universität wien

DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Evolutionary genomics of avian sex chromosomes“

verfasst von / submitted by

Luohao Xu, MSc

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doctor of Philosophy (PhD)

Wien, 2019 / Vienna 2019

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 794 685 490
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Molekulare Biologie
Betreut von / Supervisor:	Qi Zhou, PhD

Acknowledgements

I thank my supervisor Qi Zhou who kindly recruited me as his PhD student once I received my Master degree, and offered me very exciting projects. During the stay in Vienna, Qi offered great help, from instructing me how to write a research proposal, engaging in discussions about technical details, encouraging me to present my work in seminars or conferences, to revising my manuscripts including this thesis.

The publications of my projects relies on collaborations with other teams, including Martin Irested's lab at Swedish Museum of Natural History, Scott Edwards' lab at Harvard University, and Guojie Zhang's lab at BGI. Mozes Bloom from the Irested lab was very helpful in transferring the raw sequencing data from Sweden to Vienna, and helped the filtering of the datasets. Alexandra Suh and Valentina Peona from Uppsala University provided the repeat library of songbirds which was important for my research. They all participated in editing the manuscripts and the revision processes.

Timothy Sackton and Scott Edwards from Harvard University allowed me to continue to use their serve until the manuscript was accepted. Tim helped me a great deal, from whom I learnt how to work independently and think critically. He also revised my manuscript to make it publishable. Other people from Edwards lab, including Simon Sin and Philip Grayson, also contributed to the manuscript editing.

The Qi Zhou lab in Vienna is young and lovely, and I was lucky to have great colleagues Mujahid Ali, Lubna Younas, Jing Liu, Gabriel Auer, Eva Lehner and Elmira Mohandesan. During the joint seminar with Oleg Simakov lab (Elena Otalora, Hannah Schmidbaur, Nicolas Robert and Nicola Wong) I received many constructive comments. Ulrich Technau lab (Nina Znidaric, Eduard Renfer and Bob Zimmermann) was very supportive during my stay in the department which I deeply appreciate. I thank Hannah for translating the abstract of this thesis into German while travelling.

I also thank my friends who are always supportive and inspiring. Kang Kang, Shanlin Liu, Junjie Bu, Bitao Qiu and Denghui Liu managed to visit me in Vienna for a couple of days for which I am particularly grateful. The company of them, as well as Shiping Liu, Hao Yu, Chunxue Guo, Yuan Deng, Chenran Zhou, Senwei Tang and Rashmit Kaur, has made the past years colorful and joyful.

I give my greatest thanks to my family who firmly support my study and research abroad. My mother had been fighting against illness for years, but never discourage me from studying far away from home. I am deeply sadden that she can no longer see me go through the rest part of my PhD career. My father fully supported my stay in Europe when I decided to pursue an oversee Master degree, and continued to encourage me in my PhD study. My sisters have been very caring. I enjoyed having video chats with them and my nephew Shuoshuo.

The cube cluster from the Dept. of Computational Systems Biology provided most of the computational resources I needed for my research. Thomas Rattei and Florian Goldenberg patiently provided technical supports. Vienna Scientific Cluster (VSC) has also provided some computational resources.

My research and stay in Vienna had been supported by Qi Zhou via his ERC grant. I was later supported by the uni:docs fellowship programme.

List of Papers

I. **Xu, L.**, Auer, G., Peona, V., Suh, A., Deng, Y., Feng, S., Zhang, G., Blom, M.P., Christidis, L., Prost, S. and Irestedt, M. & Zhou, Q. (2019). Dynamic evolutionary history and gene content of sex chromosomes across diverse songbirds. *Nature ecology & evolution*, 3(5), 834.

II. **Xu, L.**, Sin, W. S. Y., Grayson, P., Edwards, S. V., & Sackton, T. B. (2019). Evolutionary dynamics of sex chromosomes of paleognathous birds. *Genome biology and evolution*, 11(8), 2376-2390.

III. **Xu, L.**, & Zhou, Q. (2019). Female-specific and dosage selections restore genes through transpositions onto the degenerated songbird W chromosomes. *bioRxiv*, 692194. *Manuscript submitted*

Table of Contents

Abstract

The processes and underlying forces of the evolution of avian sex chromosomes remain to be elucidated, despite decades of cytogenetic and molecular studies. In recent years, with the increasing availability of avian genomes, it becomes possible to revisit classical questions in avian sex chromosome evolution, including what are the causes and consequences of recombination suppression, at a fine-scale genomic level. In this thesis, I start tackling this task by focusing on two important clades: songbirds and paleognathous birds. The former clade represents more than half of the bird species of diverse morphological, ecological and behavioral traits, while the latter is a basal and unique clade with unusually homomorphic sex chromosomes. Through a comparative analysis of 13 genomes of paleognathous birds, I uncovered various stages of sex chromosome evolution, from complete degeneration in some tinamous to nearly stalled evolution in most ratites that show a large pseudoautosomal region (PAR) that is still recombining. Unexpectedly, I found evidence of reduced efficacy of selection for PAR-linked genes in species with large PARs, likely due to a reduced recombination rate. On the contrary, all the 11 songbird genomes analyzed here have fully differentiated sex chromosomes. I dated each event of recombination suppression in songbird sex chromosomes, and found there are in total four such events and they all occurred before the rapid speciation of songbirds. Interestingly, I found that the genes survived on the heterochromatic W chromosomes, despite in small numbers, are very conserved across songbirds, and their retention is likely due to the selection for dosage balance and their regulatory roles in the genomes. I further discovered 3 Z-to-W transposition events involving 7 haploinsufficient and house-keeping genes. All together, my work on diverse paleognathous birds and songbirds provides new insights into the dynamic evolutionary history of avian sex chromosomes.

Zusammenfassung

Die Prozesse und Kräfte, die der Evolution von Geschlechtschromosomen in Vögeln zugrundeliegen, sind trotz jahrzehntelanger zytogenetischer und molekularer Studien noch nicht geklärt. Durch die Verfügbarkeit von immer mehr Vogelgenomen wird es möglich, klassische Fragen der Evolution von Geschlechtschromosomen in Vögeln auf genomischer Ebene erneut zu untersuchen, einschließlich der Ursachen und Folgen von Rekombinationsunterdrückung.
In dieser Arbeit konzentriere ich mich auf zwei wichtige Vogelgruppen: Singvögel und Urkiefervögel (palaeognathe Vögel). Singvögel repräsentieren mehr als die Hälfte der Vogelarten mit unterschiedlichsten morphologischen, ökologischen und verhaltensbezogenen Merkmalen, während Urkiefervögel eine basale und außergewöhnliche Gruppe mit ungewöhnlich homomorphen Geschlechtschromosomen sind. Durch eine vergleichende Analyse von 13 Genomen von paläontologischen Vögeln entdeckte ich verschiedene Stadien der Evolution von Geschlechtschromosomen. Einige Steißhühner zeigen eine vollständige Degeneration der Geschlechtschromosome, während die meisten Laufvögel fast einem evolutionären Stillstand gleich sind. Sie besitzen eine große pseudoautosomal Region (PAR), die sich immer noch rekombiniert. Überraschenderweise unterliegen PAR-verknüpfte Gene bei Arten mit großen PARs einer geringeren Selektion, wahrscheinlich aufgrund reduzierter Rekombinationsraten.
Alle Singvogel-Genome auf der anderen Seite verfügen über vollständig differenzierte Geschlechtschromosome. Durch Datieren jedes Ereignisses von Rekombinationsunterdrückung in Singvogel-Geschlechtschromosomen konnte ich zeigen, dass es insgesamt vier solcher Ereignisse gab und sie alle vor der Speziation von Singvögeln auftraten. Interessanterweise fand ich heraus, dass die Gene auf den heterochromatischen W-Chromosomen, trotz ihrer geringen Anzahl, bei Singvögeln sehr konserviert sind. Ihre Erhaltung ist wahrscheinlich auf ihre Rolle in Dosisbalance und Regulation im Genom zurückzuführen. Ich entdeckte weiterhin 3 Z-to-W-Transpositionsereignisse die sowohl 7 haplo-insuffiziente als auch „house-keeping“ Gene involvierten.
Zusammenfasssend erlaubt meine Arbeit an verschiedenen palaeognathen Vögeln und Singvögeln neue Einblicke in die dynamische Evolutionsgeschichte der Geschlechtschromosomen von Vögeln.

Chapter 1

Introduction

1.1 A brief history of studies on the avian sex chromosome

Birds have a female-heterogametic sex chromosome system, that is, females have one Z chromosome and one female-specific W chromosome while males have two Z chromosomes. Following the early discovery of the ZW sex chromosome system at the beginning of the 20th century, detailed characterization of avian ZW chromosomes relied on cytogenetic methods in the last century (H. Ellegren 2000). Through comparative chromosomal mapping, researchers found that the avian sex chromosome evolved from an autosomal pair (A. K. Fridolfsson et al. 1998) that is not homologous to the mammalian XY chromosomes (Ezaz et al. 2006); throughout more than 100 million years' evolution of birds, the sex chromosomes have been particularly stable and conserved (Nanda et al. 2008; Shetty, Griffin, and Graves 1999; Nanda et al. 1999).

After the split of two major bird clades, Neognathae and Paleognathae, about 102 million years (MY) ago, they followed two diverged evolutionary paths of sex chromosome evolution. Neognathae contains more than 99% of extant bird species, including Neoaves and Galloanserae (e.g. chicken and duck), in which the sex chromosomes are highly differentiated (Rutkowska, Lagisz, and Nakagawa 2012). In most species, the size of the Z chromosome is similar to the fourth or fifth chromosome. On the contrary, the W chromosomes are gene-poor and heterochromatic, often cytogenetically indistinguishable from other microchromosomes (Graves 2014). Despite almost complete differentiation between the Z and W, a small part of the chromosome is still recombining during meiosis. The recombining part on the sex chromosomes, typically less than 1 Mb, is called the pseudoautosomal region (PAR) (Sarah P. Otto et al. 2011).

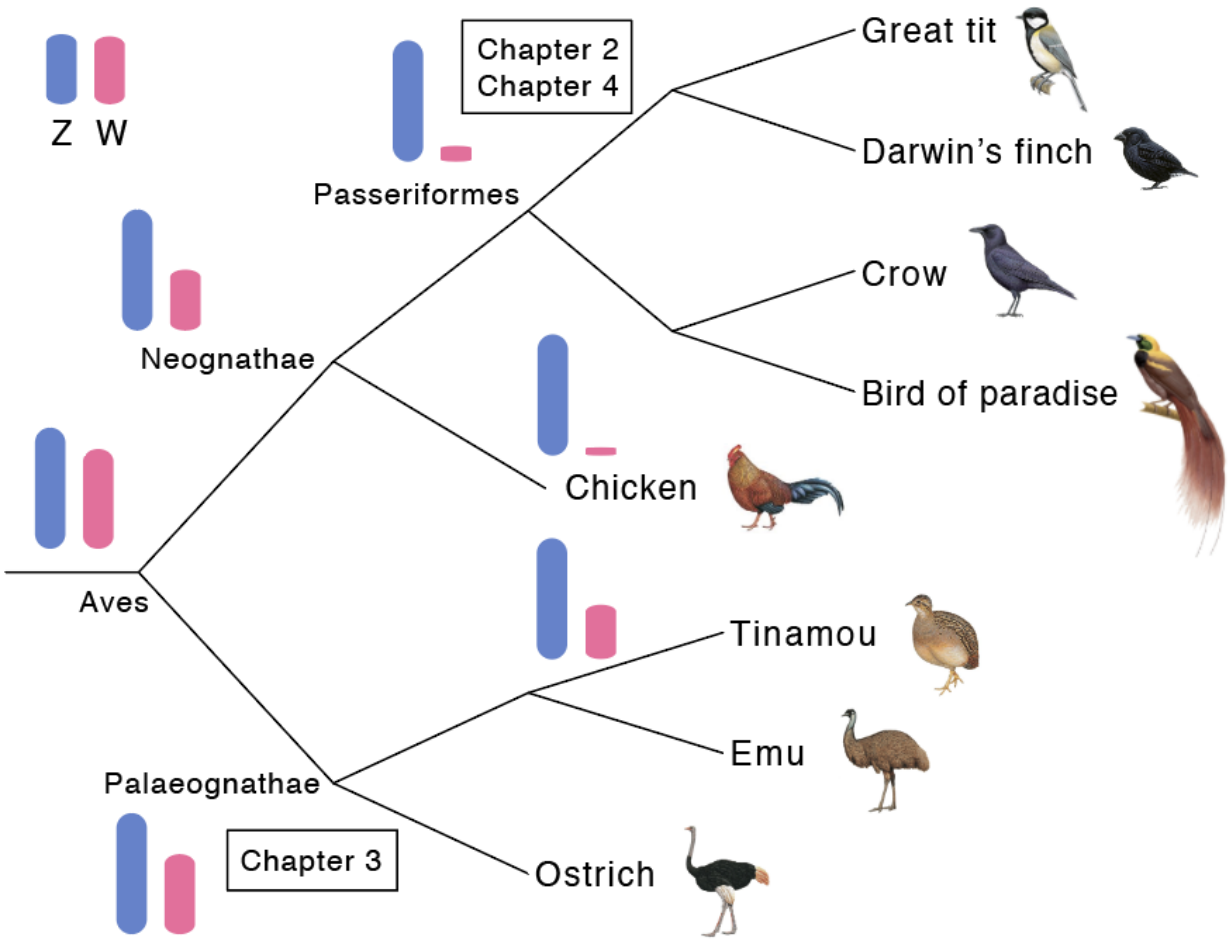


Figure 1. The evolutionary history and pattern of sex chromosome differentialtion in birds. All birds (Aves) share a female-heterogamety (ZW) system. Palaeognathae has a pair of homomorphic sex chromosome except for the tinamou lineage. The published paper in Chapter 3 will address this part. Most songbirds (in Passeriformes) have highly differantiated sex chromosomes. One published paper in Chapter 2 and one manuscript in Chapter 4 will discuss this lineage. The bird illustrations were ordered from https://www.hbw.com/.

The Palaeognathae consists of flightless ratites (e.g. ostrich and emu) and volitant tinamous. In most ratites, the W chromosomes are largely homomorphic to the Z chromosomes (Ansari, Takagi, and Sasaki 1988; Ogawa, Murata, and Mizuno 1998; M. I. Pigozzi and Solari 1999; Stiglec, Ezaz, and Graves 2007), with about two-thirds of the Z chromosome being pseudoautosomal. The tinamous, on the other hand, exhibit different degrees of W chromosome degeneration: while many of them have an intermediate degree of W degeneration (Tsuda et al. 2007; María Inés Pigozzi 2011), some show completely degenerated W chromosomes (Zhou et al. 2014), similar to that of the Neognathae. This suggests that there must be independent

degeneration of the W chromosome in tinamous, paralleling to that in the Neognathae (Judith E. Mank and Ellegren 2007).

Although cytogenetic studies have revealed an overview of avian sex chromosome evolution, a detailed evolutionary history could not be accurately inferred without the tool of molecular evolution. Starting from this century, DNA sequencing has demonstrated its power in uncovering a finer picture of the evolution of the avian sex chromosome. In 2001, Ellegren and Carmichael inferred independent restriction of recombination between the Z and W in different bird lineages, by comparing the ZW divergence of a single gene (*ATP5A1*) across species (H. Ellegren and Carmichael 2001). Later on, by including four more ZW gene pairs (gametologs), Handley et al. (2004) identified two 'evolutionary strata' on the chicken Z chromosome, due to the occurrence of recombination suppression between sex chromosomes at two different evolutionary timepoints (Handley, Ceplitis, and Ellegren 2004).

The first genome of a bird species, chicken (*Gallus gallus*), became available in 2004 (Consortium and International Chicken Genome Sequencing Consortium 2004), following which researchers were able to identify more W-linked gametologs, leading to the identification of more than three evolutionary strata in chicken (Nam and Ellegren 2008). More recently, with more avian genomes becoming available, including mallard duck (*Anas platyrhynchos*) and turkey (*Meleagris gallopavo*), independent and multiple formations of evolutionary strata have been characterized (Wright et al. 2014).

A more dedicated effort on depicting the evolutionary trajectories of avian sex chromosome (Zhou et al. 2014) has been made after the genomes of 48 birds were sequenced (Zhang et al. 2014; Jarvis et al. 2014). The cost of next generation sequencing (NGS) that reduced dramatically since 2008 has made this effort possible. In their study, Zhou and colleagues carefully demarcated the boundaries of evolutionary strata on the Z chromosomes of 17 birds, and revealed between two to four evolutionary strata across bird taxa, with the oldest stratum shared by all birds (Zhou et al. 2014).

With the cost of NGS further coming down and more bird genomes becoming available, the study on avian sex chromosome has entered the genomic era. However, more research questions have been raised than satisfactorily addressed. For instance, how did the recombination between the sex chromosomes become suppressed? Why do some avian lineage, e.g. ratites, retain a pair of homomorphic sex chromosome? Why haven't complete dosage compensation evolved in birds? What is the genetic response to W-linked gene loss?

Below I will briefly summarise some of the current research topics on avian sex chromosome evolution that I will address in this thesis, under a framework of evolutionary genomics.

1.2 Recombination suppression

In both eutherian mammals and birds (except for ratites), the sex chromosome pairs are highly differentiated (Bachtrog et al. 2014; Cortez et al. 2014). This is mainly due to a lack of recombination between the sex chromosomes in the heterogametic sex, through several processes (Bachtrog 2013), including genetic hitchhiking (W. R. Rice 1987), Muller's ratchet (B. Charlesworth 1978) and Ruby in the rubbish (Orr and Kim 1998). It has been suggested that the suppression of recombination between sex chromosome pairs is needed to maintain the linkage of sex determining gene and sexually antagonistic genes on the sex-limited chromosome (William R. Rice 1987; S. P. Otto 2014; D. Charlesworth, Charlesworth, and Marais 2005). The sexually antagonistic genes are those that benefit one sex but may harm the other sex, therefore its restriction within the non-recombining region of the Y or W chromosome is essential to avoid sexual conflicts (Charlesworth 1996). The suppression of recombination between sex chromosomes can occur multiple times to allow for the expansion of non-recombining regions and the addition of sexually antagonistic loci (Bergero and Charlesworth 2009). Each time a sex-linked region is suppressed for recombination, a new evolutionary stratum is formed.

While the sexual antagonism (SA) model can almost perfectly explain the evolution of recombination suppression, Charlesworth et al. (2014) suggested only when the sexually antagonistic selection is very strong will the recombination suppression be favored (Charlesworth, Jordan, and Charlesworth 2014). Moreover, empirical evidence for the SA model is still limited and mixed (Wright et al. 2016; Ponnikas et al. 2018). A recent study in guppies provided new evidence that recombination suppression is favored to maintain the linkage of male-coloration locus and male-determining locus, supporting the SA model (Wright et al. 2017). However, another research group suggested there is occasional recombination between male XY while most recombination events are concentrated at the chromosomal tips in males; the researchers further suggested the very low male recombination rate helps maintain the high frequency of male-beneficial coloration allele on the Y chromosome (Bergero et al. 2019). Another study in Ranidae tree frog, similarly, argues against the role of sexual antagonism in driving the restriction of recombination (Rodrigues et al. 2018). In addition, recent theoretical work suggested that in many lineages, particularly in lower vertebrates, the suppression of recombination in XY males may even be harmful to males (Cavoto et al. 2018).

Another often-debated topic over recombination suppression is how it takes place. Considering the strata-like pattern of ZW divergence along the Z chromosome, physical barriers of recombination such as inversions, seem to be a plausible cause of recombination suppression (Wright et al. 2016; Ross et al. 2005). However, again, empirical evidence is difficult to obtain. One of the difficulties is in most well studied systems, the sex chromosomes are old and already fully degenerated, therefore the inversions that we observe now on the chromosome could have occurred after the suppression of recombination (Bergero and Charlesworth 2009). In birds, Zhou et al. observed a large-scale inversion on the Z chromosome at the ancestor Neognathae, coincident with the onset of recombination suppression of the Neognathae-specific evolutionary stratum (Zhou et al. 2014). However, it is still unclear if the inversion was the direct trigger of recombination suppression.

Alternatively, recombination can be halted without chromosomal rearrangements. Instead, graduate loss of recombination can be achieved through a genetic modifier of the recombination rate (Choi and Henderson 2015) or changes in chromatin structure (Marand et al. 2017). This scenario of recombination loss has been supported by the study in threespine sticklebacks (Natri, Shikano, and Merilä 2013), *Silene* (Bergero et al. 2013) and a fungus (Sun et al. 2017). More recently, there is increasing awareness of the role of transposable elements in the regulation of recombination (Kent, Uzunović, and Wright 2017). Particularly, epigenetic modifications of transposable elements can be associated with the suppression of recombination (Underwood and Choi 2019).

1.3 Evolutionary strata

Once the recombination between the Z and W is halted, the W chromosome is expected to degenerate. Such a course can take place multiple times in a punctuated manner (Lahn and Page 1999b). Because of these processes, different regions on the Z chromosome may have different ages of recombination suppression with the W chromosomes, therefore different degrees of divergence between the Z and W. In many taxa, Z- (or X-) linked regions with similar divergence levels tend to cluster together. This pattern of spatial clusters of sequence divergence on the Z chromosome is called ‘evolutionary strata’. In humans, at least four strata have been identified, and they exhibit a linear organization on the X chromosome (Skaletsky et al. 2003).

In birds, the first stratum (S0) evolved at the ancestor of all birds (Aves) at least 102 MY ago. This stratum is about 18 Mb long, containing the candidate sex determining gene *Dmrt1 (Zhou et al. 2014)*. This is in line with the canonical model of sex chromosome evolution that predicts the involvement of recombination restriction at the sex-determining loci and its surrounding regions, at the early stage of sex chromosome evolution (D. Charlesworth, Charlesworth, and Marais 2005). In Palaeognathae, this stratum is located at the end of the Z chromosome. In Neognathae, however, it has been relocated and scattering along the middle of the Z chromosome. This is likely due to a large-scale inversion at the ancestor of Neognathae followed by frequent smaller-scale inversions that reshuffled the organization of S0. The post-recombination-suppression inversions is probably fixed by genetic drift due to reduced efficacy of selection for gene synteny on the Z chromosome (Wright et al. 2016).

Following their divergence, the Paleognathae and Neognathae evolved additional evolutionary strata on their Z chromosome independently, at a different rate. In most ratites, only one additional stratum has been formed during their almost 100 MY's evolution. Moreover, this stratum is relatively smaller, about only 10 Mb (Zhou et al. 2014; B. Vicoso, Kaiser, and Bachtrog 2013). The overall picture in tinamous is unclear, but in white-throated tinamou three evolutionary strata have been demarcated (Zhou et al. 2014). In Neognathae, the second stratum (S1) is a bit larger than S0, likely formed by a Z-linked inversion - the same inversion that brought the S0 into the middle of the Z chromosome. This stratum was estimated to occur 89 MY ago, at the ancestor of Neoaves (Zhou et al. 2014). Similar to the scenario in S0, frequent rearrangements have drastically disrupted its synteny with the ancestral Z chromosome.

After the split of the sister groups Galloanserae and Neoaves 89 MA ago, the third stratum (S2) seems to have independently formed in the two clades (Zhou et al. 2014). The S2 in Galloanserae spreads into almost the entire remaining recombining part of the Z chromosome, and appears to evolve at a very early branching of the Galloanserae. The size of Neoaves S2 is similar to Neognathae S1, about 20 Mb. The S2 of both Galloanserae and Neoaves were formed without a Z-linked inversion, but the contribution of a W-linked inversion have not been ruled out.

Finally, the formation of the last stratum (S3), likely independently in various Neoaves lineages, leaves only a very small part of the Z chromosome as the PAR (Zhou et al. 2014). This stratum is perhaps the only one showing size variations among Neoaves birds. For instance, the S3 in white-tailed tropicbird is absent, making it one of the very few Neoaves birds having a relatively

large PAR. Most songbirds, including collared flycatcher (Smeds et al. 2014) and zebra finch (Singhal et al. 2015), seem to possess a very small PAR shorter than 700 kb. However, while more than half of extant bird species belong to songbirds, very few of them have been investigated for their sex chromosome evolution.

1.4 Pseudoautosomal region

The PAR is the only part of the bird ZW chromosomes that is still recombining in females. While no differentiation between sexes is expected, the PAR shows distinct features compared with autosomes. One of the prominent features of PAR is perhaps its usually high recombination rate. However, this is likely due to the fact that most well studied PARs are very short, and that there is at least one obligate crossover (Mohandas et al. 1992) needed in the heterogametic sex that is restricted to the small PAR. For instance, researchers found the PAR in collared flycatcher is only 630 kb in size, and reported a more than 30 times increase of recombination rate relative to autosomes (Smeds et al. 2014). The high recombination rate, in turn, leads to high GC content, low repeat density, high gene density and a low evolutionary rate of the PAR (Smeds et al. 2014).

The Palaeognathae, usually having a large PAR, is not expected to display such a pattern. Indeed, the recombination rate was found not particularly high in the females of ostrich (Yazdi and Ellegren 2018). Considering the even lower recombination rate in the males, the sex-average recombination rate is likely lower than autosomes (Yazdi 2019). However, more effort is needed to estimate the recombination rate of the PAR in ratites.

The PAR has often been a subject for the study of sexually antagonistic selection (Sarah P. Otto et al. 2011). If different alleles are favored by males versus females (sexual antagonism), particularly for those close to the PAR boundary, selection for reduced recombination is needed to preserve the linkage of the sexually antagonistic allele and the fully sex-linked regions (Kirkpatrick and Guerrero 2014; D. Charlesworth and Charlesworth 1980; Charlesworth, Jordan, and Charlesworth 2014). This ongoing process can leave a signal of a higher-than-expected genetic diversity in the PAR. Studies on the PAR of a *Silene* species have provided empirical evidence supporting the role of sexual antagonism in maintaining an excess of polymorphisms in the PAR (Guirao-Rico, Sánchez-Gracia, and Charlesworth 2017; Qiu et al. 2016). However, evidence from other organisms is limited, and a recent theoretical study suggested only under certain conditions may sexually antagonistic selection play a role(Sarah P. Otto 2019). In birds, studies have failed to demonstrate the role of sexually antagonistic selection in shaping the

nucleotide diversity of the PAR, in species with both small (collared flycatcher) (Smeds et al. 2014) and large (ostrich) (Yazdi and Ellegren 2018) PARs. Particularly, the high recombination rate observed close to the PAR boundary in ostrich females can hinder the formation of full linkage of the SA allele and sex-determining region, thus preventing the shrinking of the PAR (Yazdi and Ellegren 2018).

1.5 Faster-Z evolution

One of the consequences of recombination suppression is that the Z chromosome ultimately becomes hemizygous in females. On one hand, the recessive hemizygous alleles on the Z are more likely to be selected if they are beneficial to females (B. Charlesworth, Coyne, and Barton 1987; Beatriz Vicoso and Charlesworth 2006), leading to their faster rate of fixation(faster-Z evolution). This has been supported by the result in silk moth (Sackton et al. 2014). A similar scenario can also apply to male-heterogametic systems. For instance, in *Drosophila*, faster-X evolution due to positive selection has been frequently observed (Meisel and Connallon 2013; Connallon 2007).

On the other hand, the degeneration of the W chromosome reduces the number of the carriers of Z-linked genes, therefore in a population with a balanced sex ratio, the number of Z chromosome becomes ¾ of that of autosomes. This leads to a ¼ smaller effective population size of the Z chromosome, and ultimately reduced the efficacy of selection on Z-linked genes (Judith E. Mank et al. 2010). As a consequence, genetic drift has a greater effect on fixing slightly deleterious mutations on the Z, causing accelerated nonsynonymous substitution rates relative to synonymous substitution rates. In birds, this has been suggested as the major driving force behind the faster-Z evolution (J. E. Mank, Nam, and Ellegren 2010; Wang et al. 2014). Moreover, the faster-Z effect driven by genetic drift can be stronger when there is more variance in male mating success, such as in promiscuous species (Wright et al. 2015).

Despite the prevalence of faster-Z (or faster -X) in diverse taxa (Bechsgaard et al. 2019), when other evolutionary processes are at play, the faster-Z effect can be balanced. For instance, Rousselle et al. detected enhanced purifying selection against slightly deleterious mutations on the hemizygous Z chromosome, resulting in no detectable faster-Z effect (Rousselle et al. 2016). Moreover, the strength of the faster-Z effect in birds seems to be associated with the age of sex chromosome strata. One study that involves a comparative analysis of 48 birds shows no faster-Z effect of genes from the old stratum (Wang et al. 2014), while a recent study reports

very weak support of faster-Z evolution on the neo-sex chromosome in Sylvioidea (Leroy et al. 2019).

1.6 Transposable elements

The compact bird genome, in general, contains a small portion of repetitive sequences, typically less than 10% (Kapusta, Suh, and Feschotte 2017; Zhang et al. 2014). Transposable elements (TEs) are mobile repeat elements in the genomes, including LTRs (long terminal repeats), LINEs (long interspersed nuclear elements), SINEs (short interspersed nuclear elements) and DNA transposons. A typical avian genome has a low content of SINEs and DNA transposons, while LTRs and LINEs are relatively more abundant (Weissensteiner and Suh 2019). Because of the absence of recombination (thus reduced efficacy of purging TEs), the W chromosome is usually highly repetitive and heterochromatic. The most abundant TE family on the W chromosome appears to be the LTR (Kapusta and Suh 2017). However, since most avian W chromosomes are old, the general pattern and rate of TE accumulation on the W chromosome are unclear.

While the Z chromosome has homologous recombination in males, it has accumulated TE at a higher rate compared with autosomes (Kapusta and Suh 2017). This is likely due to its reduced efficacy of selection of the Z chromosomes. The distribution of TEs on the Z chromosome is highly heterogeneous (Kapusta and Suh 2017), suggesting the presence of other evolutionary forces in regulating TE proliferation, such as local recombination rate and chromatin states. It is unclear if the landscape of TE distribution on the Z chromosome is stable over time, and if the heterogeneous distribution is a derived pattern after the suppression recombination or an ancestral pattern. If the latter were true, the locally accumulated TE might have a role in modulating the distribution of recombination which can facilitate the evolution of sex chromosomes.

The interplay between TEs and W (or Y) chromosome degeneration is more complicated (Śliwińska, Martyka, and Tryjanowski 2016; Chalopin et al. 2015). Rapid expansions of TEs may have occurred in the early-stage of sex chromosome evolution, revealed by a study on young sex chromosomes (Mahajan et al. 2018). The accumulation of TEs in turn increased the chance of TE-mediated rearrangements, including deletions. This process can not only delete coding sequences of the W, but also promote further recombination suppression through chromosomal changes, for instance, inversions. However, again this hypothesis is difficult to test in most Neoaves birds where the W chromosome is already fully degenerated.

1.7 W-chromosome gene content

The first W-linked gametolog *CHD1W* was identified 23 years ago (H. Ellegren 1996), which is later found conserved among bird W chromosomes. Due to the length difference of some introns between Z- and W-linked gametologs of *CHD1*, it has since been widely used as a molecular marker for sexing various bird species (Griffiths, Daan, and Dijkstra 1996; A.-K. Fridolfsson and Ellegren 1999). Genome sequencing, including sequencing of the transcriptomes and BAC clones (mainly the euchromatic parts), has led to the identification of about 28 W-gametologs in chicken (Bellott et al. 2017; Consortium and International Chicken Genome Sequencing Consortium 2004; Wright et al. 2014). Most of those genes are single-copy with an intact open reading frame, except for *HINT* which has been amplified into multiple copies (Backström et al. 2005; Bellott et al. 2017). Interestingly, a similar pattern has also been reported in collared flycatcher (Smeds et al. 2015), despite a long divergence time since the split of Neoaves and Galloanserae. On the W chromosome of collared flycatcher, 43 single-copy genes and amplicon *HINT* have been identified (Smeds et al. 2015). Although the number is slightly higher than that in chicken, it is much less than the homologous Z chromosome which harbors more than 700 genes, indicating massive gene loss of the W chromosome.

The comparison of the gene content of W chromosomes between chicken and collared flycatcher showed another pattern, that is, the convergent retention of W-gametologs. This suggests the retention of genes on the W chromosome is not random and is governed by selection. By comparing the dosage sensitivity of the retained and lost genes, Bellott et al. found the retained genes show significantly higher dosage sensitivity (measured by haploinsufficiency scores) than the lost genes, in both mammals (Bellott et al. 2014) and chicken (Bellott et al. 2017). Retention of dosage-sensitive genes is particularly important in birds which have not evolved a mechanism of global dosage compensation.

1.8 Sex-specific selection

The cessation of recombination between the Z and W chromosome makes them favorable genomic regions to accumulate genes with sex-specific functions. The W chromosome is inherited only in females, so female-beneficial alleles can be accumulated and expressed without affecting males. Similarly, since the Z chromosome spends more time in males than in females, it is expected to be ‘masculinized’ (Beatriz Vicoso and Charlesworth 2006). These theoretical predictions have been frequently supported by studies of the XY systems in *Drosophila* (Beatriz Vicoso and Bachtrog 2015; Zhou and Bachtrog 2012) and mammals (Graves 2006) (in the opposite way), but empirical evidence for birds is limited.

As mentioned above, most W-linked genes identified so far are gametologs, which means they have a homologous copy on the Z. So far a W-chromosome specific gene has not been reported in birds. In contrast, novel Y-linked genes or gene families have been reported in human (Lahn and Page 1999a), cat (Li et al. 2013), dog (Li et al. 2013), horse (Janečka et al. 2018) and *Drosophila* (Koerich et al. 2008; Tobler, Nolte, and Schlötterer 2017). Moreover, in many mammalian species, some Y-linked genes are highly amplified, and are usually testis-specific (Soh et al. 2014; Bachtrog 2013; Hughes and Page 2015). Those Y-specific genes or amplicons are likely a result of male-specific selection, as suggested by their testis-specific or testis-biased expression. Although the avian W chromosomes also harbor an amplified gene *HINTW*, there is limited evidence supporting the effect of female-specific selection on this gene (C. A. Smith, Roeszler, and Sinclair 2009; Smeds et al. 2015).

On the contrary, the avian Z chromosome seems to be enriched for male-biased genes (Wright, Moghadam, and Mank 2012). A frequent movement of male-biased genes into the Z chromosome is also reported in chicken, despite a generally low frequency of inter-chromosome gene movement in birds (Hans Ellegren 2011). A more complete assembly of chicken Z chromosome uncovered amplification of four genes at the end of the Z chromosome that show testis-specific expression (Bellott et al. 2010), a similar pattern that has been seen on the human X chromosome (Ross et al. 2005; Saifi and Chandra 1999). Whether this pattern can also be found in other birds remains to be tested. Particularly, recent gene duplication is difficult to be detected in Illumina-based genome assembly (Peona, Weissensteiner, and Suh 2018), but the recent development of long-read sequencing has potentials to help reveal hidden genes and genomic sequences of the sex chromosomes.

1.9 Dosage compensation

In most taxa with differentiated sex chromosomes, there is a need to compensate for imbalanced gene dosage in the heterogametic sex. Interestingly, diverse mechanisms have been evolved to tackle this issue, including random inactivation of one of the female X chromosomes in mammals (Nguyen and Disteche 2006; Pessia, Engelstädter, and Marais 2014), up-regulation of male X chromosome in *Drosophila* (Meiklejohn et al. 2011), and down-regulation of female X chromosomes in *C. elegans (Meyer and Casson 1986)*. In birds, global dosage compensation is probably absent (Graves 2014; Gu and Walters 2017). Instead, gene-by-gene dosage compensation has been reported in a number of bird species (Itoh et al. 2007, 2010; Uebbing et al. 2013; Wolf and Bryk 2011; Moghadam et al. 2013; Adolfsson and Ellegren

2013). Partial dosage compensation has also been confirmed at the protein level in a study in chicken which also reported post-transcriptional regulation of dosage compensation, in a gene-by-gene manner (Uebbing et al. 2015).

It is still unclear why complete dosage compensation has not evolved in birds. This is probably not linked to female-heterogamety, as complete dosage compensation has been observed in other female-heterogametic taxa (Gu, Walters, and Knipple 2017; Huylmans, Macon, and Vicoso 2017; G. Smith et al. 2014; Walters and Hardcastle 2011) and a lack of complete dosage compensation has also been found in male-heterogametic taxa (Julien et al. 2012; White, Kitano, and Peichel 2015; Hough et al. 2014). As mentioned above, the avian Z chromosome has been 'masculinized' due to male-specific selection, therefore up-regulation of the Z is perhaps not favoured by females (Naurin et al. 2010). This hypothesis has been supported by a study in chicken by showing conflicting effects of dosage selection and male-specific selection on gene expression of the Z chromosome (Wright, Moghadam, and Mank 2012). Furthermore, theoretical modeling suggests the extent of dosage compensation is influenced by the relative strength of sexual selection, and this has been supported with empirical evidence that in tissues with stronger sexual selection by females, dosage compensation is more effective (Mullon et al. 2015).

Although under debate, it is generally assumed that the sex determining gene in birds is *Dmrt1*, and sex is determined by the dose of *Dmrt1* (Hirst et al. 2017; C. A. Smith et al. 2009). This is maybe one of the reasons why global complete dosage compensation is not selected, because it would skew the sex ratio if the dose of *Dmrt1* is balanced between sexes. Given not all Z-linked genes are dosage sensitive, chromosome-wise complete dosage compensation is maybe not necessary, as long as the dosage balance can be achieved for dosage-sensitive genes (White, Kitano, and Peichel 2015; Judith E. Mank 2009; J. E. Mank and Ellegren 2009). This notion becomes more compelling when it is found that the majority of the retained gametologs on chicken W chromosome are dosage sensitive (Bellott et al. 2017).

Reference

Adolfsson, Sofia, and Hans Ellegren. 2013. “Lack of Dosage Compensation Accompanies the Arrested Stage of Sex Chromosome Evolution in Ostriches.” *Molecular Biology and Evolution* 30 (4): 806–10.

Ansari, H. A., N. Takagi, and M. Sasaki. 1988. “Morphological Differentiation of Sex Chromosomes in Three Species of Ratite Birds.” *Cytogenetic and Genome Research* 47 (4): 185–88.

Bachtrog, Doris. 2013. “Y-Chromosome Evolution: Emerging Insights into Processes of Y-Chromosome Degeneration.” *Nature Reviews. Genetics* 14 (2): 113–24.

Bachtrog, Doris, Judith E. Mank, Catherine L. Peichel, Mark Kirkpatrick, Sarah P. Otto, Tia-Lynn Ashman, Matthew W. Hahn, et al. 2014. “Sex Determination: Why so Many Ways of Doing It?” *PLoS Biology* 12 (7): e1001899.

Backström, Niclas, Helene Ceplitis, Sofia Berlin, and Hans Ellegren. 2005. “Gene Conversion Drives the Evolution of HINTW, an Ampliconic Gene on the Female-Specific Avian W Chromosome.” *Molecular Biology and Evolution* 22 (10): 1992–99.

Bechsgaard, Jesper, Mads Fristrup Schou, Bram Vanthournout, Frederik Hendrickx, Bjarne Knudsen, Virginia Settepani, Mikkel Heide Schierup, and Trine Bilde. 2019. “Evidence for Faster X Chromosome Evolution in Spiders.” *Molecular Biology and Evolution* 36 (6): 1281–93.

Bellott, Daniel W., Jennifer F. Hughes, Helen Skaletsky, Laura G. Brown, Tatyana Pyntikova, Ting-Jan Cho, Natalia Koutseva, et al. 2014. “Mammalian Y Chromosomes Retain Widely Expressed Dosage-Sensitive Regulators.” *Nature* 508 (7497): 494–99.

Bellott, Daniel W., Helen Skaletsky, Ting-Jan Cho, Laura Brown, Devin Locke, Nancy Chen, Svetlana Galkina, et al. 2017. “Avian W and Mammalian Y Chromosomes Convergently Retained Dosage-Sensitive Regulators.” *Nature Genetics* 49 (3): 387–94.

Bellott, Daniel W., Helen Skaletsky, Tatyana Pyntikova, Elaine R. Mardis, Tina Graves, Colin Kremitzki, Laura G. Brown, et al. 2010. “Convergent Evolution of Chicken Z and Human X Chromosomes by Expansion and Gene Acquisition.” *Nature* 466 (7306): 612–16.

Bergero, Roberta, and Deborah Charlesworth. 2009. “The Evolution of Restricted Recombination in Sex Chromosomes.” *Trends in Ecology & Evolution* 24 (2): 94–102.

Bergero, Roberta, Jim Gardner, Beth Bader, Lengxob Yong, and Deborah Charlesworth. 2019. “Exaggerated Heterochiasmy in a Fish with Sex-Linked Male Coloration Polymorphisms.” *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.1818486116.

Bergero, Roberta, Suo Qiu, Alan Forrest, Helen Borthwick, and Deborah Charlesworth. 2013. “Expansion of the Pseudo-Autosomal Region and Ongoing Recombination Suppression in the Silene Latifolia Sex Chromosomes.” *Genetics* 194 (3): 673–86.

Cavoto, E., S. Neuenschwander, J. Goudet, and N. Perrin. 2018. “Sex-Antagonistic Genes, XY Recombination and Feminized Y Chromosomes.” *Journal of Evolutionary Biology* 31 (3): 416–27.

Chalopin, Domitille, Jean-Nicolas Volff, Delphine Galiana, Jennifer L. Anderson, and Manfred Schartl. 2015. “Transposable Elements and Early Evolution of Sex Chromosomes in Fish.” *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology* 23 (3): 545–60.

Charlesworth, B. 1978. “Model for Evolution of Y Chromosomes and Dosage Compensation.” *Proceedings of the National Academy of Sciences of the United States of America* 75 (11): 5618–22.

Charlesworth, B., J. A. Coyne, and N. H. Barton. 1987. “The Relative Rates of Evolution of Sex Chromosomes and Autosomes.” *The American Naturalist* 130 (1): 113–46.

Charlesworth, Brian. 1996. “The Evolution of Chromosomal Sex Determination and Dosage Compensation.” *Current Biology* 6 (2): 149–62.

Charlesworth, Brian, Crispin Y. Jordan, and Deborah Charlesworth. 2014. “The Evolutionary Dynamics of Sexually Antagonistic Mutations in Pseudoautosomal Regions of Sex Chromosomes.” *Evolution; International Journal of Organic Evolution* 68 (5): 1339–50.
Charlesworth, D., and B. Charlesworth. 1980. “Sex Differences in Fitness and Selection for Centric Fusions between Sex-Chromosomes and Autosomes.” *Genetical Research* 35 (2): 205–14.
Charlesworth, D., B. Charlesworth, and G. Marais. 2005. “Steps in the Evolution of Heteromorphic Sex Chromosomes.” *Heredity* 95 (2): 118–28.
Choi, Kyuha, and Ian R. Henderson. 2015. “Meiotic Recombination Hotspots - a Comparative View.” *The Plant Journal: For Cell and Molecular Biology* 83 (1): 52–61.
Connallon, Tim. 2007. “Adaptive Protein Evolution of X-Linked and Autosomal Genes in Drosophila: Implications for Faster-X Hypotheses.” *Molecular Biology and Evolution* 24 (11): 2566–72.
Consortium, International Chicken Genome Sequencing, and International Chicken Genome Sequencing Consortium. 2004. “Sequence and Comparative Analysis of the Chicken Genome Provide Unique Perspectives on Vertebrate Evolution.” *Nature* 432 (7018): 695–716.
Cortez, Diego, Ray Marin, Deborah Toledo-Flores, Laure Froidevaux, Angélica Liechti, Paul D. Waters, Frank Grützner, and Henrik Kaessmann. 2014. “Origins and Functional Evolution of Y Chromosomes across Mammals.” *Nature* 508 (7497): 488–93.
Ellegren, H. 1996. “First Gene on the Avian W Chromosome (CHD) Provides a Tag for Universal Sexing of Non-Ratite Birds.” *Proceedings. Biological Sciences / The Royal Society* 263 (1377): 1635–41.
———. 2000. “Evolution of the Avian Sex Chromosomes and Their Role in Sex Determination.” *Trends in Ecology & Evolution* 15 (5): 188–92.
Ellegren, Hans. 2011. “Emergence of Male-Biased Genes on the Chicken Z-Chromosome: Sex-Chromosome Contrasts between Male and Female Heterogametic Systems.” *Genome Research* 21 (12): 2082–86.
Ellegren, H., and A. Carmichael. 2001. “Multiple and Independent Cessation of Recombination between Avian Sex Chromosomes.” *Genetics* 158 (1): 325–31.
Ezaz, Tariq, Rami Stiglec, Frederic Veyrunes, and Jennifer A. Marshall Graves. 2006. “Relationships between Vertebrate ZW and XY Sex Chromosome Systems.” *Current Biology* 16 (17): R736–43.
Fridolfsson, A. K., H. Cheng, N. G. Copeland, N. A. Jenkins, H. C. Liu, T. Raudsepp, T. Woodage, B. Chowdhary, J. Halverson, and H. Ellegren. 1998. “Evolution of the Avian Sex Chromosomes from an Ancestral Pair of Autosomes.” *Proceedings of the National Academy of Sciences of the United States of America* 95 (14): 8147–52.
Fridolfsson, Anna-Karin, and Hans Ellegren. 1999. “A Simple and Universal Method for Molecular Sexing of Non-Ratite Birds.” *Journal of Avian Biology* 30 (1): 116.
Graves, Jennifer A. Marshall. 2006. “Sex Chromosome Specialization and Degeneration in Mammals.” *Cell* 124 (5): 901–14.
———. 2014. “Avian Sex, Sex Chromosomes, and Dosage Compensation in the Age of Genomics.” *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology* 22 (1): 45–57.
Griffiths, R., S. Daan, and C. Dijkstra. 1996. “Sex Identification in Birds Using Two CHD Genes.” *Proceedings. Biological Sciences / The Royal Society* 263 (1374): 1251–56.
Guirao-Rico, Sara, Alejandro Sánchez-Gracia, and Deborah Charlesworth. 2017. “Sequence Diversity Patterns Suggesting Balancing Selection in Partially Sex-Linked Genes of the plantSilene Latifoliaare Not Generated by Demographic History or Gene Flow.” *Molecular Ecology* 26 (5): 1357–70.
Gu, Liuqi, and James R. Walters. 2017. “Evolution of Sex Chromosome Dosage Compensation in Animals: A Beautiful Theory, Undermined by Facts and Bedeviled by Details.” *Genome Biology and Evolution* 9 (9): 2461–76.

Gu, Liuqi, James R. Walters, and Douglas C. Knipple. 2017. “Conserved Patterns of Sex Chromosome Dosage Compensation in the Lepidoptera (WZ/ZZ): Insights from a Moth Neo-Z Chromosome.” *Genome Biology and Evolution* 9 (3): 802–16.
Handley, Lori-Jayne Lawson, Helene Ceplitis, and Hans Ellegren. 2004. “Evolutionary Strata on the Chicken Z Chromosome: Implications for Sex Chromosome Evolution.” *Genetics* 167 (1): 367–76.
Hirst, Claire E., Andrew T. Major, Katie L. Ayers, Rosie J. Brown, Mylene Mariette, Timothy B. Sackton, and Craig A. Smith. 2017. “Sex Reversal and Comparative Data Undermine the W Chromosome and Support Z-Linked DMRT1 as the Regulator of Gonadal Sex Differentiation in Birds.” *Endocrinology* 158 (9): 2970–87.
Hough, J., J. D. Hollister, W. Wang, S. C. H. Barrett, and S. I. Wright. 2014. “Genetic Degeneration of Old and Young Y Chromosomes in the Flowering Plant Rumex Hastatulus.” *Proceedings of the National Academy of Sciences* 111 (21): 7713–18.
Hughes, Jennifer F., and David C. Page. 2015. “The Biology and Evolution of Mammalian Y Chromosomes.” *Annual Review of Genetics* 49 (1): 507–27.
Huylmans, Ann Kathrin, Ariana Macon, and Beatriz Vicoso. 2017. “Global Dosage Compensation Is Ubiquitous in Lepidoptera, but Counteracted by the Masculinization of the Z Chromosome.” *Molecular Biology and Evolution* 34 (10): 2637–49.
Itoh, Yuichiro, Esther Melamed, Xia Yang, Kathy Kampf, Susanna Wang, Nadir Yehya, Atila Van Nas, et al. 2007. “Dosage Compensation Is Less Effective in Birds than in Mammals.” *Journal of Biology* 6 (1): 2.
Itoh, Yuichiro, Kirstin Replogle, Yong-Hwan Kim, Juli Wade, David F. Clayton, and Arthur P. Arnold. 2010. “Sex Bias and Dosage Compensation in the Zebra Finch versus Chicken Genomes: General and Specialized Patterns among Birds.” *Genome Research* 20 (4): 512–18.
Janečka, Jan E., Brian W. Davis, Sharmila Ghosh, Nandina Paria, Pranab J. Das, Ludovic Orlando, Mikkel Schubert, et al. 2018. “Horse Y Chromosome Assembly Displays Unique Evolutionary Features and Putative Stallion Fertility Genes.” *Nature Communications* 9 (1). https://doi.org/10.1038/s41467-018-05290-6.
Jarvis, Erich D., Siavash Mirarab, Andre J. Aberer, Bo Li, Peter Houde, Cai Li, Simon Y. W. Ho, et al. 2014. “Whole-Genome Analyses Resolve Early Branches in the Tree of Life of Modern Birds.” *Science* 346 (6215): 1320–31.
Julien, Philippe, David Brawand, Magali Soumillon, Anamaria Necsulea, Angélica Liechti, Frédéric Schütz, Tasman Daish, Frank Grützner, and Henrik Kaessmann. 2012. “Mechanisms and Evolutionary Patterns of Mammalian and Avian Dosage Compensation.” *PLoS Biology* 10 (5): e1001328.
Kapusta, Aurélie, and Alexander Suh. 2017. “Evolution of Bird Genomes-a Transposon’s-Eye View.” *Annals of the New York Academy of Sciences* 1389 (1): 164–85.
Kapusta, Aurélie, Alexander Suh, and Cédric Feschotte. 2017. “Dynamics of Genome Size Evolution in Birds and Mammals.” *Proceedings of the National Academy of Sciences of the United States of America* 114 (8): E1460–69.
Kent, Tyler V., Jasmina Uzunović, and Stephen I. Wright. 2017. “Coevolution between Transposable Elements and Recombination.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 372 (1736). https://doi.org/10.1098/rstb.2016.0458.
Kirkpatrick, Mark, and Rafael F. Guerrero. 2014. “Signatures of Sex-Antagonistic Selection on Recombining Sex Chromosomes.” *Genetics* 197 (2): 531–41.
Koerich, Leonardo B., Xiaoyun Wang, Andrew G. Clark, and Antonio Bernardo Carvalho. 2008. “Low Conservation of Gene Content in the Drosophila Y Chromosome.” *Nature* 456 (7224): 949–51.
Lahn, B. T., and D. C. Page. 1999a. “Retroposition of Autosomal mRNA Yielded Testis-Specific Gene Family on Human Y Chromosome.” *Nature Genetics* 21 (4): 429–33.
———. 1999b. “Four Evolutionary Strata on the Human X Chromosome.” *Science* 286 (5441):

964–67.
Leroy, Thibault, Yoann Anselmetti, Marie-Ka Tilak, Sèverine Bérard, Laura Csukonyi, Maëva Gabrielli, Céline Scornavacca, Borja Milá, Christophe Thébaud, and Benoit Nabholz. 2019. “A Bird’s White-Eye View on Neosex Chromosome Evolution.” *bioRxiv*. https://doi.org/10.1101/505610.
Li, Gang, Brian W. Davis, Terje Raudsepp, Alison J. Pearks Wilkerson, Victor C. Mason, Malcolm Ferguson-Smith, Patricia C. O’Brien, Paul D. Waters, and William J. Murphy. 2013. “Comparative Analysis of Mammalian Y Chromosomes Illuminates Ancestral Structure and Lineage-Specific Evolution.” *Genome Research* 23 (9): 1486–95.
Mahajan, Shivani, Kevin H. -C. Wei, Matthew J. Nalley, Lauren Gibilisco, and Doris Bachtrog. 2018. “De Novo Assembly of a Young Drosophila Y Chromosome Using Single-Molecule Sequencing and Chromatin Conformation Capture.” *PLOS Biology* 16 (7): e2006348.
Mank, J. E., and H. Ellegren. 2009. “All Dosage Compensation Is Local: Gene-by-Gene Regulation of Sex-Biased Expression on the Chicken Z Chromosome.” *Heredity* 102 (3): 312–20.
Mank, J. E., K. Nam, and H. Ellegren. 2010. “Faster-Z Evolution Is Predominantly Due to Genetic Drift.” *Molecular Biology and Evolution* 27 (3): 661–70.
Mank, Judith E. 2009. “The W, X, Y and Z of Sex-Chromosome Dosage Compensation.” *Trends in Genetics* 25 (5): 226–33.
Mank, Judith E., and Hans Ellegren. 2007. “Parallel Divergence and Degradation of the Avian W Sex Chromosome.” *Trends in Ecology & Evolution* 22 (8): 389–91.
Mank, Judith E., Beatriz Vicoso, Sofia Berlin, and Brian Charlesworth. 2010. “Effective Population Size and the Faster-X Effect: Empirical Results and Their Interpretation.” *Evolution; International Journal of Organic Evolution* 64 (3): 663–74.
Marand, Alexandre P., Shelley H. Jansky, Hainan Zhao, Courtney P. Leisner, Xiaobiao Zhu, Zixian Zeng, Emily Crisovan, et al. 2017. “Meiotic Crossovers Are Associated with Open Chromatin and Enriched with Stowaway Transposons in Potato.” *Genome Biology* 18 (1): 203.
Meiklejohn, Colin D., Emily L. Landeen, Jodi M. Cook, Sarah B. Kingan, and Daven C. Presgraves. 2011. “Sex Chromosome-Specific Regulation in the Drosophila Male Germline But Little Evidence for Chromosomal Dosage Compensation or Meiotic Inactivation.” *PLoS Biology* 9 (8): e1001126.
Meisel, Richard P., and Tim Connallon. 2013. “The Faster-X Effect: Integrating Theory and Data.” *Trends in Genetics: TIG* 29 (9): 537–44.
Meyer, Barbara J., and Lawrence P. Casson. 1986. “Caenorhabditis Elegans Compensates for the Difference in X Chromosome Dosage between the Sexes by Regulating Transcript Levels.” *Cell* 47 (6): 871–81.
Moghadam, Hooman K., Peter W. Harrison, Gergely Zachar, Tamás Székely, and Judith E. Mank. 2013. “The Plover Neurotranscriptome Assembly: Transcriptomic Analysis in an Ecological Model Species without a Reference Genome.” *Molecular Ecology Resources* 13 (4): 696–705.
Mohandas, T. K., R. M. Speed, M. B. Passage, P. H. Yen, A. C. Chandley, and L. J. Shapiro. 1992. “Role of the Pseudoautosomal Region in Sex-Chromosome Pairing during Male Meiosis: Meiotic Studies in a Man with a Deletion of Distal Xp.” *American Journal of Human Genetics* 51 (3): 526–33.
Mullon, Charles, Alison E. Wright, Max Reuter, Andrew Pomiankowski, and Judith E. Mank. 2015. “Evolution of Dosage Compensation under Sexual Selection Differs between X and Z Chromosomes.” *Nature Communications* 6 (1). https://doi.org/10.1038/ncomms8720.
Nam, Kiwoong, and Hans Ellegren. 2008. “The Chicken (Gallus Gallus) Z Chromosome Contains at Least Three Nonlinear Evolutionary Strata.” *Genetics* 180 (2): 1131–36.
Nanda, I., K. Schlegelmilch, T. Haaf, M. Schartl, and M. Schmid. 2008. “Synteny Conservation of the Z Chromosome in 14 Avian Species (11 Families) Supports a Role for Z Dosage in Avian Sex Determination.” *Cytogenetic and Genome Research* 122 (2): 150–56.

