
Article:
Darolti, I., Almeida, P., Wright, A.E. orcid.org/0000-0003-2479-5250 et al. (1 more author) (2022) A comparison of methodological approaches to the study of young sex chromosomes : a case study in Poecilia. Journal of Evolutionary Biology, 35 (12). pp. 1646-1658. ISSN 1010-061X

https://doi.org/10.1111/jeb.14013

Received: 29 November 2021 | Revised: 21 March 2022 | Accepted: 11 April 2022

DOI: 10.1111/jeb.14013

JOURNAL OF Evolutionary Biology eseb

RESEARCH ARTICLE

Comparison of methodological approaches to the study of young sex chromosomes: A case study in *Poecilia*

Iulia Darolti[1] | Pedro Almeida[2] | Alison E. Wright[3] | Judith E. Mank[1,4]

[1]Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, British Columbia, Canada

[2]Department of Genetics, Evolution and Environment, University College London, London, UK

[3]Ecology and Evolutionary Biology, School of Biosciences, University of Sheffield, Sheffield, UK

[4]Centre for Ecology and Conservation, College of Life and Environmental Sciences, University of Exeter, Cornwall, UK

Correspondence
Iulia Darolti and Judith E. Mank, Department of Zoology and Biodiversity Research Centre, University of British Columbia, Vancouver, BC, Canada.
Emails: darolti@zoology.ubc.ca; mank@zoology.ubc.ca

Funding information
Canada 150 Research Chairs; Natural Sciences and Engineering Research Council of Canada; H2020 European Research Council, Grant/Award Number: 680951

Abstract

Studies of sex chromosome systems at early stages of divergence are key to understanding the initial process and underlying causes of recombination suppression. However, identifying signatures of divergence in homomorphic sex chromosomes can be challenging due to high levels of sequence similarity between the X and the Y. Variations in methodological precision and underlying data can make all the difference between detecting subtle divergence patterns or missing them entirely. Recent efforts to test for X-Y sequence differentiation in the guppy have led to contradictory results. Here, we apply different analytical methodologies to the same data set to test for the accuracy of different approaches in identifying patterns of sex chromosome divergence in the guppy. Our comparative analysis reveals that the most substantial source of variation in the results of the different analyses lies in the reference genome used. Analyses using custom-made genome assemblies for the focal population or species successfully recover a signal of divergence across different methodological approaches. By contrast, using the distantly related *Xiphophorus* reference genome results in variable patterns, due to both sequence evolution and structural variations on the sex chromosomes between the guppy and *Xiphophorus*. Changes in mapping and filtering parameters can additionally introduce noise and obscure the signal. Our results illustrate how analytical differences can alter perceived results and we highlight best practices for the study of nascent sex chromosomes.

KEYWORDS
best practices, genomics, homomorphic sex chromosomes, replication, Type 2 error

1 | INTRODUCTION

Substantial recent attention in sex chromosome research has focussed on the earliest stages of X-Y divergence in order to glean the initial processes of recombination suppression (Wright et al., 2016). Studies of nascent sex chromosome divergence will by definition result in subtle patterns of X-Y sequence differentiation as substantial differences have not yet sufficiently accumulated. Given the expected subtlety, methodology and underlying data can be quite important, and small changes may make all the difference between identifying a delicate pattern or missing it entirely.

For example, several recent tests for divergence between the guppy X and Y chromosomes have revealed contradictory results. Full genomic analysis of the *Poecilia reticulata* sex chromosomes was originally presented in Wright et al. (2017) based on comparisons between male and female genomes (Figure 1). Initial linkage

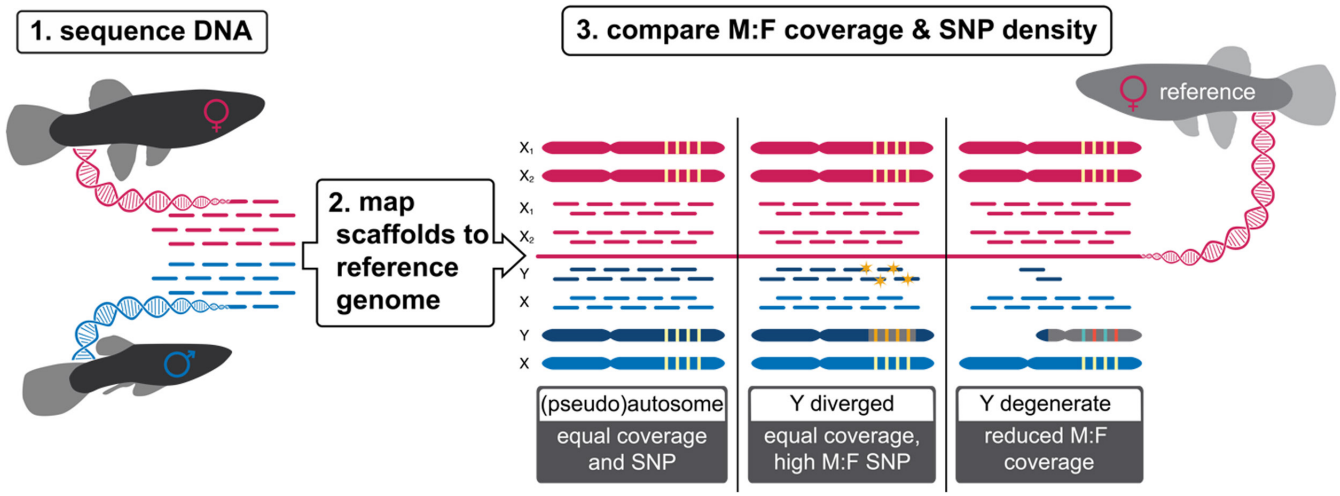

FIGURE 1 Genomic comparisons of male and female DNA data can be used to identify X-Y divergence. Step (1) Multiple males and females are sequenced, and female reads (red) are assembled, with resulting scaffolds ordered and oriented to the nearest available full reference genome. Step (2) Male (blue) and female (red) reads are mapped to this assembly. Step (3) Y divergence leads to male-specific SNPs and, therefore, elevated male:female SNP density. As the Y degenerates, Y reads will no longer map to the X chromosome assembly, leading to reduced male:female coverage. Method adapted from Vicoso and Bachtrog (2013) and Vicoso and Bachtrog (2015). Figure courtesy of Jacelyn Shu (jacelyndesigns.com)

mapping had identified the distal end of guppy chromosome 12 as containing the sex-determining locus (Tripathi et al., 2009). Based on this, genomic comparisons can be used to identify what, if any, regions of the Y chromosome are diverged from the X and to compare across populations to determine intra-specific variation. Wright et al. (2017) found a relatively small region (10 Mb) of significant Y degeneration, designated Stratum I, corresponding to the previous linkage map (Tripathi et al., 2009). This region was characterized by a reduction in the number of male reads that mapped compared with females, consistent with the concept of extensive Y divergence or degeneration. Moreover, the same pattern was observed in all six of the Trinidadian natural populations assayed as well as a captive lab population originally collected from the island, and the rules of parsimony, therefore, suggest that Stratum I is ancestral to the colonization of Trinidad from mainland South America. Wright et al. (2017) also observed evidence of a second region of nascent divergence, Stratum II, which appeared to have formed independently in three upstream populations, but was smaller in downstream populations. This region was characterized by an increase in male single-nucleotide polymorphism (SNP) density compared with females but did not show differences between male and female read depth. This pattern is consistent with either a great reduction or complete loss of male recombination recently in this region or selection against recombinant males.

Almeida et al. (2021) built on these initial findings with a greatly expanded data set, again recovering concordant patterns of Stratum I across the same six natural populations of *P. reticulata*. The expanded data set incorporated 10X Genomics linked reads, allowing for more precision and haplotype phasing. Namely by phasing X and Y haplotypes, it was possible to discern that Stratum I is comprised of two smaller separate regions of reduced male:female read depth. These regions are also enriched for male-specific sequences, male-specific SNPs and repetitive elements, and the presence of these features necessitates recombination suppression from the X chromosome (Almeida et al., 2021; Lin et al., 2022). Importantly, there was also evidence of phylogenetic clustering of phased Y sequence in these regions, indicating ancestral recombination suppression. Finally, Almeida et al. (2021) recovered evidence of parallel expansion of Stratum II in upstream populations. This work was replicated on a different underlying data set with comparable approaches by Sigeman et al. (2021).

