A paternal bias in germline mutation is widespread in amniotes and can arise independently of cell divisions

Marc de Manuel[1*†], Felix L. Wu[1,2*†] and Molly Przeworski[1,2*]

[1]Department of Biological Sciences, Columbia University, New York, New York, USA.
[2]Department of Systems Biology, Columbia University, New York, New York, USA.

*To whom correspondence should be addressed: md3914@columbia.edu; flw2113@cumc.columbia.edu; mp3284@columbia.edu.
†These authors contributed equally to this work.

Abstract

In humans and other mammals, germline mutations are more likely to arise in fathers than in mothers. Although this sex bias has long been attributed to DNA replication errors in spermatogenesis, recent evidence from humans points to the importance of mutagenic processes that do not depend on cell division, calling into question our understanding of this basic phenomenon. Here, we infer the ratio of paternal-to-maternal mutations, $\boldsymbol{\alpha}$, in 42 species of amniotes, from putatively neutral substitution rates of sex chromosomes and autosomes. Despite marked differences in gametogenesis, physiologies and environments across species, fathers consistently contribute more mutations than mothers in all the species examined, including mammals, birds and reptiles. In mammals, $\boldsymbol{\alpha}$ is as high as 4 and correlates with generation times; in birds and snakes, $\boldsymbol{\alpha}$ appears more stable around 2. These observations are consistent with a simple model, in which mutations accrue at equal rates in both sexes during early development and at a higher rate in the male germline after sexual differentiation, with a conserved paternal-to-maternal ratio across species. Thus, $\boldsymbol{\alpha}$ may reflect the relative contributions of two or more developmental phases to total germline mutations, and is expected to depend on generation time even if mutations do not track cell divisions.

1 Introduction

Humans tend to inherit more *de novo* mutations (DNMs) from their fathers than from their mothers. This phenomenon was first noted over 70 years ago, when JBS Haldane relied on the population frequency of hemophilia in order to infer that the de novo mutation rate at the disease locus is substantially higher in fathers [1]. Work since then, particularly in molecular evolution, has confirmed a "male bias" in mutation (henceforth paternal bias) [2–9], with estimates from human pedigrees indicating that, genome-wide, DNMs occur roughly four times more often on the paternal genome than on the maternal one [10, 11].

The textbook explanation for the paternal mutation bias is that it arises as a consequence of the vastly different numbers of cell divisions–and hence DNA replication cycles–necessary to produce sperm compared to oocytes [12–15]. Indeed, in humans as in other mammals, oocytes are arrested in meiotic prophase I at birth, with no subsequent DNA replication in the mother's life, whereas spermatogonia start dividing shortly before puberty and divide continuously throughout the reproductive life of the father [13, 16]. The observation that the number of DNMs increases with paternal age has been widely interpreted in this light, as evidence for DNA replication errors being the predominant source of germline mutation [10, 11, 17, 18].

A number of recent findings have called this view into question, however. First, analyses of large numbers of human pedigrees revealed an effect of maternal age on the number of maternal DNMs [19, 20], with an additional ∼0.4 mutations accrued per year. Given the lack of mitotic cell division in oocytes after birth, this observation indicates that by typical reproductive ages, at least half of maternal DNMs arise from DNA damage [18]. Second, despite highly variable rates of germ cell division over human ontogenesis, germline mutations accumulate with absolute time in both sexes, resulting in a ratio of paternal-to-maternal germline mutation, α, of around 3.5 at puberty and very little increase with parental ages [21]. Third, studies in a dozen other mammals suggest that α ranges from 2 to 4 whether the species reproduces months, years or decades after birth [22–24], when estimates of germ cell division numbers at time of reproduction would predict a much wider range in α [13, 22, 25, 26].

Explaining the observations in humans under a model in which most mutations are due to replication errors, and thus track cell divisions, would call for an exquisite balance of cell division and mutation rates across developmental stages in both sexes [27]. In males, the constant accumulation of mutations with absolute time would require varying rates of germ cell divisions over ontogenesis to be precisely countered by reciprocal differences in the per cell division mutation rates. In females, it would necessitate that the mutation rate per unit of time be identical whether mutations arise from replication errors or damage. In turn, the similarity of α across mammals that differ drastically in their reproductive ages would entail two distinct sources of mutation—replication error in males and damage in females—covarying in tight concert with generation times.

A more parsimonious alternative is that most germline mutations arise from the interplay between damage and repair rather than from replication errors [28], and that the balance results in more mutations on the paternal than the maternal genome [27]. Assuming repair is inefficient relative to the length of the cell cycle or, perhaps more plausibly, that repair is efficient but inaccurate [29, 30], mutations that arise from damage will not track cell divisions [27]. Damage-induced mutations must underlie the observed maternal age effect on DNMs in humans; they could also account for the accumulation of germline mutations in proportion to absolute time in males, assuming fixed rates of damage and repair machinery errors in germ cells.

Multiple lines of evidence have emerged in support of damage-induced mutations being predominant in the human germline. Analyses of the mutation spectrum in humans indicate that 75% of DNMs and 80% of mutations in adult seminiferous tubules are due to mutation "signatures" SBS5/40 [31, 32], which are clock-like, uncorrelated with cell division rates in the soma [33, 34], and prevalent in post-mitotic tissues [30, 35]. In addition, most substitutions in post-pubertal germ cell tumors are attributed to SBS5/40, in both females and males [36]. More generally, cell division rates do not appear to be a major determinant of mutation rates across somatic tissues [37]: notably, post-mitotic neurons accumulate mutations at a similar rate as granulocytes, which are the product of continuous cell divisions [30]. A decoupling between cell division numbers and mutation burden has also been described in colonic crypts across mammals [38], and in yeast, up to 90% of mutations have been estimated to be non-replicative in origin [39]. Altogether, these results suggest an important role, for both germline and soma, of mutagenic processes that accumulate with absolute time, as expected from damage-induced mutations [27].

In undermining the prevailing understanding of the paternal bias in human germline mutations, these observations revive the question of how the bias arises, as well as of the influences of life history traits and exogenous or endogenous environments. To investigate them, we took a broad taxonomic view, characterizing the paternal mutation bias across amniotes, including mammals but also birds and snakes, which differ in potentially salient dimensions. As two examples, in birds as in mammals, oogenesis is arrested by birth in females, while spermatogenesis is ongoing throughout male reproductive life [40, 41], but birds have internal testes whereas some mammals have external testes. In addition, mammals and birds are endotherms, in contrast to ectothermic reptiles such as snakes. More generally, the taxa considered vary widely in their life histories, physiologies, and natural habitats.

2 Results

2.1 Estimating sex differences in germline mutation rates across amniotes

To estimate α in each lineage, we based ourselves on the evolutionary rates at putatively neutrally-evolving sites of sex chromosomes compared to the

autosomes [42]. The more direct approach of detecting *de novo* mutations in pedigrees requires them to be available for each species, and in large numbers for the estimates not to be imprecise. In contrast, the evolutionary method is in principle applicable to any set of species with high quality genome assemblies and a stable sex karyotype. It takes advantage of the fact that at the population level, sex chromosomes spend different numbers of generations in each sex (e.g., the X chromosome spends twice as many generations in females as in males), whereas autosomes spend an equal number in both (Figure 1A). Thus, all else being equal, if there is a paternal mutation bias, an autosome with greater exposure to the more mutagenic male germline will accumulate more neutral substitutions than the X over evolutionary timescales (Figure 1A); the inverse will be true for the autosomes compared to the Z chromosome [42].

Such evolutionary approaches have been widely applied, but until recently they were limited in the number of loci or species (e.g. [6–8, 43–45]) and did not take into account the influence of sex differences in generation times on the estimation of α [23]. An additional complication to consider is that X (Z) and autosomes differ not only in their exposures to male and female germlines but in a number of technical and biological features (notably, GC content) that may need to be controlled for [46–48]. Moreover, analyses involving closely related species can be confounded by the effects of ancestral polymorphism: for example, lower ancestral diversity in the X chromosome relative to the autosomes reduces the X-to-autosome divergence ratio, leading to overestimation of α [5] (Figure 1B). In birds, unresolved branches within the phylogeny present an additional difficulty in estimating substitution rates [49, 50].

Here, we designed a pipeline for estimating the paternal mutation bias systematically across a wide range of species, mindful of these issues. To these ends, we employed existing whole genome alignments [51, 52] or produced our own (for snakes, see Sequence alignments in Materials and Methods), focusing on assemblies with high quality and contiguity and, where possible, those based on a homogametic individual. To handle the confounding effects of ancestral polymorphism on divergence, we thinned species in the phylogeny to ensure a minimum level of divergence between them, relative to polymorphism levels (see Species selection criteria in Materials and Methods). This stringent filtering procedure resulted in three whole genome alignments including 20 mammals, 17 birds and five snake species, respectively (Table S2).

In order to estimate neutral substitution rates from the alignments and compare X (Z) and autosomes while minimizing confounding factors, we focused on non-repetitive, non-exonic regions that were orthologous across all species in an alignment and did not overlap with pseudo-autosomal regions with orthologs on the Y (W) chromosome (see Selecting non-repetitive and putatively neutral sequences in Materials and Methods; see Figure S1F for a more stringent masking of all conserved regions). To account for differences between X (Z) and autosomes in features other than their exposure to each sex, we regressed putatively-neutral substitution rates in the 1Mb genomic windows against GC content and GC content squared (Figure 1B). We took this

approach because GC content is readily obtained from any genome sequence and is highly correlated with known modifiers of the mutation rate such as replication timing and the fraction of CpG dinucleotides [48, 53]. We then obtained substitution rate estimates for the X (Z) chromosome and autosomes from the regression fit. Finally, we inferred α for the terminal branches leading to the 42 amniote species from the ratio of the substitution rate estimates for the X (Z) versus the autosomes (Figure 2), taking into account sampling error as well as uncertainty in the ratio of paternal-to-maternal generation times [54] (see Estimating α from X-to-autosome substitution rate ratios in Materials and Methods).

Overall, our evolutionary-based estimates, $\hat{\alpha}_{\text{evo}}$, are consistent with estimates from pedigree sequencing studies, $\hat{\alpha}_{\text{dnm}}$ (Figure 2). Notably, and reassuringly, the point estimates for species with the largest amount of available DNM data (e.g., humans, mice and cattle) are in very close agreement. Importantly, this finding is not necessarily expected, as $\hat{\alpha}_{\text{evo}}$ is an average over many thousands of generations of evolution, whereas estimates from DNMs are based on small numbers of families at present. In principle, differences between the estimates could therefore arise if α evolves rapidly (as may have happened in the lineage leading to macaque), or if the ages of the parents in the pedigree are quite unrepresentative of average paternal-to-maternal generation times in evolution (Figure 2) [54]. Disagreement between the two estimates could also arise from mutation rate modifiers that differ between sex chromosomes and autosomes. For example, the low $\hat{\alpha}_{\text{evo}}$ compared to $\hat{\alpha}_{\text{dnm}}$ in cats [55] could be due to unusual features of the X chromosome. In this respect, we note that the feline X chromosome is known to harbor a large recombination coldspot spanning over 50Mb [56], visible in its effects on GC substitution rates (Figure S2) [57], features that may have influenced the rate of substitution of the X chromosome relative to the autosomes. Overall, the general concordance between $\hat{\alpha}_{\text{evo}}$ and $\hat{\alpha}_{\text{dnm}}$ suggests that, with the possible exception of cats, the evolutionary approach is providing reliable estimates and the paternal bias in mutation is not rapidly-evolving.

2.2 A paternal bias in mutation is widespread in amniotes

A paternal bias in mutation is seen across amniotes, with a range of 1–4 in the species considered (Figure 2). The $\hat{\alpha}_{\text{evo}}$ estimates remain similar if we exclude hypermutable CpG sites (Figure S1B), or focus only on mutation types that are not subject to the effects of GC-biased gene conversion (gBGC) (Figure S1F and Figure S3). More generally, $\hat{\alpha}_{\text{evo}}$ are robustly above 1 and their ranking across species remains similar for different choices of conservation filters (e.g., excluding all conserved regions, not just exons) and different substitution types (see Figure S1 for details). The robustness of $\hat{\alpha}_{\text{evo}}$ across conditions and filters suggests that, while our pipeline may not account for all the differences between autosomes and X (Z) chromosomes unrelated to sex differences in mutation, the qualitative patterns are reliable. These results therefore establish that the paternal bias in mutation is not a feature of long lived humans

or of mammals, but is instead ubiquitous across species that vary markedly in their gametogenesis, physiology and life history.

The effects of gBGC track recombination rates and result in greater selection for GC in regions of higher recombination. Therefore, if α is similar for different types of de novo mutations, as has been found in humans [18, 21], the greater population recombination rate of autosomes relative to the sex chromosomes should lead the the X-to-autosome substitution rate ratio of gBGC-favored mutation types (T>C and T>G) to be somewhat lower than that of mutation types unaffected by gBGC (C>G and T>A). Consistent with expectation, $\hat{\alpha}_{\text{evo}}$ estimates in mammals using only gBGC-favored mutation types are inflated relative to estimates from mutation types unaffected by gBGC (Figure S3). Also as expected, bird and snake species with ZW sex determination exhibit the opposite pattern (i.e., a deflated ratio of Z-to-autosome substitution rate leads to a decreased estimate of $\hat{\alpha}_{\text{evo}}$; Figure S3). The behavior of the different mutation types therefore provides a further sanity check on our estimates. While the estimation of $\hat{\alpha}_{\text{evo}}$ could be further partitioned into single mutation classes, such estimates are noisier and–given the lack of ground truth–harder to interpret; we therefore focused on α for all substitution types combined.

Within mammals, the mean value of $\hat{\alpha}_{\text{evo}}$ is 2.7, with a range 1.0 to 4.1 and a coefficient of variation of 0.29. In birds, $\hat{\alpha}_{\text{evo}}$ is lower on average but also seemingly more stable, ranging from 1.5 to 2.7 (mean = 1.8, coefficient of variation = 0.19). In the handful of snake species sampled, the mean is similar to that of birds and $\hat{\alpha}_{\text{evo}}$o ranges from 1.3 to 2.2 (mean = 1.7, coefficient of variation = 0.23), in agreement with a previous evolutionary estimate for rattlesnake ($\alpha = 2.0$; [58]).

In mammals, variation in α has long been known to be associated with generation times, and has been consistently interpreted as resulting from greater numbers of replication errors in species with longer-lived fathers (e.g., [4, 9, 23, 59]). We confirmed the observation here: after accounting for the phylogenetic relationship between species, mammals reproducing at older ages show a stronger paternal bias in mutation (p-value = 0.01, $r^2 = 29\%$; Figure 3). Statistically significant relationships also exist between $\hat{\alpha}_{\text{evo}}$ and other life history traits (Figure S4), but these traits are strongly correlated with one another (Figure S5) and generation time is the strongest single predictor (Figure S4; see Testing relationships between α and life history traits in Materials and Methods). In contrast, a significant relationship between generation time and $\hat{\alpha}_{\text{evo}}$ is not seen in birds (p-value = 0.30, $r^2 = 7\%$; Figure 3; [60]), despite similar numbers of species and a comparable range of generation times to mammals. Moreover, we could reject the null model of a slope in birds equal to or greater than that of mammals (p-value = 10^{-5}). (Given the paucity of generation time and α estimates for snakes, we could not test the relationship in reptiles.) In light of more recent evidence that most mutations depend on absolute time and not cell division rates, the standard explanation for this generation time effect no longer holds. These observations therefore

raise the question of how else the relationship between generation times and α in mammals can be explained.