Nanda, I., Z. Shan, M. Schartl, D. W. Burt, M. Koehler, H. Nothwang, F. Grützner, et al. 1999. "300 Million Years of Conserved Synteny between Chicken Z and Human Chromosome 9." *Nature Genetics* 21 (3): 258–59.
Natri, Heini M., Takahito Shikano, and Juha Merilä. 2013. "Progressive Recombination Suppression and Differentiation in Recently Evolved Neo-Sex Chromosomes." *Molecular Biology and Evolution* 30 (5): 1131–44.
Naurin, Sara, Bengt Hansson, Staffan Bensch, and Dennis Hasselquist. 2010. "Why Does Dosage Compensation Differ between XY and ZW Taxa?" *Trends in Genetics: TIG* 26 (1): 15–20.
Nguyen, Di Kim, and Christine M. Disteche. 2006. "Dosage Compensation of the Active X Chromosome in Mammals." *Nature Genetics* 38 (1): 47–53.
Ogawa, A., K. Murata, and S. Mizuno. 1998. "The Location of Z- and W-Linked Marker Genes and Sequence on the Homomorphic Sex Chromosomes of the Ostrich and the Emu." *Proceedings of the National Academy of Sciences* 95 (8): 4415–18.
Orr, H. A., and Y. Kim. 1998. "An Adaptive Hypothesis for the Evolution of the Y Chromosome." *Genetics* 150 (4): 1693–98.
Otto, Sarah P. 2019. "Evolutionary Potential for Genomic Islands of Sexual Divergence on Recombining Sex Chromosomes." *The New Phytologist*, July. https://doi.org/10.1111/nph.16083.
Otto, Sarah P., John R. Pannell, Catherine L. Peichel, Tia-Lynn Ashman, Deborah Charlesworth, Adam K. Chippindale, Lynda F. Delph, Rafael F. Guerrero, Samuel V. Scarpino, and Bryant F. McAllister. 2011. "About PAR: The Distinct Evolutionary Dynamics of the Pseudoautosomal Region." *Trends in Genetics* 27 (9): 358–67.
Otto, S. P. 2014. "Selective Maintenance of Recombination between the Sex Chromosomes." *Journal of Evolutionary Biology* 27 (7): 1431–42.
Peona, Valentina, Matthias H. Weissensteiner, and Alexander Suh. 2018. "How Complete Are 'complete' Genome Assemblies?-An Avian Perspective." *Molecular Ecology Resources* 18 (6): 1188–95.
Pessia, Eugénie, Jan Engelstädter, and Gabriel A. B. Marais. 2014. "The Evolution of X Chromosome Inactivation in Mammals: The Demise of Ohno's Hypothesis?" *Cellular and Molecular Life Sciences* 71 (8): 1383–94.
Pigozzi, María Inés. 2011. "Diverse Stages of Sex-Chromosome Differentiation in Tinamid Birds: Evidence from Crossover Analysis in Eudromia Elegans and Crypturellus Tataupa." *Genetica* 139 (6): 771–77.
Pigozzi, M. I., and A. J. Solari. 1999. "The ZW Pairs of Two Paleognath Birds from Two Orders Show Transitional Stages of Sex Chromosome Differentiation." *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology* 7 (7): 541–51.
Ponnikas, Suvi, Hanna Sigeman, Jessica K. Abbott, and Bengt Hansson. 2018. "Why Do Sex Chromosomes Stop Recombining?" *Trends in Genetics: TIG* 34 (7): 492–503.
Qiu, S., R. Bergero, S. Guirao-Rico, J. L. Campos, T. Cezard, K. Gharbi, and D. Charlesworth. 2016. "RAD Mapping Reveals an Evolving, Polymorphic and Fuzzy Boundary of a Plant Pseudoautosomal Region." *Molecular Ecology* 25 (1): 414–30.
Rice, William R. 1987. "THE ACCUMULATION OF SEXUALLY ANTAGONISTIC GENES AS A SELECTIVE AGENT PROMOTING THE EVOLUTION OF REDUCED RECOMBINATION BETWEEN PRIMITIVE SEX CHROMOSOMES." *Evolution; International Journal of Organic Evolution* 41 (4): 911–14.
Rice, W. R. 1987. "Genetic Hitchhiking and the Evolution of Reduced Genetic Activity of the Y Sex Chromosome." *Genetics* 116 (1): 161–67.
Rodrigues, Nicolas, Tania Studer, Christophe Dufresnes, and Nicolas Perrin. 2018. "Sex-Chromosome Recombination in Common Frogs Brings Water to the Fountain-of-Youth." *Molecular Biology and Evolution* 35 (4): 942–48.
Ross, Mark T., Darren V. Grafham, Alison J. Coffey, Steven Scherer, Kirsten McLay, Donna

Muzny, Matthias Platzer, et al. 2005. “The DNA Sequence of the Human X Chromosome.” *Nature* 434 (7031): 325–37.
Rousselle, Marjolaine, Nicolas Faivre, Marion Ballenghien, Nicolas Galtier, and Benoit Nabholz. 2016. “Hemizygosity Enhances Purifying Selection: Lack of Fast-Z Evolution in Two Satyrine Butterflies.” *Genome Biology and Evolution* 8 (10): 3108–19.
Rutkowska, Joanna, Malgorzata Lagisz, and Shinichi Nakagawa. 2012. “The Long and the Short of Avian W Chromosomes: No Evidence for Gradual W Shortening.” *Biology Letters* 8 (4): 636–38.
Sackton, Timothy B., Russell B. Corbett-Detig, Javaregowda Nagaraju, Lakshmi Vaishna, Kallare P. Arunkumar, and Daniel L. Hartl. 2014. “Positive Selection Drives Faster-Z Evolution in Silkmoths.” *Evolution*. https://doi.org/10.1111/evo.12449.
Saifi, G. M., and H. S. Chandra. 1999. “An Apparent Excess of Sex- and Reproduction-Related Genes on the Human X Chromosome.” *Proceedings. Biological Sciences / The Royal Society* 266 (1415): 203–9.
Shetty, S., D. K. Griffin, and J. A. Graves. 1999. “Comparative Painting Reveals Strong Chromosome Homology over 80 Million Years of Bird Evolution.” *Chromosome Research: An International Journal on the Molecular, Supramolecular and Evolutionary Aspects of Chromosome Biology* 7 (4): 289–95.
Singhal, Sonal, Ellen Leffler, Keerthi Sannareddy, Isaac Turner, Oliver Venn, Daniel Hooper, Alva Strand, et al. 2015. “Stable Recombination Hotspots in Birds.” *Science*. https://doi.org/10.1101/023101.
Skaletsky, Helen, Tomoko Kuroda-Kawaguchi, Patrick J. Minx, Holland S. Cordum, Ladeana Hillier, Laura G. Brown, Sjoerd Repping, et al. 2003. “The Male-Specific Region of the Human Y Chromosome Is a Mosaic of Discrete Sequence Classes.” *Nature* 423 (6942): 825–37.
Śliwińska, Ewa B., Rafał Martyka, and Piotr Tryjanowski. 2016. “Evolutionary Interaction between W/Y Chromosome and Transposable Elements.” *Genetica* 144 (3): 267–78.
Smeds, Linnéa, Takeshi Kawakami, Reto Burri, Paulina Bolivar, Arild Husby, Anna Qvarnström, Severin Uebbing, and Hans Ellegren. 2014. “Genomic Identification and Characterization of the Pseudoautosomal Region in Highly Differentiated Avian Sex Chromosomes.” *Nature Communications* 5 (November): 5448.
Smeds, Linnéa, Vera Warmuth, Paulina Bolivar, Severin Uebbing, Reto Burri, Alexander Suh, Alexander Nater, et al. 2015. “Evolutionary Analysis of the Female-Specific Avian W Chromosome.” *Nature Communications* 6 (1). https://doi.org/10.1038/ncomms8330.
Smith, Craig A., Kelly N. Roeszler, Thomas Ohnesorg, David M. Cummins, Peter G. Farlie, Timothy J. Doran, and Andrew H. Sinclair. 2009. “The Avian Z-Linked Gene DMRT1 Is Required for Male Sex Determination in the Chicken.” *Nature* 461 (7261): 267–71.
Smith, Craig A., Kelly N. Roeszler, and Andrew H. Sinclair. 2009. “Genetic Evidence against a Role for W-Linked Histidine Triad Nucleotide Binding Protein (HINTW) in Avian Sex Determination.” *The International Journal of Developmental Biology* 53 (1): 59–67.
Smith, Gilbert, Yun-Ru Chen, Gary W. Blissard, and Adriana D. Briscoe. 2014. “Complete Dosage Compensation and Sex-Biased Gene Expression in the Moth Manduca Sexta.” *Genome Biology and Evolution* 6 (3): 526–37.
Soh, Y. Q. Shirleen, Jessica Alföldi, Tatyana Pyntikova, Laura G. Brown, Tina Graves, Patrick J. Minx, Robert S. Fulton, et al. 2014. “Sequencing the Mouse Y Chromosome Reveals Convergent Gene Acquisition and Amplification on Both Sex Chromosomes.” *Cell* 159 (4): 800–813.
Stiglec, R., T. Ezaz, and J. A. M. Graves. 2007. “A New Look at the Evolution of Avian Sex Chromosomes.” *Cytogenetic and Genome Research* 117 (1-4): 103–9.
Sun, Yu, Jesper Svedberg, Markus Hiltunen, Pádraic Corcoran, and Hanna Johannesson. 2017. “Large-Scale Suppression of Recombination Predates Genomic Rearrangements in Neurospora Tetrasperma.” *Nature Communications* 8 (1): 1140.
Tobler, Ray, Viola Nolte, and Christian Schlötterer. 2017. “High Rate of Translocation-Based

Gene Birth on the Drosophila Y Chromosome." *Proceedings of the National Academy of Sciences* 114 (44): 11721–26.
Tsuda, Yayoi, Chizuko Nishida-Umehara, Junko Ishijima, Kazuhiko Yamada, and Yoichi Matsuda. 2007. "Comparison of the Z and W Sex Chromosomal Architectures in Elegant Crested Tinamou (Eudromia Elegans) and Ostrich (Struthio Camelus) and the Process of Sex Chromosome Differentiation in Palaeognathous Birds." *Chromosoma* 116 (2): 159–73.
Uebbing, Severin, Anne Konzer, Luohao Xu, Niclas Backström, Björn Brunström, Jonas Bergquist, and Hans Ellegren. 2015. "Quantitative Mass Spectrometry Reveals Partial Translational Regulation for Dosage Compensation in Chicken." *Molecular Biology and Evolution* 32 (10): 2716–25.
Uebbing, Severin, Axel Künstner, Hannu Mäkinen, and Hans Ellegren. 2013. "Transcriptome Sequencing Reveals the Character of Incomplete Dosage Compensation across Multiple Tissues in Flycatchers." *Genome Biology and Evolution* 5 (8): 1555–66.
Underwood, Charles J., and Kyuha Choi. 2019. "Heterogeneous Transposable Elements as Silencers, Enhancers and Targets of Meiotic Recombination." *Chromosoma*, July. https://doi.org/10.1007/s00412-019-00718-4.
Vicoso, Beatriz, and Doris Bachtrog. 2015. "Numerous Transitions of Sex Chromosomes in Diptera." *PLoS Biology* 13 (4): e1002078.
Vicoso, Beatriz, and Brian Charlesworth. 2006. "Evolution on the X Chromosome: Unusual Patterns and Processes." *Nature Reviews Genetics* 7 (8): 645–53.
Vicoso, B., V. B. Kaiser, and D. Bachtrog. 2013. "Sex-Biased Gene Expression at Homomorphic Sex Chromosomes in Emus and Its Implication for Sex Chromosome Evolution." *Proceedings of the National Academy of Sciences* 110 (16): 6453–58.
Walters, James R., and Thomas J. Hardcastle. 2011. "Getting a Full Dose? Reconsidering Sex Chromosome Dosage Compensation in the Silkworm, Bombyx Mori." *Genome Biology and Evolution* 3 (April): 491–504.
Wang, Zongji, Jilin Zhang, Wei Yang, Na An, Pei Zhang, Guojie Zhang, and Qi Zhou. 2014. "Temporal Genomic Evolution of Bird Sex Chromosomes." *BMC Evolutionary Biology* 14 (December): 250.
Weissensteiner, Matthias H., and Alexander Suh. 2019. "Repetitive DNA: The Dark Matter of Avian Genomics." In *Avian Genomics in Ecology and Evolution*, edited by Robert H. S. Kraus, 93–150.
White, Michael A., Jun Kitano, and Catherine L. Peichel. 2015. "Purifying Selection Maintains Dosage-Sensitive Genes during Degeneration of the Threespine Stickleback Y Chromosome." *Molecular Biology and Evolution* 32 (8): 1981–95.
Wolf, Jochen Bw, and Jarosław Bryk. 2011. "General Lack of Global Dosage Compensation in ZZ/ZW Systems? Broadening the Perspective with RNA-Seq." *BMC Genomics* 12 (February): 91.
Wright, Alison E., Iulia Darolti, Natasha I. Bloch, Vicencio Oostra, Ben Sandkam, Severine D. Buechel, Niclas Kolm, Felix Breden, Beatriz Vicoso, and Judith E. Mank. 2017. "Convergent Recombination Suppression Suggests Role of Sexual Selection in Guppy Sex Chromosome Formation." *Nature Communications* 8 (January): 14251.
Wright, Alison E., Rebecca Dean, Fabian Zimmer, and Judith E. Mank. 2016. "How to Make a Sex Chromosome." *Nature Communications* 7 (1). https://doi.org/10.1038/ncomms12087.
Wright, Alison E., Peter W. Harrison, Stephen H. Montgomery, Marie A. Pointer, and Judith E. Mank. 2014. "Independent Stratum Formation on the Avian Sex Chromosomes Reveals Inter-Chromosomal Gene Conversion and Predominance of Purifying Selection on the W Chromosome." *Evolution; International Journal of Organic Evolution* 68 (11): 3281–95.
Wright, Alison E., Peter W. Harrison, Fabian Zimmer, Stephen H. Montgomery, Marie A. Pointer, and Judith E. Mank. 2015. "Variation in Promiscuity and Sexual Selection Drives Avian Rate of Faster-Z Evolution." *Molecular Ecology* 24 (6): 1218–35.
Wright, Alison E., Hooman K. Moghadam, and Judith E. Mank. 2012. "Trade-off Between Selection for Dosage Compensation and Masculinization on the Avian Z Chromosome."

Genetics 192 (4): 1433–45.
Yazdi, Homa Papoli. 2019. “The Evolution of Sex Chromosomes and Sex-Linked Sequences in Birds.” PhD, Uppsala University. https://books.google.com/books/about/The_Evolution_of_Sex_Chromosomes_and_Sex.html?hl=&id=Kv_4xAEACAAJ.
Yazdi, Homa Papoli, and Hans Ellegren. 2018. “A Genetic Map of Ostrich Z Chromosome and the Role of Inversions in Avian Sex Chromosome Evolution.” *Genome Biology and Evolution* 10 (8): 2049–60.
Zhang, Guojie, Cai Li, Qiye Li, Bo Li, Denis M. Larkin, Chul Lee, Jay F. Storz, et al. 2014. “Comparative Genomics Reveals Insights into Avian Genome Evolution and Adaptation.” *Science* 346 (6215): 1311–20.
Zhou, Qi, and Doris Bachtrog. 2012. “Sex-Specific Adaptation Drives Early Sex Chromosome Evolution in Drosophila.” *Science* 337 (6092): 341–45.
Zhou, Qi, Jilin Zhang, Doris Bachtrog, Na An, Quanfei Huang, Erich D. Jarvis, M. Thomas P. Gilbert, and Guojie Zhang. 2014. “Complex Evolutionary Trajectories of Sex Chromosomes across Bird Taxa.” *Science* 346 (6215): 1246338.

Chapter 2

Paper I: Dynamic evolutionary history and gene content of sex chromosomes across diverse songbirds

Luohao Xu[1,2], Gabriel Auer[2], Valentina Peona[3], Alexander Suh[3], Yuan Deng[4,5], Shaohong Feng[4,5], Guojie Zhang[4,6,7], Mozes P.K. Blom[8,9], Les Christidis[10,11], Stefan Prost[12,13], Martin Irestedt[8,*], Qi Zhou[1,2,*]

[1] MOE Laboratory of Biosystems Homeostasis & Protection, Life Sciences Institute, Zhejiang University, Hangzhou, China
[2] Department of Molecular Evolution and Development, University of Vienna, Vienna, Austria
[3] Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden
[4] China National Genebank, BGI-Shenzhen, Shenzhen, China
[5] BGI-Shenzhen, Shenzhen, China
[6] State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China
[7] Section for Ecology and Evolution, Department of Biology, University of Copenhagen, Copenhagen, Denmark
[8] Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden
[9] Museum für Naturkunde, Leibniz-Institut für Evolutions- und Biodiversitätsforschung, Berlin 10115, Germany
[10] National Marin Science Centre, Southern Cross University, Coffs Harbour, NSW, Australia
[11] School of BioSciences, University of Melbourne, Parkville, VIC, Australia
[12] Department of Integrative Biology, University of California, Berkeley, Berkeley, United States
[13] Department of Biology, Stanford University, Stanford, United States

Highlights

- The evolutionary history of the sex chromosome is shared by all songbirds
- Accumulation of transposable elements contributes to recombination suppression
- The gene content of the W chromosome is conserved across species

- Dosage sensitive genes are retained on the W chromosome by selection

Summary

Songbirds have a species number close to that of mammals and are classic models for studying speciation and sexual selection. Sex chromosomes are hotspots of both processes, yet their evolutionary history in songbirds remains unclear. We characterized genomes of 11 songbird species, with 5 genomes of bird-of-paradise species. We conclude that songbird sex chromosomes have undergone four periods of recombination suppression before species radiation, producing a gradient of pairwise sequence divergence termed 'evolutionary strata'. The latest stratum was probably due to a songbird-specific burst of retrotransposon CR1–E1 elements at its boundary, instead of the chromosome inversion generally assumed for suppressing sex-linked recombination. The formation of evolutionary strata has reshaped the genomic architecture of both sex chromosomes. We find stepwise variations of Z-linked inversions, repeat and guanine–cytosine (GC) contents, as well as the W-linked gene loss rate associated with the age of strata. A few W-linked genes have been preserved for their essential functions, indicated by higher and broader expression of lizard orthologues compared with those of other sex-linked genes. We also find a different degree of accelerated evolution of Z-linked genes versus autosomal genes among species, potentially reflecting the diversified intensity of sexual selection. Our results uncover the dynamic evolutionary history of songbird sex chromosomes and provide insights into the mechanisms of recombination suppression.

Keywords

Songbirds; Sex chromosomes; W degeneration; Recombination suppression; Transposable element

ARTICLES
https://doi.org/10.1038/s41559-019-0850-1

nature ecology & evolution

Dynamic evolutionary history and gene content of sex chromosomes across diverse songbirds

Luohao Xu[1,2], Gabriel Auer[2], Valentina Peona[3], Alexander Suh[3], Yuan Deng[4,5], Shaohong Feng[4,5], Guojie Zhang[4,6,7], Mozes P. K. Blom[8,9], Les Christidis[10,11], Stefan Prost[12,13], Martin Irestedt[8*] and Qi Zhou[1,2*]

Songbirds have a species number close to that of mammals and are classic models for studying speciation and sexual selection. Sex chromosomes are hotspots of both processes, yet their evolutionary history in songbirds remains unclear. We characterized genomes of 11 songbird species, with 5 genomes of bird-of-paradise species. We conclude that songbird sex chromosomes have undergone four periods of recombination suppression before species radiation, producing a gradient of pairwise sequence divergence termed 'evolutionary strata'. The latest stratum was probably due to a songbird-specific burst of retrotransposon CR1-E1 elements at its boundary, instead of the chromosome inversion generally assumed for suppressing sex-linked recombination. The formation of evolutionary strata has reshaped the genomic architecture of both sex chromosomes. We find stepwise variations of Z-linked inversions, repeat and guanine–cytosine (GC) contents, as well as W-linked gene loss rate associated with the age of strata. A few W-linked genes have been preserved for their essential functions, indicated by higher and broader expression of lizard orthologues compared with those of other sex-linked genes. We also find a different degree of accelerated evolution of Z-linked genes versus autosomal genes among species, potentially reflecting diversified intensity of sexual selection. Our results uncover the dynamic evolutionary history of songbird sex chromosomes and provide insights into the mechanisms of recombination suppression.

Songbirds (Oscines, suborder Passeri) have over 5,000 species and comprise most passerines and nearly half of all extant bird species[1]. This is because of the largest avian species radiation that occurred about 60 million years (Myr) ago[2]. With the development of genomics, many species besides zebra finch are now becoming important models for studying molecular patterns and mechanisms of speciation[3,4], supergene[5] or cognition[6], out of their long history of ecological or behavioural studies. One major reason for biologists' interest in songbirds is their diversified sexual traits. For example, their ostentatious plumage forms and colours, sophisticated songs and mating rituals, all of which can undergo rapid turnovers even between sister species. Theories predict that sex chromosomes play a disproportionately large role in speciation (the 'large X/Z' effect), sexual selection and evolution of sexually dimorphic traits[7–9]. However, the evolutionary history of songbirds' sex chromosome remains unclear because there were few genomic studies characterizing songbirds' sex chromosomes except for collared flycatcher[10]. Unlike the mammalian XY system, birds have independently evolved a pair of female heterogametic sex chromosomes that are usually heteromorphic in females (ZW) and homomorphic in males (ZZ). A recent cytological investigation of over 400 passerine species found a higher fixation rate of chromosomal inversions on the Z chromosome than autosomes within species, so that gene flow is probably more reduced by hybridization[11,12]. A significantly lower level of introgression in Z-linked genes compared to autosomal genes has been reported from studying pairs of recently diverged songbird species[13–15]. Such a large-Z pattern is probably contributed by several factors that act in an opposite manner in the XY sex system. First, Z chromosomes are more often transmitted in males, thus are expected to have a higher mutation rate than the rest of the genome, due to the 'male-driven evolution' effect[16]. Previous studies[17–19] showed this effect is less pronounced in birds than in mammals, thus the contribution of 'male-driven evolution' to the large-Z pattern may be limited. Second, as sexual selection more frequently targets males, the variation in male reproductive success will further reduce the effective population size of Z chromosomes from three-quarters that of autosomes[20]. The consequential genetic drift effect is expected to fix excessive slightly deleterious mutations on the Z chromosome and lead to its faster evolutionary rate than autosomes (the 'fast-Z' effect)[21]. This has been demonstrated in Galloanserae species (for example, chicken and duck), of which those undergoing stronger sperm competition, which is a more intensive male sexual selection, exhibit a larger difference between Z chromosome and autosomes in their evolution rates[22].

In contrast to the avian Z chromosomes, or more broadly the mammalian XY chromosomes, genomic studies of avian W chromosomes, especially those of songbirds, have only recently been performed[10,23,24]. This is because most genomic projects prefer to

[1]MOE Laboratory of Biosystems Homeostasis & Protection, Life Sciences Institute, Zhejiang University, Hangzhou, China. [2]Department of Molecular Evolution and Development, University of Vienna, Vienna, Austria. [3]Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Uppsala, Sweden. [4]China National Genebank, BGI-Shenzhen, Shenzhen, China. [5]BGI-Shenzhen, Shenzhen, China. [6]State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China. [7]Section for Ecology and Evolution, Department of Biology, University of Copenhagen, Copenhagen, Denmark. [8]Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden. [9]Museum für Naturkunde, Leibniz-Institut für Evolutions- und Biodiversitätsforschung, Berlin, Germany. [10]National Marin Science Centre, Southern Cross University, Coffs Harbour, New South Wales, Australia. [11]School of BioSciences, University of Melbourne, Parkville, Victoria, Australia. [12]Department of Integrative Biology, University of California, Berkeley, Berkeley, CA, USA. [13]LOEWE-Center for Translational Biodiversity Genomics, Senckenberg, Frankfurt, Germany. *e-mail: Martin.Irestedt@nrm.se; zhouqi1982@zju.edu.cn

choose the homogametic sex (for example, male birds or female mammals) for sequencing, to avoid the presumably gene-poor and highly repetitive Y or W chromosomes. It has been suggested but not yet experimentally shown that the Y/W chromosomes have undergone suppressions of recombination to prevent the sex-determining gene or sexual antagonistic (beneficial to one sex but detrimental to the other) genes being transmitted in the opposite sex[25]. The loss of recombination reduces the efficacy of natural selection and drives the ultimate genetic decay of non-recombining regions of Y/W chromosomes due to the effect of for example, 'Hill–Robertson interference' between linked loci[26]. The degeneration process can be accelerated by selective sweep targeting male-related genes on the Y chromosome, or by background selection, purging the deleterious mutations from highly dosage-sensitive genes[27]. Simulation showed that they play a different role at different stages of Y/W degeneration[28]. Both processes have gained evidence from the analyses of mammalian[29,30] and *Drosophila*[31,32] Y-linked genes. Although purifying selection acting on dosage-sensitive genes has been implicated to maintain the few W-linked genes retained in Galloanserae (for example, chicken and duck)[24,33] or flycatcher[10], little evidence has been found for female-specific positive selection acting on W-linked genes (but see ref. [34]).

In both birds[23] and mammals[35], as well as several plant species such as *Silene latifolia*[36], recombination suppressions have all proceeded in a stepwise manner presumably through chromosomal inversions, leaving a stratified pattern of sequence divergence between sex chromosomes termed 'evolutionary strata'. Eutherian mammalian X and Y chromosomes have been inferred to share at least three strata, with another two more recent ones shared only among catarrhines (old world monkeys and great apes)[30]. We recently discovered from a broad but sparse sampling of diverse bird genomes that the history and tempo of avian sex chromosome evolution are much more complicated than those of mammals[23]. We showed that all birds' sex chromosomes only share the first time of recombination suppression (stratum 0, Aves S0) encompassing the avian male-determining gene *DMRT1*. This was followed by the independent formation of S1 in different basal Palaeognathae species (for example, ratites and tinamou) and in the ancestor of Neognathae (for example, chicken and zebra finch). Ratites have halted any further recombination loss and maintained over two-thirds of the entire sex chromosome pair as the exceptionally long recombining pseudoautosomal regions (PAR). Therefore, their sex chromosomes are homomorphic and gene-rich on the W chromosome. All Neognathae species have suppressed recombination throughout most regions of the sex chromosomes with short and varying sizes of PAR (ref. [37]). However, overall, avian W chromosomes seem to have retained more genes and decayed at a slower rate than the mammalian Y chromosomes. Moreover, sexually monomorphic species (for example, most ratites) seem to differentiate more slowly than sexually dimorphic species (chicken and many Neoaves species) in their sex chromosomes, consistent with the hypothesis that sexual antagonistic alleles have triggered the expansion of recombination suppression between sex chromosomes[38]. However, ratites have a deep divergence from other birds and also a much lower mutation rate as expected from their larger body size. These confounding factors make the actual influence of sexual selection on the rate of sex chromosome evolution unclear. The principal group of Neognathae, Neoaves share one stratum S2, with the more recent history of sex chromosomes of songbirds being unclear. So far, only one songbird (collared flycatcher), has been characterized for its W-linked genes[10], in the range 46–90 reported W-linked genes of other Neoaves species. To explain the evolutionary history of songbirds' sex chromosomes, we produced high-quality female genomes of five bird-of-paradise (BOP) species. Together with six other published female genomes of songbird species, our analyses covered main songbird lineages (Corvida and Passerida) that diverged in the last 50 Myr (refs. [2,39]).

Results

Characterization of songbird sex chromosome sequences. We produced 36- to 150-fold genomic coverage of sequencing data for each BOP species and performed de novo genome assembly followed by chromosome mapping using the genomes of highly continuous or closely related great tit or hooded crow as reference[6]. The high continuity and completeness of the draft genomes are revealed by their scaffold N50 lengths (all longer than 3 Mb, except for Raggiana BOP) and BUSCO scores (92.9–94.0%; Supplementary Table 1). To reconstruct the evolutionary history of sampled songbirds' sex chromosomes, we first identified sequences from putative PARs by their homology to the published PAR sequence of collared flycatcher[40] and confirmed them by their similar read depth level to that of autosomes. Sequences from sexually differentiated regions (SDR) were identified as those that show half the female sequencing depth of autosomes (Fig. 1a and Supplementary Fig. 1). We then separated the Z- and W-linked sequences with the expectation that the latter would diverge much faster than the former from the reference Z chromosome sequence (Methods), and we further confirmed the W-linkage with a clear female-specific pattern in all but one species with sequencing data of both sexes (Fig. 1a and Supplementary Fig. 2). Our method cannot identify recent fusion/translocation of autosomal fragments to the sex chromosome pair (forming so-called 'neo-sex' chromosome), as in some warblers[41]. All the studied songbirds have a short putative PAR ranging from 564 to 781 kilobases (kb). The assembled lengths of the largely euchromatic parts of W chromosomes range from 1.33 to 7.24 megabases (Mb), corresponding to only 1.9–8.5% of the Z chromosome length across species (Fig. 1b and Supplementary Table 2), probably as a result of large deletions and massive invasions of repetitive elements. Indeed, the repeat content of the assembled W chromosomes is 2.5- to 4.9-fold higher than that of Z chromosomes on the chromosome-wide average (Supplementary Fig. 3 and Table 2).

Age-dependent genomic impact of evolutionary strata. If recombination was suppressed between sex chromosomes in a stepwise manner, we expect a gradient of Z/W sequence divergence levels along the Z chromosome, such as has been reported along the human X chromosome[42]. However, we have previously showed that the extant synteny of Neognathae Z chromosomes is misleading for inferring evolutionary strata, due to the marked intrachromosomal rearrangements[23]. By contrast, Palaeognathae species (for example, emu and ostrich) have maintained a highly conserved sequence synteny even with reptile species, with over two-thirds of their sex-linked regions still recombining as an approximate of the proto-sex chromosomes of all bird species[23,43]. We are able to identify the reshuffled fragments of the first and the second strata (S0 and S1) shared by all Neognathae species in the studied songbird genomes by their homology to the emu genome. They were mapped as two continuous regions on the emu Z chromosome (Fig. 2a ; Supplementary Figs. 4 and 5). Two recently formed strata (Neoaves S2 and S3) are much more conserved for their synteny across avian species and each shows a significantly different level of Z/W sequence divergence (Fig. 2b and Supplementary Fig. 6), GC3 (GC content at the third codon positions; Supplementary Fig. 7) and Z-linked long terminal repeat (LTR) content (Fig. 2c and Supplementary Fig. 7) from each other. The marked change of Z/W divergence level allows us to precisely map the boundaries between those two strata. In general, series of recombination suppressions have reshaped the genomic architecture of the Z chromosome in chronological order. Regions of younger strata exhibit much less Z-linked intrachromosomal rearrangements between species, suggesting the reduced selective constraints on gene synteny after recombination was suppressed in the older strata[44]. Alternatively, it could also reflect a neutral process that older strata have fixed more genomic rearrangements, as genetic drift has been acting for longer

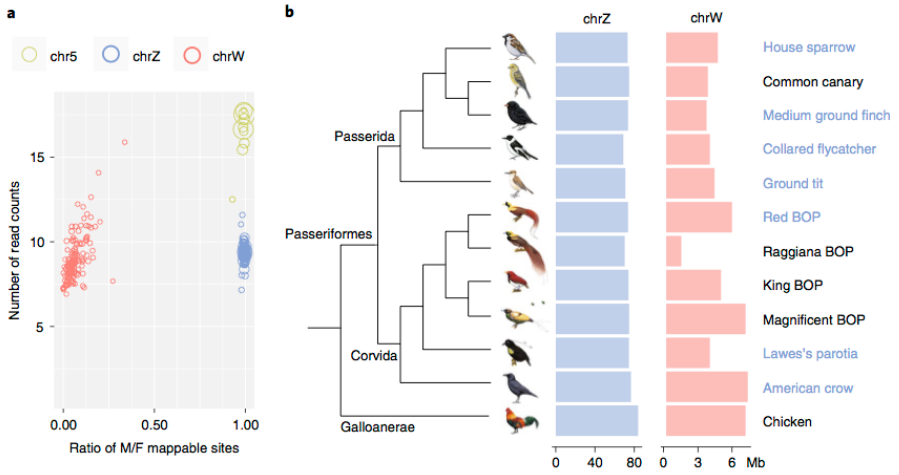


Fig. 1 | The Z and W chromosomes of different songbirds. a, We use medium ground finch as an example to demonstrate our identification and verification of sex-linked sequences. For each scaffold shown as a circle with scaled size to its length, the ratio of nucleotide sites that were mapped by male (M) versus female (F) genomic reads is plotted against the sequencing depth of this scaffold. Scaffold sequences are clustered separately by their derived chromosomes, with W-linked (red circles) and Z-linked (blue circles) sequences showing the expected half the autosome (green) sequencing depth, and W-linked sequences showing almost no mappable sites from male reads. **b**, The lengths of Z and W chromosomes across the studied songbird species. The data of chicken and collared flycatcher are derived from refs. [10,24]. The shorter length of Raggiana BOP W chromosome is probably caused by the low sequencing coverage. Species with Illumina reads of both sexes available are marked in blue. All bird illustrations were ordered from https://www.hbw.com/; ref. [91].

in these regions due to the reduced effective population size. In particular, GC3 content decreases, while the repeat content increases by age of stratum. This is probably because weaker effects of GC-biased gene conversion (gBGC; ref. [45]) and purifying selection against transposable element (TE) insertions[46] have been acting for longer in Z-linked regions of older strata with reduced recombination. Consistently, a similar pattern has also been found contrasting PAR versus the rest Z-linked regions in collared flycatcher[40].

Lineage-specific burst of retrotransposon probably has induced recombination suppression between sex chromosomes. The distribution of long interspersed elements (LINEs), mainly the retrotransposon chicken repeat 1 (CR1) elements, shows an exceptional pattern compared to that of LTR elements (Fig. 2c and Supplementary Fig. 7). The abundance of CR1 is unexpectedly similar in S3 and S0, and much higher than that of the rest of the Z-linked regions. A close examination shows that this is due to the specific accumulation of CR1 spanning the boundary between PAR and S3. Such a burst of CR1, particularly the CR1–E1 subfamily[47], is shared by all the investigated songbirds but absent in the basal passerine rifleman and other Neoaves species. It extends with gradual reduction into about one-third of the entire S3 region (Fig. 2d and Supplementary Fig. 8). The peak region of CR1 accumulation is associated with a large deletion (about 1.5 Mb) in passerines that removes a gene *DCC* (Deleted in Colorectal Carcinoma) highly conserved across other vertebrates[48]. This gene is responsible for axon guidance for brain midline crossing and has been independently lost in some but not all passerines and Galliformes[49].

In addition, the burst of CR1–E1 element seems to coincide with S3 emergence. Almost all the investigated genomes of songbirds have about two-fold more CR1–E1 elements than that of rifleman (Supplementary Table 3). Our phylogenetic reconstruction of Z- and W-linked gametologue sequences shows that only songbird-derived sequences are always grouped by chromosome instead of by species (Fig. 2e and Supplementary Figs. 9–12). This indicates that all songbirds share four evolutionary strata, with the latest S3 formed at the same time with the genome-wide expansion of CR1–E1 elements, after the divergence between all the songbirds and other passerine species. The highly conserved Z-linked synteny of S3 between songbird species and between songbirds and chicken (Fig. 2a and Supplementary Fig. 5) suggests that there was no Z-linked chromosomal inversion at S3. It is likely that the recent burst of CR1–E1 subfamily elements have led to the formation of S3, although we cannot exclude the contribution of W-linked chromosomal inversions. Interestingly, other CR1 subfamilies CR1–E(4–6) have an independent burst both genome-wide and specifically at the PAR/S3 boundary in rifleman (Fig. 2d, Supplementary Fig. 8 and Table 3). Given this boundary region has been shown to have frequent but different degrees of multiple gene loss in different lineages of birds[48,49], it is probably a hotspot for mutations or LINE insertions that have recurrently contributed to the independent formation of S3 in many bird species.

Fast-Z pattern of songbirds suggest their dynamic evolution of sexual selection. The formation of evolutionary strata has subjected the Z chromosome to male-biased transmission and a reduced effective population size, which are expected to produce faster mutation and evolution rates of Z-linked genes, respectively[20]. We found a larger branch-specific synonymous substitution rate (dS) of Z-linked genes (statistically not significant) but a significantly smaller dS of W-linked genes, compared to that of autosomal genes ($P=0.002165$, Wilcoxon rank sum test; Supplementary Fig. 13), as a result of male-driven evolution[16]. The branch-specific evolution rates (ω) measured by the ratios of non-synonymous substitution rates (dN) over dS have significantly ($P<0.003$, Wilcoxon rank sum test; Supplementary Fig. 14) increased for both Z- and W-linked gametologues relative to autosomal genes, indicating a 'fast-Z' effect and degeneration of W-linked genes (see below). Previous simulation

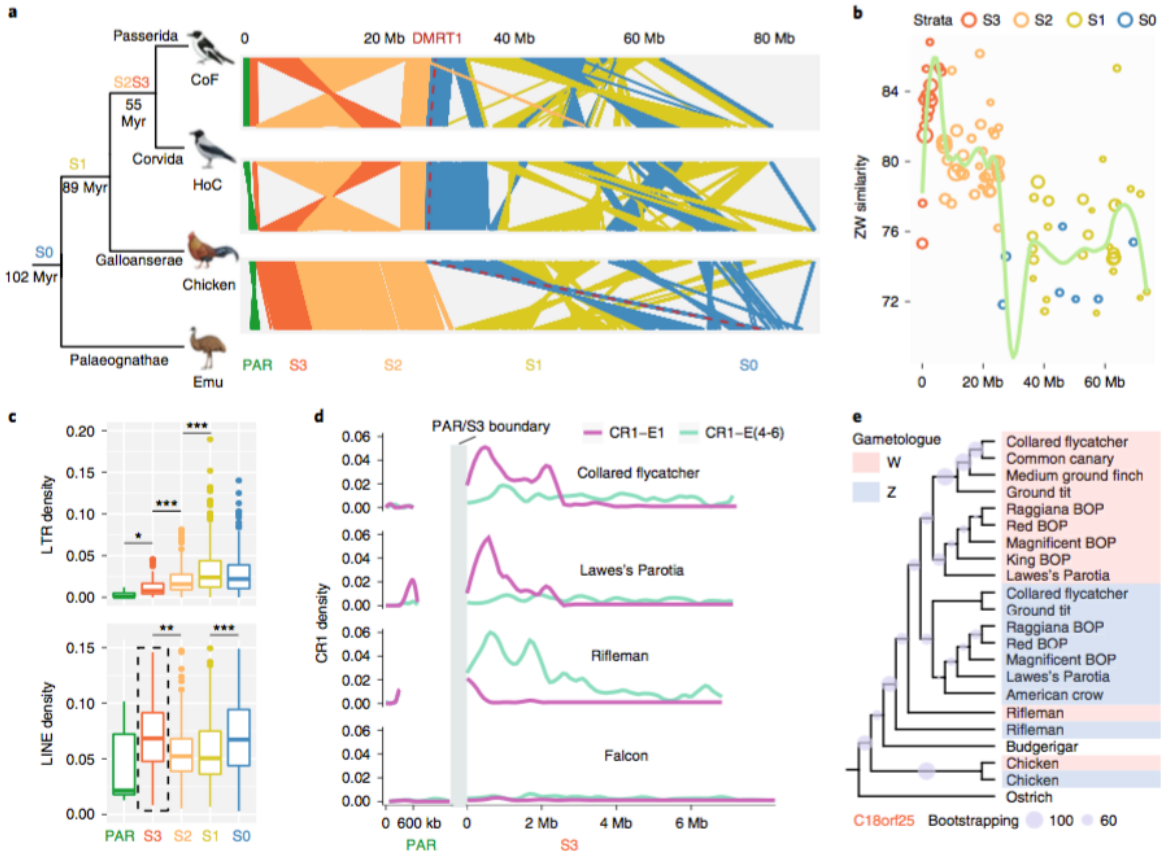


Fig. 2 | Evolutionary strata of songbirds. a, We use chromosomal or nearly chromosomal genome assemblies of four representative bird species (CoF, collared flycatcher; HoC, hooded crow; chicken; and emu) to show their rearrangements on the Z chromosome. Each line represents one pair of aligned fragments between two species, and each colour corresponds to one evolutionary stratum of songbirds. The location of *DMRT1*, the avian male-determining gene, is marked by the red dashed line. On the phylogenetic tree, we also indicate the evolutionary strata at their respective node of origination. Generally, the synteny is more conserved in younger strata between species. **b**, We use Lawes's Parotia as an example to demonstrate the pairwise sequence similarity pattern of evolutionary strata. The size of circles is scaled to the length of sequence alignments between Z/W chromosomes. **c**, Transposable elements (LINEs and LTRs) have accumulated strongly in older strata (S0 is the first stratum), except for LINE at S3. In the boxplots, the horizontal line shows the medium value, the whiskers show the 25 and 75% quartile values of the density of TEs (percentage of TE sequences in every 100 kb window). ***$P<0.001$, **$P<0.01$, *$P<0.05$. **d**, Lineage-specific burst of CR1-E1 (a subtype of CR1 LINE element, purple line) at the boundary of the PAR and S3 in songbirds, since their divergence with other passerine species. Other subtypes of CR1 elements are also plotted with the green line for comparison. **e**, Phylogenetic tree using Z- and W-linked gametologue sequences of the gene *C18orf25* located at S3. Lineages are clustered by chromosomes (red or blue), not by species, suggesting S3 independently formed in rifleman, chicken and the ancestor of songbirds. All bird illustrations were ordered from https://www.hbw.com/; ref. [91].

work and experimental evidence in Galloanserae have suggested that different degrees of sexual selection targeting males will influence the male-mating success, hence the genetic drift effect on the Z chromosome to a different degree[20,22]. Songbirds, especially BOPs have been frequently used as a textbook example of sexual selection[50,51]; however, their evolution history of sexual selection remained unclear. To reconstruct that, we approximated the intensity of sexual selection targeting males by measuring the degree of fast-Z effect (Z/A value, the ratio of branch-specific ω values of Z-linked genes versus autosomal genes) in a phylogenetic context (Fig. 3 and Supplementary Table 4). The varying Z/A values at different lineages suggest a dynamic change of intensity of sexual selection, even among the five BOP species that diverged in 15 Myr (ref. [51]). A social mating system has previously been shown to influence the degree of sexual selection in birds[52] but we did not find a significant (Wilcoxon rank sum test, $P>0.05$) difference of Z/A values between the monogamous species versus polygynous species, probably because of the few species used for comparison here. While the significant (permutation test, two-sided $P<0.05$) fast-Z pattern of the sexually monochromatic American crow may reflect the sexual selection acting on the ancestral lineage leading to Corvidae species, a lack of such a pattern in Raggiana and magnificent BOP species is unexpected. These species are known for their lekking behaviours[50,53], with which few males dominate almost all females for copulation through out-competing other males. This produces a strongly biased male-mating success and direct challenge for maintaining genetic variation in the population (the lekking paradox)[54]. Few field quantitative studies have been performed on BOP species; it will be interesting to investigate whether Raggiana and magnificent BOP female

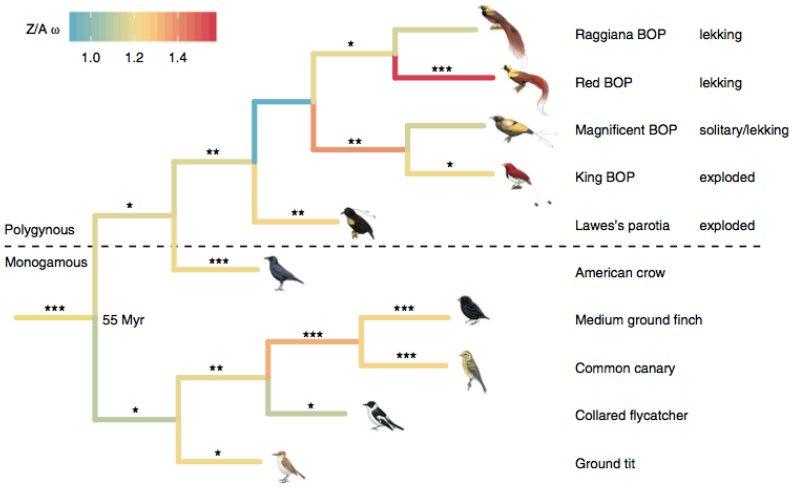


Fig. 3 | Fast-Z evolution of songbirds. Difference of evolution rates are shown for Z-linked genes versus autosome genes (Z/A value) as a measurement of fast-Z effect throughout the lineages of studied songbird species. There are on average 813 Z-linked genes. Genes from chromosome 4 and 5 are used to represent autosomal genes. Chromosome-wide *dN/dS* values are compared between chromosome Z and autosomes. The tree length and colour are scaled to the Z/A value, with lineages that show a significant (permutation test, re-sampled 1,000 times, two-sided $P<0.05$) fast-Z pattern labelled with asterisks. ***$P<0.001$, **$P<0.01$, *$P<0.05$. Social mating systems ('monogamous' versus 'polygamous') and male display type[51] ('lekking', 'exploded lekking', 'solitary display') are labelled. All bird illustrations were ordered from https://www.hbw.com/; ref. [91].

individuals solve the lekking paradox by changing mating preference and mate with more males than presumed.

Conserved gene content of the songbird W chromosomes. In contrast to the dynamic evolution of Z-linked genes and sequences, W chromosomes of all the studied songbirds have undergone marked gene loss but exhibit an unexpected conservation of the retained gene repertoire across species. The numbers of assembled W-linked genes range from 31 in house sparrow to 63 in the king BOP, compared to about 600–800 Z-linked genes (Fig. 4a and Supplementary Tables 2 and 7). These numbers are probably an underestimate because genes embedded in the highly repetitive regions may be missing from the current W chromosome assemblies. In general, Corvida species have retained more W-linked genes than Passerida species (Supplementary Table 5), probably due to their longer generation time thus lower mutation rate. Most W-linked genes are single-copy without lineage-specific expansion, except for *HINT1W* (Supplementary Fig. 15). Despite rare occasions of gene retroposition in birds[55], we find one W-linked gene that is derived through retroposition from an autosomal gene *NARF* in American crow (Supplementary Fig. 16). It will be interesting to investigate whether this gene shows signatures of female-specific selection, for example, a new pattern of ovary-specific expression, which drives its fixation on the W chromosome. Fifty-seven genes are shared by at least one Corvida and another Passerida species and 23 genes are shared between at least one songbird species and chicken[24]. This suggests they were present on the W chromosome before the divergence of passerine or Neognathae species. Despite the independent origination of S2 in chicken and Neoaves[23], all the chicken W-linked genes but one are also found in passerines, indicating similar underlying evolutionary forces governing their convergent retention since Galloanserae and Neoaves diverged from each other 89 Myr ago.

To examine such forces, we performed gene ontology analyses on the 79 genes that are present on the W chromosome of at least one songbird species. They are enriched ($P<0.01$, Fisher's exact test) for two gene ontology terms of 'DNA binding' and 'transcription factor activity, sequence-specific DNA binding' (Supplementary Table 6). This indicates that, similar to the mammalian Y-linked genes[30], some W-linked genes are retained for their important functions of regulating gene activities elsewhere in the genome. The Z-linked homologues of lost genes evolve significantly faster ($P=0.002165$, Wilcoxon rank sum test) with ω ratios higher than those of the retained genes on the W chromosome (Fig. 4b and Supplementary Fig. 17). This shows a different selective pressure acting on these two sets of genes on the proto-sex chromosomes. As this pattern maybe confounded by the 'faster-Z' effect of hemizygous Z-linked genes, we studied the autosomal orthologues of these genes in green anole lizard. We found that the lizard orthologues of retained genes have significantly higher ($P<1.497\times10^{-5}$, Wilcoxon rank sum test; Fig. 4c) expression levels in all tissues of both males and females, and also a broader expression pattern than those of the lost genes across all the tissues (Fig. 4d). The patterns are consistent among the four songbird evolutionary strata; or if we use emu to infer the ancestral expression pattern (Supplementary Fig. 18), whose sex chromosomes are largely PAR. In addition, the retained genes on the W chromosomes are more likely to be dosage-sensitive than those that have become lost. This is indicated by their significantly higher (Wilcoxon rank sum test, $P<0.001$) predicted haplo-insufficiency scores for the human orthologues[56] of the former than those of the latter (Supplementary Fig. 19). This is consistent with the patterns found for chicken or mammalian W- or Y-linked genes[24,30]. We have not found an excess of ovary-biased lizard orthologues among those of the retained W-linked genes: only 6 out of 72 (8.3%) are ovary-biased while the genome-wide proportion is about 20%. This suggests that female-specific selection may not play an important role in preventing the gene loss, or that certain genes undergo positive selection on the songbirds' W chromosomes, which is consistent with the result of collared flycatcher[10].