Expanding phylogenetically, Darolti et al. (2019) uncovered consistent patterns of sex chromosome topology in *P. wingei*. Initial karyotype analysis suggested that the X and Y chromosomes are ancestral to the common guppy (*P. reticulata*) and Endler's Guppy (*P. wingei*) (Nanda et al., 1993). Furthermore, Darolti et al. (2019) found the same small region of reduced male coverage on the Y chromosome consistent with Stratum I, although the degree of X-Y divergence in this region was more pronounced in *P. wingei* than in *P. reticulata*, as evidenced by greater reduction in male read depth. This region matched nearly perfectly with *P. reticulata*, suggesting that recombination suppression in Stratum I occurred in the common ancestor of *P. wingei* and *P. reticulata*. Consistent with this, Morris et al. (2018) found evidence of male-specific sequence shared between *P. reticulata* and *P. wingei*, possible only if recombination between the X and Y was halted in the common ancestor of these species. Additionally, Darolti et al. (2020) used SNP segregation patterns from RNA-seq data across pedigrees to determine X and Y sequence and found four genes that showed phylogenetic evidence of recombination suppression in the ancestor of *P. wingei* and *P. reticulata*. Although the bootstrap values for any one locus were not excessively high, it is telling that all four were in the areas of X-Y divergence in Stratum I observed by Almeida et al. (2021). The ancestral origin of Stratum I was further supported by conserved patterns of

male hypomethylation within this region in both species (Metzger & Mank, 2020), consistent with sexualization of gene regulation. Finally, Darolti et al. (2019) found evidence for another independent origin of Stratum II based on SNP data in *P. wingei*. Work in outgroup species revealed the same chromosome is a sex chromosome in *P. picta* and *P. parae* (Darolti et al., 2019; Sandkam et al., 2021), although with far greater X-Y divergence in both these species. This sex chromosome was not evident in outgroup species *Poecilia latipinna* or *Gambusia holbrooki* (Darolti et al., 2019), and parsimony therefore suggests the sex chromosome originated ~20 mya in the most recent common ancestor of *P. reticulata*, *P. wingei*, *P. picta* and *P. parae*.

Crucially, all of these analyses were based on custom genome or transcriptome assemblies generated bespoke from the underlying data (Almeida et al., 2021; Darolti et al., 2019, 2020; Wright et al., 2017), although they did use existing related reference genomes to physically place and orient scaffolds and lift-over annotations. This is in contrast to other studies, which have used existing resources derived from different populations or species, resulting in potential mismatches between the underlying data and the genome to which it is compared. Taking a bespoke approach is critical as it reduces the phylogenetic distance between the sequence reads and the reference to which they are mapped, which can increase the proportion of reads that are accurately mapped and reduce issues arising from structural variation and repetitive sequence. Secondly, an important step in identifying diverged regions in sex chromosomes is ensuring stringent mapping parameters (Carvalho & Clark, 2013; Palmer et al., 2019; Smeds et al., 2015; Vicoso & Bachtrog, 2013, 2015). This is particularly relevant for homomorphic sex chromosomes as they still retain extensive sequence similarity between the X and Y, and incorrectly mapped reads can mask coverage differences between the sexes and lead to the misclassification of sex-linked sequences as autosomal. Wright et al. (2017), Darolti et al. (2019) and Almeida et al. (2021) used stringent mapping limits, removed minor alleles with low frequency, which likely represent sequencing errors, and focussed on coding sequence to minimize issues with repetitive elements (Table 1). This was based on the reasoning that young sex

TABLE 1 Comparison of methods and findings for *P. reticulata* sex chromosome strata

Study	Evidence for Stratum I from M:F read depth	Phylogenetic clustering of Y sequences in Stratum I	Evidence for Stratum II from M:F F_{ST}	Genome Assembly	Read depth analysis	SNP analysis
Wright et al. (2017)	Yes	n/a	Yes	Bespoke (*P. reticulata*)	Uniquely mapping reads	Limited to coding sequence; Site coverage >10; SNP frequency >0.3 x site coverage
Bergero et al. (2019)	No	n/a	Yes	Künstner et al. (2016) (*P. reticulata*)	Default mismatch parameters, duplicate reads excluded	Quality score >30; Minimum coverage 20; Biallelic SNPs
Darolti et al. (2019)	Yes	n/a	Yes	Bespoke (*P. reticulata* and *P. wingei*)	Uniquely mapping read	Limited to coding sequence; Site coverage >10; SNP frequency >0.3 x site coverage
Charlesworth et al. (2020)	No	n/a	Yes	Künstner et al. (2016) (*P. reticulata*)	Not reported[a]	Not reported[a]
Fraser et al. (2020)	Yes	n/a	No	Bespoke (*P. reticulata*)	Default mismatch parameters	MAF > 0.05
Kirkpatrick et al. (2022)	No	Yes	No	Schartl et al. (2013) (*Xiphophorus maculatus*)	Default read mapping parameters with local argument, duplicate reads excluded	Quality score >20; Minimum read depth 3; Biallelic SNPs
Almeida et al. (2021)	Yes	Yes	Yes	Bespoke (*P. reticulata*, river-specific)	Uniquely mapping reads, duplicate reads excluded	MAF >0.1; excluding extremely high coverage sites
Sigeman et al. (2021)	Yes	n/a	Yes	Bespoke (*P. reticulata*)	Ranges of thresholds from unfiltered to uniquely mapping	Quality score >20; Minimum read depth 3; Biallelic SNPs

[a]Mapping criteria not reported in methods and bioinformatic code not publicly available.

chromosomes would exhibit subtle divergence signatures and stringency would be required to detect it (Palmer et al., 2019; Vicoso & Bachtrog, 2015).

Despite observing remarkable concordance of these patterns across multiple data sets, species and analytical methods, other recent studies have differed substantially in their approach and reported some different results (Table 1). For example, Bergero et al. (2019) did not report evidence for Stratum I in their own *P. reticulata* data, and although they did uncover a pattern that is broadly consistent with Stratum II, it was not statistically different across populations (Charlesworth et al., 2020). Qiu et al. (2022) and Fraser et al. (2020) identified small male-specific regions largely consistent with the regions identified by Almeida et al. (2021). However, because of scaffold orientation differences and population-specific inversions, the male-specific regions in Fraser et al. (2020) are in different physical locations on the X chromosome. Finally, Kirkpatrick et al. (2022), reanalysing data from Darolti et al. (2019), found evidence of Stratum I and Stratum II in *P. wingei*, but not in *P. reticulata*. Notably, although they did find phylogenetic evidence of recombination suppression in the ancestor of these two species in Stratum I, consistent with Darolti et al. (2020) and Almeida et al. (2021), although they did not report it as such.

Importantly, Bergero et al. (2019), Charlesworth et al. (2020) and Kirkpatrick et al. (2022) relied on existing reference genomes for all their analyses and did not use genomic reads to build custom-made assemblies for the target populations or species. Kirkpatrick et al. (2022) mapped reads from *P. wingei* and *P. reticulata* to *Xiphophorus*, which last shared a common ancestor 40 mya (Kumar et al., 2017), and Bergero et al. (2019) and Charlesworth et al. (2020) used the publicly available *P. reticulata* reference genome (Künstner et al., 2016), derived from a separate population. Additionally, Bergero et al. (2019) used less stringent mapping criteria, and Fraser et al. (2020) used default mapping parameters. Together, this produces two sources of potential methodological noise in replication efforts. First, noise can arise from accumulated mutations due to phylogenetic distance between the samples used to generate sequence reads and the genome that they are mapped to. Second, permissive default mapping parameters allow for mismapping and, therefore, potentially result in significant noise in genomic comparisons between males and females. In addition to this, many of these studies used different underlying data sets that varied in sample origin, number and read depth, and so it is difficult to distinguish the role of sample variation from methodological differences in these discrepancies.

The proliferation of studies on this system with different levels of analytical precision allows for a revealing comparison of the role of genomic methodology in pattern discovery. Guppies have become a useful genomic model for detecting extreme sex chromosome homomorphy (Sigeman et al., 2021) and so offer a valuable opportunity to test the sensitivity of different approaches at the earliest stages of divergence. We tested various methods on the same underlying data with the goal of determining the methodological reasons for inconsistent findings across these studies and to develop best practices moving forward to the genomic study of nascent sex chromosome systems.

2 | METHODS

2.1 | Data sets

Using the *P. reticulata* data from Almeida et al. (2021) and the *P. wingei* data from Darolti et al. (2019), we ran multiple analyses of guppy sex chromosome evolution, following the various analytical methods used by Wright et al. (2017), Bergero et al. (2019) and Kirkpatrick et al. (2022), as summarized in Table 1 and detailed below. We were unable to include the methodology of Charlesworth et al. (2020) as mapping criteria were not reported in methods and bioinformatic code is not publicly available. The data sets for *P. reticulata* and *P. wingei* included paired-end DNA-seq reads from three males and three females from the Quare upstream population (EBI ENA under BioProject PRJEB39998) and from our lab population (NCBI SRA under BioProject PRJNA528814), respectively. Although the Almeida et al. (2021) data included barcodes for linked reads, we disregarded these and assessed it only as paired-end reads. We assessed read quality using FastQC v0.11.9 (www.bioinformatics.babraham.ac.uk/projects/fastqc/, last accessed 8 November 2021), trimmed using Trimmomatic v0.36 (Bolger et al., 2014) and concatenated reads as in Darolti et al. (2019) and Almeida et al. (2021). To replicate previous studies, analyses were repeated using several different genomes and their respective gene annotations, which included the *P. reticulata* Quare genome assembly from Almeida et al. (2021), the *P. reticulata* reference genome from Künstner et al. (2016) (NCBI accession GCF_000633615.1), the *P. wingei* genome assembly from Darolti et al. (2019) and the *Xiphophorus maculatus* reference genome from Schartl et al. (2013) (NCBI accession GCF_002775205.2, v5.0).

