A bird's white-eye view on avian sex chromosome evolution

Thibault Leroy, Yoann Anselmetti, Marie-Ka Tilak, Sèverine Bérard, Laura Csukonyi, Maëva Gabrielli, Céline Scornavacca, Borja Milá, Christophe Thébaud, Benoit Nabholz

https://www.biorxiv.org/content/biorxiv/early/2019/01/23/505610.full.pdf version January 23, 2019

Submitted by Thibault Leroy 2019-01-24 14:17

Abstract

Chromosomal organization is relatively stable among avian species, especially with regards to sex chromosomes. Members of the large Sylvioidea clade however have a pair of neo-sex chromosomes which is unique to this clade and originate from a parallel translocation of a region of the ancestral 4A chromosome on both W and Z chromosomes. Here, we took advantage of this unusual event to study the early stages of sex chromosome evolution. To do so, we sequenced a female (ZW) of two Sylvioidea species, a Zosterops borbonicus and a Z. pallidus. Then, we organized the Z. borbonicus scaffolds along chromosomes and annotated genes. Molecular phylogenetic dating under various methods and calibration sets confidently confirmed the recent diversification of the genus Zosterops (1-3.5 million years ago), thus representing one of the most exceptional rates of diversification among vertebrates. We then combined genomic coverage comparisons of five males and seven females, and homology with the zebra finch genome (Taeniopygia guttata) to identify sex chromosome scaffolds, as well as the candidate chromosome breakpoints for the two translocation events. We observed reduced levels of within-species diversity in both translocated regions and, as expected, even more so on the neoW chromosome. In order to compare the rates of molecular evolution in genomic regions of the autosomal-to-sex transitions, we then estimated the ratios of non-synonymous to synonymous polymorphisms ($\pi N/\pi S$) and substitutions (dN/dS). Based on both ratios, no or little contrast between autosomal and Z genes was observed, thus representing a very different outcome than the higher ratios observed at the neoW genes. In addition, we report significant changes in base composition content for translocated regions on the W and Z chromosomes and a large accumulation of transposable elements (TE) on the newly W region. Our results revealed contrasted signals of molecular evolution changes associated to these autosome-to-sex transitions, with congruent signals of a W chromosome degeneration yet a surprisingly weak support for a fast-Z effect.

Keywords: Sex chromosome, molecular evolution, molecular dating, bird diversification, Sylvioidae, Zosterops

Round #1

Author's Reply:

Decision

by Kateryna Makova, 2019-03-18 22:24

Manuscript: https://www.biorxiv.org/content/biorxiv/early/2019/01/23/505610.full.pdf

Please revise the manuscript according to reviewers' comments

Additional requirements of the managing board:

As indicated in the 'How does it work?' section and in the code of conduct, please make sure that:

-Data are available to readers, either in the text or through an open data repository such as Zenodo (free), Dryad (to pay) or some other institutional repository. Data must be reusable, thus metadata or accompanying text must carefully describe the data.

-Details on quantitative analyses (e.g., data treatment and statistical scripts in R, bioinformatic pipeline scripts, etc.) and details concerning simulations (scripts, codes) are available to readers in the text, as appendices, or through an open data repository, such as Zenodo, Dryad or some other institutional repository. The scripts or codes must be carefully described so that they can be reused.

-Details on experimental procedures are available to readers in the text or as appendices.

-Authors have no financial conflict of interest relating to the article. The article must contain a "Conflict of interest disclosure" paragraph before the reference section containing this sentence: "The authors of this preprint declare that they have no financial conflict of interest with the content of this article." If appropriate, this disclosure may be completed by a sentence indicating that some of the authors are PCI recommenders: "XXX is one of the PCI Evol Biol recommenders."

Reviews

Reviewed by Melissa Wilson, 2019-01-30 23:17

Overall, I think the topic of the manuscript is very interesting, and is why I agreed to review. For now, however, I am cautious with any interpretations. I am waiting to evaluate the rest of the manuscript until several questions about the methodology are answered. Some are strait-forward, but just useful to have in the same place, and for clarity. Others are essential for interpreting measurements of diversity and the transcriptome analyses.

We thank Melissa Wilson for her feedback. Given that the materials and methods section is located at the end of our manuscript, some information was not explicitly stated. In this revised version, we tried to fill this gap by giving more details in order to generate a more standalone section.

Methods questions 1. The authors start by describing sampling from one female Z. borbonicus and extracting DNA from tissue - which tissue?

We used fresh tissues to generate both whole-genome sequences. It is important to note that the two bird species are very light (10 grams or less). As a consequence, there is a tradeoff between the tissue specificity and the quantity of DNA per sample. Given the objectives of the present study, we made a conscious choice of privileging quantity of DNA, since we had no objectives to study between-tissue variation. As a consequence, for the *Z. borbonicus* individual as well as for the *Z. pallidus* one, we extracted fresh tissues from their bodies, including muscles, but not exclusively.

Which tissue did the authors collect DNA from form the Zosterops pallidus? Please specify the sex of the individual in the methods, currently it just says, "individual".
The sequenced *Z. pallidus* individual is also a female. We sequenced tissues from fresh tissue, here muscle. We made this explicit (see lines 1194-1197)

1. The authors refer to liver and/or muscle - did they do multiple extractions for each individual? Given there are only two birds, please be specific about what was collected, how many extractions were done, whether these extractions were pooled, and what exactly (tissue sample and species) was sequenced.

   For each individual, we made a single extraction. We used fresh tissues (mostly muscles) to generate a whole-genome sequence. Here, our objective was only to obtain a reference genome sequence for the species. Some tissue-specific somatic mutations, albeit extremely rare, are possible and are not investigated in this study.