Comparing gene loss between avian W chromosomes and mammalian Y chromosomes. Overall, 4.6–9.2% of songbird single-copy W-linked genes, compared to 1.6–3.0% mammalian single-copy

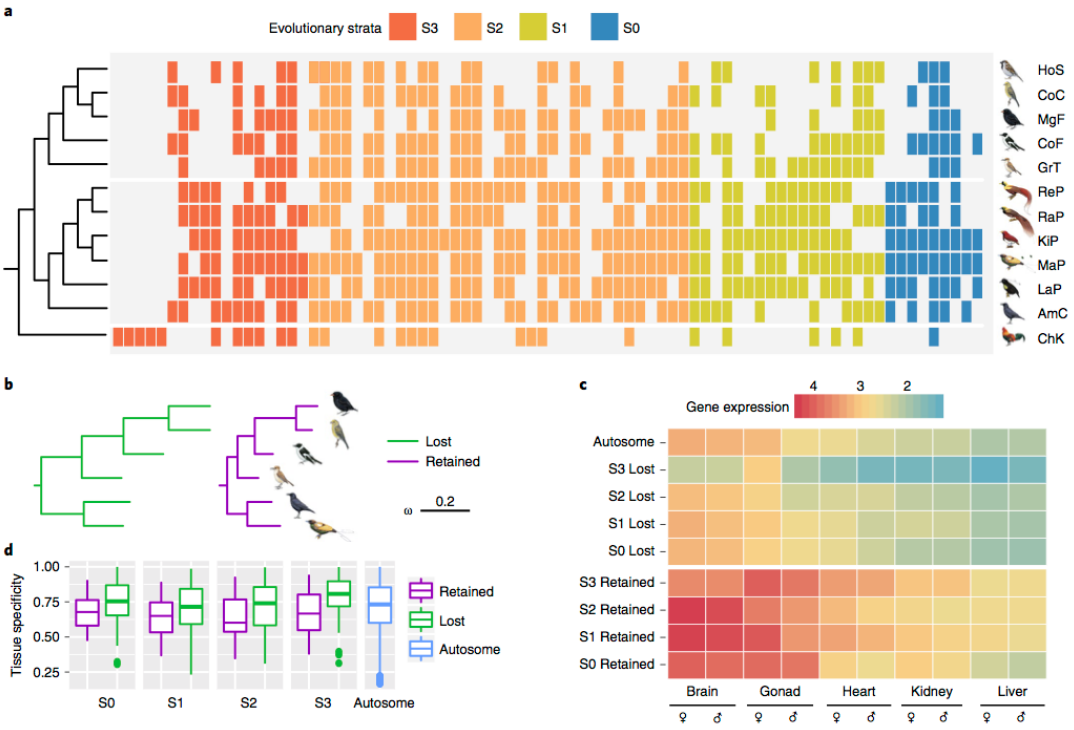


Fig. 4 | The W-linked genes are preserved by purifying selection. a, The retained W-linked genes of each studied songbird species (the species common names are shown with a three-letter code corresponding to those in Fig. 1), as well as those of chicken, with homologous genes aligned vertically. The order of genes follows that of their emu homologues along the Z chromosome. The colours represent the evolutionary strata of songbirds. Species abbreviations: HoS, house sparrow; CoC, Atlantic Canary; MgF, medium ground finch; CoF, collared flycatcher; GrT, ground tit; ReP, red BOP; RaP, Raggiana BOP; KiP, king BOP; MaP, magnificent BOP; LaP, Lawes's parotia; AmC, American crow; ChK, chicken. **b**, The Z-linked genes without W-linked homologues (green, 'Lost') evolve faster than those with W-linked homologues retained (red, 'Retained'), as indicated by their branch lengths scaled to *dN*/*dS* ratios. **c**, The Z-linked genes whose W-linked homologues have become lost (upper panel) tend to have a higher expression level (measured by TPM) in their lizard orthologues than those with W-linked homologues retained (lower panel). The genes are divided further by their resided stratum and the expression level is shown by log-transformed medium expression values of each category as colour-coded heatmap. **d**, Gene expression tissue specificity in green anole lizard for the homologous avian Z-linked genes. In the boxplots, the horizontal line shows the medium value, the whiskers show the 25% and 75% quartile values of expression tissue specificity. All bird illustrations were ordered from https://www.hbw.com/; ref. [91].

Y-linked genes[30] have been retained for their essential or sex-specific functions. A higher retention ratio of W-linked genes in birds can be partially attributed to the generally much lower mutation rate of W chromosome relative to Y chromosome by male-driven evolution effect (Supplementary Fig. 13), assuming a similar generation time between mammals and birds. In addition, a more frequent and stronger sex-specific selection acting on the Y chromosome than on the W chromosome, sometimes driving the massive expansion of Y-linked gene copies with male-related function[29], probably also contributed to a faster rate of Y chromosome gene loss by hitchhiking effect. To examine the tempo of gene loss throughout the evolution of songbirds sex chromosomes, we conservatively reconstructed the numbers of retained W-linked gametologues at each phylogenetic node of avian tree (Fig. 5a and Supplementary Table 8) by identifying the genes present on any of the studied avian W chromosomes. We found that in each stratum, the percentage of gene loss is always much larger at an earlier evolutionary time point than the recent ones and this is consistent between birds and mammals (Fig. 5b). Thus, most gene loss probably occurred during the early stages of recombination suppression, and the rate of gene loss markedly decreases by the less retained genes. Although convergent gene loss may cause an overestimate of lost genes at more ancestral time points (for example, in S0 region), this probably has little influence on our estimate in the most recent songbird-specific stratum S3 which has already lost 69.8% of the W-gametologues in 50 Myr. We also found that the retained genes of songbird W chromosomes are often close to each other (Supplementary Fig. 20), indicating large sequence deletions have contributed to marked gene loss.

The decrease of gene loss rate on Y/W chromosomes over evolutionary time can be explained by a weaker Hill–Robertson effect that the less retained genes can induce, which has been previously shown by simulation study[28]. In addition, the size, that is, the ancestral gene number of older evolutionary strata which would have undergone more serious gene loss must have a larger influence on the extant number of retained genes. Thus, a lower rate of retained mammalian Y-linked genes relative to avian W-linked genes can be attributed to the fact that the first two or three mammalian evolutionary strata occurred before the divergence of eutherians; together

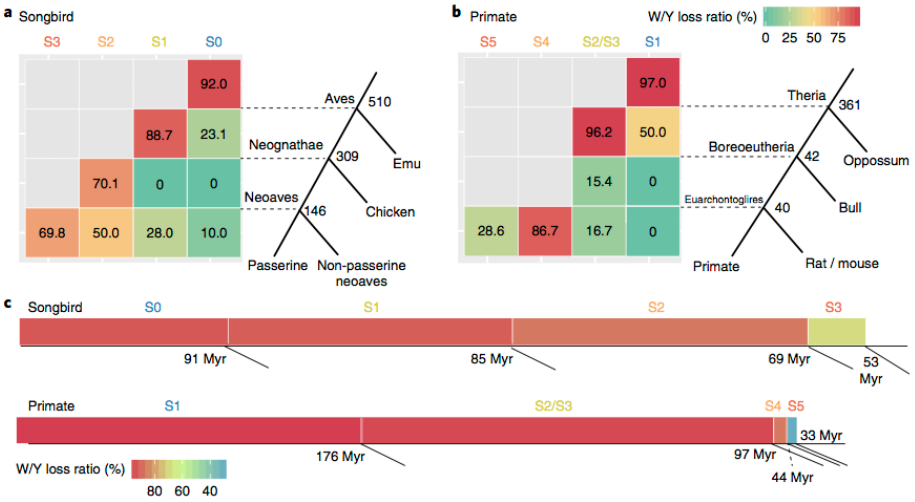


Fig. 5 | Comparison of gene loss between W chromosomes of songbirds and Y chromosomes of primates. a, Percentage of gene loss and ancestral gene number for each evolution stratum at each phylogenetic node. **b**, Similar analyses for the Y-linked gametologues of primates on the basis of the data of ref. [24], with S1 as the first stratum of eutherian mammals. **c**, Lengths of songbird W or primate Y chromosomes scaled to the ancestral gene number of each evolutionary stratum, with the colour scaled to the overall percentage of gene loss. The ages of evolutionary strata are indicated by the numbers (in million years) at the nodes below the bars. As eutherian mammals have much larger ancestral evolutionary strata than those of birds, they probably have more gene loss on the Y chromosome.

these account for over 93.2% of the entire gene content of ancestral Y chromosome, while those of birds only account for 53.3% of the entire ancestral W-linked gene content (Fig. 5c).

Discussion

The evolution of sex chromosomes is often but not always (for example, in frogs[57], ratites[23] and python[58]) marked with episodes of recombination suppressions that eventually restrict the recombining region in one or two small PARs at the end of chromosome. The resulting patterns of evolutionary strata have been widely reported in many animal and plant species, with the formation mechanism presumed to be chromosomal inversions[59]. Indeed, footprints of inversions in the latest two strata between human X and Y chromosomes have been found by examining the synteny order between X/Y, and particularly the homologous X-linked PAR boundary (PAB) sequence on the Y chromosome that has been disrupted into two dispersed sequences[60]. The Y-linked PAB is defined by an insertion of *Alu* element[61], with similar insertions of various types of TE elements at PAB reported among other mammals such as cattle and pig (reviewed in ref. [62]). Such TE insertions were probably due to the reduction of recombination rate at PAB, after chromosome inversion suppressed the recombination in the youngest stratum bordering the PAR. In the case of birds, a W-linked chromosome inversion giving birth to a young evolutionary stratum would be fixed within the population, given its advantage of linking sexual antagonistic genes (beneficial to female but detrimental to male in the case of ZW system) to the sex-determining genes. However, neither W-linked sexual antagonistic alleles nor sex-determining genes have been identified so far in birds.

Alternatively, we propose that the TE insertion may have occurred before the chromosome inversion and initiated the recombination suppression[62]. In this case, the recombination loss between sex chromosomes proceeded gradually rather than immediately by chromosome inversion. We found a genome-wide burst of CR1 subfamilies specifically concentrated at PAB (Fig. 2d and Supplementary Table 3). This songbird PAB has undergone genomic deletions or rearrangements independently in many other bird species[48,49] and exhibits a different subfamily of CR1 insertion in rifleman, therefore it is probably a mutation hotspot. It has been reported in many species that local recombination rate and abundance of TE elements generally have a negative association, with their causal relationship being difficult to disentangle[46]. However, several patterns suggest that the CR1–E accumulation is the cause rather than the result of recombination suppression at songbird PAB: first, in several species (for example, Lawes's parotia and King BOP; Fig. 2d and Supplementary Fig. 8), the CR1–E repeats are also enriched in the part of PAR close to PAB, where there is supposed to be frequent recombination. Second, the abundance of CR1–E is gradually decreased further away from the PAB. Third, only the CR1–E repeats but not any other type of CR1 or repeat families have accumulated at the PAB. These patterns are distinct from that of *Alu* insertion at the human PAB[61], which does not extend into PAR or show a specific enrichment at certain regions at a chromosome-wide level. They are unlikely if chromosome inversion occurred before the CR1–E accumulation, which predicts a more uniformly distributed accumulation of various kinds of repeat elements (for example, LTR elements in this study; Supplementary Fig. 21) that would not extend into the PAR. In addition, our comparative analyses between species indicate there was no Z-linked inversion in S3 (Fig. 2a and Supplementary Fig. 5), although we cannot exclude the possibility of a W-linked inversion that may have contributed to the formation of S3. Verification of the latter requires improvements of genome assembly using, for example, PacBio/Nanopore sequencing technology to assemble the highly repetitive W-linked sequence.

We propose that TEs probably reduced the recombination rate in PAR through, for example, changing the chromatin structure or disrupting the recombination hotspot[63,64]. TE accumulation was probably selected against at the beginning because it disrupted

gene functions. This has been demonstrated by results showing that at the PAB where CR1–E has accumulated, several genes have been partially or completely deleted in songbirds[48,49]. However, the resulting reduction of recombination rate can provide the selective advantage of accelerating the fixation of pre-existing sexual antagonistic polymorphic alleles in PAR through sex-biased transmission, or subjecting the PAR for the 'fast-Z' evolution by male-driven evolution effect (Fig. 3) and increasing its exposure for male-biased selection, so that new sexual antagonistic alleles may more frequently emerge and become fixed. The latter has been implicated by the recent findings in songbirds that male-specific trait genes, for example those related to sperm morphology[65] or plumage colours[66], which have recently diverged within or between populations, are enriched on the Z chromosome. In addition, TE accumulation is likely to increase the chance of chromosome inversions through ectopic recombination or by reducing the selective constraints on gene synteny. The latter is supported by our result that older evolutionary strata have undergone more Z-linked genomic rearrangements between songbird species than the younger ones (Fig. 2a), which creates a positive feedback once the recombination suppression was initiated. This provides a mechanistic explanation for a more frequent fixation of Z-linked inversions among passerines.

While Z chromosome is predicted to accumulate dominant male-beneficial mutations, W chromosome is expected to accumulate female-beneficial mutations responding to the female-specific transmission. However, both previous works in chicken and flycatcher[10,24], as well as our study have not found evidence for such 'feminization' of W chromosome. This is in contrast to reported 'masculinization' cases of ancestral Y chromosomes of mammals[29] or of recently evolved Y chromosome of *Drosophila miranda*[31]. Y-linked genes specifically expressed in male germline have either greatly amplified their copy numbers or upregulated their expression levels in these systems. Such a difference can be explained by the fact that, regardless of the sex chromosome type, sexual selection is more often targeting males in most species. Thus Z/Y chromosome is more frequently influenced than the W/X chromosome due to male-biased transmission, although the X chromosome is nevertheless expected to accumulate recessive male-beneficial alleles[9]. The convergently evolved pattern shared between the mammalian Y and avian W chromosomes is largely attributed to the essential genes that have important regulatory functions and are preferentially retained over the long period of recombination suppression (Fig. 4). Besides a weaker Hill–Robertson effect by the course of Y or W chromosome evolution, these essential genes probably have also contributed to the decreased rate of gene loss, as they are under stronger selective constraints than other genes that became lost at earlier stages. However, previous transcriptome comparison of chicken breeds selected for egg-laying versus fighting, that is, female-specific versus male-specific traits, has found most W-linked genes are upregulated in the former[34]. Few high-quality avian W chromosome sequences are available except for that of chicken. Songbirds provide a rich resource with many species (for example, blue crow) having a reversed direction of sexual selection and ornamented females. Application of long-read sequencing technology will help to elucidate the role of W chromosome in sexual selection and speciation of birds[67].

Methods

Genome assembly and annotation. Genomic DNA was extracted from fresh tissue samples of female BOP species *Cicinnurus regius* (museum catalogue number ANWC B24969), *Cicinnurus magnificus* (ANWC B27061), *Paradisaea raggiana* (USNM638608) and both sexes of *Paradisaea rubra* (YPM84686; ref. [68]) and *Parotia lawesii* (ANWC B26535 and ANWC B15265), using Thermo Scientific KingFisher Duo Prime purification system or EZNA SQ Tissue DNA Kit. Paired-end and mate pair libraries for these samples were prepared by SciLifeLab in Stockholm, Sweden. All libraries were sequenced on Illumina HiSeq 2500 or Hiseq X v4 at SciLifeLab or BGI. We also used the published female genomes of *Corvus brachyrhynchos*, *Serinus canaria*, *Passer domesticus*, *Geospiza fortis*, *Ficedula albicollis* and *Pseudopodoces humilis* for analysis in this work (Supplementary Table 9). The BOP genomes were assembled using ALLPATHS-LG (52488; ref. [69]) with 'HAPLOIDIFY = True'. For *P. raggiana*, due to the lack of overlapping paired-end reads, SOAPdenovo2 (v.2.04; ref. [70]) was used instead (K-mer 23). Gaps of the SOAPdenovo2 scaffolds were filled using Gapcloser (v.1.12) with default parameters. Gene models were annotated using the MAKER pipeline (v.2.31.9) in two rounds[71]. The reference protein sequences of zebra finch, great tit, hooded crow, American crow, collared flycatcher and chicken were downloaded from NCBI RefSeq (Supplementary Table 9). Using the reference protein sequences and chicken HMM (Hidden Markov Models), an initial set of gene models was obtained by using MAKER and those models were taken for SNAP (v.2013.11.29) model training[72]. In addition, 3,000 gene models with top AED (Annotation Edit Distance) scores were selected for Augustus (v.3.2.3) training[73]. The trained gene models and the protein sequences were taken as input for MAKER in the second run. To annotate repeats, first we used RepeatModeler (http://www.repeatmasker.org/, v.1.0.10) with default parameters to identify and classify repeat elements for each species. Then we combined each individual library with an avian repeat library[68] to annotate repeats using RepeatMasker (v.4.0.7) with the parameter '-a -xsmall -gccalc'. This repeat annotation pipeline was also applied to other published avian species for re-annotation to allow for a direct comparison.

Identification of sex-linked sequences. We used the published Z chromosome sequences of two species, great tit[6] and hooded crow[4], as references to maximize the identification of the sex-linked sequences in the studied species. The great tit Z chromosome assembly is the most complete among the published genomes of songbirds and hooded crow is the closest species to BOP with male genome data available. We first used nucmer from MUMMer package (v.4.0; ref. [74]) for genome-wide pairwise sequence alignment with the parameter '-b 400' between all the studied species versus the reference species. Only the best one-to-one alignments were retained (delta-filter-1). Any scaffold longer than 10 kb that has over 60% of the sequence aligned to reference Z chromosome was identified as a candidate Z-linked scaffold in that species. All the scaffolds, including the candidate Z-linked scaffolds were examined for whether their female sequencing coverage values are about half of those of autosomes. The raw female reads were mapped to the scaffold sequences by using bwa (v.0.7.16a, ref. [75]; default parameters of the maximal exact matches algorithm), subsequently the average sequencing coverage of every 50-kb window was calculated. We then plotted the distribution of coverage values of all the windows (Supplementary Fig.1) to decide whether a scaffold is a candidate Z- or W-linked scaffold by showing the half-coverage value of autosomes. Sequencing coverage of each nucleotide position was also counted using 'samtools depth' before calculating window-based coverage. Any alignment with low mapping quality (lower than 60) was not counted to exclude the effect of probable misalignments. Additionally, any site with high coverage (three times larger than average) was excluded, as it was probably derived from repetitive sequences.

To identify the W-linked scaffolds, first we focused on all the half-coverage scaffolds that aligned with the reference Z chromosome and inspected their distribution of the proportion of aligned sequences to decide a cutoff. This cutoff was used to separate the candidate Z- and W- linked scaffolds, and varies from species to species, possibly due to the varying assembly quality and/or the divergence from the reference genomes. We excluded the candidate W-linked scaffolds that over 10% of the sequences were aligned to reference autosomes, or a larger portion of sequences aligned to the autosomes than the Z. Finally, only the scaffolds longer than 10 kb were kept because shorter scaffold often show ambiguous coverage patterns. For species that have both male and female sequencing reads available, we directly verified the candidate W-linked sequences by mapping the male reads. Specifically, for each scaffold, the number of nucleotide sites that were mapped by male and female sequencing data were counted as N_m and N_f respectively, with their ratios as N_m/N_f. W-linked scaffolds are expected to have Nm/Nf ratios close to zero, while autosome or Z-linked scaffolds tend to have an expected ratio of 1 (this is the ratio of mappable sites; for the ratio of coverage between sexes, the expected value would be 2). Given the short divergence time of BOP species[51], we are able to map substantial numbers of male reads of a red BOP to the three BOP species (magnificent BOP, king BOP and Raggiana BOP; over 95% of the genomes can be mapped) lacking the male data to verify their W-linked sequences. We used the known PAR sequences of zebra finch[76] and flycatcher[40] to infer those of other species using nucmer (-b 400) and then confirmed their similar levels of female coverage value to autosomes.

Demarcation of evolutionary strata. We ordered and oriented the identified Z-linked scaffolds of all BOP species into one pseudo-chromosomal sequence (pseudo-chrZ) on the basis of their alignments against the Z chromosome of great tit. For Fig. 2a, we used chromosomal or nearly chromosomal assemblies of four species but not BOPs. For example, hooded crow has 15 Z-linked scaffolds and ten of them are larger than 1 Mb. We determined the relative order and orientation of the crow scaffolds according to their alignment with the great tit Z chromosome. We used nucmer for pairwise alignment of the Z or pseudo-Z chromosomes between species. Alignments short than 2 kb were excluded to avoid probable

misalignments. Similarly, for BOP species, we created pseudo-chromosome Z using the great tit Z chromosome as reference. The pseudo-Z chromosome of emu was built using ostrich Z chromosome[77] as reference and was available from Xu et al.[78]. The W-linked scaffolds were then aligned to the pseudo-chrZ of the same species using lastz (v.1.04; ref. [79]) with '--step = 19 --hspthresh = 2200 --inner = 2000 --ydrop = 3400 --gappedthresh = 10000', after masking the repetitive sequences. Sequence similarity of the alignments between the Z and W chromosomes was calculated by the script pslScore from UCSC Genome Browser (https://genome.ucsc.edu/). Individual alignments that have sequence identities lower than 60 or higher than 96, or alignment lengths shorter than 65% were removed as those are probably derived from misalignments or unmasked repeats. After that, we ordered the W-linked scaffolds by their aligned positions along the pseudo-chrZ. We then extracted Z/W pairwise alignments for every non-overlapping sliding window of 100 kb along the pseudo-chrZ and calculated the sequence divergence level for the windows whose lengths of ZW alignment are longer than 2 kb. The window-based sequence divergence levels were then plotted along the pseudo-chrZ, with the shift of divergence level to demarcate the boundaries between evolutionary strata. Since few W-linked sequences have been assembled for the most ancient stratum S0, we mapped its reshuffled fragments in songbirds on the basis of their homology with the emu S0. Our previous study showed that emu has a recent species-specific stratum (S1), while the oldest stratum (S0) is shared by all birds[23]. This allows for the demarcation of S1 and S0 by detecting their differential degree of Z/W differentiation in emu. Specifically, by using a relatively relaxed mapping criteria (bwa mem) to map the female sequencing reads, only S0 showed reduced coverage relative to autosomes or PAR (Supplementary Fig. 4), while S1 showed reduced coverage when stringent mapping was applied (bwa sampe –a 900 –n 1 –N 0 –o 10000). To examine the accumulated LINE (mostly CR1) elements at the PAB, we first divided them into each subtype according to the RepeatMasker annotation. Among all the subtypes, CR1–E1 was usually ranked with the highest or second highest copy number at the S3 region across all songbird species. Other high-ranking subtypes included CR1–E3, CR1–E5, CR1–E4, CR1–E6, CR1–J2 and CR1–Y2. Then we plotted each subtype's abundance with a 100 kb non-overlapping window along the Z chromosome, in all the studied songbirds, as well as outgroup species rifleman and falcon, to identify the burst of CR1–E1.

Sex-linked gene analyses. We used BLAT (v.35.1; ref. [80]) to align the annotated coding sequence of W-linked genes to the Z chromosome to search for their homologous pairs after removing the LTR genes. Then we produced pairwise gametologue alignments using MUSCLE (v.3.8.31) with the default parameters[81] and manually inspected the alignments to remove genes with short or ambiguous alignments. For species other than BOPs, gene models of the W chromosomes were directly retrieved from the RefSeq genome annotation (https://www.ncbi.nlm.nih.gov/refseq/), with some of them also subjected to manual inspections. To determine the orthologous relationship among the studied species, we first extracted the sequence of the longest protein of each gene. Those protein sequences were subjected to all-versus-all BLAST search that was implemented through the program proteinortho (v.5.16; ref. [82]). BLAST hits with sequence identity lower than 50% or aligned percentage lower than 50% were removed. We also took gene synteny information into account when grouping orthologous genes. Besides the 12 female genomes for which we studied the sex chromosomes, we also included high-quality genomes of great tit, hooded crow and ostrich (Supplementary Table 9). We retained those orthologous groups if they contain sequences of at least ten species. To estimate the substitution rates of coding sequences, first we performed multiple sequence alignment for orthologous genes. We used the guidence2 pipeline (http://guidance.tau.ac.il/ver2/source.php) which used PRANK (v.170427) to align sequences of codons, with the default parameters. To filter low-quality sites in the alignments, we ran trimal (http://trimal.cgenomics.org/) with '-gt 0.8'. The phylogeny of the birds was extracted from Jetz. et al.[83]. We used codeml from the PAML package (v.4.9e; ref. [84]) to estimate the synonymous substitution rates (d*S*) and non-synonymous substitution rates (d*N*). To estimate the chromosome-wide d*N* and d*S*, sums of synonymous or non-synonymous substitutions were divided by those of total synonymous or non-synonymous sites, as applied in Wright et al.[22]. Individual genes with abnormal d*N* (higher than 0.1, in total 179 genes) or d*S* (higher than 0.8, in total 135 genes) out of 111,748 orthologous gene groups were removed, as those were probably caused by misalignments or misassignment of orthologues. Confidence intervals were calculated by 100 bootstrappings. Chromosome-wide d*N*/d*S* (ω) was calculated by the ratios of chromosome-wide d*N* to chromosome-wide d*S*. The fast-Z effect was measured by Z/A value as the ratio of ω values of Z-linked genes to autosomal genes and we calculated the Z/A value for each terminal branch and internal branch. To determine if the difference of ω between Z-linked and autosomal genes is significant, we performed permutation test by resampling 1,000 times to calculate the two-sided *P*-values. The genes of chromosome 4 and chromosome 5 were used to represent autosomal genes as the sizes of those two chromosomes are similar to the Z chromosome. For each gametologue pair, we grouped together Z-linked genes and assembled W-linked genes and performed multiple sequence alignment. The same guidence2 pipeline was used as in sequence divergence analysis. For S3 genes, we also included rifleman[85] to infer the origination time of S3. We used IQ-TREE (v.1.6.1; ref. [86]) to construct phylogenetic trees. The best substitution model was automatically selected in by IQ-TREE. We ran bootstrapping 100 times to evaluate the confidence levels of phylogenies with ostrich as outgroup to root the tree. The gene ontology annotations for both studied gametologue-pair genes (list) and all Z-linked genes (background) of chicken was analysed by DAVID v.6.8 (ref. [87]). Gene ontology term enrichment was estimated by comparing the numbers of appearance of GO terms of 'list' gene versus 'background' gene. The GC content of the third position of codons (GC3) was calculated using codonW (http://codonw.sourceforge.net//culong.html) for the longest isoform of each gene.

Gene loss analysis. We identified 673 Z-linked orthologous genes shared between chicken and emu as the putative ancestral genes on the proto-sex chromosomes of birds. For the gene cluster that was lost in chicken at the *DCC* loci on the S3 and ancestral gene content was inferred on the basis of Fig. 3 of Patthey et al.[49]. They were then grouped into four evolutionary strata according to the strata annotation of songbird Z chromosomes. At each node of the avian phylogenetic tree, we calculated the ratio of lost genes to ancestral genes of that node. For the nodes leading to Passerida and Corvida, if there was at least one species retaining a W-linked gene, we inferred that that gene was present in their ancestor. Similarly, we defined the presence of ancestral genes in Passeriformes, Neoaves and Neognathae according to the presence/absence of W-linked genes in other published avian species[23,24].

Gene expression analysis. We downloaded the raw RNA-seq reads of green anole (brain, gonad, liver, heart and kidney) and emu (brain, gonad and spleen) from SRA (Supplementary Table 9). In addition, we collected the transcriptomes of adult emu kidneys of both sexes. We used the RSEM pipeline (v.1.3.0; ref. [88]) to quantify the gene expression levels. The RSEM pipeline used STAR (v.2.5.30; ref. [89]) with default parameters to map raw reads to the transcriptomes which was constructed on the basis of gene annotations. The expectation–maximization (EM) algorithm was used to estimate the abundance of transcripts by RSEM. The expression levels were estimated at the gene level, in the form of TPM (transcripts per million). The mean TPM value of biological replicates was calculate for each gene. Tissue specificity of gene expression were estimated by calculating tau[90].

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Genome sequencing and RNA-seq data generated in this study have been deposited in the NCBI SRA under PRJNA491255. The raw genomic reads of *Paradisaea raggiana* are available in the CNGB Nucleotide Sequence Archive (https://db.cngb.org/cnsa; accession number CNP0000186). The genome assemblies are available under NCBI BioProject portal (PRJNA491255). The IDs of W-linked scaffolds are included in Supplementary Table 10.

Code availability

Custom scripts and pipelines used in this study have been deposited at Github (https://github.com/lurebgi/BOPsexChr).

Received: 29 October 2018; Accepted: 22 February 2019;
Published online: 1 April 2019

References

1. World Bird List v.8.2 (IOC) (2018); https://www.worldbirdnames.org/ioc-lists/master-list-2/
2. Barker, F. K., Cibois, A., Schikler, P., Feinstein, J. & Cracraft, J. Phylogeny and diversification of the largest avian radiation. *Proc. Natl Acad. Sci. USA* **101**, 11040–11045 (2004).
3. Ellegren, H. et al. The genomic landscape of species divergence in Ficedula flycatchers. *Nature* **491**, 756–760 (2012).
4. Poelstra, J. W. et al. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* **344**, 1410–1414 (2014).
5. Tuttle, E. M. et al. Divergence and functional degradation of a sex chromosome-like supergene. *Curr. Biol.* **26**, 344–350 (2016).
6. Laine, V. N. et al. Evolutionary signals of selection on cognition from the great tit genome and methylome. *Nat. Commun.* **7**, 10474 (2016).
7. Mank, J. E. Sex chromosomes and the evolution of sexual dimorphism: lessons from the genome. *Am. Nat.* **173**, 141–150 (2009).
8. Coyne, J. A. Genetics and speciation. *Nature* **355**, 511 (1992).
9. Charlesworth, B., Coyne, J. & Barton, N. The relative rates of evolution of sex chromosomes and autosomes. *Am. Nat.* **130**, 113–146 (1987).
10. Smeds, L. et al. Evolutionary analysis of the female-specific avian W chromosome. *Nat. Commun.* **6**, 7330 (2015).
11. Hooper, D. M. & Price, T. D. Chromosomal inversion differences correlate with range overlap in passerine birds. *Nat. Ecol. Evol.* **1**, 1526–1534 (2017).
12. Hooper, D. M. & Price, T. D. Rates of karyotypic evolution in Estrildid finches differ between island and continental clades. *Evolution* **69**, 890–903 (2015).

13. Backström, N. et al. A high-density scan of the Z chromosome in Ficedula flycatchers reveals candidate loci for diversifying selection. *Evolution* **64**, 3461–3475 (2010).
14. Elgvin, T. O. et al. Hybrid speciation in sparrows II: a role for sex chromosomes?. *Mol. Ecol.* **20**, 3823–3837 (2011).
15. Storchová, R., Reif, J. & Nachman, M. W. Female heterogamety and speciation: reduced introgression of the Z chromosome between two species of nightingales. *Evolution* **64**, 456–471 (2010).
16. Li, W. H., Yi, S. & Makova, K. Male-driven evolution. *Curr. Opin. Genet. Dev.* **12**, 650–656 (2002).
17. Wang, Z. et al. Temporal genomic evolution of bird sex chromosomes. *BMC Evol. Biol.* **14**, 250 (2014).
18. Ellegren, H. Characteristics, causes and evolutionary consequences of male-biased mutation. *Proc. Biol. Sci.* **274**, 1–10 (2007).
19. Smeds, L., Qvarnström, A. & Ellegren, H. Direct estimate of the rate of germline mutation in a bird. *Genome Res.* **26**, 1211–1218 (2016).
20. Vicoso, B. & Charlesworth, B. Effective population size and the faster-X effect: an extended model. *Evolution* **63**, 2413–2426 (2009).
21. Mank, J. E., Nam, K. & Ellegren, H. Faster-Z evolution is predominantly due to genetic drift. *Mol. Biol. Evol.* **27**, 661–670 (2010).
22. Wright, A. E. et al. Variation in promiscuity and sexual selection drives avian rate of Faster-Z evolution. *Mol. Ecol.* **24**, 1218–1235 (2015).
23. Zhou, Q. et al. Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science* **346**, 1246338 (2014).
24. Bellott, D. W. et al. Avian W and mammalian Y chromosomes convergently retained dosage-sensitive regulators. *Nat. Genet.* **49**, 387–394 (2017).
25. Charlesworth, D., Charlesworth, B. & Marais, G. Steps in the evolution of heteromorphic sex chromosomes. *Heredity* **95**, 118–128 (2005).
26. Bachtrog, D. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat. Rev. Genet.* **14**, 113–124 (2013).
27. Charlesworth, B. & Charlesworth, D. The degeneration of Y chromosomes. *Philos. Trans. R. Soc. Lond. B* **355**, 1563–1572 (2000).
28. Bachtrog, D. The temporal dynamics of processes underlying Y chromosome degeneration. *Genetics* **179**, 1513–1525 (2008).
29. Soh, Y. Q. et al. Sequencing the mouse Y chromosome reveals convergent gene acquisition and amplification on both sex chromosomes. *Cell* **159**, 800–813 (2014).
30. Bellott, D. W. et al. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature* **508**, 494–499 (2014).
31. Zhou, Q. & Bachtrog, D. Sex-specific adaptation drives early sex chromosome evolution in *Drosophila*. *Science* **337**, 341–345 (2012).
32. Koerich, L. B., Wang, X., Clark, A. G. & Carvalho, A. B. Low conservation of gene content in the *Drosophila* Y chromosome. *Nature* **456**, 949–951 (2008).
33. Wright, A. E., Harrison, P. W., Montgomery, S. H., Pointer, M. A. & Mank, J. E. Independent stratum formation on the avian sex chromosomes reveals inter-chromosomal gene conversion and predominance of purifying selection on the W chromosome. *Evolution* **68**, 3281–3295 (2014).
34. Moghadam, H. K., Pointer, M. A., Wright, A. E., Berlin, S. & Mank, J. E. W chromosome expression responds to female-specific selection. *Proc. Natl Acad. Sci. USA* **109**, 8207–8211 (2012).
35. Cortez, D. et al. Origins and functional evolution of Y chromosomes across mammals. *Nature* **508**, 488–493 (2014).
36. Bergero, R., Forrest, A., Kamau, E. & Charlesworth, D. Evolutionary strata on the X chromosomes of the dioecious plant *Silene latifolia*: evidence from new sex-linked genes. *Genetics* **175**, 1945–1954 (2007).
37. Gorelick, R. et al. Abrupt shortening of bird W chromosomes in ancestral Neognathae. *Biol. J. Linn. Soc. Lond.* **119**, 488–496 (2016).
38. Rice, W. R. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. *Evolution* **41**, 911–914 (1987).
39. Claramunt, S. & Cracraft, J. A new time tree reveals Earth historys imprint on the evolution of modern birds. *Sci. Adv.* **1**, e1501005–e1501005 (2015).
40. Smeds, L. et al. Genomic identification and characterization of the pseudoautosomal region in highly differentiated avian sex chromosomes. *Nat. Commun.* **5**, 5448 (2014).
41. Pala, I. et al. Evidence of a neo-sex chromosome in birds. *Heredity* **108**, 264–272 (2012).
42. Ross, M. T. et al. The DNA sequence of the human X chromosome. *Nature* **434**, 325–337 (2005).
43. O'Connor, R. E. et al. Chromosome-level assembly reveals extensive rearrangement in saker falcon and budgerigar, but not ostrich, genomes. *Genome Biol.* **19**, 171 (2018).
44. Wright, A. E., Dean, R., Zimmer, F. & Mank, J. E. How to make a sex chromosome. *Nat. Commun.* **7**, 12087 (2016).
45. Duret, L. & Galtier, N. Biased gene conversion and the evolution of mammalian genomic landscapes. *Annu. Rev. Genomics Hum. Genet.* **10**, 285–311 (2009).
46. Kent, T. V., Uzunović, J. & Wright, S. I. Coevolution between transposable elements and recombination. *Philos. Trans. R. Soc. B* **372**, 20160458 (2017).
47. Suh, A. et al. Mesozoic retroposons reveal parrots as the closest living relatives of passerine birds. *Nat. Commun.* **2**, 443 (2011).
48. Friocourt, F. et al. Recurrent DCC gene losses during bird evolution. *Sci. Rep.* **7**, 37569 (2017).
49. Patthey, C., Tong, Y. G., Tait, C. M. & Wilson, S. I. Evolution of the functionally conserved DCC gene in birds. *Sci. Rep.* **7**, 42029 (2017).
50. Ligon, R. A. et al. Evolution of correlated complexity in the radically different courtship signals of birds-of-paradise. *PLoS Biol.* **16**, e2006962 (2018).
51. Irestedt, M., Jonsson, K. A., Fjeldsa, J., Christidis, L. & Ericson, P. G. An unexpectedly long history of sexual selection in birds-of-paradise. *BMC Evol. Biol.* **9**, 235 (2009).
52. Dunn, P. O., Whittingham, L. A. & Pitcher, T. E. Mating systems, sperm competition, and the evolution of sexual dimorphism in birds. *Evolution* **55**, 161–175 (2001).
53. Diamond, J. Biology of birds of paradise and bowerbirds. *Annu. Rev. Ecol. Syst.* **17**, 17–37 (1986).
54. Kirkpatrick, M. & Ryan, M. J. The evolution of mating preferences and the paradox of the lek. *Nature* **350**, 33–38 (1991).
55. Suh, A. The specific requirements for CR1 retrotransposition explain the scarcity of retrogenes in birds. *J. Mol. Evol.* **81**, 18–20 (2015).
56. Huang, N., Lee, I., Marcotte, E. M. & Hurles, M. E. Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet.* **6**, e1001154 (2010).
57. Stock, M. et al. Ever-young sex chromosomes in European tree frogs. *PLoS Biol.* **9**, e1001062 (2011).
58. Vicoso, B., Emerson, J. J., Zektser, Y., Mahajan, S. & Bachtrog, D. Comparative sex chromosome genomics in snakes: differentiation, evolutionary strata, and lack of global dosage compensation. *PLoS Biol.* **11**, e1001643 (2013).
59. Charlesworth, D. Evolution of recombination rates between sex chromosomes. *Philos. Trans. R. Soc. Lond. B* **372**, 20160456 (2017).
60. Lemaitre, C. et al. Footprints of inversions at present and past pseudoautosomal boundaries in human sex chromosomes. *Genome Biol. Evol.* **1**, 56–66 (2009).
61. Ellis, N. A. et al. The pseudoautosomal boundary in man is defined by an Alu repeat sequence inserted on the Y chromosome. *Nature* **337**, 81–84 (1989).
62. Raudsepp, T. & Chowdhary, B. P. The eutherian pseudoautosomal region. *Cytogenet. Genome Res.* **147**, 81–94 (2015).
63. Zamudio, N. et al. DNA methylation restrains transposons from adopting a chromatin signature permissive for meiotic recombination. *Genes Dev.* **29**, 1256–1270 (2015).
64. Ben-Aroya, S., Mieczkowski, P. A., Petes, T. D. & Kupiec, M. The compact chromatin structure of a Ty repeated sequence suppresses recombination hotspot activity in *Saccharomyces cerevisiae*. *Mol. Cell* **15**, 221–231 (2004).
65. Knief, U. et al. Fitness consequences of polymorphic inversions in the zebra finch genome. *Genome Biol.* **17**, 199 (2016).
66. Campagna, L. et al. Repeated divergent selection on pigmentation genes in a rapid finch radiation. *Sci. Adv.* **3**, e1602404 (2017).
67. Irwin, D. E. Sex chromosomes and speciation in birds and other ZW systems. *Mol. Ecol.* **27**, 3831–3851 (2018).
68. Prost, S. et al. Comparative analyses identify genomic features potentially involved in the evolution of birds-of-paradise. *GigaScience* https://doi.org/10.1093/gigascience/giz003 (2018).
69. Gnerre, S. et al. High-quality draft assemblies of mammalian genomes from massively parallel sequence data. *Proc. Natl Acad. Sci. USA* **108**, 1513–1518 (2011).
70. Luo, R. et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience* **1**, 18 (2012).
71. Cantarel, B. L. et al. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18**, 188–196 (2008).
72. Korf, I. Gene finding in novel genomes. *BMC Bioinformatics* **5**, 59 (2004).
73. Stanke, M. & Waack, S. Gene prediction with a hidden Markov model and a new intron submodel. *Bioinformatics* **19**, ii215–ii225 (2003).
74. Marçais, G. et al. MUMmer4: A fast and versatile genome alignment system. *PLoS Comput. Biol.* **14**, e1005944 (2018).
75. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *Genomics (q-bio.GN)* https://arxiv.org/abs/1303.3997 (2013).
76. Singhal, S. et al. Stable recombination hotspots in birds. *Science* **350**, 928–932 (2015).
77. Zhang, J., Li, C., Zhou, Q. & Zhang, G. Improving the ostrich genome assembly using optical mapping data. *GigaScience* **4**, 24 (2015).
78. Xu, L. et al. Evolutionary dynamics of sex chromosomes of palaeognathous birds. Preprint at *bioRxiv* https://doi.org/10.1101/295089 (2018).
79. Harris, R. S. Improved pairwise alignment of genomic DNA. PhD thesis, The Pennsylvania State University (2017).
80. Kent, W. J. BLAT — the BLAST-like alignment tool. *Genome Res.* **12**, 656–64 (2002).
81. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

82. Lechner, M. et al. Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinformatics* **12**, 124 (2011).
83. Jetz, W., Thomas, G. H., Joy, J. B., Hartmann, K. & Mooers, A. O. The global diversity of birds in space and time. *Nature* **491**, 444–448 (2012).
84. Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).
85. Zhang, G. et al. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**, 1311–1320 (2014).
86. Nguyen, L.-T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
87. Jiao, X. et al. DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics* **28**, 1805–1806 (2012).
88. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
89. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
90. Yanai, I. et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* **21**, 650–659 (2005).
91. del Hoyo, J, . & Elliot, A. & Sargatal, J. & Christie, D. & de Juana, E. *Handbook of the Birds of the World Alive* (Lynx Editions, 2019).

Acknowledgements

We thank the Smithsonian Institute (G. Graves), Australian National Wildlife Collection, CSIRO Sustainable Ecosystems (L. Joseph), Museum Victoria, Australia (J. Sumner), Division of Vertebrate Zoology Yale University, Peabody Museum of Natural History (K. Zyskowski) for tissue samples; and E. Scholes for discussions on BOP. We also acknowledge the support from Science for Life Laboratory, the National Genomics Infrastructure (NGI), Uppmax. L.X. is supported by the uni:docs fellowship programme from University of Vienna. M.I. is supported by the Swedish Research Council (grant no. 621-2014-5113). Q.Z. is supported by National Natural Science Foundation of China (grant nos. 31722050 and 31671319), the Fundamental Research Funds for the Central Universities (grant no. 2018XZZX002-04) and start-up funds from Zhejiang University. The computational analyses were performed on CUBE cluster from Department of Computational System Biology of University of Vienna and Vienna Scientific Cluster.

Author contributions

Q.Z. and M.I. conceived the project. L.X., Q.Z., G.A., V.P., Y.D., S.F., G.Z., M.B. and S. P. performed the analyses. Q.Z., L.X., A.S., L.C. and M.I. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

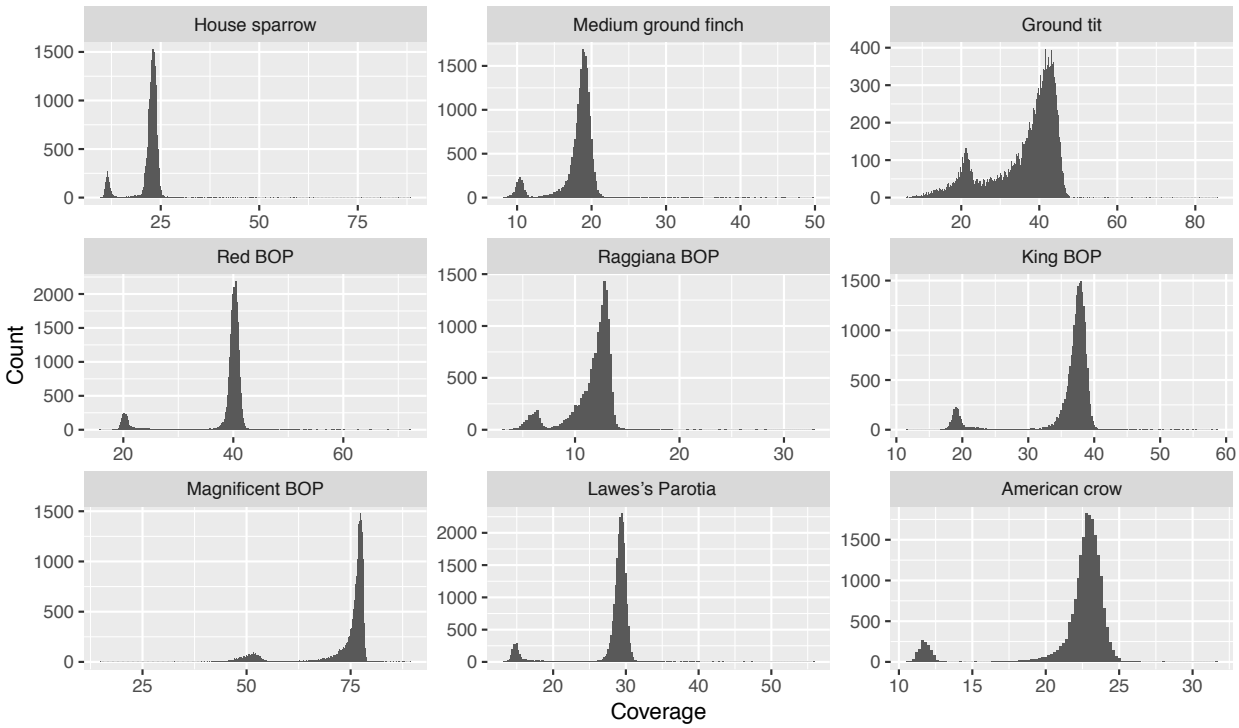
Supplementary information is available for this paper at https://doi.org/10.1038/s41559-019-0850-1.

Reprints and permissions information is available at www.nature.com/reprints.

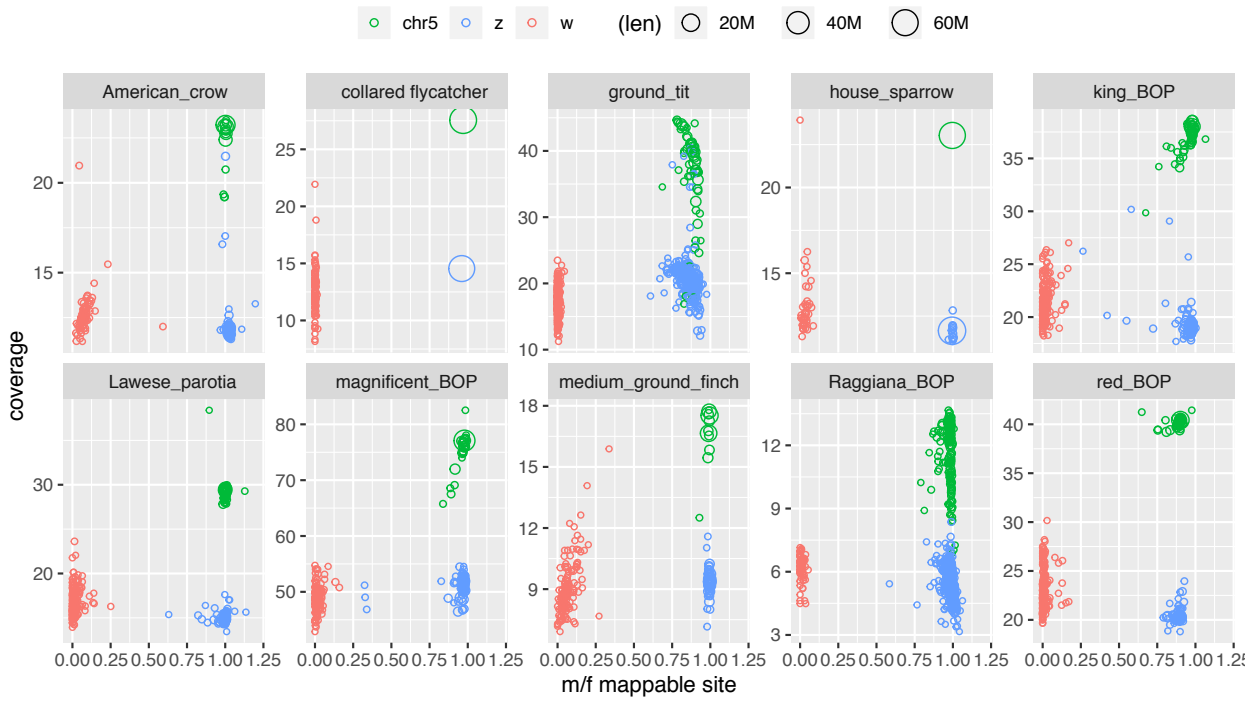
Correspondence and requests for materials should be addressed to M.I. or Q.Z.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

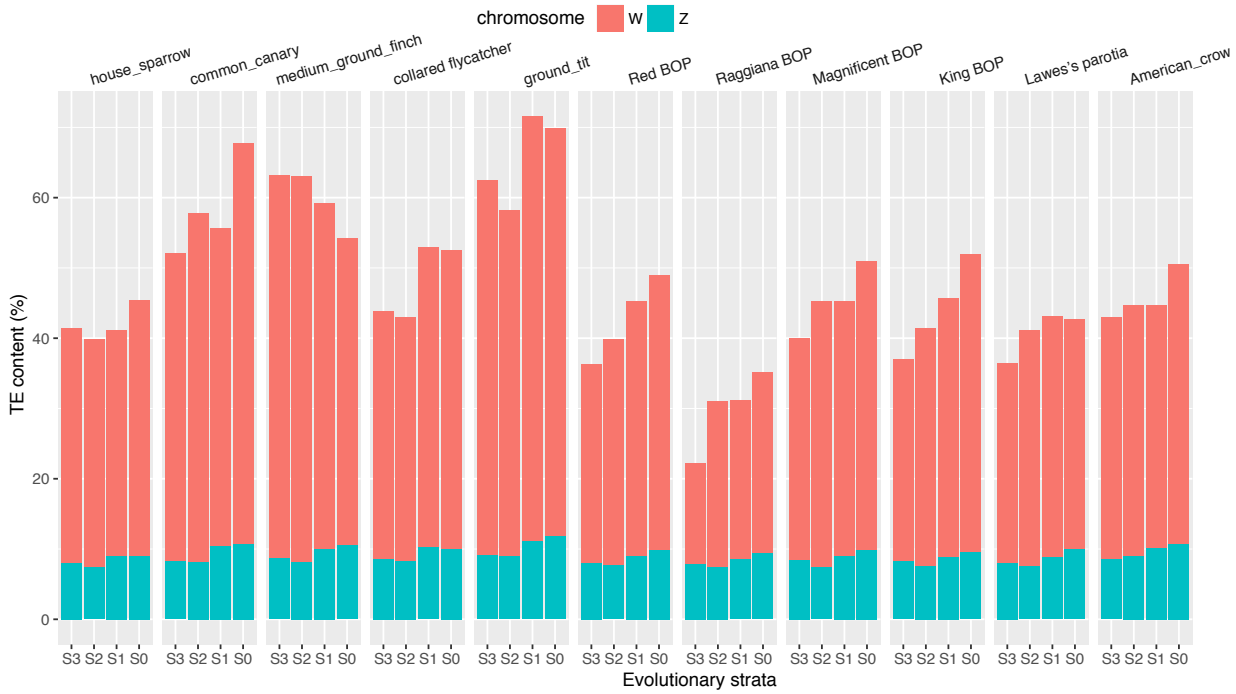
Supplementary Material



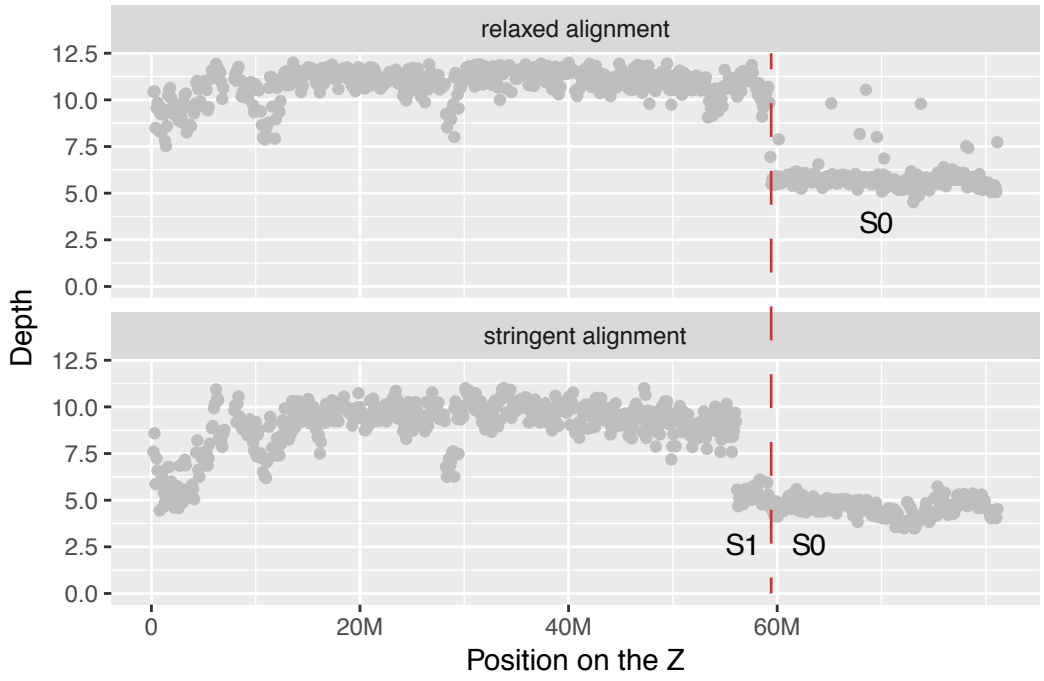
Supplementary Fig. 1 Identifying candidate sex-linked scaffolds.
The sequencing coverage was calculated for every 50-kb non-overlapping window along each scaffold. Scaffolds that are shorter than 5 kb or have less than 60% of the length covered by reads were discarded. The scaffolds showing half the coverage are expected to be either Z- or W-linked.



