Dear Editors,

We have now revised our manuscript entitled "Genetic sex determination in three closely related hydrothermal vent gastropods, one of which has intersex individuals", which we resubmit to your consideration for recommendation in PCI evolutionary biology.

We thank recommender Tanja Schwander and all reviewers for their useful, constructive ideas. Below we provide point-by-point responses, indicating how each comment was taken into account into the revised manuscript (with line numbers referring to the revised version).

In addition, we have updated the following sections at the end of the manuscript to comply with PCI requirements: the "data availability" section was updated to "data, script, and code availability", which now gives direct links to the data stored in NCBI and a zenodo repository. We added a section called "Supplementary material" in which we indicate that the supplementary material can be found on bioarxiv together with the article. The doi is thus the same as for the main paper. We refrain from moving supplementary material to yet another archive in order to avoid duplication of data storage. Finally, we have now added specific "Funding" and "Conflict of interest disclosure" sections.

---

by Tanja Schwander, 27 May 2023 12:33
Manuscript: ***https://doi.org/10.1101/2023.04.11.536409*** *version 1*

Invitation to revise the preprint "Genetic sex determination in three closely related hydrothermal vent gastropods, one of which has intersex individuals"

Dear authors,

Thank you for submitting your preprint "Genetic sex determination in three closely related hydrothermal vent gastropods, one of which has intersex individuals" for evaluation by PCI Evol Biol.

Your manuscript has now been peer reviewed by three reviewers. As you will see, the reviewers all find your manuscript interesting, but make a number of constructive suggestions that will help you to improve and/or clarify the manuscript. I would particularly encourage you to focus the discussion more on the sexual and sex determination systems rather than the finding that sex linkage generates more structure in the dataset than geography. The current phrasing makes the bulk of the discussion rather taxon specific and largely a repetition of the results. Integrating the relevant discussion paragraphs directly into the results and mostly discussing the sexual systems and potential transition

from gonochorism to hermaphroditism will make the discussion more appealing to a general audience.

We have largely revised the discussion, removing repetitions of results, and focusing much more on sex determination and sexual systems. Note that we have made a slight changed to the title of the manuscript : "Genetic sex determination in three closely related hydrothermal vent gastropods, including one species with intersex individuals".

Furthermore, I think that the analyses and discussion related to the sex-linked regions (and notably the size of the sex-linked regions) could be significantly improved by including an analysis of male-specific regions. Since your analyses are based on a de-novo assembly of RAD loci, the most strongly differentiated regions on the sex chromosomes will assemble as separate loci in males and females. All these loci are filtered out since only loci genotyped in at least 80% of inds are included (and sex-ratios are about 50:50). By adding another set of analyses with sex specific filtering (eg. retain loci genotyped in 80% of males, 80% of females, and looking for overlap vs sex-specific loci), one could assess variation in X-Y divergence across the three species as well as the size of the male-specific genome region.

Following this suggestion, we have repeated the analyses of raw sequence data using sex-specific filters (new paragraph L. 261-272). We looked for snps that are totally absent from the females while being genotyped in at least 10 male individuals, therefore pointing to genomic regions that are either specific to the Y (present in males only) or homologous but strongly divergent on X and Y, as suggested. We found 40 such male-specific RAD loci in species *A. boucheti*, out of which 37 were found on 18 scaffolds of the related *A. marisindica* genome. Eleven of these scaffolds had already been identified as sitting on the sex chromosomes by our other analyses, and the seven others were thus newly identified as bearing Y-specific sequences, thereby adding to the total length of the sex-linked region in *A. boucheti*. By contrast, no such Y-specific data were found in the two other species. This result reinforces the difference in size of the sex-linked region in *A. boucheti* (gonochory) vs *A. kojimai* (andro-dioecy or otherwise altered gonochory with intersex individuals), while the absence of Y-specific sequences in *A. strummeri* is less informative due to lack of power.

Note that in the revision process we realized that using a 75% coverage threshold for our mapping analyses might have been a bit too permissive (in particular the RAD loci found to map on the two small scaffolds that were found to be sex-linked exclusively in species *A. strummeri* had rather low coverage). Hence, we decided to upscale this threshold to 95% (i.e. the mapping results were retained only for the RAD loci that were identified by *Geneious* as identified on the reference genome using at least 95% of their length, instead of 75%). This is now better justified in L. 285-292. This new threshold did not change the results for *A. kojimai* (five sex-linked scaffolds), and it reduced the number of sex-linked

scaffolds from 11 to 8 in *A. strummeri* (removing in particular the two dubious small scaffolds previously found to be sex-linked only in this species) and 48 to 41 in *A. boucheti*. Combined with the new results from the Y-specific analysis, the overall figure of the (minimum) respective sizes of sex-linked genomic regions remained largely unchanged (see new Fig. 6). As explained below following a related comment by D. Jeffries, we stress in the paper that our analyses can only provide a minimum size for sex-linked regions (L. 291). Good quality reference genomes are required to describe further the size and structure of the sex chromosomes in these three species (in particular with regards to the distribution of gene functions, repeated elements, Y-specific regions, and recombination landscape).

Finally, for genetic variance and structure analyses, I recommend using a single SNP per RAD locus (as SNPs on the same locus are physically very close and thus in very strong LD, i.e., not independent estimates). This would also affect the simulations, which will consider too many independent SNPs.