2. Were the samples multiplexed? Or were each run on an individual lane? What sequencing depth was conducted for each of the samples?

   Each individual were independently sequenced on a single lane. For *Z. borbonicus*, the sequencing depth are 70X for pair end sequencing (~62X after read trimming), 12X for mate pair and 11X for PacBio (see Fig. 1). For *Z. pallidus*, the sequencing depth are 85X for paired-end (~72X after trimming) and 10X for mate-pair. The sequencing depth of *Z. pallidus* is now specified in the main text (l. 216-219).

3. For the 10X PacBio sequencing coverage, how many SMRT cells were used (or what was the technology) to obtain the 10X coverage? Is this a theoretical 10X coverage? How much was the coverage when aligned to the reference genome, and how did it vary on the Z and autosomes?

   We used 18 SMRT cells from a PacBio RS technology (now available in SRA : SRR8897337 – SRR8897346). These cells produced 13.8 Gb of sequencing data, thus leading to an expected coverage of 11X (13.8 Gb / 1.222 Gb = 11.3; see Fig. 1) . We agree that this remains theoretical. We can reasonably suspect that the PacBio reads will fellow the general rule with almost half of the relative coverage on the Z of as compared to the autosomes. However, such an evidence

would remain questionable given this relatively low coverage, that's why we have not even tried to map our PacBio reads against the reference genome.

4. For Z. barbonicus, describing the RNA extraction, the authors say they extracted RNA then stabilized it. I was under the impression RNA would be stabilized in a solution (like RNAlater), then extracted later. If it really was extracted immediately in the field, that is worth noting. Else, can you describe how the tissues were collected for RNA isolation - how long was the individual dead? Did you have RNAlater in the field? Or did you bring the individual to the lab to dissect?

=> Before to answer to all questions concerning our RNAseq project, it seems very important to more clearly explain our objective. We generated sets of transcript to train the gene prediction software to find genes and annotate our genome, as it is usually done for *de novo* sequence projects. We neither intended to report expression levels for the two tissues, nor to generate an exhaustive transcript catalogue.

After the accidental death of the *Z. borbonicus* individual during a field campaign, a dissection were performed in the field to remove two organs (brain and kidney). These tissues were stabilized using RNAlater solution. Samples were brought back and stored in a freezer. RNA were extracted several months later but only the brain sample was used for RNA library and sequencing that were performed soon after extraction (< 2 weeks).

5. Did you get RIN values for each of the samples before sequencing? If so, please report them.

=> Only the RNA brain was sequenced. We made this much more explicit by no longer mentioning that we had also extracted RNA from kidney. The RIN value for this RNA extraction was 7.9. We now indicate this value in the manuscript.

6. What kind of RNA library prep was performed? Did the authors do ribosomal RNA (rRNA) depletion, or was total RNA sequenced? If total RNA, what kind of effect does this have on the ability to detect transcripts and what proportion of transcripts were assembled/aligned to ribosomal sequences (and how many thus were used for annotation)?

=> The RNA library is a Truseq stranded total RNA. Expression levels, and similar aspects, are not covered in the present study. Again, our objective was to generate a set of transcripts to train the gene prediction software.

7. How many millions of reads were sequenced for each sample? The authors refer to (1 line) of Hiseq2500 sequencing. I think they mean, (1 lane)? If so, how many predicted reads per sample, and how were the samples multiplexed (or were they)?

=> We are sorry for the line/lane typo, this is now rectified (l. 1227). We generated 261,702,637 reads corresponding to 65Gb of data (run accession: SRR8887131, 60.6G after read trimming).

8. Can the authors comment on if (and then if so) why the two RNA extracts were pooled prior to sequencing rather than sequencing them separately and combining the fastq files afterwards? Each tissue is presumed to have very different genes expressed in them, and even different transcripts.

=> It has to be clarified that the objective of the study is not to study the different expression levels. We performed a RNAseq experiment to generate a sets of transcript to train the gene prediction software. As a consequence, our objective was neither to generate a tissue-specific gene expression profile nor to perform a transcriptome analysis. These analyses are almost incompletely insensitive to sequencing errors and can be non-exhaustive, but are rather motivated by the idea of helping the software to do a better job by identifying specific signatures of genes based on the *de novo* transcript catalogue, in order to be more efficient in the scanning step (scan of the whole reference genome sequence to identify candidate genes). In addition, we do not pretend to give a complete list of the *Z. borbonicus* genes, since it remains a complex task, especially for avian genomes (e.g. Botero-Castro *et al.* 2017 MBE).

9. Will the authors please provide the statistics and fastqc reports for the data pre- and post-trimming (or at least post-trimming) for the DNA and RNA sequence samples?

=> We performed runs of fastQC before and after trimming. For *Zosterops borbonicus*, the single "red flag" initially concerned the Kmer content and were removed after trimming (green flag). Whiskers for all per base quality variation (boxplot) along the read length are greater than 26 (>22 before trimming), suggesting a reliable quality sequencing. Similarly, quality was very high for the RNA data too (> 24 after trimming). All these reports are now publicly available on Figshare : https://doi.org/10.6084/m9.figshare.7999514.

10. For the DNA, the quality score threshold used was 15 (meaning somewhere between 1/10 and 1/100 of the bases in the DNA are likely to be errors). For the RNA, the quality score threshold used was just 5, meaning - as I understand PHRED scores - more than 1/10 base pairs in the RNA are expected to be errors. Perhaps this was because the authors only wanted to use transcripts to identify transcribed regions, but this low of a PHRED quality score is concerning, as it is a probability that nearly a third of base positions are incorrect in the RNA data. I would like to see more assessment of the quality of the RNAseq, and the effects of using such low quality RNA in assembling a transcriptome. Presumably there is a lot of degraded RNA, and this will make both transcriptome assembly, and gene expression analyses quite difficult. But, perhaps the authors have evidence that such a low PHRED threshold for RNAseq data still yields trustworthy transcriptome annotation?