Supplementary Fig. 2 Verification of W-linked sequences.
Five studied species have sequencing data of both sexes for verifying W-linked sequences. The numbers of nucleotide sites mapped by male versus female genomic sequencing reads were compared. The W-linked sequences are expected to be mapped by very few male (ZZ) reads, i.e., approaching 0 at the x-axis. Both W-linked (red) and Z-linked (blue) sequences have sequencing coverage (y-axis) about half of that of autosomes (green, represented by chromosome 5). The sizes of circles represent the length of scaffolds or chromosomes. A similar plot for medium ground finch is shown in Fig. 1.

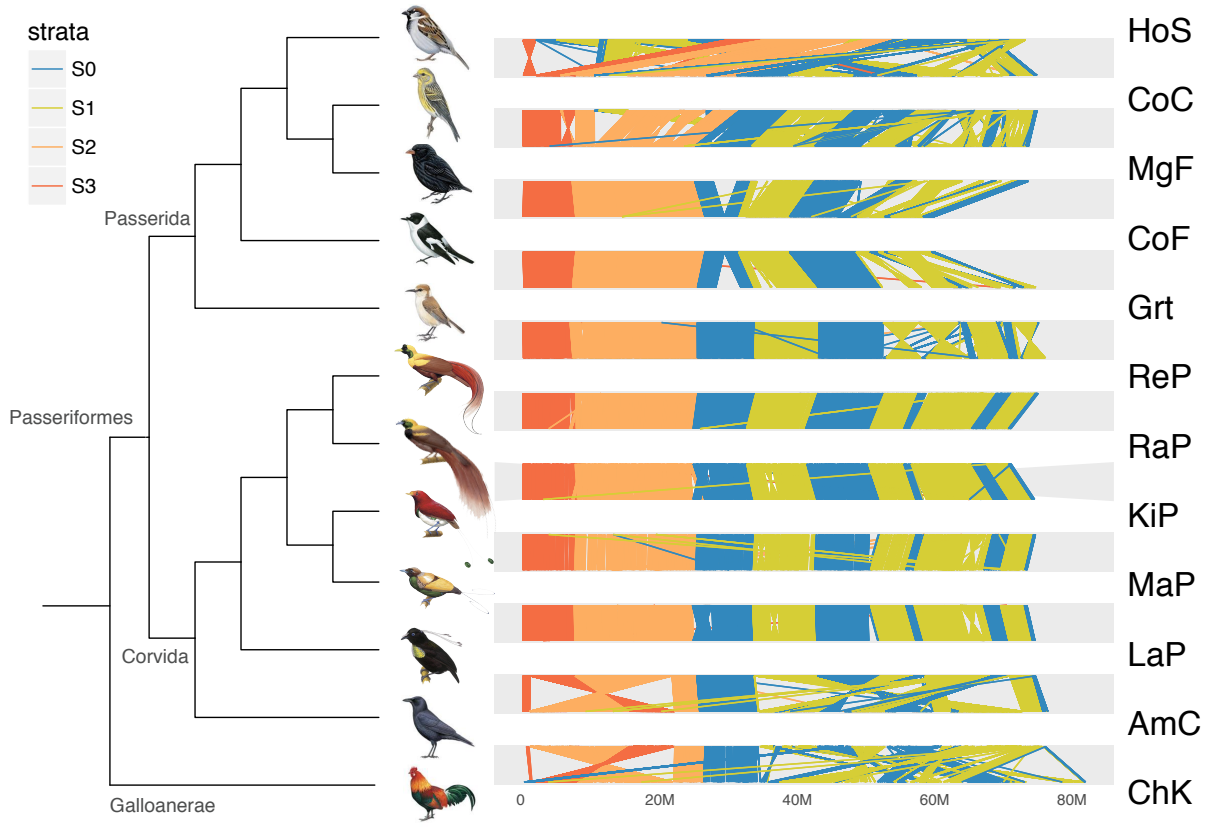


Supplementary Fig. 3 TE content of the Z and W chromosomes.
The TEs include LINEs, SINEs, LTRs and DNA transposons. TE content tends to be higher in the older evolutionary strata on both Z (cyan) and W (red) chromosomes. TE content is much higher on the W than the Z chromosomes.



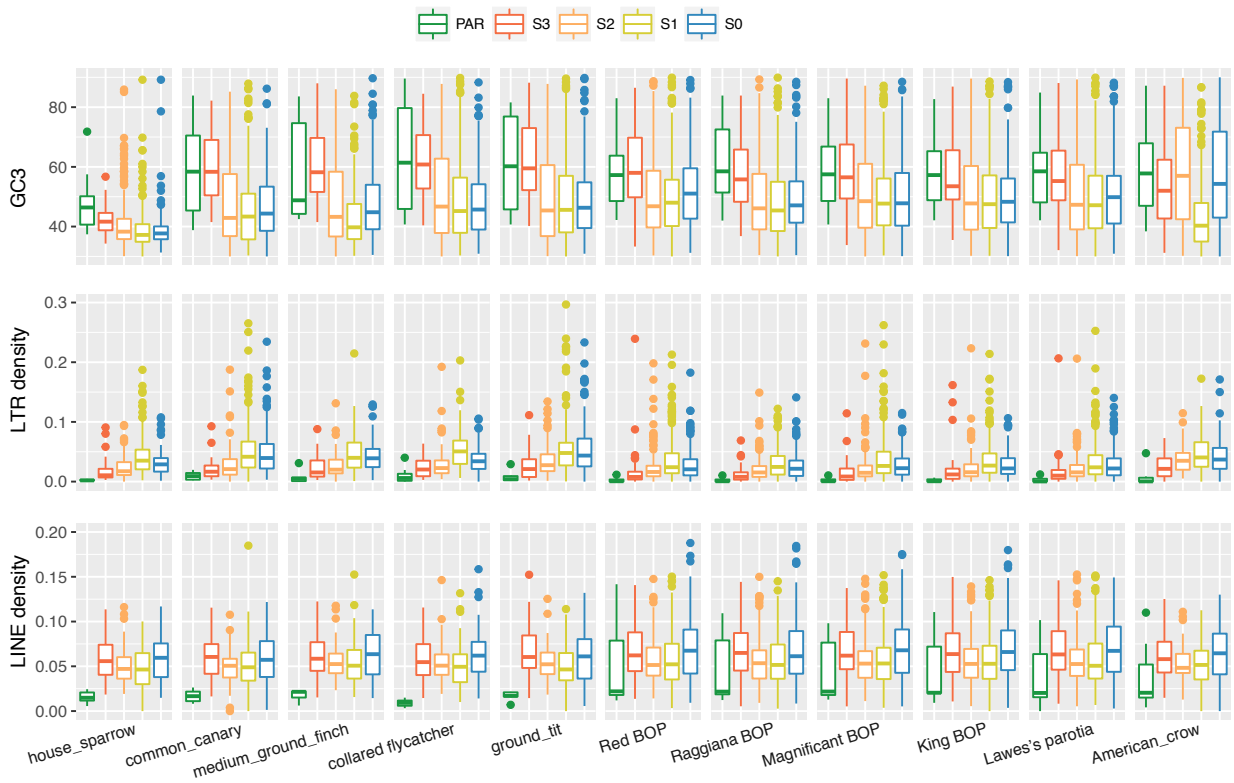
Supplementary Fig. 4 Demarcation of the boundary of emu S1/S0.

We first defined the evolutionary strata of emu by mapping the female reads to its reference genome. Sequencing depth was calculated for every 50k non-overlapping sliding windows and shown here. For the relaxed alignment, bwa mem was used (default parameters) while for the stringent alignment bwa map was used (-o 1 –e 50 –m 100000 –l 15 –k 0) with only one mismatch allowed (bwa sampe –a 900 –n 1 –N 0 –o 10000). A small region of S1 was retrieved using stringent alignment.

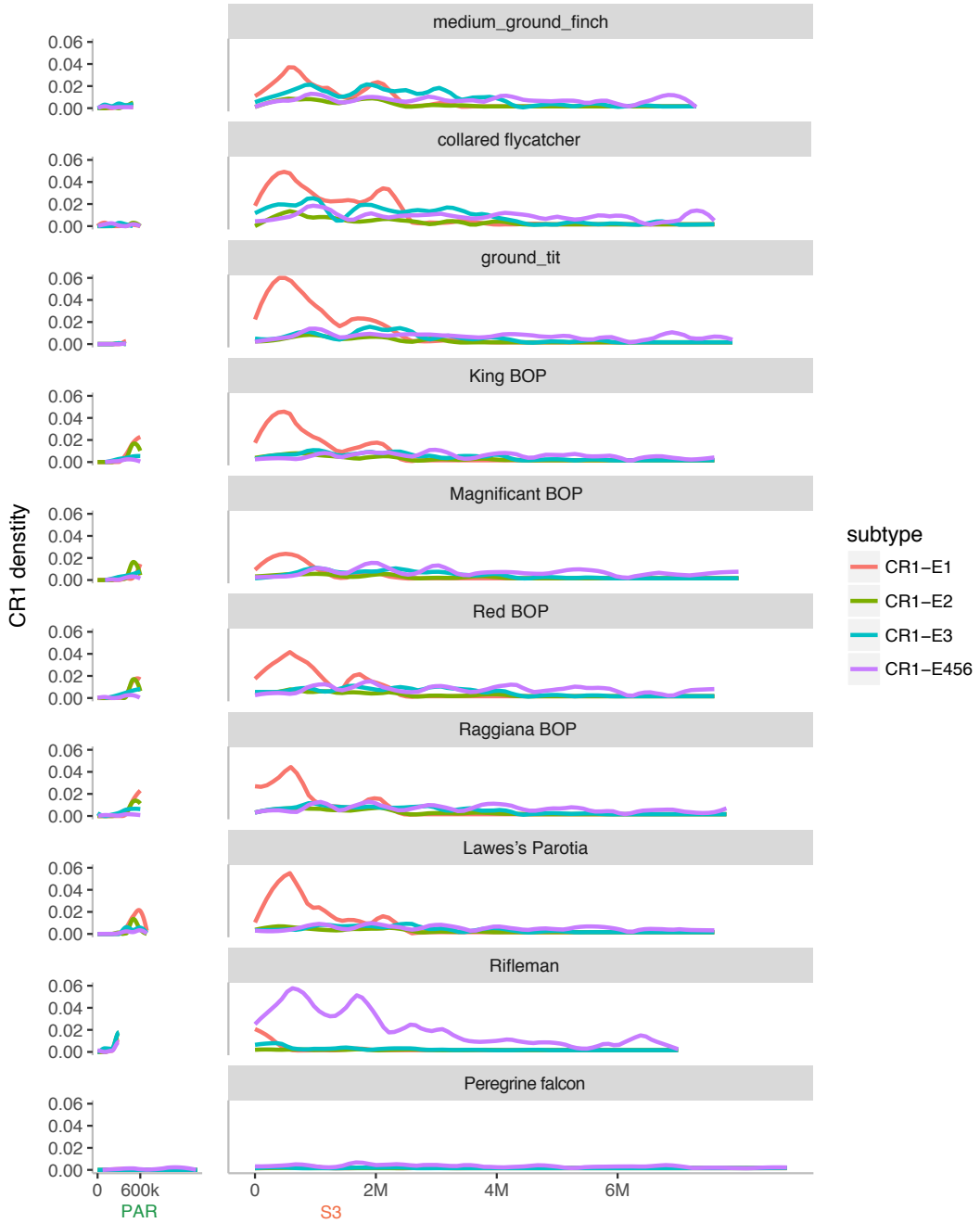


Supplementary Fig. 5 Genome synteny of the studied Z chromosomes.
The pseudo-Z chromosomes of songbirds were constructed using great tit Z chromosome as a reference. Pairwise alignments were performed using nucmer. The colors of lines represent the evolutionary strata of songbirds. There are in general more frequent genomic rearrangements on older strata. All bird illustrations were ordered from https://www.hbw.com/[1].

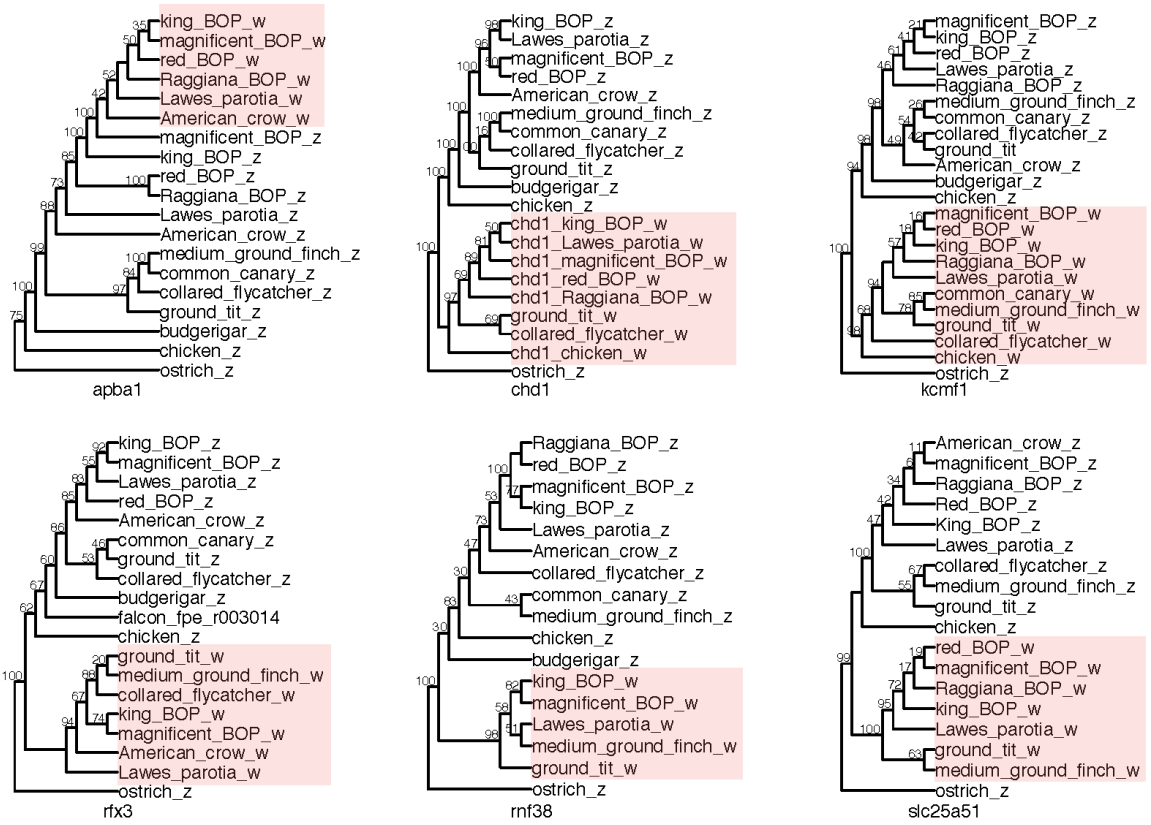
Supplementary Fig. 6 Sequence similarity of the Z and W chromosomes.
The sequence similarity was calculated for every 100-kb windows. The size of circles represents the length of sequence alignments. The smooth lines (light green) were added using the 'loess' method with 'span=0.2'. House sparrow has excessive Z-linked inversions, similar to the reported case of zebra finch. The scaffolds of American crow were ordered according to the synteny of great tit Z chromosome. Different colours represent different evolutionary strata.



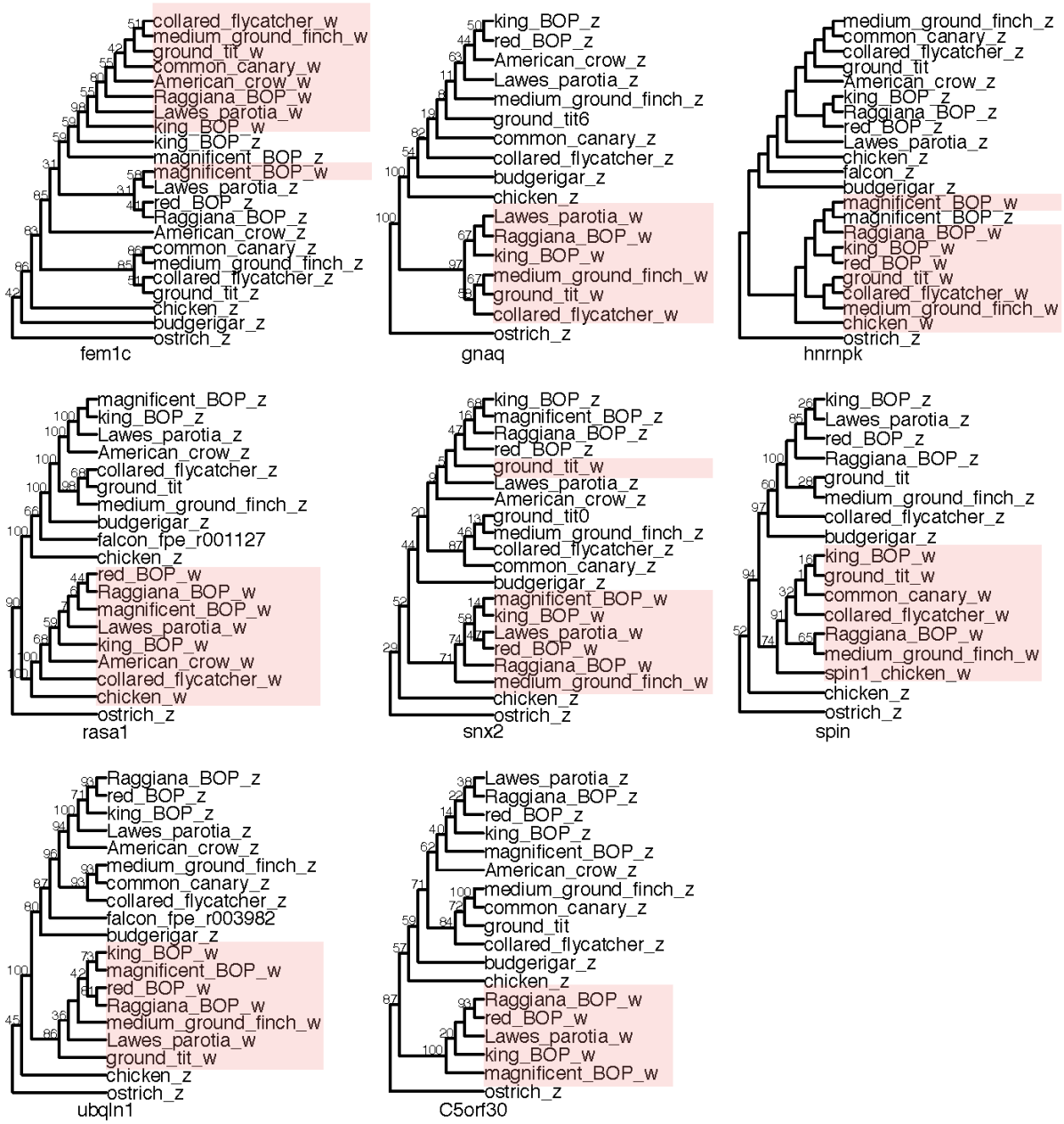
Supplementary Fig. 7 GC3 and TE density of evolutionary strata.
GC3 is the GC content of the third position of codons. The density of the transposable element (TE) groups LINE and LTR were defined as base-pairs of TE sequences per 200-kb non-overlapping window. They change by the age of strata except for LINE density at S3.



Supplementary Fig. 8 Burst of CR1-E1 elements at the boundary of PAR and S3. The density of CR1 is calculated as base-pairs of CR1 elements per 100-kb windows. Only the subtype CR1-E1 is enriched at the PAR/S3 boundary of songbirds. In Passerida (medium ground finch, collared flycatcher and ground tit) the PAR-linked sequences that enrich for CR1-E1 have been deleted. CR1-E4, CR1-E5 and CR1-6 are enriched in the homologous region of rifleman.

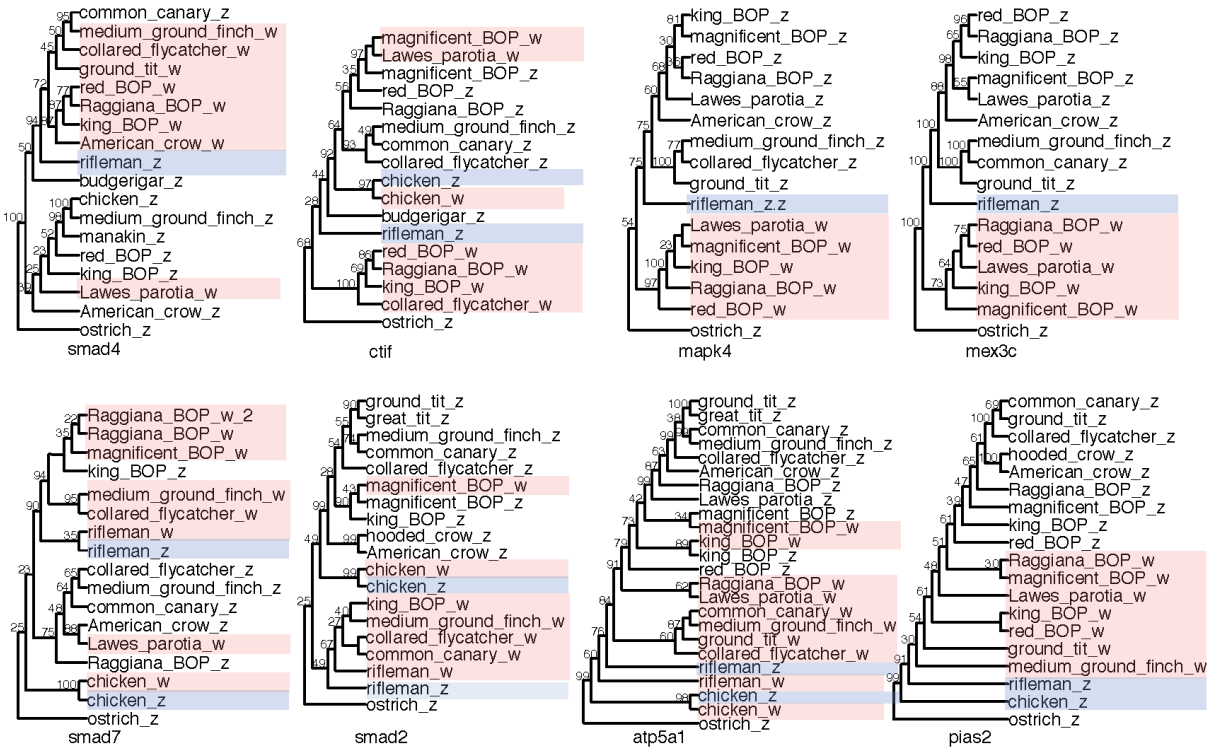


Supplementary Fig. 9 Gene trees for the Z- and W-linked gametologs of S0.
Gene name is shown under each tree. W-linked gametologs are highlighted in red. Genes are grouped by chromosome rather than species.

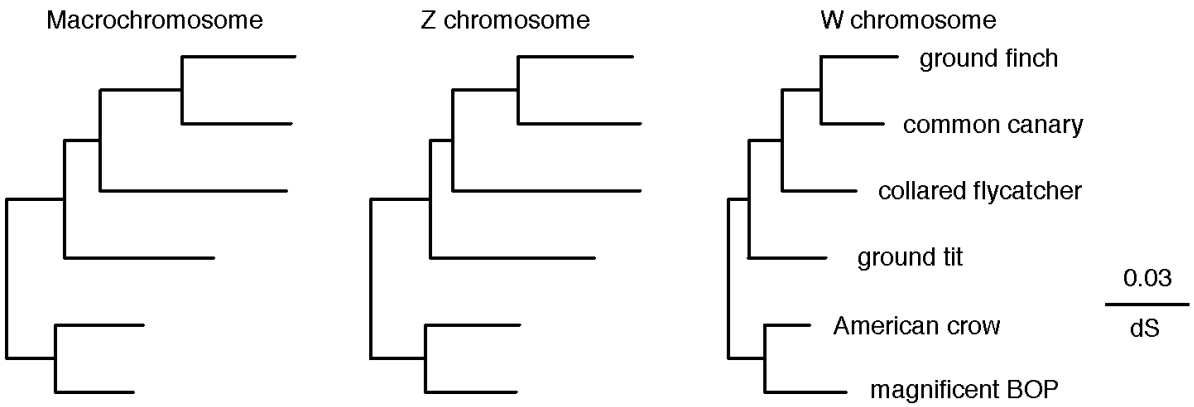


Supplementary Fig. 10 Gene trees for the Z- and W-linked gametologs of S1.
Gene name is shown under each tree. W-linked gametologs are highlighted in red. Genes are grouped by chromosome rather than species.

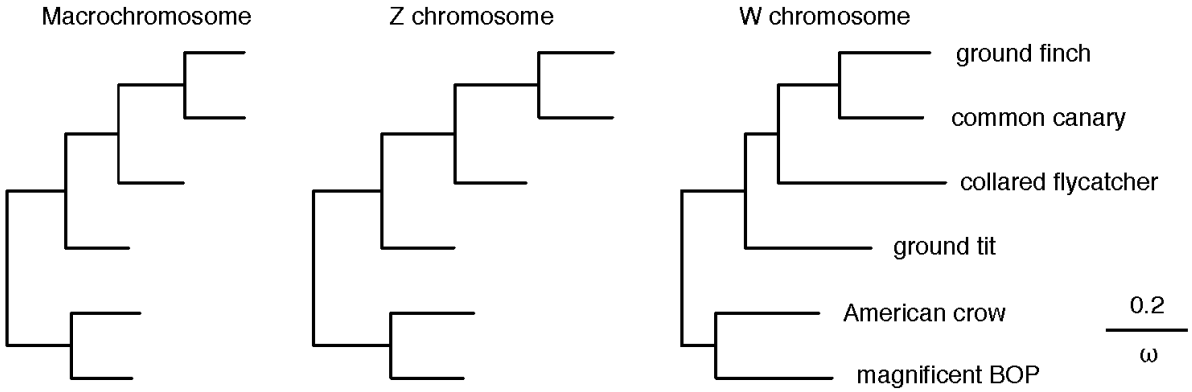
Supplementary Fig. 11 Gene trees for the Z- and W-linked gametologs of S2.
W-linked gametologs are highlighted in red, and chicken Z-linked gametologs are in blue. Chicken W-linked genes are grouped with its Z-linked gametologs instead of the songbirds' homologs, suggesting its independent origin of S2.



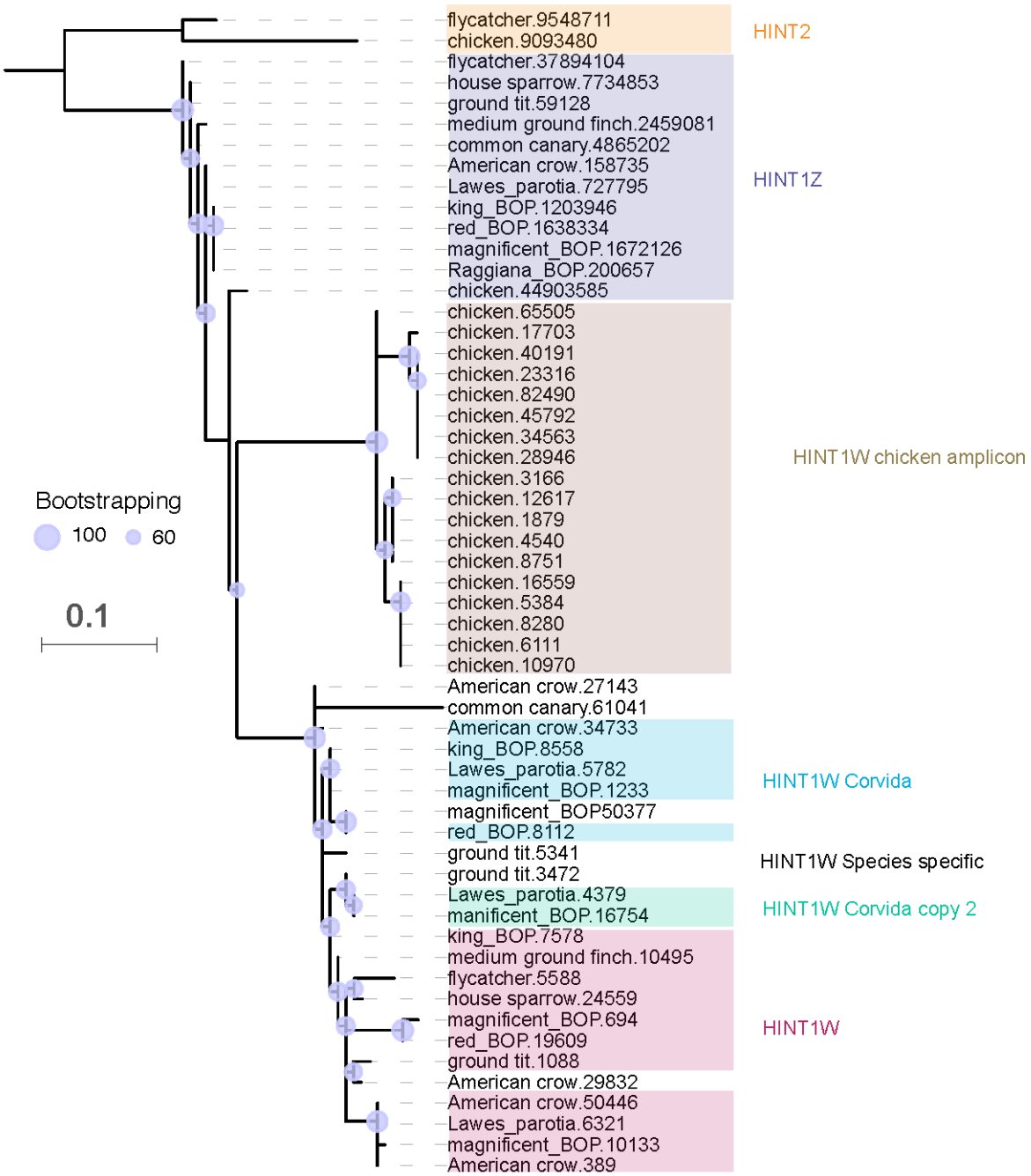
Supplementary Fig. 12 Gene trees for the Z- and W-linked gametologs of S3.
W-linked gametologs are highlighted in red, while Z-linked gametologs of chicken and rifleman are in blue. In most cases rifleman Z-linked genes (*smad4*, *smad7*, *smad2*, *atp5a1* and *pias2*) do not tend to group with songbird orthologs, and the W-linked genes (e.g. smad7) are closely related to their Z-linked gametologs. Another S3 gene c18orf5 is shown in Fig. 2.



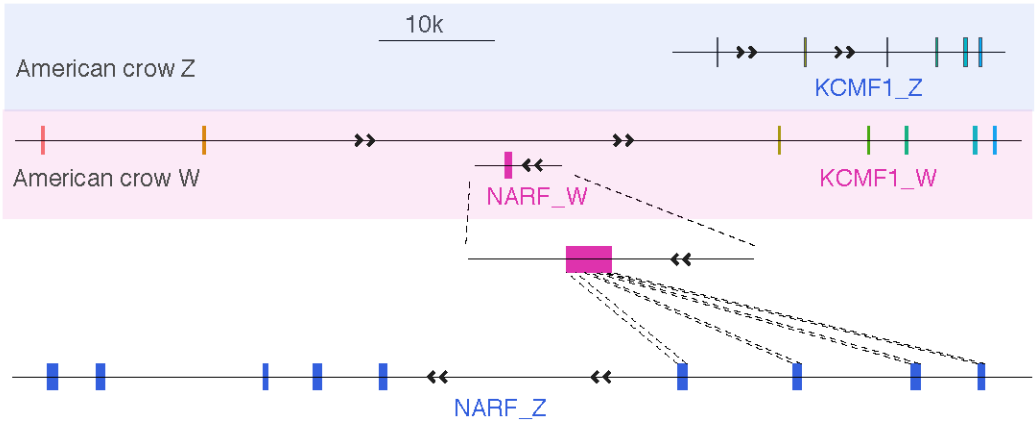
Supplementary Fig. 13 The dS values are larger for the Z and significantly smaller for the W chromosome.
Only one BOP species was selected to avoid short branches within BOP lineages. The chromosome-wise dS values are shown. The dS values (synonymous substitution rates) tend to be larger for the Z-linked genes relative to macrochromosomes though statistically insignificant. The W-linked dS values are significantly smaller, consistent with the prediction of 'male-driven' evolution.



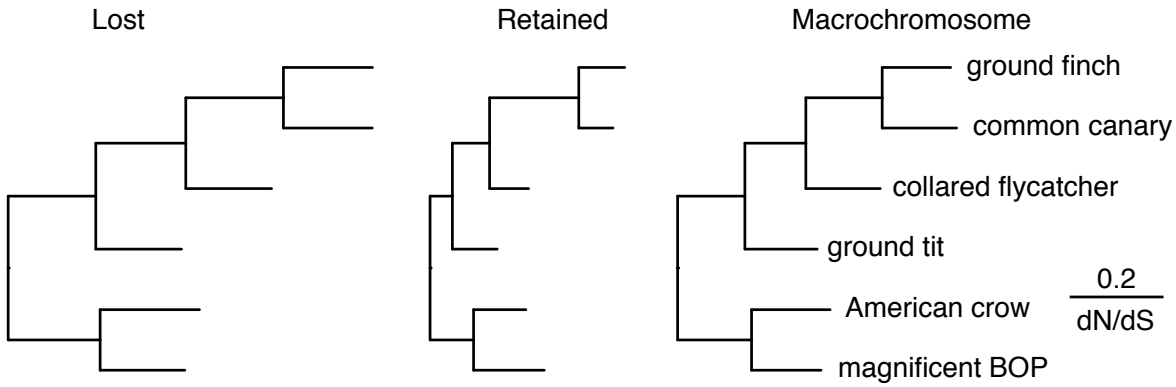
Supplementary Fig. 14 The ω values are larger in both the Z and W chromosomes. Only one BOP species was selected to avoid short branches within BOP lineages. Both sex chromosomes show elevated ω (ratio of nonsynonymous substitution rates to synonymous substitution rates), due to the 'faster-Z' effect and accumulation of deleterious mutations, respectively. The ω values were calculated by dividing chromosome-wise dN by chromosome-wise dS values.



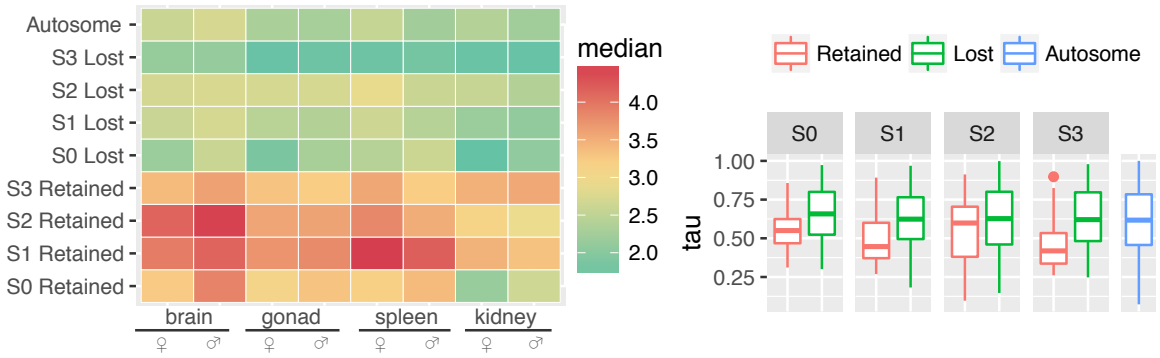
Supplementary Fig. 15 Independent amplification of HINT1 on the W chromosomes. *HINT2* (in orange) is a paralog of *HINT1* which is also on the Z but absent on the W chromosome. There were at least two independent duplication events (branches marked in blue and green) of *HINT1W* in Corvida species.



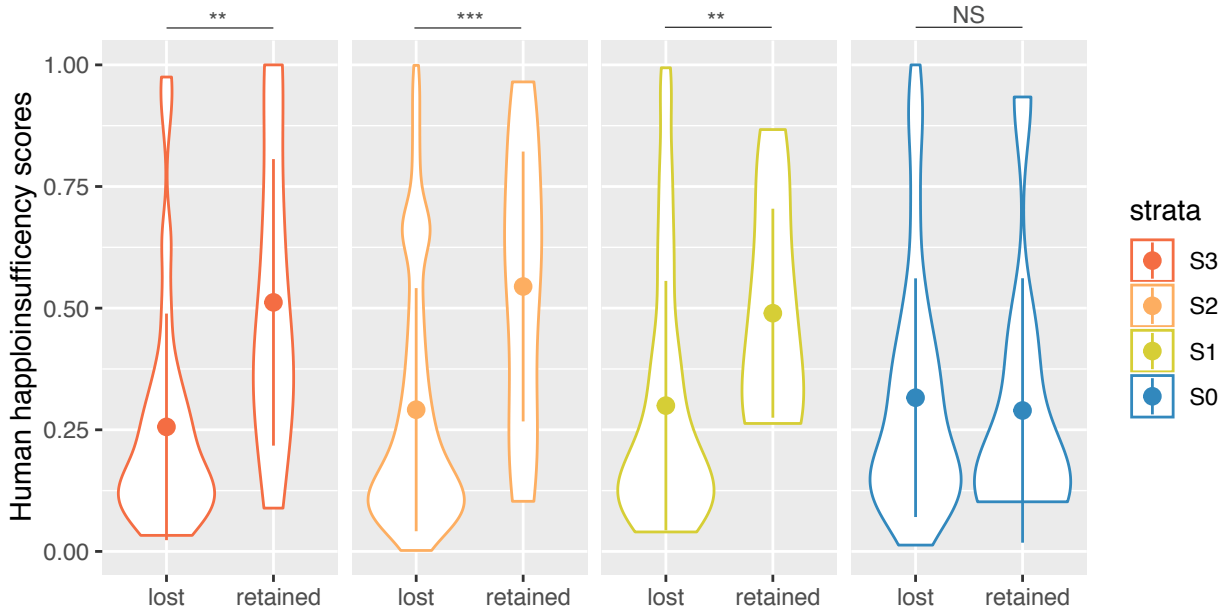
Supplementary Fig. 16 A duplicate gene of NARF on the W chromosome of American crow.
The duplicated gene has one exon, corresponding to the first four exons and a part of fifth exon of Z-linked *NARF*. It is likely to be produced by retroposition. The retrogene was inserted into the second intron of W-linked *KCFM1*, causing the intron expansion.



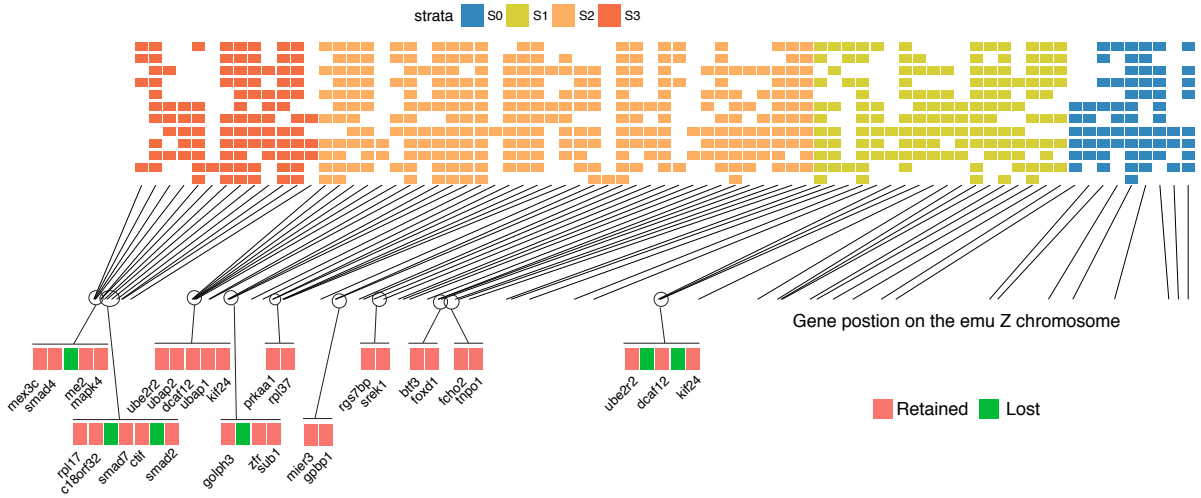
Supplementary Fig. 17 Z-linked homologs of retained W-gametologs have lower dN/dS ratios.
We show the dN/dS ratio (ω, nonsynonymous substitution rate to synonymous substitution rate) as the branch length for Z-linked genes homologous to lost or retained W-linked genes, in comparison to macrochromosomes. The Z-linked genes with retained gametologous W-linked gene show smaller ω irelative to those without a W-linked gametolog.



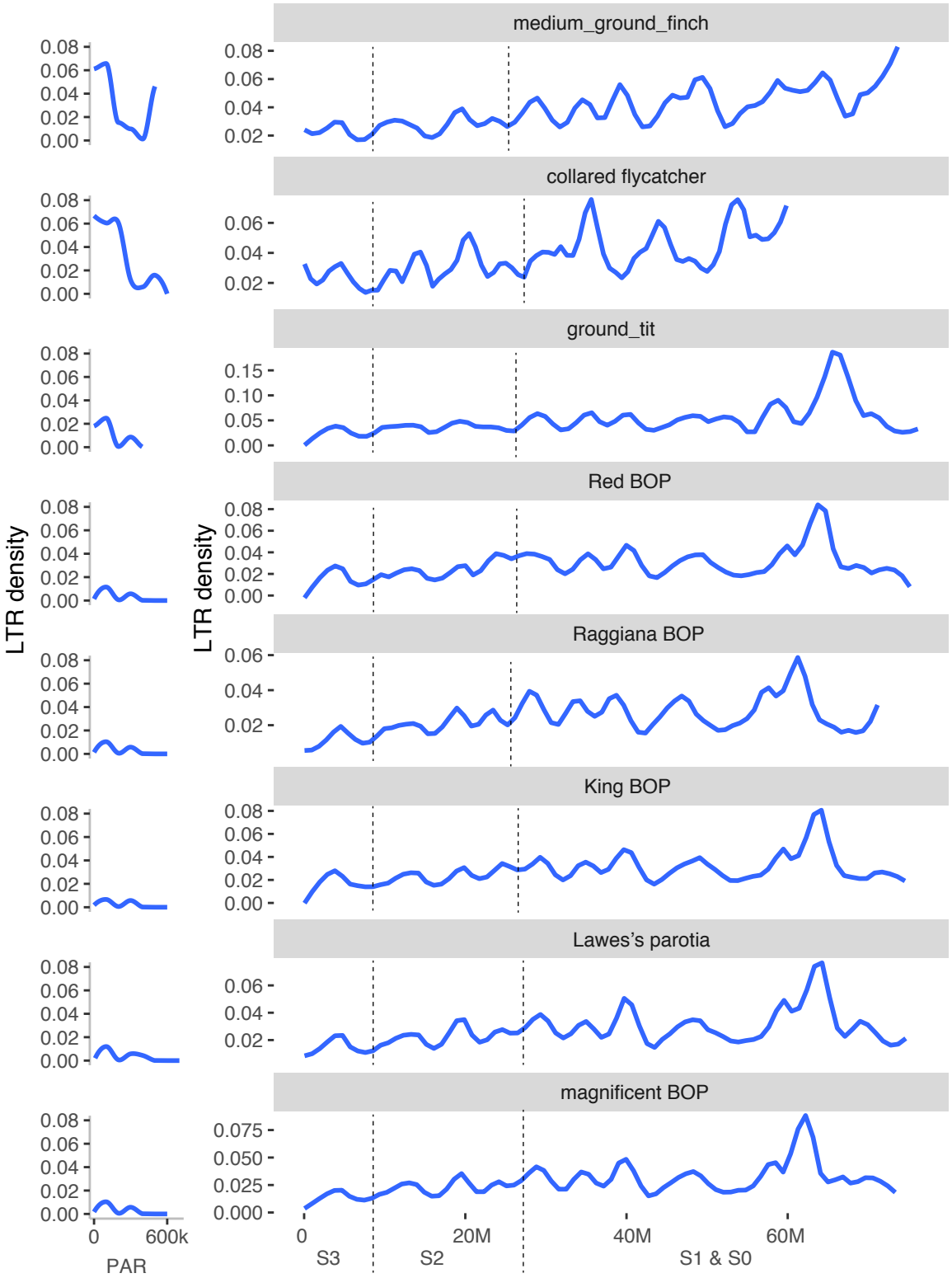
Supplementary Fig. 18 Retained W-gametologs have ancestrally higher and broader expression.
Left panel: the expression levels (measured by TPM) of emu homologous genes of those avian Z-linked genes with (denoted as 'Retained') or without ('Lost') W-gametologs. The log transformed medium expression values of each category are color-coded. Right panel: gene expression tissue specificity (measured as tau) in emu for the homologous avian Z-linked genes. A higher tau value means larger tissue specificity.



Supplementary Fig. 19 Retained W-gametologs have higher haploinsufficiency scores. We used the human orthologs of the Z-linked genes with (denoted as 'retained') and without ('lost') W-linked gametologs to search for the haploinsufficiency scores from Huang et al. (2010)[2]. significant levels (Wilcoxon rank-sum test) of 'retained' vs. 'lost' comparison is denoted with asterisks. '***': P<0.0001, '**': P<0.001.



Supplementary Fig. 20 Position of W-linked genes on emu Z chromosome. The green tiles (Lost) represent genes without a W-linked gametologs, while the red tiles (Retained) represent genes with retained W-linked gametologs. In those 10 pairs/ clusters of genes, genes are next to or close to each other on the Z chromosome.



Supplementary Fig. 21 LTR distribution across different evolution strata

We show e LTR distribution along the Z chromosome of each studied species, divided by their different evolutionary strata. Young strata, e.g. S3, have relatively low density of LTRs.

Reference

1. del Hoyo, J., Elliot, A., Sargatal, J., Christie, D. & de Juana, E. *Handbook of the Birds of the World Alive*, (Lynx Editions, Barcelona, 2019).
2. Huang, N., Lee, I., Marcotte, E.M. & Hurles, M.E. Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet* **6**, e1001154 (2010).

Supplementary table S1-S2, S4, S7-S10 that do not fit in this document can be viewer online at https://www.nature.com/articles/s41559-019-0850-1

Table S3 Sequence percentage of CR1 family in the genomes

Species	**CR1-E1**	**CR1-E2**	**CR1-E3**	**CR1-E4**	**CR1-E5**	**CR1-E6**
Magnificent BOP	0.595365	0.389128	0.695205	0.841479	0.861853	0.334868
Lawes's parotia	1.05873	0.387895	0.633715	0.538542	0.633918	0.300907
Raggiana BOP	1.15605	0.574546	0.755646	0.492042	0.640285	0.380182
King BOP	1.02302	0.498493	0.699736	0.504362	0.466591	0.334132
Red BOP	1.05207	0.4701	0.882334	0.978416	1.10483	0.416072
American crow	1.8464	0.734855	1.08384	0.633796	1.04191	0.509066
Common canary	1.79833	0.586826	1.32054	0.787536	0.551211	0.333394
Ground tit	1.43465	0.513468	0.740818	1.15296	0.893043	0.378259
House sparrow	1.56736	0.51595	1.30397	1.03209	0.846879	0.327969
Medium ground finch	1.09728	0.505809	1.34292	1.08149	0.850727	0.307169
Rifleman	0.591276	0.333366	0.72281	1.52884	1.83421	0.612583

Table S5 Number of W-linked genes on each stratum

Species	Group	S3	S2	S1	S0	Total
house sparrow	Passerida	5	16	7	3	31
common canary	Passerida	6	19	7	3	35
medium ground finch	Passerida	7	23	5	3	38
collared flycatcher	Passerida	7	22	11	6	46
ground tit	Passerida	5	26	12	3	46
Red BOP	Corvida	7	26	12	6	51
Raggiana BOP	Corvida	10	22	13	5	50
Magnificant BOP	Corvida	9	29	14	9	61
King BOP	Corvida	10	29	15	9	63
Lawes's parotia	Corvida	10	26	14	7	57
American crow	Corvida	9	26	10	5	50
chicken	Galloanerae	11	11	5	1	28

Table S6 GO term enrichment for W-linked gametologs

Category	Go term	Go term name	P-value	Benjamini
Molecular function	**3677**	**DNA binding**	**4.51E-05**	**0.00359**
Molecular function	**3700**	**transcription factor activity, sequence-specific DNA binding**	**0.00490409**	**0.179**
Biological process	6351	transcription, DNA-templated	0.0105044	0.936
Cellular component	5634	nucleus	0.0151125	0.645

Chapter 3

Paper II: Evolutionary dynamics of sex chromosomes of paleognathous birds

Luohao Xu[1], Simon Yung Wa Sin[2,3,4], Phil Grayson[2,3], Scott V. Edwards[2,3], Timothy B. Sackton[5]

[1] Department of Molecular Evolution and Development, University of Vienna, Austria
[2] Department of Organismic and Evolutionary Biology, Harvard University, USA
[3] Museum of Comparative Zoology, Harvard University, USA
[4] School of Biological Sciences, The University of Hong Kong, Pok Fu Lam Road, Hong Kong
[5] Informatics Group, Harvard University, USA

Highlights

- The sizes of pseudoautosomal regions (PAR) vary among paleognathous birds
- There is a lack of male-biased gene on the PAR
- The large PAR shows reduced recombination rate and efficacy of selection
- There is partial dosage compensation in paleognathous birds

Summary

Standard models of sex chromosome evolution propose that recombination suppression leads to the degeneration of the heterogametic chromosome, as is seen for the Y chromosome in mammals and the W chromosome in most birds. Unlike other birds, palaeognaths (ratites and tinamous) possess large non-degenerate regions on their sex chromosomes (PARs or pseudoautosomal regions). It remains unclear why these large PARs are retained over more than 100 MY of evolution, and their impact on sex chromosome evolution. To address this puzzle, we analyzed Z chromosome evolution and gene expression across 12 palaeognaths, several of whose genomes have recently been sequenced. We confirm at the genomic level that most palaeognaths retain large PARs. As in other birds, we find that all palaeognaths have incomplete dosage compensation on the regions of the Z chromosome homologous to degenerated portions of the W (differentiated regions or DRs), but we find no evidence for enrichments of male-biased genes in PARs. We find limited evidence for increased evolutionary rates (faster-Z) either across the chromosome or in DRs for most palaeognaths with large PARs, but do recover signals of faster-Z evolution in tinamou species with mostly degenerated

W chromosomes, similar to the pattern seen in neognaths. Unexpectedly, in some species, PAR-linked genes evolve faster on average than genes on autosomes, possibly due to reduced efficacy of selection in palaeognath PARs. Our analysis shows that palaeognath Z chromosomes are atypical at the genomic level, but the evolutionary forces maintaining largely homomorphic sex chromosomes in these species remain elusive.

