) and YY individuals

Whole-genome sequencing of both Pool-seq samples and YY individuals was performed at the University of Lausanne Genomic Sequencing Centre using an Illumina HiSeq 2500 platform with 101 bp paired-end reads. From the genotyping results, we selected male samples with three varying levels of X-Y differentiation (XY_{B1}^{a}, XY_{B2}^{a}; XY_{B2}^{0}; XX) and females (XX). For the fully differentiated Y haplotypes, we selected the two most abundant ones (XY_{B1}^{a} and XY_{B2}^{a}). Given the number of collected field samples, for each group we selected several individuals to make a pool, and the sequencing coverage was roughly equal to the number of individuals per pool (detailed sample number and read coverage see Supplementary Table S4). After sample selection, we performed DNA extraction for each selected sample separately and checked their precise DNA concentration using Qubit. We then equalized DNA concentration across samples by either diluting or evaporating excessive DNA solution, to avoid possible bias in the DNA representation within each pool. The $2\mu l$ DNA samples with similar concentration were taken individually to each pool group, and total pool DNA was carefully mixed thoroughly by pipetting up and down 20 times. For each pool, we used two technical replicates for DNA library preparation and genome sequencing.

Doubled-haploid YY individuals were selected based on the homozygosity of nine different sex-linked markers along the sex-linked linkage group (Supplementary Table S2). We first conducted genotyping of fathers from all collected amplexus pairs and only selected those having Y haplotypes for the androgenesis experiment. Since only the XY fathers were selected, all sex-linked markers were expected to be heterozygous, thus all homozygous individuals with Y haplotypes were considered as doubled haploid YY individuals (Supplementary Table S3). The successful rate of androgenesis was rather low and many tadpoles were selected for genotyping before the final six pools of YY individuals were detected. Since all genotypes of their fathers were obtained first, we only conducted genotyping of four markers located at the *Dmrt* region. Both YY individuals with fully differentiated Y chromosome (Y_{B1}^{a}, Y_{B1}^{a}) and proto-Y were selected for sequencing (Y_{B1}^{0}, Y_{B1}^{0}, Y_{B2}^{0}, Y_{B2}^{0}) with standard paired-end short reads. For the Y_{B1}^{a} Y_{B1}^{a} YY individual, we further conducted mate-pair sequencing with insertion sizes 3kb, 5kb, and 15kb to facilitate the (essentially) haploid genome assembly.

Mapping and variant calling

For each pool of the five pool-seq datasets, before and after read quality trimming, we assessed paired-end reads quality using FASTQC v0.11.9 [112] with default parameters and retained reads passing all the read quality criteria. We trimmed the raw reads to remove sequencing adaptors and low-quality base pairs using Trimmomatic 0.39 [113] with the following parameters: "ILLUMINACLIP: 2:30:10 SLIDINGWINDOW:4:20 MINLEN:36. The trimmed reads were mapped to the published chromosome-level genome assembly of one XX female individual [77] using bowtie2 v2.4.5 [114] with the following parameters: “--very-sensitive”. We filtered out the unplaced scaffolds (only 1.74% of the genome) from the mapping files. For each pool, we split mapping files per chromosome using SAMtools v1.16.1 [115] allowing parallel computation in downstream analysis. We removed duplicates using the following functions of SAMtools v1.16.1 107: collate, fixmate with -m option, sort, markdup with -r option. Unmapped reads were filtered out and only properly paired mapped reads with a mapping quality above 20 were retained using SAMtools v1.16.1 107 using the following parameters with the SAMtools view function: quality = 20; F = "0x4"; f ="0x2". We used BCFtools v1.15.1 107to call variant using the following functions: *mpileup* with -a “FORMAT/AD, FORMAT/DP” option and *call*. We filtered out SNPs with a minimum basecall quality below 20 and minimum depth of coverage of 5 for each pool, using the *bcftools filter* function.

Putative centromere imputation

To account for variation in sequencing depth, SNP counts were normalized within each pool using the autosomal median. The median SNP value per pool was computed from all autosomes. For all genomic windows, normalized SNP counts were calculated as log□(SNP count + 1) and smoothed using a rolling mean with a window size of 50. For each chromosome, breakpoint detection was performed on the smoothed log□-transformed SNP density. A linear regression model of SNP density against genomic position was fitted, followed by segmented regression using the segmented R package[116]. The model included up to seven breakpoints per chromosome. For each fitted segmentation, estimated breakpoint positions, segment slopes, and the corresponding Davies tests for changes in slope were extracted. All breakpoints were plotted against the smoothed SNP profiles to allow visual inspection. For each chromosome, a subset of breakpoints was manually selected based on clear inflection points corresponding to pronounced shifts in SNP density. These selected breakpoints were interpreted as boundaries of putative centromeric regions, consistent with expected reductions in polymorphism in pericentromeric domains. When only one breakpoint was selected, the second boundary was set to the end of the chromosome to ensure a contiguous interval. For each chromosome, the smallest and largest selected breakpoint positions were taken as the start and end of the putative centromere.

Read coverage ratio and SNP count ratio between male and female pools

Overlapping windows of 1,000,000 bp window length and 500,000 bp interval stride across each of 13 chromosomes from the reference genome assembly were generated using bedtools v2.30.0 [117] and its *makewindows* function. For each pool, we then used the filtered mapping files to compute coverage across overlapping window using SAMtools v1.16.1 [118]and its *bedcov* function. Similarly, we

used variant calling output files to compute SNP counts across overlapping window using the *coverage* function of bedtools v2.30.0 with the "-counts" command line option [119]. We then calculated a normalized coverage value per window by dividing its raw value by the autosomal median value. For computing the autosomal coverage, the putative centromere region was removed. For each window, we computed the log□ ratio of normalized SNP density and depth of coverage for each male pool relative to the female pool. SNP counts were incremented by one, normalized within each pool using the autosomal median derived from autosomes with centromeric regions excluded, and transformed to log□ ratios between male and female values. For each chromosome, rolling means of the log□ SNP ratios were calculated using a fixed-size sliding window of 50. Confidence intervals were obtained exclusively from autosomal windows outside centromeric regions. For each male pool, the 2.5th and 97.5th percentiles of the autosomal rolling-mean log□ ratios were used as empirical 95% confidence limits, following the percentile-based approach recommended by [120]. These confidence intervals were subsequently used to annotate the sex chromosome and autosomes in the final plots. The calculations and plots were performed in R v4.2.1 [121]using custom R scripts. All related scripts and R codes is available in the GitHub upon manuscript acceptance.