2.2 | Coverage analysis

For each focal species, we used three separate methodological pipelines to map and filter reads and to estimate read depth. The first method followed the analysis in Wright et al. (2017), which used bwa v0.7.15 aln/sampe (Li & Durbin, 2009) to map reads, removed reads that were not uniquely mapping and estimated coverage with soap.coverage v2.7.7 (http://soap.genomics.org.cn, last accessed 1 April 2019). The second method followed the pipeline in Kirkpatrick et al. (2022), which mapped reads using bowtie2 v2.2.9 with default parameters and the -local argument (Langmead & Salzberg, 2012), removed PCR duplicates using Picard v2.0.1 (http://broadinstitute.github.io/picard, last accessed 8 November 2021) and calculated coverage with BEDtools v2.26 (Quinlan & Hall, 2010). Lastly, the third method followed the analysis in Bergero et al. (2019), which mapped reads with bwa mem and the -M argument (Li & Durbin,

2009), removed PCR duplicates with BEDtools (Quinlan & Hall, 2010) and estimated coverage using SAMtools v1.3.1 (Li et al., 2009).

For all three methodological pipelines, average coverage values were calculated separately for males and females, and average male:female coverage for each non-overlapping window was calculated as $\log_2$(average male coverage) – $\log_2$(average female coverage). A window size of 50 kb was used for all *P. reticulata* analyses and *P. wingei* analyses based on the *X. maculatus* genome, whereas 10 kb windows were used for *P. wingei* analyses using the more fragmented *P. wingei* genome. Moving averages of coverage were plotted in R v4.0.5 (R Core Team, 2019) based on sliding window analyses using the *roll_mean* function. Ninety-five per cent confidence intervals for the moving average plots were obtained by randomly sampling autosomal values 1000 times without replacement.

2.3 | SNP density analysis

To further assess patterns of Y divergence, for both *P. reticulata* and *P. wingei*, we compared three methodological approaches of estimating SNP density differences between males and females. First, based on Wright et al. (2017), we mapped reads to each genome using bowtie2 with default parameters (Langmead & Salzberg, 2012). After file sorting, we used bow2pro v0.1 (http://guanine.evolbio.mpg.de, last accessed 8 November 2021) to generate a profile for each sample, representing counts for each of the four nucleotide bases at each site. We then applied a minimum site coverage threshold of 10 and kept SNPs with a frequency of 0.3 times the site coverage. We further used gene annotation information to remove SNPs from the analysis if they were not located within coding sequences. For each sample, we calculated average SNP density for each gene as the sum of all SNPs divided by the sum of filtered sites in that gene, excluding those with zero filtered sites.

Second, following Kirkpatrick et al. (2022), we called variants from files previously filtered for PCR duplicates (see *Coverage analysis* section above) using BCFtools v.1.3.1 (Li, 2011). We then filtered variants using VCFtools v0.1.12b (Danecek et al., 2011), removing indels, variants with a quality score lower than 20, singletons and private variants. We also selected for biallelic SNPs and a minimum read depth of 3. For each sample, we then used BEDtools counts to count the number of SNPs within 50 kb windows across the genome.

Third, we used the pipeline in Bergero et al. (2019) to call SNPs from the PCR duplicates filtered files (see *Coverage analysis* section above) using GATK HaplotypeCaller v4.1.9 (Poplin et al., 2017) with the parameters --emit-ref-confidence GVCF and -stand-call-conf 30. Further genotyping was done with GATK GenotypeGVCFs with default parameters and SelectVariants to keep SNPs with a minimum coverage of 20, minimum quality of 30 and selecting for biallelic SNPs only. For each sample, we then used BEDtools counts to count the number of SNPs within 50 kb windows across the genome.

Lastly, in each of these three methodological approaches, average SNP density across all males and across all females was calculated separately. For each gene or window, we calculated male:female SNP density as $\log_2$(average male SNP density) – $\log_2$(average female SNP density). We then divided male:female SNP density estimates into autosomal and sex-linked based on chromosomal position. The distributions of male:female SNP density for the autosomes and the sex chromosomes were plotted in R (R Core Team, 2019), and differences between them were tested using Wilcoxon rank-sum tests.

2.4 | Pairwise synteny analyses

We used LAST v1256 (Kiełbasa et al., 2011) to perform pairwise synteny analyses between the *P. reticulata* sex chromosome (chromosome 12) from the reference genome (Künstner et al., 2016), the *P. reticulata* sex chromosome from the Quare assembly (Almeida et al., 2021) and the *X. maculatus* syntenic chromosome 8. For alignments involving the *X. maculatus* sequence, we used LAST with the HOXD70 seeding scheme designed for a higher rate of substitution, whereas for alignments involving *P. reticulata* sequences only we used the uNEAR seeding scheme for aligning sequences with a lower rate of substitutions.

3 | RESULTS

Using the same data set across different genomes and methods, we first assessed the role of various genomic analysis parameters (Table 1) in detecting Stratum I on the *P. reticulata* and *P. wingei* sex chromosomes, previously reported in Wright et al. (2017), Darolti et al. (2019) and Almeida et al. (2021), summarized in Figures 2 and 3. Analyses of *P. reticulata* sequencing data that used the bespoke Quare *P. reticulata* female genome assembly show a significantly lower male to female coverage, indicative of X-Y differentiation, at the distal end of the chromosome, in the previously estimated location of Stratum I (Figure2a–c). This pattern is evident from all three methodological approaches that followed the pipeline from Wright et al. (2017) (Figure 2a), Kirkpatrick et al. (2022) (Figure 2b) and Bergero et al. (2019) (Figure 2c). All three analyses that relied on the *X. maculatus* reference genome also show a region with decreased male coverage relative to that in females; however, this region is shifted closer to the end of the chromosome and only partially overlaps with the syntenic region of the estimated location of *P. reticulata* Stratum I (Figure 2d–f). Pairwise alignments revealed several structural rearrangements between the *P. reticulata* sex chromosome (chromosome 12) and the syntenic *X. maculatus* chromosome 8, particularly in the region of the predicted guppy Stratum I (Figure 4), which may explain the shifted position of the region with low male coverage in analyses that use the *X. maculatus* genome. In addition, different methodological parameters can have a significant impact on the proportion of reads mapped. Mapping efficiency is substantially reduced when using the *X. maculatus* reference (Table 2), which decreases power to detect a signal of X-Y differentiation.

We find no clear pattern of Stratum I when mapping reads to the *P. reticulata* reference genome based on the methodology in Bergero