We re-analyzed the dataset by subsampling a single SNP of each RAD. We used this reduced dataset to compute the PCA of each species (new Fig. 1) and estimate all pop gen stats reported in table 1. While we did not anticipate this, using a single SNP per RAD improved the PCA in two aspects. First, it revealed more clearly the effect of a few outlier individuals (which we had left in the submitted version, but which we have now removed following suggestions from the reviewers, see below). Second, it improved the power of the PCA to identify the sex of individuals in species *A. strummeri* (axis 2 now separates the two sexes with no overlap). For this species we can now report the sex-ratio estimated purely from genetic data (L. 375).

On the other hand, in contrast with structure and variance analyses that use multi-SNP statistics, we think that statistics performed individually for each SNP should not consider only a single SNP per RAD locus, especially in the perspective of detecting sex-linkage. This is because many SNPs convey too little information (meaning that the alternative allele is too rare) to reveal whether its segregation is autosomal or not. Hence using a single SNP chosen randomly discards lots of useful information. For instance, in *A. boucheti*, by keeping only a single SNP of each RAD we would lose 121 RAD loci that we could not detect as sex-linked although they actually are (252 RAD instead of 373 found to be sex-linked). In this situation, using all SNPs per RAD produces no bias or otherwise wrong result. Instead, it provides the maximum level of information. Hence, we prefer to keep the analysis where we use all SNPs and where the number of SNPs and corresponding number of RADs are systematically given in the results, so that the reader understands the data. Also, for the simulations here we chose to use a number of SNPs dictated by the number of SNPs obtained empirically, but that is arbitrary (each simulation is for a single SNP – what is critical for the simulations is the number of individuals simulated). The more SNPs simulated the better, and we could have simulated say 500 000 or 1 million SNPs per species to get the best possible null envelop, since our point is to detect SNPs that have F_{IS}/F_{ST} combinations that can never be obtained if they had autosomal segregation. Really, as far as there are

enough independent SNPs simulated to cover the parameter space well, then the actual number does not matter here. So here again, we refrain to reduce the number of SNPs simulated because it would reduce the efficiency of the simulations to detect non-autosomal segregation (allowing for more false positives). Thus, all analyses and results leading to figures 4 and 5 are left unchanged (with regards to the number of SNPs used) in the revised version.

I invite you to revise the manuscript by addressing the comments provided.

Reviews

#1 Reviewed by Hugo Darras, 12 May 2023 10:43

In this biorxiv manuscript, Castel and colleagues explore the sexual and sex determination systems of three closely related hydrothermal vent gastropods. Dissections and histological analyses of 276 individuals revealed that two of these species are gonochoristic, while the third one appears to have both males and intersex individuals. Reanalysis of published RAD-seq data shows that all three species have an XY system and that their sex chromosomes probably share ancestry.

This study provides the first insights into the reproduction of animals living in deep marine vents, which have remained poorly studied, and opens new research perspectives on these animals. I found this very interesting. The methods are standard, the results are sound, and the introduction and discussion were enjoyable to read.

I only have minor comments:

- l140-144: This is a bit confusing. Are these new data?

We have now clarified that our dataset comes from the sampling, DNA extracting, library prep, sequencing, and demultiplexing previously performed for Castel et al. (2022). While we kept here a brief summary of these steps (L. 162-168), it is now shorter, and we made it clearer that all downstream analyses (starting from the demultiplexed reads) were specific to this new study (L. 169). We also emphasize now (L. 172) the most fundamental difference in data treatment: here the Stacks pipeline has been optimized and run independently for each species, unlike in our previous work.

- l145-165: What samples were used to build Stacks' catalog? Line 158 suggests that samples with missing data were removed, but the final data set still contains 499 samples.

The catalogs in Stacks were built using all samples before filtering for missing data (570 samples, divided into three batches, one for each species considered). This is now clarified on L. 163. The final data set filtered for missing data contained 499 samples overall (now 486 because we removed PCA outliers, see next comment). In the article, we only considered the sex of these individuals because it is already difficult to follow male and female sample sizes due to the fact that we have two sexing methods (hence 499 (now 486) was also the number reported earlier on L. 134 in the "sexing" paragraph).

- l145-165: Is there a filter for allele frequency? Knowing whether there are singletons in the final data set would be helpful. A bad quality library could increase SNP count dramatically (e.g., the individual discussed in l313 might be better removed).

Yes, we filtered out any SNP where the rare allele is represented by less than four copies, regardless of sample size (i.e. Minimum Allele Count MAC=4, as mentioned in the Sup. Mat. on p.4). Our experience (based on the systematic comparison of replicates) suggests that this allows to remove some genotyping errors while not removing a too large fraction of rare alleles that are biologically relevant. Following this comment, we looked more closely at the individual that stands out in the PCA shown in Fig. S4. With the help of an expert in multilocus PCA (Alan Le Moan, now acknowledged), we realized that several samples had a strong impact on PCA (visible when removing these samples one by one or two by two), potentially due to trace contaminations (although observed heterozygosity was high but not standing out) or perhaps stronger-than-average pairwise relatedness. We explored this in all three species, and this led us to remove 8 individuals in species *A. boucheti* and 5 in *A. kojimai* (none in *A. strummeri*). Removing these individuals did not change any of the results or conclusions (most importantly, the lists of scaffolds identified as sex-linked remained strictly unchanged in all species) but it changed the appearance of PCA for *A. boucheti* and *A. kojimai*. We have the feeling that the dataset is cleaner now, so we present the new results in this revised version (i.e. all analyses were repeated with the reduced dataset).