Sliding-window F_{ST} scanning analysis

For each pool and for each pairwise comparison between male pools and female pool, we generated pileup files using SAMtools v1.16.1[118] *mpileup* function from the filtered mapping files. Pileup files were converted into sync files using grenedalf [122,123] sync function for which we only retained position matching the following criteria: a minimum base quality of 20 (--pileup-min-base-qual), minimum base count of 2 for a nucleotide to be considered as an allele (--filter-sample-min-count), minimum read depth (--filter-sample-min-read-depth) and maximum read depth (--filter-sample-max-read-depth) per sample of 30 and 120 respectively for the pools with 120 individuals (Y_{B1}^{a}, XX females, XX males). F_{ST} was computed using grenedalf [122,123] applied to the pileup-derived sync files. For each focal–reference pool pair and each chromosome, the grenedalf *fst* module was executed with the following settings: method = *unbiased-nei*, window type = *interval*, window-average policy = *window-length*, window interval width = 10,000 bp, and window interval stride = 5,000 bp. Pool sizes were provided per chromosome, and computations were performed using the reference genome and pool-size tables as specified in the workflow. The resulting output consisted of per-window F_{ST} estimates for each chromosome. In parallel, the same sync files were supplied to the grenedalf *frequency* function to extract reference and alternate allele counts and allele frequencies genome-wide, using the quality and coverage filters defined in the configuration (minimum base quality = 30, minimum count = 2, minimum coverage = 30, maximum coverage = 120). Similar approach and filtering process and criteria were conducted for pairwise comparison among the four male pools.

d_N/d_S and π_N/π_S analysis

To estimate synonymous (d_S) and non-synonymous (d_N) substitution rates between X- and Y-linked coding sequences, we used paired-end sequencing reads from doubled haploid YY individuals representing the Y_{B1}^0, Y_{B2}^0, and Y_{B1}^a haplotypes. Reads were processed following the same quality-control, trimming, and mapping procedures described above. Variant calling was performed with bcftools call using the haploid mode ("--ploidy 1"). For each sample and chromosome, SNP-only BCF files were combined with the reference genome to generate consensus sequences using bcftools consensus. Coding sequences (CDS) were then extracted from each consensus genome with gffread based on the reference GFF annotation. Reference CDS were extracted in parallel using the same procedure. For every chromosome and sample, CDS from the reference and the corresponding sample were split into per-gene FASTA files using a custom Python script (split_cds_per_gene.py). These gene-specific alignments were used as input for pairwise dN/dS estimation.

For each gene, codon-based alignments in PHYLIP format were analysed with the yn00 program from PAML v4.9j [124], executed through Biopython's PAML interface[125] via the script run_paml_pairwise.py. The script extracted the dN, dS, and dN/dS values for each gene, and returned a summary table for every sample and chromosome. Only coding sequences with valid codon length (multiples of three) and without premature stop codons were retained. For each YY haplotype, dN, dS, and dN/dS values were summarized per gene, and results were restricted to the longest isoform. The final dataset comprised one dN, one dS, and one dN/dS estimate per gene and per Y haplotype.

Synonymous and non-synonymous nucleotide diversity (πS and πN) and their ratio (πN/πS) were estimated for each pool using SNPGenie v3.0[126]. Variant calls were first filtered and split by chromosome. For each chromosome and pool, we generated strand-specific input files consisting of (i) VCF files extracted from per-chromosome SNP BCFs using bcftools view, (ii) FASTA sequences extracted from the reference genome, and (iii) GTF annotations restricted to CDS features whose lengths were multiples of three. For the negative strand, corresponding FASTA, VCF, and GTF files were reverse complemented using custom Perl utilities to ensure correct codon orientation (see SNPGenie manual). SNPGenie was then executed separately on the positive and negative strand datasets, using its recommended settings (VCF format 3 and codon-based CDS parsing). For each pool and chromosome, SNPGenie computed per-gene and genome-wide estimates of π_N, π_S, and π_N / π_S. The resulting outputs from both strands were retained for downstream analyses.

Molecular clock and age of sex chromosomes

As mentioned in the above section, we used paired-end sequencing reads from double-haploid YY individuals representing the Y_{B1}^0, Y_{B2}^0, and Y_{B1}^a haplotypes to estimate the synonymous (d_S) substitution rate across the genome. The molecular clock derived from the Tibetan frog *Nanorana parkeri*, to estimate the age of *R. temporaria* sex chromosomes [84]. Since the molecular clock was provided as 0.776×10^{-9} substitutions per site per year ($\pm 1.34 \times 10^{-12}$ SE) 82, we calculated the age as follows: (d_S)/rate = $d_S/(0.776 \times 10^{-9})$. To encounter possible outlier d_S values to bias the age estimate, d_S values of all coding genes at the small NRR or large NRR were used after correcting for within-XX male pool polymorphism.

Defining genomic compartments from F_{ST} variation

We used the F_{ST} estimates produced by grenedalf for all focal–reference pool comparisons to infer large-scale genomic compartments along the sex chromosome. Per-chromosome F_{ST} results in CSV format were imported and parsed to extract genomic coordinates, window boundaries, and F_{ST}

values. For each pool comparison, the midpoint of each window was computed, and windows with missing or undefined F_{ST} values were excluded. To reduce local noise, we applied a rolling mean with a window size of 1,000 windows. For each male and female pool, change-point detection was applied to the smoothed F_{ST} profiles using the PELT and Binary Segmentation (BinSeg) algorithms implemented in the changepoint R package 116. The number of expected change points (Q) was specified per pool based on graphical inspection of the underlying profiles. For each pool, the genomic positions of inferred change points were converted from window midpoints to the nearest window start coordinate in the original data. These change points defined the boundaries of successive genomic compartments along the sex chromosome. For each pool, compartment start and end coordinates were defined by combining the first genomic position on the chromosome, the inferred change-point positions, and the annotated end coordinate of the sex chromosome. Each interval was recorded as a distinct genomic compartment specific to that pool. Autosomal compartments were defined separately. For all pools, autosomal regions were taken directly from the annotated genomic intervals in the reference GFF file. Each autosome was assigned a unique compartment label per pool without further segmentation. All sex-chromosome and autosomal compartments were then combined into a unified BED file, which was exported for downstream analyses.