=> For the mapping step and our DNAseq experiment, we used a minimum quality score of 20 (the most commonly used threshold in similar studies). Based on our fastQC report, these errors are expected to be rare on trimmed reads. In addition, such sequencing errors only represent a single allele among all the reads mapped at this position. As a consequence, the mapper is expected to be conservative and to generate a heterozygous call very infrequently. To be called, it would require

several errors at the same location to pass all our vcf filtering parameters. Such a situation are expected to be rare.

11. The authors use mean depth to identify autosomal, Z, and W-linked scaffolds, but this should not be the only evaluation, as depth in sequencing experiments is not evenly distributed across the genome, and especially in small scaffolds can confound estimates of Z, and W linkage. At least one additional measure (heterozygosity for example), should be applied to further support the Z and W linkage. Also, have the authors considered the effects of Z-W homology? In an effort to be conservative, it seems like they are identifying regions that are highly differentiated between Z and W, and as such, this should be acknowledged.

=> Indeed, it is true that the coverage is not evenly distributed. To take this into account this bias, we computed the median per-site coverage over all scaffolds but, to ensure reliable results, we only considered scaffolds with a minimum size of 4 kb. This estimate is expected to be robust to per-site coverage variation (see also Smeds et al. 2015 nat. comms). In addition, coverage is not used in absolute terms, but relatively between males and females. Although highly variable, the median coverage per scaffold is highly correlated between males and females. When the males coverage is plotted against the females coverage, the differences between autosomal, Z and W-linked scaffold become very clear, with the notable exception of scaffolds with very low coverage (<2X, see figure 2). Regarding heterozygosity, it is expected to poorly perform here, particularly on the W chromosomes, given both the very low diversity observed on the W chromosome, the small sizes of the W scaffolds (based on the 3 female individuals we investigated in our study, many scaffolds have not even a single polymorphism) and the highly repeated content of the W chromosome (see also correspondence with the Reviewer #3, remark #8).

12. When the authors describe measuring genome-wide genetic diversity, the description of how those BAM files were produced is not included, but should be (albeit briefly) here. If the authors used the same trimming and quality thresholds for the DNA used in these analyses, I would be cautious about interpreting the patterns of diversity. Further, are these BAM files from DNA or RNA, it isn't specified. If they are from RNA, there are additional challenges. If they are from DNA, it is helpful to know what depth, how it varies on Z and autosomes, and how this coverage will affect variant calling.

=> Again, we used a quality threshold of 20 (see also comment #10). In addition, all our scripts are freely available on figshare (https://figshare.com/s/122efbec2e3632188674 "*Scripts and programs to estimates nucleotide diversity in the Zosterops borbonicus genome*"), as well as all our raw data on SRA (BioProject PRJNA530916) to make possible additional investigations or to facilitate the ability to track potential errors. Our estimate (nucleotide diversity Pi = 0.0018) is well in line with other estimates obtained independently on other passerines such as *Ficedula* flycatchers, Darwin finches and *Saxicola* stonechats (Pi = 0.001 - 0.004; Van Doren et al. 2017. Mol. Ecol. 26:3982–3997; Vijay et al. 2017 Mol. Ecol. 26:4284–4295).

Reviewed by Gabriel Marais, 2019-01-30 23:18

Leroy et al.'s ms is focused on a neo-sex chromosome system in birds. They have sequenced 2 bird genomes and identified the scaffolds from the sex and neo-sex chromosomes. They then performed a thorough characterization of this system using comparative genomics, population genetics, molecular evolution and phylogenetics approaches. They found that the neo-sex chromosomes are recent (1-3.5 My old). They originated from a partial autosomal fusion of which they identified precisely the breakpoints using an outgroup bird genome. The neo-sex chromosomes harbour hallmarks of early sex chromosome evolution: reduced diversity (on both neo-Z and neo-W, stronger on the latter), higher dN/dS on the neo-W genes and accumulation of TEs on the neo-W.

In vertebrates, most of the studies on sex chromosomes have focused on old systems such as the mammalian XY and the avian ZW chromosomes. Very little is known on the early stages of sex chromosome evolution. In invertebrates, the study of neo-sex chromosomes (in drosophila, see for example papers from D. Bachtrog and AB Carvalho groups) has provided many key observations to understand the early steps of sex chromosome evolution and Y degeneration. To my knowledge, this is the first ms doing similar work on a vertebrate neo-sex chromosome system. Although no completely unexpected patterns have come out of this study, it filled an important gap in the literature. We have data on a vertebrate neo-sex chromosome system showing that neo-W ongoing degeneration resembles what has been observed in neo-Y systems in drosophila.

Another positive point about this ms is the amount of data and the number of analyses that have been done. Two genomes have been sequenced, one of which is of very high quality. Many analyses have been done (reconstructing the autosome-sex chromosome fusion, identifying the sex and neo-sex scaffolds, dating the system, computing the genetic diversity, looking at the dN/dS, the TE content, the GC content, etc…). It is a lot of data and findings in a single paper. For most analyses, the results are very neat (e.g figures 2, 4, 5, 8, 9). In other systems, several papers have been published to report a similar work. On the other hand, the ms reads very well and is clear and interesting all along.

I am thus very positive about this ms and I think it would merit a PCI evol biol recommendation. I only have minor comments/suggestions, which are listed below.

