Contrasted sex chromosome evolution in primates with and without sexual dimorphism

Rylan Shearn¹, Emilie Lecompte², Corinne Régis¹, Sylvain Mousset³, Simon Penel¹, Guillaume Douay⁴, Brigitte Crouau-Roy², Gabriel A.B. Marais¹*

¹Laboratoire Biométrie et Biologie Evolutive, CNRS / Univ. Lyon 1, France.

²Laboratoire Evolution et Diversité Biologique, CNRS / Univ. Toulouse, France.

³Faculty of Mathematics, University of Vienna, Austria.

⁴Zoo de Lyon, France.

*Correspondence to: gabriel.marais@univ-lyon1.fr
Abstract:

The human sex chromosomes provide a striking example of how sex chromosomes can become modified in gene content and structure (i.e. become heteromorphic). Why sex chromosomes in humans and other lineages can reach such levels of heteromorphy is not clear. One hypothesis posits that mutations with antagonistic effects between males and females can drive recombination suppression and subsequent divergence between X and Y chromosomes. Here, we tested this hypothesis by focusing on strepsirrhine primates (lemurs and lorises) that have much lower sexual dimorphism and opportunities for sexually antagonistic mutations than haplorrhines (apes and monkeys). We sequenced seven strepsirrhine species and identified the pseudoautosomal boundary (PAB), the border between recombining and non-recombining regions on the sex chromosomes. We found that during primate evolution, the PAB has remained unchanged in strepsirrhines whereas several recombination suppression events have occurred in haplorrhines, suggesting that sexually antagonistic mutations might have driven sex chromosomes heteromorphy in primates.
The human sex chromosomes are strongly heteromorphic as they exhibit extensive differences in size, gene number, DNA repeat abundance and heterochromatin composition\(^1,2\). The X chromosome comprises a large X-specific region recombining only in females whereas the Y comprises a male-specific region that does not recombine at all. Both sex chromosomes share two regions called pseudoautosomal regions (PAR1 and 2) that recombine in both males and females. Strikingly, our strongly heteromorphic sex chromosomes originated from a pair of nearly identical autosomes ~150 millions years ago\(^1,3,4\). It all started with the emergence of \(Sry\) – the master male-determining gene in mammals - from \(Sox3\) just prior to the divergence of placentals and marsupials\(^1,3,4\). Since then, at several moments throughout evolutionary history, vast regions of the Y chromosome have at once stopped recombining with the X, likely through inversions on the Y, forming the so-called evolutionary strata\(^3,5,6,7\). The most recent strata (4 and 5) have originated in the history of Catarrhini (Old World monkeys and apes) respectively, ~40 and ~25 Mya, and now only a very small PAR continues to recombine between X and Y in those primates\(^8\). In humans, PAR1 is the consequence of that process, while PAR2 is a recent addition\(^1\).

Such a process of gradual suppression of recombination between sex chromosomes, of PAR reduction, and formation of evolutionary strata has been documented in several animal and plant lineages (e.g\(^9,10,11\)). Why such a process occurred, however, is unclear. It has been proposed that sexually antagonistic mutations may have favoured the suppression of recombination\(^12,13\). Theoretical models suggest that if there are male-beneficial/female-detrimental mutations in the PAR, genetically linking those mutations to the Y chromosome, through for example an inversion, will be selected. Some evidence supporting this hypothesis has recently been found in
guppies, but in many groups including primates where it has been proposed to explain sex chromosome evolution, evidence is lacking.

While previous work on primate sex chromosomes has focused on Haplorrhini (apes, Old and New World monkeys), we studied representatives of the other main primate lineage, the Strepsirrhini (lemurs and lorises). In strepsirrhines, female social dominance (FSD), in which females dominate males, is widespread and likely ancestral. FSD is associated with increased testosterone production in females, resulting in the masculinization of females, including aspects of their social behaviour and genitalia. Some species also have rather egalitarian social systems. In addition, sexual size dimorphism is virtually absent among strepsirrhines. This is in sharp contrast with haplorrhines, where sexual dimorphism is much more pronounced and male-biased. We therefore hypothesized that if male-female differentiation and sexually antagonistic mutations were associated with the degree of X-Y recombination suppression, strepsirrhines should show evidence of less recombination suppression compared to haplorrhines. However, to date, very little is known about the sex chromosomes of strepsirrhines, except that strata 4 and 5 are missing in gray mouse lemurs (Microcebus murinus) preventing previous tests of this hypothesis.