TE insertion and abundance distribution analysis

A *de-novo* TE library was generated using RepeatModeler2 on the *Rana temporaria* reference genome aRanTem1.1 through the Dfam TE tools 1.88.5 container (https://github.com/Dfam-consortium/TETools) 117. Later, Cd-hit was used to reduce redundancy on the TE library with the following settings: -d 0 -aS 0.8 -c 0.8 -G 0 -g 1 -b 500 118. PoPoolationTE2 was used to detect TE insertion abundance on Poolseq data of one female and four male pools with varying Y chromosome differentiation 119. Briefly, the newly generated TE library and aRanTem1.1 were used as reference TE and reference genome respectively. PoPoolationTE2 requires a TE masked genome that was created through RepeatMasker masking TE sequences of the *de-novo* generated TE library [131,132]. Later, the masked genome and TE library were merged. Reads were reordered using BBMap repair and later mapped against the merged genome using bwa and processed using PoPoolationTE2 in separate mode with default parameters [133].

RNA-seq data generation, processing and major allele ratio analysis

Adult samples from the same Swiss Alpine population were previously collected in the breeding season of 2013 for transcriptome sequencing. In total, 13 adult males and 3 adult females were captured and sacrificed following standard euthanasia methods previously described [74–76]. Brain and gonad tissues were removed and flash frozen in liquid nitrogen, followed by RNA extraction using the Trizol Method. To determine the haplotype of the male Y chromosome, males were genotyped with the 12-16 sex-linked markers we described in the above section on genotyping. Three males were identified as undifferentiated, five as semi- differentiated, and five as fully differentiated Further investigation into their specific haplotype was not undertaken. Paired-end sequencing producing 100 bp reads was performed by the University of Lausanne Genomic Technologies Facility.

Quality was assessed using FASTQC v0.11.9 (www.bioinformatics.babraham.ac.uk/projects/fastqc). Adapters were removed and reads were trimmed in Trimmomatic with the parameters: ILLUMINACLIP:TruSeq2-PE.fa:2:30:10:2:keepBothReads LEADING:20 TRAILING:20 SLIDINGWINDOW:4:20 MINLEN:50

123. Trimmed reads were then aligned against the *Rana temporaria* reference genome using the 2-pass method in STAR [135], in which the first pass collects the splice junction information to be included in the genome index, and the second pass uses this information to refine the alignment. As the chromosomes are *R. temporaria* are particularly large and exceed the capability of GATK to index using the typical .bai format, all indexing steps were performed using the .csi format. We then processed the according to GATK best practices for RNA-seq data [136], including adding read groups, marking duplicates, and splitting reads. We then used Haplotype Caller to call variants and combined (using GATK CombineGVCFs) and genotyped (using GATK GenotypeGVCFs the resulting vcf files. These were then stringently filtered (QD < 2.0, FS > 30.0, SOR > 3.0, MQ < 40.0, MQRankSum < -12.0, ReadPosRankSum < -8.0, DP < 10) to produce an initial, high-confidence variant file. This variant file was then using in a second iteration of aligning reads using STAR in WASP mode [137], to further improve alignments. Again, a two-pass alignment method of STAR was performed with trimmed sequences, followed by the same pipeline of GATK processing (adding read groups, marking duplicates, and splitting reads). Again, variants were called and resulting vcfs were combined and genotyped as above. During the GATK VariantFiltration step, we included the previous high-quality variant file produced in our first round of STAR alignments, and filtered our raw vcf from our second alignment with slightly less stringent parameters (QD < 2.0 || FS > 60.0 || MQ < 30.0 || SOR > 3.0 || ReadPosRankSum < -10.0 || MQRankSum < -10.0 || DP < 8). We extracted biallelic SNPs that passed all previous filters using bcftools [138]. We implemented GATK ASEReadCounter, which produces a .tsv file for each file with the number of reads for the reference and alternative alleles [139].

We next aimed to detect evidence of allele-specific expression. First, we removed SNPs from our filtered variant site if there were 5 in a 100bp window, as this can indicate bias towards the mapping allele during the mapping step [140]. If the two alleles have equal expression, we expect a major allele ratio of 0.5. Deviations from 0.5 indicate a bias towards one allele or the other. As our filtering criteria above likely resulted in the inclusion of mapping or genotyping errors in our dataset, we then selected for polymorphisms with a minimum depth of 30, with at least 5 reads representing the reference and the alternative genotype each. Then or each SNP, we conducted a two-tailed binomial test and corrected or multiple using the Benjamin-Hochberg method. To assess significance for allelic specific expression, we permutated Major Allele Ratio (MAR) values within each genotypic class 10,000 and compared the difference between the median MAR for sex chromosomes and autosomes. P-values were computed as the proportion of replicates which exceeded the median difference in MAR between sex chromosomes and autosomes. Further, to detect evidence of allele-specific expression in differentiated regions of chromosome 1, we computed the major allele ratio in sliding windows of 100 SNPs using the rollmean function in the R package zoo [141].

Differential gene expression and enrichment on sex chromosome

Meitreile RNA-seq datasets of *R. temporaria* including gonadal (testis and ovary), and brain tissues sampled from males with three Y-differentiation levels and females were used in differential expression analysis. Raw sequencing reads were assessed using FastQC v0.11.9 both before and after trimming to evaluate read quality metrics (www.bioinformatics.babraham.ac.uk/projects/fastqc). Adapter removal and quality trimming were performed with Trimmomatic v0.39 in paired-end mode, using the following parameters: ILLUMINACLIP (adapter clipping), HEADCROP:12 (removal of sequencing bias at the start), SLIDINGWINDOW:4:15 (quality filtering), and MINLEN:36 (minimum read length). Both paired and unpaired reads were retained for downstream analysis [134].

Transcript-level quantification was performed using Salmon *v1.10.3* [142]. To prepare a decoy-aware transcriptome index, the *R. temporaria* reference transcriptome and genome (RefSeq assembly

GCF_905171775.1, aRanTem1.1) were concatenated to create a “gentrome” file. Decoy sequences were extracted from the genome FASTA headers, and the Salmon index was built using the --decoys and --validateMappings options to ensure accurate alignment. Salmon quantification was run in quasi-mapping mode using paired-end trimmed reads. For each sample, the following parameters were used: -l A for automatic library type detection, --validateMappings for selective alignment, and --gcBias correction. Transcript-level abundances were quantified as TPM (Transcripts Per Million), which normalizes for both transcript length and sequencing depth. These normalized values were used for gene-level expression comparisons, while raw read counts were used for differential expression analyses.