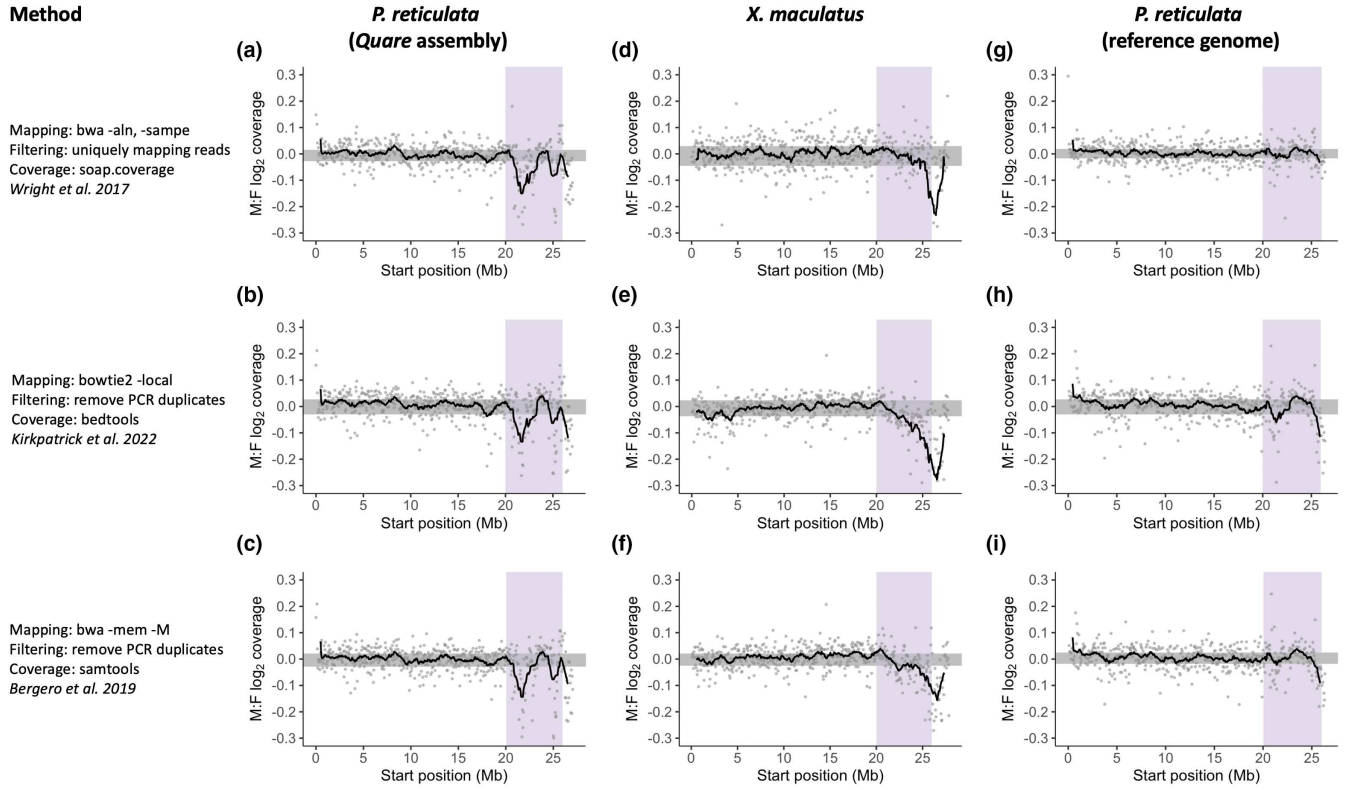

FIGURE 2 Signal for *P. reticulata* Stratum I using comparative methodological approaches. *P. reticulata* DNA-seq reads were mapped in turn to the *P. reticulata* Quare genome assembly (Almeida et al., 2021), the *X. maculatus* reference genome assembly (Schartl et al., 2013) and the *P. reticulata* reference genome assembly (Künstner et al., 2016). Moving average plots represent male to female coverage differences across the guppy sex chromosome (*P. reticulata* chromosome 12 and syntenic *X. maculatus* chromosome 8) in non-overlapping windows of 50 kb. 95% confidence intervals, based on bootstrapping autosomal values, are shown in grey, and predicted boundaries for Stratum I from Almeida et al. (2021) are highlighted in purple

et al. (2019) (Figure 2g–i). Our data suggest that the absence of a Stratum I signal is largely due to the *P. reticulata* reference genome. Specifically, previous work has reported several inversions and assembly errors on the sex chromosome of the first draft of the *P. reticulata* reference genome (Bergero et al., 2019; Charlesworth et al., 2020; Darolti et al., 2020; Fraser et al., 2020), which may be obscuring a signal of Stratum I.

The analyses for *P. wingei* also reveal a lower male to female coverage at the distal end of the chromosome; however, this pattern is only observed in analyses that mapped reads to the *P. wingei* assembly (Figure 3a–c). By contrast, the analyses that used the *X. maculatus* genome all show a significantly elevated read depth in males compared with females, similar to the results in Kirkpatrick et al. (2022) (Figure 3d–f). Previous cytogenetic work has shown that the *P. wingei* Y chromosome is the largest chromosome in the genome, having accumulated a large heterochromatin block (Nanda et al., 1993). However, in addition to the expansion of repetitive sequence, duplication events from the rest of the genome could have also contributed to the remarkable size of the *P. wingei* Y chromosome. The complex inversions and segmental duplications that have occurred between *Xiphophorus* and *Poecilia* (Figure 4) coupled with Y-specific duplications of X-linked regions could explain a signal of elevated coverage in males relative to females in this species when mapping to the *Xiphophorus* genome.

Regions of the sex chromosomes where recombination has recently been halted or greatly suppressed still retain a high degree of similarity between X and Y sequences. They are also expected to show an elevated SNP density in males compared with females, as Y-linked reads carrying Y-specific polymorphisms will still align to the homologous X region of the female reference genome (Figure 1; Vicoso et al., 2013). We observed this pattern in *P. wingei* (Darolti et al., 2019) and in replicate upstream populations of *P. reticulata* (Almeida et al., 2021; Wright et al., 2017), and we designated this as Stratum II. It is important to note that in contrast to Stratum I, Stratum II appears to have formed independently several times. Therefore, to further quantify divergence between the sex chromosomes, we investigated SNP density differences between the sexes using several methodological approaches. In *P. reticulata*, we observe a significantly elevated male SNP density on the sex chromosomes in both of the analyses that aligned reads to the Quara *P. reticulata* genome (Wilcoxon rank-sum test $p < 0.001$, Figure 5a,b). By contrast, the SNP density profiles of the autosomes and the sex chromosomes were indistinguishable in all the analyses that used the *P. reticulata* reference genome (Figure 5c) and the *X. maculatus* genome (Figure 5d–f). The *P. wingei* X and Y chromosomes have previously been suggested to be more diverged than those of *P. reticulata*, as shown through more pronounced coverage and SNP density

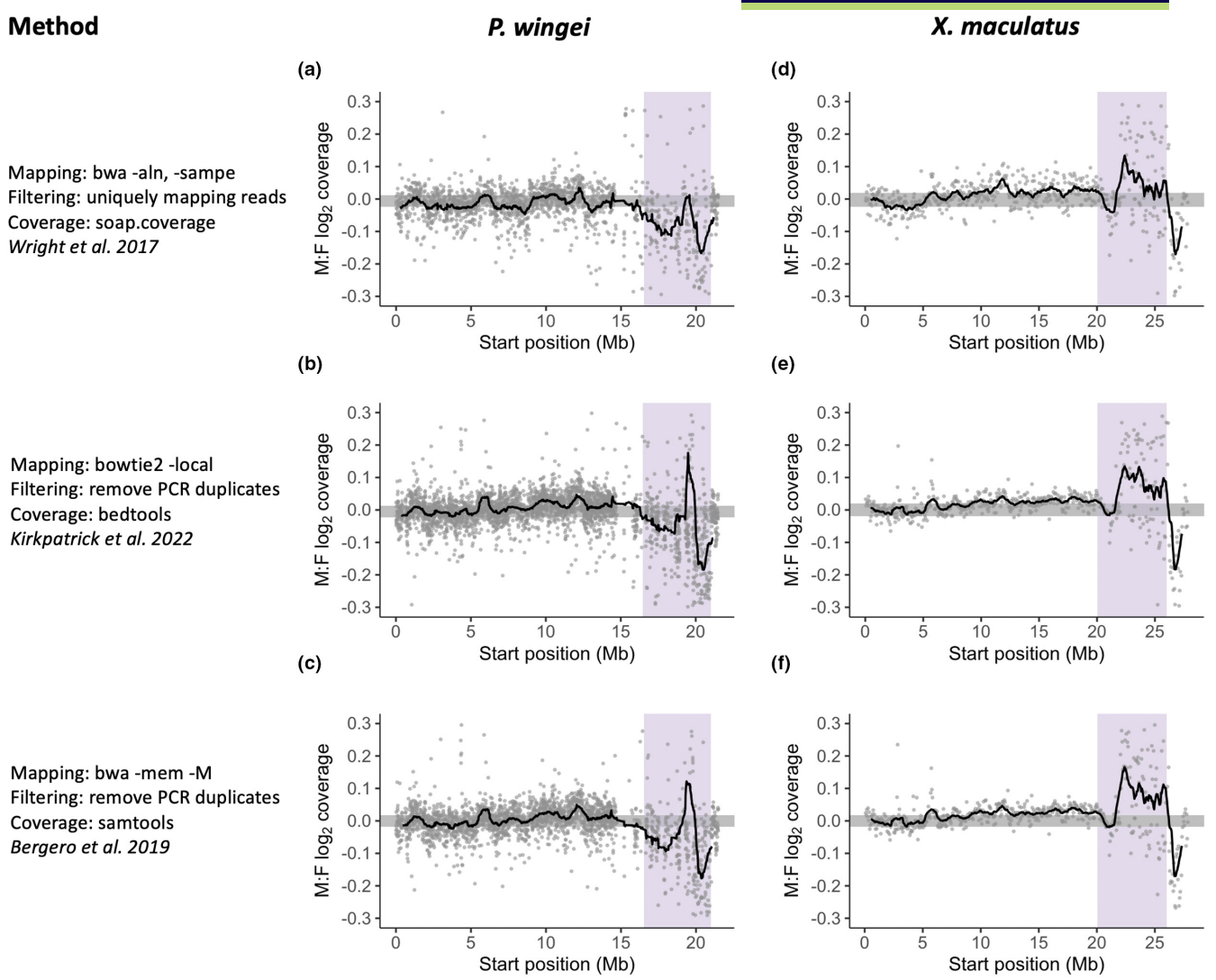

FIGURE 3 Signal for *P. wingei* Stratum I using comparative methodological approaches. *P. wingei* DNA-seq reads were mapped in turn to a *P. wingei* genome assembly (Darolti et al., 2019) and the *X. maculatus* reference genome assembly (Schartl et al., 2013). Moving average plots represent male to female coverage differences across the sex chromosome (*P. wingei* chromosome 12 and syntenic *X. maculatus* chromosome 8) in non-overlapping windows of 50 kb for the analyses that rely on the *X. maculatus* genome and windows of 10 kb for the analyses that use the *P. wingei* genome. The 95% confidence intervals, based on bootstrapping autosomal values, are shown in grey, and predicted boundaries for Stratum I from Darolti et al. (2019) are highlighted in purple

differences between the sexes (Darolti et al., 2019) and a greater accumulation of repetitive sequences on the sex chromosomes in *P. wingei* compared with *P. reticulata* (Almeida et al., 2021; Morris et al., 2018). Our results here confirm this, as we find a significantly higher male:female SNP density for the sex chromosomes compared with the autosomes across all methodological analyses, as well as when using either the *P. wingei* or *X. maculatus* genome (Wilcoxon rank-sum test $p < 0.001$, Figure 6).

