

Research



Cite this article: Dagilis AJ, Sardell JM, Josephson MP, Su Y, Kirkpatrick M, Peichel CL. 2022 Searching for signatures of sexually antagonistic selection on stickleback sex chromosomes. *Phil. Trans. R. Soc. B* **377**: 20210205.
<https://doi.org/10.1098/rstb.2021.0205>

Received: 30 October 2021
Accepted: 12 April 2022

One contribution of 15 to a theme issue 'Genomic architecture of supergenes: causes and evolutionary consequences'.

Subject Areas:
evolution, genomics

Keywords:
sex chromosome evolution, sexually antagonistic selection, Japan Sea stickleback, supergenes

Author for correspondence:
Catherine L. Peichel
e-mail: catherine.peichel@iee.unibe.ch

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5979250>.

Searching for signatures of sexually antagonistic selection on stickleback sex chromosomes

Andrius J. Dagilis^{1,3}, Jason M. Sardell¹, Matthew P. Josephson⁴, Yiheng Su², Mark Kirkpatrick¹ and Catherine L. Peichel⁴

¹Department of Integrative Biology, and ²Department of Computer Science, University of Texas, Austin TX 78712, USA

³Department of Biology, University of North Carolina, Chapel Hill, NC 27599, USA

⁴Division of Evolutionary Ecology, Institute of Ecology and Evolution, University of Bern, 3012 Bern, Switzerland

MK, 0000-0002-0039-4172; CLP, 0000-0002-7731-8944

Intralocus sexually antagonistic selection occurs when an allele is beneficial to one sex but detrimental to the other. This form of selection is thought to be key to the evolution of sex chromosomes but is hard to detect. Here we perform an analysis of phased young sex chromosomes to look for signals of sexually antagonistic selection in the Japan Sea stickleback (*Gasterosteus nipponicus*). Phasing allows us to date the suppression of recombination on the sex chromosome and provides unprecedented resolution to identify sexually antagonistic selection in the recombining region of the chromosome. We identify four windows with elevated divergence between the X and Y in the recombining region, all in or very near genes associated with phenotypes potentially under sexually antagonistic selection in humans. We are unable, however, to rule out the alternative hypothesis that the peaks of divergence result from demographic effects. Thus, although sexually antagonistic selection is a key hypothesis for the formation of supergenes on sex chromosomes, it remains challenging to detect.

This article is part of the theme issue 'Genomic architecture of supergenes: causes and evolutionary consequences'.

1. Introduction

In many taxa, non-recombining regions of sex chromosomes, termed the sex-determining region (SDR), act as supergenes [1]. Supergenes are linked loci that are inherited as a single locus and generate alternative phenotypes [2], properties common to the SDR. The mammalian Y, for example, has a large non-recombining SDR and carries multiple genes necessary for male development and fertility [3]. The boundaries of the SDR are not static, and distinct evolutionary 'strata' can form as recombination is sequentially suppressed across the chromosome [1,4,5]. It is unclear, however, why the SDR expands in some taxa but not others.

The leading hypothesis for the origin and later expansion of the SDR is sexually antagonistic selection (SAS) [6]. SAS occurs when an allele is beneficial in one sex but deleterious in the other [7]. Males and females frequently have different fitness optima, which results in SAS that drives the evolution of sexual dimorphism [8] and sex-biased gene expression [9]. While the SDR carries alleles that are restricted to the X or the Y (or Z and W in ZW systems), polymorphisms under SAS can accumulate in the recombining pseudoautosomal region (PAR) of sex chromosomes [10–12]. These can then drive the suppression of recombination between X and Y chromosomes by linking male beneficial alleles to the sex-determining locus on the Y chromosome, which prevents the deleterious fitness effects that occur when those alleles are carried by females [1,13].

The observation of genes with sex-specific fitness effects in regions of reduced recombination in or near the SDR is consistent with that idea, but there is a strong alternative hypothesis: polymorphisms subject to SAS can accumulate after recombination is suppressed [14,15]. This possibility is supported by examples of male-beneficial genes that have been transposed from autosomes to the SDR of the sex chromosomes of *Drosophila* [16–21], mammals [22,23], and the threespine stickleback [24]. The second alternative hypothesis is even simpler—some sex chromosomes may simply never accrue loci under SAS in a fashion disproportionate to the rest of the genome [25]. Indeed, recent work has demonstrated that population genetic approaches may be biased to identifying SAS on the X [26], and there have been several calls to search for alternative explanations for the evolution of sex chromosomes in general [27,28]. Finding evidence of SAS in PARs is therefore a key goal of sex chromosome research.

Models show that polymorphisms under SAS in the PAR will generate characteristic patterns of molecular variation that can be more conspicuous than those on autosomes or in the SDR [29]. Because sites under SAS on autosomes are unlinked to the SDR, any build up in linkage disequilibrium (LD) between loci under SAS and the sex chromosomes is erased in each generation. SAS can still be detected as minor allele frequency differences between males and females (e.g. [30,31]), but these signals are quite weak. On the other hand, the PAR is linked to the SDR and LD can build up over multiple generations, leading to a stronger signal of SAS. In particular, the X and Y are expected to show peaks of differentiation around sites under SAS, allowing them to be mapped with higher resolution than is possible elsewhere in the genome. Allele frequency differences between the sexes in the PARs of the white campion [32,33] and hops [34] are consistent with SAS, but it has not been possible to rule out alternative hypotheses. Guppies show broad peaks of divergence between males and females in regions of the PAR with very low recombination [35], but these span large numbers of genes.

The sex chromosomes of the Japan Sea stickleback (*Gasterosteus nipponicus*) are highly suited to studies of sexually antagonistic selection. The Japan Sea stickleback diverged from its sister species, the threespine stickleback (*Gasterosteus aculeatus*), between 0.6 and 1.2 Ma [36]. Sex is determined by chromosome 19 (Chr 19) in both species [37–39]. The threespine stickleback has a PAR of 2.5 Mb and a non-recombining SDR of 17.5 Mb that contains three evolutionary strata that correspond to three inversions [24,38]. In the Japan Sea stickleback, the ancestral Y on Chr 19 fused with an autosome, chromosome 9 (Chr 9), to produce a ‘neo-Y’ chromosome [39,40] (electronic supplementary material, figure S1a). The unfused homologue of the neo-Y co-segregates with the ancestral X, and so is referred to as a ‘neo-X’. A large (7 Mb) region of the neo-Y adjacent to the fusion has ceased to recombine, resulting in an expanded SDR [39–41]. The remaining 13.7 Mb of Chr 9 continues to recombine as a second PAR. A quantitative trait locus (QTL) study of Japan Sea stickleback mapped two chromosome regions affecting phenotypes involved with courtship (and so are potential targets of SAS) to the PAR on Chr 9 [39].