=> Before to do a point-by-point response to the Gabriel Marais's comments, we want to thank him for his time and the positive feedback. We also want to make clearer that the neo-sex chromosome formation is older than reported in his review. These two translocation events are not only tractable in *Zosterops* but also in all present-day Sylvioidea (including larks, swallows and warblers dating to a common ancestor dating ~ 15-30 Myrs ago, see also Dierickx et al. 2019 & Sigeman et al. 2019 for two very recent and interesting reports). In Fig 3, the neo-sex chromosome formation occurred in the ancestral branch of *Zosterops* and *Phylloscopus*.

1) I think the title could be improved. At present it does not tell a lot about what's is in the ms. I think a title such as "The young neo-sex chromosome system in Zosterops provides insight on the early steps of

ZW chromosome evolution in birds" or something similar (including the key words neo-sex chromosome and early steps of ZW chromosome evolution or ongoing W degeneration) should be preferred.

As compared to the original sex chromosome, we agree that this event occurred quite recently. But, we preferred not to use "early steps", since it can be quite confusing regarding the timing of this event. Two new studies reported some additional and more recent translocation events in some other Sylvioidea species (Dierickx et al. 2019 & Sigeman et al. 2019, two preprints publicly shared over the last month). Our previous title was: "A bird's white-eye view on avian sex chromosome evolution". We have now changed to "A bird's white-eye view on neo-sex chromosome evolution" to include the keyword "neo-sex".

2) In the literature, the term neo-sex chromosome is restricted to the part that has been fused to the original sex chromosomes (see for example Bachtrog 2013 figure 2). In Leroy et al. they call the neo-sex chromosomes the fused autosomal part + the original sex chromosomes (see figure 4 and throughout the ms). I think this needs to be either corrected or at least acknowledged, to avoid readers getting confused.

=> On the one hand, newly translated regions are expected to exhibit very clear signals - much more than the original sex chromosome regions - of the autosome-to-sex translocation, with considerably high impacts on its diversity, local recombination rates and selection footprints. On the other hand, the chromosomal context also changed for the original sex chromosome, with a 10 Mb increase in chromosome size (~13% increase in size for the Z; >150% for the W?) probably inducing a change in recombination rate (Z) and potentially impacting the efficacy of positive and background selections (W or Z). That's why we consider that it is better to consider the autosomal part and the original sex chromosome as the neo-sex chromosome. We made this explicit (see lines 80-85).

3) The authors may want to cite work done on plant neo-sex chromosomes (and speciation) such as Weingartner & Delph 2014.

=> We added this reference to the list.

4) Perhaps, it would be interesting to speculate on why such an A-ZW fusion has taken place. Is there something special about Zosterops, in terms for example of sexual dimorphism (one possible driver of neo-sex chromosome evolution, see Kitano et al. 2009).

=> This is an interesting point. As already pointed out, the fusion predates the *Zosterops*, and occurred in the common ancestor of all Sylvioidea species. From a phenotypic point of view, the study of the sexual dimorphism requires to reconstruct the ancestral state of this branch. For the genetic point of view, one way to tackle this question would be test whether the translocated region contains genes with sex-biased expression and sex-specific genes suggesting the existence of sexual antagonistic genes. Both analyses are beyond the scope of this study. We only have, for example, one RNA-seq sample that was used for gene annotation purpose. Moreover, the group of Bengt Hansson is currently investigating this question and has already found evidence that the translocated regions of neo-sex chromosome are

enriched in sex-specific genes (see for example Sigeman et al. 2018 doi:10.3390/genes9080373 and Sigeman et al. 2019 bioRxiv doi.org/10.1101/585059).

References

Bachtrog D. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. Nat Rev Genet. 2013 Feb;14(2):113-24.

Weingartner LA, Delph LF. Neo-sex chromosome inheritance across species in Silene hybrids. J Evol Biol. 2014 Jul;27(7):1491-9.

Kitano J, Ross JA, Mori S, Kume M, Jones FC, Chan YF, Absher DM, Grimwood J, Schmutz J, Myers RM, Kingsley DM, Peichel CL. A role for a neo-sex chromosome in stickleback speciation. Nature. 2009 Oct 22;461(7267):1079-83.

Reviewed by anonymous reviewer, 2019-01-30 23:19

The manuscript “A bird's white-eye view on avian sex chromosome evolution” by Leroy et al. reports a high-quality genome assembly of the “great speciator”, the songbird genus Zosterops. This group of species may have the fastest rate of speciation after East African cichlids. The authors use the chromosome-level information of this genome assembly and additional Zosterops species plus outgroups to convincingly infer a) the timing of Zosterops diversification, b) the evolution of remarkable neo-sex chromosome of sylvioid songbirds through fusion of the Z and W chromosomes with half of chromosome 4A, c) the molecular evolution of genes since neo-sex chromosome emergence, and d) the accumulation of transposable elements (TEs) since neo-sex chromosome emergence.

My impression is that this manuscript is a well-written and comprehensive genome analysis of an interesting study system. The methods appear sound and the main conclusions are well supported by the data. I also commend the authors for their attention to detail in the methods section, and for providing all scripts and datasets on figshare. However, I have some suggestions for how to further improve the manuscript. Most of these are related to clarification of specific statements or methods as a service to the reader.

We want to thank the anonymous reviewer for his detailed reading, interesting and meaningful comments and positive feedback.

Main comments:

1. Throughout the text: I wonder why the authors use the term “translocation” instead of “fusion” for the mechanism leading to a neo-sex chromosome? Please explain this choice at the first mentioning of the Sylvioidea neo-sex chromosome system, or replace with the term “fusion”. As far as I am aware of, other papers on this system usually use the term “fusion”.