2.3 A cell-division-independent explanation for the correlation between α and generation time

In eutherian mammals, embryo development is likely independent of sex until primordial germ cell (PGC) specification and subsequent development of the gonads [61]. As a result, mutations arising during early embryogenesis (*Early*) are expected to occur at a similar rate in males and females ($\alpha_{Early} = 1$), as has been inferred in the few pedigree studies in which DNMs during parental early embryogenesis are distinguished from mutations later in development, namely in humans [62], cattle [26] and mice [25] (Figure 4A). While sex differences in early development may exist [63], differences in male and female mutation rates at such an early stage are likely modest in mammals [64, 65]. At some point after sexual differentiation of the germline, however, (in what we term the *Late* stage) mutation rates in the two sexes need no longer be the same: sources and rates of DNA damage could differ between germ cells, as could the efficiency and accuracy of repair. Indeed, human fathers that recently reached puberty contribute over three times more mutations than similarly aged mothers [21]. Intriguingly, the magnitude of paternal bias for mutations that occurred long after sexual differentiation of the PGCs appears to be similar in mice, cattle and humans, at approximately 4:1 [25, 26, 62] (Figure 4A).

In light of these observations, we considered a simple model in which α in mammals is the outcome of two developmental stages with distinct ratios of paternal-to-maternal mutations. In the *Early* stage until germline sex differentiation, we assumed a paternal-to-maternal mutation ratio of 1 and an expected number of mutations (M_e) on par with what is observed in humans (i.e., 5 mutations per haploid genome; [62, 66, 67]) (Figure 4B). In the *Late* developmental stage after germline sex differentiation, which varies in length among species, we assumed mutations arise at a constant rate per year, μ_s in sex s ($s \in \{f, m\}$). If we assume the length of *Early* to be negligible relative to the generation time, G_s in sex s, then the expectation of α can be written as:

$$\alpha = \frac{M_e + \mu_m G_m}{M_e + \mu_f G_f}. \quad (1)$$

If the ratio μ_m/μ_f is 4 across species, as suggested by DNM data [25, 26, 62] (Figure 4A and Table S3), this model yields a relationship between α and generation time bounded below by 1 and with a plateau at 4, assuming the same generation times in the two sexes (Figure 4C); more generally, the height of the plateau depends on the ratio of paternal-to-maternal generation times (Figure S6). The rapidity with which α reaches this asymptote is determined

by the magnitude of μ_m (and μ_f) in the *Late* stage (Figure 4C). Most pertinent, a positive relationship between α and the sex-averaged generation time is expected as long as $\mu_m G_m > \mu_f G_f$.

Using this model, we then predicted α for the terminal branches in the mammalian tree. To estimate the number of mutations occurring in *Late* for each branch, we used the evolutionary rates in Figure 2A. Specifically, we calculated a sex-averaged substitution rate per generation by multiplying the autosomal yearly substitution rate in each branch (μ_y) by a generation time estimate for its tip (Table S2). Given a fixed ratio of paternal-to-maternal mutation rates of 4 in the *Late* stage, the mutation rate for each sex can be calculated for any given ratio of paternal-to-maternal generation times:

$$\mu_f = \frac{\mu_y \left(G_f + G_m\right) - 2M_e}{G_f + 4G_m}. \tag{2}$$

From the parental mutation rates and assuming a fixed M_e, we obtained an estimate of α that we can use to predict $\hat{\alpha}_{\text{evo}}$ using equation 1 (see Modeling the effects of germline developmental stages on α in Materials and Methods). This model explains a significant proportion of the variance in $\hat{\alpha}_{\text{evo}}$ in mammals ($r^2 = 37\%$; p-value = 0.005; Figure 4D). After taking into account sampling error in our $\hat{\alpha}_{\text{evo}}$ estimates (see Modeling the effects of germline developmental stages on α in Materials and Methods)), it explains 42% of the variance in α across species. Moreover, the fit of the model remains good regardless of the precise number of *Early* mutations assumed (see Modeling the effects of germline developmental stages on α in Materials and Methods). The two clear outliers are carnivores, for which $\hat{\alpha}_{\text{evo}}$ may be an underestimate, given the higher estimate from DNMs in cats (Figure 2).

These predictions rely on evolutionary estimates that are uncertain, due for instance to inaccuracies in split time estimates and the use of contemporary generation times as proxies for past ones. If we instead predict α using parameters derived from pedigree data in the nine mammalian species for which at least 30 DNMs have been phased and more than one trio has been studied (Modeling the effects of germline developmental stages on α in Materials and Methods), the model explains 86% of the variance in $\hat{\alpha}_{\text{dnm}}$ (p-value $= 3 \times 10^{-4}$; Figure 4C). We caution that this assessment is based on few phylogenetically-independent contrasts, however, and so while the fit of the model again appears quite good, the variance explained may be deceivingly high.

In any case, this phenomenological model clarifies that the increased α seen in long-lived mammals may simply reflect a reduction in the fraction of early embryonic mutations relative to total number of mutations per generation–consistent with the higher proportion of *Early* mutations in mice and cattle compared to humans (Figure 4A). This model can also explain the only modest increase in α with parental ages observed in humans [21].

Given this explanation for the effect of generation times on α in mammals, why is a relationship not seen in birds (Figure 2)? One interpretation is simply

a lack of statistical power: since the ratio of paternal-to-maternal age effects in the *Late* stage is lower in birds than in mammals (e.g., 2 instead of 4), under our model, bird generation times would influence α within a narrower range (i.e., between 1 and 2). Alternatively, the lack of a relationship between α and generation times in birds could reflect their distinct germ cell development: Unlike mammals, avian sexual phenotype is directly determined by the sex chromosome content of individual cells [68, 69] and PGCs are determined by inheritance of maternally derived gene products [70]. Given these features, it seems plausible that sex differences in mutation rates appear earlier in ontogenesis in birds than in mammals, consistent with reported sex differences in the cellular phenotypes of PGCs prior to gonad development [71]. If indeed the developmental window when both sexes have a similar mutation rate is short in birds then, under our model, generation times are expected to have little to no influence on α.

2.4 Discussion

Analyzing diverse species with the same pipeline, we found that, far from being a feature of species with long-lived males, a paternal bias in germline mutation is ubiquitous across amniotes that differ markedly in their life history, physiology and gametogenesis. Moreover, by considering the different development stages over which germline mutations arise, we provide a new and simple explanation for variation in the degree of sex bias across mammals that does not require dependence on the number of cell divisions. While our findings do not account for *why* male germ cells might accumulate more mutations than female ones, the observation that paternal bias varies little across species exposed to disparate physical environments, and presumably exogenous mutagens, hints at sex differences in endogenous sources of DNA damage or repair (e.g., [72]). Another question raised by our findings is why, after sexual differentiation of the germline, mutation appears to be more paternally-biased in mammals (∼4:1) than in birds and snakes (∼2:1).

More generally, our results recast long standing questions about the source of sex bias in germline mutations as part of a larger puzzle about why certain cell types (here, spermatogonia versus oocytes) accrue more mutations than others. Intriguingly, the relative mutagenicity of different tissues appears to be conserved across species: for instance, in mammals, the balance of damage and repair results in an approximately four-fold increase in mutation rates per unit of time in spermatogonia compared to oocytes (Figure 4A). Similarly, comparing yearly mutation rates in colonic crypts [38] to estimates for spermatogonia, the ratio of crypt-to-sperm mutation rate appears relatively stable across four mammalian species (Figure S7).

These findings raise the possibility that mutation rates in different cell types are somehow coupled across mammalian species, either because of natural selection to maintain specific rates in each cell type, or because changes to the repair machinery in one tissue have pleiotropic consequences on mutation rates in other tissues. Regardless, our observations point to a role of stabilizing

selection in maintaining the relative rates at which mutations accumulate in different tissues over evolutionary timescales.

3 Figures

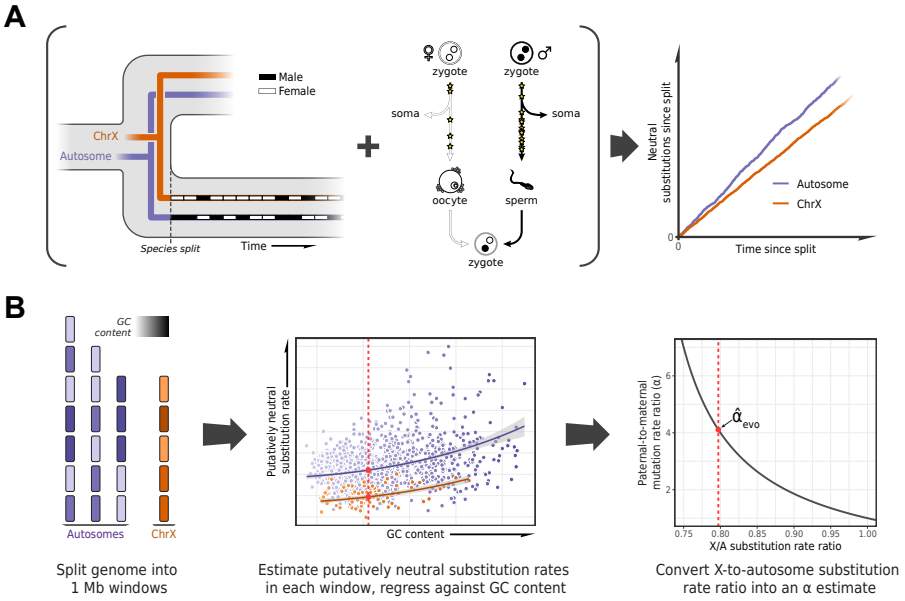

Fig. 1: **Estimating the paternal bias in mutation from neutral substitution rates of sex chromosomes and autosomes.** (A) On average, the lineage of an X chromosome spends fewer generations in males than females. Given a higher mutation rate in males than in females and all else being equal, this leads to lower rates of neutral substitutions on the X chromosome compared to autosomes [42]. (B) Procedure for estimating the ratio of paternal-to-maternal mutation rates, α, from substitution rates in sex chromosomes and autosomes. The autosomes and the X chromosome are partitioned into 1 Mb windows, depicted in purple and orange, respectively. Each window is filtered to focus on putatively neutrally-evolving sequences (see Selecting non-repetitive and putatively neutral sequences in Materials and Methods), and its GC content is calculated (represented by shading). The putatively neutral substitution rates per window are then regressed against the GC content (center panel, see Estimating α from X-to-autosome substitution rate ratios in Materials and Methods). Substitution rate estimates for the X chromosome and autosomes are obtained from the regression fit (red points). Finally, the ratio of the point estimates is converted to an estimate of α (right panel). An analogous procedure applies to comparisons of the Z chromosome and autosomes in a ZW sex-determination system.

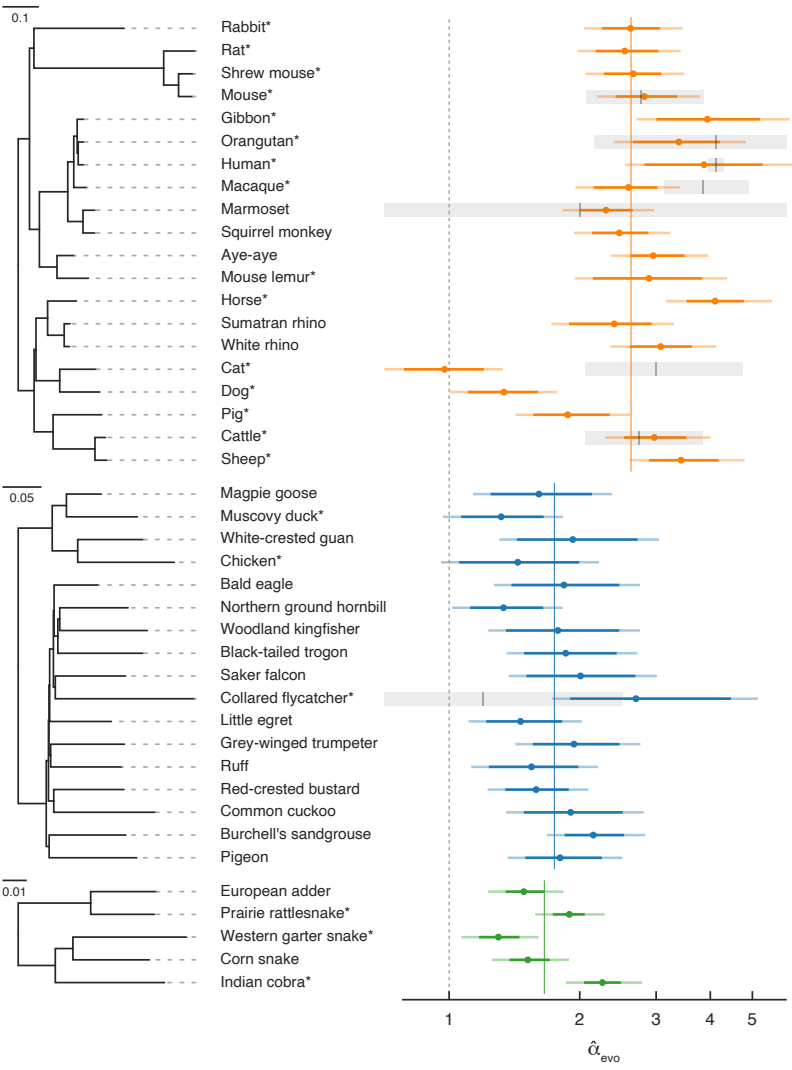
0.1
Rabbit*
Rat*
Shrew mouse*
Mouse*
Gibbon*
Orangutan*
Human*
Macaque*
Marmoset
Squirrel monkey
Aye-aye
Mouse lemur*
Horse*
Sumatran rhino
White rhino
Cat*
Dog*
Pig*
Cattle*
Sheep*
0.05
Magpie goose
Muscovy duck*
White-crested guan
Chicken*
Bald eagle
Northern ground hornbill
Woodland kingfisher
Black-tailed trogon
Saker falcon
Collared flycatcher*
Little egret
Grey-winged trumpeter
Ruff
Red-crested bustard
Common cuckoo
Burchell's sandgrouse
Pigeon
0.01
European adder
Prairie rattlesnake*
Western garter snake*
Corn snake
Indian cobra*
1
2
3
4
5
$\hat{\alpha}_{evo}$

Fig. 2 *(previous page)*: **Estimates of the paternal bias in mutation across 42 amniote lineages.** Colored points denote estimates of α from X(Z)-to-autosome substitution rate ratios, $\hat{\alpha}_{evo}$, in mammals (top, orange), birds (middle, blue), and snakes (bottom, green). Note that the x-axis is log-scaled, and CIs extending past 6 are truncated (see Table S2). Vertical colored lines denote the mean $\hat{\alpha}_{\text{evo}}$ for each group and the vertical gray dotted line denotes $\alpha = 1$ (i.e., no sex bias in mutation). Species in each group are plotted according to their phylogenetic relationships, with branch lengths scaled by the neutral substitution rate (see the legends for the divergence per base-pair). Asterisks indicate species with chromosome-level assemblies. Points denote the point estimate of $\hat{\alpha}_{\text{evo}}$. Darker colored horizontal lines represent 95% CIs, which include variation in substitution rates across genomic windows; lighter colored horizontal lines incorporate additional uncertainty in the ratio of paternal-to-maternal generation times, assuming it ranges from 0.9 to 1.1 (see Estimating α from X-to-autosome substitution rate ratios in Materials and Methods). Short vertical black lines denote point estimates of α from published pedigree studies of *de novo* mutations, $\hat{\alpha}_{\text{dnm}}$, and the surrounding horizontal gray boxes represent their 95% binomial CIs. For more detail, see the extended Figure 2 legend in Supplementary tables and figures and Materials and Methods.