To identify the PAB of strepsirrhines, we used an approach relying on sequencing a male and a female at low/moderate depth, mapping the reads to a reference genome and computing the male:female depth ratio. For autosomes, a M:F depth ratio of 1 is expected as males and females have the same number of autosomes. On the X chromosome, a ratio of 1 should indicate the PAR that is shared among sexes, a ratio of 0.5 the X-specific region as males have only one such region and females two, and the boundary between both would indicate the PAB. Using Illumina short-read sequencing technology, we sequenced a male and a female genome in
seven species covering the main strepsirrhine lineages representing 65 My of evolution: four Lemuriformes (aye-ayes, gray mouse lemur, red-bellied lemur, greater bamboo lemur) and three Lorisiformes (northern greater galago, senegal bushbaby, slow loris). The sequencing depth of each sample was between 11.8X and 39.1X (of human genome equivalent) with 78% of the samples being between 20X and 40X, i.e. moderate sequencing depth (Table S1). We then mapped the reads onto publicly available reference genomes of two strepsirrhhines (using the human X to scaffold the strepsirrhine X chromosome) and computed a normalized M:F depth ratio to identify the X-specific region and the PAR on the X chromosome (see Methods).

Results

Figure 1A-B shows the results for the gray mouse lemur. Using the human X chromosome to order the gray mouse lemur X scaffolds, we found that the scaffolds corresponding to human strata 4, 5 and PAR1 have a M:F depth ratio around 1 (Fig. 1B), indicating that these regions have remained pseudoautosomal in gray mouse lemurs and agreeing with older cytogenetic data. The rest of the gray mouse lemur X is X-specific with a ratio close to 0.5. However, five regions in the X-specific region show an elevated ratio. Detailed analysis of these five regions showed that they are fragments of autosomes (see Text S1). It is not clear, however, whether this comes from contamination of the assembly of the X chromosome by autosomal scaffolds or if this has resulted from fusion of autosomal DNA fragments to the PAR during evolution, which are misplaced in the current assembly of the X chromosome. Only an improved assembly of the X chromosome in the gray mouse lemur could confirm one of these alternatives. With the fragmented assembly that is available our approach only reliably identifies the PAB, not the size of the PAR. If some autosomal material were translocated to the PAR, and thus enlarging it, it
would not be possible to detect it with our approach. Despite these limitations, it is nonetheless clear that the regions homologous to strata 4, 5 and PAR1 of humans are still recombining in gray mouse lemur and that the PAB is still the same PAB as in the ancestor of all primates.

We repeated the same analysis for the other six species (Fig. 1C). For the lemurs, we used the gray mouse lemur reference genome for the mapping because it is the only available one, and for the lorises, we used the northern greater galago reference genome for the same reason (see Methods and Fig. S1 for the dot plot with the human X). Some species are quite distant from the species with the reference genome and mapping was consequently more difficult. This explains why in some cases the M:F depth ratio is more variable. The results of the aye-ayes analyses are especially noisy because of the large phylogenetic distance to the species with the reference genome (Fig. 1C). However, in all seven species studied here, the pattern is very similar. All studied strepsirrhines harbor a large pseudoautosomal region including the genes that are in PAR1, strata 4 and 5 in humans. We can therefore conclude that no recombination suppressing event has occurred in strepsirrhines since the origin of the group >65 millions years ago (Fig. 2).