Gene annotation was manually curated from the *R. temporaria* reference genome assembly (GCF_905171775.1, aRanTem1.1). Custom annotation tables were constructed to support downstream gene-level analyses. Two annotation sets were generated: i) Annotation_CDS: based on coding sequence, incorporating gene name, product, chromosome, coordinates, length, and effective length, merged with Salmon quantification results using the CDS reference. ii) Annotation_Transcriptome: based on mRNA and non-coding transcript identifiers, excluding rRNA, snRNA, snoRNA, and tRNA, and similarly merged with Salmon outputs to extract accurate transcript length and effective length estimates. Salmon RNA-seq quantifications were used to refine effective lengths, accounting for transcript-specific sequence characteristics. Duplicate FASTA entries were filtered out to ensure consistency between quantification and annotation. Chromosome names were cleaned and standardized. The resulting annotation tables were integrated with expression data enabling positional and functional analyses of differentially expressed genes.

Differential expression analysis was performed using the EdgeR *v4.2.2* package in R [143,144]. Genes with low expression were removed through a two-step filtering process: first, genes with average log□-counts per million (CPM) ≤ 0 were excluded; second, genes were retained only if at least two samples exhibited CPM ≥ 1. Read counts were normalized using the trimmed mean of M-values (TMM) method to correct for differences in library size and composition. Gene-wise dispersion estimates were obtained using quasi-likelihood methods (*glmQLFit*), and differential expression was tested using quasi-likelihood F-tests (*glmQLFTest*). Experimental contrasts were specified via a custom contrast matrix, and p-values were adjusted for multiple testing using the Benjamini-Hochberg false discovery rate (FDR) procedure. Exploratory analyses were conducted prior to model fitting. Multidimensional scaling (MDS) plots based on $\log_2$ fold change (log□FC) and biological coefficient of variation (BCV) were used to assess sample clustering and detect potential outliers. Mean-variance plots of log2 CPM values were used to evaluate expression dispersion. Sex-biased gene expression was defined using a combined threshold of FDR < 0.05 and |log□FC| ≥ 1, corresponding to at least a two-fold expression difference.

Differential expression results obtained from EdgeR *v4.2.2* were used to investigate whether sex-biased gene expression was non-randomly distributed across chromosomes[144,145]. For each chromosome and each level (undifferentiated, semi-differentiated, and fully differentiated), whether the distribution of log□FC values differed from the rest of the genome was tested using permutation tests. Comparisons involving sex chromosome (Chr1) were performed against all autosomes (Chr2-13), while autosomes were compared to all other autosomes excluding the focal one. Resulting p-values were adjusted using the Benjamini-Hochberg procedure to control the false discovery rate.

To examine the distribution of sex-biased genes across chromosomes, gene count summaries were generated stratified by chromosome, differential expression status, and direction of expression bias. Using DE gene lists identified by *EdgeR* (FDR < 0.05) across each contrast, we defined sex-biased genes as those with |log□FC| > 1 and classified them as male-biased or female-biased. Genes not meeting the DE criteria were labelled as non-differentially expressed (non-DE). For each chromosome and contrast, we recorded the total number of tested genes, the number of DE genes, and

the number of sex-biased genes. To determine whether sex chromosome was significantly enriched for sex-biased genes, we performed permutation tests. Similarly, we also performed permutation tests to assess compartment enrichment on Chromosome 1 for PAR and NRR regions compared to each other and autosomes. P-values were adjusted for multiple testing using the Benjamini-Hochberg method.

Gene expression analysis in the sex-determining region

The expression profiles of 29 putative candidate genes located within the semi-differentiated region of the proto-Y chromosome were examined to investigate their potential involvement in sex determination. Differential expression results from the EdgeR pipeline were used to extract log□FC values and associated false discovery rates for each gene across the three Y differentiation levels. Genes were classified as sex-biased if they showed a significant expression difference (FDR < 0.05) in either direction (log□FC $\neq$ 0); genes without significant differential expression were labelled as "Not significant," and those with missing log□FC values due to expression below detection thresholds were labelled as "Not expressed." To visualize these patterns, absolute log□FC values were plotted for each gene, ordered by genomic position, with separate panels for each Y haplotype. Additionally, six genes exhibiting strong sex bias (|log□FC| > 4) were selected for detailed visualization of expression levels. Transcripts per million (TPM) values were plotted across gonadal (ovary and testis) and somatic (brain) tissues. This approach enabled the assessment of tissue specificity in expression patterns, supporting the identification of candidate master sex-determining genes.

Dendrogram clusters along the sex chromosome

For all the RNA-seq datasets, filtered VCF files from the Swiss Alpine population Metreille were further processed using VCFtools (v0.1.16) to generate region-specific VCFs corresponding to previously defined pseudoautosomal and nonrecombining regions of chromosome 1 in 10 Mb and 50 Mb windows, respectively, and specifically the region including *Dmrt1*, *Dmrt3* and *Dmrt2*. Trimmed VCF files, genome assembly, and annotation were imported into RStudio (v.4.5.0) using the vcfR package [146]. Genotypes were extracted and recoded into numeric format to represent genotype classes (homozygotes coded as 0 and heterozygotes as 1). Additional filtering removed monomorphic sites and sites with missing data. As RNA seq data can only be genotyped for samples with gene expression, we examined two different datasets: one with all individuals and one limited to males (as females did not express some genes of interest, including *Dmrt1*). Dendrograms were generated by hierarchical clustering of samples using Euclidean distance calculated from genotypic data using the pheatmap package (https://github.com/raivokolde/pheatmap). To examine SNP patterns of homozygosity and heterozygosity, the dplyr package was used to extract sites where females and/or XX males show homozygosity and all other samples showed heterozygosity. These sites were then referred to the annotated Rana temporaria genome to record corresponding gene names and locations.

Author contributions

W-JM, NP, PV designed the study. PV, W-JM collected frog samples, conducted genotyping analysis and performed all androgenesis and related experiments. FC, PV, W-JM, EU, KLD, RF and MAT performed bioinformatics analyses. AB designed and generated RNA-seq dataset. NP and W-JM obtained the fundings for the project. W-JM and PV supervised the work. W-JM drafted the

manuscript with the help of FC, with significant contribution from PV, which was improved and commented by all co-authors. All authors agreed on the final version of the manuscript.

Capture and ethical permits

Capture permits were delivered by the division Biodiversité et Paysage (DGE Vaud), and ethical permit delivered by the Veterinary office of the Canton Vaud (authorization 2287).