- l234-237: Theoretically, performing reciprocal blast using the whole RAD catalog would be better to avoid forcing matches.

The e-value used to decide what blast result is useful is not relative to the number of potential targets. Unless we have misunderstood this comment, we don't see how the reciprocal blast performed here could force matches (if the sequences are different, they will not match, regardless of the total number of possible targets).

- l239: Do we know the sex of the reference genome?

No, it was not reported by the authors, and we could not obtain an answer about it.

- l245: How were multiple hits handled?

They were not. Or not really. We used only the best hit (now corrected on L. 286, there was a mistake there left from a previous version). Yet we know that some of these loci map repeatedly, and it is clear that, at least in *A. boucheti*, the Y might be full of repeated elements. We stress this in the discussion, emphasizing that the length of the sex-linked genomic regions reported in Fig. 6B is a lower bound to the real sizes (L.292 and L.607-612). Yet, exploring further the genomic structure of the sex chromosomes in any efficient way will require good reference genomes and resequencing data for sexed individuals, as we also mention on L. 613.

- l255: Given that sex assignment proved error-prone, should the samples be described as "putative" males and "putative" females instead?

The sex inferred from external observation of the gonad seems to be roughly 9% wrong. We added "putative" in the first sentence of this paragraph (L. 316) but we did not use it repeatedly after that to avoid making the text too cumbersome.

- l255-265: Sample sizes differed from those in Table 1 and l168.

Sample sizes are correct, but they differ for reasons that are not easy to follow: using the revised version of the article, we start with 570 individuals (L. 163), down to 499 that passed missing data filtering (L. 182), down to 486 (L. 190) after PCA (because we removed outliers at this stage). In these 486, we had sex identification for 271 (table S1), but the sex of three *A. boucheti* individuals had been wrongly identified (visible in PCA). We reclassified these as "unknown", so that we end up with 268 sexed individuals (e.g. L. 193) used in all downstream analyses (anything but PCA). We have now made sure to clarify this throughout.

- l264-265: Sample sizes differed from those found in l267.

No mistake there either, but again, we understand that it is not easy to keep track of sample sizes. In the revised version, the results of histological analyses for *A. kojimai* are now: 48 males, 11 females, 31 "morphological hermaphrodites". These numbers are not easily compared to the results from the external observation of the gonads because samples for these two methods were only partially overlapping. We tried to be as clear as possible in this section despite the complexity of the dataset.

- l303-316: Only PC1 and PC2 are shown. Sexes may show strong clustering on other axes. An AMOVA could eventually be used to determine the total percentage of genetic variation explained by sex in each species and to disentangle sex from geography.

Since we have completely removed the section on sex and geography from the discussion, we are reluctant to add more analyses on this topic. Another reason is that we don't have enough data with species *A. strummeri* to disentangle these effects in any useful way (although we agree that AMOVA would have been the best way to do it).

- l313: Cropping a PCA plot seems unconventional. Ideally, it would be better to perform PCA without the problematic individual or choose other axes for plotting.

Yes, as detailed in our response to the editor's comment above, re-examining this outlier individual and using a single SNP per RAD led us to remove 8 *A. boucheti* individuals and 5 *A. kojimai* (including the problematic individual mentioned in this comment) from all analyses. The new PCA are visible in Fig. 3. This is indeed a better way to treat outliers rather than cropping the PCA.

- l345-347: Did all females have the same genotype at each of the 553 SNPs?

Yes, good point (confirming that we are looking at SNPs that have one allele fixed on the Y and an alternate allele fixed on the X). We now clarified this point L. 411

- l361-362: Could this be driven by a few errors in sex assignment or low coverage in some females?

No, with 48 males and 31 females and hard-filtered SNPs it is not possible that this result is artefactual. Moreover, this result was confirmed multiple times: first when repeating these analyses using only the individuals that had their gonads examined under a microscope (a reliable sexing method, see sup. mat. figure S7), second when using F_{ST}/F_{IS} analyses, and then when realizing that these SNPs were all located in 5 scaffolds that were also sex-linked in *A. boucheti* and *A. strummeri* (a result impossible to obtain by chance).

- l469-470: Could high differentiation also lead to smaller Fst values? Less sex-linked RAD loci will be recovered if the two chromosomes are highly divergent.

This section has been removed from the discussion, which has been substantially rewritten. For the sake of the argument, it is probably true that there is no simple relationship between the "sex effect" on a genome-wide multilocus PCA and the landscape of recombination and divergence on sex chromosomes. Measuring differentiation at homologous loci through F_{ST} is probably a better indicator of sex chromosome differentiation, but then strong divergence could be problematic when many loci are in fact non homologous but X-specific and F_{ST} alone is then useless (which is why we used a combination of F_{ST} and F_{IS} in downstream locus-specific analyses). In any case, the genome-wide genetic statistics have only indicative value. None of our conclusions depend on these synthetic results.