In prior work, we phased the entire genome of the Japan Sea stickleback using a small pedigree design (electronic supplementary material, figure S1b) [42], giving us phased X and

Y sequences. This significantly increases our power to detect differences between the X and Y relative to the more common strategy of indirectly comparing the sex chromosomes by comparing XY males against XX females (e.g. [32,34,35]). Here, we first use these phased data to describe patterns of divergence and diversity on the sex chromosomes of the Japan Sea stickleback. We then look for windows with elevated X-Y divergence, which would be consistent with evidence of SAS in the PAR of its neo-sex chromosome pair.

2. Methods

(a) Sampling and crosses

We sampled 15 Japan Sea males and 13 threespine females from Akkeshi Bay and Biwase on Hokkaido, Japan. To obtain phased X and Y sequences, 15 hybrid crosses were made from these individuals, using one threespine female three times. DNA was extracted from fin clips taken from each father and mother, plus one son and one daughter per cross (electronic supplementary material, figure S1b). Paired end whole-genome sequencing was conducted on all 58 individuals. Eight crosses were sequenced at the Genomics Shared Resources of the Fred Hutchinson Cancer Research Center on an Illumina HiSeq 2500, and the remainder were sequenced at the Next Generation Sequencing Platform of the University of Bern on an Illumina HiSeq 3000. The average coverage was 28X. Full details of the crosses and sequencing are given in Sardell *et al.* [42].

(b) Genotyping and phasing

We used *FastQC* v. 11.5 [43] to remove poor quality reads. Reads were mapped to the Glazer *et al.* [44] threespine stickleback reference genome using *bwa mem* v. 7.12 [45]. SNPs were sorted using *SAMtools*, per sample depth was calculated using *SAMtools depth*, and genotypes were called using *mpileup* v. 1.3 [46]. We filtered variants for minimum quality scores of 999, a minimum genotype quality of 20, and a minimum mean depth of 10X using *VCFTools* v. 1.15 [47]. To reduce genotyping errors resulting from mapping of paralogues, we also removed variants with mean depth greater than 1.5 times the median coverage on Chr 9 and 19. Lastly, variants were filtered to biallelic single nucleotide polymorphisms (SNPs) using *VCFTools* v. 1.15.

We used a custom *R* script (previously made available in [48]) to phase the data. Briefly, at all heterozygous sites in an offspring, we identified the paternal allele as that which was present in the father and absent in the mother. Crossing different species gave us increased power to phase, as the parents were fixed for alternate variants for many of the heterozygous sites in the offspring. We removed any sites in which the offspring and both parents were all heterozygous because they cannot be phased. Sites at which an offspring is homozygous for an allele that is not carried by one of its parents were discarded as they result from sequencing or other type of error. We filtered out any sites where more than five offspring across all families were missing phased genotypes. This pipeline resulted in phased sequences of the sex chromosomes, with 112 463 SNPs on Chr 9 and 81 061 SNPs on Chr 19 across the paternally inherited sex chromosomes from Japan Sea stickleback and the maternally inherited sex chromosomes from threespine stickleback. Of these, 35 409 sites on Chr 9 and 58 284 sites on Chr 19 were polymorphic among the 15 X chromosomes and 15 Y chromosomes sampled from the Japan Sea stickleback.

Population genetic analyses used to describe the evolution of the Japan Sea sex chromosome are detailed in the electronic supplementary material. Briefly, we calculated divergence (*K*) between the threespine and Japan Sea stickleback on Chr 9 and

Chr 19. We also calculated divergence (K), nucleotide diversity (π) and Tajima's D between the Japan Sea X and Y in 10 kb windows. On average, we observed 10 SNPs per 10 kb window across the PAR on Chr 9. We therefore built gene trees in 100 kb non-overlapping windows to have a larger number of phylogenetically informative sites. In the recombining PAR, the X and Y are intermingled on the gene trees, while in the non-recombining SDR, all Y chromosomes form a monophyletic clade to the exclusion of the X chromosomes [48,49]. Gene trees with this property are said to be 'XY gene tree consistent'. Finally, we calculated d_n , d_s and d_n/d_s on the Japan Sea X and the Y relative to the threespine stickleback using gene annotations from Glazer *et al.* [44].

(c) Detecting sexually antagonistic selection

The region on the PAR close to the SDR is an attractive region to search for signals of SAS using the phased sequencing data from the Japan Sea X and Y chromosomes. This is because peaks of F_{ST} resulting from SAS are more visible on the PAR [29]. However, in the region of the PAR that is extremely close to the SDR (within several ρ , where $\rho = 4N_e r$, N_e is the effective population size, and r is the rate of recombination in males between the focal site and the SDR), high F_{ST} between the X and the Y is expected even in the absence of SAS, owing to LD between the site and the SDR. To determine the region of the PAR on Chr 9 suitably distant from the SDR, we modified previously published neutral coalescent simulations [29,50] to account for two kinds of demographic effects: the demographic history of the species [36], and the selective sweep that occurred when the neo-Y was established. For the latter, we assumed the neo-Y originated just after the Japan Sea and threespine sticklebacks diverged 306 000 generations ago [36]. (Further details are given in the electronic supplementary material, and the code is available at <https://github.com/adagilis/SexCoal>). We simulated 10^4 replicates, each producing samples of sequences from 15 X and 15 Y chromosomes. Low density linkage maps [39,41,42] and coalescent simulations (electronic supplementary material, figure S2) suggest that the region in the PAR of Chr 9 starting at 7.5 Mb, which lies at 0.6 Mb (= 0.6 cM = 60 ρ) from the SDR, has sufficiently high recombination with the SDR to avoid high F_{ST} in the absence of SAS. While windows as close as 40 ρ show reduced F_{ST} , we conservatively chose a 50% greater distance given the uncertainty in recombination rates within the region. There are also reasons to limit how far from the SDR we should seek signals of SAS. SAS is less likely to maintain polymorphisms when recombination with the SDR is high and correcting for more multiple comparisons takes an increasing toll on statistical power. We therefore defined a focal region for the search for SAS starting at 7.5 Mb and ending at 11.5 Mb (roughly 460 ρ from the SDR).

Theory shows that a peak in F_{ST} caused by SAS is expected to be several ρ wide [29]. We estimate that ρ corresponds to about 10 kb in the Japan Sea stickleback, given estimates for N_e of 10^5 [36] and a local recombination rate of 1 cM/Mb on Chr 9 in males near the SDR [41]. We therefore estimated F_{ST} between the X and Y chromosomes in non-overlapping windows of 10 kb. A total of 178 10 kb windows meeting our filtering criteria (greater than 10 sites polymorphic among Japan Sea haplotypes) within the focal region of Chr 9 were used for further analysis.