Keywords
Paleognathae; Sex chromosome; Pseudoautosomal region; Recombination rate; Dosage compensatoin

Evolutionary Dynamics of Sex Chromosomes of Paleognathous Birds

Luohao Xu[1], Simon Yung Wa Sin[2,3,4], Phil Grayson[2,3], Scott V. Edwards [2,3], and Timothy B. Sackton [5,*]

[1]Department of Molecular Evolution and Development, University of Vienna, Austria

[2]Department of Organismic and Evolutionary Biology, Harvard University

[3]Museum of Comparative Zoology, Harvard University

[4]School of Biological Sciences, The University of Hong Kong, Hong Kong

[5]Informatics Group, Division of Science, Harvard University

*Corresponding author: E-mail: tsackton@g.harvard.edu.

Accepted: July 10, 2019

Data deposition: This project has been deposited at NCBI under BioProject PRJNA450676.

Abstract

Standard models of sex chromosome evolution propose that recombination suppression leads to the degeneration of the heterogametic chromosome, as is seen for the Y chromosome in mammals and the W chromosome in most birds. Unlike other birds, paleognaths (ratites and tinamous) possess large nondegenerate regions on their sex chromosomes (PARs or pseudoautosomal regions). It remains unclear why these large PARs are retained over >100 Myr, and how this retention impacts the evolution of sex chromosomes within this system. To address this puzzle, we analyzed Z chromosome evolution and gene expression across 12 paleognaths, several of whose genomes have recently been sequenced. We confirm at the genomic level that most paleognaths retain large PARs. As in other birds, we find that all paleognaths have incomplete dosage compensation on the regions of the Z chromosome homologous to degenerated portions of the W (differentiated regions), but we find no evidence for enrichments of male-biased genes in PARs. We find limited evidence for increased evolutionary rates (faster-Z) either across the chromosome or in differentiated regions for most paleognaths with large PARs, but do recover signals of faster-Z evolution in tinamou species with mostly degenerated W chromosomes, similar to the pattern seen in neognaths. Unexpectedly, in some species, PAR-linked genes evolve faster on average than genes on autosomes, suggested by diverse genomic features to be due to reduced efficacy of selection in paleognath PARs. Our analysis shows that paleognath Z chromosomes are atypical at the genomic level, but the evolutionary forces maintaining largely homomorphic sex chromosomes in these species remain elusive.

Key words: sex chromosomes, genomics, molecular evolution, paleognaths.

Introduction

Sex chromosomes are thought to evolve from autosomes that acquire a sex determination locus (Bull 1983). Subsequent suppression of recombination between the X and Y (or the Z and W) chromosomes leads to the evolutionary degeneration of the sex-limited (Y or W) chromosome (Bergero and Charlesworth 2009; Bachtrog 2013). Theoretical models predict that suppression of recombination will be favored so that the sexually antagonistic alleles that are beneficial in the heterogametic sex can be linked genetically to the sex determination locus (Rice 1987; Ellegren 2011). Recombination suppression leads to the formation of evolutionary strata, which can occur multiple times in the course of sex chromosome evolution (Lahn and Page 1999; Bergero and Charlesworth 2009; Cortez et al. 2014; Zhou et al. 2014; Wright et al. 2016; Xu et al. 2019). Despite differences in their autosomal origins and heterogamety, eutherian mammals and neognathous birds followed similar but independent trajectories of sex chromosome evolution (Graves 2016; Bellott et al. 2017).

Although this model of sex chromosome evolution has a clear theoretical basis, it is inconsistent with empirical patterns in many vertebrate lineages. Henophidian snakes (boas) are thought to have ZW chromosomes that have remained

homomorphic for ~100 Myr (Vicoso, Emerson, et al. 2013), although a recent study suggests a transition from ZW to XY system may have occurred (Gamble et al. 2017). Many lineages in fish and nonavian reptiles also possess homomorphic sex chromosomes, in most cases, because the sex chromosomes appear to be young due to frequent sex chromosome turnover (Bachtrog et al. 2014). In some species of frogs, homomorphic sex chromosomes appear to be maintained by occasional XY recombination in sex-reversed XY females (the "fountain of youth" model), which is possible if recombination suppression is independent of genotype and instead a consequence of phenotypic sex, such that XY females experience normal recombination (Perrin 2009; Dufresnes et al. 2015; Rodrigues et al. 2018).

Paleognathous birds (Paleognathae), which include the paraphyletic and flightless ratites and the monophyletic tinamous, and comprise the sister group to Neognathae (all other extant birds), also retain largely or partially homomorphic sex chromosomes (de Boer 1980; Ansari et al. 1988; Ogawa et al. 1998; Nishida-Umehara et al. 1999; Pigozzi and Solari 1999; Stiglec et al. 2007; Tsuda et al. 2007; Janes et al. 2009; Pigozzi 2011), albeit with some exceptions (Zhou et al. 2014). These species share the same ancestral sex determination locus, *DMRT1*, with all other birds (Bergero and Charlesworth 2009; Yazdi and Ellegren 2014), and do not fit the assumptions of the "fountain of youth" model (viable and fertile ZW males), requiring an alternative explanation for the retention of homomorphic sex chromosomes. Vicoso, Kaiser, et al. (2013), studying the emu, suggested that sexual antagonism is resolved by sex-biased expression without recombination suppression, based on an excess of male-biased gene expression in the pseudoautosomal region. Alternatively, lack of dosage compensation, which in mammals and other species normalizes expression of genes on the hemizygous chromosome between the homogametic and heterogametic sex, could arrest the degeneration of the W chromosome due to selection to maintain dosage-sensitive genes (Adolfsson and Ellegren 2013). Although these hypotheses are compelling, they have only been tested in single-species studies and without high quality genomes. A broader study of paleognathous birds is therefore needed for comprehensive understanding of the unusual evolution of their sex chromosomes.

Degeneration of sex-limited chromosomes (the W or the Y) leads to the homologous chromosome (the Z or the X) becoming hemizygous in the heterogametic sex. Numerous studies have shown that one common consequence of this hemizygosity is that genes on the X or Z chromosome typically evolve faster on average than genes on the autosomes (Charlesworth et al. 1987; Meisel and Connallon 2013).The general pattern of faster-X or faster-Z protein evolution has been observed in many taxa, including *Drosophila* (Charlesworth et al. 1987, 2018; Baines et al. 2008; Avila et al. 2014), birds (Mank et al. 2007; Mank, Nam, et al. 2010), mammals (Torgerson and Singh 2003; Lu and Wu 2005; Kousathanas et al. 2014), and moths (Sackton et al. 2014). One primary explanation for faster-X/Z evolution is that recessive beneficial mutations are immediately exposed to selection in the heterogametic sex, leading to more efficient positive selection (Charlesworth et al. 1987; Vicoso and Charlesworth 2006; Mank, Vicoso, et al. 2010). Alternatively, the degeneration of the Y or W chromosomes results in the reduction of the effective population size of the X or Z chromosomes relative to the autosomes (because there are three X/Z chromosomes for every four autosomes in a diploid population with equal sex ratios). This reduction in the effective population size can increase the rate of fixation of slightly deleterious mutations due to drift (Mank, Nam, et al. 2010; Mank, Vicoso, et al. 2010). In both scenarios, faster evolution of X- or Z-linked genes is expected.

The relative importance of these explanations varies across taxa. In both *Drosophila* and mammals, faster evolutionary rates of X-linked genes seem to be driven by more efficient positive selection for recessive beneficial alleles in males (Connallon 2007; Meisel and Connallon 2013). However, for young XY chromosomes in plants, reduced efficacy of purifying selection seems to be the cause for the faster-X effect (Krasovec et al. 2018). For female-heterogametic taxa, the evidence is also mixed. In Lepidoptera there is evidence that faster-Z evolution is also driven by positive selection (Sackton et al. 2014) or is absent entirely (Rousselle et al. 2016), whereas in birds, increased fixation of slightly deleterious mutations due to reduced N_e is likely a major factor driving faster-Z evolution (Mank, Nam, et al. 2010; Wang et al. 2014; Wright et al. 2015). The nonadaptive effects of faster-Z in birds seem to decrease over time, and the signals of fast-Z effects mostly come from recent nonrecombining regions (Wang et al. 2014).

For many paleognaths, a large proportion of the sex chromosomes retain homology and synteny between the Z and the W; these regions are referred to as pseudoautosomal regions (PARs) because they recombine in both sexes and are functionally not hemizygous in the heterogametic sex. In PARs, no effect of dominance on evolutionary rates is expected, and because the population size of the PAR is not different from that of autosomes, an increase in fixations of weakly deleterious mutations is also not expected. Therefore, neither the positive selection hypothesis nor the genetic drift hypothesis is expected to lead to differential evolutionary rates in the PAR compared with autosomes, although other selective forces such as sexually antagonistic selection may impact evolutionary rates in the PAR (Otto et al. 2011; Charlesworth et al. 2014). Moreover, many paleognaths (mainly tinamous) show intermediate or small PARs, implying multiple evolutionary strata, (Zhou et al. 2014) and providing a good system to study the cause of faster-Z evolution at different time scales.

With numerous new paleognath genomes now available (Zhou et al. 2014; Le Duc et al. 2015; Zhang et al. 2015;

Sackton et al. 2019), a re-evaluation of sex chromosome evolution in paleognaths is warranted. Here, we investigate faster-Z evolution, dosage compensation, and sex-biased expression, to gain a better understanding of the slow evolution of sex chromosomes in ratites. Surprisingly, we did not find evidence for widespread patterns of faster-Z evolution for most paleognaths with large PARs, even when analyzing only differentiated regions (DRs) that are functionally hemizygous in the heterogametic sex. Instead, in a few species, we find limited evidence that PARs tend to evolve faster than autosomes. Indirect evidence from the accumulation of transposable elements and larger introns suggests reduced efficacy of selection in both PARs and DRs, potentially because of lower recombination rates compared with similarly sized autosomes. Based on new and previously published RNA-seq data, we find a strong dosage effect on gene expression, suggesting incomplete dosage compensation as in other birds (Itoh et al. 2010; Adolfsson and Ellegren 2013; Uebbing et al. 2013, 2015), but do not recover a previously reported excess of male-biased expression in the PAR (Vicoso, Kaiser, et al. 2013). Our results suggest that simple models of sex chromosome evolution probably cannot explain the evolutionary history of paleognath sex chromosomes.

Materials and Methods

Identification of the Z Chromosome, PARs, and DRs

The repeat-masked sequence of ostrich Z chromosome (chrZ) (Zhang et al. 2015) was used as a reference to identify the homologous Z-linked scaffolds in recently assembled paleognath genomes (Sackton et al. 2019). We used the nucmer program (v3.0) from MUMmer package (Kurtz et al. 2004) to first align the ostrich Z-linked scaffolds to emu genome; an emu scaffold was defined as Z-linked if >50% of the sequence was aligned. The Z-linked scaffolds of emu were further used as reference to infer the homologous Z-linked sequences in the other paleognaths because of the more continuous assembly of emu genome and closer phylogenetic relationships, and in these cases 60% coverage of alignment was required. During this process, we found that a ~12Mb genomic region of ostrich chrZ (scf347, scf179, scf289, scf79, scf816, and a part of scf9) aligned to chicken autosomes. The two breakpoints can be aligned to a single scaffold of lesser rhea (scaffold_0) (supplementary fig. S1, Supplementary Material online), so we checked whether there could be a misassembly in ostrich by mapping the 10k and 20k mate-pair reads from ostrich to the ostrich assembly. We inspected the read alignments around the breakpoint and confirmed a likely misassembly (supplementary fig. S2, Supplementary Material online). The homologous sequences of this region were subsequently removed from paleognathous Z-linked sequences. When a smaller ostrich scaffold showed discordant orientation and/or order, but its entire sequence was contained within the length of longer scaffolds of other paleognaths (supplementary fig. S1, Supplementary Material online), we manually changed the orientation and/or order of that scaffold for consistency. After correcting the orientations and orders of ostrich scaffolds of chrZ, a second round of nucmer alignment was performed to determine the chromosomal positions for paleognathous Z-linked scaffolds.

One way to infer the boundary between the PAR and DR is to compare the differences in genomic sequencing depth of female DNA. Because the DR does not recombine in females and W-linked DRs will degenerate over time and thus diverge from Z-linked DRs, the depth of sequencing reads from the Z-linked DR is generally expected to be half of that for the PAR or autosomes. This approach was applied to cassowary, whose sequence is derived from a female individual. For emu, female sequencing was available from Vicoso, Kaiser, et al. (2013). To facilitate annotation of the PAR, we generated additional DNA-seq data from a female for each of lesser rhea, Chilean tinamou, and thicket tinamou. Default parameters of BWA (v0.7.9) were used to map DNA reads to the repeat-masked genomes with BWA-MEM algorithm (Li 2013), and mapping depth was calculated by SAMtools (v1.2) (Li et al. 2009). A fixed sliding window of 50 kb was set to calculate average mapping depths along the scaffolds. Any windows containing <5 kb were removed. Along the pseudo-Z chromosome, the genomic coverage of female reads is usually either similar to that of autosomes (PAR) or reduced to half relative to autosomes (DR). We designated the PAR/DR boundary as the position where a half-coverage pattern starts to appear. For North Island brown kiwi, however, this boundary is unclear, likely due to relatively low quality of the genome assembly. For this reason, as well as a lack of genome annotation for this species, we did not include this species in analyses of molecular evolution.

Another independent method for annotation of the PAR is based on differences in gene expression between males and females for PAR- and DR-linked genes. Because global dosage compensation is lacking in birds and <5% of DR-linked genes have homologous W-linked homologs, most DR-linked genes are expected to have higher expression in males. To reduce the effect of transcriptional noise and sex-biased expression, 20-gene windows were used to calculate the mean male-to-female ratios. Increases in male-to-female expression ratios were used to annotate approximate PAR/DR boundaries. This method was applied to little spotted kiwi, Okarito brown kiwi, emu, and Chilean tinamou. Given the small divergence between little spotted kiwi and great spotted kiwi, it is reasonable to infer that the latter should have a similar PAR size. Neither female genomic reads nor RNA-seq reads are available for greater rheas and elegant crested tinamou, so the PAR/DR boundaries of lesser rhea and Chilean tinamou were used to estimate the boundaries, respectively.

Because the DR is not expected to show heterozygosity in females, we verified the DR annotation by identifying SNPs

derived from female sequencing data. To do so, we used GATK (v3.8) pipeline (HaplotypeCaller) following best practices (DePristo et al. 2011). The variants were filtered using parameters "QD < 2.0 || FS > 60.0 || MQRankSum < −12.5 || RedPosRankSum < −8.0 || SOR > 3.0 || MQ < 40.0" and "-window 15 -cluster 2" of the GATK program VariantFiltration. We only retained variants that were heterozygous (allele frequency between 0.2 and 0.8). To calculate the density of female heterozygous sites, the number of variants was counted for every sliding window of 50 kb along Z chromosomes. For little spotted kiwi and Okarito brown kiwi, for which only RNA-seq data were available, we called the variants using a similar GATK pipeline, but instead calculated SNPs densities over exons only.

Comparison of Genomic Features

To estimate GC content of synonymous sites of the third position of codons (GC3s), codonW (http://codonw.sourceforge.net) was used with the option "-gc3s." The exon density was calculated by dividing the total length of an exon over a fixed 50 kb windows by the window size. Similarly, we summed the lengths of transposable elements (TEs, including LINEs, SINEs, LTRs, and DNA transposons) based on RepeatMasker outputs (A. Kapusta and A. Suh personal communication) to calculate density for 50 kb windows. Intron sizes were calculated from gene annotations (GFF file) using a custom script. Codon usage bias was quantified by the effective number of codons (ENC) using ENCprime (Novembre 2002). We extracted the intronic sequence of each gene for ENCprime to estimate background nucleotide frequency to further reduce the effect of local GC content on codon usage estimates. Wilcoxon sum rank test were used to assess statistical significance.

Divergence Analyses

Estimates of synonymous and nonsynonymous substitutions per site were extracted from PAML (Yang 2007) outputs generated by free-ratio branch models, based on previously produced alignments (Sackton et al. 2019). For a given chromosome, the overall synonymous substitution rate (dS) was calculated as the ratio of the number of synonymous substitutions to the number of synonymous sites over the entire chromosome. Outliers (genes showing >1,500 substitutions) were removed prior to calculations. Similarly, the chromosome-wide dN was calculated using the numbers of nonsynonymous substitutions and sites over the entire chromosome (this is effectively a length-weighted average of individual gene values). The dN/dS values (ω) were calculated by the ratios of dN to dS values. Confidence intervals for dN, dS, and dN/dS were estimated using the R package "boot" with 1,000 replicates of bootstrapping. P values were calculated by taking 1,000 permutation tests.

Gene Expression Analyses

Three biological replicates of samples from emu brains, gonads, and spleens of both adult sexes were collected by Daniel Janes from Songline Emu farm (specimen numbers: Museum of Comparative Zoology, Harvard University Cryo 6597-6608). For Chilean tinamou, RNA samples were collected from brains and gonads of both sexes of adults with one biological replicate (raw data from Sackton et al. 2019, but reanalyzed here). RNA-seq reads for both sexes of ostrich brain and liver (Adolfsson and Ellegren 2013), emu embryonic brains of two stages (Vicoso, Kaiser, et al. 2013), and blood of little spotted kiwi and Okarito brown kiwi (Ramstad et al. 2016) were downloaded from NCBI SRA.

For the newly generated samples (emu brains, gonads, and spleens), RNA extraction was performed using RNeasy Plus Mini kit (Qiagen). The quality of the total RNA was assessed using the RNA Nano kit (Agilent). Poly-A selection was conducted on the total RNA using PrepX PolyA mRNA Isolation Kit (Takara). The mRNA was assessed using the RNA Pico kit (Agilent) and used to make transcriptome libraries using the PrepX RNA-Seq for Illumina Library Kit (Takara). HS DNA kit (Agilent) was used to assess the library quality. The libraries were quantified by performing qPCR (KAPA library quantification kit) and then sequenced on a NextSeq instrument (High Output 150 kit, PE 75 bp reads). Each library was sequenced to a depth of ~30M reads. The quality of the RNA-seq data was assessed using FastQC. Error correction was performed using Rcorrector; unfixable reads were removed. Adapters were removed using TrimGalore!. Reads of rRNAs were removed by mapping to the Silva rRNA database.

We used RSEM (v1.2.22) (Li and Dewey 2011) to quantify the gene expression levels. RSEM implemented bowtie2 (v2.2.6) to map the RNA-seq raw reads to transcripts (based on a GTF file for each species). Default parameters were used for bowtie2 mapping and expression quantification in RSEM. Both the reference genomes and annotations are from (Sackton et al. 2019). All reference genomes except the cassowary are derived from male individuals. TPM (Transcripts Per Million) on the gene level were used to represent the normalized expression. The expected reads counts rounded from RSEM outputs were used as inputs for DESeq2 (Love et al. 2014) for differential expression analysis between sexes. We used a 5% FDR cutoff to define sex-biased genes.

Results

Most Paleognaths Have Large PARs

To identify Z-linked scaffolds from paleognath genomes, we used nucmer (Kurtz et al. 2004) to first align the published ostrich Z chromosome (Zhang et al. 2015) to assembled emu scaffolds (Sackton et al. 2019), and then aligned additional paleognaths (fig. 1) to emu. We then ordered and oriented putatively Z-linked scaffolds in nonostrich assemblies into

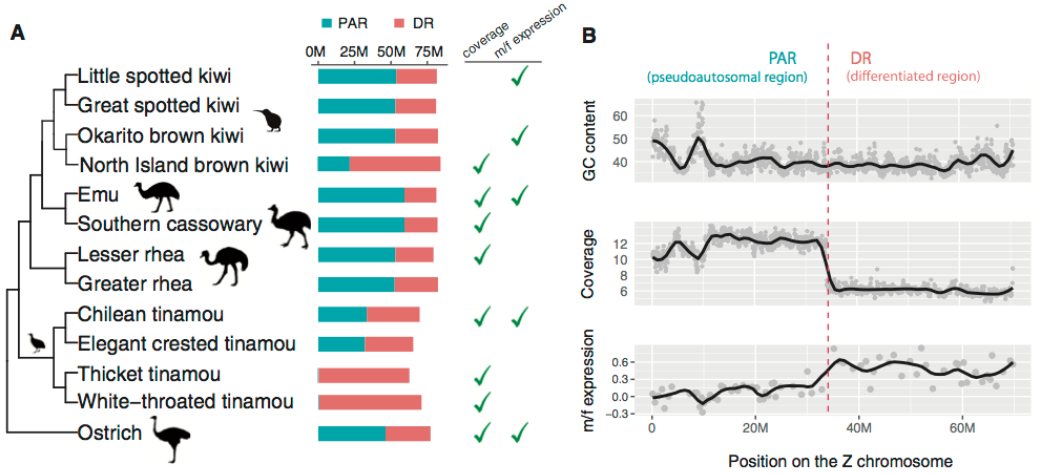


FIG. 1.—Overview of PAR/DR annotation. (*A*) The phylogeny of Palaeognathae based on Sackton et al. (2019) and Cloutier et al. (2019). The sizes of the PARs (pseudoautosomal regions) and DRs (differentiated regions) are indicated by the bars in cyan and tomato. The check marks indicate whether the PAR/DR boundaries were annotated by female read coverage and/or male-to-female expression ratios; species with no checks were annotated by homology to closest relatives. (*B*) An example of PAR/DR annotation for Chilean tinamou. In the panels of GC content and coverage depth, each dot represents a 50k window. In the panel of m/f expression, each dot represents log2-transformed mean m/f expression ratio of ten consecutive genes.

pseudochromosomes using the ostrich Z chromosome as a reference (supplementary fig. S1, Supplementary Material online). Consistent with earlier work (Chapus and Edwards 2009), visualization of pseudochromosome alignments (supplementary fig. S1, Supplementary Material online) showed little evidence for interchromosomal translocations, as expected based on the high degree of synteny across birds (Ellegren 2010); an apparent 12 Mb autosomal translocation onto the ostrich Z chromosome is a likely misassembly (supplementary fig. S2, Supplementary Material online). This assembly error has been independently spotted using a new linkage map of ostrich (Yazdi and Ellegren 2018).

We next annotated the PAR and DR of the Z chromosome in each species. In the DR, reads arising from the W in females will not map to the homologous region of the Z (due to sequence divergence associated with W chromosome degeneration), whereas in the PAR, reads from both the Z and the W will map to the Z chromosome. Thus, we expect coverage of sequencing reads mapped to the Z chromosome in the DR to be ½ that of the autosomes or PAR in females, logically similar to the approach used to annotate Y and W chromosomes in other species (Chen et al. 2012; Carvalho and Clark 2013; Tomaszkiewicz et al. 2017). We also annotated PAR/DR boundaries using gene expression data. If we assume that global dosage compensation is absent, as it is in all other birds studied to date (Graves 2014), M/F expression ratios of genes on the Z with degenerated W-linked gametologs in the DR should be larger than that of genes with intact W-linked gametologs in the PAR. There are other processes that can generate a reduced M/F expression ratio in the absence of W chromosome degeneration (e.g., sex-biased expression) or a "PAR-like" M/F expression ratio close to 1 even when the W chromosome is degenerated, such as gene-specific dosage compensation (Naurin et al. 2012) or incomplete degradation of W-linked gametolog. Although these likely account for local departures in expression patterns for individual genes, they are unlikely to explain chromosomal shifts in the means of expression in sliding windows. Nonetheless, we only use expression data when no other method for annotating PAR/DR boundaries is available.

For seven species with DNA (re)sequencing data from females, either newly reported in this study (lesser rhea [*Rhea pennata*], thicket tinamou [*Crypturellus cinnamomeus*], and Chilean tinamou [*Nothoprocta perdicaria*]) or previously published (emu [*Dromaius novaehollandiae*], ostrich [*Struthio camelus*], cassowary [*Casuarius casuarius*], North Island brown kiwi [*Apteryx mantelli*], and white-throated tinamou [*Tinamus guttatus*]), we annotated PAR and DR regions using genomic coverage alone (fig. 1*B* and supplementary fig. S3, Supplementary Material online), or in the case of the white-throated tinamou used previously published coverage-based annotations (Zhou et al. 2014). Although some variation in coverage attributable to differences in GC content is apparent, the coverage reduction in the DR region is robust (fig. 1*B*). We used expression ratios alone to demarcate the DR/PAR boundaries in little spotted kiwi (*Apteryx owenii*) and

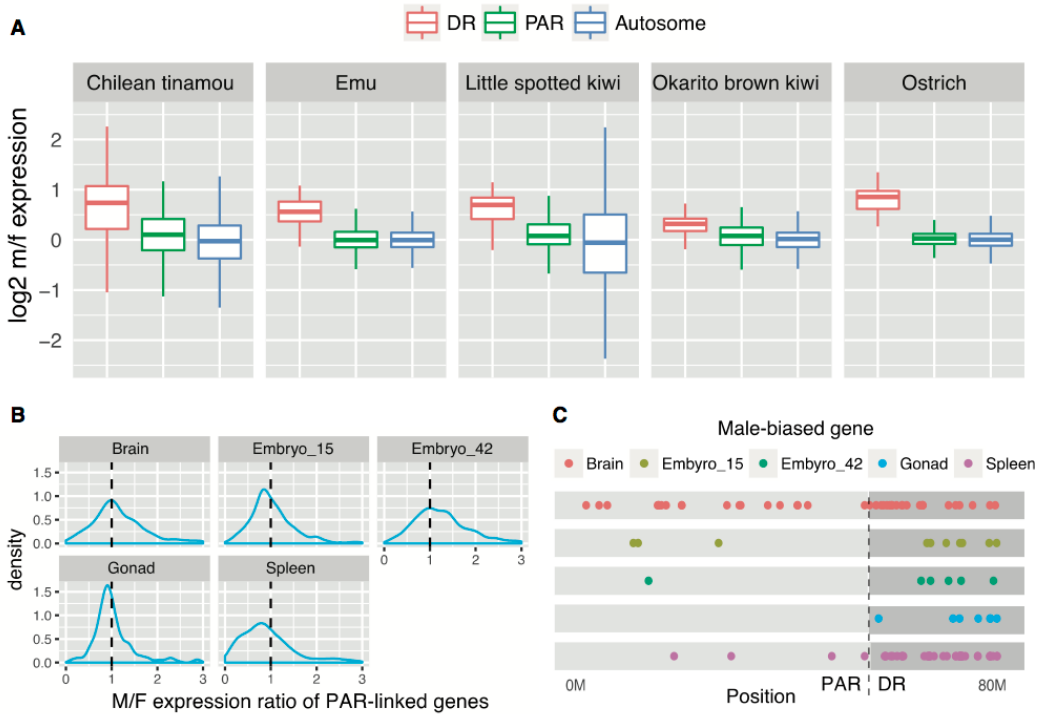


Fig. 2.—Transcriptomic analyses for five paleognathous species. (*A*) Incomplete dosage compensation in emu, kiwi, and tinamou. For each species, only one sample is shown: Chilean tinamou (brain), emu (gonad), ostrich (brain), and both kiwis have only blood samples. Log2 m/f expression ratios of DR-linked are >0 but <1. (*B*) No excess of male expression levels of PAR-linked genes in most emu tissues, despite slight male-biased expression for 42-day embryo. (*C*) No overrepresentation of male-biased genes in emu PAR. Most Z-linked male-biased genes are located on the DR.

Large PARs Are Associated with Lack of Faster-Z Evolution in Paleognaths

The unusually large PARs and the variation in PAR size make Palaeognathae a unique model to study faster-Z evolution. To test whether Z-linked genes evolve faster than autosomal genes, we computed branch-specific d*N*/d*S* ratios (the ratio of nonsynonymous substitution rate to synonymous substitution rate) using the PAML free-ratio model for protein coding genes (Yang 2007), based on previously published alignments (Sackton et al. 2019). Because macrochromosomes and microchromosomes differ extensively in the rates of evolution in birds (Gossmann et al. 2014; Zhang et al. 2014) (supplementary fig. S6, Supplementary Material online), we include only the macrochromosomes (chr1 to chr10) in our comparison, and further focus on only chromosome 4 (97 Mb in chicken) and chromosome 5 (63 Mb) to match the size of the Z chromosome (75 Mb), unless otherwise stated.

We included 23 neognaths and 12 paleognaths in our analysis. Overall, in neognaths, Z-linked genes, with few exceptions, have a significantly higher d*N*/d*S* ratio than autosomal (chr 4/5) genes, suggesting faster-Z evolution (fig. 3). This result is consistent with a previous study involving 46 neognaths (Wang et al. 2014). We further divided Z-linked genes into those with presumed intact W-linked gametologs (PAR genes) and those with degenerated or lost W-linked gametologs (DR genes) to repeat the analysis, because we only expect faster-Z evolution for DR-linked genes. Surprisingly, we do not see widespread evidence for faster-Z evolution in paleognaths for DR genes: only in cassowary, thicket tinamou and white-throated tinamou do DR genes show accelerated d*N*/d*S* and d*N* relative to autosomes (fig. 4 and supplementary fig. S7, Supplementary Material online). Thicket tinamou and white-throated tinamou possess small PARs typical of neognaths, and faster-Z has also been observed for white-throated tinamou in a previous study (Wang et al. 2014), so faster-DR in these species is expected. The observation of faster-DR evolution in cassowary ($P = 0.009$, two-sided permutation test) suggests that

Okarito kiwi (*Apteryx owenii*) (supplementary fig. S3, Supplementary Material online), which we found to be in similar genomic locations in both species. For three species (greater rhea [*Rhea americana*], elegant crested tinamou [*Eudromia elegans*], and great spotted kiwi [*Apteryx haastii*]) with neither female sequencing data nor expression data, we projected the DR/PAR boundary from a closely related species (lesser rhea, Chilean tinamou, and little spotted kiwi, respectively) using shared annotations and synteny.

An alternate approach to identifying the PAR/DR boundary is to rely on SNP densities in females: since the DR is hemizygous in females, we would expect to observe no heterozygous SNPs in the DR (except for those which arise from mapping of partially degenerated W reads to the Z, which should instead cause an increase in the number of SNPs observed). For most species, SNP data corroborate our PAR/DR boundaries (supplementary fig. S3, Supplementary Material online). The exception is the kiwis, where the polymorphism data are ambiguous and suggest the possibility of a recent expansion of the DR and/or a second PAR (supplementary fig. S3, Supplementary Material online). We note that the kiwi variation data are based on RNA-seq data from several individuals (Ramstad et al. 2016), and thus it is difficult to rule out biases arising from the interaction between sex chromosome degeneration and transcriptional patterns across the Z chromosome. Thus, we suggest caution in interpreting results from kiwi.

Nonetheless, overall our results corroborate prior cytogenetic studies across paleognaths and support a large PAR in all species except the Tinaminae (thicket tinamou and white-throated tinamou), which have small PARs and heteromorphic sex chromosomes. PAR sizes in large-PAR paleognaths range from ~20 Mb (23.5% of Z chromosome in North Island brown kiwi) to 59.3 Mb (73% of Z chromosome, in emu); in contrast, PAR sizes in two of the four tinamous and in typical neognaths rarely exceed ~1 Mb (~1.3% of Z chromosome size) (supplementary table S1, Supplementary Material online).

Genes with Male-Biased Expression Are Not Overrepresented in Paleognath PARs

Several possible explanations for the maintenance of old, homomorphic sex chromosomes are related to gene dosage (Adolfsson and Ellegren 2013; Vicoso, Kaiser, et al. 2013). We analyzed RNA-seq data from males and females from five paleognath species, including newly collected RNA-seq data from three tissues from emu (brain, gonad, and spleen; three biological replicates from each of males and females), as well as previously published RNA-seq data from Chilean tinamou (Sackton et al. 2019), ostrich (Adolfsson and Ellegren 2013), kiwi (Ramstad et al. 2016), and additional embryonic emu samples (Vicoso, Kaiser, et al. 2013). For each species, we calculated expression levels for each gene with RSEM (Li and Dewey 2011), and computed male/female ratios with DESeq2 (Love et al. 2014) to assess the extent of dosage compensation, although we note that this measure does not always reflect retention of ancestral sex chromosome expression levels in the hemizygous sex (Gu and Walters 2017). Consistent with previous studies in birds (Graves 2014), we find no evidence for complete dosage compensation by this measure. Instead, we see evidence for partial compensation with M/F ratios ranging from 1.19 to 1.68 (fig. 2*A*). The extent of dosage compensation seems to vary among species, but not among tissues within species (supplementary fig. S4, Supplementary Material online). Retention of divergent W-linked gametologs could appear consistent with incomplete dosage compensation, if the reads arising from the W-linked copy no longer map to the Z-linked copy and are thus invisible in the absence of a W assembly. However, previous work in birds suggest that only a very small fraction of Z-linked genes in the DR retain W gametologs (Zhou et al. 2014; Xu et al. 2019), making this explanation unlikely to account for the bulk of expression differences between sexes in the DR.

Incomplete dosage compensation poses a challenge for detection of sex-biased genes: higher expression levels of DR-linked genes in males may be due to the incompleteness of dosage compensation rather than sex-biased expression per se. With substantially improved genome assemblies and PAR/DR annotations, as well as data from a greater number of species, we re-evaluated the observation that there is an excess of male-biased genes in the emu PAR (Vicoso, Kaiser, et al. 2013). We find that most emu Z-linked male-biased genes are located on the DR (fig. 2*C*), and when DR genes are excluded, we no longer detect an excess of male-biased genes on the Z chromosome of emu ($P > 0.05$ in all tissues, comparing to autosomes, Fisher's exact test, supplementary table S2, Supplementary Material online and fig. 2*C*). We similarly do not detect an excess of female-biased genes, either on the Z as a whole or in the PAR only ($P > 0.05$ in all tissues, Fisher's exact test, supplementary table S2, Supplementary Material online). For PAR-linked genes, although there was a slight shift of expression toward male-bias in 42-day-old emu embryonic brain (fig. 2*B*), only one gene was differentially expressed in male (fig. 2*C*). This dearth of genes with male-biased expression in the PAR is largely consistent across other paleognaths with large PARs, including Chilean tinamou, ostrich, and little spotted kiwi, with one exception in the Okarito brown kiwi (supplementary fig. S5, Supplementary Material online). Overall, we see little evidence for accumulation of either male- or female-biased genes in paleognath PARs, and suggest that the lack of degeneration of the emu W chromosome and other paleognathous chromosomes is probably not due to resolution of sexual antagonism through acquisition of sex-biased genes.

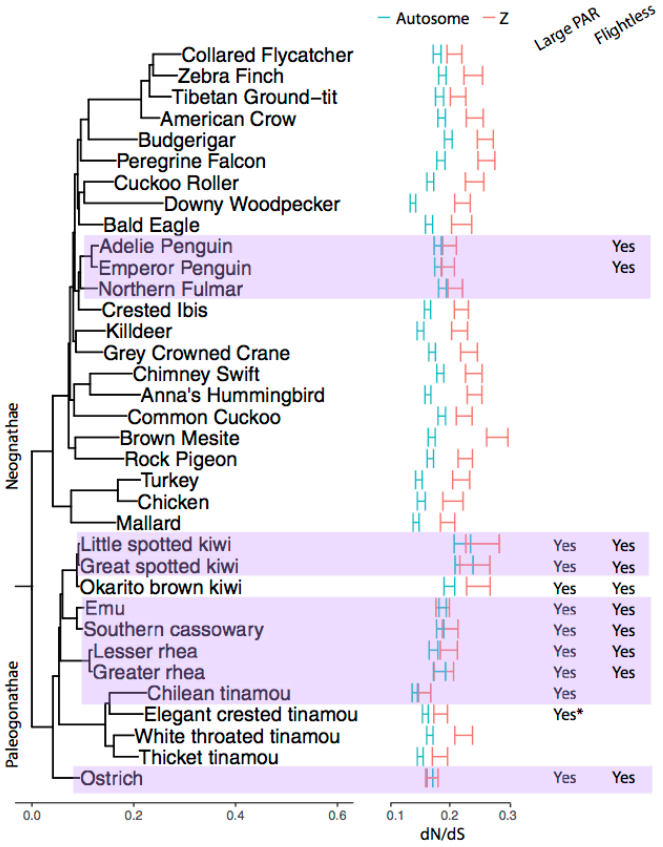


FIG. 3.—A lack of faster-Z evolution in most Paleognaths. Autosomes were represented by chromosome 4 and chromosome 5 (chr4/5) which have similar sizes compared with the Z chromosomes. The confidence intervals of *dN*/*dS* ratios were determined by 1,000 bootstraps. Species without faster-Z effect (permutation test, $P > 0.05$) are highlighted in purple. The asterisk after "Yes" or "No" indicates uncertainty.

faster-DR evolution may not be limited to species with extensive degeneration of the W chromosome (e.g., with small PARs). However, an important caveat is that the cassowary genome (alone among the large-PAR species) was derived from a female individual, which means that some W-linked sequence could have been assembled with the Z chromosome, especially for the region with recent degeneration. This would cause an artefactual increase in apparent rate of divergence.

Unexpectedly, in three tinamous and one kiwi (white-throated tinamou, Chilean tinamou, elegant-crested tinamou, and Okarito brown kiwi), we find evidence that genes in the PAR evolve faster than autosomal genes on chromosomes of similar size (chr4/5), which is not predicted by either the positive selection or genetic drift hypothesis for faster-Z evolution (fig. 4). All those species have higher *dN* in the PAR than autosomes, although not significantly so for the elegant-crested tinamou (fig. 4). Moreover, the faster-PAR effect is not likely to be caused by genes in the newly formed DRs but falsely identified as PARs, because our results are consistent if we remove genes near the inferred PAR/DR boundary (supplementary fig. S8, Supplementary Material online). The faster-PAR in white-throated tinamou is particularly unexpected because previous studies suggest that genes on small PARs evolve slower in birds than non-PAR genes (Smeds et al. 2014). Interestingly, we find the GC content of PAR-linked genes in white-throated tinamou (the only species with both a small PAR and faster-PAR evolution in our analysis) is

FIG. 4.—Relative evolutionary rates of Z-linked and autosomal (chr4/5) genes. Confidence intervals were estimated by 1,000 bootstraps. The label chr4/5 stands for chromosome 4 + chromosome 5, and the median value for chr4/5 is also shown as a dotted line. Asterisks indicate the significant levels of PAR/DR versus chr4/5 comparison (two-sided permutation test), *<0.05, **<0.01, ***<0.001.

significantly biased toward GC (supplementary fig. S9, Supplementary Material online), suggesting GC-biased gene conversion might have contributed to the elevated divergence rate. The small number of PAR-linked genes in white-throated tinamou ($N = 9$), however, suggests some caution in interpreting this trend is warranted.

Evidence for Reduced Efficacy of Selection on the Z Chromosome

The signatures of higher d*N* and d*N*/d*S* we observe in the PARs of tinamous and some other species could be driven by increased fixation of weakly deleterious mutations, if the efficacy of selection is reduced in PARs despite homology with the nondegenerated portion of the W chromosome. One potential marker of the efficacy of selection is the density of transposable elements (TEs), which are thought to increase in frequency when the efficacy of selection is reduced (Rizzon et al. 2002; Lockton et al. 2008). We find that chromosome size, which is inversely correlated with recombination rates in birds (Kawakami et al. 2014), shows a strong positive correlation with TE density (lowest in Okarito brown kiwi, $r = 0.90$; highest in white-throated tinamou, $r = 0.98$) (supplementary fig. S10 and table S3, Supplementary Material online). Extrapolating from autosomal data, we would expect PARs (<50 Mb in all species) to have lower TE density than chr5 (~63 Mb in paleognaths) or chr4 (~89 Mb in paleognaths) if similar evolutionary forces are acting on them to purge TEs. Strikingly, we find that all paleognaths with large PARs harbor significantly higher TE densities on the PAR than autosomes (fig. 5), which suggests reduced purging of TEs on PARs.

Intron size is probably also under selective constraint (Carvalho and Clark 1999), and in birds, smaller introns are likely favored (Zhang and Edwards 2012; Zhang et al. 2014). If this is also the case in paleognaths, an expansion of intron sizes could suggest reduced efficacy of selection. We compared the intron sizes among PARs, DRs, and autosomes across all paleognaths in our study. Like TE densities, intron sizes show strong positive correlation with chromosome size (lowest in Okarito brown kiwi, $r = 0.74$; highest in thicket tinamou, $r = 0.91$) (supplementary fig. S10 and table S3, Supplementary Material online). Except for white-throated tinamou and thicket tinamou, intron sizes of the PARs are larger than those of chr4/5 ($P < 8.8e\text{-}10$, Wilcoxon rank sum test, fig. 4C). The pattern of larger intron sizes in the PARs remains unchanged when all macrochromosomes were included for comparison (supplementary fig. S10, Supplementary Material online). Similar to PARs, DRs also show larger intron sizes relative to chr4/5 ($P < 0.00081$, Wilcoxon rank sum test).

Finally, codon usage bias is often used as proxy for the efficacy of selection and is predicted to be larger when selection is more efficient (Shields et al. 1988). To assess codon usage bias, we estimated ENC values, accounting for local nucleotide composition. ENC is lower when codon bias is stronger, and thus should increase with reduced efficacy of selection. As expected, ENC values showed a strong positive correlation with chromosome sizes (supplementary table S3, Supplementary Material online), and are higher for DR-linked genes in most species (although not rheas, the little spotted kiwi, or the Okarito brown kiwi) (supplementary fig. S11, Supplementary Material online). However, for PAR-linked

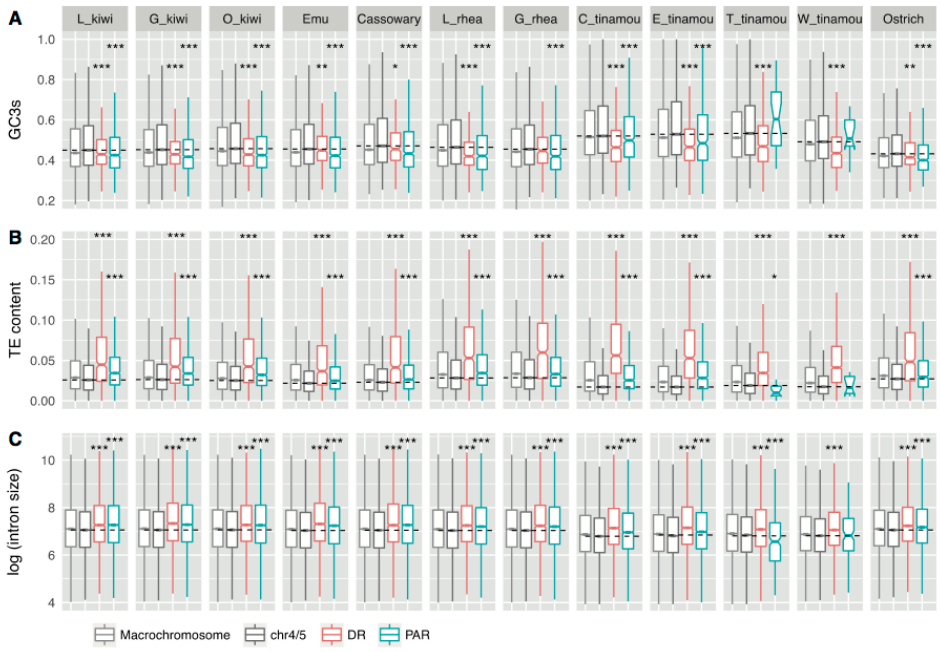


FIG. 5.—The comparison of PAR/DR versus chr4/5 and macrochromosomes of three genomic features. Median values from chr4/5 are shown as a dotted horizontal line. Asterisks indicate the significant levels of PAR/DR versus chr4/5 comparison (Wilcoxon sum rank test), *<0.05, **<0.01, ***<0.001. (*A*) GC content of the synonymous sites. (*B*) TE content, including SINE, LINE, LTR, and DNA element. (*C*) Log-transformed intron size. Abbreviation for species names: L_kiwi, little spotted kiwi; G_kiwi, great spotted kiwi; O_kiwi, Okarito brown kiwi; L_rhea, Lesser rhea; G_rhea, Greater rhea; C_tinamou, Chilean tinamou; E_tinamou, elegant crested tinamou; T_tinamou, thicket tinamou; W_tinamou, white-throated tinamou.

genes, ENC does not suggest widespread reductions in the efficacy of selection: only cassowary and Chilean tinamou exhibited significantly higher ENC values in the PAR, although a trend of higher ENC values can be seen for most species (supplementary fig. S11, Supplementary Material online).

One possible cause of changes in the efficacy of selection in the absence of W chromosome degeneration is a reduction in the recombination rate of the PAR of some species with a large PAR, although a previous study on the collared flycatcher (a neognath species with a very small PAR) showed that the PAR has a high recombination rate (Smeds et al. 2014). Previous work (Bolivar et al. 2016) has shown that recombination rate is strongly positively correlated with GC content of synonymous third positions in codons (GC3s) in birds, so we used GC3s as a proxy for recombination rate in the absent of pedigree or population samples to estimate the rate directly. We find that GC3s are strongly negatively correlated with chromosome size in all paleognaths ($-0.78 \sim -0.91$, P value ≤ 0.0068) except for ostrich ($r = -0.51$, $P = 0.11$) (supplementary fig. S10 and table S3, Supplementary Material online), similar to what was observed in mammals (Romiguier et al. 2010). Recombination rates are also negatively correlated with chromosome sizes in birds (Gossmann et al. 2014; Kawakami et al. 2014) and other organisms (Jensen-Seaman et al. 2004) suggesting that GC3s are at least a plausible proxy for recombination rate. In contrast to the results for collared flycatcher, GC3s of paleognath PARs were significantly lower than those of chr4/5s ($P < 2.23\text{e-}05$, Wilcoxon sum rank test) (fig. 5*A* and supplementary fig. S10, Supplementary Material online), except for white-throated tinamou and thicket tinamou. Inclusion of the other macrochromosome does not change the pattern ($P < 0.0034$). Moreover, distribution of GC3s along the PAR is more homogeneous compared with chr4 or chr5, except for the 5′-prime chromosomal ends (supplementary fig. S12, Supplementary Material online).

Discussion

Old, homomorphic sex chromosomes have long been an evolutionary puzzle because they defy standard theoretical expectations about how sexually antagonistic selection drives recombination suppression of the Y (or W) chromosome and eventual degradation. The Palaeognathae are a classic

example where previous cytogenetic and genomic studies have clearly demonstrated the persistence of largely homomorphic sex chromosomes. Our results extend previous studies, and confirm at the genomic level that all ratites have large, nondegenerate PARs, whereas, in at least some tinamous, degradation of the W chromosome has proceeded, resulting in typically small PARs.

Evolutionary Forces Acting on Sex Chromosomes

Several studies have reported evidence for faster-Z evolution in birds, probably driven largely by increased fixation of weakly deleterious mutations due to reduced N_e of the Z chromosome (Mank, Nam, et al. 2010; Wright et al. 2015). However, these studies have focused on neognaths, with fully differentiated sex chromosomes. Here, we show that paleognath sex chromosomes, which mostly maintain large PARs, do not have consistent evidence for faster-Z evolution, although we confirm the pervasive faster-Z effect in neognaths. Notably, the two species in our data set that presumably share heteromorphic sex chromosomes derived independently from neognaths (white-throated tinamou and thicket tinamou) do show evidence for faster-Z evolution, and in particular faster evolution of DR genes. In contrast, paleognaths with small DR and large PAR do not tend to show evidence for faster-DR, even though hemizygosity effects should be apparent (the exception is cassowary, which may be an artifact due to W-linked sequence assembling as part of the Z).

A previous study on neognaths showed that the increased rate of divergence of the Z is mainly contributed by recent strata, whereas the oldest stratum (S0) does not exhibit the faster-Z effect (Wang et al. 2014). Neognaths and paleognaths share the S0, and, since their divergence, only a small secondary stratum has evolved in paleognaths (Zhou et al. 2014). The absence of a faster-Z effect in paleognath DRs where S0 dominates is therefore largely consistent with the results of the study on the neognath S0. A possible mechanism to explain the lack of faster-Z in the DR is that, in S0, the reduced effective population size (increasing fixation of deleterious mutations) is balanced by the greater efficacy of selection in removing recessive mutations (due to hemizygosity). A recent study on ZW evolution in butterflies suggests a similar model, where purifying selection is acting on the hemizygous DR genes to remove deleterious mutations (Rousselle et al. 2016). Although this model would account for the pattern we observe, it remains unclear why the shared S0 stratum should have a different balance of these forces than the rest of the DR in both neognaths and paleognaths with large DRs. Nonetheless, the evolutionary rates of the DR genes in the older strata are probably the net results of genetic drift and purifying selection against deleterious mutation, with little contribution of positive selection for recessive beneficial mutations.

We also detect evidence for faster evolution of genes in the PAR than for autosomes for three tinamous and one species of kiwi. Because the PAR is functionally homomorphic and recombines with the homologous region of the W chromosome, it is not clear why this effect should be observed in these species. However, a common feature of tinamous and kiwis is that the PARs in some species of these two clades are intermediate or small, for example, the PARs of North Island brown kiwi and most tinamous. This raises at least two possible explanations for the faster-PAR effect in tinamous and kiwis: 1) the differentiation of the sex chromosomes is more rapid compared with other paleognaths, and at least some parts of the PARs may have recently stopped recombining but are undetectable by using the coverage method; or 2) the PARs are still recombining but at lower rate, resulting in weaker efficacy of selection against deleterious mutations. Tinamous are well-known for an increased genome-wide substitution rate compared with other paleognaths (Harshman et al. 2008; Zhang et al. 2014; Zhou et al. 2014; Sackton et al. 2019), but why rates of evolution in the PAR should be so high remains unclear.