- l479: "on the multi-locus PCA" -> on the first two axes of the multi-locus PCA

Writing "on the first two axes" would give the reader the impression that we have seen an effect of sex on other axes. When performing PCA we looked systematically at the proportion of variance explained by the 50 first principal components (using the eigenvalues reported by Adegenet), and we explored visually the first 6 principal components. So, it is fair to say that "the sex of individuals did not drive any obvious pattern of genetic differentiation on the multi-locus PCA".

- l483: The percentage of total genetic variation explained by geography is likely higher (there are other PC axes).

This section of the discussion was removed.

-l556: "share the same sex-linked genomic region" -> share some sex-linked genomic regions.

We changed to "this sex-linked genomic region" to be more specific but leave open the possibility that other regions do exist (L. 567).

#2 Reviewed by anonymous reviewer, 09 May 2023 16:56

In this manuscript Castel et al analyse a RAD-seq dataset in three species of gasteropodes and study sex determinism of the species. They newly identify hermaphrodite individuals (females also producing male gametes) in A. kojimai. The study is well conducted and the paper is detailed and well written. I am mostly suggesting further analyses for future investigations or slight modification of the text (especially the discussion on A. kojimai XY system).

I am not a 100% convinced of the XY sex chromosomes in A. kojimai because heterozygosity in males is not equal to 1 and I think maybe the discussion should slightly tone down this conclusion. There could be partial X-Y recombination or a recent breakdown of the non-recombining region in A. kojimai. I understand that the observed patterns were not obtained in autosomal simulations, but could the authors please add more discussion on why heterozygosity is not 1 in A. kojimai males? One explanation may be that the non recombining region in A. kojimai is very small and the SNPs detected as "sex-linked" are in fact only partially sex-linked, outside of the non recombining region but strongly linked to it.

We fully agree with this interpretation. The results suggest that there has always been partial X-Y recombination, or a recent breakdown of the non-recombining region in *A. kojimai*, so that all the loci that we have detected are not linked strictly to the sex-determining region (i.e. they are not in a region where recombination is never happening). This is what we had in mind when mentioning that "recombination is also strongly reduced in *A. kojimai* (albeit over a much smaller genomic region)", but we agree that this was not clear enough. This section has now been largely rewritten. In particular, we have now clarified throughout the article that we use the words "XY chromosomes" to refer to the chromosomes that bear the sex-determining locus, and "sex-linked" for any pattern of variation that is not independent of sex. We do not limit "sex-linkage" to regions where recombination is completely arrested or has been so for some long period of time. As this comment suggests, our featured SNPs could lie is a region of the XY sex chromosomes where recombination was never completely suppressed, or has recently changed. Also, we do not imply that the sex chromosomes are strongly divergent and heteromorphic. We have now clarified that our interpretation of the results in the three species is that they share a male-heterogametic XY genetic system of sex determination but with sex-linked regions that have probably not the same size or structure. This interpretation was also reinforced by a new phylogenetic analysis presented in Figure 7 (see second-next comment below), where we see that all Y haplotypes are more similar to conspecific X haplotypes rather than to heterospecific Y, and X and Y chromosomes share more haplotypes in *A. kojimai* than in other species. This suggests that (for the three loci that we consider in this analysis) recombination has not been completely arrested since speciation events (in any species), and that it has been highest in species *A. kojimai* (see e.g. L. 620-640).

I was wondering if the authors would be interested in using the program SD-pop (Kafer et al 2021)? This would allow to statistically test for the presence of sex chromosomes with a BIC approach in each species. I appreciated the simulation approach taken by the authors to validate their sex-linked SNPs, however the simulations did not include a genotyping error nor duplicated genes, which SDpop does.

We preferred to design our own ad hoc simulations to detect loci that might be located on young sex chromosomes or on sex-linked genomic regions that have either recently stopped recombining or have not completely stopped recombining. It is correct that SD-pop is the program that comes closest to this mark among published methods, but it is not perfect in our particular case, since it does not allow for shared polymorphism between gametologs (it assumes that an allele is fixed in at least one of the two sex chromosomes, which is not the case for many of the sex-linked SNPs that we detected, and in particular for the sex-linked SNPs in *A. kojimai*).

Are the Y alleles identified in A. kojimai identical to the Y alleles in other species? Said otherwise, do the three species share Y alleles in the RAD loci that are commonly sex-linked? This would be a further argument to say that the species share the same chromosome pair (otherwise the same chromosome region may have become sex-linked independently in each species, although less likely, this is still possible).

This comment prompted us to examine into more details the nature of the SNPs and alleles at the RAD-seq loci that we found to be common between species (remember that RAD-sequencing provides only a reduced representation of the genomes and that divergence between species generates differences in the identity of the RADs independently sequenced in each species). We looked at all SNPs and alleles of the three RAD-seq loci that were genotyped and sex-linked in all three species (intersection of the Venn diagram in Fig. 6A). Doing so, we realized that we had access to phased haploid sequences (218 to 360 bp depending on the locus considered) and that it was easy to infer whether the sequence obtained was located on the X or Y chromosome of each diploid individual (because in most cases the males shared a single haplotype that was different from all others). Hence, we now present in new Fig. 7 a phylogeny of all haplotypes observed for the individuals that were sexed with confidence (i.e. used in histological analyses), and where the association to the X or Y is shown. This new analysis is the best way to answer the comment above: at these three loci, the Y chromosomes did not share the same haplotypes across species, and they were even always more closely related to conspecific X chromosomes than to allospecific Y chromosomes. This means that the shared XY chromosomes have undergone some recombination since speciation events (at least in the genomic region we consider here with three RADs), but not too much so that X and Y have since fixed different haplotypes within *A. boucheti* and within *A. strummeri* (see terminal branches in Figure 7). It also shows that X and Y divergence is highest in *A. boucheti* (longer branches at all three loci) and lowest (recombination strongest) in *A. kojimai*, where the X and Y chromosomes share common haplotypes.