=> The difference between the two terminologies are quite subtle. In general, “fusions” refers to the merging of the two complete chromosomes (as the exact opposite of chromosomal fissions, e.g. Fertin et al. 2009 in “Combinatorics of genome rearrangements”) and therefore induces a change in the number of chromosomes. Here, only a part of the 4A were translocated (and fused) on the W and Z chromosomes without inducing a change in chromosome number. That’s why, the term “translocation” was preferred all along the manuscript.

2. Throughout the text/figures: Please always add the term “contig” or “scaffold” in context of N50. Scaffold N50 and contig N50 are quite different metrics and should therefore be explicitly mentioned.

=> We agree and apologise for the confusion. All our reports were scaffold N50. We made this explicit throughout the manuscript.

3. Throughout the text: The authors mention a megabase-scale deletion on the neoW-4A gametolog. How many genes are located in this region in the neoZ-4A gametologous region? It might be interesting to state the number of genes lost in this deletion event, highlighting that a single large intrachromosomal rearrangement can lead to significant differences between Z/W gametologous regions.

=> This is an interesting point. In the *T. guttata* reference genome, the 4A chromosome contains 332 genes, including 142 genes on the region corresponding to the Sylvioidea neo-sex translocations (9.6 Mb, *i.e.* 14.8 genes/Mb). Among these 142 genes, only 2 (ENSTGUG00000002395 & ENSTGUG00000002396) were found on the region corresponding to this 1.7 Mb neoW deletion (*i.e.* 1.1 genes/Mb, as compared to 14.8 genes/Mb for the genome space corresponding to the whole translocated region). This result therefore supports that this large deletion occurred in a gene-poor region, and that the gene content of the translocated region was only slightly affected. We added this new result to the main text (lines 337-351).

4. Throughout the text: Strictly spoken, gametologs are not paralogs (those arose through gene duplication) but homologs that arose through sex chromosome differentiation.

=> This is correct. We modified the text accordingly.

5. Lines 171-190, Figure 1, Figure S2: It is not clear to me what the TE content in genes refers to – are these TEs nested inside introns/exons/UTRs or even part of the protein-coding sequence? Figure S2 shows that some genes have a TE content of nearly 100%, so I wonder if these could be either TE-derived genes or gene models incorporating genes? Please clarify.

=> We agree. To guarantee greater clarity, we computed the TE content only on coding regions (see the new Fig. S2) in order to identify TE mistakenly annotated as genes. In addition, we defined three sets of genes, namely genes with a low within-CDS TE content and supported by at least one transcript based on our RNAseq dataset (high_reliability), genes with either a low TE content or non-zero expression levels (moderate_reliability) and all other predicted genes (low_reliability). This is now clearly stated in the Fig. 1 and in the materials and methods section (l. 1366-1375).

6. Lines 226-234, lines 658-667: I commend the authors for improving the assemblies of 25 other bird species with DeCoStar and making these data freely available on figshare. The question is: How informative are these – do the authors have thoughts on whether these assemblies might be biased towards those assemblies with higher quality and potentially underestimating the number of intra/interchromosomal rearrangements? Please briefly comment on these possibilities.

=> The approach implemented in DeCoSTAR and used in this paper for the phylogeny-guided scaffolding of the 48 extant avian genomes (assisted by paired-sequencing data) has been described in Anselmetti et al., 2018 doi:https://doi.org/10.1186/s12864-018-4466-7. Comparing 3 well-assembled Anopheles genomes and some highly fragmented de novo assemblies of the same individuals, we tested the accuracy of the predicted adjacencies of DeCoSTAR. The result highly depends of the phylogenetic placement of a given species. Indeed, DeCoStar performs less well for isolated species that had more time to evolve unique rearrangements. We found that a species with close relatives had better precision (~97%) than species near the root of the species tree or with fewer closely-related species (~92%) . So, the precision is high, but remains variable.

Based on the 19 avian species for which paired-sequencing data were available to guide DeCoSTAR, 28.1 to 83.42% (average = 50.25%) of new adjacencies predicted are supported by paired-reads data (except *Acanthisitta chloris* for which less than 0.5% of new predicted adjacencies are supported by paired-read data, but this genome is also the most fragmented, N50=64kb, between 4 and 258-fold lower, mean=85X). *Ficedulla albicollis* and *Gallus gallus* for which no paired-end reads helped to improve scaffolding (no improving at all for *Gallus gallus*). For these predicted adjacencies, we have a strong support given by the phylogenetic and paired-sequencing signals.

7. Line 409 (Figure 4): a) How are the LASTZ hits ordered on the Y axis? What about regions with hits to other chromosomes – are these translocated regions from autosomes or spurious LASTZ hits due to repeats? b) Is it known where the pseudoautosomal region (PAR) is on the neoZ and neoW chromosomes? If this is the case, please indicate the chromosome ends containing the PAR in this figure.

=> The different hits are randomly sorted just to avoid the overlap. Hits that starts at the same location are likely spurious LASTZ hits that were non-assigned as TE (because LASTZ do not start any alignment in the softmasked regions). Some cleaning would be possible here, but we try to keep this as conservative as possible even if this choice lead to a somewhat blurred picture for the neoW chromosome because of the very high content in TE content. Currently, we do not know where are the PAR.

8. Lines 482-485: Have the authors looked into whether there are “heterozygous” sites in those neoW-W windows with higher nucleotide diversity? Could some ambiguous mapping to repeat regions on neoW-W play a role here?