Efficacy of Selection and Recombination Rate

Multiple lines of evidence suggest a possible reduction in the efficacy of selection in the PAR across all paleognaths with a large PAR. Specifically, we find both an increase in TE density and an increase in intron size in PARs. In contrast, we do not find clear evidence for a reduction in the degree of codon bias in PARs. However, it is possible that GC-biased gene conversion (Galtier et al. 2018) and/or mutational bias (Szövényi et al. 2017) may also affect the codon bias, which may weaken the correlation between codon usage bias and the strength of natural selection.

It is unclear, however, why the efficacy of selection may be reduced in PARs. One possible cause is that the PARs may recombine at lower rates than autosomes. This is a somewhat unexpected prediction because in most species PARs have higher recombination rates than autosomes (Otto et al. 2011). In birds, direct estimates of recombination rates of the PARs are available in both collared flycatcher and zebra finch, and in both species PARs recombines at much higher rates than most macrochromosomes (Smeds et al. 2014; Singhal et al. 2015). This is probably due to the need for at least one obligate crossover in female meiosis, combined with the small size of the PAR in both collared flycatcher and zebra finch.

In paleognaths where PARs are much larger, direct estimates of recombination rate from pedigree or genetic cross data are not available. Our observation that GC3s are significantly lower in large paleognath PARs than similarly sized autosomes is at least consistent with reduced recombination rates in these species, although the lower GC3 may alternatively be due to AT mutational bias (Lipinska et al. 2017). A recent study on greater rhea shows that the recombination

rate of the PAR does not differ from similarly sized autosomes in females (del Priore and Pigozzi 2017), but this study did not examine males and it cannot exclude the possibilities that the recombination rate in males is lower. A recent study in ostrich, indeed found that the PAR recombines at much lower rate in males than females (Yazdi and Ellegren 2018). If this pattern held true for greater rhea, the sex-average recombination rate of the PAR could potentially be lower relative to similarly sized autosomes. A previous study of emu conducted prior to the availability of an emu genome assembly suggested that the PAR has a higher population recombination rate than autosomes (Janes et al. 2009). However, of 22 loci in that study, seven appear to be incorrectly assigned to the sex chromosomes based on alignment to the emu genome assembly (supplementary table S4, Supplementary Material online), potentially complicating that conclusion. The relatively small size of that study and recently improved resources and refined understanding of recombination rates across chromosome types provide opportunities for a new analysis. Further direct tests of recombination rate on ratite Z chromosomes are needed to resolve these discrepancies.

Sexual Antagonism and Sex Chromosome Degeneration

A major motivation for studying paleognath sex chromosomes is that, unusually, many paleognaths seem to maintain old, homomorphic sex chromosomes. We have shown that previously proposed hypotheses do not seem to fully explain the slow degeneration of paleognath sex chromosomes. RNA-seq expression data from both males and females from multiple species suggest dosage compensation is partial in paleognaths, consistent with what has been seen in neognaths. If the absence of complete dosage compensation is the reason for the arrested sex chromosome degeneration in paleognaths, it is not clear why some paleognaths (thicket tinamou and white-throated tinamou) and all neognaths have degenerated W chromosomes and small PARs. The other hypothesis, derived from a previous study on emu (Vicoso, Kaiser, et al. 2013), implies an excess of male-biased genes on the PAR as resolution of sexual antagonism. However, gene expression data from multiple tissues and stages of emu in this study show that male-biased genes are only enriched on the DR, presumably attributable to incomplete dosage compensation and with very few such genes on the PAR. We find similar patterns in other species.

Classic views on the evolution of sex chromosomes argue that recombination suppression ultimately leads to the complete degeneration of the sex-limited chromosomes (Charlesworth et al. 2005; Bachtrog 2006). However, recent theoretical work suggests suppression of recombination is not always favored, and may require strong sexually antagonistic selection (Charlesworth et al. 2014) or other conditions (Otto 2014). Thus, there may be conditions which would have driven tight linkage of the sex-determining locus and sex-specific beneficial loci via the suppression of recombination in neognaths (Gorelick et al. 2016; Charlesworth 2017), but not in paleognaths, although the exact model that could produce this pattern remains unclear, given that it would require, for example, fewer sexually antagonistic mutations in paleognaths than in neognaths. While theoretically possible, there is little evidence to support such a hypothesis, and indeed some paleognaths (e.g., rheas) have complex mating systems that are at least consistent with extensive sexual conflict (Handford and Mares 1985).

Alternatively, the suppression of recombination between sex chromosomes may be unrelated to sexually antagonistic selection (Rodrigues et al. 2018), and nonadaptive. Simulations suggest that complete recombination suppression can sometimes be harmful to the heterogametic sex, and sex chromosomes are not favorable locations for sexually antagonistic alleles in many lineages (Cavoto et al. 2017). An alternative evolutionary explanation for loss of recombination in the heterogametic sex is then needed. Perhaps the rapid evolution of the sex-limited chromosome may facilitate the expansion of the nonrecombining region on the sex chromosome. For instance, once recombination ceases around the sex-determination locus, the W or Y chromosome rapidly accumulate TEs, particularly LTRs, and the spread of LTRs in the nonrecombining region may in turn increase the chance of LTR-mediated chromosomal rearrangements, including inversions, leading to the suppression of recombination between the W and Z (or Y and X). Further definition and study of the W chromosomes of paleognaths and neognaths, including patterns of substitution and divergence across genes and noncoding regions, is needed to elucidate the role the W in the evolution of avian sex chromosomes.

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

We thank John Parsch, Jamie Walters, Beatriz Vicoso, Daniel Janes, and Qi Zhou for their useful comments, the two anonymous reviewers and Judith Mank for helpful comments on the article, and Dan Janes for collecting emu samples. The computations in this article were performed on the Odyssey cluster at Harvard University and supported by Harvard University Research Computing. This work was supported by NSF grant DEB-135343/EAR-1355292 to S.V.E and NSF grant DEB-1754397 to T. B. S.

Literature Cited

Adolfsson S, Ellegren H. 2013. Lack of dosage compensation accompanies the arrested stage of sex chromosome evolution in ostriches. Mol Biol Evol. 30(4):806–810.

Ansari HA, Takagi N, Sasaki M. 1988. Morphological differentiation of sex chromosomes in three species of ratite birds. Cytogenet Cell Genet. 47(4):185–188.
Avila V, et al. 2014. Faster-X effects in two Drosophila lineages. Genome Biol Evol. 6(10):2968–2982.
Bachtrog D. 2006. A dynamic view of sex chromosome evolution. Curr Opin Genet Dev. 16(6):578–585.
Bachtrog D. 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. Nat Rev Genet. 14:113–124.
Bachtrog D, et al. 2014. Sex determination: why so many ways of doing it? PLoS Biol. 12(7):e1001899.
Baines JF, Sawyer SA, Hartl DL, Parsch J. 2008. Effects of X-linkage and sex-biased gene expression on the rate of adaptive protein evolution in Drosophila. Mol Biol Evol. 25(8):1639–1650.
Bellott DW, et al. 2017. Avian W and mammalian Y chromosomes convergently retained dosage-sensitive regulators. Nat Genet. 49:387–394.
Bergero R, Charlesworth D. 2009. The evolution of restricted recombination in sex chromosomes. Trends Ecol Evol. 24(2):94–102.
Bolivar P, Mugal CF, Nater A, Ellegren H. 2016. Recombination rate variation modulates gene sequence evolution mainly via GC-biased gene conversion, not Hill-Robertson interference, in an avian system. Mol Biol Evol. 33(1):216–227.
Bull JJ. 1983. Evolution of sex determining mechanisms. Menlo Park, California: Benjamin-Cummings Publishing Company.
Carvalho AB, Clark AG. 1999. Genetic recombination: intron size and natural selection. Nature 401(6751):344–344.
Carvalho AB, Clark AG. 2013. Efficient identification of Y chromosome sequences in the human and Drosophila genomes. Genome Res. 23(11):1894–1907.
Cavoto E, Neuenschwander S, Goudet J, Perrin N. 2017. Sex-antagonistic genes, XY recombination and feminized Y chromosomes. J Evol Biol. 65:123.
Chapus C, Edwards SV. 2009. Genome evolution in Reptilia: in silico chicken mapping of 12,000 BAC-end sequences from two reptiles and a basal bird. BMC Genomics 10(Suppl. 2):S8.
Charlesworth B, Campos JL, Jackson BC. 2018. Faster-X evolution: theory and evidence from Drosophila. Mol Ecol. 27(19):3753–3771.
Charlesworth B, Coyne JA, Barton NH. 1987. The relative rates of evolution of sex chromosomes and autosomes. Am Nat. 130(1):113–146.
Charlesworth B, Jordan CY, Charlesworth D. 2014. The evolutionary dynamics of sexually antagonistic mutations in pseudoautosomal regions of sex chromosomes. Evolution 68(5):1339–1350.
Charlesworth D. 2017. Evolution of recombination rates between sex chromosomes. Philos Trans R Soc B. 372(1736):20160456.
Charlesworth D, Charlesworth B, Marais G. 2005. Steps in the evolution of heteromorphic sex chromosomes. Heredity. 95(2):118–128.
Chen N, Bellott DW, Page DC, Clark AG. 2012. Identification of avian W-linked contigs by short-read sequencing. BMC Genomics 13:183.
Cloutier A, et al. 2019. Whole-genome analyses resolve the phylogeny of flightless birds (Paleognathae) in the presence of an empirical anomaly zone. Syst Biol. [Internet]. Available from: https://academic.oup.com/sysbio/advance-article/doi/10.1093/sysbio/syz019/5477409
Connallon T. 2007. Adaptive protein evolution of X-linked and autosomal genes in Drosophila: implications for faster-X hypotheses. Mol Biol Evol. 24(11):2566–2572.
Cortez D, et al. 2014. Origins and functional evolution of Y chromosomes across mammals. Nature 508(7497):488–493.
de Boer LE. 1980. Do the chromosomes of the kiwi provide evidence for a monophyletic origin of the ratites? Nature 287(5777):84–85.
del Priore L, Pigozzi MI. 2017. Broad-scale recombination pattern in the primitive bird *Rhea americana* (Ratites, Paleognathae). PLoS One 12(11):e0187549.
DePristo MA, et al. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet. 43(5):491–498.
Dufresnes C, et al. 2015. Sex-chromosome homomorphy in palearctic tree frogs results from both turnovers and X–Y recombination. Mol Biol Evol. 32(9):2328–2337.
Ellegren H. 2010. Evolutionary stasis: the stable chromosomes of birds. Trends Ecol Evol. 25(5):283–291.
Ellegren H. 2011. Sex-chromosome evolution: recent progress and the influence of male and female heterogamety. Nat Rev Genet. 12(3):157–166.
Galtier N, et al. 2018. Codon usage bias in animals: disentangling the effects of natural selection, effective population size and GC-biased gene conversion. Mol Biol Evol 35(5):1092–1103.
Gamble T, et al. 2017. The discovery of XY sex chromosomes in a Boa and Python. Curr Biol. 27:2148–2153.
Gorelick R, et al. 2016. Abrupt shortening of bird W chromosomes in ancestral Neognathae. Biol J Linn Soc. 119(2):488–496.
Gossmann TI, Santure AW, Sheldon BC, Slate J, Zeng K. 2014. Highly variable recombinational landscape modulates efficacy of natural selection in birds. Genome Biol Evol. 6(8):2061–2075.
Graves J. 2014. Avian sex, sex chromosomes, and dosage compensation in the age of genomics. Chromosome Res. 22(1):45–57.
Graves J. 2016. Evolution of vertebrate sex chromosomes and dosage compensation. Nat Rev Genet. 17(1):33–46.
Gu L, Walters JR. 2017. Evolution of sex chromosome dosage compensation in animals: a beautiful theory, undermined by facts and bedeviled by details. Genome Biol Evol. 9(9):2461–2476.
Handford P, Mares MA. 1985. The mating systems of ratites and tinamous: an evolutionary perspective. Biol J Linn Soc. 25(1):77–104.
Harshman J, et al. 2008. Phylogenomic evidence for multiple losses of flight in ratite birds. Proc Natl Acad Sci U S A. 105(36):13462–13467.
Itoh Y, et al. 2010. Sex bias and dosage compensation in the zebra finch versus chicken genomes: general and specialized patterns among birds. Genome Res. 20(4):512–518.
Janes DE, Ezaz T, Marshall Graves JA, Edwards SV. 2009. Recombination and nucleotide diversity in the sex chromosomal pseudoautosomal region of the emu, *Dromaius novaehollandiae*. J Hered. 100(2):125–136.
Jensen-Seaman MI, et al. 2004. Comparative recombination rates in the rat, mouse, and human genomes. Genome Res. 14(4):528–538.
Kawakami T, et al. 2014. A high-density linkage map enables a second-generation collared flycatcher genome assembly and reveals the patterns of avian recombination rate variation and chromosomal evolution. Mol Ecol. 23(16):4035–4058.
Kousathanas A, Halligan DL, Keightley PD. 2014. Faster-X adaptive protein evolution in house mice. Genetics 196(4):1131–1143.
Krasovec M, Nevado B, Filatov DA. 2018. A comparison of selective pressures in plant X-linked and autosomal genes. Genes 9(5):234.
Kurtz S, et al. 2004. Versatile and open software for comparing large genomes. Genome Biol. 5(2):R12.
Lahn BT, Page DC. 1999. Four evolutionary strata on the human X chromosome. Science 286(5441):964–967.
Le Duc D, et al. 2015. Kiwi genome provides insights into evolution of a nocturnal lifestyle. Genome Biol. 1–15.
Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12:323.
Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv13033997 Q-Bio [Internet]. Available from: http://arxiv.org/abs/1303.3997
Li H, et al. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16):2078–2079.

Lipinska AP, et al. 2017. Multiple gene movements into and out of haploid sex chromosomes. Genome Biol. 18(1):104.
Lockton S, Ross-Ibarra J, Gaut BS. 2008. Demography and weak selection drive patterns of transposable element diversity in natural populations of *Arabidopsis lyrata*. Proc Natl Acad Sci U S A. 105(37):13965–13970.
Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15(12):550.
Lu J, Wu C-I. 2005. Weak selection revealed by the whole-genome comparison of the X chromosome and autosomes of human and chimpanzee. Proc Natl Acad Sci U S A. 102(11):4063–4067.
Mank JE, Axelsson E, Ellegren H. 2007. Fast-X on the Z: rapid evolution of sex-linked genes in birds. Genome Res. 17(5):618–624.
Mank JE, Nam K, Ellegren H. 2010. Faster-Z evolution is predominantly due to genetic drift. Mol Biol Evol. 27(3):661–670.
Mank JE, Vicoso B, Berlin S, Charlesworth B. 2010. Effective population size and the faster-X effect: empirical results and their interpretation. Evolution 64(3):663–674.
Meisel RP, Connallon T. 2013. The faster-X effect: integrating theory and data. Trends Genet. 29(9):537–544.
Naurin S, Hasselquist D, Bensch S, Hansson B. 2012. Sex-biased gene expression on the Avian Z chromosome: highly expressed genes show higher male-biased expression. PLoS One 7(10):e46854.
Nishida-Umehara C, et al. 1999. Differentiation of Z and W chromosomes revealed by replication banding and FISH mapping of sex-chromosome-linked DNA markers in the cassowary (Aves, Ratitae). Chromosome Res. 7:635–640.
Novembre JA. 2002. Accounting for background nucleotide composition when measuring codon usage bias. Mol Biol Evol. 19(8):1390–1394.
Ogawa A, Murata K, Mizuno S. 1998. The location of Z- and W-linked marker genes and sequence on the homomorphic sex chromosomes of the ostrich and the emu. Proc Natl Acad Sci U S A. 95(8):4415–4418.
Otto SP. 2014. Selective maintenance of recombination between the sex chromosomes. J Evol Biol. 27(7):1431–1442.
Otto SP, et al. 2011. About PAR: the distinct evolutionary dynamics of the pseudoautosomal region. Trends Genet. 27(9):358–367.
Perrin N. 2009. Sex reversal: a fountain of youth for sex chromosomes? Evolution 63(12):3043–3049.
Pigozzi MI. 2011. Diverse stages of sex-chromosome differentiation in tinamid birds: evidence from crossover analysis in *Eudromia elegans* and *Crypturellus tataupa*. Genetica 139(6):771–777.
Pigozzi MI, Solari AJ. 1999. The ZW pairs of two paleognath birds from two orders show transitional stages of sex chromosome differentiation. Chromosome Res. 7(7):541–551.
Ramstad KM, Miller HC, Kolle G. 2016. Sixteen kiwi (*Apteryx* spp) transcriptomes provide a wealth of genetic markers and insight into sex chromosome evolution in birds. BMC Genomics 17:410.
Rice WR. 1987. The accumulation of sexually antagonistic genes as a selective agent promoting the evolution of reduced recombination between primitive sex chromosomes. Evolution 41(4):911–914.
Rizzon C, Marais G, Gouy M, Biémont C. 2002. Recombination rate and the distribution of transposable elements in the *Drosophila melanogaster* genome. Genome Res. 12(3):400–407.
Rodrigues N, Studer T, Dufresnes C, Perrin N. 2018. Sex-chromosome recombination in common frogs brings water to the fountain-of-youth. Mol Biol Evol. 67:1.
Romiguier J, Ranwez V, Douzery EJP, Galtier N. 2010. Contrasting GC-content dynamics across 33 mammalian genomes: relationship with life-history traits and chromosome sizes. Genome Res. 20(8):1001–1009.
Rousselle M, Faivre N, Ballenghien M, Galtier N, Nabholz B. 2016. Hemizygosity enhances purifying selection: lack of fast-Z evolution in two satyrine butterflies. Genome Biol Evol. 8(10):3108–3119.
Sackton TB, et al. 2014. Positive selection drives faster-Z evolution in silkmoths. Evolution 68(8):2331–2342.
Sackton TB, et al. 2019. Convergent regulatory evolution and loss of flight in paleognathous birds. Science 364(6435):74–78.
Shields DC, Sharp PM, Higgins DG, Wright F. 1988. "Silent" sites in Drosophila genes are not neutral: evidence of selection among synonymous codons. Mol Biol Evol. 5:704–716.
Singhal S, et al. 2015. Stable recombination hotspots in birds. Science 350(6263):928–932.
Smeds L, et al. 2014. Genomic identification and characterization of the pseudoautosomal region in highly differentiated avian sex chromosomes. Nat Commun. 5:5448.
Stiglec R, Ezaz T, Graves J. 2007. A new look at the evolution of avian sex chromosomes. Cytogenet Genome Res. 117(1–4):103–109.
Szövényi P, et al. 2017. Selfing in Haploid Plants and Efficacy of Selection: Codon Usage Bias in the Model Moss Physcomitrella patens. Genome Biol. Evol. 9(6):1528–1546.
Tomaszkiewicz M, Medvedev P, Makova KD. 2017. Y and W chromosome assemblies: approaches and discoveries. Trends Genet. 33(4):266–282.
Torgerson DG, Singh RS. 2003. Sex-linked mammalian sperm proteins evolve faster than autosomal ones. Mol Biol Evol. 20(10):1705–1709.
Tsuda Y, Nishida-Umehara C, Ishijima J, Yamada K, Matsuda Y. 2007. Comparison of the Z and W sex chromosomal architectures in elegant crested tinamou (*Eudromia elegans*) and ostrich (*Struthio camelus*) and the process of sex chromosome differentiation in paleognathous birds. Chromosoma 116(2):159–173.
Uebbing S, et al. 2015. Quantitative mass spectrometry reveals partial translational regulation for dosage compensation in chicken. Mol Biol Evol. 32(10):2716–2725.
Uebbing S, Künstner A, Mäkinen H, Ellegren H. 2013. Transcriptome sequencing reveals the character of incomplete dosage compensation across multiple tissues in flycatchers. Genome Biol Evol. 5(8):1555–1566.
Vicoso B, Charlesworth B. 2006. Evolution on the X chromosome: unusual patterns and processes. Nat Rev Genet. 7(8):645–653.
Vicoso B, Emerson JJ, Zektser Y, Mahajan S, Bachtrog D. 2013. Comparative sex chromosome genomics in snakes: differentiation, evolutionary strata, and lack of global dosage compensation. PLoS Biol. 11(8):e1001643–15.
Vicoso B, Kaiser VB, Bachtrog D. 2013. Sex-biased gene expression at homomorphic sex chromosomes in emus and its implication for sex chromosome evolution. Proc Natl Acad Sci U S A. 110(16):6453–6458.
Wang Z, et al. 2014. Temporal genomic evolution of bird sex chromosomes. BMC Evol. Biol. 14:250.
Wright AE, Dean R, Zimmer F, Mank JE. 2016. How to make a sex chromosome. Nat Commun. 7:1–8.
Wright AE, et al. 2015. Variation in promiscuity and sexual selection drives avian rate of Faster-Z evolution. Mol Ecol. 24(6):1218–1235.
Xu L, et al. 2019. Dynamic evolutionary history and gene content of sex chromosomes across diverse songbirds. Nat Ecol Evol. 3(5):834.
Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol. 24(8):1586–1591.
Yazdi HP, Ellegren H. 2014. Old but not (so) degenerated–slow evolution of largely homomorphic sex chromosomes in ratites. Mol Biol Evol. 31(6):1444–1453.
Yazdi HP, Ellegren H. 2018. A genetic map of ostrich Z chromosome and the role of inversions in avian sex chromosome evolution. Genome Biol Evol. 10(8):2049–2060.

Zhang G, et al. 2014. Comparative genomics reveals insights into avian genome evolution and adaptation. Science 346(6215):1311–1320.
Zhang J, Li C, Zhou Q, Zhang G. 2015. Improving the ostrich genome assembly using optical mapping data. GigaScience 4(1):298–293.
Zhang Q, Edwards SV. 2012. The evolution of intron size in amniotes: a role for powered flight? Genome Biol Evol. 4(10):1033–1043.
Zhou Q, et al. 2014. Complex evolutionary trajectories of sex chromosomes across bird taxa. Science 346(6215):1246338.

Associate editor: Judith Mank

Supplementary material

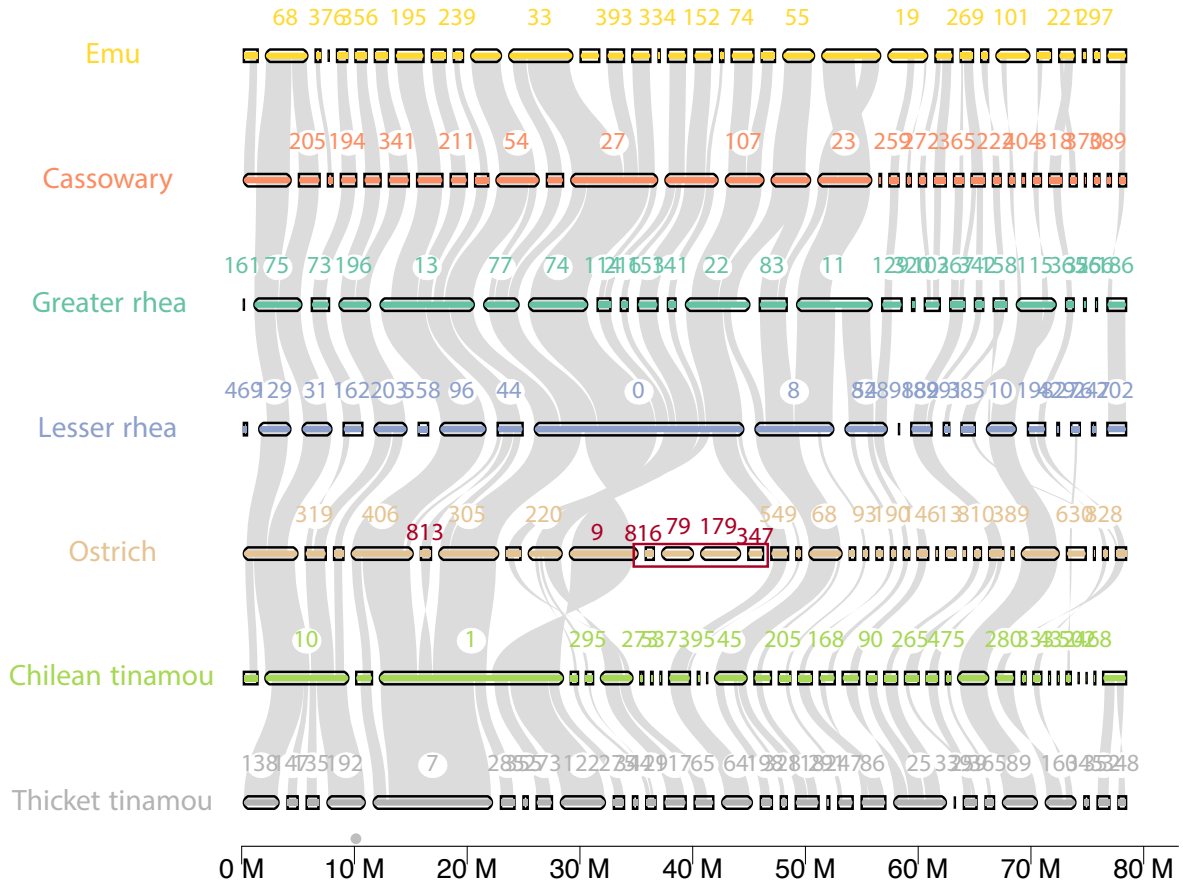


FIG. S1. Gene synteny among the Z chromosomes. The alignment of coding sequence and plot were implemented by the python package jcvi (MCscan). Only scaffolds (bars) longer than 50k were shown. As a showcase, the orientation of scaffold 813 of ostrich was corrected. The ~12M containing scaffolds 816, 79, 179, 347 and a part of scaffold 9 were removed from the ostrich. Mate-pair reads alignment for the breakpoint on the scaffold 9 is shown in S2.

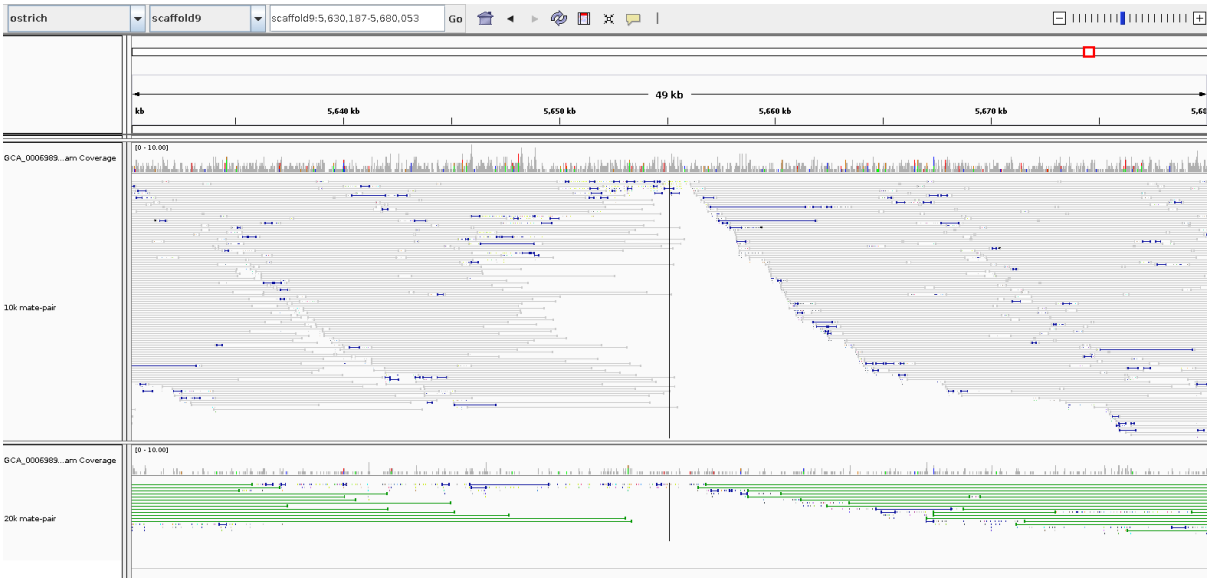


FIG. S2. Alignments of mate-pair reads against the scaffold9 of ostrich. The breakpoint is located at the near-end of the scaffold (~5.6M). The upper panel shows the alignments of 10k mate-pair reads and the bottom panel is for 20k mate-pair reads.

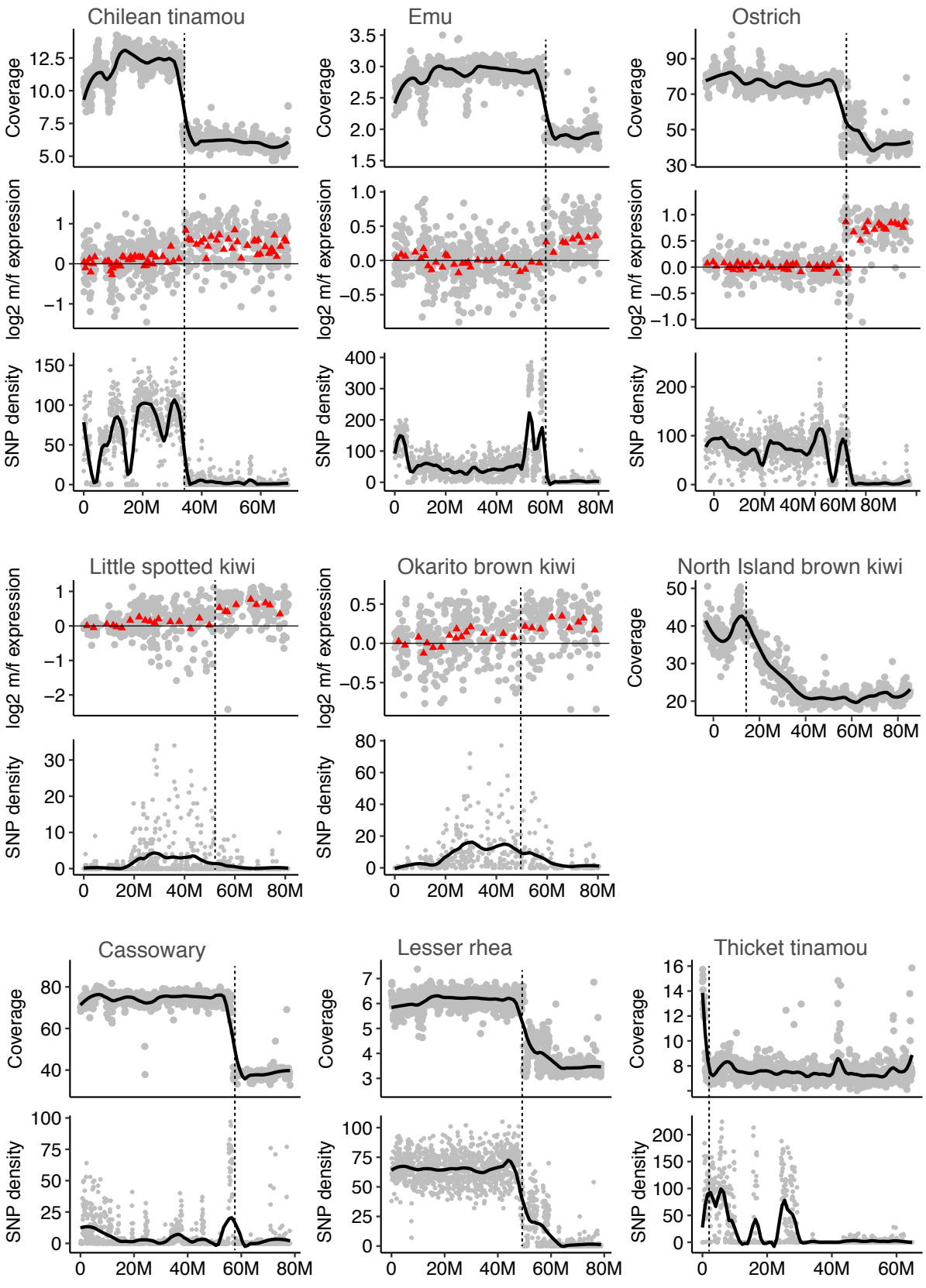


FIG. S3. Annotation of PAR/DR boundary. In the 'coverage' panels, each dot represents a 50k window. The black dashed line denotes the boundary of the pseudoautosomal region (PAR) and the differentiated region (DR). In kiwis, an addition dashed line at ~18M show a putative

PAR boundary. In ‘m/f expression’ panels, the red triangle represents the mean m/f expression ratio of 20 genes. The ‘SNP density’ shows the density of female heterozygous sites or SNPs over 50k windows. For both kiwi, female RNA-seq reads were used to call SNPs and the density of SNPs were calculated by dividing the number of SNPs over the length of exonic sequences for every 50k windows. Heterozygous sites were called using GATK pipelines based on female-reads alignments. Note that white-throated tinamou is not shown, as previously published PAR and DR annotations were used for this species.

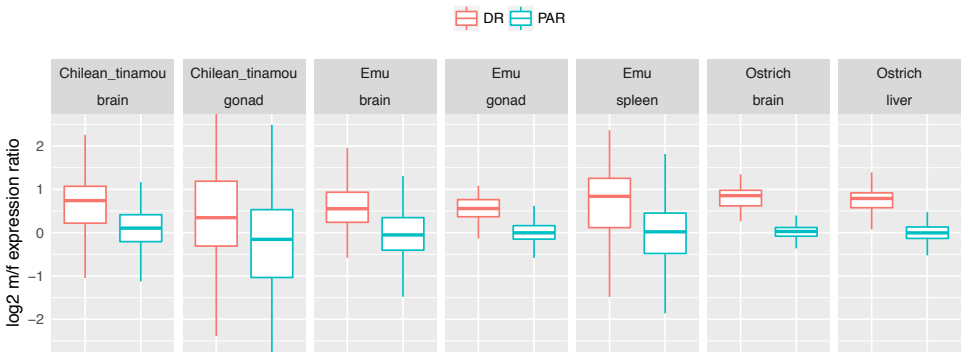


FIG. S4. Male-to-female expression ratios for DR- and PAR-linked genes. For Chilean tinamou, emu and ostrich, RNA-seq data of multiple tissues of both sexes are available. The m/f ratios (log2 transformed) of DR-linked genes are larger than 1 but less than 2, suggesting incomplete dosage compensation, but show limited variation within species.

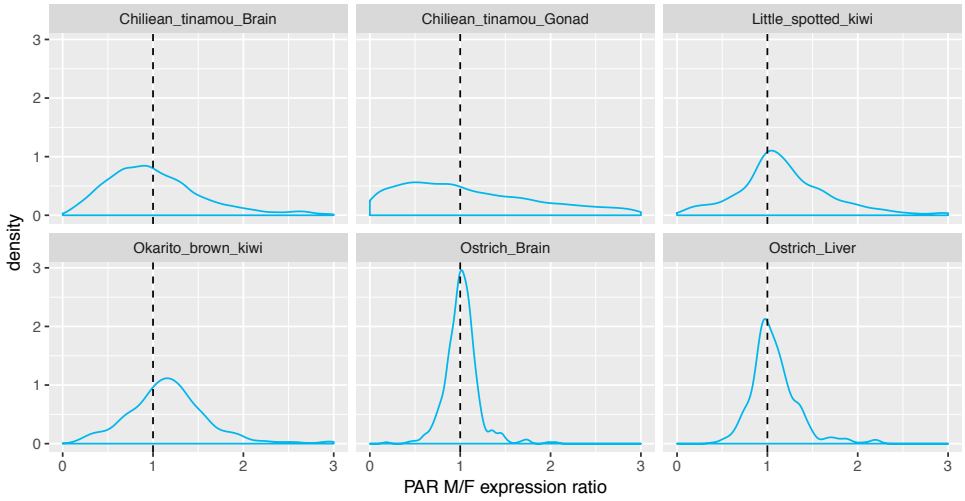


FIG. S5. Distribution of male-to-female expression ratios for PAR-linked genes. In most samples m/f expression ratios do most deviate from 1, suggest similar expression levels of PAR-linked genes between males and females. Only in Okarito brow kiwi, however, male expression levels are slightly higher than for females.

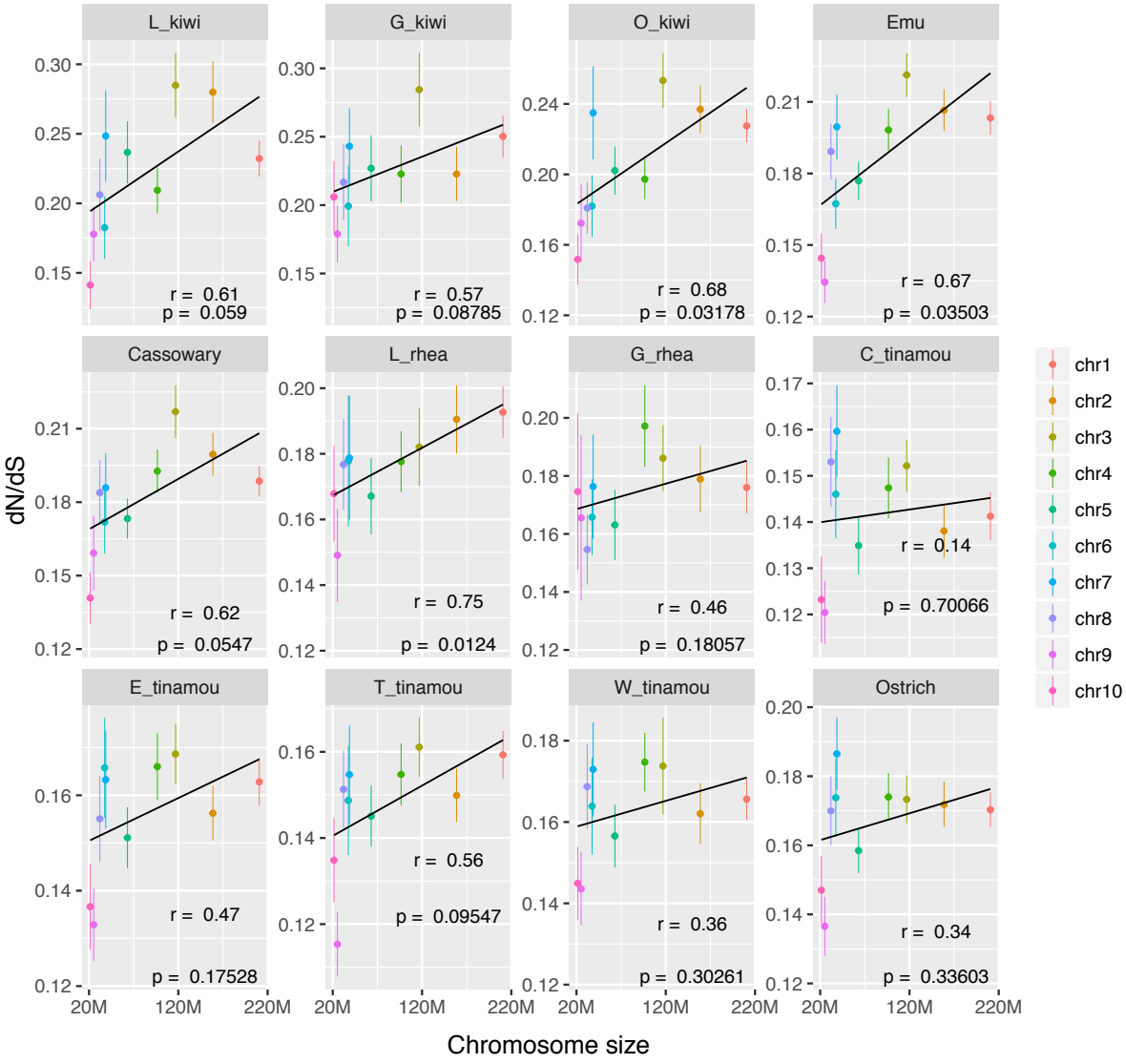


FIG. S6. Positive correlation of dN/dS ratios and chromosome size among macro-chromosomes. Among macro-chromosome (chr1 – chr10), chromosome size positively correlates with dN/dS ratios. The chromosome size of the Z is about 75M, between the sizes of chr4 (~97M) and chr5 (~63M). The 'r' strands for Pearson's correlation coefficient. Abbreviation for species names: L_kiwi, little spotted kiwi; G_kiwi, great spotted kiwi; O_kiwi, Okarito brown kiwi; L_rhea, Lesser rhea; G_rhea, Greater rhea; C_tinamou, Chilean tinamou; E_tinamou, elegant crested tinamou; T_tinamou, thicket tinamou; W_tinamou, white-throated tinamou.

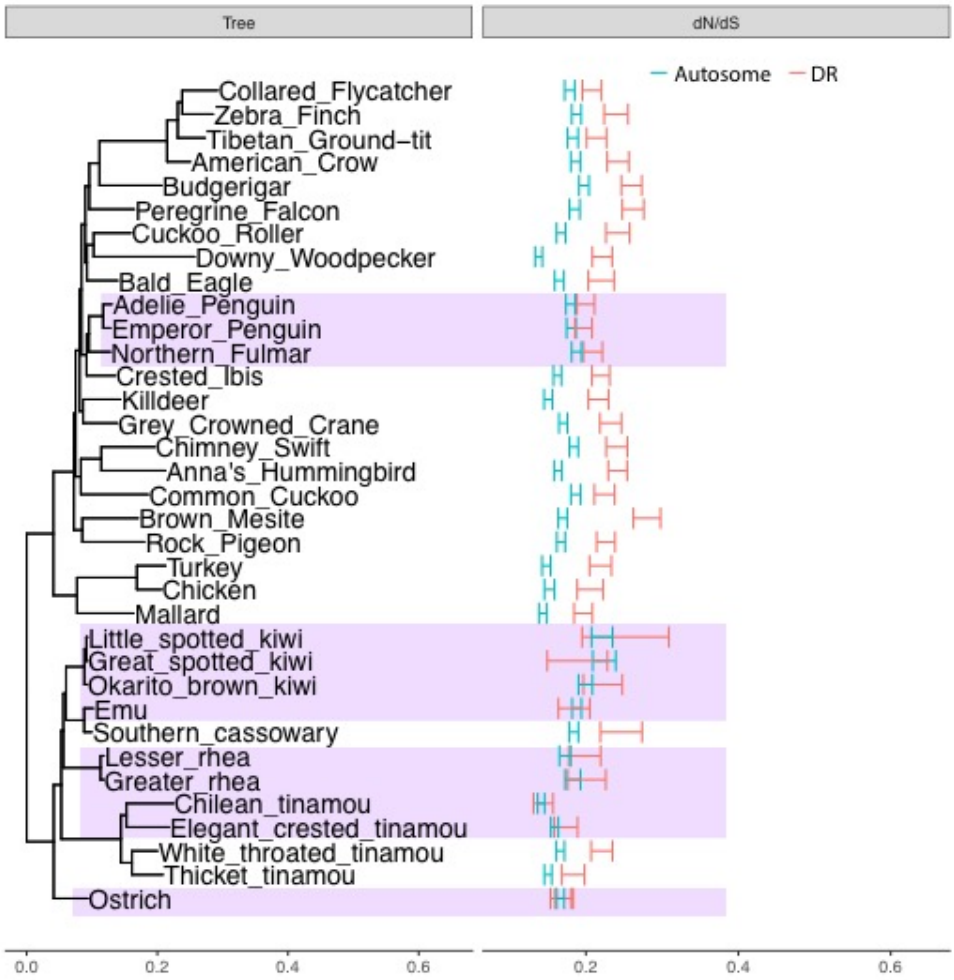


FIG. S7. A lack of faster-DR in most palaeognaths. The PAR-linked genes were removed from the analysis. Species without faster-DR effect (permutation test, P > 0.05) were highlighted by purple colour. The faster-Z effect is no longer observed in Okarito brown kiwi, elegant crested tinamou and thicket tinamou after PAR-linked genes were removed.

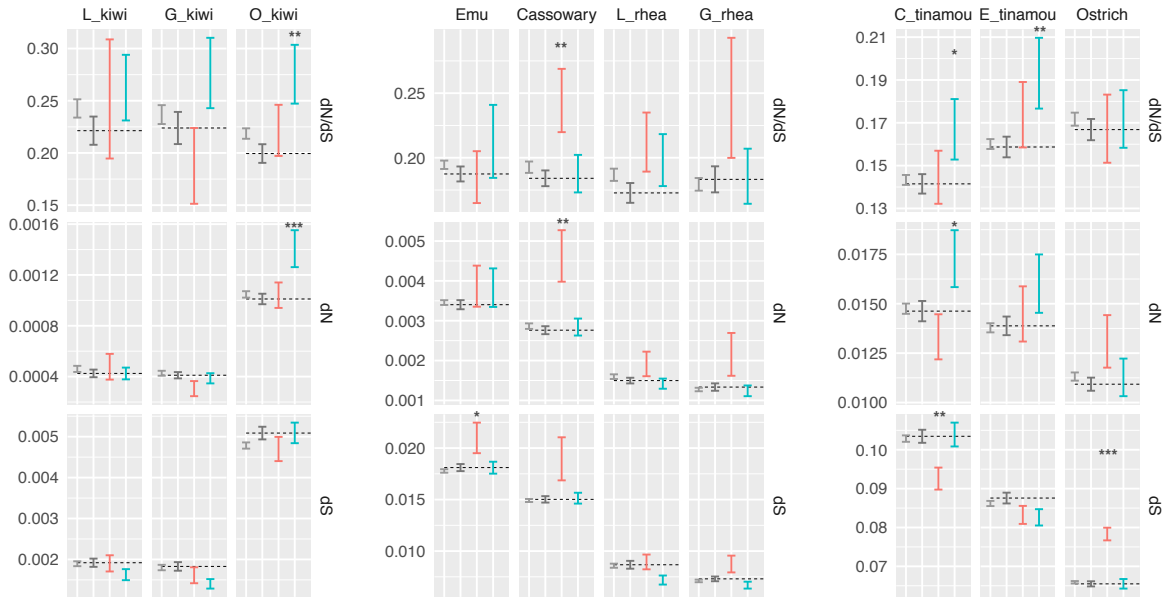


FIG.S8. The boundary of PAR/DR does not show faster-Z effect. dN (nonsynonymous substitution rate), dS (synonymous substitution rate) and their ratios (dN/dS) are shown for PAR (cyan), DR (red), chr4/5 (dark grey) and macro-chromosome (grey) genes. The test for faster-Z evolution was repeated after the exclusion of PAR-linked gene close to PAR boundaries (less then 5 Mb away). Similar to Fig. 4 in the main text, Confidence intervals were estimated by 1,000 bootstraps. Asterisks indicate the significant levels of PAR/DR *vs.* chr4/5 comparison (two-sided permutation test), * <0.05, ** <0.01, *** <0.001.

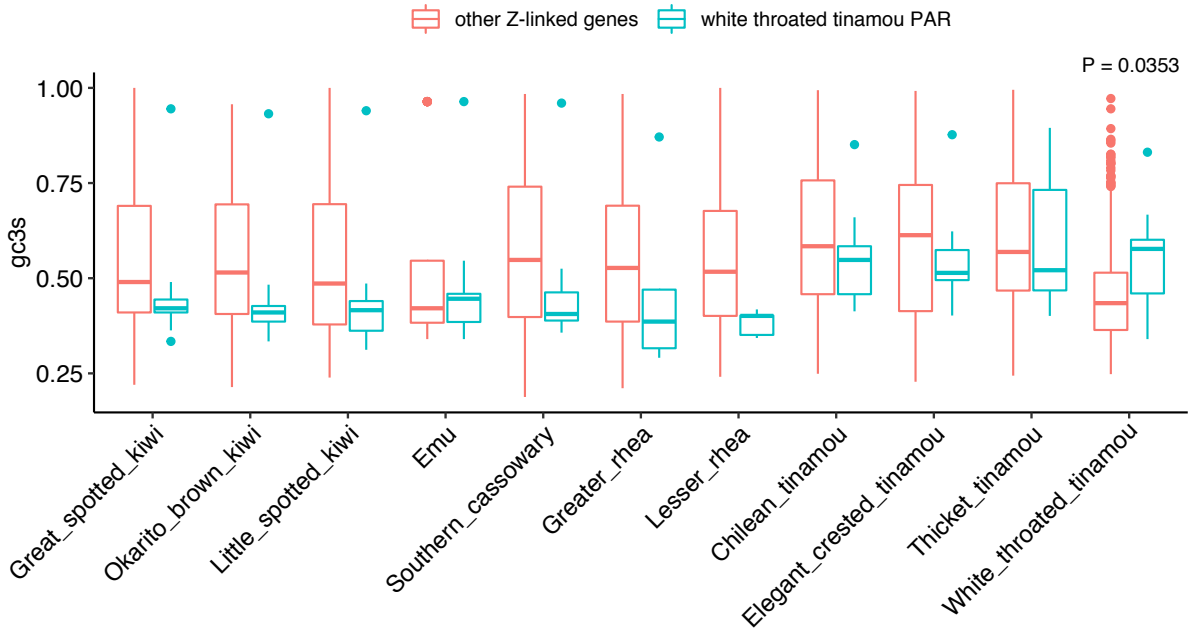


FIG. S9. The PAR-linked genes in white throated tinamous is GC-biased. The boxplots show the median of gc3s of nine PAR-linked (genBank ID 104571644, 104571645, 104571646, 104571647, 104571642, 104571648, 104571649, 104571643 and 104571650) gene and the

rest Z-linked genes. Their homologous genes in other species are also shown for comparison. The PAR-linked genes of white throated tinamou are GC-biased only in white throated tinamou (P = 0.0353, Wilcoxon rank-sum test).

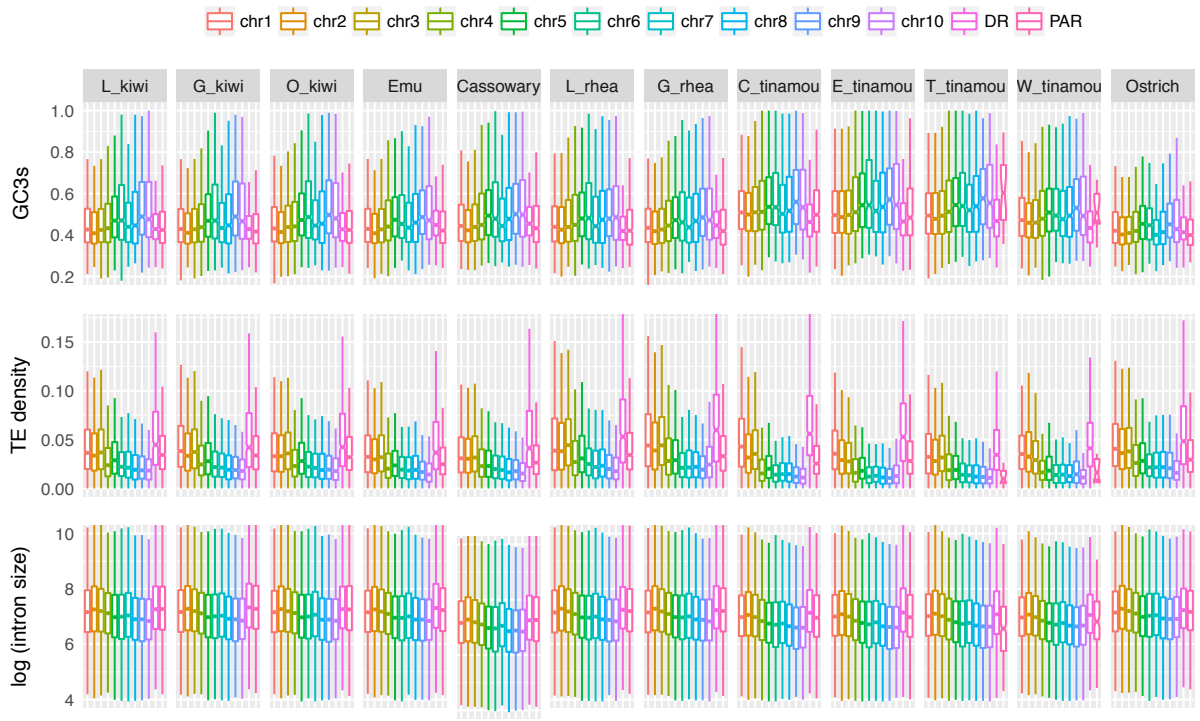


FIG. S10. Comparison of genomic feature among macro-chromosomes. GC3s (GC content of synonymous site of the third codon) and exon density show negative correlation with chromosome size, while TE (transposable element) density and intron size show positive correlation with chromosome size.