Our findings are consistent with the hypothesis that recombination suppression between X and Y chromosomes is driven by sexually antagonistic mutations. However, the rate of strata formation is generally low and additional strata in strepsirrhines might be absent just by chance. In haplorrhines, at least two strata (4 and 5) were formed in apes and Old World monkeys, and one was formed independently in New World monkeys (4’) based on the species studied so far (8, 24 and see Fig. 2). We designed a statistical test to compare the rate of strata formation in both suborders using the phylogeny of the species in which the presence of strata was examined. Our test points to a significant difference between both lineages (see Text S2). It is possible that the M:F read depth approach missed recently evolved strata in strepsirrhines. Recent strata are
indeed more difficult to detect with the M:F read depth approach as sex chromosome divergence can be so low that both X and Y reads map onto the X chromosome and the ratio is close to 1. To identify recent strata, we computed the male:female SNP density ratio, which is expected to more effectively detect the PAB when recent strata are present. The M:F SNP density ratio is predicted to be 1 for the PAR, 0.5 for the old strata (using the same rationale as for the M:F read depth ratio) and to be increased (>1) for the recent strata due to accumulation of fixed X-Y differences. However, our analyses revealed no recent strata in the seven strepsirrhines species studied here (Figure S2).

Discussion
Evidence for the sexually antagonistic mutations hypothesis has been found in other organisms. In guppies, populations exhibiting stronger sexual dimorphism seem to have more strata. In the brown alga Ectocarpus, sexual dimorphism is extremely low and as expected sex chromosomes are homomorphic, with a small non-recombining region, despite being very old. It should be noted, however, that other forces might be driving the process of strata formation in some lineages. In ruminants, the PAB seems to have moved through a process of erosion due to DNA repeats accumulation. Moreover, in fungal species having mating types instead of sexes, strata have formed when genes important for a given mating type were linked to the mating locus. Thus, sexually antagonistic mutation may not be a ubiquitous explanation of strata formation in all organisms.

Although sexual dimorphism is generally low in strepsirrhines, there are some differences among species in this lineage, with the genus Eulemur exhibiting the most pronounced sexual dimorphism. In these species, including the red-bellied lemur (Eulemur rubriventer), which
was analysed here, males and females exhibit striking sexual dichromatism, i.e. they differ in pelage colouration \(^{27}\). The red-bellied lemur showed no more evidence for recombination suppression than the other species studied here. Sexual dichromatism may imply sexually antagonistic mutations. The antagonism might have been solved not through Y-linkage but instead through sex-biased expression for example \(^{28}\). Similarly, in Emu (\textit{Dromaius novaehollandiae}), sex chromosomes are homomorphic and sexual dimorphism is thought to be have evolved mainly through sex-biased expression \(^{22}\). Future research could focus on sex-biased expression in strepsirrhines to test these ideas.

**Methods**

**Research plan.** To examine whether strepsirrhine sex chromosomes have undergone less recombination suppression compared to haplorrhines, we selected strepsirrhine species that would maximise the representation of this group’s diversity, and that were also readily accessible. We then sequenced a male and female of each species and mapped the obtained male and female reads to a reference X chromosome. The male to female depth ratio was then computed along the length of the X chromosome and the PAB was identified as the boundary between zones with a ratio of one (indicative of the PAR) and zones with a ratio of 0.5 (indicative of the non-recombining region).

**Sampling.** We selected seven species covering as much phylogenetic diversity of Strepsirrhini as possible (see Table S1). Both infra-orders (Lemuriformes and Lorisiformes) are equally represented. A male and a female individual were sampled for all species (except \textit{Otolemur ...
garnetti, the northern greater galago, for which sequence data from a female individual were retrieved from NCBI, see Table S1). Blood samples of *Eulemur rubriventer* (red-bellied lemur) and *Prolemur simus* (greater bamboo lemur) were collected from living animals at Zoo de Lyon on EDTA blood collection tubes to avoid coagulation. Hair samples (with follicles and roots) of the female *Daubentonia madagascariensis* (aye-aye) were collected from a living animal at Zoo Frankfurt. Samples of *Microcebus murinus* belong to the Bruyno laboratory (UMR7179, France; agreement E91-114-I from the Direction Départementale de la Protection des Populations de l’Essonne): the biopsies were obtained from muscle tissues after the animals’ natural death.

Tissues samples of a male *Daubentonia madagascariensis*, and samples of *Galago senegalensis* (Senegal bushbaby), *Nycticebus coucang* (slow loris) and of a male *Otolemur garnetti* were obtained from the tissues and cryopreserved cells collection of the National Museum of Natural History (MNHN, Paris, see Table S1).