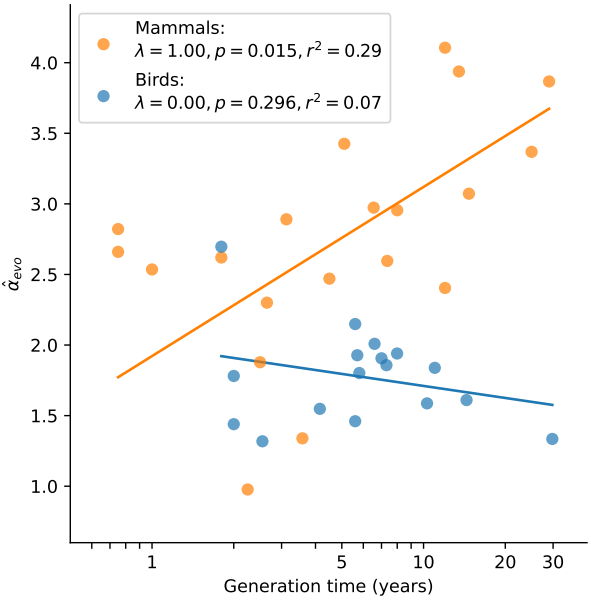

Fig. 3: **Relationship between $\hat{\alpha}_{\text{evo}}$ and generation time estimates in mammals and birds.** Estimates of α from X (Z)-to-autosome comparisons are plotted against generation times from the literature (see Table S2), on a log scale. Lines denote the phylogenetic generalized least squares regression fits in mammals (orange) and birds (blue). λ refers to Pagel's λ [73], a measure of the strength of phylogenetic signal, which was inferred via maximum likelihood (see Testing relationships between α and life history traits in Materials and Methods). Fixing λ to 1 in birds, as estimated for mammals, did not meaningfully improve the fit (p-value = 0.282, $r^2 = 0.08$).

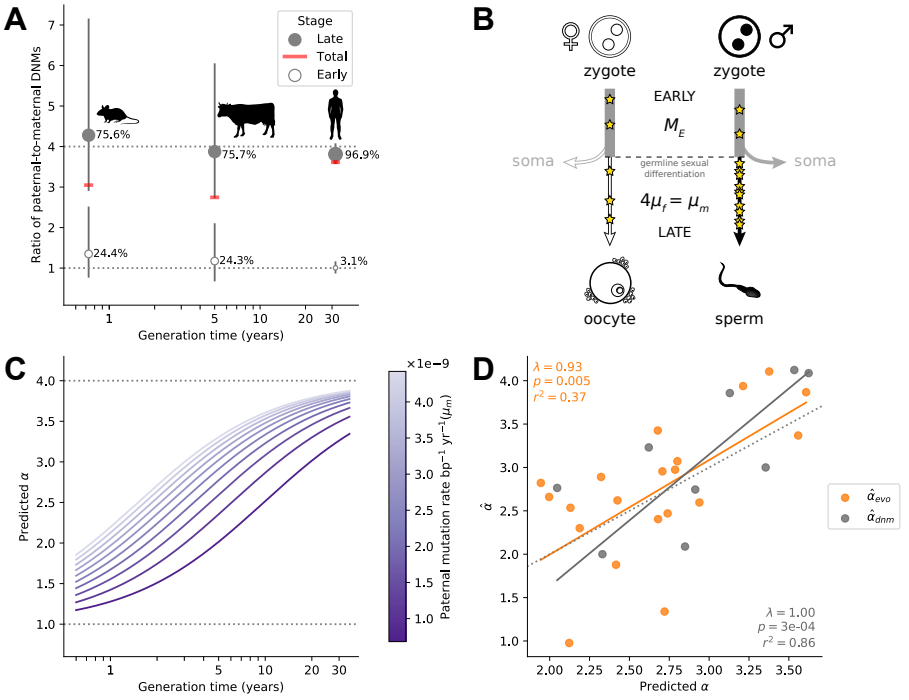
A
Stage
Late
Total
Early
Ratio of paternal-to-maternal DNMs
75.6%
75.7%
96.9%
24.4%
24.3%
3.1%
Generation time (years)
B
zygote
zygote
EARLY
M_E
soma
soma
germline sexual differentiation
$4\mu_f = \mu_m$
LATE
oocyte
sperm
C
Predicted α
×1e−9
Paternal mutation rate bp^{-1} yr^{-1} (μ_m)
Generation time (years)
D
$\lambda = 0.93$
$p = 0.005$
$r^2 = 0.37$
$\hat{\alpha}_{evo}$
$\hat{\alpha}_{dnm}$
$\lambda = 1.00$
$p = 3e\text{-}04$
$r^2 = 0.86$
$\hat{\alpha}$
Predicted α

Fig. 4 *(previous page)*: **Variation in α among mammals may reflect varying exposures to different developmental stages.** (A) Ratio of paternal-to-maternal *de novo* mutations (DNMs) occurring in early embryogenesis (*Early*, white points), after the sexual differentiation of the germline (*Late*, grey points) and in both of these stages combined (*Total*, red line), for the three mammalian species in which this classification is available (mouse [25], cattle [26], and human [62]). For each species, the percentage of DNMs occurring at each stage are indicated and used to scale the size of points. Vertical lines show the 95% binomial CIs. Since the phasing rate is not equal across developmental stages, point estimates for α in *Total* were computed by extrapolating the proportion of paternally and maternally phased DNMs in each stage to all the DNMs in that stage (i.e., assuming full phasing)(see Estimating α from pedigree studies in vertebrates in Materials and Methods). (B) Schematic representation of a model in which α is the outcome of mutation in two developmental stages (see Modeling the effects of germline developmental stages on α in Materials and Methods). (C) Expected relationship between α and generation time under the model outlined in B, assuming generation times are the same in both sexes. The increase of α with generation time depends on the paternal mutation rate per year in *Late*, μ_m, as illustrated by the purple gradient. (D) Fits of predicted α values to $\hat{\alpha}_{\text{evo}}$ (orange) and $\hat{\alpha}_{\text{dnm}}$ (gray). In each species, α is predicted with equation 1 assuming $M_e = 1.66 \times 10^{-9}$ and using μ_f and μ_m, the latter estimated from autosomal branch-specific substitution rates per year ($\hat{\alpha}_{\text{evo}}$) or as estimated from pedigree sequencing data ($\hat{\alpha}_{\text{dnm}}$) (see Modeling the effects of germline developmental stages on α in Materials and Methods). The orange and gray lines denote the regression fit using phylogenetic generalized least squares (PGLS). PGLS statistics are shown for the two models (see Figure 3 legend for details).

4 Materials and Methods

4.1 Sequence alignments

In mammals, we obtained sequence alignments from the 241-way multi-alignment generated by the Zoonomia Project (https://zoonomiaproject.org/) [51]. To assess the effect of reference sequence selection on our α estimates, we considered two alignments, one using the *Homo sapiens* genome as reference sequence, and the other using the *Mus musculus* genome as reference (Figure S1A).

In birds, we subdivided the 363-way alignment generated by the B10K project (https://b10k.genomics.cn/) [52] into six subgroups, avoiding the inclusion of ancestral nodes with high uncertainty within Neoaves [49, 52, 74]. Since a species topology is required to accurately infer branch-specific substitution rates, we built species sets by combining monophyletic groups that are well supported across data types and studies [50] (Table S4). In all cases, we used the *Gallus gallus* genome as the reference sequence.

In snakes, we built our own multiple genome alignments using whole genome assemblies downloaded from the National Center for Biotechnology Information (NCBI) database (Table S1). To speed up computation, we removed repetitive regions—which are ignored in all downstream analyses—from the whole genome FASTA files prior to alignment by converting lowercase bases (i.e., a, t, c, g) to N bases. We ran the Cactus program (version 1.2.5, https://github.com/ComparativeGenomicsToolkit/cactus) to align the genomes in each clade using topologies generated by TimeTree as our guide trees (see `trees/Snakes.TimeTree.nwk` at https://github.com/flw88/mut_sex_bias_amniotes/). For subsequent analyses, we used *Thamnophis elegans* as the reference sequence in snakes.

For each taxon, we converted the HAL file into a Multiple Alignment Format (MAF) file and split the alignment into non-overlapping windows of 1 Mb using the `hal2maf` tool in halTools (https://github.com/ComparativeGenomicsToolkit/hal/):

```
hal2maf $hal $maf --targetGenomes $species_list --refGenome \
  $reference --refSequence $reference_chrom --start $start \
  --length $end-$start --onlyOrthologs --noDupes --noAncestors
```

4.2 Species selection criteria

To estimate α, we aimed to measure differences in the rates of neutral substitution in X (Z) versus autosomes that are directly attributable to differences in the mutation rate of males and females. However, X (Z) and autosomes also differ in a number of other technical and biological features that must first be taken into account.

One important source of technical bias is the unequal sequence coverage of the X (Z) and autosomes in heterogametic individuals. To minimize any potential issues due to systematic differences in assembly quality between X

(Z) and autosomes, we excluded non-chromosome level genomes known to be assembled exclusively from DNA of the heterogametic sex. In addition, we discarded any species belonging to a genus in which a complex system of chromosomal sex determination has been identified (annotated as "complex XY" or "complex ZW" in the Tree of Sex database https://coleoguy.github.io/tos/data.vert.csv, with the exception of the *Mus* genus). In mammals, out of a total of 241 genomes, this approach led us to exclude 50 male-based assemblies and nine species with at least one case of complex XY in the same genus. In birds, out of a total 363 genomes, we excluded 186 female-based assemblies and two species with at least one species with a complex ZW in the same genus.

The quality of the genome assembly is an additional potential confounder. Given that we relied on higher quality, chromosome-level assemblies to categorize alignments as X (Z) or autosomal, we would be more likely to miscategorize alignments (i.e., as X/Z or autosomal) in species with lower quality genome assemblies that are highly diverged from the nearest chromosome-level assembly. To address this issue, in mammals, we removed species if their genomes were $> 15\%$ diverged from the nearest chromosome-level assembly. We relaxed the divergence threshold to 30% in birds, in which fewer genomes are assembled at chromosome-level and across which karyotypes are believed to be relatively stable [75]. In both mammals and birds, we relied on published divergence estimates inferred from the same multi-alignments used in this study (see http://cgl.gi.ucsc.edu/data/cactus/241-mammalian-2020v2.phast-242.nh [51] and http://cgl.gi.ucsc.edu/data/cactus/363-avian-2020-phast.nh [52]). We also discarded species with low quality scaffold-level assemblies, i.e., where scaffold N50 < 350 kb and contig N50 < 25 kb. These filters led to the removal of 120 and 76 species in mammals and birds, respectively.

Given the paucity of genomes in snakes, we relaxed our filtering criteria to allow the inclusion of a larger number of species. Specifically, we allowed scaffold-level assemblies from the heterogametic sex and reduced the scaffold and contig N50 thresholds to 100 kb and 10 kb, respectively. These changes allowed the inclusion of *Vipera berus* and *Pantherophis obsoletus*. We estimated divergence between species using `phyloFit` (see Estimating putatively-neutral substitution rates) in the largest chromosome in *Thamnophis elegans* (NC_045541.1). As in mammals, we removed any species with distance to nearest chromosome-level assembly $> 15\%$ and confirmed that none of the species belong to a genus with a complex ZW system in the Tree of Sex database. This procedure excluded one of the nine snake species (*Laticauda laticaudata*).

Another important consideration comes from the differing evolutionary histories of sex chromosomes and autosomes. Under neutrality and assuming equal variances in reproductive success, the X (Z) chromosome is expected to have a lower effective population size, N_e, than the autosomes [76]. For closely related species, this implies a deeper coalescence time of autosomes than X (Z) in their ancestral population and therefore an unequal contribution of ancestral polymorphisms to the substitution rates; e.g., if $N_e^{\mathrm{X}} < N_e^{\mathrm{A}}$, then the X-to-autosome substitution rate ratio will be deflated relative to the

expectation under mutational male bias alone, and consequently α will be overestimated [5]. To minimize this problem, we sought to keep a subset of species that were sufficiently distantly related such that the contribution of ancestral polymorphism to divergence is small and the bias in α estimates is negligible. Specifically, we proceeded as follows: under simplifying assumptions, the expected neutral divergence attributable to ancestral polymorphisms is given by the heterozygosity, π, in the ancestral species. Since π in the ancestral population of a species pair is unknown, we used estimates for π from present-day species as a proxy. We pruned the phylogeny of each taxon so to retain only species pairs with a combined (summed) substitution rate of at least 15π, where π is the higher value of the pair.

We collected mammalian π estimates from the individual heterozygosities in the Zoonomia Project ("Overall heterozygosity" in Supplementary Table 3 in [51]), complemented with the nucleotide diversities in [77] ("log10_diversity" in `data/combined_data.tsv` at https://github.com/vsbuffalo/paradox_variation/ commit 14366fe), obtaining π values for 16 of the remaining mammalian species. For any species lacking a value in both databases, we assigned the π of the closest species in the mammalian phylogeny as inferred with PHAST (45 species). Finally, in the one case in which π from both databases were available (*Daubentonia madagascariensis*), we took the average π.

In birds, we used π estimates in [78], obtaining direct estimates for 13 of the remaining species. For species not present in the database, we assigned the π of the closest species in the bird phylogeny (85 species). In snakes, we collected π values from the literature (Table S1).

Because initiatives like the Zoonomia Project or B10k may preferentially select species at risk of extinction [51], some of the present-day π values may underestimate the diversity levels in the ancestor. We thus set an extra requirement of at least a combined 2% substitution rate between any pair of species. In species pairs where the rate was below either of these two thresholds (15π or 2%), we preferentially retained the species that met the following criteria, considered in this order: (I) more phased *de novo* mutation count data from pedigree sequencing (count of 0 if not available), (II) a chromosome-level assembly, and (III) a higher scaffold N50. Altogether, 20 out of 241 mammalian species, 17 out of 363 bird species, and five out of 9 snakes species remained after the complete filtering procedure.

A list of the species kept after filtering, together with other genome statistics and results from our analyses, can be found in Table S2. The code to reproduce the filtering procedure described above can be found in `notebooks/Filter_species.ipynb` at https://github.com/flw88/mut_sex_bias_amniotes.

4.3 Selecting non-repetitive and putatively neutral sequences

In the absence of natural selection and/or GC-biased gene conversion (gBGC), the substitution rate is equal to the mutation rate [79]. To minimize the effects of selection, we limited our analyses to non-coding regions by removing all exons annotated in the given reference sequence as well as the 1 kb of sequence flanking each exon. As a check, we also estimated α in mammals and birds after masking conserved elements identified by `phastCons` [80] (http://hgdownload.cse.ucsc.edu/goldenpath/hg38/database/phastConsElements100way.txt.gz and http://hgdownload.cse.ucsc.edu/goldenpath/galGal6/database/phastConsElements77way.txt.gz, respectively). Since the $\hat{\alpha}_{\text{evo}}$ are similar (Figure S1F), we based our analyses on the larger data set based on masking only exons and their 1 kb flanking sequences.

The effect of gBGC mutation on the substitution process is analogous to that of selection for specific base pairs, in that the process increases the probability of fixation of strong (G/C) over weak (A/T) alleles [81]. To explore the effects of gBGC, we estimated specific rates for each single-nucleotide substitution type (see Estimating putatively-neutral substitution rates for details). To remove the effects of gBGC, we estimated α for the subset of mutation types that are not subject to gBGC (i.e., substitutions from strong to strong and weak to weak nucleotides) (Figure S3 and Figure S1G).

In addition, to ensure the high quality of the alignment data for analysis, we removed repetitive regions, keeping only those genomic positions at which the reference sequence in a given analysis group (mammals, birds, and snakes) carries an uppercase nucleotide.