4 | DISCUSSION

Replication is fundamental to scientific pursuits, and confirmation is necessary to build a robust understanding of the natural world. The expansion of public data efforts has greatly aided transparency and replication efforts, and this rapid shift in the scientific culture is exemplified by genomics research, where most of the major journals require the deposition of sequencing data as a condition of publication. Failures to replicate results are concerning and necessitate further work to validate or nullify. However, it is important to understand that different replication approaches will have different risks of Type II errors or erroneous negative results. This is especially problematic for the detection of subtle, small effect patterns, such as with initial divergence between X and Y chromosomes. Guppies represent an excellent example of sex chromosomes where differentiation is at the extreme lower end of what is detectable using molecular genomic approaches (Sigeman et al., 2021) and are, therefore, a useful test case for evaluating the efficacy and sensitivity of different methods.

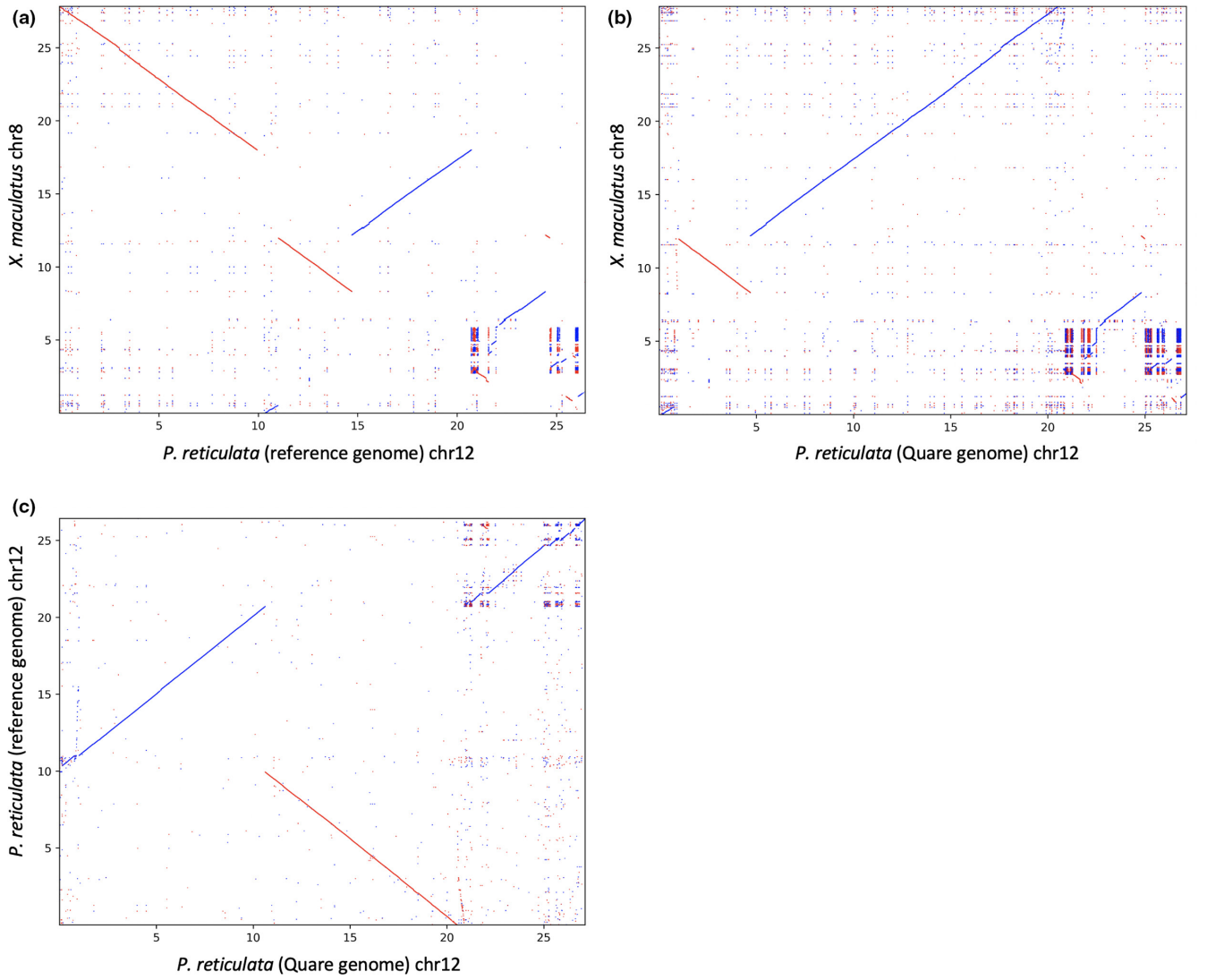

FIGURE 4 Structural rearrangements and duplications between *P. reticulata* Quare, reference and *X. maculatus* genomes. Dot plots of alignments between (a) *X. maculatus* chromosome 8 and *P. reticulata* chromosome 12 (the sex chromosome) from the reference genome assembly, (b) *X. maculatus* chromosome 8 and *P. reticulata* chromosome 12 from the Quare genome assembly and (c) *P. reticulata* chromosome 12 from the reference genome assembly and chromosome 12 from the Quare genome assembly. Forward alignments are shown in blue and reverse alignments in red

TABLE 2 Percentage of concordant, properly paired[a] read alignments

	Sequencing data	*P. reticulata*			*P. wingei*	
Data	Genome assembly	*Quare*	*X. maculatus*	Reference	*P. wingei*	*X. maculatus*
Method	bwa -aln/-sampe (Wright et al., 2017)	59	5	51	75	17
	bowtie2 -local (Kirkpatrick et al., 2022)	83	56	80	87	71
	bwa -mem -M (Bergero et al., 2019)	87	72	83	84	79

[a]Both mates of a read pair map to the same chromosome or scaffold, with the expected insert size and read orientation.

Here, we used the same data set across various methodologies and genome assemblies to test the sensitivity and accuracy of different approaches. Our results show how small changes in the precision of mapping methods can lead to the failure to detect patterns of sex chromosome differentiation in the guppy. The low overall divergence between the X and Y can make detection

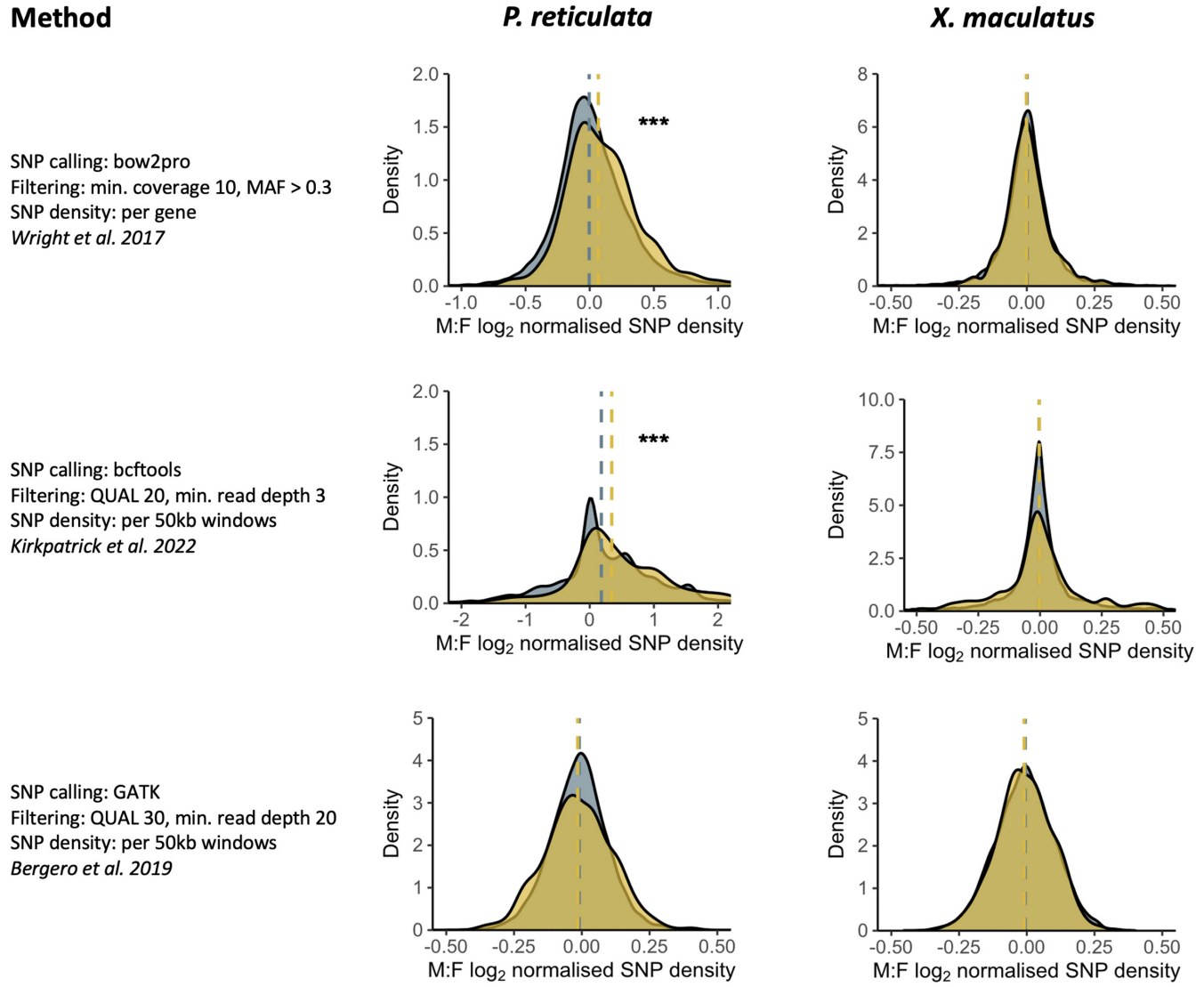

FIGURE 5 Distribution of *P. reticulata* male:female SNP density for the autosomes (grey) and the sex chromosomes (yellow). Dashed vertical lines indicate median SNP densities, and significant differences between the autosomes and the sex chromosomes are shown (****p*-value <0.001). The *P. reticulata* reference genome (Künstner et al., 2016) was used in the analysis based on the method from Bergero et al. (2019), whereas the Quare assembly (Almeida et al., 2021) was used in the other two analyses