**DNA extraction and sequencing.** DNA from *Eulemur rubriventer*, *Prolemur simus* and female *Daubentonia madagascariensis* were extracted using two different Macherey Nagel kits. Blood sample were treated with NucleoSpin Blood Quickpure kit. Hair samples were treated with NucleoSpin DNA trace kit after a mechanical crushing of hair bulbs. DNA from the tissues and cells samples (for other species) was extracted using the DNeasy Blood and Tissue kit (Qiagen) following the manufacturer’s instructions. DNA was stored at -20°C and sent onto dry ice to the sequencing platform.

A DNAg library was constructed for each sample using Illumina kits (TruSeq nano LT for Hiseq 2500 and 3000 sequencing). Paired-end sequencing was conducted using an Illumina Hiseq 2500 (2 x 125 bp) or 3000 (2 x 150 bp) with 1 or 2 individuals per lane at Genotoul, the
INRA sequencing platform in Toulouse. Sequences were all found to be of high quality and without contamination. Sequence data and coverage are shown in Table S1.

**Chromosome assembly.** Reference X chromosomes were not available for any species and genome assemblies were only available for two species that were 1) closely related to, or the same as the species being studied, and 2) assembled to an extent that it would be possible to construct a *de novo* X chromosome. These were *Microcebus murinus* (gray mouse lemur) and *Otolemur garnettii* (northern greater galago).

*De novo* X chromosomes were constructed for these species using scaffolds from whole genome assemblies on NCBI, which were selected, ordered and oriented against the human X chromosome. This was achieved using SynMap, an online software pipeline within the CoGe toolkit\(^\text{29,30}\) that identified putative homologous genes between potential X scaffolds and the human X chromosome with a blast comparison\(^\text{31}\) using the Last algorithm (a variant of Blastz;\(^\text{32}\)). An algorithm within the SynMap pipeline then identified a colinear series of homologous genes between potential X scaffolds and the human X chromosome as regions of synteny, and these were arranged in order accordingly. The relative gene order DAGChainer option was used, with a maximum distance of 20 genes between two matches and a minimum of five aligned pairs of genes. The human X chromosome reference was sourced from the GRCh37.p13 Primary Assembly on NCBI (Reference Sequence: NC_000023.10).

As the results of some of the analyses in this study required normalisation using an autosome from the corresponding species, a reference autosome was constructed using the same process. In this case, the human chromosome four was used to construct a *de novo* chromosome
four for *Microcebus murinus* and *Otolemur garnettii*, which was selected for its similar size to the X chromosome).

**Read mapping.** Male and female reads for each species were aligned separately to their most closely related *de novo* X chromosome using Bowtie version 2.2.7. The reads were then sorted according to their position on the *de novo* X chromosome using Samtools version 1.3.1.

**Coverage analysis.** Read depth was calculated for each sex at each position from the mapped reads on the *de novo* X using Samtools. The coverage for each sex was then normalised by dividing the depth at each position by the mean coverage depth for that species and sex on an autosome (chromosome four). The ratio of normalised male to female coverage was then calculated at each position and the data was summarised as a sliding window average using a window size of 150 kb sliding at increments of 10 kb. This data manipulation was performed using AWK version 4.1.3. Scripts for the entire coverage analysis pipeline (suitable for compute clusters using Torque job scheduling) are available on GitHub ([https://github.com/rylanshearn/sex-read-depth](https://github.com/rylanshearn/sex-read-depth)).

**SNP density analysis.** To detect potential areas that may have stopped recombining between strepsirrhine X and Y chromosomes relatively recently, the difference in male to female SNP diversity was also examined for all species. For each sex of each species, SNPs were called from the mapped reads using Samtools mpileup and then converted to profiles using sam2pro version 0.8 from the mlRho package. Average SNP density was then calculated at each position.
following as the sum of SNPs divided by four, and one was added to allow for a Log2 transformation. The male to female SNP density was then calculated at each site as Log2(average male SNP density) – Log2(average female SNP density). This calculation was performed using AWK version 4.1.3 and R version 3.4.0. Scripts for the SNP density analysis are also available on the GitHub repository (https://github.com/rylanshearn/sex-read-depth).