4.4 Filtering idiosyncratic genomic regions

We excluded sequences aligned to known pseudo-autosomal regions (PAR) in the sex chromosomes, which have homologs on both X and Y (or both Z and W) and thus behave like autosomes in terms of their ploidy (see Table S5 for PAR definitions). For snakes, we aligned sequenced reads from a female *Thamnophis sirtalis* individual (NCBI accession SAMN02402779) to the *Thamnophis elegans* reference genome using BWA-MEM v0.7.17-r1188 (http://bio-bwa.sourceforge.net/), with default parameters. We removed PCR duplicates with the `markdup` tool in samtools v1.10 (http://www.htslib.org/) and calculated the mean depth of coverage along the Z and the largest autosome in 1 Mb windows using *mosdepth* (https://github.com/brentp/mosdepth). We then determined regions of the Z chromosome in which the depth of coverage was significantly different to that in the autosomes, assuming depth is Poisson distributed with λ equal to the mean depth in the autosome, potentially indicative of the region being in a PAR and having homologs on the W chromosome (Figure S8).

The genome of birds and snakes are organized into two types of autosomes, macro- and microchromosomes, which differ in their length, gene content,

density of hypomethylated CpG islands, recombination rates and replication timing [82]. Given the idiosyncrasies of microchromosomes, which may affect the substitution rate estimates [60], we excluded sequences aligned to microchromosomes in birds and snakes (chromosomes 10 to 28 in *Gallus gallus* and chromosomes 13 to 18 in *Crotalus viridis*). The fraction of base pairs in microchromosomes is relatively small, comprising 20% and 5.1% of the autosomal genome in *Gallus* and *Crotalus*, respectively. We checked that $\hat{\alpha}_{\text{evo}}$ are similar whether or not microchromosomes are excluded ($r > 0.9$ between $\hat{\alpha}_{\text{evo}}$ estimates obtained after excluding or including microchromosomes, in both birds and snakes, Figure S1E).

An additional concern is that genomic translocations between X (Z) and the autosomes could lead to sequence misclassification in species without a chromosome-level assembly. To alleviate this potential issue, we only kept sequences that exclusively mapped to chromosomes of the same kind (i.e., X or Z versus autosome) in all species for which chromosome-level assemblies were available. In other words, we removed all alignments in which chromosome-level assemblies indicated a mapping between an X (Z) sequence of one species with an autosomal or Y (W) sequence of another.

To summarize, each 1 Mb MAF file in each taxon was first filtered with the `maf_parse` tool in PHAST (http://compgen.cshl.edu/phast/), using a thinned set of species obtained as described in Species selection criteria and a BED file with the regions to be excluded as indicated by the reference genome (i.e., exons ±1 kb and the PARs, if known). The python scripts `filter_PARs_micros_CpGs.py` and `filter_species_gaps_maf_XYA.py` (available at https://github.com/flw88/mut_sex_bias_amniotes) were then used to filter any gaps, annotated PARs, as well as regions that mapped to known chromosomes of a different kind:

```
maf_parse --features $regions_to_exclude_bed -M $reference \
  --seqs $(cat $species_list_thinned) $maf | \
  python filter_PARs_micros_CpGs -p data/Species_to_PARs.tsv | \
  python filter_species_gaps_maf_XYA.py \
  -l $species_list_thinned -c data/Species_to_chromosomes.txt \
  -b $filtered_regions_bed -a > $filtered_maf
```

4.5 GC content and replication timing estimates

The framework provided by Miyata et al. [42] to infer α assumes that the generation time is the same for both sexes, as well as that the substitution rates on autosomes versus X (Z) are solely determined by the sex-specific mutation rates and the ploidy difference between sexes. However, other genomic features, such as GC content and replication timing, are known to differentially influence the mutation rate of sex-linked and autosomal chromosomes [48, 53]. To account for these differences, we collected measures of species-specific GC content. Specifically, for every filtered 1 Mb MAF in each

taxon, we calculated the fraction of G/C base-pairs in each genome with:

```
cat $filtered_maf |  \
    python gc_content_from_maf.py -s $species_list_thinned
```

We additionally obtained replication timing data in human embryonic stem cells from the UCSC genome browser (http://hgdownload.cse.ucsc.edu/goldenpath/hg19/encodeDCC/wgEncodeFsuRepliChip/wgEncodeFsuRepliChipH1hescWaveSignalRep1.bigWig). We converted the data from bigWig format to BED using bigWigToBedGraph (https://hgdownload.soe.ucsc.edu/admin/exe/linux.x86_64/bigWigToBedGraph) and lifted the coordinates from the hg19 reference genome to hg38 using the liftOver tool (http://hgdownload.soe.ucsc.edu/admin/exe/linux.x86_64/liftOver).

To explore the relationship between replication timing and substitution rates in humans, we calculated an average replication timing value across the unfiltered bases in each 1 Mb window of the mammalian alignment (*Homo sapiens* as reference). Specifically, we used the mean replication timing value weighted by the number of bases associated with each replication timing datum.

4.6 Estimating putatively-neutral substitution rates

To estimate putatively neutral substitution rates on X (Z) and autosomes, we used `phyloFit`, a program within the PHAST software suite [83, 84] (http://compgen.cshl.edu/phast/). For every 1 Mb window of aligned sequence in each taxon with $\geq$ 10 kb of sequence remaining after filtering, we estimated substitution rates using the general, unrestricted single nucleotide model (`--subst-mod UNREST`) with the expectation maximization algorithm with medium precision for convergence (`--EM --precision MED`). We also obtained the number of expected counts at each node for each substitution type (option `-Z`). For mammals and birds, we used the relevant tree topology defined in the Newick files in http://cgl.gi.ucsc.edu/data/cactus/; for snakes, we used a topology from TimeTree (http://timetree.org/, `trees/Snakes.TimeTree.nwk`). To avoid local maxima in the likelihood surface, we ran six independent `phyloFit` runs with random initialization of the parameters (option `-r`) and kept the replicate with the highest likelihood. We note that `phyloFit` estimates the expected substitution counts for type $A_1 > A_2$ by inferring the expected number of times allele A_1 is found at the internal node of a branch in the tree and allele A_2 is observed at the terminal node. However, the overall branch lengths are maximum likelihood estimates of the expected rate of substitution in continuous time along the branches. Thus, the rate of substitution estimated by summing substitution counts and dividing by the genome size is slightly smaller than the maximum likelihood branch-length estimate (as the latter allows back mutation but the former

does not include them).

```
phyloFit -r --EM --precision MED --subst-mod UNREST -Z \
  --msa-format MAF $filtered_maf --tree $newick \
  -e $phylofit_errors -o $phylofit_output
```

4.7 Estimating α from X-to-autosome substitution rate ratios

We took a regression approach to estimate α from ratios of X (Z)-to-autosome substitution rates. This approach allowed us to control for the effect of GC content g on the substitution rates (see Estimating putatively-neutral substitution rates). For each species, we performed a Poisson regression with a log link function on the number of substitutions Y_i in the terminal branch (as inferred from `phyloFit`):

$$\log\left[\mathrm{E}\left(Y_i \,\middle|\, n_i, x_i, g_i\right)\right] = \log(n_i) + \beta_0 + \beta_1 x_i + \beta_2 g_i + \beta_3 g_i^2 \qquad (3)$$

where the subscript denotes the ith window, n denotes the number of bases at which a substitution could have occurred, x is an indicator variable denoting whether the window is on the X (Z) or the autosomes, and the β variables denote the regression coefficients. Modeling the relationship between substitution rate and GC content as a quadratic function captures effects of hypermutable CpG sites via the squared term [85, 86]. Note that for the overall substitution count, the number of substitution opportunities n is the total number of sites left in the window after filtering; however, when applying the regression model to a specific substitution type $A_1 > A_2$, we only considered sites where the ancestral allele was inferred by `phyloFit` to be A_1 (or its complementary base, see Estimating putatively-neutral substitution rates).

We used the fitted regression models to estimate α in each species. To this end, we first obtained point estimates of the substitution rates on the X (Z) and autosomes calculated at the mean GC content values of the X (Z) windows. We then converted the resulting X(Z)-to-autosome substitution rate ratio to an estimate of α using Miyata's equations [42].

We note that this approach infers α from the ratio of the expectations of the X (Z) and autosomal substitution rates rather than the expectation of the ratios. To check whether that makes a difference, we re-estimated α in each species using a modified procedure in which we repeatedly sampled a pair of X (Z) and autosome windows with GC content values in a narrow range (mean GC content value of the X (Z) chromosome $\pm 1.5\%$) and calculated a X (Z)-autosome substitution rate ratio. Estimating α from the mean ratio across 1,000 resamples yielded highly similar estimates to those obtained from our regression approach ($r = 0.93$ across species, Figure S1).

To understand whether controlling for replication timing in addition to GC content might affect our α estimates, we modified Equation 3 to include an extra term for the average replication timing of each window t_i (see GC

content and replication timing estimates). We applied this modified regression framework to mammals and obtained X-autosome substitution rate ratios for each species at the mean GC content and replication timing values of the X windows. Converting the X-autosome substitution rate ratios to α estimates using Miyata's equations [42] yielded values that were highly similar to those obtained when controlling for GC content only ($r > 0.99$, see Figure S1B). Given the observed agreement and the lack of replication timing data for most species, in subsequent analyses, we relied on evolutionary estimates obtained from the regression model described in Equation 3.

To assess the uncertainty in our α estimates, we bootstrap resampled windows on the X (Z) and autosomes 500 times. For each replicate, we fit the regression model and calculated the X (Z)-to-autosome ratio as described above to obtain an empirical distribution from which we could compute the central 95% interval. We note that because of the functional form describing the relationship between α and the X (Z)-to-autosome substitution rate ratio (Figure 1A), confidence intervals on α tend to be wider at larger values of α. In other words, in the regime of large α, a small shift in the X (Z)-to-autosome substitution rate ratio will have a larger impact on the inferred α estimate. We implemented our regression and α estimation framework in the R script, `alpha_from_unrest.regression.R`.

Although ignored in the original Miyata et al. approach and subsequent applications (e.g., [23, 58, 60]), recent modeling work shows that sex differences in generation times can also affect the relative ratio of substitution rates on the X (Z) and autosome by altering the amount of time that a sex chromosome lineage spends in males versus females compared to autosomes [54]. Thus, sex differences in generation times modulate how sex biases in mutations are reflected in substitution rates of X (Z) versus autosomes. Unfortunately, sex-specific generation time estimates are rarely available for extant species, let alone ancestral lineages, and likely evolve over time. To incorporate uncertainty in sex differences in generation times, we re-computed our uncertainty intervals on α under the assumption that the male-to-female ratio of the generation times for any particular lineage lies between 0.9 and 1.1, using formulas derived by Amster and Sella [54].

4.8 Estimating α from pedigree studies in vertebrates

In order to obtain estimates of α from extant vertebrate species, we identified 15 *de novo* mutation (DNM) studies with published counts of parentally-phased DNMs [18, 22, 25, 26, 55, 87–94]. For each species in each study, we calculated point estimates of α by dividing the number of DNMs phased to the paternal chromosome by the number phased to the maternal chromosome (Table S2). We measured uncertainty by computing binomial confidence intervals on the proportion of all phased DNMs that were paternal and then converting the resulting interval bounds back to a paternal-to-maternal ratio.

From this list, we excluded one study from mouse lemur (*Microcebus murinus*), which reported an anomalously high mutation rate per year for

a primate species ($> 3.5 \times 10^{-9}$ per site) and unusually low rates of transitions at CpG sites [89]. The authors suggested C to T substitutions in the branch leading to mouse lemur occurred at a similar rate irrespective of their dinucleotide context (CpG or non-CpG), in contrast to what is seen in other primates [95]. However, analyzing our substitution data, we find the C>T substitution rate in mouse lemur to be over five-fold higher at CpG sites compared to non-CpG sites. Specifically, we estimated substitution rates from our filtered autosomal mammalian alignments as described in Estimating putatively-neutral substitution rates with the following modifications: (I) CpG islands, as defined in http://hgdownload.cse.ucsc.edu/goldenpath/hg38/database/cpgIslandExtUnmasked.txt.gz, were masked following [89]; (II) CpG dinucleotide substitution rates were estimated using a context-dependent model (`--subst-mod U2S`). This study also reports the weakest mammalian paternal bias in mutation described to date ($\alpha = 1.18$). This value is out of sync with reports for other primates and far from what we estimate from substitution rates, $\hat{\alpha}_{\text{evo}}$ (Figure 2 and Table S2). One possibility is that a substantial rate of false positive DNMs biased $\hat{\alpha}_{\text{dnm}}$ towards 1 (since errors are likely placed with equal probability on the maternal or paternal haplotype). Given the uncertainty surrounding how to interpret these DNM data, we do not include this $\hat{\alpha}_{\text{dnm}}$ in our analyses.

4.9 Estimating α for different developmental stages

DNM studies typically quantify the number of mutations in the offspring that are not found in some somatic tissue (usually blood) of the parents. This approach can mistakenly include DNMs that occurred in the early development of the offspring, as well as mistakenly exclude DNMs that occurred early in the development of the parents [27]. DNMs that occurred in early development of the parents can be distinguished by patterns of "incomplete linkage" with nearby informative constitutive heterozygous positions, as well as incomplete transmission to the offspring [26, 62]. Moreover, DNMs that occurred right after or during primordial germ cell specification (PGCS) will not be present in the soma of the parents but may be transmitted to multiple offspring [25, 62].

To examine if α varies across developmental stages, we considered studies that distinguish between DNMs in the early development of the parent (i.e., mutations detectable in the parental soma but showing patterns of "incomplete linkage", as well as DNMs transmitted to multiple offspring), versus DNMs that occurred in later stages after PGCS (i.e., not present in the parental soma and transmitted to a single offspring). Counts for early DNMs were obtained: in mice [25], where we counted the number of mutations phased to each parental haplotype in "Early Embryonic" and "Peri-PGC" categories (Supplementary Table 1 at https://doi.org/10.1038/s41467-019-12023-w); in cattle [26], where we counted mutations classified as "Sire Mosaic" or "Dam Mosaic" (Supplementary Table 1 at https://doi.org/10.1101/079863); and in humans [62], where we counted the number of mutations phased to each parental haplotype in "Gonosomal mutations" and "Post-PGCS" (Tables

in https://github.com/quinlan-lab/ceph-dnm-manuscript/tree/master/data). DNM counts for phases later in development were obtained from the same publications, under the categories "Late post-PGCS", "Sire/Dam non Mosaic", and "Third-generation" in mice, cattle, and humans, respectively. All three studies also employed strategies to discard DNMs in the early development of the offspring. The combined counts for each species and mutation timings can be found in Table S3.

Since the paternal bias in mutation varies among developmental stages, as does the fraction of mutations that were successfully phased (Table S3), simply summing over DNM counts from different stages would result in a biased point estimates of the overall α. We therefore computed α by extrapolating the proportion of paternally and maternally phased DNMs in each stage to all the DNMs identified in that stage (i.e., extrapolating to what would be expected given complete phasing). Given this extrapolation, the measures of uncertainty associated with "Total" are not shown in Figure 4A. For DNMs within a single developmental stage, we calculated binomial confidence intervals, as described above.

4.10 Testing relationships between α and life history traits

In mammals, we collected life history traits from the AnAge database (https://genomics.senescence.info/species/dataset.zip), including maximum longevity, gestation time, adult weight, and birth weight. We also obtained generation time estimates from the literature (Table S2). Thus, in total, we collected data on five traits. Four species were not represented in the AnAge dataset; in these cases, we substituted the trait values of closely related species of the same genus (see Table S2 for species substitutions). We additionally performed principal component analysis (PCA) on the four traits, generation time, gestation time, adult weight, and birth weight (Figure S5), and treated PC1 and PC2 as meta-traits to be tested alongside the others. Only the 17 mammalian species annotated for all four traits were included in the PCA procedure. The first principal component captured 90% of the variance in the traits and was highly correlated with generation time ($r^2 = 86\%$). In birds, we focused on the life history trait of generation time, taking estimates from the literature (Table S2).