Hence, although the three species share a common, sex-linked, genomic region of about 2.8 Mb that is linked to XY genetic sex determination, the recombination landscape in this region has been variable across species. The most parsimonious interpretation of the results is that the three species share an ancestral pair of XY chromosomes, but as mentioned in the comment we cannot exclude that the same region became sex-linked independently (hence, we talk only about shared, not ancestral, XY chromosomes in the article).

All this is now presented in new sections in the methods and discussion.

I wonder if the data could allow a GWAS analysis to look for variants specific to hermaphrodites? This may help understand how hermaphroditic sex is determined in A. kojimai.

Yes, that is part of our plan for the future. We are now performing more detailed histological analyses (more cross sections of the gonads, and estimation of the proportion of male and female tissues within each section) in the hope that we can relate these observations to multi-locus genotypes and refine our understanding of genetic sex determination in *A. kojimai*. But that is work under progress, and it is likely that we don't have enough genomic coverage with a RADseq data set (only 16 SNPs on 14 RAD loci). For this objective, WGS will be more promising.

What are the functions of the genes located in sex-linked scaffolds? This analysis could provide sex-determining gene candidates.

The reference genome for *A. marisindica* is not annotated, so that we can't easily answer this question. Before writing the article, we had been looking for any gene annotation (for any taxonomic group) for the sex-linked RAD loci that we identified, but with limited success: the only potentially interesting hits were for a gene involved in flagellated sperm motility, a flavin-containing monoamine oxidase A (dopamine degradation) that is located on the X chromosome of vertebrates. We decided this information was too fragmentary and too far from sex determination to be of interest at this stage.

Line 580: if selfing was possible in A. kojimai, wouldn't it be detected with higher Fis values? If selfing was not detected, then there must be a mechanism preventing self-fertilization in hermaphrodite individuals. Or maybe the male gametes are not functional?

This is correct, selfing should be reflected in increased F_{IS} levels (but the power of F_{IS} to detect small selfing rates is not great). We are planning to investigate the genetic signatures of selfing (including F_{IS} and multilocus individual heterozygosity) in another study, along with additional anatomical analyses to determine whether or not the male tissues of intersex *A. kojimai* individuals are actually functional. This is work under progress, but the relevant hypotheses are discussed on L. 643-703.

line 523 identified AS having.

Corrected L. 585

#3 Reviewed by Daniel Jeffries, 07 May 2023 12:46

Pradillon et al: Genetic sex determination in three closely related hydrothermal vent gastropods, one of which has intersex individuals

https://doi.org/10.1101/2023.04.11.536409

General comments

Here Pradillon et al present new data on the sex determination systems of three gastropod species. Gastropod sex determination systems are severely understudied so this is a welcome contribution to the field in that regard! The paper's strengths lie in the size and power in the dataset, and in the robust analyses performed to find the sex linked loci. In particular I very much like the simulation analyses to contextualise the Fst and Fis results from the real data. While I suggest a few minor improvements for the analyses (below), I have good confidence in the results and conclusions as they currently stand.

Note that although Florence Pradillon has made a major contribution to this work, the first author of this manuscript was and remains former PhD student Jade Castel, who has been at the core of this work.

The main weakness of the paper, as I see it, is the framing of the paper around the observation that "sex is the primary driver of the genetic structure" in these species. I suppose this was the finding that turned the authors onto exploring sex determination in these species, but I do not think this result has much value in the paper. The reason being, it confounds autosomal population structure with sex linkage. The most interesting results for the paper (given the title) are mainly the description of the sex determination systems, rather than the relative strength of signal from sex linkage vs geography in population structure analyses. The latter being difficult to interpret as it could be due to either highly differentiated sex chromosomes, or very weak population structure (likely both of course).

Following this comment and a similar comment by T. Schwander, the discussion is now focusing more on the description of sexual systems and sex determination.

Further, the population structure is perhaps of some interest too here, but in the present state it is obscured by the sex linkage. Now that the sex linked loci are known, they can be removed them, and

the population structure analyses can be re-run. This will undoubtedly give more accurate and high-resolution results, perhaps beyond the basin-level.

Thus, my suggestion would be to move Figs 3, S4, S5 and table 1 to the supplementary, and streamline the main text to focus on the sex determination systems.

We have been working on fine-scale genetic structure and local adaptation in these three species, and for that we have indeed removed the sex-linked RADs. This is the subject of another paper in preparation. As suggested here and in other comments, we have now reduced the discussion on genetic structure to focus more on sexual systems and sex determination. Figures S4 and S5 were already in the supplementary material. We kept Fig. 3 and table 1 in the main text because they relate directly to sex-linked patterns. Figure 3 is essential to give an overview of the effect of sex on genetic variation and how that differs between species, and table 1 is purely about sex-linkage (e.g. the F_{ST} in this table refers to the amount of genetic variance explained by sex, not geography).