difficult, but it has nonetheless been observed across multiple data sets, spanning DNA, RNA and methylation data, as well as multiple methods, including comparisons of male and female coverage and SNP density (Almeida et al., 2021; Darolti et al., 2019; Sigeman et al., 2021; Wright et al., 2017), identification of male-specific sequence (Almeida et al., 2021; Lin et al., 2022; Morris et al., 2018), phylogenetic analysis of recombination suppression (Almeida et al., 2021; Darolti et al., 2020) and comparative epigenomics (Metzger & Mank, 2020) (Figure 7).

By using the same sample data across multiple methods and genomes, our results illustrate how important methodological differences can alter perceived results and highlight the need for replication studies to at minimum replicate the analysis using identical methods to the original or equivalent data set. When possible, replication efforts should go beyond this minimum and expand the analysis by employing more precise methods on existing or expanded data sets. Despite this, some replication efforts use less sensitive approaches, and in these cases, there is a real concern that a perceived failure of replication is instead the result of a lack of precision or statistical power. This is particularly problematic in the field of genomics, as there is little consensus about the gold standard in methodologies, particularly with regard to data processing and filtering procedures. The lack of standardized practices, coupled with the rich nature of genomic data and the complexity of genomes, can make it difficult to discern subtle but important patterns.

Our approach of evaluating the same underlying data with multiple methods and genomes does not account for natural variation across samples and populations, which is substantial (Almeida et al., 2021; Wright et al., 2017). For our *P. reticulata* samples, we chose individuals from an upstream low predation Quare population which we have previously shown to have an intermediate signal of sex chromosome divergence (Almeida et al., 2021; Wright et al., 2017). Samples from

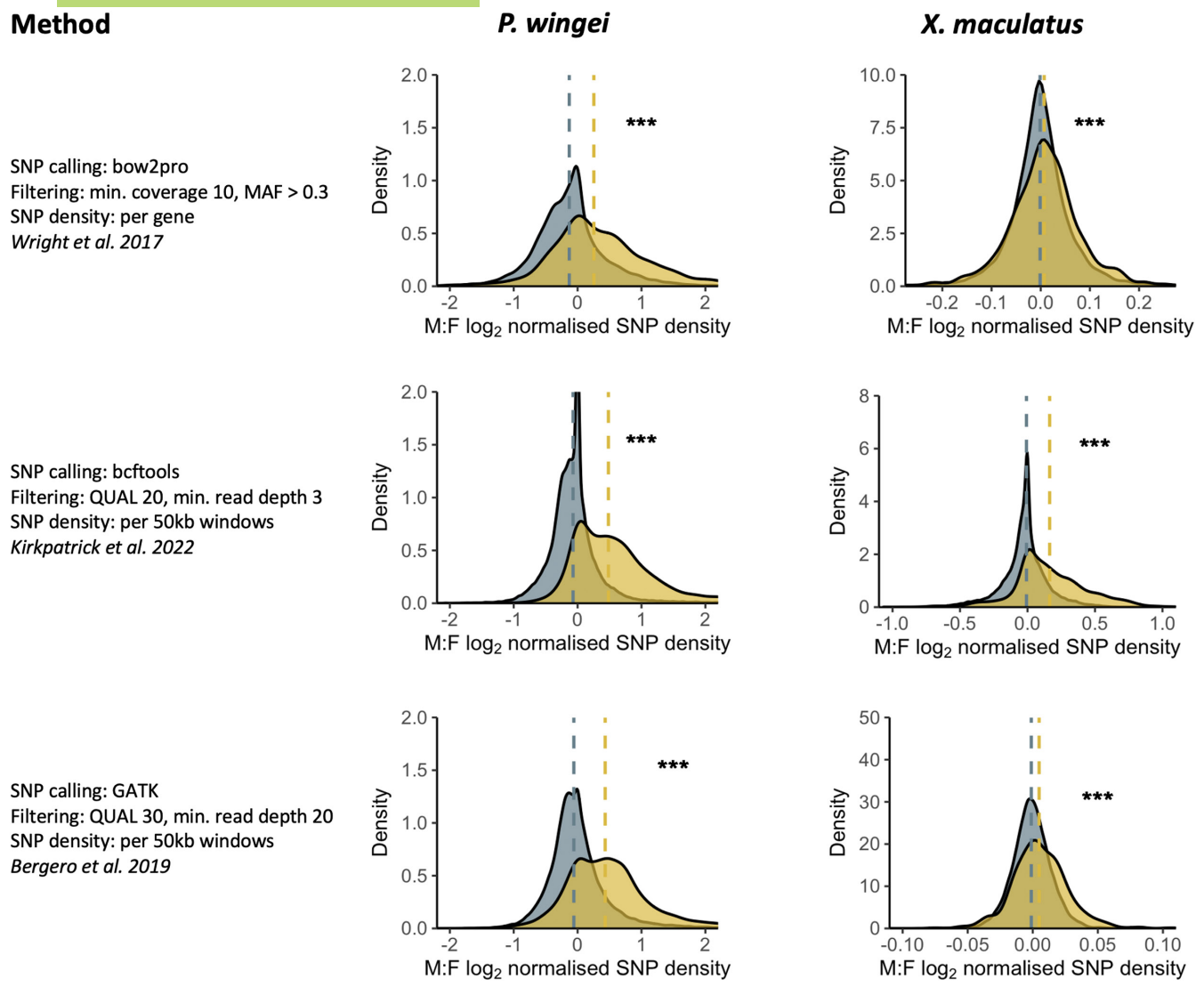

FIGURE 6 Distribution of *P. wingei* male:female SNP density for the autosomes (grey) and the sex chromosomes (yellow). Dashed vertical lines indicate median SNP densities, and significant differences between the autosomes and the sex chromosomes are shown (***p-value <0.001)

populations with greater or lesser signal, or sampling variation due to differences in inversions, duplications and divergence among individuals may also contribute to observed differences, although broadly consistent patterns have been observed across numerous lab and natural populations when accounting for predation (Almeida et al., 2021; Sigeman et al., 2021; Wright et al., 2017).

4.1 | Stratum I

We have previously observed evidence for a small region of ancestral recombination suppression in *P. wingei* and *P. reticulata* (Almeida et al., 2021; Darolti et al., 2019; Wright et al., 2017). This has been replicated in some studies, for example, Fraser et al. (2020) also found evidence of small regions of Y divergence, and Kirkpatrick et al. (2022) recovered phylogenetic clustering of Y sequence in this stratum although they did not report it as such. However, other studies (Bergero et al., 2019; Charlesworth et al., 2020; Kirkpatrick et al., 2022) did not fully replicate these findings.

It is worth noting that Stratum I region of the guppy Y chromosome is enriched for repetitive elements (Almeida et al., 2021), and reads from this region may, depending on the parameters used, map to repetitive elements across the genome, obscuring real read depth differences between males and females if non-coding sequence is included in the analysis. Focussing on uniquely mapping reads when comparing coverage differences between males and females can minimize issues associated with Y repetitive regions. However, our comparative analysis revealed that a pattern of X-Y differentiation can still be recovered without restricting the analysis to uniquely mapping reads (Figures 2 and 3). More stringent SNP filtering parameters can also help eliminate noise in genomic comparisons, and this is particularly important when studying young sex chromosomes as they are expected to exhibit subtle divergence signatures. We were, however, able to identify a signal of elevated male SNP density on