To test for relationships between life history traits and α while accounting for phylogenetic non-independence in our data, we used phylogenetic generalized least squares (PGLS) [96]. Ordinary least squares is unsuitable for species trait comparisons, because shared phylogenetic history can create correlation structure in the residuals [97]. PGLS addresses this issue by considering the covariance structure of the residuals as a covariate, assuming that the traits evolve under Brownian motion on the phylogeny [73, 96]. We implemented the analysis using the `pgls` function in the caper R package, which provides the option of fitting Pagel's λ [73], a scalar multiplier of the off-diagonal elements of the expected covariance matrix of the residuals. Briefly, λ denotes the amount

of phylogenetic "signal" in the data. If λ is 0, there is no phylogenetic signal; when λ is 1, the regression model is equivalent to the method of phylogenetic independent contrasts (PIC) [73, 97, 98]. In practice, we found that the `pgls` R function would occasionally fail to converge or converge on a local maximum during maximum likelihood estimation of λ; to address this issue, we initialized the likelihood optimization algorithm with a variety of starting values for λ and retained the model with the highest overall likelihood, which required a minor modification of the base `pgls` function from the caper package.

For each predictive trait (Figure S4), we used our $\hat{\alpha}_{\text{evo}}$ estimates from X (Z)-to-autosome comparisons as the response variable and a time-calibrated phylogeny from TimeTree to estimate the covariance matrix (http://timetree.org/). Following what had been done previously to analyze these relationships [23], we log10-transformed each life history trait prior to performing PGLS. *Canis lupus familiaris*, *Ceratotherium simum cottoni*, and *Pterocles burchelli* were not named in the TimeTree database and so we used split times for *Canis lupus*, *Ceratotherium simum*, and *Pterocles gutturalis* instead, respectively, (Table S2). In all comparisons, we calculated p-values under a model in which λ was set to its maximum likelihood estimate and used default values for the remaining arguments of the `pgls` program. In birds, in which the MLE for λ was 0, we also considered a model in which λ was fixed at 1. To test whether the slope of the $\hat{\alpha}_{\text{evo}}$ versus generation time relationship is the same in birds as in mammals, we performed a modified PGLS regression on the bird data with the slope fixed to the maximum likelihood value obtained for mammals (i.e., slope = 1.20) and the intercept (and λ) as the free parameter. After fitting this model with PGLS, we performed a likelihood ratio test (df = 1) to compare it to an alternative model in which the slope was not fixed (i.e., including intercept, slope, and λ parameters).

4.11 Modeling the effects of germline developmental stages on α

To model variation in α among species, we considered the expected number of mutations that arise in two developmental stages: an early embryonic period, *Early*, which loosely encompasses the time between the zygote and the sexual differentiation of the germline, and a second period, *Late*, that refers to the remaining time until reproduction (Figure 4B). In mammals, the expected number of mutations in the *Early* stage, M_e, is approximately the same in both sexes, as observed in the three cases in which there are data (Figure 4A). In the *Late* stage, we assume mutations arise at a constant rate per year, μ_s in sex s ($s \in \{f, m\}$). If we assume the length of *Early* to be negligible relative to the generation time, G_s in sex s, then the expected number of mutations in sex s equals $\mu_s G_s$. Therefore, the expectation of the ratio of paternal-to-maternal mutations at reproduction, α, can be obtained using equation 1.

To predict α in species lacking estimates of the sex-specific mutation rates for the *Late* stage (i.e., μ_m and μ_f), we made two further assumptions, namely that:

- The expected number of mutations per base-pair M_e in the *Early* stage is constant across species and the same in the two sexes. We used a M_e of 1.66×10^{-9} per base-pair, which equates to 5 early embryonic mutations in an haploid genome of 3 Gb. This value was chosen based on observations in humans, notably a study showing that monozygotic twins differ on average by 5.2 mutations that arose between the twinning event and PGCS (1.3 mutations per haploid set of chromosomes) [66]. Given that 75-80% twinning events occur around the 8–16 cell stage [99], approximately 4 mutations are expected to have arisen during the first few divisions in the embryo (assuming ∼1 extra mutation per cell division [67]). This rate is also in rough agreement with a pedigree study in humans, which estimated that ∼5% of DNMs arise during early development [62]. Varying the expected number from 3 to 7 yielded similar results (see below for more details).
- The ratio μ_m/μ_f is fixed across species. We assumed a ratio of 4, consistent with the ratio of paternal-to-maternal DNMs occurring post-PGCS in humans [62], mice [25] and cattle [26] (Figure 4A).

Using derivations from Amster and Sella [54], the yearly substitution rate μ_Y for a given lineage is:

$$\mu_Y = \frac{2M_e + \mu_f G_f + \mu_m G_m}{G_f + G_m}.$$

If $\mu_m/\mu_f = 4$ and M_e is known, we can solve for μ_f using equation 2 and α can be estimated using equation 1.

We used the PGLS method described in Testing relationships between α and life history traits to assess the fit of α values predicted by our model to the α values estimated from X-to-autosome comparisons ($\hat{\alpha}_{\text{evo}}$) and from DNM studies ($\hat{\alpha}_{\text{dnm}}$)(Figure 4C). We applied the model to mammals using estimates of G from the literature (Table S2). When testing the fit of the model to $\hat{\alpha}_{\text{evo}}$, we estimated μ_Y by dividing the autosomal substitution rates in a lineage (see Estimating putatively-neutral substitution rates) by the split time for that lineage reported in the TimeTree database (http://timetree.org/). When testing the fit to $\hat{\alpha}_{\text{dnm}}$, we obtained α from yearly mutation rates obtained from pedigree sequencing studies, given the parental ages in the study (see Table S2). We note that $\hat{\alpha}_{\text{dnm}}$ can be noisy if not based on a large amount of DNMs and trios. To overcome this limitation, we focused on species with at least 30 phased DNMs and more than one trio sequenced (which excluded three species out of 14, namely *Pongo abelii*, *Callithrix jacchus* and *Ursus arctos*, see Table S2).

We note the model remains a significant predictor for a range of M_e values. As examples, using a λ of 1, as inferred by maximum likelihood in Figure 4D, for a $M_e = 1 \times 10^{-9}$, the model for $\hat{\alpha}_{\text{evo}}$ explains $r^2 = 0.33$ (p-value= 0.008) and for $\hat{\alpha}_{\text{dnm}}$, $r^2 = 0.90$ (p-value= 1×10^{-4}). Instead using $M_e = 2.33 \times 10^{-9}$, the model for $\hat{\alpha}_{\text{evo}}$ accounts for $r^2 = 0.35$ ($p$value= 0.006) and for $\hat{\alpha}_{\text{dnm}}$, $r^2 = 0.79$ ($p$value= 0.001).

Following [10], we sought to determine the extent to which variation in $\hat{\alpha}_{\text{evo}}$ in mammals is attributable to sampling error. To that end, we made use of the empirical distribution of $\hat{\alpha}_{\text{evo}}$, which we obtained by bootstrap resampling genomic windows (see Estimating α from X-to-autosome substitution rate ratios). For each bootstrap replicate, we regressed the α estimates against our original $\hat{\alpha}_{\text{evo}}$ using ordinary least squares and obtained the r^2 value. Across the 500 bootstrap replicates, the median r^2 value was 89%, suggesting that 11% of the variance in $\hat{\alpha}_{\text{evo}}$ is due to sampling error. Combining this value with the estimated proportion of variance in $\hat{\alpha}_{\text{evo}}$ explained by our model yielded an estimate of $37\%/89\% = 42\%$ of the variance explained after accounting for sampling error.

The code to reproduce the modelling described above can be found in the `scripts/2exposure_model.ipynb` Jupyter notebook.

Supplementary tables and figures

Extended Fig. 2: Estimates of the paternal bias in mutation across 42 amniote lineages. Colored points denote estimates of α from X(Z)-autosome substitution rate ratios ($\hat{\alpha}_{\text{evo}}$) in mammals (top, orange), birds (middle, blue), and snakes (bottom, green). Vertical colored lines denote the mean $\hat{\alpha}_{\text{evo}}$ for each group, while the vertical gray dashed line denotes $\alpha = 1$ (i.e., no sex bias in mutation). Species in each group are plotted by their phylogenetic relationships and branch lengths are scaled by the neutral substitution rate estimated from autosomes (see Estimating putatively-neutral substitution rates in Materials and Methods). Note that branch lengths are comparable within the phylogeny of each taxon but not across taxa, as the scaling differs (see the legend for each group). In mammals, $\hat{\alpha}_{\text{evo}}$ was estimated from neutral substitutions along the lineage from the tip to the most recent common ancestor indicated in the phylogeny. In birds, where phylogenetic relationships are more tenuous, we divided species into six subgroups (Table S5) to avoid highly uncertain ancestral nodes in Neoaves; thus, some $\hat{\alpha}_{\text{evo}}$ estimates in Neoaves average over deeper splits than suggested by the full phylogeny, which we plot for clarity. Asterisks indicate species with chromosome-level assemblies. Darker colored horizontal lines behind the points represent 95% CIs, which were computed by bootstrap resampling of the 1 Mb genomic windows across 500 replicates; the central 95% interval across bootstrap replicates is shown. Lighter colored horizontal lines include uncertainty in the ratio of paternal-to-maternal generation times, allowing the ratio to range between 0.9 and 1.1 [54]. Short vertical red lines denote point estimates of $\hat{\alpha}_{\text{dnm}}$ from published pedigree mutation studies of *de novo* mutations, and the surrounding horizontal gray boxes represent the 95% binomial CI for those estimates.

Table S1: Genome assembly statistics and heterozygosity estimates for 241 mammals, 365 birds and 9 snakes.

Table S2: Estimates of α obtained from the ratios of X(Z)-autosome substitution rates and from pedigree data, as well as life history traits, and genome assembly statistics for 46 mammal, bird, and snake species.

Table S3: Phased *de novo* mutation counts categorized by developmental stage.

Species	Paternal	Maternal	Total	Stage	Parental age	Ref
Mus musculus	27	20	140	Early	0.73	[25]
Mus musculus	107	25	434	Late	0.73	[25]
Homo sapiens	352	350	745	Early	31.63	[62]
Homo sapiens	4158	1091	23399	Late	31.63	[62]
Bos taurus	27	23	50	Early	5.00	[26]
Bos taurus	124	32	156	Late	5.00	[26]

Table S4: Sets of bird species used to estimate $\hat{\alpha}_{\text{evo}}$. To avoid uncertain phylogenetic relationships within Neoaves, birds were split into separate analysis sets before estimating substitution rates (see Sequence alignments in Materials and Methods). In groups 1–5, *Gallus gallus* was included as an outgroup when estimating substitution rates.

Analysis Group	Species
1	*Calidris pugnax*, *Columba livia*, *Pterocles burchelli*
2	*Cuculus canorus*, *Egretta garzetta*, *Lophotis ruficrista*
3	*Bucorvus abyssinicus*, *Halcyon senegalensis*, *Trogon melanurus*
4	*Falco cherrug*, *Ficedula albicollis*
5	*Haliaeetus leucocephalus*, *Psophia crepitans*
6	*Anseranas semipalmata*, *Cairina moschata*, *Gallus gallus*, *Penelope pileata*

Table S5: Pseudo-autosomal region intervals.

Species	Chromosome	Regions	Reference
Bos taurus	CM000206.6	145000000-150266161	[100]
Homo sapiens	chrX	10001-2781479, 155701383-156030895	NCBI assembly GCF_000001405.38
Canis lupus familiaris	chrX	1-6642728	[101]
Felis catus	CM001396.2	1-6813699	[101]
Ovis aries	CM001608.2	1-6883022	[102]
Sus scrofa	CM000830.4	1-6900000	[103]
Macaca mulatta	CM002997.3	1-1810963	[101]
Microcebus murinus	CM007693.1	1-5000000	[104]
Mus musculus	chrX	168000000-171031299	[101]
Crotalus viridis	CM012323.1	106784505-113984505	[105]
Gallus gallus	chrZ	1-2000000	[106]
Thamnophis elegans	NC_045558.1	1-6000000, 125000000-145130870	This study

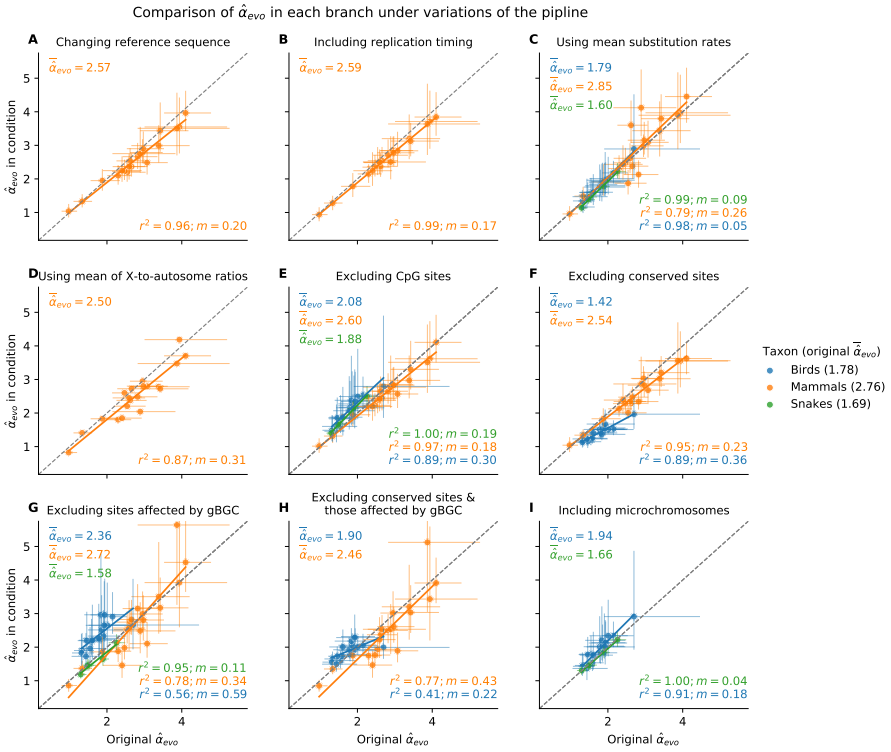
Comparison of $\hat{\alpha}_{evo}$ in each branch under variations of the pipline
A
Changing reference sequence
$\overline{\hat{\alpha}}_{evo} = 2.57$
$r^2 = 0.96; m = 0.20$
B
Including replication timing
$\overline{\hat{\alpha}}_{evo} = 2.59$
$r^2 = 0.99; m = 0.17$
C
Using mean substitution rates
$\overline{\hat{\alpha}}_{evo} = 1.79$
$\overline{\hat{\alpha}}_{evo} = 2.85$
$\overline{\hat{\alpha}}_{evo} = 1.60$
$r^2 = 0.99; m = 0.09$
$r^2 = 0.79; m = 0.26$
$r^2 = 0.98; m = 0.05$
D
Using mean of X-to-autosome ratios
$\overline{\hat{\alpha}}_{evo} = 2.50$
$r^2 = 0.87; m = 0.31$
E
Excluding CpG sites
$\overline{\hat{\alpha}}_{evo} = 2.08$
$\overline{\hat{\alpha}}_{evo} = 2.60$
$\overline{\hat{\alpha}}_{evo} = 1.88$
$r^2 = 1.00; m = 0.19$
$r^2 = 0.97; m = 0.18$
$r^2 = 0.89; m = 0.30$
F
Excluding conserved sites
$\overline{\hat{\alpha}}_{evo} = 1.42$
$\overline{\hat{\alpha}}_{evo} = 2.54$
$r^2 = 0.95; m = 0.23$
$r^2 = 0.89; m = 0.36$
G
Excluding sites affected by gBGC
$\overline{\hat{\alpha}}_{evo} = 2.36$
$\overline{\hat{\alpha}}_{evo} = 2.72$
$\overline{\hat{\alpha}}_{evo} = 1.58$
$r^2 = 0.95; m = 0.11$
$r^2 = 0.78; m = 0.34$
$r^2 = 0.56; m = 0.59$
H
Excluding conserved sites & those affected by gBGC
$\overline{\hat{\alpha}}_{evo} = 1.90$
$\overline{\hat{\alpha}}_{evo} = 2.46$
$r^2 = 0.77; m = 0.43$
$r^2 = 0.41; m = 0.22$
I
Including microchromosomes
$\overline{\hat{\alpha}}_{evo} = 1.94$
$\overline{\hat{\alpha}}_{evo} = 1.66$
$r^2 = 1.00; m = 0.04$
$r^2 = 0.91; m = 0.18$
$\hat{\alpha}_{evo}$ in condition
Original $\hat{\alpha}_{evo}$
Taxon (original $\overline{\hat{\alpha}}_{evo}$)
Birds (1.78)
Mammals (2.76)
Snakes (1.69)