Acknowledgments

We thank Karim Ghali, Guillaume Lavanchy, Glib Mazepa, Yvan Vuille, Charlotte Karsegaard, Nathalie Jollien, Maud Baudraz and Kim Schalcher for their help during the Meitreile fieldwork. We are grateful for Roberto Sermier for his significant help in the wet lab. We thank Nicolas Rodrigues for sharing some additional frog samples from this population. We thank James Hagan for consultation with various statistical aspects including helping a more reasonable way of computing the 95% confidence interval. We are grateful for Chengde Chang for insightful suggestions on the pool-seq experimental design, to Glib Mazepa for continuous help with androgenesis experimental design, and to Mark Kirkpatrick for discussions and advice throughout the development of this project. The computations were performed at the Vital-IT (http://www.vital-it.ch) Centre for high-performance computing of the SIB Swiss Institute of Bioinformatics, the VSC (Flemish Supercomputer Centre), funded by the Research Foundation - Flanders (FWO) and the Flemish Government, and Louisiana Optical Network Infrastructure (LONI: http://www.loni.org). This work was supported by the Swiss National Science Foundation (Sinergia grant CRSII3_147625 to NP), and by the European Union (ERC starting grant, FrogWY, 101039501) and a starting grant from Vrije Universiteit Brussel (OZR4049) to W-J M.

Data and code availability

Raw sequencing reads have been deposited in the NCBI Sequence Read Archive under BioProject PRJNA1399088 (BioSample accessions SAMN54473431 – SAMN54473446). All code used for the analyses is available at https://github.com/TheWMaLab/poolseq and has been archived on Zenodo (doi: 10.5281/zenodo.17599170). All datasets and code will be made publicly available upon acceptance of the manuscript.

Supplementary files

Supplementary File 1 contains Figures S1–S27. Supplementary File 2 contains Tables S1–S21. All supplementary materials are provided as separate figure and table files, compiled into two documents.

References

1. Bachtrog, D. *et al.* Sex determination: why so many ways of doing it? *PLoS Biol* 12, e1001899 (2014).
2. Bell, G. *The Masterpiece of Nature: The Evolution and Genetics of Sexuality*. (University of California Press, Berkeley, 1982).

3. Beukeboom, L. W. & Perrin, N. *The Evolution of Sex Determination*. (Oxford University Press, Oxford, 2014). doi:10.1086/682619.

4. Kirkpatrick, M. & Guerrero, R. F. Signatures of sex-antagonistic selection on recombining sex chromosomes. *Genetics* 197, 531–541 (2014).

5. Ohno, S. *Sex Chromosomes and Sex-Linked Genes*. (Springer-Verlag, Berlin Heidelberg, 1966). doi:10.1007/978-3-642-88178-7.

6. Bachtrog, D. Adaptation shapes patterns of evolution on sexual and asexual chromosomes in Drosophila. *Nat Genet* 34, 215–219 (2003).

7. Connallon, T. *et al.* Local adaptation and the evolution of inversions on sex chromosomes and autosomes. *Phil. Trans. R. Soc. B* 373, 20170423 (2018).

8. Bachtrog, D. *et al.* Are all sex chromosomes created equal? *Trends in Genetics* 27, 350–357 (2011).

9. Ming, R., Bendahmane, A. & Renner, S. S. Sex chromosomes in land plants. *Annu Rev Plant Biol* 62, 485–514 (2011).

10. Wilson Sayres, M. A. Genetic diversity on the sex chromosomes. *Genome Biol Evol* 10, 1064–1078 (2018).

11. Charlesworth, B. The evolution of sex chromosomes. *Science* 251, 1030–1033 (1991).

12. Fisher, R. A. The evolution of dominance. *Biological Reviews and Biological Proceedings of the Cambridge Philosophical Society* 6, 345–368 (1931).

13. Rice, W. R. Evolution of the Y Sex chromosome in animals. *Bioscience* 46, 331–343 (1996).

14. Ohno, S. Evolution of sex chromosomes in mammals. *Annu Rev Genet* 44, 91–112 (2010).

15. Graves, J. A. M. Sex chromosome specialization and degeneration in mammals. *Cell* 124, 901–914 (2006).

16. Larson, E. L., Kopania, E. E. K. & Good, J. M. Spermatogenesis and the Evolution of Mammalian Sex Chromosomes. *Trends in Genetics* 34, 722–732 (2018).

17. Xu, L., Wa Sin, S. Y., Grayson, P., Edwards, S. V & Sackton, T. B. Evolutionary Dynamics of Sex Chromosomes of Paleognathous Birds. *Genome Biol Evol* 11, 2376–2390 (2019).

18. Presgraves, D. C. Sex chromosomes and speciation in Drosophila. *Trends in Genetics* 24, 336–343 (2008).

19. Kaiser, V. B. & Bachtrog, D. Evolution of sex chromosomes in insects. *Annu. Rev. Genet.* 44, 91–112 (2010).

20. Blackmon, H., Ross, L. & Bachtrog, D. Sex determination, sex chromosomes, and karyotype evolution in insects. *Journal of Heredity* 108, 78–93 (2017).

21. Ma, W. J., Carpentier, F., Giraud, T. & Hood, M. E. Differential gene expression between fungal mating types Is associated with sequence degeneration. *Genome Biol Evol* 12, 243–258 (2020).

22. Furman, B. L. S. *et al.* Sex chromosome evolution: so many exceptions to the rules. *Genome Biol Evol* 12, 750–763 (2020).

23. Ironside, J. E. No amicable divorce? Challenging the notion that sexual antagonism drives sex chromosome evolution. *BioEssays* 32, 718–726 (2010).

24. Bazzicalupo, A. L., Carpentier, F., Otto, S. P. & Giraud, T. Little evidence of antagonistic selection in the evolutionary strata of fungal mating-type chromosomes (Microbotryum lychnidis-dioicae). *Genes|Genomes|Genetics* 9, 1987–1998 (2019).

25. Perrin, N. Sex-chromosome evolution in frogs: What role for sex-antagonistic genes? *Philosophical Transactions of the Royal Society B: Biological Sciences* 376, 20200094 (2021).

26. Rice, W. R. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution (N Y)* 41, 911–914 (1987).

27. Rice, W. R. Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome. *Genetics* 116, 161–167 (1987).

28. Charlesworth, D. The timing of genetic degeneration of sex chromosomes. *Phil. Trans. R. Soc. B* 376, 20200093 (2021).

29. Bachtrog, D., Mahajan, S. & Bracewell, R. Massive gene amplification on a recently formed Drosophila Y chromosome. *Nat Ecol Evol* 3, 1587–1597 (2019).