Hopefully these are constructive, as they are intended to be! And I look forward to seeing the follow-ups in A. kojimai, there is clearly lots of interesting stuff to be found there!

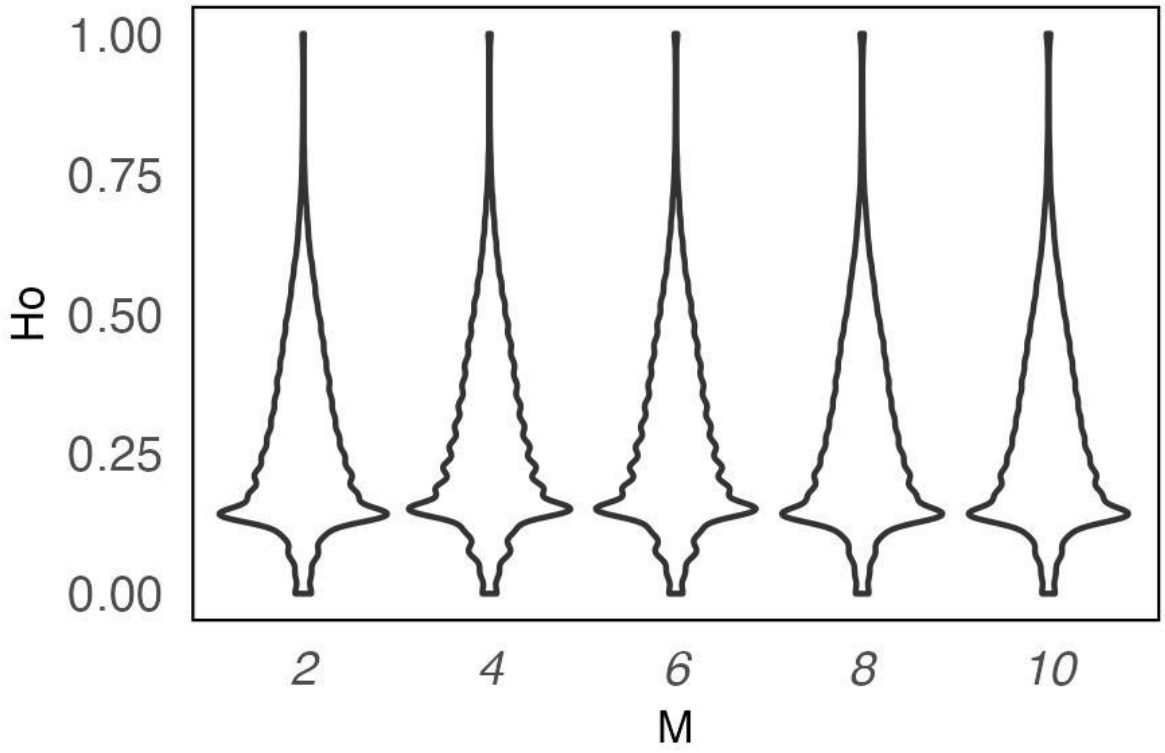
Best Wishes

Daniel Jeffries

Specific comments to authors

L. 156 (and supp. mat.) - The rationale for the high M value makes sense, but I am concerned that this will lead to over-merging loci. To address this concern, perhaps the authors could include a plot of M value vs heterozygosity (which increases when overmerging) in a subset of test samples.

We have now computed the heterozygosity at each locus for five of the test datasets (M=2, 4, 6, 8, 10) used to identify the best Stacks parameters as described in the supplementary material. We did that only for *A. kojimai* because it is the species for which we used the highest M value. Although these datasets are not perfect because they include a number of replicates (which will artificially inflate heterozygosity), they show that M had no effect on heterozygosity in the range of values considered.



Also it would be good to see the test results for M=10 for A. kojami given that this is the value you chose for this species.

The test results for M=10 are now provided in supplementary material Figs. S2 panels C and D.

L. 162 - Could you give some coverage statistics for the final dataset? And again a plot of coverage vs heterozygosity would be useful to tell whether coverage is high enough for accurate SNP calling.

We now provide a new table (Table S2) showing the mean depth per SNP in each species (between 24.8x and 27.5x) and the 2.5% and 97.5% quantiles to show the range of depths in which 95% of the SNPs lie (between 17.2x and 40.6x depending on the species). We also show the whole distribution of mean depth per SNP in new figure S4, and observed heterozygosity as a function of depth in new figure S5. This is all in the supplementary material, and referred to in the main text on L. 175.

L. 170 - Were these pop gene analyses performed after removing the sex-linked loci? Although there are not many, there may be few loci overall in the dataset that are informative for some of these calculations, thus a few loci may have an effect.

The purpose of these calculations is precisely to assess whether sex has a visible effect in global statistics (over all loci), looking at females, males and differentiation between sexes. We thought this was interesting because it shows the extent to which the PCA patterns are supported by quantitative statistics in terms of the effect of sex on genetic structure prior to the identification of sex-linked loci. These statistics are also here to allow us to see the difference between species in the effect of sex. It shows how much of the effect of sex-linked loci is visible, while still being buried in the bulk of the genetic data.