Fig. S1 *(previous page)*: $\hat{\alpha}_{\text{evo}}$ **for each species, obtained under variants of the pipeline presented in the main text.** In each plot, the $\hat{\alpha}_{\text{evo}}$ inferred by the pipeline described in Figure 1 and used throughout the manuscript (shown on the x-axis) is compared to $\hat{\alpha}_{\text{evo}}$ estimates from the same pipeline with one modified condition (y-axis): In (A), the mammalian $\hat{\alpha}_{\text{evo}}$ is obtained using *Mus musculus* as reference sequence instead of *Homo sapiens* (see Sequence alignments in Materials and Methods); in (B) replication timing is included as a predictor in the regression (see Estimating α from X-to-autosome substitution rate ratios in Materials and Methods); in (C) $\hat{\alpha}_{\text{evo}}$ is obtained from the ratio of X (Z)-to-autosome mean substitution rates across genomic windows (see Estimating α from X-to-autosome substitution rate ratios in Materials and Methods); in (D) $\hat{\alpha}_{\text{evo}}$ is obtained from the mean of the ratio of X (Z)-to-autosome substitution rates, in windows with similar GC content (see Estimating α from X-to-autosome substitution rate ratios in Materials and Methods). Here, $\hat{\alpha}_{\text{evo}}$ CIs in the y-axis are extremely wide and not shown; in (E) $\hat{\alpha}_{\text{evo}}$ is obtained after masking all CpG sites in the genome sequence alignment (see Selecting non-repetitive and putatively neutral sequences in Materials and Methods); in (F) $\hat{\alpha}_{\text{evo}}$ is obtained after masking all conserved sites identified by phastCons, in addition to exons (see Selecting non-repetitive and putatively neutral sequences in Materials and Methods); in (G) $\hat{\alpha}_{\text{evo}}$ is estimated for sites unaffected by GC-biased gene conversion (see Estimating putatively-neutral substitution rates in Materials and Methods); in (H) $\hat{\alpha}_{\text{evo}}$ is estimated by both excluding conserved regions as in F, and using sites unaffected by GC-biased gene conversion as in G; and in (I) $\hat{\alpha}_{\text{evo}}$ in birds and snakes is obtained including microchromosomes rather than masking them (see Selecting non-repetitive and putatively neutral sequences in Materials and Methods). r^2 and mean absolute error (m) are shown in the bottom right corner of each subplot. For each taxon, the mean $\hat{\alpha}_{\text{evo}}$ obtained using the modified pipeline are annotated in the top-left corner, and the mean $\hat{\alpha}_{\text{evo}}$ in the original pipeline can be found in the figure legend. Lines indicate ordinary least squares regression fits.

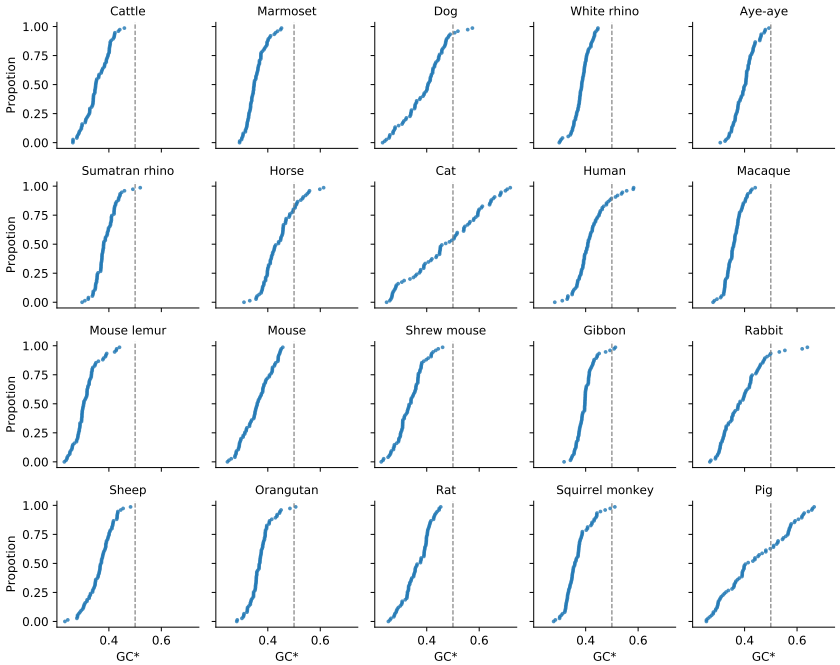

Fig. S2: **Expected equilibrium GC content (GC*) in the mammalian X chromosomes.** GC* is calculated as the fraction of the rate of substitutions favored by GC-biased gene conversion out of the total rate of substitutions affected by GC-biased gene conversion (i.e., $\frac{AT \to GC}{AT \to GC + GC \to AT}$). To calculate GC* in 1 Mb genomic windows across the X chromosome of mammals, we used estimated rates of putatively neutral substitution for single mutation classes as described in Estimating putatively-neutral substitution rates in Materials and Methods. The vertical gray dashed lines at GC* = 0.5 are plotted to facilitate comparison between the distributions. The lineages leading to cats and pigs have unusually wide distributions of GC*, consistent with their idiosyncratic recombination landscapes in the X chromosome [56].

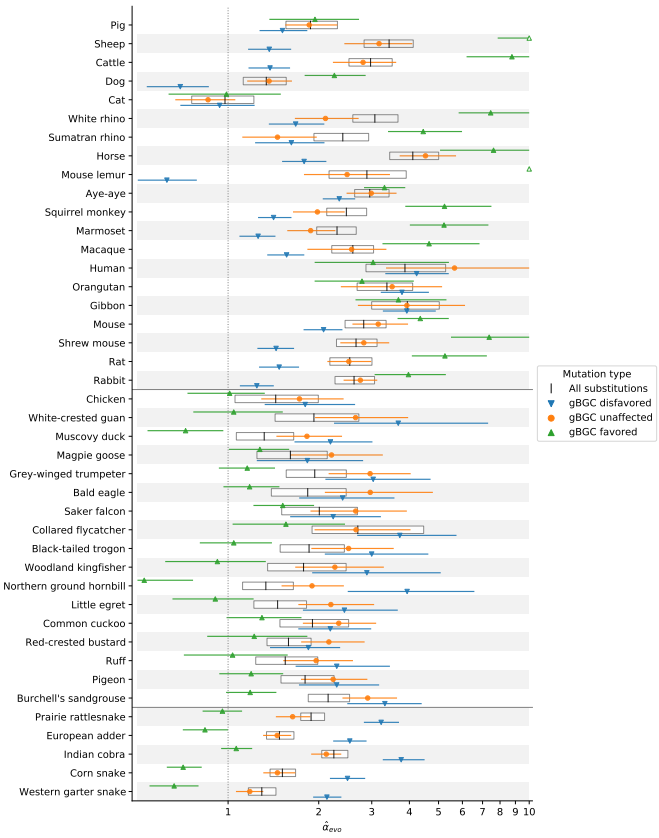

Fig. S3: **Estimation of $\hat{\alpha}_{\text{evo}}$ for mutation types affected or unaffected by GC-biased gene conversion.** α estimates were obtained from the ratio of X (Z)-to-autosome substitution rates of four mutation types (see Estimating putatively-neutral substitution rates in Materials and Methods): all substitutions (white boxes), substitutions that GC-biased gene conversion (gBGC) acts against (i.e., changes from strong [S] nucleotides [C & G] to weak [W] nucleotides [A & T], blue), substitutions favoured by gBGC (W>S, green), and substitutions unaffected by gBGC (W>W and S>S, orange). Note that the x-axis is log-scaled, and values extending below 0.5 or past 10 are truncated, with hollow markers showing point estimates that are outside of these bounds. Horizontal black lines separate species belonging to mammals, birds and snakes (from top to bottom).

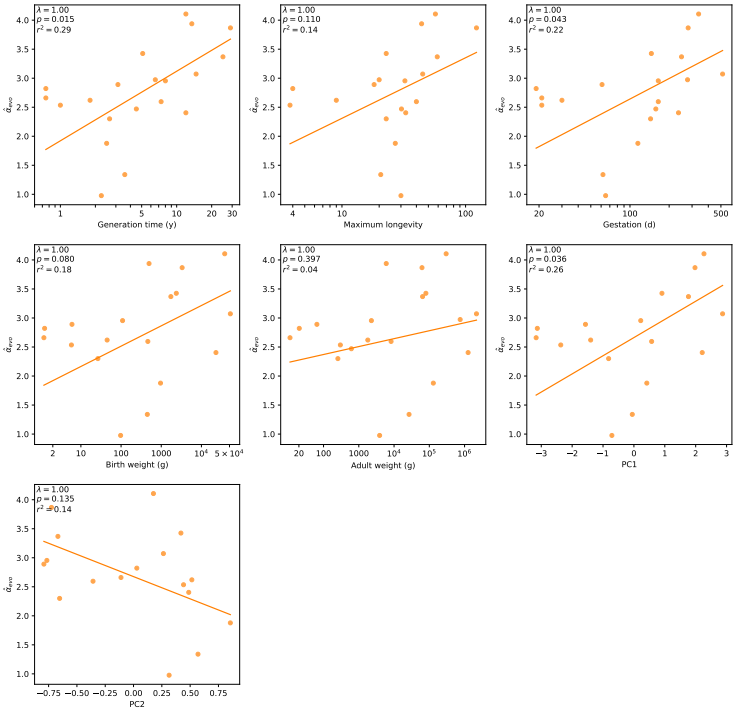

Fig. S4: **Relationship between mammalian $\hat{\alpha}_{\text{evo}}$ and various life history traits.** α estimates from the ratio of X-to-autosome substitution rates were regressed against five life history traits separately, using phylogenetic generalized least squares. The estimates were also regressed against the first two principal components (PCs, see Figure S5 and Testing relationships between α and life history traits in Materials and Methods) obtained in a PC analysis of the traits gestation, birth weight, adult weight, and generation time. (For reference, the $\hat{\alpha}_{\text{evo}}$ comparison with generation time from Figure 3 is reproduced here.) Best fit lines are drawn in orange; regression statistics are indicated in the upper left corner of each subplot. Pagel's λ [73] was estimated by maximum likelihood (see Testing relationships between α and life history traits in Materials and Methods).

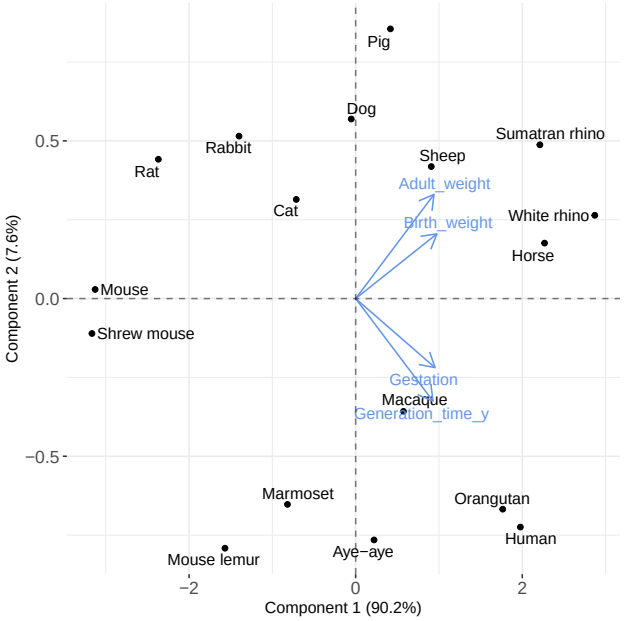

Fig. S5: **Principal component analysis of four life history traits.** Gestation, birth weight, adult weight, and generation time measurements in mammals were gathered from the literature (see Testing relationships between α and life history traits in Materials and Methods and Table S2) and subjected to principal component (PC) analysis. Species are projected onto the two leading PCs, which are shown with the amount of variance captured by each in parentheses. Blue arrows denote loadings for the trait variables.

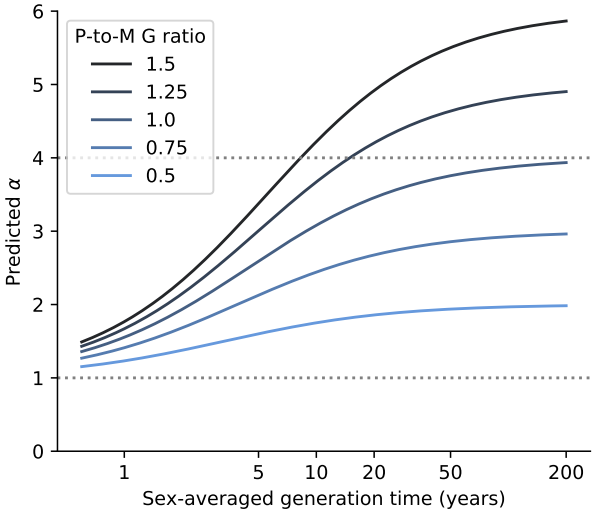

Fig. S6: **The maximal value of α depends on the ratio of paternal-to-maternal generation times.** The relationship between generation time and α that is expected using the model depicted in Figure 4B, for different ratios of paternal-to-maternal generation times ("P-to-M G times") (blue hues). Namely, we assume that (i) each sex accrues the same expected number of *Early* mutations per base pair ($M_e = 1.66 \times 10^{-9}$), (ii) the ratio of paternal-to-maternal mutation rate per unit of time is 4 in the *Late* stage ($4\mu_f = \mu_f$), and (iii) μ_m equals 1×10^{-9} per base pair per year (see Modeling the effects of germline developmental stages on α in Materials and Methods for more details). The maximum generation time (x-axis) was unrealistically set to 200 years to better show the plateau in α.