30. Stöck, M. *et al.* Ever-young sex chromosomes in European tree frogs. *PLoS Biol* 9, e1001062 (2011).

31. Jeffries, D. L. *et al.* A rapid rate of sex-chromosome turnover and non-random transitions in true frogs. *Nat Commun* 9, 4088 (2018).

32. Rovatsos, M., Rehák, I., Velenský, P. & Kratochvíl, L. Shared ancient sex chromosomes in varanids, beaded lizards, and alligator lizards. *Mol Biol Evol* 36, 1113–1120 (2019).

33. 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).

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

35. Yazdi, H. P. & Ellegren, H. Old but not (so) degenerated-slow evolution of largely homomorphic sex chromosomes in ratites. *Mol Biol Evol* 31, 1444–1453 (2014).

36. Gamble, T. *et al.* Restriction site-associated DNA sequencing (RAD-seq) reveals an extraordinary number of transitions among gecko sex-determining systems. *Mol Biol Evol* 32, 1296–1309 (2015).

37. Kostmann, A., Kratochvíl, L. & Rovatsos, M. Poorly differentiated XX/XY sex chromosomes are widely shared across skink radiation. *Proceedings of the Royal Society B: Biological Sciences* 288, 20202139 (2021).

38. Charlesworth, D. Evolution of recombination rates between sex chromosomes. *Phil. Trans. R. Soc. B* 372, 20160456 (2017).

39. Rifkin, J. L. *et al.* Recombination landscape dimorphism and sex chromosome evolution in the dioecious plant Rumex hastatulus. *Phil. Trans. R. Soc. B* 377, 20210226 (2022).

40. Veltsos, P., Shinde, S. S. & Ma, W.-J. Sex Chromosome Evolution: The Classical Paradigm and so Much Beyond. in *Reference Module in Life Sciences* (Elsevier, 2024). doi:https://doi.org/10.1016/B978-0-443-15750-9.00043-4.

41. Bergero, R., Gardner, J., Bader, B., Yong, L. & Charlesworth, D. Exaggerated heterochiasmy in a fish with sex-linked male coloration polymorphisms. *Proceedings of National Academy of Sciences* 116, 6924–6931 (2019).

42. Berset-Brändli, L., Jaquiéry, J., Broquet, T., Ulrich, Y. & Perrin, N. Extreme heterochiasmy and nascent sex chromosomes in European tree frogs. *Proceedings of the Royal Society B: Biological Sciences* 275, 1577–1585 (2008).

43. Dufresnes, C., Brelsford, A., Baier, F. & Perrin, N. When sex chromosomes recombine only in the heterogametic sex: heterochiasmy and heterogamety in Hyla tree frogs. *Mol Biol Evol* 38, 192–200 (2021).

44. Sardell, J. M. & Kirkpatrick, M. Sex differences in the recombination landscape. *American Naturalist* 195, 361–379 (2020).

45. Bachtrog, D. Sex chromosome evolution: molecular aspects of Y-chromosome degeneration in Drosophila. *Genome Res* 15, 1393–1401 (2005).

46. Steinemann, M. & Steinemann, S. Enigma of Y chromosome degeneration: Neo-Y and Neo-X chromosomes of Drosophila miranda a model for sex chromosome evolution. *Genetica* 102/103, 409–420 (1998).

47. Ma, W.-J. & Veltsos, P. The diversity and evolution of sex chromosomes in frogs. *Genes (Basel)* 12, 483 (2021).

48. Schmid, M. *et al.* Sex-Determining Mechanisms and Sex Chromosomes in Amphibia. in *Amphibian Cytogenetics and Evolution* (eds. Green, D. M. & Sessions, S. K.) 393–430 (Academic Press, Inc., 1991). doi:10.1016/b978-0-12-297880-7.50020-2.

49. Kitano, J. & Peichel, C. L. Turnover of sex chromosomes and speciation in fishes. *Environ Biol Fishes* 94, 549–558 (2012).

50. Lindholm, A. & Breden, F. Sex chromosomes and sexual selection in poeciliid fishes. *Am Nat* 160, S214–S224 (2002).

51. Phillips, R. B. Evolution of the sex chromosomes in salmonid fishes. *Cytogenet Genome Res* 141, 177–185 (2013).

52. Gammerdinger, W. J., Kocher, T. D., Gammerdinger, W. J. & Kocher, T. D. Unusual Diversity of Sex Chromosomes in African Cichlid Fishes. *Genes (Basel)* 9, 480 (2018).

53. Kottler, V. A. & Schartl, M. The colorful sex chromosomes of teleost fish. *Genes (Basel)* 9, 233 (2018).

54. Charlesworth, D. Young sex chromosomes in plants and animals. *New Phytologist* 224, 1095–1107 (2019).

55. Muyle, A., Shearn, R. & Marais, G. A. B. The evolution of sex chromosomes and Dosage compensation in plants. *Genome Biol Evol* 9, 627–645 (2017).

56. Veltsos, P. *et al.* Early sex-chromosome evolution in the diploid dioecious plant Mercurialis annua. *Genetics* 212, 815–835 (2019).

57. Meisel, R. P. & Connallon, T. The faster-X effect: integrating theory and data Richard. *Trends in Genetics* 29, 537–544 (2013).

58. Vicoso, B. Molecular and evolutionary dynamics of animal sex-chromosome turnover. *Nat Ecol Evol* 3, 1632–1641 (2019).

59. Van Doorn, G. S. & Kirkpatrick, M. Turnover of sex chromosomes induced by sexual conflict. *Nature* 449, 909–912 (2007).

60. Dufresnes, C. *et al.* Sex-chromosome homomorphy in palearctic tree frogs results from both turnovers and X-Y recombination. *Mol Biol Evol* 32, 2328–2337 (2015).

61. Saunders, P. A. Sex Chromosome Turnovers in Evolution. *eLS* John Wiley, Ltd: Chichester (2019).

62. Blaser, O., Neuenschwander, S. & Perrin, N. Sex-chromosome turnovers□: the hot-potato model. *Am Nat* 183, 140–146 (2014).

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

64. Rodrigues, N., Studer, T., Dufresnes, C. & Perrin, N. Sex-chromosome recombination in common frogs brings water to the fountain-of-youth. *Mol Biol Evol* 35, 942–948 (2018).

65. Veltsos, P. *et al.* No evidence that Y-chromosome differentiation affects male fitness in a Swiss population of common frogs. *J Evol Biol* 33, 401–409 (2020).