We first used a model-free approach to detect SAS in the focal region. We identify potential SAS targets as windows with higher F_{ST} than expected by chance. To generate a null distribution for F_{ST} in each 10 kb window we randomized the labels of X and Y haplotypes, keeping equal numbers of Xs and Ys, 10^6 times. The p -value for each window was calculated as the proportion of randomizations for that window with an F_{ST} value equal to or greater than the observed value. To correct for

multiple comparisons, we used the *qvalue* package in *R*, which returns adjusted p -values, following Storey & Tibshirani [51]. Since we are only testing 178 windows and only expected a few outliers at most, we chose a less conservative false discovery rate (FDR) of 0.2 [52].

We used a second approach to look for SAS that incorporates the effects of demographic history and the selective sweep of the Y in generating high F_{ST} . We first performed 1000 coalescent neutral simulations, each yielding 15 X and 15 Y sequences, under the demographic assumptions described above. To obtain an FDR of 0.2, we used a more stringent multiple comparison correction and adjusted the critical p -value so that at least one significant peak of F_{ST} was obtained in 20% of the simulations. Further details are given in the electronic supplementary material.

Lastly, we implemented two further approaches to detect SAS: approximate Bayesian computation, and a random forest algorithm. Further details and justifications are provided in the electronic supplementary material. Neither approach could reliably distinguish between PARs containing loci with SAS and those without.

3. Results and discussion

(a) Evolution of the Japan Sea sex-determining region

On the ancestral sex chromosome (Chr 19), we find that divergence (K) between the X and Y is higher than divergence between the Japan Sea and threespine sticklebacks along the entirety of the SDR shared with threespine stickleback (figure 1). These data are consistent with previous work showing that the SDR comprises three strata that are older than the divergence between Japan Sea and threespine stickleback [36], suggesting the ancestral sex chromosome is shared between species. We therefore define three strata on the ancestral sex chromosome, S1, S2 and S3, at boundaries identified by Peichel *et al.* [24]. We find significant departures from equal read depths in males and females. The read depth ratio is about one half in stratum 1, which is the oldest, and jumps to much higher values in strata 2 and 3 (figure 1). A ratio of one half is expected when much of the Y has been deleted or has diverged to the point that reads from it no longer map to the reference sequence [53]. In the threespine stickleback, 82% of genes have been deleted or have highly degenerated in this stratum of the Y [24,38,54–56], and because of its shared ancestry with the Japan Sea Y it is likely that a similar situation also occurs in that species.

Because stratum 1 is so degenerate, some aspects of its molecular variation are difficult to interpret. We find elevated F_{ST} between the X and Y, despite the lack of significant difference in d_N/d_S , with substitutions measured relative to the threespine stickleback (electronic supplementary material, figure S3). However, a real difference may be obscured because sequences on the Y that have diverged too much from the X do not map to the reference and so are missing from our data. Both d_N and d_S are, however, elevated on the Y compared to the X.

Strata 2 and 3 show less signs of degeneration, but nonetheless show substantial differentiation in allele frequencies (F_{ST}) between X and Y sequences (electronic supplementary material, figure S4). This may be because they are substantially younger than stratum 1 [24]. Along these strata, the ratio of read depth in males versus females is close to 0.85 (figure 1). The values of d_N , d_S , and d_N/d_S are all significantly