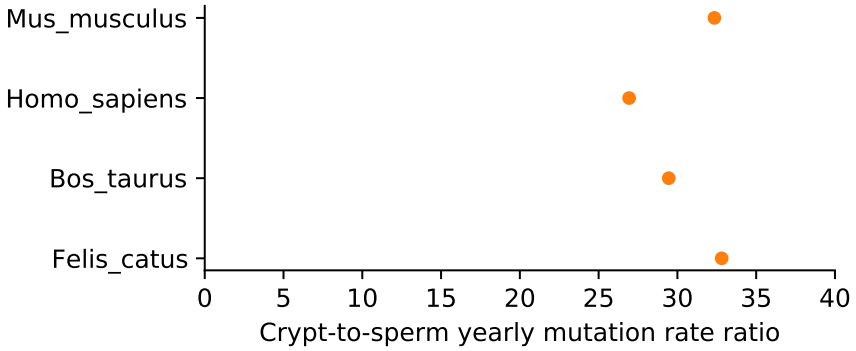

Fig. S7: **Ratio of crypt-to-sperm mutation rate per unit of time in four mammals.** The yearly mutation rate in colonic crypts was obtained from [38] for four species in which pedigree sequencing estimates of the mutation rate per generation are also available [18, 25, 26, 55]. To estimate the mutation rate per year in sperm, we used the paternal-to-maternal generation time ratio in each pedigree study and assumed that: (i) each sex accrues the same expected number of *Early* mutations per base pair ($M_e = 1.66 \times 10^{-9}$) and (ii) a ratio of paternal-to-maternal mutation rates per unit of time of 4 in the *Late* phase ($4\mu_f = \mu_m$) (see Modeling the effects of germline developmental stages on α in Materials and Methods).

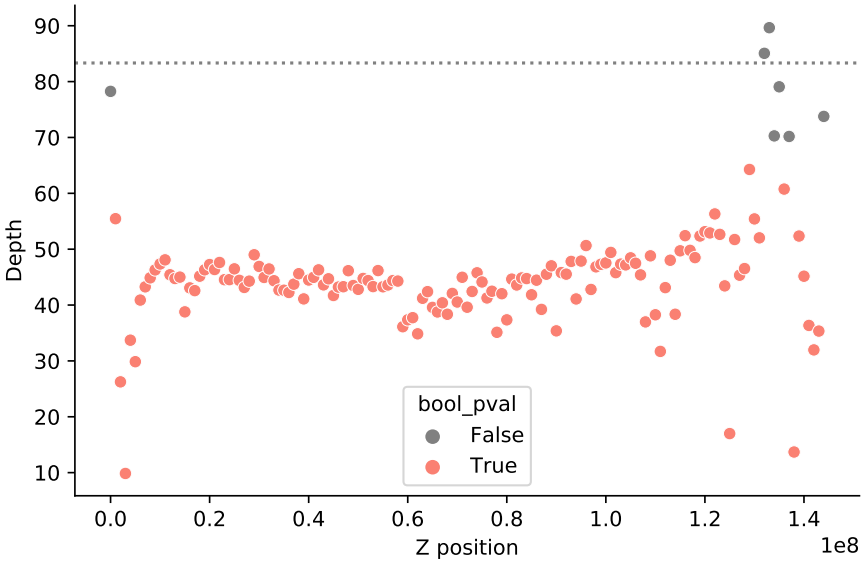

Fig. S8: **Identification of pseudo-autosomal regions in *Thamnophis*.** Depth of coverage along the Z chromosome of a heterogametic *Thamnophis sirtalis* individual. Each data point represents a 1 Mb window, colored by p-value (red if $\leqq 0.05$) assuming depth of coverage is Poisson distributed with $\lambda = 83.3$ (the mean coverage in the autosomes of the same individual)(see Selecting non-repetitive and putatively neutral sequences in Materials and Methods). We considered the Z genomic windows with similar coverage to that in the autosomes (gray points) to be pseudo-autosomal regions with homologs on the W chromosome.

Data availability

Scripts and data for reproducing the analyses and figures may be found at `https://github.com/flw88/mut_sex_bias_amniotes`.

Acknowledgements

We thank Ziyue Gao, Guy Sella, and the Coop and Schierup labs for their comments on earlier versions of the manuscript. We thank Rusty Lansford, Mike McGrew, and Daniel Hooper for discussions about avian development and evolution; Turk Rhen for discussions about reptile sex determination; Carla Hoge and Zach Fuller for sharing their corn snake genome assembly; Anne Bronikowski and the Vertebrate Genome Project for sponsoring and generating the *Thamnophis elegans* assembly; Alex Cagan for early access to data of mutation burdens in colonic crypts across mammals; Richard Wang and Matthew Hahn for sharing data on *de novo* mutation in cats; Carole Charlier and Michel Georges for sharing data on *de novo* mutation in cattle; Adam Siepel for help with applying the phyloFit program; and Peter Andolfatto, Michael B. Eisen, Priya Moorjani, as well as William R. Milligan, Anna Yoney and other members of the Andolfatto, Przeworski, and Sella labs for helpful discussions. This work was funded by GM122975 to MP and a HFSP postdoctoral fellowship to MdM.

Competing interests

Molly Przeworski is a Senior editor at eLife.

References

[1] Haldane, J.: The mutation rate of the gene for haemophilia, and its segregation ratios in males and females. Annals of eugenics **13**(1), 262–271 (1946)

[2] Makova, K.D., Li, W.-H.: Strong male-driven evolution of dna sequences in humans and apes. Nature **416**(6881), 624–626 (2002)

[3] Wolfe, K.H., Li, W.-H.: Molecular evolution meets the genomics revolution. Nature genetics **33**(3), 255–265 (2003)

[4] Li, W.-H., Ellsworth, D.L., Krushkal, J., Chang, B.H.-J., Hewett-Emmett, D.: Rates of nucleotide substitution in primates and rodents and the generation–time effect hypothesis. Molecular phylogenetics and evolution **5**(1), 182–187 (1996)

[5] Presgraves, D.C., Soojin, V.Y.: Doubts about complex speciation between humans and chimpanzees. Trends in ecology & evolution **24**(10), 533–540 (2009)

[6] Nachman, M.W., Crowell, S.L.: Estimate of the mutation rate per nucleotide in humans. Genetics **156**(1), 297–304 (2000)

[7] Huang, W., Chang, B.H.-J., Gu, X., Hewett-Emmett, D., Li, W.-H.: Sex differences in mutation rate in higher primates estimated from amg intron sequences. Journal of molecular evolution **44**(4), 463–465 (1997)

[8] Shimmin, L.C., Chang, B.H.-J., Li, W.-H.: Male-driven evolution of dna sequences. Nature **362**(6422), 745–747 (1993)

[9] Chang, B., Shimmin, L.C., Shyue, S.-K., Hewett-Emmett, D., Li, W.-H.: Weak male-driven molecular evolution in rodents. Proceedings of the National Academy of Sciences **91**(2), 827–831 (1994)

[10] Kong, A., Frigge, M.L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., Gudjonsson, S.A., Sigurdsson, A., Jonasdottir, A., Jonasdottir, A., *et al.*: Rate of de novo mutations and the importance of father's age to disease risk. Nature **488**(7412), 471–475 (2012)

[11] Francioli, L.C., Polak, P.P., Koren, A., Menelaou, A., Chun, S., Renkens, I., Van Duijn, C.M., Swertz, M., Wijmenga, C., Van Ommen, G., *et al.*: Genome-wide patterns and properties of de novo mutations in humans. Nature genetics **47**(7), 822–826 (2015)

[12] Crow, J.F.: The origins, patterns and implications of human spontaneous mutation. Nature Reviews Genetics **1**(1), 40–47 (2000)

[13] Drost, J.B., Lee, W.R.: Biological basis of germline mutation: comparisons of spontaneous germline mutation rates among drosophila, mouse, and human. Environmental and molecular mutagenesis **25**(S2), 48–64 (1995)

[14] Penrose, L.: Parental age and mutation. The Lancet **266**(6885), 312–313 (1955)

[15] Strachan, T., Read, A.: Human molecular genetics. garland. New York (2018)

[16] Guo, J., Nie, X., Giebler, M., Mlcochova, H., Wang, Y., Grow, E.J., Kim, R., Tharmalingam, M., Matilionyte, G., Lindskog, C., *et al.*: The dynamic transcriptional cell atlas of testis development during human puberty. Cell stem cell **26**(2), 262–276 (2020)

[17] Goldmann, J., Veltman, J., Gilissen, C.: De novo mutations reflect development and aging of the human germline. Trends in Genetics **35**(11), 828–839 (2019)

[18] Jónsson, H., Sulem, P., Kehr, B., Kristmundsdottir, S., Zink, F., Hjartarson, E., Hardarson, M.T., Hjorleifsson, K.E., Eggertsson, H.P., Gudjonsson, S.A., *et al.*: Parental influence on human germline de novo mutations in 1,548 trios from iceland. Nature **549**(7673), 519–522 (2017)

[19] Wong, W.S., Solomon, B.D., Bodian, D.L., Kothiyal, P., Eley, G., Huddleston, K.C., Baker, R., Thach, D.C., Iyer, R.K., Vockley, J.G., *et al.*: New observations on maternal age effect on germline de novo mutations. Nature communications **7**(1), 1–10 (2016)

[20] Goldmann, J.M., Wong, W.S., Pinelli, M., Farrah, T., Bodian, D., Stittrich, A.B., Glusman, G., Vissers, L.E., Hoischen, A., Roach, J.C., *et al.*: Parent-of-origin-specific signatures of de novo mutations. Nature genetics **48**(8), 935–939 (2016)

[21] Gao, Z., Moorjani, P., Sasani, T.A., Pedersen, B.S., Quinlan, A.R., Jorde, L.B., Amster, G., Przeworski, M.: Overlooked roles of dna damage and maternal age in generating human germline mutations. Proceedings of the National Academy of Sciences **116**(19), 9491–9500 (2019)

[22] Wu, F.L., Strand, A.I., Cox, L.A., Ober, C., Wall, J.D., Moorjani, P., Przeworski, M.: A comparison of humans and baboons suggests germline mutation rates do not track cell divisions. PLoS biology **18**(8), 3000838 (2020)

[23] Sayres, M.A.W., Venditti, C., Pagel, M., Makova, K.D.: Do variations in substitution rates and male mutation bias correlate with life-history traits? a study of 32 mammalian genomes. Evolution: International Journal of Organic Evolution **65**(10), 2800–2815 (2011)

[24] Wang, R.J., Pena-Garcia, Y., Bibby, M., Raveendran, M., Harris, R.A., Jansen, H.T., Robbins, C.T., Rogers, J., Kelley, J.L., Hahn, M.W.: Hibernation shows no apparent effect on germline mutation rates in grizzly bears. bioRxiv (2022)

[25] Lindsay, S.J., Rahbari, R., Kaplanis, J., Keane, T., Hurles, M.E.: Similarities and differences in patterns of germline mutation between mice and humans. Nature communications **10**(1), 1–12 (2019)

[26] Harland, C., Charlier, C., Karim, L., Cambisano, N., Deckers, M., Mni, M., Mullaart, E., Coppieters, W., Georges, M.: Frequency of mosaicism points towards mutation-prone early cleavage cell divisions in cattle. BioRxiv, 079863 (2017)

[27] Gao, Z., Wyman, M.J., Sella, G., Przeworski, M.: Interpreting the dependence of mutation rates on age and time. PLoS biology **14**(1), 1002355 (2016)

[28] Seplyarskiy, V.B., Soldatov, R.A., Koch, E., McGinty, R.J., Goldmann, J.M., Hernandez, R.D., Barnes, K., Correa, A., Burchard, E.G., Ellinor, P.T., *et al.*: Population sequencing data reveal a compendium of mutational processes in the human germ line. Science **373**(6558), 1030–1035 (2021)

[29] Vilenchik, M.M., Knudson, A.G.: Endogenous dna double-strand breaks: production, fidelity of repair, and induction of cancer. Proceedings of the National Academy of Sciences **100**(22), 12871–12876 (2003)

[30] Abascal, F., Harvey, L.M., Mitchell, E., Lawson, A.R., Lensing, S.V., Ellis, P., Russell, A.J., Alcantara, R.E., Baez-Ortega, A., Wang, Y., *et al.*: Somatic mutation landscapes at single-molecule resolution. Nature **593**(7859), 405–410 (2021)

[31] Rahbari, R., Wuster, A., Lindsay, S.J., Hardwick, R.J., Alexandrov, L.B., Al Turki, S., Dominiczak, A., Morris, A., Porteous, D., Smith, B., *et al.*: Timing, rates and spectra of human germline mutation. Nature genetics **48**(2), 126–133 (2016)

[32] Moore, L., Cagan, A., Coorens, T.H., Neville, M.D., Sanghvi, R., Sanders, M.A., Oliver, T.R., Leongamornlert, D., Ellis, P., Noorani, A., *et al.*: The mutational landscape of human somatic and germline cells. Nature **597**(7876), 381–386 (2021)

[33] Alexandrov, L.B., Jones, P.H., Wedge, D.C., Sale, J.E., Campbell, P.J., Nik-Zainal, S., Stratton, M.R.: Clock-like mutational processes in human somatic cells. Nature genetics **47**(12), 1402–1407 (2015)

[34] Alexandrov, L.B., Kim, J., Haradhvala, N.J., Huang, M.N., Ng, A.W.T., Wu, Y., Boot, A., Covington, K.R., Gordenin, D.A., Bergstrom, E.N., *et al.*: The repertoire of mutational signatures in human cancer. Nature **578**(7793), 94–101 (2020)

[35] Lodato, M.A., Rodin, R.E., Bohrson, C.L., Coulter, M.E., Barton, A.R., Kwon, M., Sherman, M.A., Vitzthum, C.M., Luquette, L.J., Yandava, C.N., *et al.*: Aging and neurodegeneration are associated with increased mutations in single human neurons. Science **359**(6375), 555–559 (2018)

[36] Oliver, T.R.W., Chappell, L., Sanghvi, R., Deighton, L., Ansari-Pour, N., Dentro, S.C., Young, M.D., Coorens, T.H.H., Jung, H., Butler, T., Neville, M.D.C., Leongamornlert, D., Sanders, M., Hooks, Y., Cagan, A., Mitchell, T.J., Cortes-Ciriano, I., Warren, A.Y., Wedge, D.C., Heer, R., Coleman, N., Murray, M.J., Campbell, P.J., Rahbari, R., Behjati, S.: Clonal diversification and histogenesis of malignant germ cell tumours. bioRxiv (2022)

[37] Blokzijl, F., De Ligt, J., Jager, M., Sasselli, V., Roerink, S., Sasaki, N., Huch, M., Boymans, S., Kuijk, E., Prins, P., *et al.*: Tissue-specific mutation accumulation in human adult stem cells during life. Nature **538**(7624), 260–264 (2016)

[38] Cagan, A., Baez-Ortega, A., Brzozowska, N., Abascal, F., Coorens, T.H., Sanders, M.A., Lawson, A.R., Harvey, L.M., Bhosle, S., Jones, D., et al.: Somatic mutation rates scale with lifespan across mammals. Nature, 1–8 (2022)

[39] Zhou, Z.-X., Lujan, S.A., Burkholder, A.B., St Charles, J., Dahl, J., Farrell, C.E., Williams, J.S., Kunkel, T.A.: How asymmetric dna replication achieves symmetrical fidelity. Nature structural & molecular biology **28**(12), 1020–1028 (2021)

[40] Guraya, S.S.: Ovarian Follicles in Reptiles and Birds. Springer, Berlin (1989)

[41] Deviche, P., Hurley, L.L., Fokidis, H.B.: Avian testicular structure, function, and regulation. In: Hormones and Reproduction of Vertebrates, pp. 27–70. Academic Press, London (2011)

[42] Miyata, T., Hayashida, H., Kuma, K., Mitsuyasu, K., Yasunaga, T.: Male-driven molecular evolution: a model and nucleotide sequence analysis. In: Cold Spring Harbor Symposia on Quantitative Biology, vol. 52, pp. 863–867 (1987). Cold Spring Harbor Laboratory Press

[43] Slattery, J.P., O'Brien, S.J.: Patterns of y and x chromosome dna sequence divergence during the felidae radiation. Genetics **148**(3), 1245–1255 (1998)