=> We agree. We observed a quite high rate in our heterozygous calls on the neoW-W, even for the ZZ individuals where it is obviously genotyping errors (see Fig. 1 below). As expected, the high repeat content lead to ambiguous mapping in repeated regions that could lead to mapping reads Z or autosomal reads onto neoW-W chromosomes. Unfortunately, even after excluding SNPs in regions previously detected as repeated (Fig. 2 below), this pattern still holds. This is likely due to the fact that our TE list is probably far from being exhaustive even if we performed a *de novo* TE identification with RepeatModeler. As an alternative, we only took into account the summary statistics of the 10-kb genomic windows computed on the 3 females, for windows where no SNPs were detected for the 3 male individuals. To analyze the W/Y chromosomes, this is probably the best solution to circumvent this high complexity and ensure interpretable results (see lines 1662-1671).

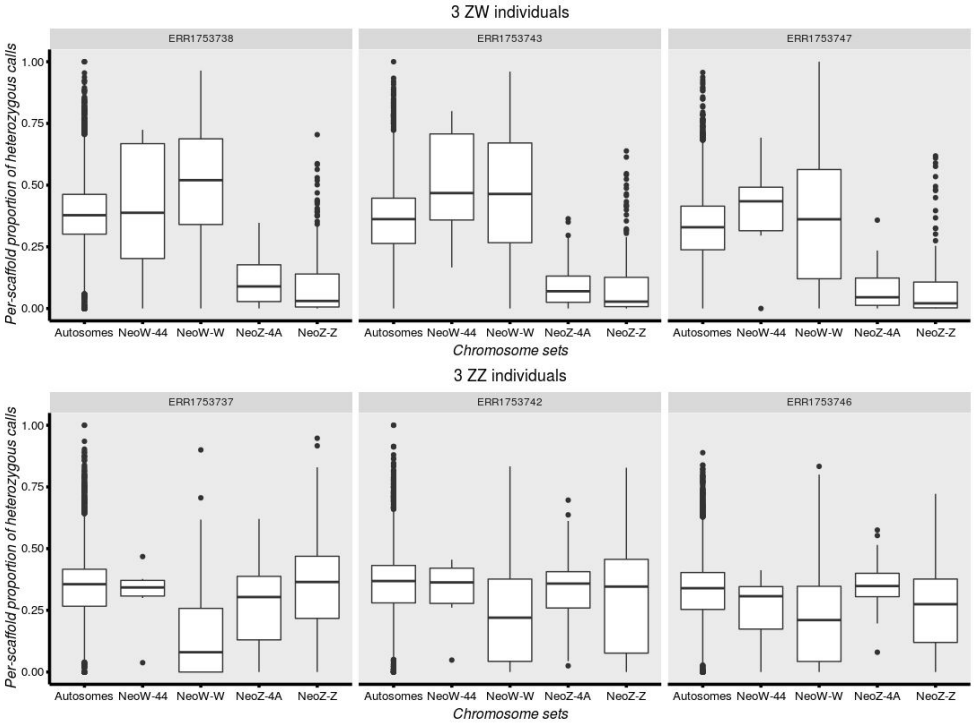

Fig1: Per-scaffold proportion of heterozygous calls at autosomal, neoW and neoZ SNPs for the 3 ZZ and 3ZW individuals. Computations were restricted to scaffolds >10 kb with at least one SNP in one individual (ZW: autosomes 2,175; NeoW-W 134; NeoW-4A 9; NeoZ-Z 308 & neoZ-4A 43 vs. ZZ: autosomes 2,022; NeoW-W 94; NeoW-4A 7; NeoZ-Z 313 & neoZ-4A 43).