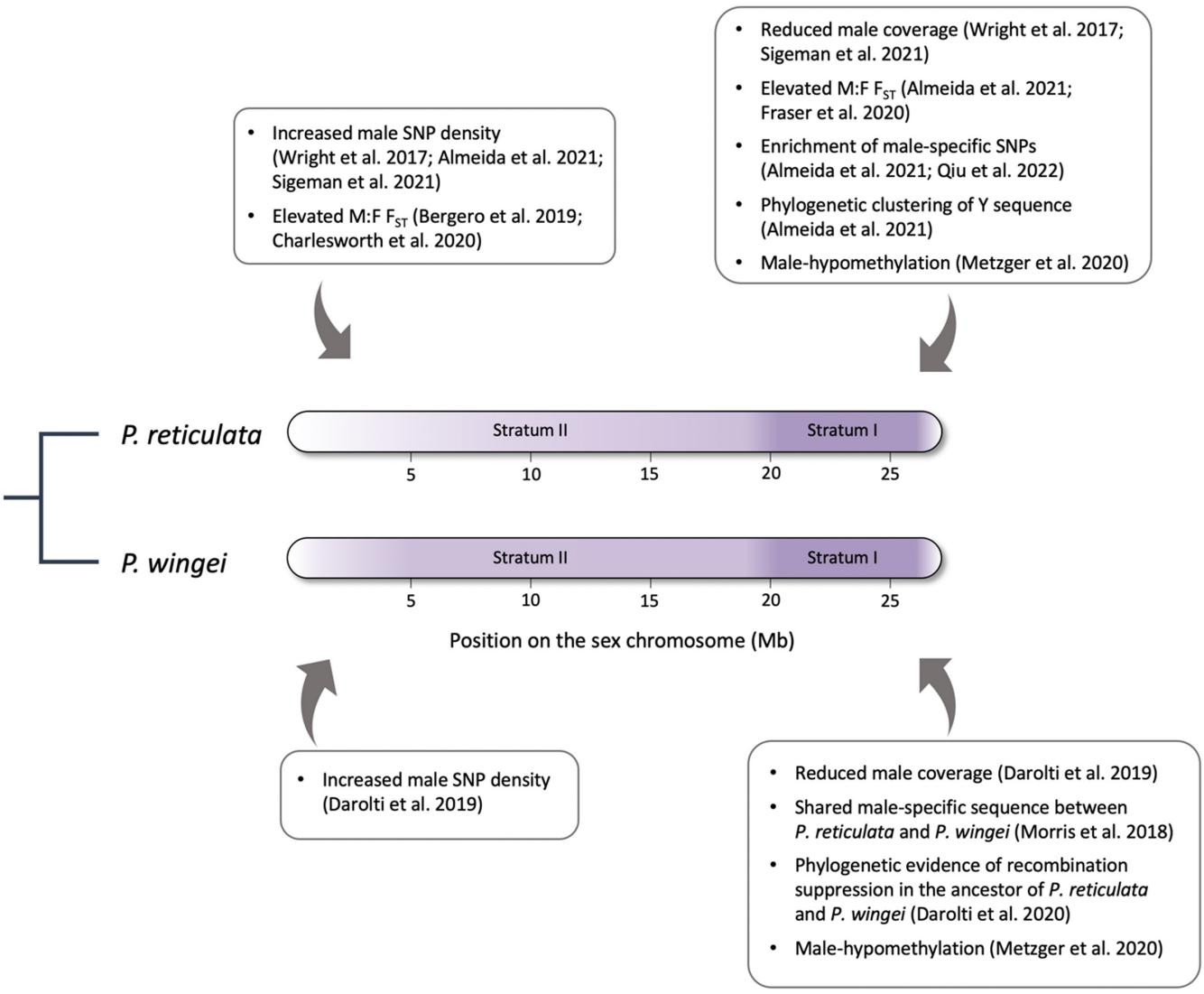

FIGURE 7 Structure of the *P. reticulata* and *P. wingei* sex chromosomes as predicted by multiple methods, including comparisons of male and female coverage, SNP density and phylogenetic and methylation analyses

the sex chromosomes relative to the autosomes, indicative of Y divergence, across several methodological approaches using different degrees of filtering stringency (Figures 5 and 6).

Beyond mapping parameters, by far the most substantial source of variation in the results of the different pipelines we compared lies in the reference genome used. This is in part due to the extensive structural variation across populations and species (Figure 4), but also due to sequence evolution. These two factors combined mean that error compounds over phylogenetic distances, and as the distance between the samples and the genome they are mapped to increases, the ability to detect reduced male:female read depth decreases. This is most evidenced in the strategy by Kirkpatrick et al. (2022), who mapped reads from *Poecilia* species to the *Xiphophorus* genome. They argued that changes over the 40 my phylogenetic distance separating these genera were outweighed by the fact that the *Xiphophorus* genome is more complete. However, the read mapping rate in Table 2 reveals instead that this strategy is less accurate than using less-complete species- or population-specific genome assemblies, as a significantly smaller proportion of *Poecilia* reads map to the *Xiphophorus* genome across all methods, thereby reducing usable data and so increasing noise and decreasing sensitivity. This problem is exacerbated by the substantial structural differences between *Xiphophorus* and *Poecilia* on the sex chromosome (Figure 4), further complicating the comparison. Interestingly, their mapping and filtering methods would have detected Stratum I if they had mapped to a con-specific genome (Figures 2 and 3). To a lesser extent, this is also a problem when mapping data to genomes assembled on different *P. reticulata* populations.

4.2 | Stratum II

As recombination is increasingly suppressed in nascent regions of a sex chromosome, we expect the accumulation of Y-specific SNPs, and we observed this in replicate upstream populations of *P. reticulata* (Almeida

et al., 2021; Wright et al., 2017) and in *P. wingei* (Darolti et al., 2019), consistent with convergent evolution across populations and species (Darolti et al., 2020). Whether this is due to the important environmental effects on recombination rate (Grell, 1971; Plough, 1917; Stevison et al., 2017), sexual conflict (Wright et al., 2017), neutral shifts in male recombination hotspots (Bergero et al., 2019; Wright et al., 2016) or selection against recombinants in the wild remains an important area of further work.

Additionally, given that many mechanisms of recombination suppression only accumulate over time (Furman et al., 2020 and references therein), it also remains unclear how complete recombination suppression is in this region, and whether rare recombination events observed in this region in lab-reared males (Bergero et al., 2019) occur in wild populations. Regardless, it is important to note that suppressed recombination does not necessarily mean that recombination never occurs between the X and Y chromosomes, but rather that it is at least exceedingly rare or recombinant individuals are selected against.

Because of the expected heterogeneity observed in the initial stages of the divergence process (Bergero et al., 2013; Natri et al., 2013; Reichwald et al., 2015), sliding window approaches may be insufficient to reveal overall patterns of elevated male SNP density expected in these regions. Density distributions or direct statistical comparisons between species may be required. This is evidenced by our observation of elevated male:female SNP density across nearly all methods (Figures 5 and 6), with the exception of *P. reticulata* data mapped to the *Xiphophorus* genome, again illustrating the problems with mapping over vast evolutionary distances.

5 | CONCLUDING REMARKS

Here, we have used the same data to compare methods and genomes in the discovery of nascent sex chromosomes. We hope that our results provide a gold standard for future work in other study systems and resolve some of the recent controversy over the sex chromosomes in *Poecilia*.

ACKNOWLEDGEMENTS

We thank Jacelyn Shu of jacelyndesigns.com for drafting Figure 1. We gratefully acknowledge funding from the ERC (grant agreement 680951), a Canada 150 Research Chair and an NSERC grant to JEM. We thank two anonymous reviewers for helpful comments and suggestions on the manuscript and W-J Ma and M Rovatsos for organizing the ESEB Satellite Symposium and Special Issue on sex chromosome evolution.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

All authors designed the research; I.D. and P.A. performed the research; I.D., J.E.M. and A.E.W. analysed the data; all authors wrote the manuscript.

DATA AVAILABILITY STATEMENT

The code and scripts used for data processing, analysis and figure generation are available on GitHub (https://github.com/manklab/Darolti_etal_2022_guppy_sexchromo) and are archived on Zenodo (https://doi.org/10.5281/zenodo.6450120).

ORCID

Judith E. Mank https://orcid.org/0000-0002-2450-513X

REFERENCES

Almeida, P., Sandkam, B. A., Morris, J., Darolti, I., Breden, F., & Mank, J. E. (2021). Divergence and remarkable diversity of the Y chromosome in guppies. *Molecular Biology & Evolution*, *38*, 619–633. https://doi.org/10.1093/molbev/msaa257

Bergero, R., Gardner, J., Bader, B., Yong, L., & Charlesworth, D. (2019). Exaggerated heterochiasmy in a fish with sex-linked male coloration polymorphisms. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 6924–6931. https://doi.org/10.1073/pnas.1818486116

Bergero, R., Qui, S., Forrest, A., Borthwick, H., & Charlesworth, D. (2013). Expansions of the pseudo-autosomal region and ongoing recombination suppression in the *Silene latifolia* sex chromosomes. *Genetics*, *194*, 673–686.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Carvalho, A. B., & Clark, A. G. (2013). Efficient identification of Y chromosome sequences in the human and Drosophila genomes. *Genome Research*, *23*, 1894–1907.