[44] Ellegren, H., Fridolfsson, A.-K.: Male–driven evolution of dna sequences in birds. Nature genetics **17**(2), 182–184 (1997)

[45] Carmichael, A.N., Fridolfsson, A.-K., Halverson, J., Ellegren, H.: Male-biased mutation rates revealed from z and w chromosome-linked atp synthase α-subunit (atp5a1) sequences in birds. Journal of Molecular Evolution **50**(5), 443–447 (2000)

[46] Shimmin, L.C., Chang, B.H.-J., Hewett-Emmett, D., Li, W.-H.: Potential problems in estimating the male-to-female mutation rate ratio from dna sequence data. Journal of molecular evolution **37**(2), 160–166 (1993)

[47] Pink, C.J., Hurst, L.D.: Timing of replication is a determinant of neutral substitution rates but does not explain slow y chromosome evolution in rodents. Molecular biology and evolution **27**(5), 1077–1086 (2010)

[48] Agarwal, I., Przeworski, M.: Signatures of replication timing, recombination, and sex in the spectrum of rare variants on the human x chromosome and autosomes. Proceedings of the National Academy of Sciences **116**(36), 17916–17924 (2019)

[49] Jarvis, E.D., Mirarab, S., Aberer, A.J., Li, B., Houde, P., Li, C., Ho, S.Y., Faircloth, B.C., Nabholz, B., Howard, J.T., *et al.*: Whole-genome analyses resolve early branches in the tree of life of modern birds. Science **346**(6215), 1320–1331 (2014)

[50] Reddy, S., Kimball, R.T., Pandey, A., Hosner, P.A., Braun, M.J., Hackett, S.J., Han, K.-L., Harshman, J., Huddleston, C.J., Kingston, S., *et al.*: Why do phylogenomic data sets yield conflicting trees? data type influences the avian tree of life more than taxon sampling. Systematic Biology **66**(5), 857–879 (2017)

[51] Consortium, Z., *et al.*: A comparative genomics multitool for scientific discovery and conservation. Nature **587**(7833), 240 (2020)

[52] Feng, S., Stiller, J., Deng, Y., Armstrong, J., Fang, Q., Reeve, A.H., Xie, D., Chen, G., Guo, C., Faircloth, B.C., *et al.*: Dense sampling of bird diversity increases power of comparative genomics. Nature **587**(7833), 252–257 (2020)

[53] Koren, A., Polak, P., Nemesh, J., Michaelson, J.J., Sebat, J., Sunyaev, S.R., McCarroll, S.A.: Differential relationship of dna replication timing to different forms of human mutation and variation. The American Journal of Human Genetics **91**(6), 1033–1040 (2012)

[54] Amster, G., Sella, G.: Life history effects on the molecular clock of autosomes and sex chromosomes. Proceedings of the National Academy of Sciences **113**(6), 1588–1593 (2016)

[55] Wang, R.J., Raveendran, M., Harris, R.A., Murphy, W.J., Lyons, L.A., Rogers, J., Hahn, M.W.: De novo mutations in domestic cat are consistent with an effect of reproductive longevity on both the rate and spectrum of mutations. bioRxiv (2021)

[56] Li, G., Hillier, L.W., Grahn, R.A., Zimin, A.V., David, V.A., Menotti-Raymond, M., Middleton, R., Hannah, S., Hendrickson, S., Makunin, A., *et al.*: A high-resolution snp array-based linkage map anchors a new domestic cat draft genome assembly and provides detailed patterns of recombination. G3: Genes, Genomes, Genetics **6**(6), 1607–1616 (2016)

[57] Meunier, J., Duret, L.: Recombination drives the evolution of gc-content in the human genome. Molecular biology and evolution **21**(6), 984–990 (2004)

[58] Schield, D.R., Perry, B.W., Nikolakis, Z.L., Mackessy, S.P., Castoe, T.A.: Population genomic analyses confirm male-biased mutation rates in snakes. Journal of Heredity **112**(2), 221–227 (2021)

[59] Li, W.-H., Yi, S., Makova, K.: Male-driven evolution. Current opinion in genetics & development **12**(6), 650–656 (2002)

[60] Wang, Z., Zhang, J., Yang, W., An, N., Zhang, P., Zhang, G., Zhou, Q.: Temporal genomic evolution of bird sex chromosomes. BMC evolutionary biology **14**(1), 1–12 (2014)

[61] Lin, Y.-T., Capel, B.: Cell fate commitment during mammalian sex determination. Current opinion in genetics & development **32**, 144–152 (2015)

[62] Sasani, T.A., Pedersen, B.S., Gao, Z., Baird, L., Przeworski, M., Jorde, L.B., Quinlan, A.R.: Large, three-generation human families reveal post-zygotic mosaicism and variability in germline mutation accumulation. Elife **8**, 46922 (2019)

[63] Engel, N.: Sex differences in early embryogenesis: Inter-chromosomal regulation sets the stage for sex-biased gene networks: The dialogue between the sex chromosomes and autosomes imposes sexual identity soon after fertilization. BioEssays **40**(9), 1800073 (2018)

[64] Spiller, C., Koopman, P., Bowles, J.: Sex determination in the mammalian germline. Annual review of genetics **51**, 265–285 (2017)

[65] Hancock, G.V., Wamaitha, S.E., Peretz, L., Clark, A.T.: Mammalian primordial germ cell specification. Development **148**(6), 189217 (2021)

[66] Jonsson, H., Magnusdottir, E., Eggertsson, H.P., Stefansson, O.A., Arnadottir, G.A., Eiriksson, O., Zink, F., Helgason, E.A., Jonsdottir, I., Gylfason, A., *et al.*: Differences between germline genomes of monozygotic twins. Nature Genetics **53**(1), 27–34 (2021)

[67] Ju, Y.S., Martincorena, I., Gerstung, M., Petljak, M., Alexandrov, L.B., Rahbari, R., Wedge, D.C., Davies, H.R., Ramakrishna, M., Fullam, A., *et al.*: Somatic mutations reveal asymmetric cellular dynamics in the early human embryo. Nature **543**(7647), 714–718 (2017)

[68] Zhao, D., McBride, D., Nandi, S., McQueen, H., McGrew, M., Hocking, P., Lewis, P., Sang, H., Clinton, M.: Somatic sex identity is cell autonomous in the chicken. Nature **464**(7286), 237–242 (2010)

[69] Ioannidis, J., Taylor, G., Zhao, D., Liu, L., Idoko-Akoh, A., Gong, D., Lovell-Badge, R., Guioli, S., McGrew, M.J., Clinton, M.: Primary sex

determination in birds depends on dmrt1 dosage, but gonadal sex does not determine adult secondary sex characteristics. Proceedings of the National Academy of Sciences **118**(10) (2021)

[70] Extavour, C.G., Akam, M.: Mechanisms of germ cell specification across the metazoans: epigenesis and preformation (2003)

[71] Soler, L., Alves, S., Brionne, A., Jacques, A., Guérin, V., Cherif-Feildel, M., Combes-Soia, L., Fouchécourt, S., Thélie, A., Blesbois, E., *et al.*: Protein expression reveals a molecular sexual identity of avian primordial germ cells at pre-gonadal stages. Scientific reports **11**(1), 1–19 (2021)

[72] Smith, T.B., Dun, M.D., Smith, N.D., Curry, B.J., Connaughton, H.S., Aitken, R.J.: The presence of a truncated base excision repair pathway in human spermatozoa that is mediated by ogg1. Journal of cell science **126**(6), 1488–1497 (2013)

[73] Pagel, M.: Inferring the historical patterns of biological evolution. Nature **401**(6756), 877–884 (1999)

[74] Prum, R.O., Berv, J.S., Dornburg, A., Field, D.J., Townsend, J.P., Lemmon, E.M., Lemmon, A.R.: A comprehensive phylogeny of birds (aves) using targeted next-generation dna sequencing. Nature **526**(7574), 569–573 (2015)

[75] Ellegren, H.: Evolutionary stasis: the stable chromosomes of birds. Trends in ecology & evolution **25**(5), 283–291 (2010)

[76] Amster, G., Sella, G.: Life history effects on neutral diversity levels of autosomes and sex chromosomes. Genetics **215**(4), 1133–1142 (2020)

[77] Buffalo, V.: Quantifying the relationship between genetic diversity and population size suggests natural selection cannot explain lewontin's paradox. Elife **10**, 67509 (2021)

[78] Brüniche-Olsen, A., Kellner, K.F., Belant, J.L., DeWoody, J.A.: Life-history traits and habitat availability shape genomic diversity in birds: implications for conservation. Proceedings of the Royal Society B **288**(1961), 20211441 (2021)

[79] Kimura, M.: The Neutral Theory of Molecular Evolution. Cambridge University Press, New York (1983)

[80] Siepel, A., Bejerano, G., Pedersen, J.S., Hinrichs, A.S., Hou, M., Rosenbloom, K., Clawson, H., Spieth, J., Hillier, L.W., Richards, S., *et al.*: Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome research **15**(8), 1034–1050 (2005)

[81] Duret, L., Galtier, N.: Biased gene conversion and the evolution of mammalian genomic landscapes. Annual review of genomics and human genetics **10**, 285–311 (2009)

[82] Waters, P.D., Patel, H.R., Ruiz-Herrera, A., Álvarez-González, L., Lister, N.C., Simakov, O., Ezaz, T., Kaur, P., Frere, C., Grützner, F., et al.: Microchromosomes are building blocks of bird, reptile, and mammal chromosomes. Proceedings of the National Academy of Sciences **118**(45) (2021)

[83] Siepel, A., Haussler, D.: Phylogenetic estimation of context-dependent substitution rates by maximum likelihood. Molecular biology and evolution **21**(3), 468–488 (2004)

[84] Hubisz, M.J., Pollard, K.S., Siepel, A.: Phast and rphast: phylogenetic analysis with space/time models. Briefings in bioinformatics **12**(1), 41–51 (2011)

[85] Hardison, R.C., Roskin, K.M., Yang, S., Diekhans, M., Kent, W.J., Weber, R., Elnitski, L., Li, J., O'Connor, M., Kolbe, D., *et al.*: Covariation in frequencies of substitution, deletion, transposition, and recombination during eutherian evolution. Genome research **13**(1), 13–26 (2003)

[86] Hellmann, I., Prüfer, K., Ji, H., Zody, M.C., Pääbo, S., Ptak, S.E.: Why do human diversity levels vary at a megabase scale? Genome research **15**(9), 1222–1231 (2005)

[87] Bergeron, L.A., Besenbacher, S., Bakker, J., Zheng, J., Li, P., Pacheco, G., Sinding, M.-H.S., Kamilari, M., Gilbert, M.T.P., Schierup, M.H., *et al.*: The germline mutational process in rhesus macaque and its implications for phylogenetic dating. GigaScience **10**(5), 029 (2021)

[88] Besenbacher, S., Hvilsom, C., Marques-Bonet, T., Mailund, T., Schierup, M.H.: Direct estimation of mutations in great apes reconciles phylogenetic dating. Nature ecology & evolution **3**(2), 286–292 (2019)

[89] Campbell, C.R., Tiley, G.P., Poelstra, J.W., Hunnicutt, K.E., Larsen, P.A., Lee, H.-J., Thorne, J.L., Dos Reis, M., Yoder, A.D.: Pedigree-based and phylogenetic methods support surprising patterns of mutation rate and spectrum in the gray mouse lemur. Heredity **127**(2), 233–244 (2021)

[90] Smeds, L., Qvarnström, A., Ellegren, H.: Direct estimate of the rate of germline mutation in a bird. Genome research **26**(9), 1211–1218 (2016)

[91] Tatsumoto, S., Go, Y., Fukuta, K., Noguchi, H., Hayakawa, T., Tomonaga, M., Hirai, H., Matsuzawa, T., Agata, K., Fujiyama, A.: Direct

estimation of de novo mutation rates in a chimpanzee parent-offspring trio by ultra-deep whole genome sequencing. Scientific reports **7**(1), 1–12 (2017)

[92] Thomas, G.W., Wang, R.J., Puri, A., Harris, R.A., Raveendran, M., Hughes, D.S., Murali, S.C., Williams, L.E., Doddapaneni, H., Muzny, D.M., *et al.*: Reproductive longevity predicts mutation rates in primates. Current Biology **28**(19), 3193–3197 (2018)

[93] Wang, R.J., Thomas, G.W., Raveendran, M., Harris, R.A., Doddapaneni, H., Muzny, D.M., Capitanio, J.P., Radivojac, P., Rogers, J., Hahn, M.W.: Paternal age in rhesus macaques is positively associated with germline mutation accumulation but not with measures of offspring sociability. Genome research **30**(6), 826–834 (2020)

[94] Yang, C., Zhou, Y., Marcus, S., Formenti, G., Bergeron, L.A., Song, Z., Bi, X., Bergman, J., Rousselle, M.M.C., Zhou, C., *et al.*: Evolutionary and biomedical insights from a marmoset diploid genome assembly. Nature **594**(7862), 227–233 (2021)

[95] Moorjani, P., Amorim, C.E.G., Arndt, P.F., Przeworski, M.: Variation in the molecular clock of primates. Proceedings of the National Academy of Sciences **113**(38), 10607–10612 (2016)

[96] Grafen, A.: The phylogenetic regression. Philosophical Transactions of the Royal Society of London. B, Biological Sciences **326**(1233), 119–157 (1989)

[97] Felsenstein, J.: Phylogenies and the comparative method. The American Naturalist **125**(1), 1–15 (1985)

[98] Blomberg, S.P., Lefevre, J.G., Wells, J.A., Waterhouse, M.: Independent contrasts and pgls regression estimators are equivalent. Systematic biology **61**(3), 382–391 (2012)

[99] Hall, J.G.: Twinning. The Lancet **362**(9385), 735–743 (2003)

[100] Liu, R., Low, W.Y., Tearle, R., Koren, S., Ghurye, J., Rhie, A., Phillippy, A.M., Rosen, B.D., Bickhart, D.M., Smith, T.P., *et al.*: New insights into mammalian sex chromosome structure and evolution using high-quality sequences from bovine x and y chromosomes. BMC genomics **20**(1), 1–11 (2019)

[101] Raudsepp, T., Chowdhary, B.P.: The eutherian pseudoautosomal region. Cytogenetic and Genome Research **147**(2-3), 81–94 (2015)

[102] Das, P., Chowdhary, B., Raudsepp, T.: Characterization of the bovine

pseudoautosomal region and comparison with sheep, goat, and other mammalian pseudoautosomal regions. Cytogenetic and genome research **126**(1-2), 139–147 (2009)

[103] Skinner, B.M., Lachani, K., Sargent, C.A., Affara, N.A.: Regions of xy homology in the pig x chromosome and the boundary of the pseudoautosomal region. BMC genetics **14**(1), 1–7 (2013)

[104] Shearn, R., Wright, A.E., Mousset, S., Régis, C., Penel, S., Lemaitre, J.-F., Douay, G., Crouau-Roy, B., Lecompte, E., Marais, G.A.: Evolutionary stasis of the pseudoautosomal boundary in strepsirrhine primates. Elife **9**, 63650 (2020)

[105] Schield, D.R., Card, D.C., Hales, N.R., Perry, B.W., Pasquesi, G.M., Blackmon, H., Adams, R.H., Corbin, A.B., Smith, C.F., Ramesh, B., *et al.*: The origins and evolution of chromosomes, dosage compensation, and mechanisms underlying venom regulation in snakes. Genome research **29**(4), 590–601 (2019)

[106] Smeds, L., Kawakami, T., Burri, R., Bolivar, P., Husby, A., Qvarnström, A., Uebbing, S., Ellegren, H.: Genomic identification and characterization of the pseudoautosomal region in highly differentiated avian sex chromosomes. Nature communications **5**(1), 1–7 (2014)