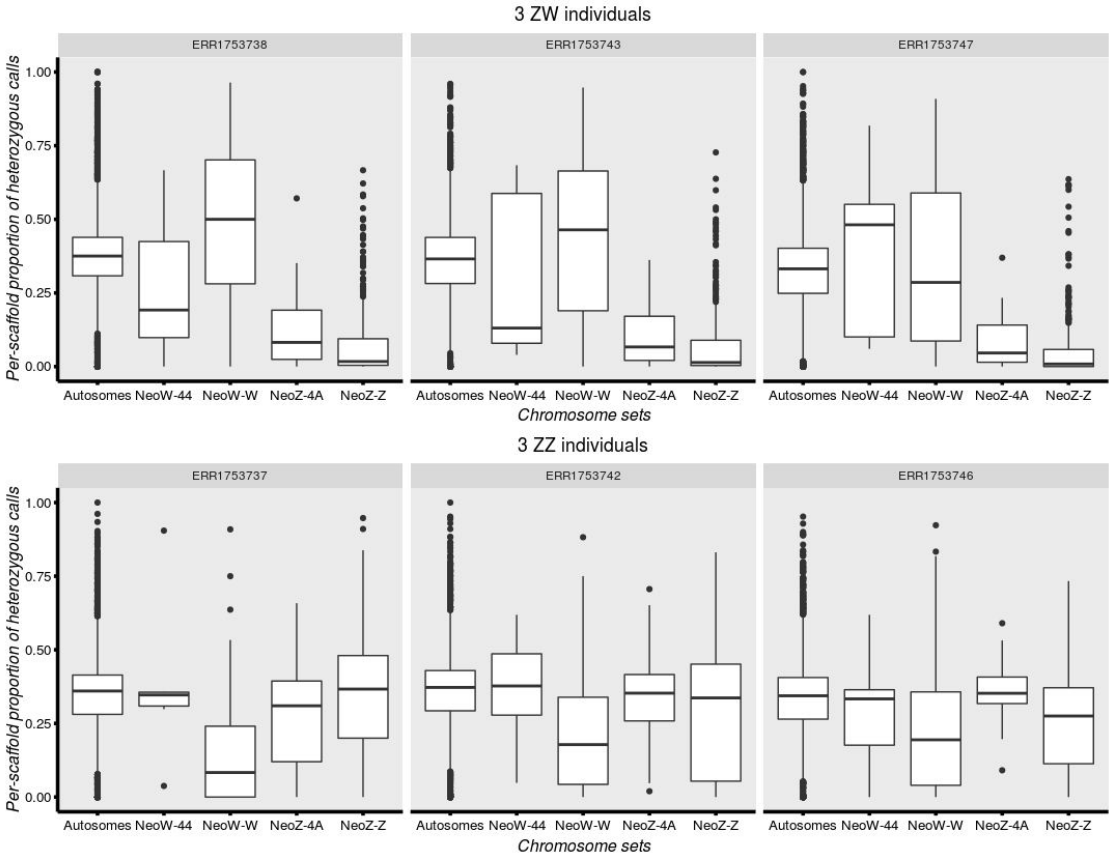

Fig2: Per-scaffold proportion of heterozygous calls at autosomal, neoW and neoZ SNPs for the 3 ZZ and 3ZW individuals, after excluding all SNPs in TE regions. Computations were restricted to scaffolds >10 kb with at least one SNP in one individual (ZW: autosomes 1,957; NeoW-W 99; NeoW-4A 8; NeoZ-Z 305 & neoZ-4A 42 vs. ZZ: autosomes 1,891; NeoW-W 88; NeoW-4A 7; NeoZ-Z 313 & neoZ-4A 42).

9. Lines 556-557: Some of the high dN/dS values might be due to positive selection instead of "ongoing pseudogenization". Did the authors check for premature stop codons and/or frameshifts in these genes?

=>This is a valid point and we can not formally exclude the action of positive selection using only the dN/dS estimates. We ran two additional analyses to test for the existence of pseudogene and to explicitly test for the action of positive selection. First, we re-aligned the raw sequences using macse (v2, http://bioweb.supagro.inra.fr/macse/ ; Ranwez et al. 2018. Mol. Biol. Evol. 35:2582–2584.). Macse is designed to take into account potential frameshift and to efficiently align pseudogene. We aligned the six genes with a dN/dS close to 1.0 (dN/dS > 0.5) with macse but we failed to find any premature stop codon or indels inducing frame-shift indicating that this gene might still be functional. Second, we explicitly test for the action of positive selection on the three genes with dN/dS > 1.0 in the branch of the neoW-W formation. To do that, we performed a likelihood-ratio test (LRT) to compare a model with a free dN/dS ratio with a model where the dN/dS is fixed to 1.0. The LRT were all non significant (p-value > 0.4) indicating a dN/dS not significantly different from 1.0 and, therefore, not evidence for the action of positive selection. We have now added this two analyses in the main text (lines 648-658).

10. Lines 625-626: Note that contig N50 might be an even better predictor for assembly quality in this context.

=> A reference genome of optimal quality is a complete and contiguous assembly, with one sequence per chromosome (end-to-end). By moving away from this ideal situation, the level of quality of a given draft sequences will depend on the scientists' interests, sequence organization or sequence composition in particular. In this study, our main objective was to generate a bird assembly with an excellent sequence organization. Consistently, our key selection criteria in the selection of the different assemblies were statistics such as scaffold N50 or total size of scaffolds, rather than N% or contig N50. In addition, such contig N50-based comparisons are more difficult to do in general, because these statistics are less frequently reported. To provide this information to the reader through this letter, we provided a simple test by using assemblathon2 to compute a "N50-like contig length". In brief, this does not represent a true "contig N50" since we computed these statistics on the scaffold assemblies. We used a simple rule to split scaffolds into contigs (≥100 N bases) for all the assemblies. Among all bird species we investigated, the lower N50-like statistic was observed for *Z. borbonicus* (21kb) as compared to all other bird species (range: 27-17490kb; median 39kb). For this purpose, the *Z. lateralis* genome (another *Zosterops* species) performs much better (42 kb). However, this must be nuanced since the total assembly size of *Z. lateralis* genome sequence is up to 200Mb lower than expected, which therefore another limit regarding the genome completeness.

11. Lines 831-832: Please explain this statement. Shouldn't a fission of a chromosome into two parts (no matter how close/distant to the centromere) still change the overall recombination rate?

=>In this sentence, we assume that at least one chiasma per chromosome arm per meiosis is required to stabilize homologous chromosome pairs. This is the case in mammals (Santucci-Darmanin and Baudat . 2010. Meiotic recombination in mammals. Oogenesis Univers. Process:141) and we assume it is also the case in birds. If so, it is not the chromosome size but the arm size that influences recombination rate (assuming everything else equal). As a consequence, a translocation close to the centromere will not change the arm size of the remaining arm of the autosomal 4A.

12. Lines 879-882: I fear obtaining a "full sequence" of any W chromosome is not feasible as long as read lengths are below the megabase-scale. I suggest using the term "high-quality sequence" instead.

=> We agree. This line was edited following the reviewer's comment (see lines 1020-1027).

13. Lines 941-942: Similar TE densities were also reported for a partial W sequence of white-throated sparrow (Davis, J.K., P.J. Thomas & J.W. Thomas. 2010. A W-linked palindrome and gene conversion in New World sparrows and blackbirds. Chromosome Res. 18: 543–553.).

=> This work is now cited in the discussion.