L. 196 - "heterogamous" should be "heterogametic"

Changed L. 212

L. 204. If I understand correctly, here you are simulating a dataset in order to estimate the expected variance around Fst and Fis for their dataset. I like this approach, and agree that the biases in heterozygosity could certainly help find loci proximate to, but not within the completely sex linked region. My one thought here is that the results of sampling from such a simulated dataset could depend on the distribution of simulated allele frequencies. Did the authors consider simulating this dataset based on the site frequency spectrum of the real dataset?

Yes, at first, we did exactly that. But then with a finite number of simulations the variance around the rare events (i.e., the allelic frequency values that are not very common in the spectrum) would be poorly estimated. We thought it is important that we see accurately the effect of sample size for any kind of SNP. The philosophy of our simulations is thus to simulate a large number of all possible outcomes for any allelic frequency, so that we eliminate the risk of false positives when identifying SNPs that present F_{IS}/F_{ST} values outside that expected for autosomal segregation. We think it is better done by sampling from a uniform distribution. This comes at the cost of increased conservatism (more false negatives), but we prefer to miss some sex-linked loci rather than the opposite.

L. 240 - could you give a rough estimate of the distance to this reference if it is known? Timetree.org is good for this.

We lack data to estimate the divergence between our species and *A. marisindica* (no data available for Timetree.org). The best data available comes from variation on CO1 and a few nuclear genes (Breusing *et al.*, 2020), but no quantitative data are available. Perhaps an indication of the pertinence of using *A. marisindica* as a reference comes from the result of the mapping of sex-linked RAD loci on

this genome: 14/14 in *A. kojimai*, 20/64 in *A. strummeri*, and 259/373 in *A. boucheti*, that is, at least 30% of the RADs mapped with a coverage of at least 95% and an e-value threshold of 10^{-5}.

L. 245 - I couldn't find it in the Stacks methods descriptions - did you assemble the 1st and 2nd end reads or treat each end as a separate RADtag? Having ~300bp fragments to map will greatly improve alignment rates to distant genomes.

Yes, we used paired-end reads (mentioned L. 166), assembled to produce longer RADtags. We have now added "paired" L. 167 to be more specific.

L. 286 - small typo - in the figure your oocyte abbreviation is "oo" but in the legend you use "ovo"

Now corrected on L. 351

L. 313 - when you say the anomalous sample in the A. kojami PCA is "not shown" in fig 3, does this mean you removed it and re-ran the PCA, or just removed before plotting the PCA? The former would be better. Also it looks like you removed 2 samples, not just one here (the blue Manus male in the top right as well). Same for Fig. S5. And do you have any idea why these samples are anomalous? You might check coverage for these samples, this is often the culprit.

As explained in our responses to Editor and ref #1 above, we have now re-analyzed the data using a single SNP per RAD and removing outlier samples. These outliers have now been entirely removed from the analyses (although it clearly improved the PCA, removing outliers from downstream analyses had no impact on further results, most importantly the identification of sex-linked scaffolds).

L. 348 - at the risk of seeming too self promoting, perhaps you would find our recent analyses useful here (Jeffries, D. L., Mee, J. A., & Peichel, C. L. (2022). Identification of a candidate sex determination gene in Culaea inconstans suggests convergent recruitment of an Amh duplicate in two lineages of stickleback. Journal of Evolutionary Biology. https://doi.org/10.1111/jeb.14034). Specifically we used some simple probability calculations to assign a p-value to each locus, using exactly the same heterozygosity analyses that you do here. I think that would be useful, especially in A. strummeri & A. kojami.

We have tested this suggestion. We calculated the probability of obtaining exactly the observed number of female heterozygotes (H_f) and male heterozygotes (H_m) for each locus given the number of female (N_f) and male individuals (N_m) sampled as:

$$p = \frac{\binom{N_f}{H_f}\binom{N_m}{H_m}}{\binom{N}{H}}$$

where $N = N_f + N_m$ and $H = H_f + H_m$

This follows Jeffries et al. (2022, see in particular supplementary file JN_03), although there may be a typo in the equation presented in the main text of this article (where the binomial coefficients seem to appear as fractions (?)).

We found that the vast majority of the SNPs identified as sex-linked with our method had low probability *p*, as expected (e.g. in *A. boucheti*, 995 out of 1011 sex-linked SNPs had Jeffries'$p < 0.05$).

However, we also pointed two caveats. First, sex-specific observed heterozygosities are less informative than F_{ST} combined with F_{IS}. For instance, a SNP that sits in a region with low recombination may have identical Hom and Hof but for different allelic frequencies in males vs females (a difference that will be captured by F_{ST} between sexes). This is what explains that some of our sex-linked SNPs had high Jeffries'*p* (up to 0.17). Second, as correctly pointed by the authors of this method, it is difficult to decide of a p-value threshold that would allow us to decide whether a SNP is sex-linked or not.