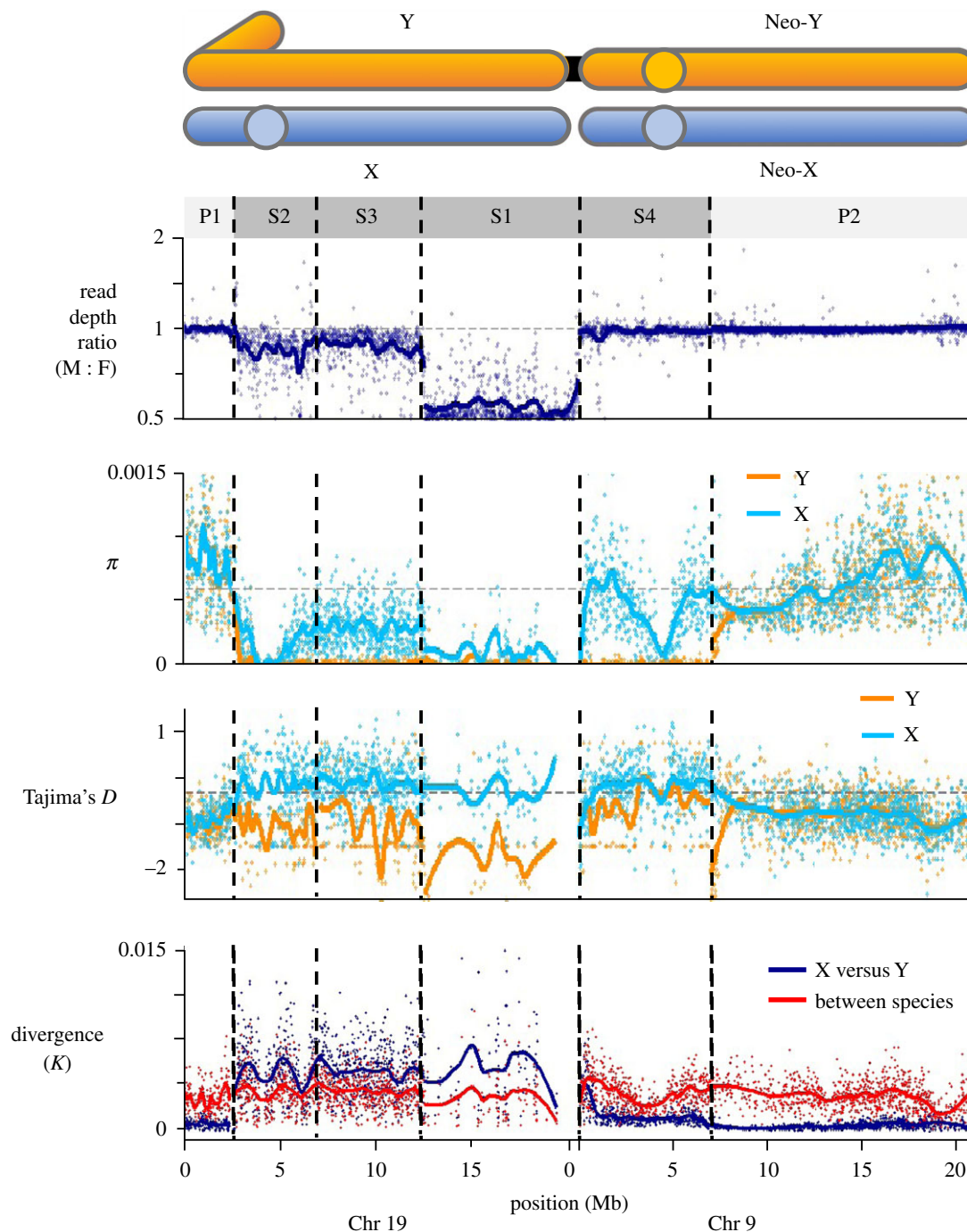


Figure 1. Population genomics of the X and Y chromosomes in Japan Sea stickleback. Four statistics are shown for the Japan Sea stickleback sex chromosomes. Dashed vertical lines give the boundaries of the evolutionary strata. Circles in the chromosome schematics show the locations of centromeres. The two PARs are labelled as P1 and P2, and the four strata in the SDR of the Y chromosome are S1, S2, S3 and S4 (from oldest to youngest). Dashed horizontal lines indicate autosomal averages. Loess curves were fitted separately for each PAR and stratum. Read depth ratio was calculated between sons and daughters, while π and Tajima's D were calculated for the Japan Sea X and Y. Divergence (K) was calculated between the Japan Sea X and Y and between the Japan Sea and threespine haplotypes.

larger on the Y than on the X in this region (electronic supplementary material, figure S3). The lack of degeneration in this region of the ancestral Y contrasts strongly with the closely related blackspotted stickleback (*Gasterosteus wheatlandi*) [48]. In that species, a substantial proportion of the Y on Chr 19 has nearly completely degenerated following a fusion with an autosome, with read depth ratios close to half across all strata of the SDR.

Stratum 4, which is on the neo-sex chromosome (Chr 9), is the youngest. To map the boundary between stratum 4 of the SDR and the PAR on the neo-sex chromosome, we used three statistics: divergence (K), XY gene tree consistency, and

nucleotide diversity on the phased Y chromosome (π_Y). All three statistics place the boundary at the 6.9 Mb position of Chr 9 (figures 1 and 2), which is consistent with earlier findings [40,41].

There is very little evidence of degeneration on the Y along stratum 4: there is no difference in read depth between males and females, and F_{ST} is considerably smaller than in strata 1–3 (figure 1, electronic supplementary material, figure S4). F_{ST} reaches values close to 1 near the fusion and near the centromere of Chr 9, a pattern driven at least in part by reduced polymorphism on the X (figure 1, electronic supplementary material, figure S4). There is no significant

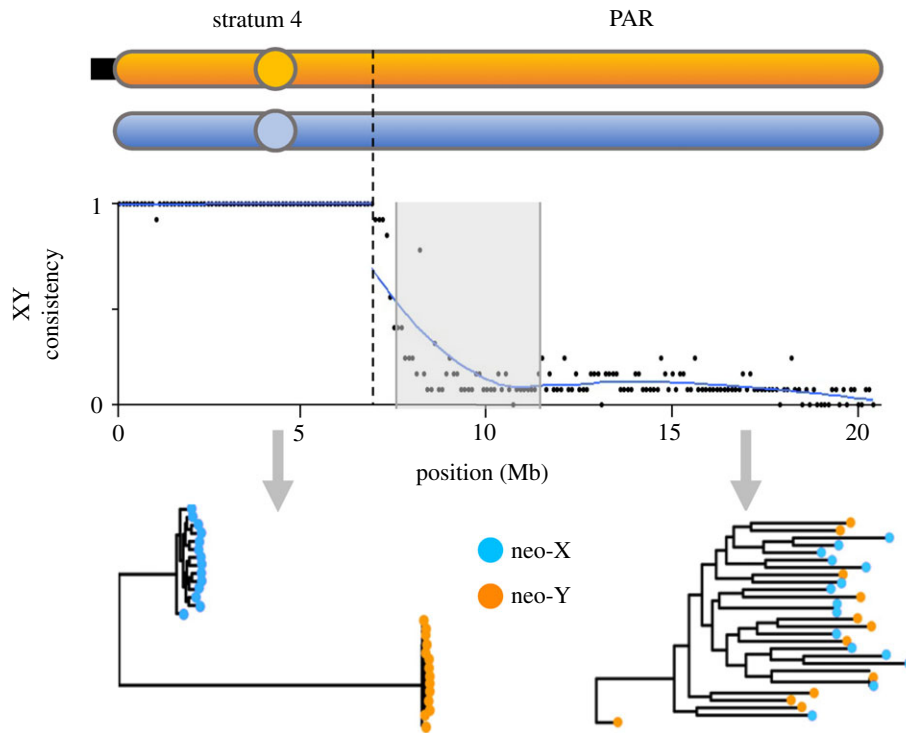


Figure 2. XY gene tree consistency of stratum 4 on Chr 9. XY gene tree consistency maps the boundary between the SDR and PAR to 6.9 Mb (dashed vertical line). The shaded area represents the focal region of the PAR investigated for evidence of SAS (figure 3). At the bottom are examples of a gene tree from the SDR (which shows XY gene tree consistency) and one from the PAR (which does not). Only Japan Sea X and Y data are shown.

difference in d_N/d_S between the X and Y in this stratum, although rates of synonymous substitution on the Y are elevated (electronic supplementary material, figure S3).

The divergence (K) between the Japan Sea and threespine sticklebacks is much greater than the divergence between the X and Y across much of stratum 4 (figure 1). That suggests that suppression of recombination in this region occurred much more recently than the speciation event (and more recently than suggested by [39]). The speciation event has been dated to between 0.6 and 1.2 million years ago [36], which suggests recombination was lost only within the last few hundred thousand years (or a few hundred thousand generations). It is plausible that recombination ceased in stratum 4 when Chr 9 fused to the S1 end of the ancestral sex chromosome (Chr 19). Consistent with this idea, divergence (K) between the X and Y in stratum 4 is highest near to the fusion point, and roughly equal to the divergence between the species (figure 1). This is the pattern expected if recombination was only suppressed in the region immediately adjacent to the fusion when it first appeared. An alternative hypothesis is that the fusion occurred around the time of the speciation event and recombination was suppressed later by some other mechanism. However, we have no evidence of other structural variants on the Japan Sea Y that would have later suppressed recombination [57]. The overall low divergence and degeneration of the Y observed in this region is similar to the neo-sex chromosome of the blackspotted stickleback, in which a fusion occurred between the ancestral Y and Chr 12 [48].

(b) Signals of sexually antagonistic selection in the pseudoautosomal region

We first used a model-free approach to search for signals of SAS along our focal region of 7.5–11.5 Mb on Chr 9.