**Data availability:** all the data generated in this study is available at NCBI (project # PRJNA482296).

### References

1. Skaletsky H, *et al.* The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **423**, 825 - 837 (2003).

2. Ross MT, *et al.* The DNA sequence of the human X chromosome. *Nature* **434**, 325-337 (2005).

3. Lahn BT, Page DC. Four evolutionary strata on the human X chromosome. *Science* **286**, 964-967 (1999).

4. Hughes JF, Rozen S. Genomics and genetics of human and primate y chromosomes. *Annu Rev Genomics Hum Genet* **13**, 83-108 (2012).

5. Van Laere A-S, Coppieters W, Georges M. Characterization of the bovine pseudoautosomal boundary: Documenting the evolutionary history of mammalian sex chromosomes. *Genome Research* **18**, 1884-1895 (2008).

6. Lemaitre C, Braga MD, Gautier C, Sagot MF, Tannier E, Marais GA. Footprints of inversions at present and past pseudoautosomal boundaries in human sex chromosomes. *Genome Biol Evol* **1**, 56-66 (2009).

7. Pandey RS, Wilson Sayres MA, Azad RK. Detecting evolutionary strata on the human x chromosome in the absence of gametologous y-linked sequences. *Genome Biol Evol* **5**, 1863-1871 (2013).

8. Hughes JF, *et al.* Strict evolutionary conservation followed rapid gene loss on human and
rhesus Y chromosomes. *Nature* **483**, 82-86 (2012).

9. Nicolas M, *et al.* A gradual process of recombination restriction in the evolutionary history of the sex chromosomes in dioecious plants. *PLoS Biol* **3**, e4 (2005).

10. Zhou Q, *et al.* Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science* **346**, 1246338 (2014).

11. White MA, Kitano J, Peichel CL. Purifying Selection Maintains Dosage-Sensitive Genes during Degeneration of the Three-spine Stickleback Y Chromosome. *Mol Biol Evol* **32**, 1981-1995 (2015).

12. Rice WR. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex-chromosomes. *Evolution* **41**, 911-914 (1987).

13. Charlesworth D, Charlesworth B, Marais G. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* **95**, 118-128 (2005).

14. Wright AE, *et al.* Convergent recombination suppression suggests role of sexual selection in guppy sex chromosome formation. *Nature communications* **8**, 14251 (2017).

15. Kappeler PM, Fichtel C. Eco-evo-devo of the lemur syndrome: did adaptive behavioral plasticity get canalized in a large primate radiation? *Frontiers in zoology* **12 Suppl 1**, S15 (2015).

16. Petty JM, Drea CM. Female rule in lemurs is ancestral and hormonally mediated. *Sci Rep* **5**, 9631 (2015).

17. Pereira ME, Kappeler PM. Divergent systems of agonistic behaviour in lemurid primates. *Behaviour* **134**, 225-274 (1997).

18. Plavcan JM. Sexual selection, measures of sexual selection, and sexual dimorphism in primates. *Sexual selection in primates*, 230 (2004).

19. Glaser B, *et al.* Transposition of SRY into the ancestral pseudoautosomal region creates a new pseudoautosomal boundary in a progenitor of simian primates. *Hum Mol Genet* **8**, 2071-2078 (1999).

20. Vicoso B, Bachtrog D. Lack of global dosage compensation in Schistosoma mansoni, a female-heterogametic parasite. *Genome Biol Evol* **3**, 230-235 (2011).

21. Vicoso B, Emerson JJ, Zektser Y, Mahajan S, Bachtrog D. Comparative sex chromosome genomics in snakes: differentiation, evolutionary strata, and lack of global dosage compensation. *PLoS Biol* **11**, e1001643 (2013).
22. Vicoso B, Kaiser VB, Bachtrog D. Sex-biased gene expression at homomorphic sex chromosomes in emus and its implication for sex chromosome evolution. *Proc Natl Acad Sci U S A* **110**, 6453-6458 (2013).

23. Pozzi L, Hodgson JA, Burrell AS, Sterner KN, Raaum RL, Disotell TR. Primate phylogenetic relationships and divergence dates inferred from complete mitochondrial genomes. *Mol Phylogenet Evol* **75**, 165-183 (2014).