66. Guerrero, R. F., Kirkpatrick, M. & Perrin, N. Cryptic recombination in the ever-young sex chromosomes of Hylid frogs. *J Evol Biol* 25, 1947–1954 (2012).

67. Perrin, N. Sex-chromosome evolution in frogs: what role for sex-antagonistic genes? *Philosophical Transactions of the Royal Society B* 376, (2021).

68. Toups, M. A., Rodrigues, N., Perrin, N. & Kirkpatrick, M. A reciprocal translocation radically reshapes sex-linked inheritance in the common frog. *Mol Ecol* 28, 1877–1889 (2019).

69. Ma, W.-J., Rodrigues, N., Sermier, R., Brelsford, A. & Perrin, N. Dmrt1 polymorphism covaries with sex-determination patterns in Rana temporaria. *Ecol Evol* 6, 5017–5117 (2016).

70. Rodrigues, N., Betto-Colliard, C., Jourdan-Pineau, H. & Perrin, N. Within-population polymorphism of sex-determination systems in the common frog (Rana temporaria). *J Evol Biol* 26, 1569–1577 (2013).

71. Rodrigues, N., Merilä, J., Patrelle, C. & Perrin, N. Geographic variation in sex-chromosome differentiation in the common frog (Rana temporaria). *Mol Ecol* 23, 3409–3418 (2014).

72. Rodrigues, N. *et al.* Dmrt1 polymorphism and sex-chromosome differentiation in Rana temporaria. *Mol Ecol* 26, 4897–4905 (2017).

73. Phillips, B. C., Rodrigues, N., Rensburg, A. J. van & Perrin, N. Phylogeography, more than elevation, accounts for sex-chromosome differentiation in Swiss populations of the common frog. *Evolution (N Y)* 74, 644–654 (2020).

74. Rodrigues, N. *et al.* Sex-chromosome differentiation and ‘sex races’ in the common frog (Rana temporaria). *Proc R Soc B* 282, 20142726 (2015).

75. Ma, W.-J., Veltsos, P., Sermier, R., Parker, D. J. & Perrin, N. Evolutionary and developmental dynamics of sex-biased gene expression in common frogs with proto-Y chromosomes. *Genome Biol* 19, 156 (2018).

76. Ma, W.-J. *et al.* Tissue specificity and dynamics of sex-biased gene expression in a common frog population with differentiated, yet homomorphic, sex chromosomes. *Genes (Basel)* 9, 294 (2018).

77. Streicher, J. W. The genome sequence of the common frog, Rana temporaria Linnaeus 1758. *Wellcome Open Res* 6, 286 (2021).

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

79. Darolti, I. *et al.* Extreme heterogeneity in sex chromosome differentiation and dosage compensation in livebearers. *Proceedings of National Academy of Sciences* 116, 19031–19036 (2019).

80. Gammerdinger, W. J., Toups, M. A. & Vicoso, B. Disagreement in Fst estimators: a case study from sex chromosomes. *Mol Ecol Resour* https://doi.org/10.1111/1755-0998.13210 (2020) doi:10.1111/1755-0998.13210.

81. Wright, A. E. *et al.* Convergent recombination suppression suggests role of sexual selection in guppy sex chromosome formation. *Nat Commun* 8, 14251 (2017).

82. Augstenová, B. & Ma, W. Decoding Dmrt1: Insights into vertebrate sex determination and gonadal sex differentiation. *J Evol Biol* 38, 811–831 (2025).

83. Montgomery, S. H. & Mank, J. E. Inferring regulatory change from gene expression: the confounding effects of tissue scaling. *Mol Ecol* 25, 5114–5128 (2016).

84. Sun, Y. *et al.* Whole-genome sequence of the Tibetan frog Nanorana parkeri and the comparative evolution of tetrapod genomes. *Proceedings of National Academy of Sciences* 112, E1257–E1262 (2015).

85. Prentout, D. *et al.* Plant genera Cannabis and Humulus share the same pair of well-differentiated sex chromosomes. *New Phytologist* 231, 1599–1611 (2021).

86. Perrin, N. Random sex determination: When developmental noise tips the sex balance. *BioEssays* 38, 1218–1226 (2016).

87. Xie, S. *et al.* Extreme heterochiasmy and high rates of sex-reversed recombination result in large yet homomorphic sex chromosomes in the Emei moustache toad. *Genome Res* 35, 1325–1336 (2025).

88. Evans, B. J. *et al.* New insights into Xenopus sex chromosome genomics from the Marsabit clawed frog X . borealis. *J Evol Biol* 35, 1777–1790 (2022).

89. Furman, B. L. S. *et al.* A frog with three sex chromosomes that co-mingle together in nature: Xenopus tropicalis has a degenerate W and a Y that evolved from a Z chromosome. *PLoS Genet* 16, e1009121 (2020).

90. Brelsford, A., Rodrigues, N. & Perrin, N. High-density linkage maps fail to detect any genetic component to sex determination in a Rana temporaria family. *J Evol Biol* 29, 220–225 (2016).

91. Furman, B. L. S. & Evans, B. J. Divergent evolutionary trajectories of two young, homomorphic, and closely related sex chromosome systems. *Genome Biol Evol* 10, 742–755 (2018).

92. Charlesworth, D. Evolution of recombination rates between sex chromosomes. *Phil. Trans. R. Soc. B* 372, 20160456 (2017).

93. Stöck, M. *et al.* Maintenance of ancestral sex chromosomes in palearctic tree frogs: Direct evidence from hyla orientalis. *Sexual Development* 7, 261–266 (2013).

94. Brelsford, 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 (Edinb)* 116, 177–181 (2016).

95. Chalopin, D., Volff, J., Galiana, D., Anderson, J. L. & Schartl, M. Transposable elements and early evolution of sex chromosomes in fish. *Chromosome Research* 23, 545–560 (2015).

96. Muyle, A., Bachtrog, D., Marais, G. A. B. & Turner, J. M. A. Epigenetics drive the evolution of sex chromosomes in animals and plants. *Philosophical Transactions of the Royal Society B: Biological Sciences* 376, 20200124 (2021).

97. Mank, J. E. The W, X, Y and Z of sex-chromosome dosage compensation. *Trends in Genetics* 25, 226–233 (2009).

98. Otto, S. P. *et al.* About PAR: The distinct evolutionary dynamics of the pseudoautosomal region. *Trends in Genetics* 27, 358–367 (2011).

99. Ross, J. A., Urton, J. R., Boland, J., Shapiro, M. D. & Peichel, C. L. Turnover of sex chromosomes in the stickleback fishes (Gasterosteidae). *PLoS Genet* 5, e1000391 (2009).