To evaluate significance, we randomized X and Y labels to generate a null distribution of F_{ST} for each window. We find four peaks in F_{ST} between the X and Y chromosomes that are significantly greater than expected at an FDR of 0.2 (figure 3; electronic supplementary material, table S1). Under the assumption of independence, the probability of all four peaks are false positives with an FDR of 0.2 is 0.0016. The values of Tajima's D among the combined X and Y chromosomes in all four windows are positive (mean = 0.17), while the average for the entire PAR is negative (−0.17). This observation is consistent with SAS acting on sites in those windows. Two of the windows overlap with coding regions, while the other two windows are very close (less than 10 kb) to coding regions. Their proximity to coding regions is not by itself very surprising, however, as more than half of the 178 windows in the PAR fall within that distance of a coding region.

More surprising is that all four of the genes near these F_{ST} peaks are associated with mental disorders and neurological development in humans (electronic supplementary material, table S1). Peak I contains an SNP with elevated F_{ST} that lies within an exon of a TUBB-like protein, and TUBB is associated with autism and schizophrenia [58]. Peak II is near DNAJB5, a gene implicated in schizophrenia [59]. Peak III overlaps most of the coding region of a gene most similar to human ADD2 and ADD3, associated with bone mineral density [60] and bipolar depression [61], respectively. Two of the SNPs in this peak result in nonsynonymous changes, but neither has a known phenotypic effect. Last, peak IV is upstream of TUSC3, a protein associated with schizophrenia and unipolar depression [62].

Both schizophrenia and depression have been proposed as phenotypes subject to SAS in humans [63]. While we do not know the phenotypic effects of these genes in stickleback, their association with genes for traits potentially under

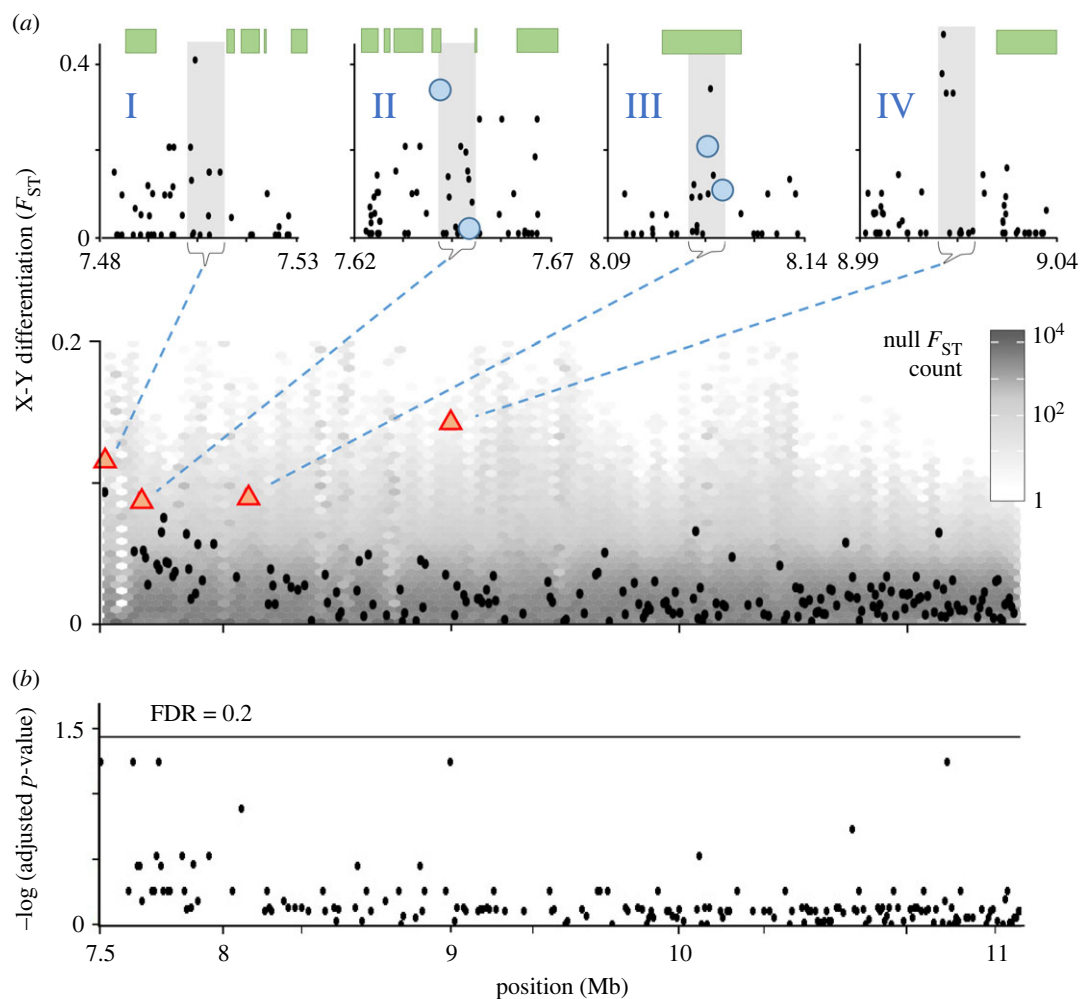


Figure 3. F_{ST} peaks between the X and Y in the PAR. (a) Bottom: F_{ST} between X and Y chromosomes of Japan Sea stickleback in 10 kb windows. The boundary between the SDR and PAR is at 6.9 Mb, to the left of the focal region shown here. The grey heatmap is the null distribution for F_{ST} generated by randomizing X and Y labels. Black dots are the average F_{ST} for each window. Red triangles are F_{ST} values in four windows showing significance with an FDR of 0.2 (for their F_{ST} quantiles and Q -values see the electronic supplementary material, table S1). Top: the four significant windows for F_{ST} , showing values of individual SNPs (black dots). The shaded regions show the windows, green rectangles show coding regions for individual genes, and blue circles show nonsynonymous SNPs in coding regions. (b) Adjusted p -values for each window compared to FDR cut-offs determined by coalescent simulations that incorporate a recent demographic sweep of the Y.

SAS in humans lends further support to the hypothesis that they are under SAS in stickleback. Some 5% of genes in our focal region have orthologues in the human Ensembl GRCh38.p12 database that are included in the gene ontology category 'nervous system processes'. The probability that four genes chosen at random would be associated with that category is therefore 0.05^4 , or less than 10^{-5} . While the proportion of genes that are actually involved in nervous system functions in stickleback is unknown, it does seem plausible that there is a positive association between the windows with high F_{ST} between the X and Y in the Japan Sea PAR and nervous system functions subject to SAS.