Charlesworth, D., Bergero, R., Graham, C., Gardner, J., & Yong, L. (2020). Locating the sex determining region of linkage group 12 of guppy (*Poecilia reticulata*). *G3 10*, 3639–3649.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R., 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Darolti, I., Wright, A. E., & Mank, J. E. (2020). Guppy Y chromosome integrity maintained by incomplete recombination suppression. *Genome Biology & Evolution*, *12*, 965–977. https://doi.org/10.1093/gbe/evaa099

Darolti, I., Wright, A. E., Sandkam, B. A., Morris, J., Bloch, N. I., Farré, M., Fuller, R. C., Bourne, G. R., Larkin, D. M., Breden, F., & Mank, J. E. (2019). Extreme heterogeneity in sex chromosome differentiation and dosage compensation in livebearers. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 19031–19036. https://doi.org/10.1073/pnas.1905298116

Fraser, B. A., Whiting, J. R., Paris, J. R., Weadick, C. J., Parsons, P. J., Charlesworth, D., Bergero, R., Bemm, F., Hoffmann, M., Kottler, V. A., Liu, C., Dreyer, C., & Weigel, D. (2020). Improved reference genome uncovers novel sex-linked regions in the guppy (*Poecilia reticulata*). *Genome Biology & Evolution*, *12*, 1789–1805. https://doi.org/10.1093/gbe/evaa187

Furman, B. L. S., Metzger, D. C. H., Darolti, I., Wright, A. E., Sandkam, B. A., Almeida, P., Shu, J. J., & Mank, J. E. (2020). Sex chromosome evolution: So many exceptions to the rules. *Genome Biology & Evolution*, *12*, 750–763. https://doi.org/10.1093/gbe/evaa081

Grell, R. F. (1971). Heat induced exchange in fourth chromosome of diploid females of *Drosophila melanogaster*. *Genetics*, *69*, 523–527.

Kiełbasa, S. M., Wan, R., Sato, K., Horton, P., & Frith, M. C. (2011). Adaptive seeds tame genomic sequence comparison. *Genome Research*, *21*, 487–493. https://doi.org/10.1101/gr.113985.110

Kirkpatrick, M., Sardell, J. M., Pinto, B. J., Dixon, G., Peichel, C. L., & Schartl, M. (2022). Evolution of the canonical sex chromosomes of the guppy and relatives. *G3*, *12*, jkab435. https://doi.org/10.1093/g3journal/jkab435

Kumar, S., Stecher, G., Suleski, M., & Hedges, S. B. (2017). TimeTree: A resource for timelines, timetrees, and divergence times. *Molecular Biology & Evolution*, *34*, 1812–1819. https://doi.org/10.1093/molbev/msx116

Künstner, A., Hoffmann, M., Fraser, B. A., Kottler, V. A., Sharma, E., Weigel, D., & Dreyer, C. (2016). The genome of the Trinidadian guppy, *Poecilia reticulata*, and variation in the Guanapo population. *PLoS One*, *11*, e0169087. https://doi.org/10.1371/journal.pone.0169087

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nature Methods*, *9*, 357–359. https://doi.org/10.1038/nmeth.1923

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics*, *27*, 2987–2993. https://doi.org/10.1093/bioinformatics/btr509

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*, 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R., 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*, 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lin, Y., Darolti, I., Furman, B. L. S., Almeida, P., Sandkam, B. A., Breden, F., Wright, A. E., & Mank, J. E. (2022). Gene duplication to the Y chromosome in Trinidadian guppies. *Molecular Ecology*, *31*, 1853–1863. https://doi.org/10.1111/mec.16355

Metzger, D. C. H., & Mank, J. E. (2020). Conserved sex-biased DNA methylation patterns target key developmental genes and non-recombining region of the guppy sex chromosome. Biorxiv, https://doi.org/10.1101/2020.08.21.261792

Morris, J., Darolti, I., Bloch, N. I., Wright, A. E., & Mank, J. E. (2018). Shared and species-specific patterns of nascent Y chromosome evolution in two guppy species. *Genes*, *9*, 238. https://doi.org/10.3390/genes9050238

Nanda, I., Schartl, M., Epplen, J. T., Feichtinger, W., & Schmid, M. (1993). Primitive sex chromosomes in poecilid fishes harbour simple repetitive DNA sequences. *Journal of Experimental Biology*, *265*, 301–308.

Natri, H. M., Shikano, T., & Merila, J. (2013). Progressive recombination suppression and differentiation in recently evolved neo-sex chromosomes. *Molecular Biology & Evolution*, *30*, 1131–1144. https://doi.org/10.1093/molbev/mst035

Palmer, D. H., Rogers, T. F., Dean, R., & Wright, A. E. (2019). How to identify sex chromosomes and their turnover. *Molecular Ecology*, *28*, 4709–4724. https://doi.org/10.1111/mec.15245

Plough, H. H. (1917). The effect of temperature on crossing over in *Drosophila*. *Journal of Experimental Zoology*, *24*, 147–209. https://doi.org/10.1002/jez.1400240202

Poplin, R., Ruano-Rubio, V., DePristo, M. A., Fennell, T. J., Carneiro, M. O., Van der Auwera, G. A., Kling, D. E., Gauthier, L. D., Levy-Moonshine, A., Roazen, D., Shakir, K., Thibault, J., Chandran, S., Whelan, C., Lek, M., Gabriel, S., Daly, M. J., Neale, B., MacArthur, D. G., … Banks, E. (2017). Scaling accurate genetic variant discovery to tens of thousands of samples. *BioRxiv*, 201178. https://doi.org/10.1101/201178

Qiu, S., Yong, L., Wilson, A., Croft, D. P., Graham, C., & Charlesworth, D. (2022). Partial sex linkage and linkage disequilibrium on the guppy sex chromosome. Biorxiv, https://doi.org/10.1101/2022.01.14.476360

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Reichwald, K., Petzold, A., Koch, P., Downie, B. R., Hartmann, N., Pietsch, S., Baumgart, M., Chalopin, D., Felder, M., Bens, M., Sahm, A., Szafranski, K., Taudien, S., Groth, M., Arisi, I., Weise, A., Bhatt, S. S., Sharma, V., Kraus, J. M., … Platzer, M. (2015). Insights into sex chromosome evolution and agine from the genome of a short-lived fish. *Cell*, *163*, 1527–1538.

Sandkam, B. A., Almeida, P., Darolti, I., Furman, B. L. S., van der Bijl, W., Morris, J., Bourne, G. R., Breden, F., & Mank, J. E. (2021). Extreme Y chromosome polymorphism corresponds to five male reproductive morphs in a freshwater fish. *Nature Ecology & Evolution*, *5*, 939–948.

Schartl, M., Walter, R. B., Shen, Y. J., Garcia, T., Catchen, J., Amores, A., Braasch, I., Chalopin, D., Volff, J. N., Lesch, K. P., Bisazza, A., Minx, P., Hillier, L., Wilson, R. K., Fuerstenberg, S., Boore, J., Searle, S., Postlethwait, J. H., & Warren, W. C. (2013). The genome of the platyfish, *Xiphophorus maculatus*, provides insights into evolutionary adaptation and several complex traits. *Nature Genetics*, *45*, 567–572. https://doi.org/10.1038/ng.2604

Sigeman, H., Sinclair, B., & Hansson, B. (2021). FindZX: An automated pipeline for detecting and visualising sex chromosomes using whole-genome sequencing data. Biorxiv, https://doi.org/10.1101/2021.10.18.464774

Smeds, L., Warmuth, V., Bolivar, P., Uebbing, S., Burri, R., Suh, A., Nater, A., Bureš, S., Garamszegi, L. Z., Hogner, S., Moreno, J., Qvarnström, A., Ružić, M., Sæther, S.-A., Sætre, G.-P., Török, J., & Ellegren, H. (2015). Evolutionary analysis of the female-specific avian W chromosome. *Nature Communications*, *6*, 7330. https://doi.org/10.1038/ncomms8330

Stevison, L. S., Sefich, S., Chase, R., & Graze, R. M. (2017). Recombination rate plasticity: Revealing mechanisms by design. *Philosophical Transactions of the Linnean Society, B*, *372*, 20160459. https://doi.org/10.1098/rstb.2016.0459

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

Vicoso, B., & Bachtrog, D. (2013). Reversal of an ancient sex chromosome to an autosome in *Drosophila*. *Nature*, *499*, 332–335. https://doi.org/10.1038/nature12235

Vicoso, B., & Bachtrog, D. (2015). Numerous transitions of sex chromosomes in Diptera. *PLoS Biology*, *13*, e1002078. https://doi.org/10.1371/journal.pbio.1002078

Vicoso, B., Emerson, J. J., Zektser, Y., Mahajan, S., & Bachtrog, D. (2013). Comparative sex chromosome genomics in snakes: Differentiation, evolutionary strata, and lack of global dosage compensation. *PLoS Biology*, *11*, e1001643. https://doi.org/10.1371/journal.pbio.1001643

Wright, A. E., Darolti, I., Bloch, N. I., Oostra, V., Sandkam, B., Buechel, S. D., Kolm, N., Breden, F., Vicoso, B., & Mank, J. E. (2017). Convergent recombination suppression suggests a role of sexual selection in guppy sex chromosome formation. *Nature Communications*, *8*, 14251.

Wright, A. E., Dean, R., Zimmer, F., & Mank, J. E. (2016). How to make a sex chromosome. *Nature Communications*, *7*, 12087. https://doi.org/10.1038/ncomms12087

How to cite this article: Darolti, I., Almeida, P., Wright, A. E., & Mank, J. E. (2022). Comparison of methodological approaches to the study of young sex chromosomes: A case study in *Poecilia*. *Journal of Evolutionary Biology*, *00*, 1–13. https://doi.org/10.1111/jeb.14013