On the other hand, observing Jeffries'*p* for some of our data (in particular the cases where Jeffries'*p* was very low despite our labelling of SNPs as unlinked with sex) made us realize that our simulations are quite conservative because they do not take into account the likelihood of each simulation output. That is, we use a hard threshold (if a situation happens once in the simulations, then we deem it possible for an autosome regardless of the likelihood of this happening). We think that this is conservative especially in the cases of strong male F_{IS} (neutral autosomal segregation allows F_{IS} to reach 1 in some cases). Because our analysis does not use the comparison of F_{IS} in both sexes, we may discard X-specific SNPs that were pointed out by Jeffries'*p*. But again, there is no easy way around the p-value issue, unless building a more sophisticated statistical test that would take into account all parameters. We think a thorough comparison of methods will bring us too far from the objectives of this paper, but we did mention that our simulation method is conservative (L. 573).

L. 377 - Could you add here how much overlap there was between the loci identified by your simulation-based approach and the heterozygosity approach?

No, we can't really do that because the heterozygosity approach was not associated with any statistical procedure (and when we tried as explained in the previous comment, we could not find a good way to set up a p-value threshold to identify sex-linked loci purely from observed heterozygosity and

compare the methods quantitatively). Our thinking with the F_{ST}/F_{IS} approach is that F_{IS} (shown for both males and females) captures the excess/deficit in Ho in both sexes, so that we don't really need to go back to the Hom/Hof approach (used only to identify the sex determination system). Once we have identified the sex determination system, we prefer to move on to what we think is a more powerful approach to identify locus sex-linkage.

L. 411 - Did you specifically look at how the species-specific sex-linked loci look in species where they are not sex linked? For example, locating the 348 markers that are sex linked in A. boucheti on the A. strummeri plot in Fig. 5 would perhaps give some indication as to whether you missed them due to power issues, or that they are truly autosomal in this species.

No, we cannot do that, because sex-linked loci can show F_{ST}/F_{IS} patterns anywhere in the "autosomal parameter region" very easily with the low sample sizes of *A. strummeri* (note that the reverse is not true, fortunately). Also, as explained in the next comment's response, it is difficult to relate individual SNPs from RAD catalogs that were built independently.

Also, if I understand correctly, you are talking in terms of RAD loci here. So you are able to say that the same genomic regions are sex linked among samples. But are we looking at the same SNPs within these tags? I.e. is it shared ancestral polymorphism that you are seeing, or independently derived sex linked SNPs within the same genomic region. Knowing this might help you narrow down which scaffold/region contains the sex determination locus.

This is an interesting point, but almost impossible to assess with RAD data (where one catalog per species was built), as it would require aligning each pair of identical RADs between species individually, and checking the alignment of each SNP (they are named according to their position along the RAD, but this can differ between species due to indels).

However, we were able to perform this analysis for the three RAD-seq loci that were sequenced and sex-linked in the three species. As detailed in previous responses above, this analysis provided interesting results now presented in the manuscript.

L. 436 - First, this statement is not accurate, as it is only true for the sex chromosome. Instead the statement should read something like "In PCA analyses, signals of sex linkage within the genome were stronger than signals of geographic isolation".

The discussion was largely rewritten, and this sentence was removed.

L. 456 - sex chromosomes are "sufficiently differentiated", or population structure is sufficiently weak.

Idem

L. 469/470 - or that population structure is stronger.

Idem

L. 496-505 - please avoid repeating results in the discussion

We have now completely removed such repetitions from the discussion, except here for *A. kojimai* for this particularly important observation, which we want to emphasize (now L. 547-553).

L. 511,512 - I'm not sure how much we can infer about the recombination on the sex chromosome in A. kojimai. Given the propensity for intersex individuals, the lack of sex linked identified could just be linked to errors in assigning samples to male or female/hermaphrodites, which would greatly reduce power to identify sex linked loci. Again here it would be good to see how the loci identified in A. boucheti behave in this species. You might even try clustering A. kojimai samples based on the A. boucheti loci.

It is correct that mixing males and females/hermaphrodites would limit the power of the analyses quite dramatically. But we have no problem identifying males vs females/hermaphrodites, and the intersex individuals clearly form a homogeneous group with females (see argument on L. 553-558), and cannot be taken for males or vice-versa.

We do have a problem with the identification of females (we are not certain that there are any, because all might conceal some male tissue somewhere in their gonads; in fact, it would be quite surprising if this species really was trioecious, which we don't know so far), but not with males. Repeating all analyses using only the individuals identified through histological analyses (which are reliable to identify males) proved this quite convincingly (as shown in sup mat). Also, with *A. kojimai* we have enough power to detect sex-linked loci, so there is no question that the size of the sex-linked region in *A. kojimai* is far smaller than that of *A. boucheti*. Now, we agree that it will take more than RADseq to quantify this precisely, as we mention clearly in the manuscript (L. 571).

L. 513-516 - I agree with your conclusion here, that some of these loci could represent polymorphism between Y chromosomes in the population. But keep in mind (and maybe mention) that the most common genotyping error is missing heterozygous calls, which would also give the same pattern.

This section has been removed from the discussion

L. 530 - As I mentioned for the results section, I think a more detailed comparison of the Fst/Fis analyses against the heterozygosity analysis would be useful here.

See responses to this comment above

L. 549 - I think "unique" should be "single" here

Changed accordingly

L. 552 - do those 3 loci co-localise on a scaffold?

These three loci are found on two of the five scaffolds that are sex-linked in all three species (now mentioned L. 511).