100. 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).

101. Augstenová, B. & Ma, W. Decoding Dmrt1: Insights into vertebrate sex determination and gonadal sex differentiation. *J Evol Biol* 1–21 (2025).

102. Matsuda, M. *et al.* DMY is a Y-specific DM-domain gene required for male development in the medaka fish. *Nature* 417, 559–563 (2002).

103. Smith, C. A. *et al.* The avian Z-linked gene DMRT1 is required for male sex determination in the chicken. *Nature* 461, 267–271 (2009).

104. Yoshimoto, S. *et al.* A W-linked DM-domain gene, DM-W, participatees in primary ovary development in Xenopus laevis. *Proceedings of National Academy of Sciences* 105, 2469–2474 (2008).

105. Ge, C. *et al.* The histone demethylase KDM6B regulates temperature-dependent sex determination in a turtle species. *Science (1979)* 360, 645–648 (2018).

106. Ge, C. *et al.* Dmrt1 induces the male pathway in a turtle species with temperature-dependent sex determination. *Development (Cambridge)* 144, 2222–2233 (2017).

107. Veltsos, P. *et al.* Y-chromosome haplotypes of varying differentiation to the X are not associated with male fitness in common frogs. (2019).

108. Vieites, D. R. *et al.* Post-mating clutch piracy in an amphibian. *Nature* 431, 305–308 (2004).

109. Roco, Á. S. *et al.* Coexistence of Y, W, and Z sex chromosomes in *Xenopus tropicalis*. *Proceedings of the National Academy of Sciences* 112, E4752–E4761 (2015).

110. Khokha, M. K. *et al.* Rapid gynogenetic mapping of Xenopus tropicalis mutations to chromosomes. *Developmental Dynamics* 238, 1398–1346 (2009).

111. Rodrigues, N. Geographic variation in sex-chromosome differentiation in the common frog (Rana temporaria). 3409–3418 (2014) doi:10.1111/mec.12829.

112. Andrews, S. *et al.* FastQC. Preprint at (2012).

113. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120 (2014).

114. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357–359 (2012).

115. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* 10, (2021).

116. Mueggo, V. M. R. segmented: an R package to fit regression models with broken-line relationships. *R News* 8, 20–25 (2008).

117. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842 (2010).

118. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079 (2009).

119. Quinlan, A. R. & Hall, I. M. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842 (2010).

120. Kirkpatrick, M. *et al.* Evolution of the canonical sex chromosomes of the guppy and its relatives. *G3: Genes, Genomes, Genetics* 12, jkab435 (2022).

121. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, Vienna, Austria, 2017). doi:10.1038/sj.hdy.6800737.

122. Czech, L., Spence, J. P. & Expósito-Alonso, M. grenedalf: population genetic statistics for the next generation of pool sequencing. *Bioinformatics* 40, (2024).

123. Kofler, R., Pandey, R. V. & Schlötterer, C. PoPoolation2: Identifying differentiation between populations using sequencing of pooled DNA samples (Pool-Seq). *Bioinformatics* 27, 3435–3436 (2011).

124. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24, 1586–1591 (2007).

125. Cock, P. J. A. *et al.* Biopython: Freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* 25, 1422–1423 (2009).

126. Nelson, C. W., Moncla, L. H. & Hughes, A. L. SNPGenie: Estimating evolutionary parameters to detect natural selection using pooled next-generation sequencing data. *Bioinformatics* 31, 3709–3711 (2015).

127. Killick, R. & Eckley, I. A. Changepoint: An R package for changepoint analysis. *J Stat Softw* 58, 1–19 (2014).

128. Flynn, J. M. *et al.* RepeatModeler2 for automated genomic discovery of transposable element families. *Proc Natl Acad Sci U S A* 117, 9451–9457 (2020).

129. Ã, W. L. & Godzik, A. Cd-hit□: a fast program for clustering and comparing large sets of protein or nucleotide sequences. 22, 1658–1659 (2006).

130. Kofler, R., Daniel, G. & Schlo, C. PoPoolationTE2□: Comparative Population Genomics of Transposable Elements Using Pool-Seq Brief communication. *Mol Biol Evol* 33, 2759–2764 (2016).

131. Tarailo-Graovac, M. & Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinformatics* Mar:Chapte, 4.10.1-4.10.14 (2009).

132. Smit, A. F. A., Hubley, R. & Green, P. RepeatMasker Open-4.0. 2013-2015 AFA Smit, R Hubley, P Green. Current Ve, open-4.0.9 (2015).

133. Li, H. Aligning sequence reads , clone sequences and assembly contigs with BWA-MEM. *ArXiv* https://doi.org/10.48550/arXiv.1303.3997 (2013).

134. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120 (2014).

135. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21 (2013).

136. McKenna, A. *et al.* The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20, 1297–1303 (2010).

137. Asiimwe, R. & Alexander, D. STAR+WASP reduces reference bias in the allele-specific mapping of RNA-seq reads. *bioRxiv* 2024.01.21.576391 (2024).

138. Narasimhan, V. *et al.* BCFtools/RoH: A hidden Markov model approach for detecting autozygosity from next-generation sequencing data. *Bioinformatics* 32, 1749–1751 (2016).

139. Castel, S. E., Levy-Moonshine, A., Mohammadi, P., Banks, E. & Lappalainen, T. Tools and best practices for data processing in allelic expression analysis. *Genome Biol* 16, 1–12 (2015).

140. Stevenson, K. R., Coolon, J. D. & Wittkopp, P. J. Sources of bias in measures of allele-specific expression derived from RNA-seq data aligned to a single reference genome. *BMC Genomics* 14, 536 (2013).

141. Zeileis, A. & Grothendieck, G. zoo: S3 infrastructure for regular and irregular time series. *J Stat Softw* 14, 1–27 (2005).

142. Rob Patro, Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference. *Nat Methods* 14, 417–419 (2017).

143. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* 11, R25 (2010).

144. Robinson, M. D., Mccarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Application Note* 26, 139–140 (2010).

145. Chen, Y., McCarthy, D., Ritchie, M., Robinson, M. & Smyth, G. edgeR: differential analysis of sequence read count data user's guide. *R package* Bioconduct, https://www.bioconductor.org/packages/devel/bioc/v (2020).

146. Knaus, B. J. & Grünwald, N. J. vcfr: a package to manipulate and visualize variant call format data in R. *Mol Ecol Resour* 17, 44–53 (2017).