It is conceivable that one or more of the F_{ST} peaks result from errors in which a sequence lying within the SDR is incorrectly mapped to the PAR in the reference genome [64–66]. We evaluated this possibility using windows in the autosomes with F_{ST} between paternal haplotypes greater than 0.08, the smallest average value of F_{ST} among the four significant windows in the PAR. Based on the frequency of 0.061% of those windows in autosomes, a calculation shows that the probability that mapping errors would cause four

or more windows with high F_{ST} to occur in the 178 windows of the focal region is less than 10^{-5} . Furthermore, no SNPs with high F_{ST} in the four candidate windows lie in regions that have high similarity to regions in the SDR, as assessed by blasting 1 kb windows of the peaks against the SDR (although one low F_{ST} window in peak I blasts to the SDR). This source of error, therefore, does not seem to account for the patterns seen in the PAR.

It is also possible that the one or more of the F_{ST} peaks result from SNPs in LD with the SDR increasing in frequency as the result of demographic effects of the recent population expansion [36] or the selective sweep that occurred when the neo-Y was established. We find that none of our four peaks remain significant when measured against an FDR set by coalescent simulations that include both a sweep of the Y and recent population expansion (figure 3). However, a weakness of this simulation approach is that it relies on several parameters for which we do not have accurate estimates. These include the timing of the sweep on the Y, local recombination rates along the PAR, and recombination rate differences between males and females (see the electronic supplementary material for further details).

Several other caveats apply to our interpretation of the signatures of SAS we find. Despite focusing only on windows with at least 10 high quality SNPs, all four peaks are driven by a small number of SNPs with high F_{ST} (figure 3). Peaks I and II could be false positives that show elevated F_{ST} simply because of very tight linkage to the SDR, rather than the action of SAS. We estimate that peak I lies about 60ρ from the SDR, which is beyond where the effects of tight linkage to the SDR should affect allele frequencies neutrally evolving chromosomes (electronic supplementary material, figure S2). That estimate depends, however, on estimates of linkage and population size that include uncertainty. Peaks III and IV are much further from the SDR and so are much less likely to result in the absence of SAS. In fact, peak IV is so far away (2.1 cM) that it is somewhat surprising that a signal of SAS is visible there.

4. Conclusion

Identifying genes under SAS on sex chromosomes is complex. It is difficult in old sex chromosomes with small PARs, as meiosis requires at least one crossover, and this cross-over occurs in a small region. As a result, the PAR experiences high recombination rates not more conducive to maintaining polymorphisms under SAS than on autosomes [10–12,67–69]. Our study shows that very young PARs also present difficulties. In particular, the effects of demographic history can generate patterns of differentiation between the X and Y that look very much like SAS. Our simulations suggest that vastly increased sample sizes might allow us to distinguish peaks arising from SAS from alternative

hypotheses with more certainty. The genes around the peaks in F_{ST} we identify make them plausible candidates for SAS, but functional validation is needed to examine whether these polymorphisms underlie fitness differences between males and females to definitively test for SAS. Despite these limitations, the bioinformatics approaches like those in this study are a powerful strategy to identify candidate genes that contribute to the evolution of supergenes.

Data accessibility. All sequence data were previously published in Sardell *et al.* [42] and were previously deposited in the NCBI Short Read Archive, with reference number SRP135745. Coalescent simulation code is available at <https://github.com/adagilis/SexCoal> and all scripts used to analyse the data are available at <https://github.com/adagilis/JapanSeaSAS>. Data are provided in the electronic supplementary material [70].

Authors' contributions. A.J.D.: data curation, formal analysis, methodology, writing—original draft; J.M.S.: formal analysis, methodology, writing—review and editing; M.P.J.: investigation, writing—review and editing; Y.S.: formal analysis; M.K.: conceptualization, funding acquisition, methodology, project administration, resources, supervision, writing—review and editing; C.L.P.: conceptualization, funding acquisition, methodology, project administration, resources, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This work was supported by a grant from the National Institutes of Health (R01 GM116853) to M.K. and C.L.P.; A.J.D. was partially supported by the National Science Foundation (1737752).

Acknowledgements. We are very grateful to Doris Bachtrog for suggesting the crossing design, Jun Kitano and Asano Ishikawa for sampling the fish and performing the crosses, and the Kirkpatrick laboratory for very useful comments on the manuscript.

References

- Charlesworth D. 2017 Evolution of recombination rates between sex chromosomes. *Phil. Trans. R. Soc. B* **372**, 20160456. (doi:10.1098/rstb.2016.0456)
- Thompson MJ, Jiggins CD. 2014 Supergenes and their role in evolution. *Heredity* **113**, 1–8. (doi:10.1038/hdy.2014.20)
- Hughes JF, Page DC. 2015 The biology and evolution of mammalian Y chromosomes. *Annu. Rev. Genet.* **49**, 507–527. (doi:10.1146/annurev-genet-112414-055311)
- Lahn BT, Page DC. 1999 Four evolutionary strata on the human X chromosome. *Science* **286**, 964–967. (doi:10.1126/science.286.5441.964)
- Handley LJ, Cepelis H, Ellegren H. 2004 Evolutionary strata on the chicken Z chromosome: implications for sex chromosome evolution. *Genetics* **167**, 367–376. (doi:10.1534/genetics.167.1.367)
- Charlesworth D. 2021 When and how do sex-linked regions become sex chromosomes? *Evolution* **75**, 569–581. (doi:10.1111/evo.14196)
- Rowe L, Chenoweth SF, Agrawal AF. 2018 The genomics of sexual conflict. *Am. Nat.* **192**, 274–286. (doi:10.1086/698198)
- Arnqvist G, Rowe L. 2005 *Sexual conflict*. Princeton, NJ: Princeton University Press.
- Mank JE. 2017 The transcriptional architecture of phenotypic dimorphism. *Nat. Ecol. Evol.* **1**, 6. (doi:10.1038/s41559-016-0006)
- Rice WR. 1987 The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution* **41**, 911–914. (doi:10.1111/j.1558-5646.1987.tb05864.x)
- Otto SP *et al.* 2011 About PAR: the distinct evolutionary dynamics of the pseudoautosomal region. *Trends Genet.* **27**, 358–367. (doi:10.1016/j.tig.2011.05.001)
- Charlesworth B, Jordan CY, Charlesworth D. 2014 The evolutionary dynamics of sexually antagonistic mutations in pseudoautosomal regions of sex chromosomes. *Evolution* **68**, 1339–1350. (doi:10.1111/evo.12364)
- Ponnikas S, Sigeman H, Abbott JK, Hansson B. 2018 Why do sex chromosomes stop recombining? *Trends Genet.* **34**, 492–503. (doi:10.1016/j.tig.2018.04.001)
- Rice WR. 1984 Sex chromosomes and the evolution of sexual dimorphism. *Evolution* **38**, 735–742. (doi:10.1111/j.1558-5646.1984.tb00346.x)
- Patten MM. 2019 The X chromosome favors males under sexually antagonistic selection. *Evolution* **73**, 84–91. (doi:10.1111/evo.13646)
- Levine MT, Jones CD, Kern AD, Lindfors HA, Begun DJ. 2006 Novel genes derived from noncoding DNA in *Drosophila melanogaster* are frequently X-linked and exhibit testis-biased expression. *Proc. Natl Acad. Sci. USA* **103**, 9935–9939. (doi:10.1073/pnas.0509809103)
- Chen ST, Cheng HC, Barbash DA, Yang HP. 2007 Evolution of *hydra*, a recently evolved testis-expressed gene with nine alternative first exons in *Drosophila melanogaster*. *PLoS Genet.* **3**, e107. (doi:10.1371/journal.pgen.0030107)
- Arguello JR, Chen Y, Yang S, Wang W, Long M. 2006 Origination of an X-linked testes chimeric gene by illegitimate recombination in *Drosophila*. *PLoS Genet.* **2**, e77. (doi:10.1371/journal.pgen.0020077)
- Carvalho AB, Koerich LB, Clark AG. 2009 Origin and evolution of Y chromosomes: *Drosophila* tales. *Trends Genet.* **25**, 270–277. (doi:10.1016/j.tig.2009.04.002)
- Hall AB, Qi Y, Timoshevskiy V, Sharakhova MV, Sharakhov IV, Tu Z. 2013 Six novel Y chromosome genes in *Anopheles* mosquitoes discovered by independently sequencing males and females. *BMC Genomics* **14**, 273. (doi:10.1186/1471-2164-14-273)
- Tobler R, Nolte V, Schlötterer C. 2017 High rate of translocation-based gene birth on the *Drosophila* Y chromosome. *Proc. Natl Acad. Sci. USA* **114**, 11 721–11 726. (doi:10.1073/pnas.1706502114)
- Emerson JJ, Kaessmann H, Betrán E, Long M. 2004 Extensive gene traffic on the mammalian X