14. Lines 948-950: Similar to zebra finch, collared flycatcher also appears to have recently active LTR elements (Suh, A., L. Smeds & H. Ellegren. 2018. Abundant recent activity of retrovirus-like retrotransposons within and among flycatcher species implies a rich source of structural variation in songbird genomes. Mol. Ecol. 27: 99–111.

=> This reference was also added to the manuscript (line 1105). We thank the referee for pointing out these references.

15. Lines 951-952: Is this a "recent burst of LTR elements on autosomes" or rather an overall burst of LTR elements with an increased retention of insertions in low-recombination regions, such as the neoW-4A region?

=> Again, we agree with the reviewer's comment. We added a line to explain this scenario (see l. 1106-1115).

16. Lines 1485-1486: Why were Repbase repeats used only from chicken and not also those from zebra finch (i.e., by selecting taxon "Aves")? As the authors state in the manuscript, there was recent LTR activity in zebra finch and it is possible that zebra finch might be sharing some TEs with Zosterops. This might possibly reduce the percentage of "No Category" (i.e., "unknown" or "unclassified") repeats in the Zosterops annotation.

=> We agree that the repbase repeat for the Zebra finch could be used as an alternative or complement. By using the chicken reference database, our main objective was to use the consensus sequences for all avian species and then generate a *de novo* TE library for our focal species. We agree that it remains possible to reduce the proportion of unclassified transposable elements. More broadly, we think that far more detailed analyses remain to be done on this topic in future.

Additional minor comments:

17. Lines 4-5: I assume this statement applies to the recombining sex chromosome (Z or X), but not the sex-limited one (W and Y)? Or does this statement only apply to synteny of sex chromosomes (but not to intrachromosomal rearrangements)? Please clarify.

Right. At this step, we do not have sufficient information about the non-recombining sex chromosomes.

18. Line 27: The term "transposed regions" is unclear here. How about using the term "sex-linked", to keep it independent of the mechanism leading to sex linkage/sex chromosome differentiation (e.g., transposition, translocation, fusion)?

Edit done.

19. Lines 44-45: I assume also purifying selection should be more efficient in the heterogametic sex.

Edit done.

20. Line 52: The authors could use the more general term “linked selection” instead of “background selection” here.

Edit done.

21. Line 65: I suggest replacing the term “all birds” with “most birds” here.

Edit done.

22. Line 118: I suggest replacing the term “chromosomes” with “pseudochromosomes” here.

Edit done.

23. Line 135: Please replace “surprising” with “surprisingly”.

Edit done.

24. Line 244: This approach not only assumes “perfect synteny” but also “collinearity”, which I suggest adding here.

Edit done.

25. Line 319: I suggest adding “onset of the” before “diversification”.

Edit done.

26. Line 422: It is not clear to me how “high activity of transposable elements” can lead to homology (only) between this scaffold and chromosome 4A. Please clarify.

Local alignments unambiguously support a breakpoint near 9.603Mb. This sentence was to precise that special attention need to be made on additional hits generated by the recent TE activity or some other local biases. This sentence unnecessarily complicate the reading and were removed from the new version to ensure a proper understanding.

27. Line 579: Please add “p=” before “0.096”.

Edit done.

28. Line 638: Are GC-rich regions mainly difficult because of assembly issues or rather underrepresentation in sequencing data?

We made the sentence clearer (l. 748-752).

29. Line 807-808: I assume the authors mean GC content negatively correlates with chromosome size?

Edit done.

30. Line 813-814: In addition to “instantaneous changes in chromosome sizes”, the new recombination regime (male-specific) might also play a role here, assuming that recombination quickly ceased (or was quickly reduced) between neoW-4A and neoZ-4A.

The divergence between the neoW-4A and neoZ-4A homologous copies indicate that the recombination has ceased quite early in the Sylvioidea evolution. Indeed, male specific recombinaison might also participate to the GC content evolution of the neoZ-4A sequences.

31. Line 912: What do the authors mean with the term “non-recombinant W chromosome”? Please clarify.

Here, we only discussed the non-recombinant W region (excluding the PAR region of the W chromosome)

32. Line 1044: Typo in “lane”.

Edit done.

33. Lines 1243-1245: If de-novo prediction was done, I assume the authors mean “RepeatModeler” instead of “RepeatMasker” here?

Edit done.

34. Lines 1286-1289: If there was disagreement between scaffold orders, was preference given to LASTZ or DeCoSTAR? Please clarify.

=> 31 adjacencies predicted by DeCoSTAR were in conflicts with LASTZ syntenies. Most of them were manually solved by further re-analysis of raw LASTZ alignments (before filtering. For remaining conflicts (where raw LASTZ alignments couldn’t help to solve conflicts) preference was given to DeCoSTAR as syntenies inferred are guided by syntenies from multiple close related genomes instead of just one. Moreover, some of adjacencies predicted by DeCoSTAR, in conflicts with LASTZ syntenies, were supported by paired-reads data validating DeCoSTAR predictions.

35. Lines 1309-1311: Were these distances to chromosome ends measured in zebra finch chromosomes or Z. borbonicus pseudochromosomes?

This distances were measured in zebra finch chromosomes. We have added this information (lines 1533-1538).

36. Line 1338: Please remove “that”.

Edit done.

37. Lines 1425-1430: Please clarify how a single base on the neoW chromosome can have non-zero coverage in males.

=> Given the highly repeated nature of the neoW chromosome, read mismapping is particularly expected (particularly in regions detected as TE, but non only, since this detection remains non-exhaustive) generating a non-zero coverage (and in the worst case, SNP calls, see also #8). We made this point explicit.

38. Lines 1467-1478: In contrast to most parts of the manuscript, here the English names for the study species are used. Consider simplifying this as a service to the reader, for example by only using the scientific name.

Edit done.