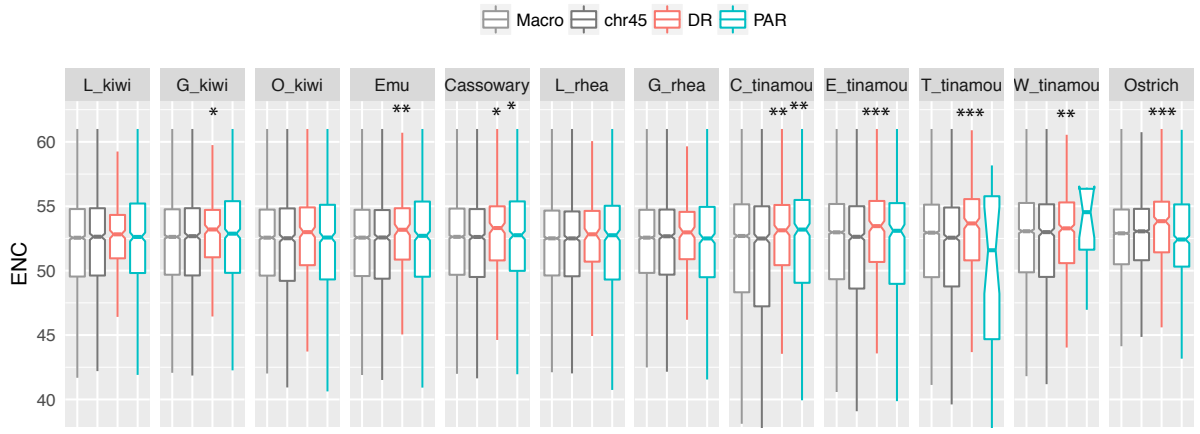


FIG. S11. Comparison of ENC between PAR/DR and autosomes. The ENC (Effective Number of Codons) values are higher in DRs for many species, but only for cassowary and Chilean tinamou ENC values are higher in PAR than for autosomes. Asterisks indicate the significant levels of PAR/DR *vs.* chr4/5 comparison (Wilcoxon sum rank test), * <0.05, ** <0.01, *** <0.001.

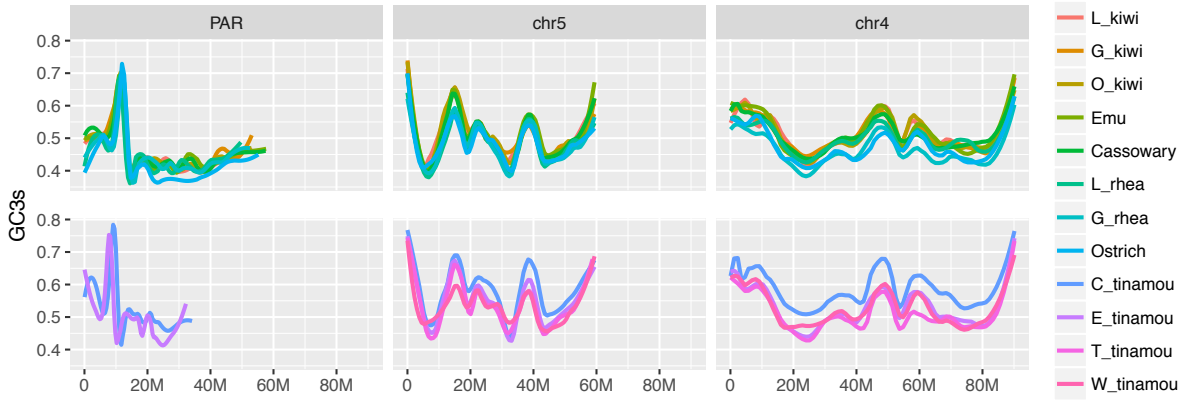


FIG. S12. Reduced GC3s on the PARs compared to chr5 and chr4. The location of the PAR-linked genes is based on the pseudo-chromosome Z, and the location of genes of chr4 and chr5 are based on the homologous genes of the chicken genomes. The abbreviation for species names is the same as in FIG S8.

Table S1. The length of pseudoautosomal region (PAR) and differentiated region (DR) in palaeognaths and selected neognaths

Species	PAR		DR		Reference
	Length (bp)	#gene	Length (bp)	#gene	
Little spotted kiwi	53,648,137	644	27,858,477	315	This study
Great spotted kiwi	53,103,935	639	27,917,301	295	This study
Okarito brown kiwi	53,052,411	655	29,311,255	345	This study
North Island brown kiwi	~20M	-	~65M	-	This study
Emu	59,302,072	695	21,929,632	234	This study
Southern cassowary	59,264,808	749	22,782,997	295	This study
Lesser rhea	54,869,205	508	26,064,611	207	This study
Great rhea	52,553,322	604	29,742,736	214	This study
Chilean tinamou	34,050,901	485	36,483,903	380	This study
Elegant created tinamou	32,217,551	419	33,168,615	350	This study
Thicket tinamou	250,000	10	71,263,047	831x	This study
White-throated tinamou	685,144	14	62,454,206	736	(Zhou et al. 2014)
Ostrich	52,483,918	704	31,782,146	391	(Zhou et al. 2014), this study
Collared flycatcher	630,000	17	68,355,977	591	(Smeds et al. 2014)
Zebra finch	450,000	16	75,826,118	653	(Singhal et al. 2015)
Chicken	10,000	0	82,519,921	826	(Bellott et al. 2017)
Pekin duck	1,050,000	-	76,500,000	-	(Zhou et al. 2014)

Large PAR species are shaded in gray

Table S2. P-values of Fisher's exact test for overrepresentation of sex-biased on the Z chromosome and PAR.

Tissue	Z chromosome	PAR

	Male biased	Female biased	Male biased	Female biased
Spleen	**1.26E-23**	0.062	0.303	0.122
Gonad	**2.72E-07**	1.000	1.000	1.000
Brain	**0.00848**	0.211	0.637	0.910
Embryo day15	**1.52E-05**	1.000	0.214	0.747
Embryo day42	**0.00676**	0.524	1.000	0.174

Table S3. Correlation between chromosome sizes and genomic features

Species		GC3s	TE density	Intron size	Exon density	ENC	Intergenic size
L_kiwi	r	-0.86	0.90	0.83	-0.68	0.79	0.53
	p-value	0.00143	0.00033	0.0033	0.03091	0.00633	0.11574
G_kiwi	r	-0.86	0.93	0.81	-0.75	0.74	0.50
	p-value	0.00157	0.00009	0.00434	0.01307	0.01534	0.1394
O_kiwi	r	-0.86	0.90	0.74	-0.63	0.87	0.19
	p-value	0.00134	0.00045	0.01365	0.0532	0.00117	0.60513
Emu	r	-0.86	0.92	0.86	-0.71	0.90	0.71
	p-value	0.00158	0.00019	0.00125	0.02259	0.00035	0.02122
Cassowary	r	-0.87	0.92	0.84	-0.84	0.87	0.76
	p-value	0.00105	0.00013	0.00256	0.00209	0.00119	0.01071
L_rhea	r	-0.88	0.90	0.87	-0.77	0.77	0.62
	p-value	0.00069	0.00039	0.0011	0.00957	0.00951	0.05399
G_rhea	r	-0.91	0.92	0.82	-0.81	0.77	0.86
	p-value	0.00021	0.00017	0.00406	0.00439	0.00893	0.00124
C_tinamou	r	-0.79	0.95	0.89	-0.86	0.91	0.81
	p-value	0.00681	0.00003	0.00059	0.00131	0.00022	0.00447
E_tinamou	r	-0.90	0.94	0.91	-0.79	0.94	0.74
	p-value	0.00039	0.00006	0.00022	0.00681	0.00006	0.01391
T_tinamou	r	-0.89	0.97	0.91	-0.85	0.88	0.78
	p-value	0.00066	0	0.00023	0.0017	0.00068	0.00735
W_tinamou	r	-0.82	0.94	0.87	-0.80	0.80	0.71
	p-value	0.00396	0.00006	0.00119	0.00513	0.00555	0.02226
Ostrich	r	-0.54	0.95	0.75	-0.77	0.51	0.75
	p-value	0.10623	0.00004	0.01255	0.00933	0.1314	0.01212

Table S4. Location of Janes et al 2009 BAC sequence in the emu genome assembly.

GENBANK RECORD	CHROMOSOME (JANES ET AL 2009)	EMU GENOME LOCATION	EMU GENOME CHROMOSOME
EU200931	Autosome	not determined	not determined
EU200931	Autosome	not determined	not determined
ET041500	Autosome	not determined	not determined
ET041501	Autosome	not determined	not determined
ET041502	Autosome	not determined	not determined
ET041515	Autosome	not determined	not determined
ET041512	Autosome	not determined	not determined
ET041513	Autosome	not determined	not determined
AB002056	PAR	presumed assembly gap	Z (PAR) (1)
AB006694	PAR	scaffold_221: 152067-154756	Z (DR)
AY095498	PAR	scaffold_13:6582073-6583283	Z (DR)
AB006695	PAR	scaffold_239:1028675-1030884	Z (PAR)
ET041507	PAR	scaffold_16:5843173-5843899	chr5
ET041520	PAR	scaffold_66:2150084-2282118	chr7
ET041521	PAR	scaffold_66:2150084-2282118	chr7
ET041516	PAR	scaffold_14:4429102-4300170	chr4
ET041517	PAR	scaffold_14:4429102-4300170	chr4
ET041508	PAR	scaffold_19: 1980997-2079372	Z (DR)
ET041509	PAR	scaffold_19: 1980997-2079372	Z (DR)
ET041510	PAR	scaffold_19: 1980997-2079372	Z (DR)
ET041518	PAR	scaffold_106:822019-944625	chr8
ET041519	PAR	scaffold_106:822019-944625	chr8

(1) determined by alignment to other palaeognaths

Chapter 4

Paper III: Female-specific and dosage selections restore genes through transpositions onto the degenerated songbird W chromosomes

Luohao Xu[1,2], Martin Irestedt[3], Qi Zhou[1,2,4]

[1] MOE Laboratory of Biosystems Homeostasis & Protection, Life Sciences Institute, Zhejiang University, Hangzhou, China

[2] Department of Molecular Evolution and Development, University of Vienna, Vienna, Austria

[3] Department of Bioinformatics and Genetics, Swedish Museum of Natural History, Stockholm, Sweden

[4] Center for Reproductive Medicine, The 2nd Affiliated Hospital, School of Medicine, Zhejiang University

Highlights

- Transposition from the Z to W chromosome occurred three times in songbirds
- Seven transposition-derived genes survived on the W chromosome
- Survived genes are generally dosage-sensitive or housekeeping genes
- One ovary-biased gene has been transposed due to female-specific selection

Summary

Homologous recombination is usually suppressed between sex chromosomes, which leads to the loss of functional genes on the W and Y chromosomes. It remains unclear how species like birds with a ZW sex system (male ZZ, female ZW) cope with the consequential gene dosage imbalance, in the absence of global dosage compensation mechanism. Here we tackle this conundrum by reporting 14 genes recently duplicated from the Z to the W chromosomes of three songbird lineages, after analyzing a total of 12 songbird species' genomes. These Z-to-W transpositions are estimated to have occurred within 9 million years. Besides the expected signatures of functional degeneration in some genes on the non-recombining W chromosomes, many other retained genes after transposition are putative haploinsufficient genes or housekeeping genes. Several genes show biased expression in ovaries of birds or lizard, or function in female germ cells. These results, together with the reported X-to-Y transpositions,

strongly suggest that sex-specific and dosage selections may have recurrently driven the restoration of genes on the W or Y chromosomes, and their evolutionary processes are more dynamic than simply becoming completely degenerated

Keywords
Transposition; Sex chromosome; Songbird; Dosage sensitivity; Sex-specific selection

Abstract

Homologous recombination is usually suppressed between sex chromosomes, which leads to the loss of functional genes on the W and Y chromosomes. It remains unclear how species like birds with a ZW sex system (male ZZ, female ZW) cope with the consequential gene dosage imbalance, in the absence of global dosage compensation mechanism. Here we tackle this conundrum by reporting 14 genes recently duplicated from the Z to the W chromosomes of three songbird lineages, after analyzing a total of 12 songbird species' genomes. These Z-to-W transpositions are estimated to have occurred within 9 million years. Besides the expected signatures of functional degeneration in some genes on the non-recombining W chromosomes, many other retained genes after transposition are putative haploinsufficient genes or housekeeping genes. Several genes show biased expression in ovaries of birds or lizard, or function in female germ cells. These results, together with the reported X-to-Y transpositions, strongly suggest that sex-specific and dosage selections may have recurrently driven the restoration of genes on the W or Y chromosomes, and their evolutionary processes are more dynamic than simply becoming completely degenerated.

The female-specific W or male-specific Y chromosomes very often embark on an irreversible trajectory of functional degeneration, at regions where their homologous recombination with the Z or X chromosomes was suppressed (Charlesworth and Charlesworth 2000; Bachtrog 2013). The recombination suppression between sex chromosome pair was proposed to be driven by the selection for restricting the sex-determining (SD) genes, or genes beneficial to one sex but detrimental to the other (so-called 'sexual antagonistic', SA genes) within one sex from being inherited in the opposite sex through recombination (Ponnikas, et al. 2018). The consequential cost of maintaining the SD and SA genes within one sex is essentially much less effective natural selection on the W/Y chromosome due to the lack of recombination (Charlesworth and Charlesworth 2000). Although some genes with important regulatory functions or high dosage-sensitivity have been demonstrated to be degenerating much slower than others on the mammalian Y (Bellott, et al. 2014; Cortez, et al. 2014) or the avian W chromosomes (Smeds, et al. 2015; Bellott, et al. 2017; Xu, et al. 2019), due to a much higher level of selective constraints. This nevertheless creates a conundrum that when recombination was initially suppressed, the affected regions must contain a great number of sex-linked genes with important functions besides the SD/SA genes.

A direct resolution to such 'collateral damage' is evolution of dosage compensation on the Z/X chromosome, so that the balance of expression level can be restored. In addition, studies showed that the W/Y chromosomes come up with various strategies to 'rescue' functions of certain genes during their complex and dynamic evolutionary course. The human Y chromosome contains palindromic sequence structures that are thought to have been favored by natural selection, because they help repair deleterious mutations and facilitate gene conversions between Y-linked genes (Rozen, et al. 2003). Other ways of rescuing or even innovating the gene functions on the Y chromosomes include escaping onto the autosomes (Hughes, et al. 2015), or recruiting novel genes via various resources. Emerging cases of gene restorations on the Y chromosome after the complete loss of original copies have been reported since the characterization of 'X-transposed' region (XTR) on the male-specific region of human Y chromosome (MSY) over 30 years ago (Page, et al. 1984; Schwartz, et al. 1998; Skaletsky, et al. 2003). The XTR was duplicated from the X chromosome onto the Y chromosome within 4.7 million years (MY) (Ross, et al. 2005) after the human-chimpanzee split, and subsequently disrupted into two blocks by a Y-linked inversion (Schwartz, et al. 1998). The enclosed *PCDH11* X-Y gene pair has been suggested to contribute to the human-specific cerebral asymmetry and language development (Crow 2002; Speevak and Farrell 2011). More cases of transposition from the X chromosome or autosomes to the Y chromosome have been reported in Drosophila (Koerich, et al. 2008; Carvalho, et al. 2015; Tobler, et al. 2017) or other Diptera species (Mahajan and Bachtrog 2017), dog (Li, et al. 2013), cat (Li, et al. 2013; Brashear, et al. 2018)

and horse (Janečka, et al. 2018), suggesting such transposition events are not rare during the Y chromosome evolution.

Little is known about whether and how the avian W chromosome resolves the conundrum of losing dosage-sensitive genes long after the recombination was suppressed, which is particularly important given that global dosage compensation has never evolved on the homologous Z chromosome (Itoh, et al. 2007; Graves 2014; Gu and Walters 2017). A previous study showed that palindromic sequence structures also exist on the W chromosomes of sparrows and blackbirds (Davis, et al. 2010). This suggests that birds and mammals, despite their independent origins of sex chromosomes, can convergently evolve sequence structures to retard the functional degeneration of their W or Y chromosomes. However, one might expect that DNA-mediated transposition or RNA-mediated retrotransposition events are scarce in avian genomes due to their compact structures with a much lower repeat content to mediate these events, particularly the L1 retroposons relative to mammals (International Chicken Genome Sequencing 2004; Suh 2015). Indeed, there are only 51 retrogenes identified in chicken, compared to over 8,000 cases in human (Zhang, et al. 2003; International Chicken Genome Sequencing 2004). So far no transposed genes have been reported on the avian W chromosomes, and we have recently reported one retrotransposed gene on the W chromosome of American crow (Xu, et al. 2019). Of course, these results are far from being conclusive regarding the role of transposition or retrotransposition in the evolution of avian W chromosomes, because only a few out of over 10,000 bird species have been investigated. In addition, the degree of sexual selection, which is known to dramatically vary across bird species, must have a different impact shaping the evolution of sex chromosomes.

Here we sought to address the question of how birds cope with their W-linked gene loss without global dosage compensation, by studying 12 songbird genomes whose male and female sequencing data are both available. We reasoned that these Illumina-based genomes do not contain complete information of complex and repetitive sequence structures (e.g., palindromes) or traces of ancient transposition events, if any on the W chromosome. We therefore focused on searching for the recent duplicative Z-to-W transpositions, similar to the XTR of human (for simplicity, referred as transpositions or transposed genes hereafter) that were manifested as female-specific elevations of both read coverage and heterozygosity level (i.e. Z/W sequence divergence level), relative to other Z-linked regions that have become hemizygous. Those located at the end of the chromosome with an elevation of female coverage to the hemizygous Z-linked regions, but without sex-specific patterns of heterozygosity, were inferred as pseudoautosomal regions (PAR) that maintained recombination between sex chromosomes (**Figure 1, Supplementary Fig. S1**).

Intriguingly, we identified four Z-to-W transpositions involving 14 genes, with 6 genes subsequently deleted (see below), among 4 songbird species great tit (*Parus major*), medium ground finch (*Geospiza fortis*), red bird-of-paradise (*Paradisaea rubra*) and Raggiana bird-of-paradise (*P. raggiana*). We also identified a very recent Z-linked duplication, which showed elevations of read coverage and heterozygosity in both sexes (**Figure 2d, Supplementary Fig. S1c**), a pattern distinguishable from that of transpositions. Recombination with the W chromosome has been suppressed in the Z-linked regions involved in the transpositions at least 85 million years (MY) ago, where most primary W-linked gene copies have become completely lost (Zhou, et al. 2014). We further confirmed that none of the 8 retained genes after transposition can be found from any of the previously assembled W-linked genomic sequences for the studied species (Xu, et al. 2019). Therefore, these transpositions probably occurred after the original W-linked genes had become lost. To verify these identified recent transpositions, we randomly selected and amplified 10 sex-linked genomic fragments in both sexes of great tit, and genotyped 60 SNPs within and near the transposition loci. We confirmed that female-specific heterozygous sites were only present within the transposed regions (see two examples in **Supplementary Fig. S2**). The two birds-of-paradise species share the same transposition (**Supplementary Fig. S3**), and for simplicity hereafter we used red bird-of-paradise to represent this lineage. The lengths of detected transposed regions range from 67kb in great tit to 1.3Mb in bird-of-paradise species. We dated the transposition of medium ground finch about 8.3 MY ago, as the same transpositions have been found in all the other Coerebinae (Darwin's finches and their relatives) but absent in their sister group Sporophilinae (Lamichhaney, et al. 2015) (**Supplementary Fig. S4**). Similarly, we dated the transpositions of bird-of-paradise species within 4 MY (**Supplementary Fig. S3**) and that of great tit about 7 MY ago, after examining their sister species.

These very recent Z-to-W transpositions provided us a unique window to examine the evolution of W-linked genes at their early stages. They show clear evidence of functional degeneration. For instance, among the five genes transposed in medium ground finch, at least one (*THBS4*) has become a probable pseudogene due to frameshift mutations (**Supplementary Fig. S5**). The most prominent case of gene loss was found in bird-of-paradise species (**Figure 2a-c**). A 1.3Mb-long region on the Z chromosome shows clear signatures of transposition, except for a large encompassing 583kb region and a nearby 2kb region (**Figure 2c**, **Supplementary Fig. S6**). The involved 8 Z-linked genes and their residing scaffold sequence show a conserved synteny across multiple bird species (**Supplementary Fig. S7**), suggesting there were no intrachromosomal rearrangements on the Z chromosome. Based on these results, we inferred that there was one large Z-to-W transposition, followed by two deletion events on the W chromosome. This is more parsimonious a scenario than multiple independent

transpositions occurred in the same region. This scenario is also supported by the similar level of female heterozygosity, i.e., Z/W pairwise divergence level surrounding the deleted regions (**Supplementary Fig. S8**). The large 583kb-long deletion has removed 4 complete genes and 2 partial genes on the W chromosome after the transposition (**Figure 2e**)**.** We have not detected any large-scale insertions into the transposed regions, based on analyses of insert size of mate-pair libraries.

While such gene losses are expected because of the lack of recombination, the retained genes, essentially the recently restored genes that had previously become lost on the W chromosomes, are more informative for the driving forces that originally fixed these transpositions. We reasoned that two types of selection, i.e., female-specific selection for the female reproductive genes, as well as dosage selection for the haploinsufficient genes probably account for the restoration of W-linked genes. The first type of selection is demonstrated by a previous study showing that the chicken breeds selected for higher female fecundity exhibit an increased W-linked gene expression than other breeds (Moghadam, et al. 2012). Indeed, the only two retained genes *ANXA1* and *ALDH1A1* after the transposition in bird-of-paradise species (**Figure 2**), and the great tit transposed gene *MELK* all have a biased or specific expression pattern in ovary in many examined bird species (**Supplementary Fig. S9**), and also their outgroup species green anole lizard (**Figure 3**). Although *ALDH1A1* has a relatively lower expression level in ovary than in testis, it has been recently shown in mice that the disruption of this gene delays the onset of meiosis in ovary (Bowles, et al. 2016). Besides, *ANXA1* and *CDK7* probably have been restored by strong dosage selection, indicated by their much higher levels of predicted haploinsufficiency (HP score) than most other genes on the Z chromosome (**Supplementary Fig. S9**) (Huang, et al. 2010), as well as a lack of any nonsynonymous changes compared to their Z-linked homologs (**Supplementary Table 1**). Several medium ground finch genes, for example, *SERINC5* and *MTX3*, have a low HP score, but a very broad expression pattern across tissues measured by tissue-specificity matrix *tau*, thus are likely restored as housekeeping genes (**Figure 3**). In fact, the restored genes tend to have on average a higher HP score (although not significantly, P=0.051, Wilcoxon test) than those that have become lost after the transpositions.

These results together strongly suggested that the female-specific and dosage selections have driven the frequent restoration of W-linked genes through transpositions among songbird species. Because similar X-to-Y transpositions have been reported in insects and mammals (Page, et al. 1984; Mahajan and Bachtrog 2017; Tobler, et al. 2017; Janečka, et al. 2018), we propose that restoration of once-lost genes onto the non-recombining sex chromosomes is probably a general feature in sex chromosomes evolution. Such restoration is not expected to alter the evolutionary trajectories of W or Y chromosomes toward complete

functional degeneration. In fact, we found some transposed genes have already become lost or shown signatures of functional degeneration (e.g., *THBS4*). Such loss-and-restoration cycles may recurrently occur throughout the evolution of sex chromosomes, particularly in ZW systems that usually do not have global dosage compensation to cope with the imbalance of gene expression. We have to point out that our method can only identify recent transpositions, and probably has missed ancient transpositions that have become too divergent in sequence between Z and W chromosomes. The genes involved in the such cases nevertheless have probably already become pseudogenes. Our results are in line with the reported cases in avian W or mammalian Y chromosomes that dosage-sensitive genes are retarded for their functional degeneration due to the strong selective constraints (Bellott, et al. 2014; Smeds, et al. 2015; Bellott, et al. 2017; Xu, et al. 2019). We also provided new evidence that sex-specific selection is shaping the evolution of the W chromosome, which was assumed to be less frequent than that shaping the Y chromosome, due to the more frequent and intensive male-targeted sexual selection.

Materials and Methods

The genomic, transcriptomic and resequencing data used in this study are listed in **Supplementary Table 2-4**. For the studied 12 songbird species, genomic data are available for both sexes except for three species. Genome assemblies were derived from female samples, except for great tit. We first used the published Z chromosome sequence of great tit (Laine, et al. 2016) to identify and order the Z-linked sequences among the investigated species. To calculate the read coverage, we first mapped the reads to the reference genomes using BWA-MEM (0.7.16a-r1181) with default parameters. We used the function 'depth' in samtools (1.9) to calculate coverage for every nucleotide site, subsequently removed those sites with mapping quality (-Q) lower than 60 or depth 3 times higher than average. Then we calculated genomic coverage of every 50 kb sliding window by using 'bedtools map' function. Any windows with less than 60% of the region (30 kb) mapped by reads were excluded. We used the GATK (3.8.0) pipeline (HaplotypeCaller) to call variants. Raw variants were filtered by this criteria: -window 10 -cluster 2 "FS > 10.0", "QD < 2.0", "MQ < 50.0", "SOR > 1.5", "MQRankSum < -1.5", " RedPosRankSum < -8.0". We expected the allele frequency to be 0.5 for one individual, thus further required the variants to show an allele frequency ranging between 0.3 and 0.7. The SNP density was defined by the number of SNPs over a 50 kb window. To genotype the W-derived alleles, we used the FastaAlternateReferenceMaker function of the GATK to create W-linked sequences for the transposed regions. The gene models on the W were then predicted by genewise (2.4.1). To remove potential chimeric W-derived alleles in the Z-linked regions (due to the collapse of genome assembly), if any, we used male sequencing reads to polish the Z-linked

sequence using pilon (1.22). To estimate pairwise substitution rated between sex-linked alleles, we used the guidance program (v2.02) and PRANK (170427) to align the Z- and W-linked coding sequences. Then we used the 'free ratio' model in codeml from PAML package (4.9e) to estimate the substitution rates. We used the program RSEM (1.3.0) to estimate gene expression levels. Details of the method is described in Xu et al. (2019). Codes used in this study has been deposited at Github (https://github.com/lurebgi/ZWtransposition). We measured the probability of haploinsufficiency of avian genes, with published HP scores (Huang, et al. 2010) for their human orthologs. Haploinsufficiency is defined as one single copy of genes is not sufficient to accomplish normal gene functions. Huang et al. predicted HP score for each human gene, based on known haploinsufficient genes identified from disease studies, and haplosufficient genes which show copy number variations among healthy human individuals.

Acknowledgment
We would like to thank Fumin Lei and Yalin Chen at the Institute of Zoology, Chinese Academy of Sciences for sharing their unpublished genomic data of the green-backed tit (*Parus monticolus*) for inferring the origination time of transposition in great tit. We are grateful to the Swedish Museum of Natural History (Ulf Johansson) to access to female and male samples of *Parus major*. L.X. is supported by the uni:docs fellowship programme from University of Vienna. Q.Z. is supported by the National Natural Science Foundation of China (31722050, 31671319), the Fundamental Research Funds for the Central Universities (2018XZZX002-04), and start-up funds from Zhejiang University. M.I. is supported by the Swedish Research Council (grant number 621-2014-5113).

Figure legend
Figure 1 Transpositions from the Z to W chromosomes in songbirds.
a) We show seven representative species out of the 12 studied songbirds, including the signatures of Z-to-W duplicative transpositions for three species. We labelled the phylogenetic node when the transposition occurred with red asterisks. b) For each of the three species, genomic regions on the Z chromosome showing female-specific elevations of SNP density (f/m SNP density) and read mapping coverage were inferred as recent transpositions, and were marked by red vertical bars. PAR and Z-linked duplications (marked in purple vertical bars) are not expected to show a female-specific elevation of SNP density level. c) We showed the SNP density of male and female calculated in 50kb windows at the Z-linked region that generated the transpositions, relative to the rest Z-linked regions, the PAR, and autosomes. SNP density is calculated as number of SNPs every 50kb window, and indicates levels of sequence divergence

between the Z- and W-linked homologous regions in female, or those between the two Z chromosomes in males, or between the two Z-linked duplications in both sexes.

Figure 2 The Z-to-W transposition in red bird-of-paradise. a) The loci of transposition (at ~60 Mb) on the Z chromosome shows an elevated heterozygosity and coverage in females. b) Since the transposition was found in two bird-of-paradise species, it was inferred to emerge before their speciation. c) A zoom-in view of the Z-to-W transposed region and d) the Z-linked duplication region. The Z-linked duplicate show a similar level of coverage and SNP density between sexes. e) The 1.3 Mb transposed sequence involves 8 genes, but 4 compete and 2 partial genes probably have become lost through a 583 kb sequence deletion, where the female coverage becomes lower than the rest transposed regions in c). Only *ANXA1* and *ALDH1A1* are retained on the W.

Figure 3 Female-specific and dosage selections restore avian W-linked genes. The seven restored functional genes through transposition on the W chromosomes tend to show a higher expression level or a broader (larger 1-tau value) expression pattern across tissues than the lost genes and putative pseudogenes. Most of restored genes also have a higher degree of dosage sensitivity (higher predicted haploinsufficiency scores) than the lost genes and putative pseudogenes, with some genes (e.g., *ANXA1*) showing an ovary-biased expression pattern.

Fig. 1 Transpositions from the Z to W chromosomes in songbirds.

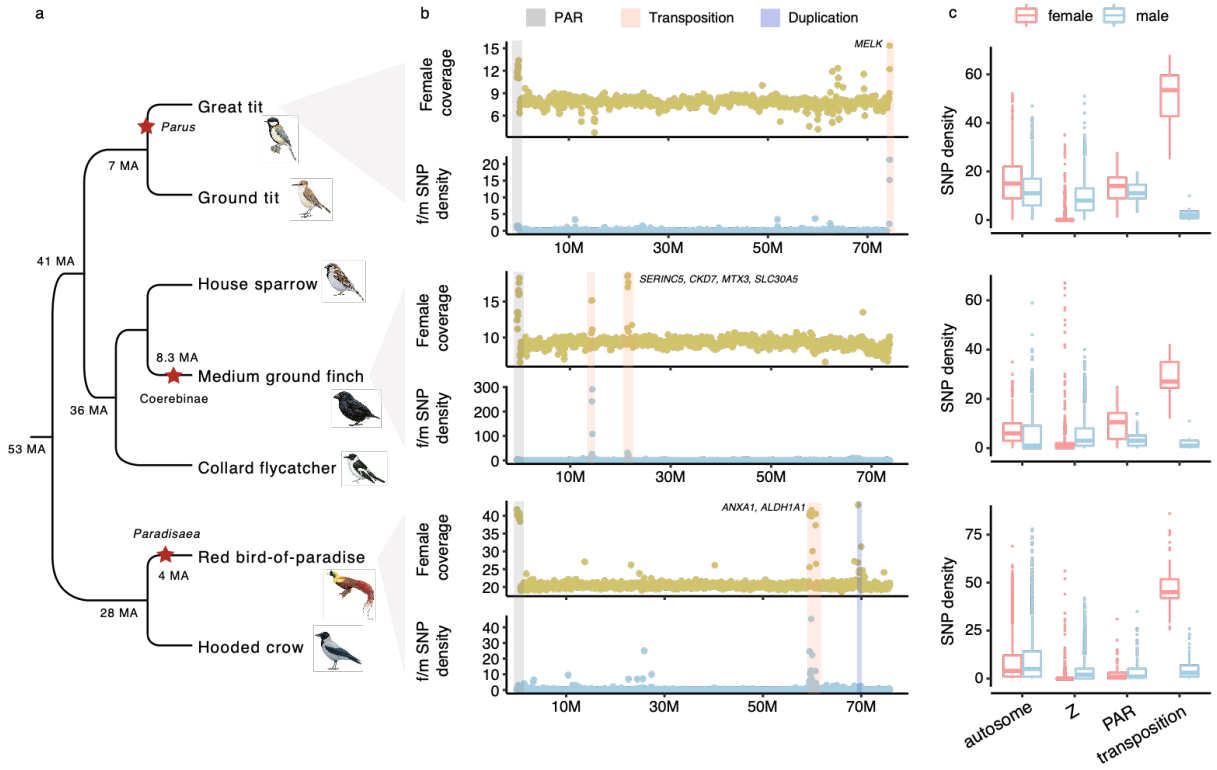


Fig. 2 The Z-to-W transposition in red bird-of-paradise.

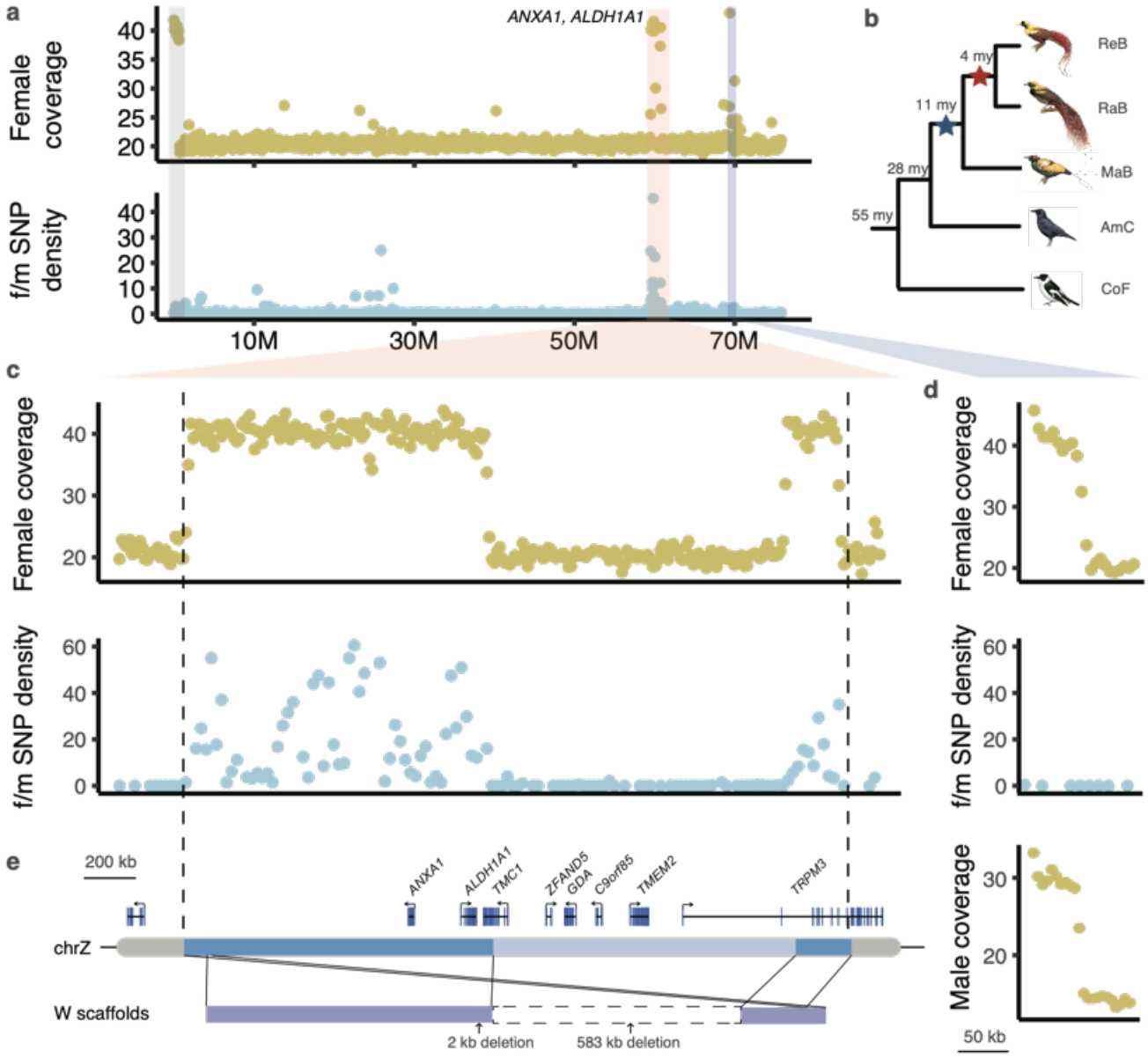
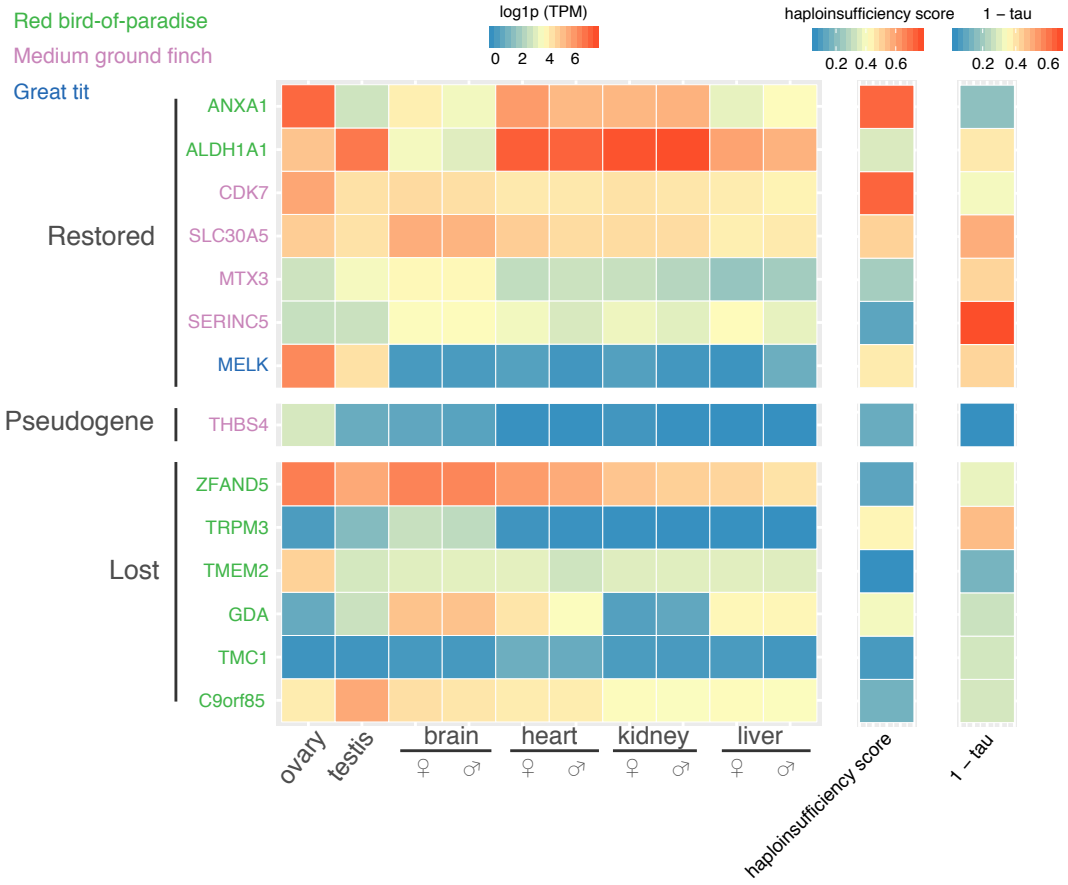


Fig. 3 Female-specific and dosage selections restore avian W-linked genes.



References

Bachtrog D. (WXU0 co-authors). 2013. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. Nat. Rev. Genet. 14:113-124.

Bellott DW, Hughes JF, Skaletsky H, Brown LG, Pyntikova T, Cho T-J, Koutseva N, Zaghlul S, Graves T, Rock S, et al. (noCG co-authors). 2014. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. Nature 508:494-499.

Bellott DW, Skaletsky H, Cho T-J, Brown L, Locke D, Chen N, Galkina S, Pyntikova T, Koutseva N, Graves T, et al. (mR99 co-authors). 2017. Avian W and mammalian Y chromosomes convergently retained dosage-sensitive regulators. Nat. Genet. 49:387-394.

Bowles J, Feng C-W, Miles K, Ineson J, Spiller C, Koopman P. (jhpQ co-authors). 2016. ALDH1A1 provides a source of meiosis-inducing retinoic acid in mouse fetal ovaries. Nat. Commun. 7:10845.

Brashear WA, Raudsepp T, Murphy WJ. (sISi co-authors). 2018. Evolutionary conservation of Y Chromosome ampliconic gene families despite extensive structural variation. Genome Res. 28:1841-1851.

Carvalho AB, Vicoso B, Russo CAM, Swenor B, Clark AG. (GMwb co-authors). 2015. Birth of a new gene on the Y chromosome of Drosophila melanogaster. Proc. Natl. Acad. Sci. 112:12450-12455.

Charlesworth B, Charlesworth D. (vUzJ co-authors). 2000. The degeneration of Y chromosomes. Philos. Trans. R. Soc. Lond. B Biol. Sci. 355:1563-1572.

Cortez D, Marin R, Toledo-Flores D, Froidevaux L, Liechti A, Waters PD, Grützner F, Kaessmann H. (TsSo co-authors). 2014. Origins and functional evolution of Y chromosomes across mammals. Nature 508:488-493.

Crow TJ. (Ysmz co-authors). 2002. Handedness, language lateralisation and anatomical asymmetry: relevance of protocadherin XY to hominid speciation and the aetiology of psychosis. Point of view. Br. J. Psychiatry 181:295-297.

Davis JK, Thomas PJ, Program NCS, Thomas JW. (TpKL co-authors). 2010. A W-linked palindrome and gene conversion in New World sparrows and blackbirds. Chromosome Res. 18:543-553.

Graves JAM. (ugPH co-authors). 2014. Avian sex, sex chromosomes, and dosage compensation in the age of genomics. Chromosome Res. 22:45-57.

Gu L, Walters JR. (JKg7 co-authors). 2017. Evolution of sex chromosome dosage compensation in animals: a beautiful theory, undermined by facts and bedeviled by details. Genome Biol. Evol. 9:2461-2476.

Huang N, Lee I, Marcotte EM, Hurles ME. (90KG co-authors). 2010. Characterising and predicting haploinsufficiency in the human genome. PLoS Genet. 6:e1001154.

Hughes JF, Skaletsky H, Koutseva N, Pyntikova T, Page DC. (MrGt co-authors). 2015. Sex chromosome-to-autosome transposition events counter Y-chromosome gene loss in mammals. Genome Biol. 16:104.

International Chicken Genome Sequencing C. (2MFK co-authors). 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432:695-716.

Itoh Y, Melamed E, Yang X, Kampf K, Wang S, Yehya N, Van Nas A, Replogle K, Band MR, Clayton DF, et al. (Q1uV co-authors). 2007. Dosage compensation is less effective in birds than in mammals. J. Biol. 6:2.

Janečka JE, Davis BW, Ghosh S, Paria N, Das PJ, Orlando L, Schubert M, Nielsen MK, Stout TAE, Brashear W, et al. (Klw6 co-authors). 2018. Horse Y chromosome assembly displays unique evolutionary features and putative stallion fertility genes. Nat. Commun. 9.

Koerich LB, Wang X, Clark AG, Carvalho AB. (pL5C co-authors). 2008. Low conservation of gene content in the Drosophila Y chromosome. Nature 456:949-951.

Laine VN, Gossmann TI, Schachtschneider KM, Garroway CJ, Madsen O, Verhoeven KJF, de Jager V, Megens H-J, Warren WC, Minx P, et al. (3yE9 co-authors). 2016. Evolutionary signals of selection on cognition from the great tit genome and methylome. Nat. Commun. 7:10474.
Lamichhaney S, Berglund J, Almén MS, Maqbool K, Grabherr M, Martinez-Barrio A, Promerová M, Rubin C-J, Wang C, Zamani N, et al. (1zR7 co-authors). 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature 518:371-375.
Li G, Davis BW, Raudsepp T, Pearks Wilkerson AJ, Mason VC, Ferguson-Smith M, O'Brien PC, Waters PD, Murphy WJ. (8hnY co-authors). 2013. Comparative analysis of mammalian Y chromosomes illuminates ancestral structure and lineage-specific evolution. Genome Res. 23:1486-1495.
Mahajan S, Bachtrog D. (s5jZ co-authors). 2017. Convergent evolution of Y chromosome gene content in flies. Nat. Commun. 8.
Moghadam HK, Pointer MA, Wright AE, Berlin S, Mank JE. (P97h co-authors). 2012. W chromosome expression responds to female-specific selection. Proc. Natl. Acad. Sci. U. S. A. 109:8207-8211.
Page DC, Harper ME, Love J, Botstein D. (DyNx co-authors). 1984. Occurrence of a transposition from the X-chromosome long arm to the Y-chromosome short arm during human evolution. Nature 311:119-123.
Ponnikas S, Sigeman H, Abbott JK, Hansson B. (mFlx co-authors). 2018. Why do sex chromosomes stop recombining? Trends Genet. 34:492-503.
Ross MT, Grafham DV, Coffey AJ, Scherer S, McLay K, Muzny D, Platzer M, Howell GR, Burrows C, Bird CP, et al. (8n41 co-authors). 2005. The DNA sequence of the human X chromosome. Nature 434:325-337.
Rozen S, Skaletsky H, Marszalek JD, Minx PJ, Cordum HS, Waterston RH, Wilson RK, Page DC. (bMEr co-authors). 2003. Abundant gene conversion between arms of palindromes in human and ape Y chromosomes. Nature 423:873-876.
Schwartz A, Chan DC, Brown LG, Alagappan R, Pettay D, Disteche C, McGillivray B, de la Chapelle A, Page DC. (yjaR co-authors). 1998. Reconstructing hominid Y evolution: X-homologous block, created by X-Y transposition, was disrupted by Yp inversion through LINE--LINE recombination. Hum Mol Genet. 7:1-11.
Skaletsky H, Kuroda-Kawaguchi T, Minx PJ, Cordum HS, Hillier L, Brown LG, Repping S, Pyntikova T, Ali J, Bieri T, et al. (GmFT co-authors). 2003. The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. Nature 423:825-837.
Smeds L, Warmuth V, Bolivar P, Uebbing S, Burri R, Suh A, Nater A, Bureš S, Garamszegi LZ, Hogner S, et al. (ZSEn co-authors). 2015. Evolutionary analysis of the female-specific avian W chromosome. Nat. Commun. 6:7330.
Speevak MD, Farrell SA. (0MBN co-authors). 2011. Non-syndromic language delay in a child with disruption in the Protocadherin11X/Y gene pair. Am J Med Genet B Neuropsychiatr Genet. 156:484-489.
Suh A. (trGu co-authors). 2015. The Specific Requirements for CR1 Retrotransposition Explain the Scarcity of Retrogenes in Birds. J. Mol. Evol. 81:18-20.
Tobler R, Nolte V, Schlötterer C. (gh8f co-authors). 2017. High rate of translocation-based gene birth on the Drosophila Y chromosome. Proc. Natl. Acad. Sci. 114:11721-11726.
Xu L, Auer G, Peona V, Suh A, Deng Y, Feng S, Zhang G, Blom MPK, Christidis L, Prost S, et al. (nfad co-authors). 2019. Dynamic evolutionary history and gene content of sex chromosomes across diverse songbirds. Nat Ecol Evol 3:834-844.
Zhang Z, Harrison PM, Liu Y, Gerstein M. 2003. Millions of years of evolution preserved: a comprehensive catalog of the processed pseudogenes in the human genome. Genome Res 13:2541-2558.

Zhou Q, Zhang J, Bachtrog D, An N, Huang Q, Jarvis ED, Gilbert MTP, Zhang G. (WPBI co-authors). 2014. Complex evolutionary trajectories of sex chromosomes across bird taxa. Science 346:1246338.

Supplementary material

Supplementary Note

Bioinformatic verification of the Z-to-W Transpositions

In red bird-of-paradise, the entire transposed sequence is located in a single scaffold (scaffold_234). This scaffold shows strong synteny relationship with Z-linked sequence of other birds (Supplementary Fig. S7). Moreover, this scaffold (the transposed and retained part) shows a female-specific increase in SNP density (heterozygosity). This suggests this scaffold has not been translocated to autosomes, as it would predict equal heterozygosity in males and females. Similarly, we show the transposed sequence (scaffold NW_005054440.1, NW_005054526.1 and NW_005055028.1) in medium ground finch is Z-linked as the transposed sequence, as well as the flanking non-transposed sequences of those scaffolds, have good synteny with the Z chromosomes of other birds (Supplementary Fig. S7). The Z chromosome assembly in great tit is supported by linkage map (Laine et al. 2016), therefore the Z-linked of the transposed region is supported. Moreover, the transposed region show a similar location in the Z chromosome of collared flycatcher (Supplementary Fig. S7).

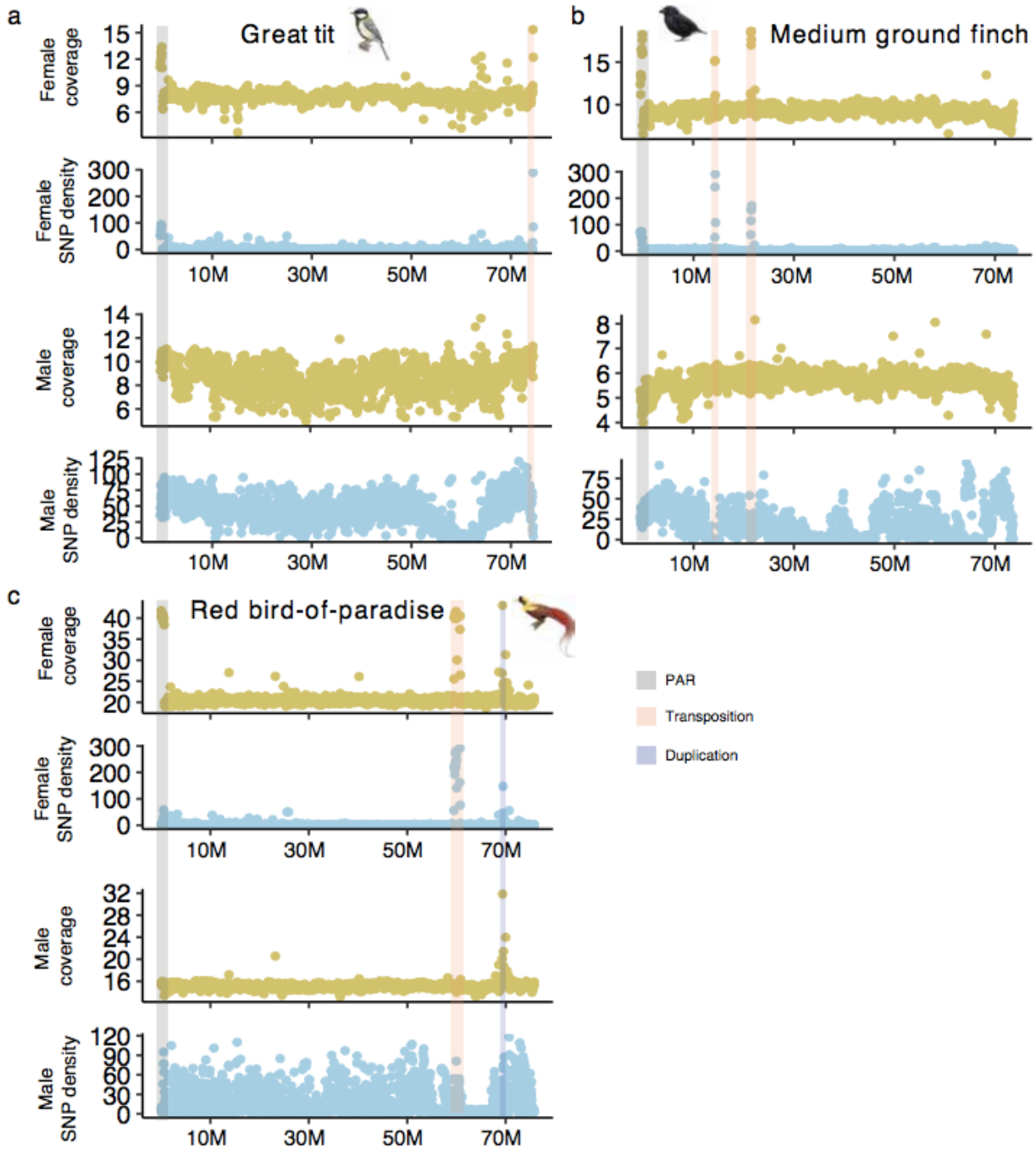


Fig. S1 Identifying Z-to-W transpositions from coverage and SNP patterns. The transposed regions show substantially increased SNP density in female, but not in male, and the coverage level in female become twice of that of other Z-linked sequences. The patterns of coverage and SNP density are different for Z-linked duplications in which, as seen in red bird-of-paradise, male coverage also become twice of the rest Z-linked regions.