- chromosome. *Science* **303**, 537–540. (doi:10.1126/science.1090042)
23. Zhang YE, Vibranovski MD, Landback P, Marais GAB, Long M. 2010 Chromosomal redistribution of male-biased genes in mammalian evolution with two bursts of gene gain on the X chromosome. *PLoS Biol.* **8**, e1000494. (doi:10.1371/journal.pbio.1000494)
24. Peichel CL *et al.* 2020 Assembly of the threespine stickleback Y chromosome reveals convergent signatures of sex chromosome evolution. *Genome Biol.* **21**, 177. (doi:10.1186/s13059-020-02097-x)
25. Dufresnes C, Crochet P-A. 2022 Sex chromosomes as supergenes of speciation: when amphibians defy the rules. *Phil. Trans. R. Soc. B* **377**, 20210202. (doi:10.1098/rstb.2021.0202)
26. Ruzicka F, Connallon T. 2020 Is the X chromosome a hot spot for sexually antagonistic polymorphisms? Biases in current empirical tests of classical theory. *Proc. R. Soc. B* **287**, 20201869. (doi:10.1098/rspb.2020.1869)
27. Jeffries DL, Gerchen JF, Scharmann M, Pannell JR. 2021 A neutral model for the loss of recombination on sex chromosomes. *Phil. Trans. R. Soc. B* **376**, 20200096. (doi:10.1098/rstb.2020.0096)
28. Kratochvil L *et al.* 2021 Expanding the classical paradigm: what we have learnt from vertebrates about sex chromosome evolution. *Phil. Trans. R. Soc. B* **376**, 20200097. (doi:10.1098/rstb.2020.0097)
29. Kirkpatrick M, Guerrero RF. 2014 Signatures of sex-antagonistic selection on recombining sex chromosomes. *Genetics* **197**, 531–541. (doi:10.1534/genetics.113.156026)
30. Kasimatis KR, Ralph PL, Phillips PC. 2019 Limits to genomic divergence under sexually antagonistic selection. *Genes Genomes Genet.* **9**, 3813–3824. (doi:10.1534/g3.119.400711)
31. Cheng C, Kirkpatrick M. 2016 Sex-specific selection and sex-biased gene expression in humans and flies. *PLoS Genet.* **12**, e1006170. (doi:10.1371/journal.pgen.1006170)
32. Qiu S, Bergero R, Guirao-Rico S, Campos JL, Cezard T, Gharbi K, Charlesworth D. 2016 RAD mapping reveals an evolving, polymorphic and fuzzy boundary of a plant pseudoautosomal region. *Mol. Ecol.* **25**, 414–430. (doi:10.1111/mec.13297)
33. Guirao-Rico S, Sánchez-Gracia A, Charlesworth D. 2017 Sequence diversity patterns suggesting balancing selection in partially sex-linked genes of the plant *Silene latifolia* are not generated by demographic history or gene flow. *Mol. Ecol.* **26**, 1357–1370. (doi:10.1111/mec.13969)
34. Hill ST, Coggins J, Liston A, Hendrix D, Henning JA. 2016 Genomics of the hop pseudo-autosomal regions. *Euphytica* **209**, 171–179. (doi:10.1007/s10681-016-1655-9)
35. Bergero R, Gardner J, Bader B, Yong L, Charlesworth D. 2019 Exaggerated heterochiasmy in a fish with sex-linked male coloration polymorphisms. *Proc. Natl Acad. Sci. USA* **116**, 6924–6931. (doi:10.1073/pnas.1818486116)
36. Ravinet M, Yoshida K, Shigenobu S, Toyoda A, Fujiyama A, Kitano J. 2018 The genomic landscape at a late stage of stickleback speciation: high genomic divergence interspersed by small localized regions of introgression. *PLoS Genet.* **14**, e1007358. (doi:10.1371/journal.pgen.1007358)
37. Peichel CL *et al.* 2004 The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. *Curr. Biol.* **14**, 1416–1424. (doi:10.1016/j.cub.2004.08.030)
38. Ross JA, Peichel CL. 2008 Molecular cytogenetic evidence of rearrangements on the Y chromosome of the threespine stickleback fish. *Genetics* **179**, 2173–2182. (doi:10.1534/genetics.108.088559)
39. Kitano J *et al.* 2009 A role for a neo-sex chromosome in stickleback speciation. *Nature* **461**, 1079–1083. (doi:10.1038/nature08441)
40. Yoshida K *et al.* 2014 Sex chromosome turnover contributes to genomic divergence between incipient stickleback species. *PLoS Genet.* **10**, e1004223. (doi:10.1371/journal.pgen.1004223)
41. Natri HM, Shikano T, Merila J. 2013 Progressive recombination suppression and differentiation in recently evolved neo-sex chromosomes. *Mol. Biol. Evol.* **30**, 1131–1144. (doi:10.1093/molbev/mst035)
42. Sardell JM, Cheng C, Dagilins AJ, Ishikawa A, Kitano J, Peichel CL, Kirkpatrick M. 2018 Sex differences in recombination in sticklebacks. *Genes Genomes Genet.* **8**, 1971–1983. (doi:10.1534/g3.118.200166)
43. Andrews S. 2010 FastQC: a quality control tool for high throughput sequence data. See <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
44. Glazer AM, Killingbeck EE, Mitros T, Rokhsar DS, Miller CT. 2015 Genome assembly improvement and mapping convergently evolved skeletal traits in sticklebacks with genotyping-by-sequencing. *Genes Genomes Genet.* **5**, 1463–1472. (doi:10.1534/g3.115.017905)
45. Li H, Durbin R. 2010 Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595. (doi:10.1093/bioinformatics/btp698)
46. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009 Genome Project Data Processing Subgroup. 2009 The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079. (doi:10.1093/bioinformatics/btp352)
47. Danecek P *et al.* 2011 The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158. (doi:10.1093/bioinformatics/btr330)
48. Sardell JM, Josephson MP, Dalziel AC, Peichel CL, Kirkpatrick M. 2021 Heterogeneous histories of recombination suppression on stickleback sex chromosomes. *Mol. Biol. Evol.* **38**, 4403–4418. (doi:10.1093/molbev/msab179)
49. Dixon G, Kitano J, Kirkpatrick M. 2019 The origin of a new sex chromosome by introgression between two stickleback fishes. *Mol. Biol. Evol.* **36**, 28–38. (doi:10.1093/molbev/msy181)
50. Kirkpatrick M, Guerrero RF, Scarpino SV. 2010 Patterns of neutral genetic variation on recombining sex chromosomes. *Genetics* **184**, 1141–1152. (doi:10.1534/genetics.109.113555)
51. Storey JD, Tibshirani R. 2003 Statistical significance for genomewide studies. *Proc. Natl Acad. Sci. USA* **100**, 9440–9445. (doi:10.1073/pnas.1530509100)
52. Verhoeven KJF, Simonsen KL, McIntyre LM. 2005 Implementing false discovery rate control: increasing your power. *Oikos* **108**, 643–647. (doi:10.1111/j.0030-1299.2005.13727.x)
53. Vicoso B, Bachtrog D. 2011 Lack of global dosage compensation in *Schistosoma mansoni*, a female-heterogametic parasite. *Genome Biol. Evol.* **3**, 230–235. (doi:10.1093/gbe/evr010)
54. White MA, Kitano J, Peichel CL. 2015 Purifying selection maintains dosage-sensitive genes during degeneration of the threespine stickleback Y chromosome. *Mol. Biol. Evol.* **32**, 1981–1995. (doi:10.1093/molbev/msv078)
55. Roesti M, Moser D, Berner D. 2013 Recombination in the threespine stickleback genome—patterns and consequences. *Mol. Ecol.* **22**, 3014–3027. (doi:10.1111/mec.12322)
56. Schultheiss R, Viitaniemi HM, Leder EH. 2015 Spatial dynamics of evolving dosage compensation in a young sex chromosome system. *Genome Biol. Evol.* **7**, 581–590. (doi:10.1093/gbe/evw013)
57. Cech JN, Peichel CL. 2016 Centromere inactivation on a neo-Y fusion chromosome in threespine stickleback fish. *Chrom. Res.* **24**, 437–450. (doi:10.1007/s10577-016-9535-7)
58. Autism Spectrum Disorders Working Group of The Psychiatric Genomics Consortium. 2017 Meta-analysis of GWAS of over 16,000 individuals with autism spectrum disorder highlights a novel locus at 10q24.32 and a significant overlap with schizophrenia. *Mol. Autism* **8**, 21. (doi:10.1186/s13229-017-0137-9)
59. Yu H *et al.* 2018 Five novel loci associated with antipsychotic treatment response in patients with schizophrenia: a genome-wide association study. *Lancet Psychiat.* **5**, 327–338. (doi:10.1016/S2215-0366(18)30049-X)
60. Kichaev G, Bhatia G, Loh PR, Gazal S, Burch K, Freund MK, Schoech A, Pasaniuc B, Price AL. 2019 Leveraging polygenic functional enrichment to improve GWAS power. *Am. J. Hum. Genet.* **104**, 65–75. (doi:10.1016/j.ajhg.2018.11.008)
61. Charney AW *et al.* 2017 Evidence for genetic heterogeneity between clinical subtypes of bipolar disorder. *Transl. Psychiat.* **7**, e993. (doi:10.1038/tp.2016.242)
62. Li Q, Wineinger NE, Fu DJ, Libiger O, Alphas L, Savitz A, Gopal S, Cohen N, Schork NJ. 2017 Genome-wide association study of paliperidone efficacy. *Pharmacogenet. Genomics* **27**, 7–18. (doi:10.1097/FPC.0000000000000250)
63. Mokkonen M, Crespi BJ. 2015 Genomic conflicts and sexual antagonism in human health: insights from