24. Cortez D, *et al.* Origins and functional evolution of Y chromosomes across mammals. *Nature* **508**, 488-493 (2014).

25. Ahmed S, *et al.* A haploid system of sex determination in the brown alga Ectocarpus sp. *Curr Biol* **24**, 1945-1957 (2014).

26. Branco S, *et al.* Evolutionary strata on young mating-type chromosomes despite the lack of sexual antagonism. *Proc Natl Acad Sci U S A* **114**, 7067-7072 (2017).

27. Rakotonirina H, Kappeler PM, Fichtel C. Evolution of facial color pattern complexity in lemurs. *Sci Rep* **7**, 15181 (2017).

28. Ellegren H, Parsch J. The evolution of sex-biased genes and sex-biased gene expression. *Nat Rev Genet* **8**, 689-698 (2007).

29. Lyons E, Freeling M. How to usefully compare homologous plant genes and chromosomes as DNA sequences. *Plant J* **53**, 661-673 (2008).

30. Lyons E, *et al.* Finding and comparing syntenic regions among Arabidopsis and the outgroups papaya, poplar, and grape: CoGe with rosids. *Plant Physiol* **148**, 1772-1781 (2008).

31. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* **215**, 403-410 (1990).

32. Schwartz S, *et al.* Human-mouse alignments with BLASTZ. *Genome Res* **13**, 103-107 (2003).

33. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* **10**, R25 (2009).

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

35. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987-2993 (2011).
36. Haubold B, Pfaffelhuber P, Lynch M. mlRho - a program for estimating the population mutation and recombination rates from shotgun-sequenced diploid genomes. *Mol Ecol* **19 Suppl 1**, 277-284 (2010).

37. Horvath JE, *et al*. Development and application of a phylogenomic toolkit: resolving the evolutionary history of Madagascar's lemurs. *Genome Res* **18**, 489-499 (2008).
Acknowledgments: We thank the Muséum National d'Histoire Naturelle de Paris for access to their collections of tissues and cryopreserved cells, Fabienne Aujard (UMR7179 MECADEV – CNRS, MNHN, Brunoy, France) for providing the gray mouse lemur samples and Christina Geiger from Zoo Frankfurt for providing the female aye-ayes samples. We thank Stéphane Delmotte and Bruno Spataro (LBBE computing unit) for maintenance of the cluster on which all the bioinformatics was done. We thank Alison E Wright for advice with the M:F SNP density analysis. We thank Philippe Faivre for providing pictures of primates for figure 2. We thank Peter Kappeler for suggestions to improve this manuscript and Laurent Duret and Hervé Philippe for discussions. RS, EL, BCR and GABM acknowledge the financial support by ANR (grant number ANR-12-BSV7-0002-04).

Author contributions: conceptualization, methodology and funding acquisition: GABM (lead), EL, BCR; supervision and project administration: GABM, formal analysis: RS (lead), SM, SP; investigation: RS (lead), EL, CR, SM, SP, BCR, GABM; resources: GD, writing- original draft preparation: GABM (lead), RS, SM; writing – review and editing: EL and all authors; visualization: RS.

Competing financial interests: Authors declare no competing financial interests
**Figure legends**

**Fig. 1.** Identification the PAB in seven strepsirrhine species. (A) synteny plot of the human and gray mouse lemur X chromosomes. The human X was used to order the gray mouse lemur scaffolds (see Methods). Black dots represent orthologous genes between the human and gray mouse lemur X chromosomes. Strata in humans using $^1,^4$ definition of strata are shown (note that old strata have been split into smaller strata in $^7$). (B) M:F read depth ratio along the gray mouse lemur X chromosome. (C) M:F read depth ratio for all seven strepsirrhine species. See Text S1 for the detailed analysis of the regions with elevated M:F coverage ratio shown in grey.

**Fig. 2.** Strata formation in primates. Data on strata in haplorrhines are from $^1,^2,^3,^4,^8,^24$; data on strepsirrhines are from this study. The phylogenetic tree and divergence times are from $^{23,37}$.