- oxytocin and testosterone. *Evol. Appl.* **8**, 307–325. (doi:10.1111/eva.12244)
64. Mank JE, Shu JJ, Wright AE. 2020 Signature of sexual conflict is actually conflict resolved. *Mol. Ecol.* **29**, 215–217. (doi:10.1111/mec.15311)
65. Cheng C, Kirkpatrick M. 2020 The signal of sex-specific selection in humans is not an artefact: reply to Mank *et al.* *Mol. Ecol.* **29**, 1406–1407. (doi:10.1111/mec.15420)
66. Bisseger M, Laurentino TG, Roesti M, Berner D. 2020 Widespread intersex differentiation across the stickleback genome – The signature of sexually antagonistic selection? *Mol. Ecol.* **29**, 262–271. (doi:10.1111/mec.15255)
67. Jordan CY, Charlesworth D. 2012 The potential for sexually antagonistic polymorphism in different genome regions. *Evolution* **66**, 505–516. (doi:10.1111/j.1558-5646.2011.01448.x)
68. Clark AG. 1988 The evolution of the Y chromosome with X-Y recombination. *Genetics* **119**, 711–720. (doi:10.1093/genetics/119.3.711)
69. Connallon T, Clark AG. 2010 Sex linkage, sex-specific selection, and the role of recombination in the evolution of sexually dimorphic gene expression. *Evolution* **64**, 3417–3442. (doi:10.1111/j.1558-5646.2010.01136.x)
70. Dagilis AJ, Sardell JM, Josephson MP, Su Y, Kirkpatrick M, Peichel CL. 2022 Searching for signatures of sexually antagonistic selection on stickleback sex chromosomes. Figshare. (doi:10.6084/m9.figshare.c.5979250)