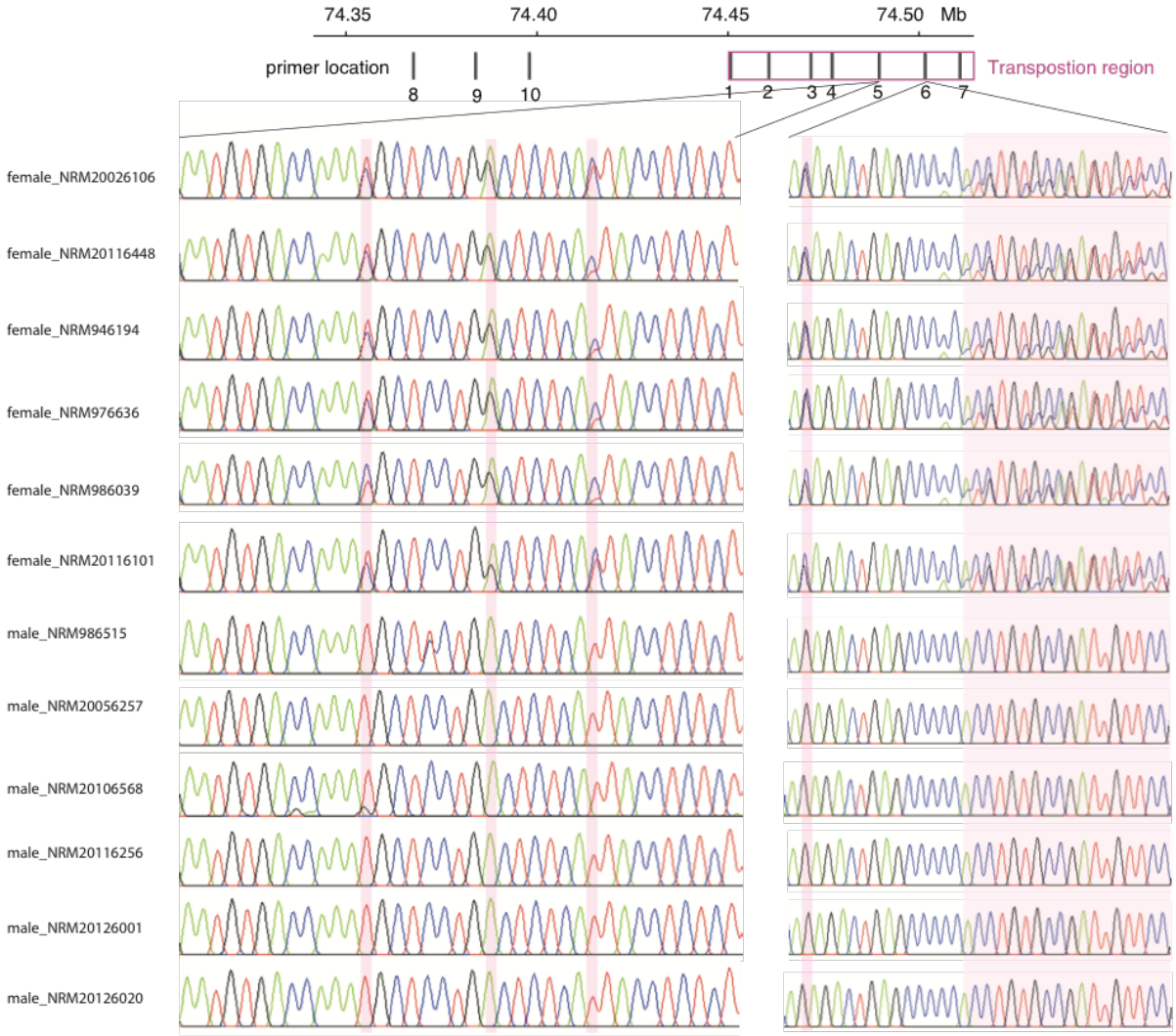


Fig. S2. Verification of female-specific SNPs in the Z-transposed regions in great tit. The fragments have been amplified from 6 females and 6 males, with 7 of them at the Z-transposed regions (ZTR) and 3 near the ZTR. One region from the product 5 shows three female-specific SNPs, and one region of the product 6 shows one female-specific SNP and signals of female-specific indels.

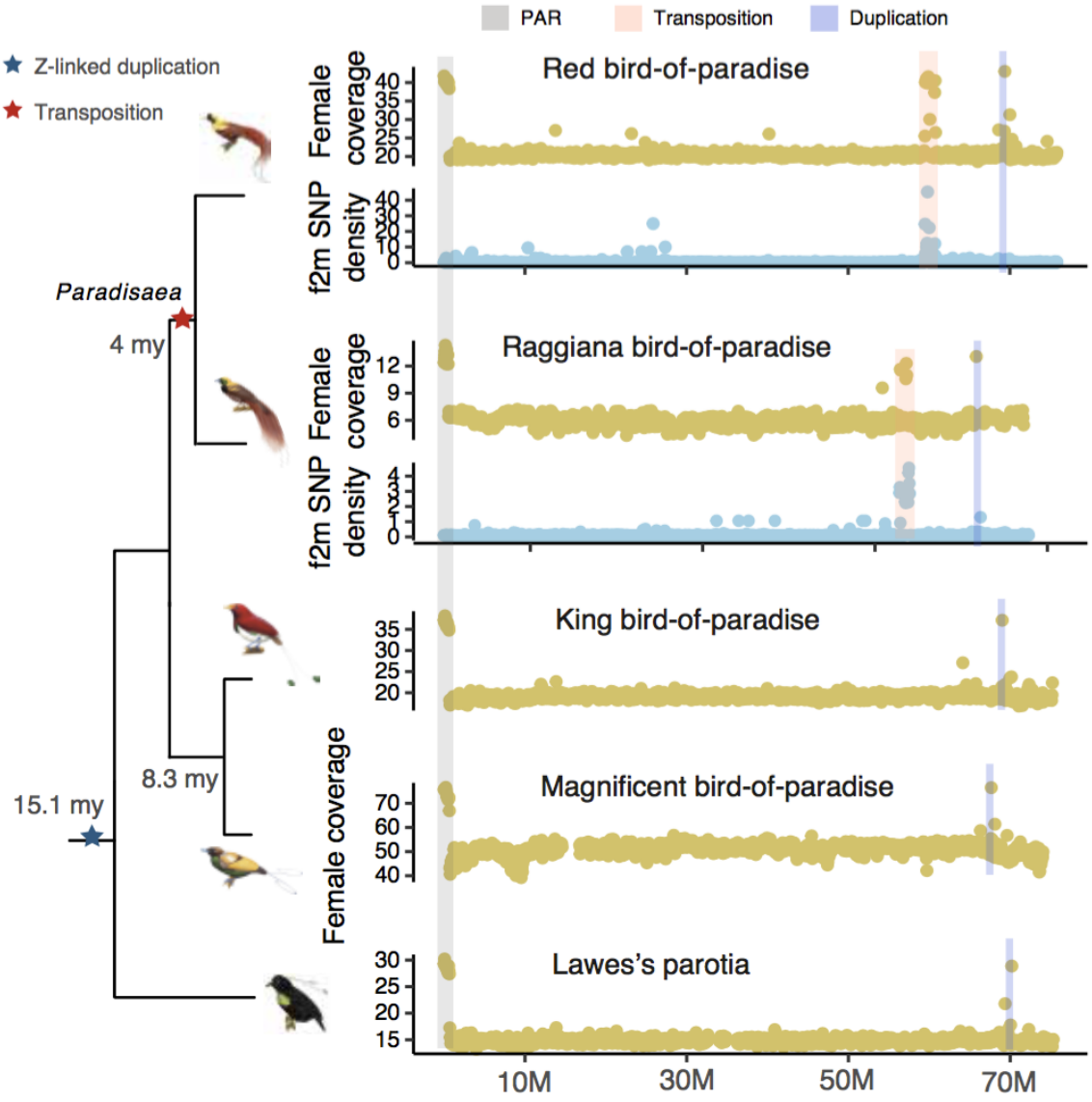


Fig. S3 Z-to-W transposition of red bird-of-paradise originated at the ancestor of *Paradisaea*. Signals of Z-to-W transpositions were examined in five genomes of birds-of-paradise. The 1.3 Mb transposition discovered in red bird-of-paradise and Raggiana bird-of-paradise are homologous, but are not found in the other species. The origin of the transposition is therefore about 4 my ago when the *Paradisaea* lineage diverged from other birds-of-paradise. The duplication of a ~50k sequence is shared by all species.

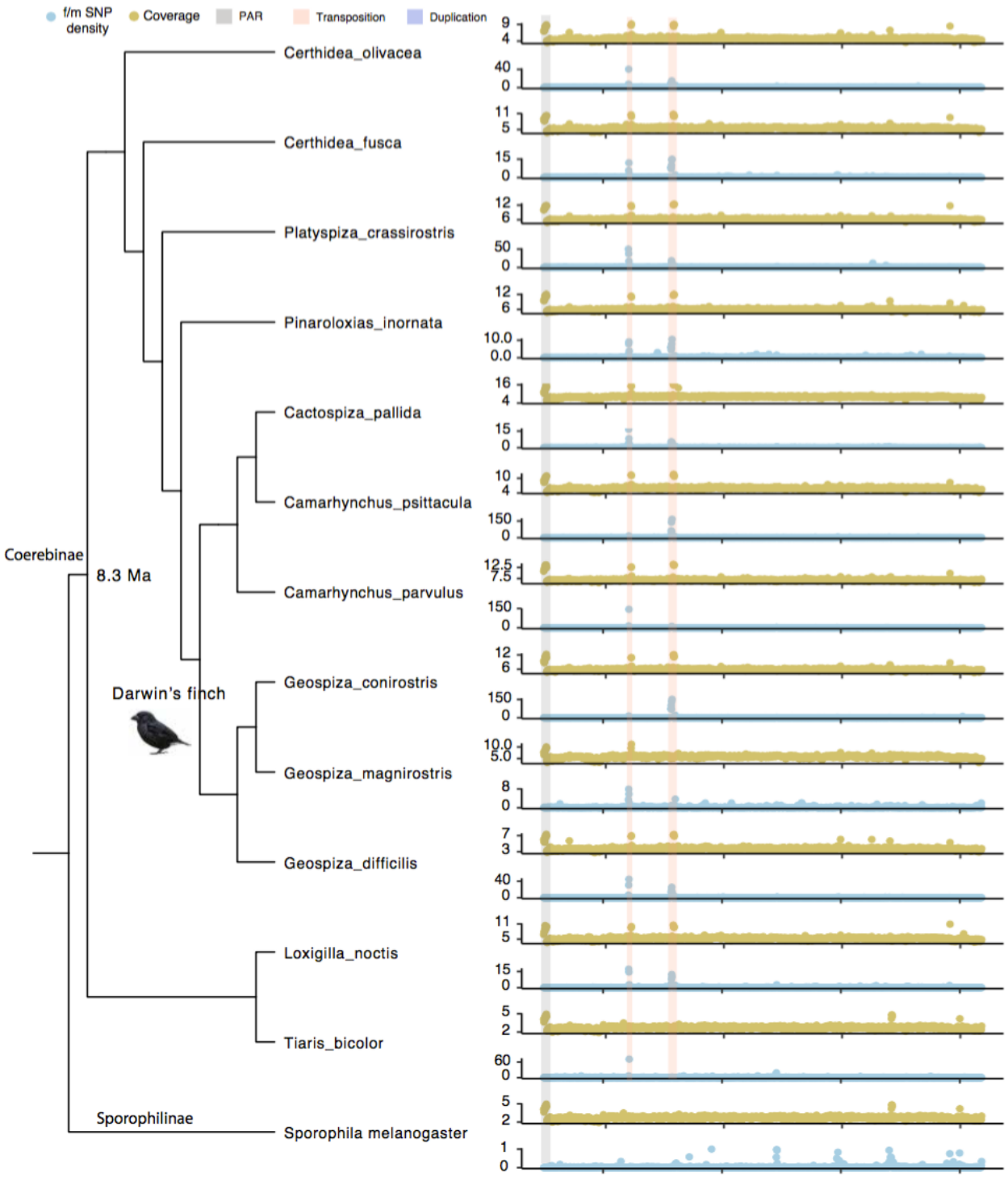


Fig. S4 Shared Z-to-W transposition across Darwin's finches and their close relatives. We used the re-sequencing data of 12 Coerebinae (including 6 Darwin's finches) and one Sporaphilinae species (sister group to Coerebinae) to screen for signals of transpositions. Reads were mapped against the genome of medium ground finch. All Coerebinae species share the two transpositions as seen in medium ground finch.

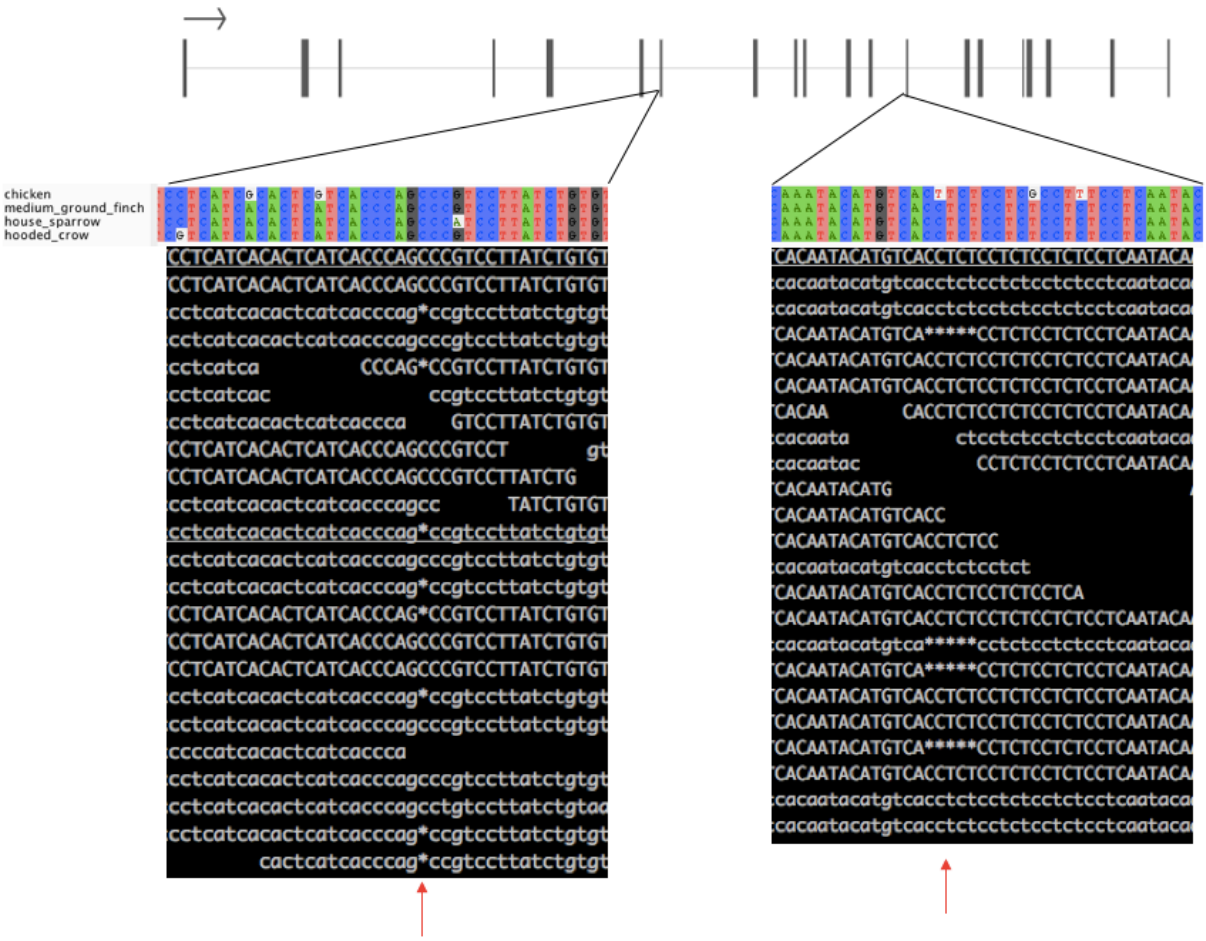


Fig. S5. Two frame shift mutations of *THBS4* on the W-linked transposed sequence. The panel in the middle shows the sequence alignments of four birds, and panels below show the alignment of female reads ('samtools tview' visualisation). A deletion of one basepair at the 7th exon and a deletion of 5 basepairs at the 13th exon have probably disrupted the open reading frame of the coding sequences.

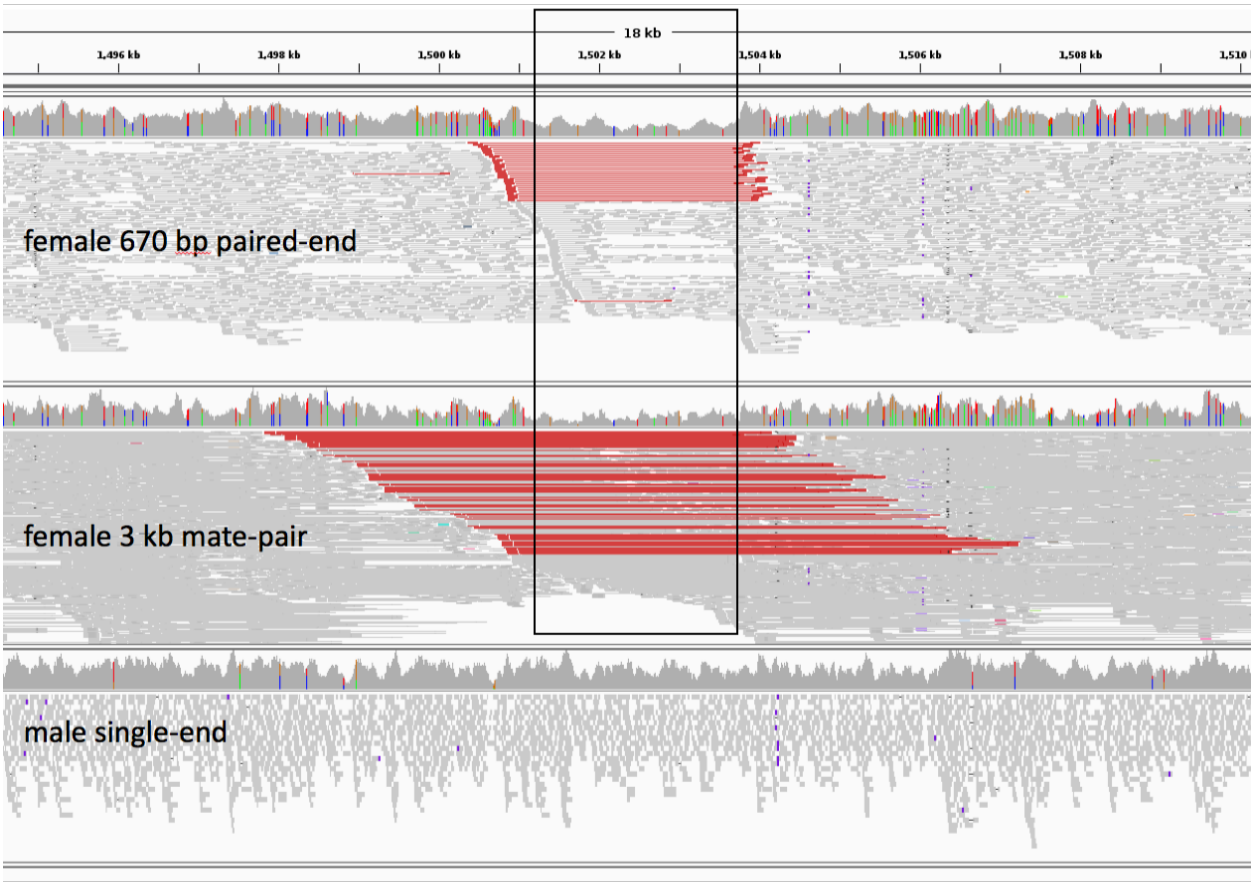


Fig. S6 A W-linked 2-kb deletion in red bird-of-paradise transposed region. We showed paired-end read pairs with an insert size of 670bp, and mate-pair read pairs with an insert size of 3kb as color-coded lines. Red lines indicated a deletion with a length about 2kb: for example, in the 670bp track, the red read pairs spanned a region over 2kb long. The grey peaks above the read pair tracks showed the heterozygous sites with colored vertical lines, and the heights of peaks indicated the read coverage. At the deleted region, there are reduced levels of heterozygosity and read coverage.

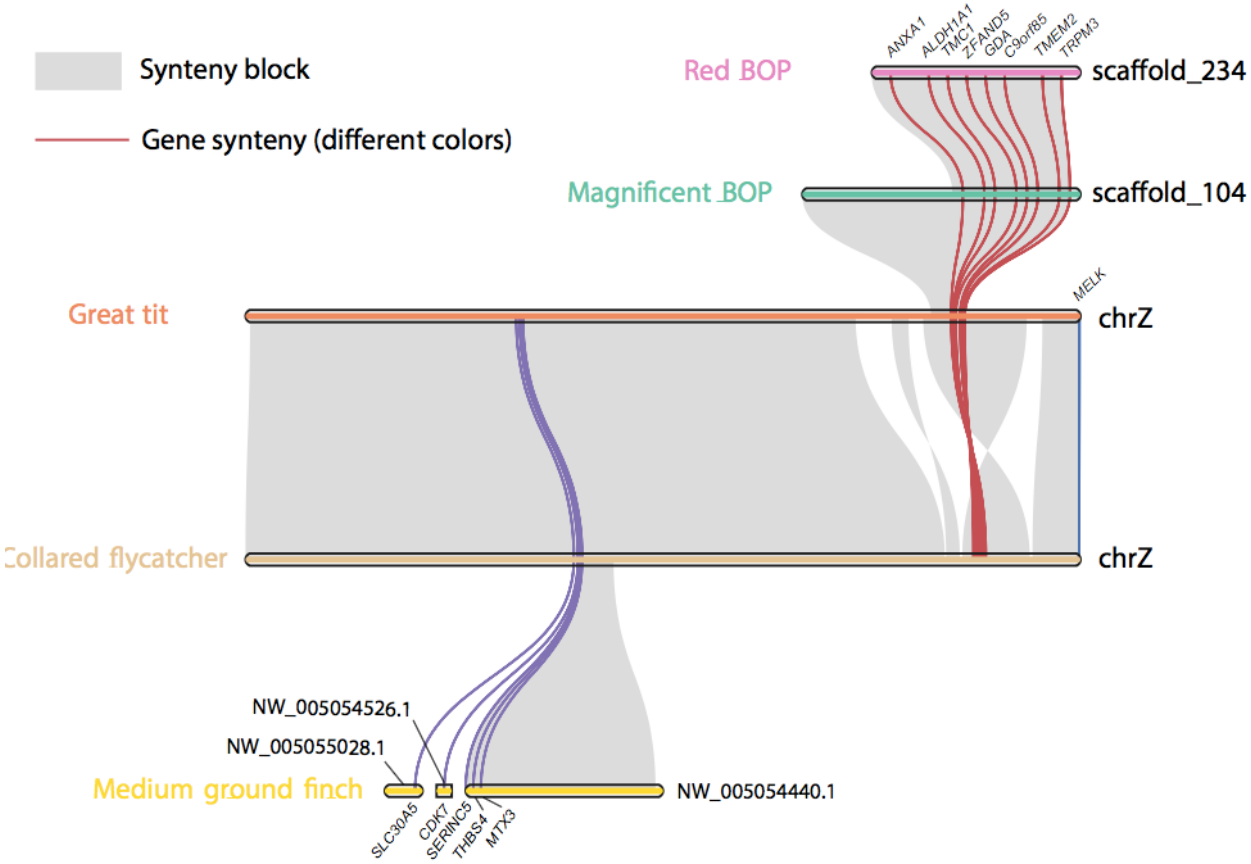


Fig. S7 Synteny of transposed sequences. Great tit and collared flycatcher have a chromosome assembly. Synteny blocks are shown in grey, while the synteny of individual gene is shown by a connecting line. For red bird-of-paradise (BOP) and medium ground finch, we show the synteny of the scaffolds containing transposed sequences to the Z chromosomes of great tit and collared flycatcher. The synteny is analyzed with MCscan (a jcvi tool).

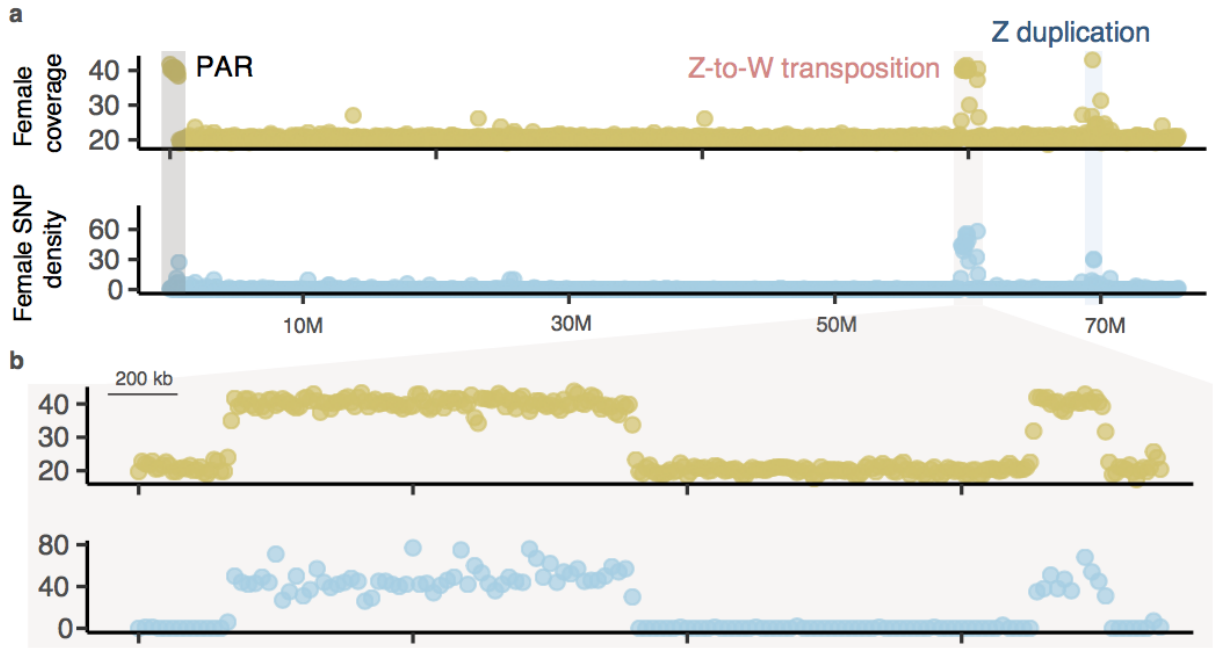


Fig. S8 Similar heterozygosity levels in two separate blocks of Z-transposed region in red bird-of-paradise. The ZTR (1.3Mb) is separated by a 583kb sequence deletion, and the two separated region show a similar level of female heterozygosity, suggesting a similar age of the transposition. This supports the scenario of one single transposition event followed by sequence deletions instead of two independent transpositions.

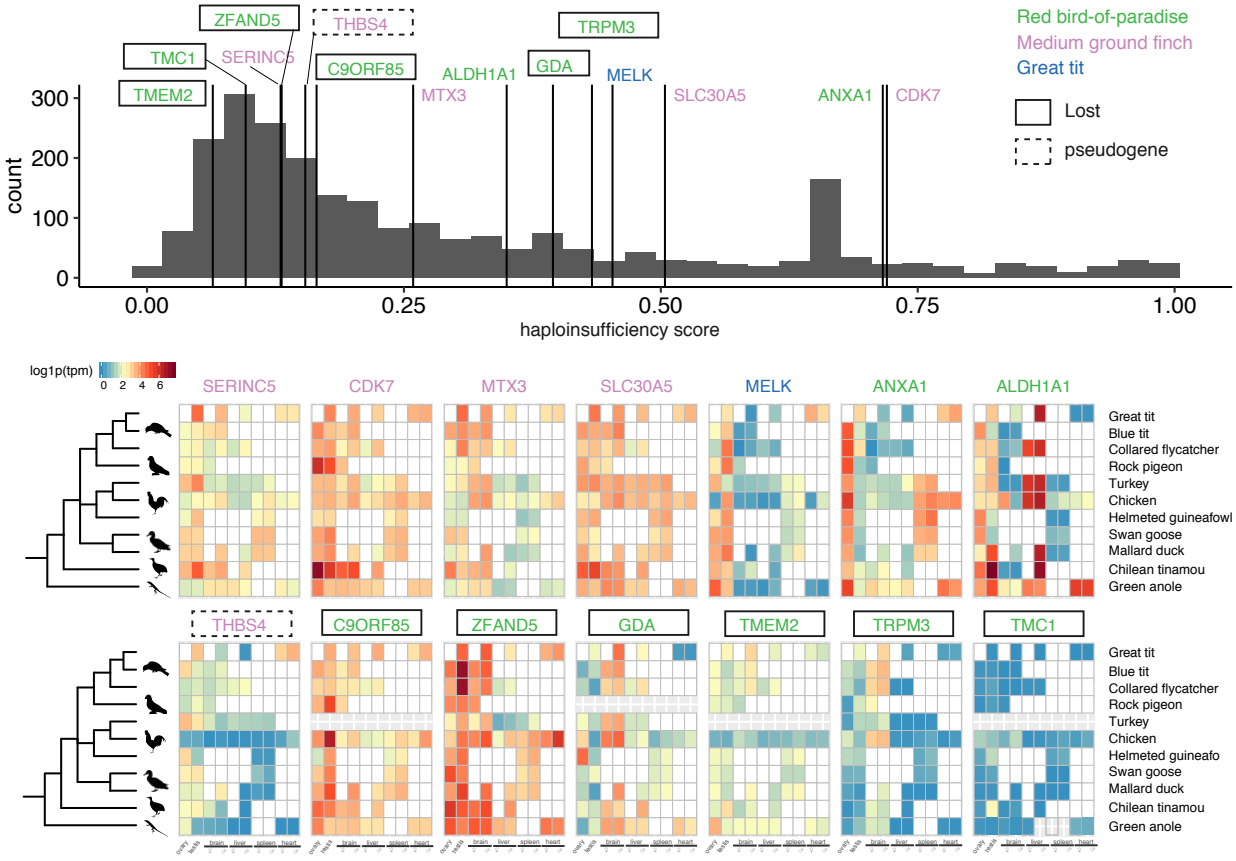


Fig. S9 Gene expression patterns and haploinsufficiency of transposed genes. The upper panel shows the distribution of haploinsufficiency scores of Z-linked genes. The vertical lines indicate haploinsufficiency scores of transposed genes, included those that are already lost (in box) or become pseudogene (dashed box). The lower panel shows the color-coded expression levels (log1p transformed TPM, transcripts per million) of transposed genes in 10 birds and green anole. The blank tiles indicate the data that is not available. *ANXA1* shows biased expression in ovaries across species.

Reference

Laine VN, Gossmann TI, Schachtschneider KM, Garroway CJ, Madsen O, Verhoeven KJF, de Jager V, Megens H-J, Warren WC, Minx P, et al. 2016. Evolutionary signals of selection on cognition from the great tit genome and methylome. Nature Communications 7:10474.

Table S-1 Divergence rate of the transposed genes measured by pairwise synonymous (dS) and nonnsynonymous (dN) substitution rates between Z/W gametologs

Gene	Species	Z			W			Status
		dS	dN	dN/dS	dS	dN	dN/dS	
ANXA1	Red bird-of-paradise	0.003995	0	0.0001	0.003992	0	0.0001	Retainec
ALDH1A1	Red bird-of-paradise	0.010218	0.000001	0.0001	0.005625	0.000888	0.15792	Retainec
MELK	Great tit	0.007387	0.000751	0.101655	0.005241	0.001507	0.287433	Retainec
SLC30A5	Medium ground finch	0.001119	0	0.0001	0.000621	0	0.0001	Retainec
CDK7	Medium ground finch	0.000005	0	0.0001	0	0.000002	66.4915	Retainec
SERINC5	Medium ground finch	0.000005	0	0.0001	0.002758	0	0.0001	Retainec
MTX3	Medium ground finch	0.003784	0.001664	0.439907	0.008356	0.000001	0.0001	Retainec
THBS4	Medium ground finch	0.004096	0.000528	0.128935	0.002846	0.001047	0.367974	Pseudog ene

Table S2 Accessions of NCBI genome assemblies

Species	Common name	Assembly	Sex	Reference
Pseudopodoces humilis	Tibetan ground tit	GCF_000331425.1	Female	[1,2]
Geospiza fortis	Medium ground finch	GCF_000277835.1	Female	[3]
Passer domesticus	House sparrow	GCA_001700915.1	Female	[4]
Corvus brachyrhynchos	American crow	GCF_000691975.1	Female	[3]
Corvus cornix	Hooded crow	GCF_000738735.1	Male	[6]
Ficedula albicollis	Collared flycatcher	GCF_000247815.1, GCA_900067835.1	Female	[7]
Parus major	Great tit	GCF_001522545.2	Male	[9]
Parotia lawesii	Lawes's Parotio	GCA_003713295.1	Female	[8]
Cicinnurus magnificus	Magnificent BOP	GCA_003713285.1	Female	[8]
Paradisaea raggiana	Raggiana BOP	GCA_003713265.1	Female	[8]
Cicinnurus regius	King BOP	GCA_003713305.1	Female	[8]
Paradisaea rubra	Red BOP	GCA_003713215.1	Female	[8]

1 Qu, Y. *et al.* Ground tit genome reveals avian adaptation to living at high altitudes in the Tibetan plateau. *Nature Communications* **4,** 2071 (2013).
2 Cai, Q. *et al.* Genome sequence of ground tit Pseudopodoces humilis and its adaptation to high altitude. *Genome Biol* **14,** R29 (2013).
3 Zhang, G. *et al.* Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346,** 1311–1320 (2014).
4 Elgvin, T. O. *et al.* The genomic mosaicism of hybrid speciation. *Sci Adv* **3,** e1602996 (2017).
5 Frankl-Vilches, C. Using the canary genome to decipher the evolution of hormone-sensitive gene regulation in seasonal singing birds. *Genome Biol* **16,** 1–25 (2015).
6 Poelstra, J. W. *et al.* The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* **344,** 1410–1414 (2014).
7 Ellegren, H. *et al.* The genomic landscape of species divergence in Ficedula flycatchers. *Nature* **491,** 756–760 (2013).
8 Xu, L. *et al.* Dynamic evolutionary history and gene content of sex chromosomes across diverse songbirds. Nat Ecol Evol **3**:834–844 (2019).
9 Laine, V. N. *et al.* Evolutionary signals of selection on cognition from the great tit genome and methylome. *Nature Communications* **7,** 10474 (2016).

Table S3 Resequencing data of related species

Taxa	Species	SRA	Sex	Reference
Coerebinae (Darwin's finch)	*Cactospiza_pallida*	SRR1607494	M	1
Coerebinae (Darwin's finch)	*Cactospiza_pallida*	SRR1607498	F	1
Coerebinae (Darwin's finch)	*Camarhynchus_parvulus*	SRR1607504	M	1
Coerebinae (Darwin's finch)	*Camarhynchus_parvulus*	SRR1607506	F	1
Coerebinae (Darwin's finch)	*Camarhynchus_psittacula*	SRR1607543	F	1
Coerebinae (Darwin's finch)	*Camarhynchus_psittacula*	SRR1607545	M	1
Coerebinae (Darwin's finch)	*Geospiza_conirostris*	SRR1607300	M	1
Coerebinae (Darwin's finch)	*Geospiza_conirostris*	SRR1607318	F	1
Coerebinae (Darwin's finch)	*Geospiza_difficilis*	SRR1607400	F	1
Coerebinae (Darwin's finch)	*Geospiza_difficilis*	SRR1607403	M	1
Coerebinae (Darwin's finch)	*Geospiza_magnirostris*	SRR1607485	M	1
Coerebinae (Darwin's finch)	*Geospiza_magnirostris*	SRR1607488	F	1
Coerebinae (Darwin's finch)	*Pinaroloxias_inornata*	SRR1607512	F	1
Coerebinae (Darwin's finch)	*Pinaroloxias_inornata*	SRR1607514	M	1
Coerebinae (Darwin's finch)	*Platyspiza_crassirostris*	SRR1607532	M	1
Coerebinae (Darwin's finch)	*Platyspiza_crassirostris*	SRR1607541	F	1
Coerebinae	*Tiaris_bicolor*	SRR1607551	F	1
Coerebinae	*Tiaris_bicolor*	SRR1607554	M	1
Coerebinae	*Loxigilla_noctis*	SRR1607474	F	1
Coerebinae	*Loxigilla_noctis*	SRR1607478	M	1
Coerebinae	*Certhidea_olivacea*	SRR1607385	F	1
Coerebinae	*Certhidea_olivacea*	SRR1607390	M	1
Coerebinae	*Certhidea_fusca*	SRR1607327	M	1
Coerebinae	*Certhidea_fusca*	SRR1607330	F	1
Sporophilinae	*Sporophila melanogaster*	SRR5447379	F	2

1 Lamichhaney, Sangeet, et al. "Evolution of Darwin's finches and their beaks revealed by genome sequencing." *Nature*518.7539 (2015): 371.
2 Campagna, Leonardo, et al. "Repeated divergent selection on pigmentation genes in a rapid finch radiation." *Science advances* 3.5 (2017): e1602404.

Table S4 RNA-seq datasets analyzed in this study

Species	Comman name	SRA	Reference
Parus major	Great tit	SRR1847223, SRR1847228, SRR1847415; SRR2170826, SRR2170832	1, 2
Cyanistes caeruleus	Blue tit	PRJNA284903	3
Ficedula albicollis	collared flycatcher	PRJEB2984	4
Columba livia	rock pigeon	PRJEB16136, PRJNA427400	5, 6
Meleagris gallopavo	turkey	PRJNA271731, PRJNA259229	7,8
Gallus gallus	chicken	PRJEB8390, PRJNA381064, PRJNA171809, PRJNA171809	9, 10, 11, 12
Numida meleagris	helmeted guineafowl	PRJNA271731	7
Anser cygnoides	swan goose	PRJNA271731	7
Anas platyrhynchos	mallard duck	PRJNA419583, PRJNA271731	7, 13
Nothoprocta perdicaria	Chilean tinamou	PRJNA433114	14
Anolis carolinensis	great anole	PRJNA78917	15

1 Laine, Veronika N., et al. "Evolutionary signals of selection on cognition from the great tit genome and methylome." *Nature Communications* 7 (2016): 10474.

2 Qu, Yanhua, et al. "Ground tit genome reveals avian adaptation to living at high altitudes in the Tibetan plateau." *Nature communications* 4 (2013): 2071.

3 Mueller, Jakob C., et al. "Characterization of the genome and transcriptome of the blue tit C yanistes caeruleus: polymorphisms, sex-biased expression and selection signals." *Molecular ecology resources* 16.2 (2016): 549-561.

4 Uebbing, Severin, et al. "Divergence in gene expression within and between two closely related flycatcher species." *Molecular ecology* 25.9 (2016): 2015-2028.

5 MacManes, Matthew D., et al. "Widespread patterns of sexually dimorphic gene expression in an avian hypothalamic–pituitary–gonadal (HPG) axis." *Scientific reports* 7 (2017): 45125.

6 Gazda, Małgorzata A., et al. "Signatures of selection on standing genetic variation underlie athletic and navigational performance in racing pigeons." *Molecular biology and evolution* 35.5 (2018): 1176-1189.

7 Harrison, Peter W., et al. "Sexual selection drives evolution and rapid turnover of male gene expression." *Proceedings of the National Academy of Sciences* 112.14 (2015): 4393-4398.

8 Dalloul, Rami A., et al. "Multi-platform next-generation sequencing of the domestic turkey (Meleagris gallopavo): genome assembly and analysis." *PLoS biology* 8.9 (2010): e1000475.

9 Uebbing, Severin, et al. "Quantitative mass spectrometry reveals partial translational regulation for dosage compensation in chicken." *Molecular biology and evolution*32.10 (2015): 2716-2725.

10 Marin, Ray, et al. "Convergent origination of a Drosophila-like dosage compensation mechanism in a reptile lineage." *Genome research* 27.12 (2017): 1974-1987.

11 Ayers, Katie L., et al. "RNA sequencing reveals sexually dimorphic gene expression before gonadal differentiation in chicken and allows comprehensive annotation of the W-chromosome." *Genome biology* 14.3 (2013): R26.

12 Ayers, Katie L., et al. "Identification of candidate gonadal sex differentiation genes in the chicken embryo using RNA-seq." *BMC genomics* 16.1 (2015): 704.
13 Zhang, Zebin, et al. "Whole-genome resequencing reveals signatures of selection and timing of duck domestication." *GigaScience* 7.4 (2018): giy027.
14 Xu, Luohao, et al. "Evolutionary dynamics of sex chromosomes of palaeognathous birds." *bioRxiv* (2018): 295089.
15 Alföldi, Jessica, et al. "The genome of the green anole lizard and a comparative analysis with birds and mammals." *Nature*477.7366 (2011): 587.

Discussion

The avian chromosomes are generally stable with few interchromosomal rearrangements, so as the sex chromosomes (Zhang et al. 2014; Zhang 2018). As the recent efforts in the characterization of chromosome evolution in the bird lineage revealed frequently chromosomal rearrangements, mostly interchromosomal (Damas et al. 2018), my evolutionary genomic study also revealed that the evolution of sex chromosomes in birds is more dynamic than previously thought (Xu, Wa Sin, et al. 2019; Xu, Auer, et al. 2019; Xu and Zhou 2019). First, we showed the disparity of the rate of sex chromosome differentiation in two major clades of birds: Palaeognathae and Neognathae. This disparity further extends to within the paleognathous lineage. Second, During the course of sex chromosome differentiation, the pace of evolutionary changes varies at different time points, and different evolutionary forces may at play. Third, we identified multiple occasions of gene acquisitions on the W chromosomes of birds, suggesting the avian W chromosome can also be evolutionarily active, likely through female-specific selection or selection for dosage-sensitive genes. Below I will elaborate on each argument.

Palaeognathae versus Neognathae

Paleognaths are thought to be slow-evolving and maintain many ancestral features of birds (Yonezawa et al. 2017), including the karyotypes (Damas et al. 2018). This is in part in line with the primordial status of sex chromosome evolution in paleognaths. The chromosomal inversion is one of the main mechanisms of recombination suppression, and in neognaths, a large inversion involving ~20 Mb sequence is probably the direct cause of the formation of the second stratum (Zhou et al. 2014). Paleognaths did not undergo this inversion; though many paleognaths independently evolved a second stratum, they are much smaller than that of neognaths. This suggests the differentiation of sex chromosomes in paleognaths is not completely halted, but at a much slower rate and to a much smaller scale.

Moreover, within paleognaths, the tinamou lineage shows an accelerated rate of sex chromosome evolution, relative to the rest of paleognaths, ratites (Zhou et al. 2014; Wang et al. 2019). In some lineages, including *Crypturellus* and *Tinamus*, the W chromosome has substantially differentiated from the Z, to an extent similar to that in neognaths. However, it seems the third stratum that spans nearly half of the Z chromosome can account for tinamous' nearly complete differentiation of the sex chromosome. In fact, kiwis which also show a small

PAR, has only two strata but the latest stratum spans more than half of the length of the Z chromosome (Wang et al. 2019). This suggests the differential degrees of sex chromosomes degeneration among birds stems from the occurrence of one or two large strata (likely due to chromosomal rearrangements), rather than differential rates of accumulations of graduate changes (small strata) between the Z and W chromosomes.

However, it's unclear why no large chromosomal rearrangements have been fixed in ratites which lack additional large-scale sex chromosome differentiation. A previous study on emu suggests the sexual antagonism can be solved through male-biased expression in the PAR without the need of restricting the recombination of the PAR (Vicoso, Kaiser, and Bachtrog 2013). This hypothesis is however not supported by our more extensive and sophisticated study (Xu, Sin, et al. 2019). We also showed that the absence of global dosage compensation is probably not a sensible explanation for the slower degeneration of ratite sex chromosome, as suggested by a study in ostrich (Yazdi and Ellegren 2014). Recently, a study suggests the unique paternal care of ratites can be responsible for the slower evolution of sex chromosomes in this lineage, though no direct evidence has been provided (Wang et al. 2019).

Nevertheless, in ratites which have nearly homomorphic sex chromosome, the PAR already display features resembling a hemizygous Z-chromosome, including lower recombination rates and accumulation of TEs (Transposable elements) (Xu, Sin, et al. 2019; Yazdi 2019). The reduced efficacy of selection due to lower recombination rate, as well as the accumulation of TE, may increase the chance of chromosomal rearrangements and/or chromatin structure alternations, thereby recombination suppression. Given sufficient time, the sex chromosomes of ratites may ultimately be as fully differentiated as in other neognaths.

Temporal evolution of sex chromosome

The rate of sex chromosome differentiation shows not only an interspecies variation, but also a temporal variation over more than 100 million years' evolution of birds. In chapter 2, we reveal that at each stratum, the rate of gene loss of the W chromosome slows down over time (Xu, Auer, et al. 2019). This suggests the genes became loss more rapidly at the earlier stages of sex chromosome differentiation, a pattern that is also seen in mammalian (Bellott et al. 2014) and Drosophila (Bachtrog 2008) sex chromosomes. This is perhaps because there was a larger portion of the gene repertoire affordable to be lost immediately after recombination suppression, and over time that portion becomes smaller. We have identified two evolutionary forces that retain a certain pool of genes, despite being in a small number, on the W chromosome, namely

the purifying selection for important regulatory genes and selection to maintain dosage-sensitive genes. Those two forces do not seem to differ among the strata.

Additionally, we reveal a differential rate of sex chromosome degeneration among strata, with the younger strata losing fewer proportions of their gene content. This might reflect the biased spatial distribution of dosage-sensitive or regulatory genes towards younger-strata regions of the Z chromosome. Alternatively, it could simply be that the younger strata have undergone less time for the genes to decay. Since the sex chromosomes of tinamous and kiwis have young strata of independent origin, a detailed study on those two lineages may provide insight into the evolutionary forces governing the gene retention on the W chromosomes.

Interestingly, in chapter 3 we also reveal a different pattern of the faster-Z effect among strata of different ages. In particular, the oldest stratum of birds rarely displays a faster-Z effect, in both paleognaths and neognaths (Xu, Sin, et al. 2019; Wang et al. 2014). We argue that in the oldest stratum the purifying selection on the hemizygous alleles may have counterbalanced the effect of genetic drift that normally predicts a faster-Z in the non-recombining part of the Z chromosome. In the younger strata, however, the fixation of slightly deleterious mutations may be more tolerable, though we don't completely understand why. More study is needed to confirm this trend, and I figure the tinamou lineage will be an excellent study model since it contains strata with multiple various ages.

We have also provided evidence that different mechanisms of recombination suppression have been involved for different strata. It is most likely that a large inversion led to the formation of the second stratum in neognaths (Zhou et al. 2014), but additional evidence for the role of inversion in other strata is absent. In chapter 2, we present evidence the specific accumulation of a CR1, a family of TE, is the most likely explanation for the formation of the latest stratum in songbirds (Xu, Auer, et al. 2019). One implication of the TE-induced recombination suppression is, the formation of a new stratum may not necessarily be an adaptive outcome, or a result of sexually antagonistic selection, but simply a deleterious byproduct of TE proliferation. Nevertheless, the first phase of recombination suppression that contained *Dmrt1* was probably favored by the selection for maintaining the linkage of *Dmrt1* and other sexual antagonistic loci, though empirical evidence is still lacking.

W-chromosome innovations

Finally, in chapter 2 and chapter 4 we discovered and characterized novel gene acquisitions on the W chromosome through retroposition and transposition, respectively. In both cases, such reports are the first of its kind in the bird sex chromosome system. Retroposition is usually mediated by transposable elements (Moran, DeBerardinis, and Kazazian 1999; Tan et al. 2016), leaving a duplicated copy that contains no introns at a new locus. We found one such case on the W chromosome of American crow, and the gene *Narf* was retroposed from an autosome (Xu, Auer, et al. 2019). On the contrary, the mechanism of transposition is less clear, but likely due to nonallelic homologous recombination (Veerappa, Padakannaya, and Ramachandra 2013). We found much more frequent transpositions relative to retrotranspositions from the Z to W chromosomes, involving 14 genes in three songbird lineages (Xu, Auer, et al. 2019). The recent transpositions essentially created new young strata on the Z chromosomes.

It is unclear about the function and adaptive relevance of *Narf*, but we have identified two evolutionary forces that fixed the transposed genes on the songbird W chromosomes. The first is purifying selection for dosage-sensitive or housekeeping genes, which is also at play for other retained W-linked genes (Bellott et al. 2017; Xu, Auer, et al. 2019; Bellott et al. 2014). It appears that the need to maintain dosage-sensitive genes is probably a more dominant force in shaping the gene content of the avian W chromosomes, while the mammalian or Drosophila Y chromosomes are enriched for male beneficial genes driven by sex-specific selection.

Importantly, we provided new evidence for a novel force, that is, the female-specific selection that maintains the function of transposed genes. Particularly, the gene Anxa1 has been shown to have ovary-biased genes in all birds investigated as well as in lizard (Xu, Auer, et al. 2019). It is unclear if the expression of this gene may be harmful to males, but transposing this gene (and likely its regulatory regions as well) to the female-specific W chromosome is perhaps a more direct way to avoid sexual conflicts this gene may cause. It remains to be investigated whether the W-linked *Anxa1* has rewired its regulatory domains since its rebirth on the W chromosome.

Future perspectives

Besides what I have covered in this thesis on the topics of avian sex chromosome evolution, there are many other interesting processes that have been made in recent years. In particular, neo-sex chromosomes through chromosome fusions have been reported in multiple lineages of birds. The first reported neo-sex chromosome has an ancestral origin in Sylvioidea warblers, caused by a chromosomal fusion between the sex chromosomes and a part of chromosome 4

(Pala et al. 2012; Leroy et al. 2019). Furthermore, within this clade, additional fusions of chromosome 5 and a part of chromosome 3 in larks have made the Z chromosomes the largest chromosome in their genomes (Dierickx et al. 2019; Sigeman et al. 2019). More recently, an independent formation of the neo-sex chromosome was reported in eastern yellow robin that was derived from the fusion between the sex chromosomes and chromosome 1A (Gan et al. 2019). The neo-sex chromosome may also exist in lineages other than passerines, for instance, the chromosomal assembly of budgerigar indicates a fusion of the Z chromosome and chromosome 11, though the sex-linkage needs further verification (Cooke et al. 2017). In most of these cases, it seems the sex chromosomes started the differentiation process once the fusions took place, therefore creating young strata on the Z chromosomes. Those young strata will also be useful models to study the evolutionary forces that shape the evolution of bird sex chromosomes. For instance, the effect of faster-Z evolution in the young stratum of reunion grey white-eye seems to be curiously weak (Leroy et al. 2019), implying another form of selection likely at play at the nascent stage of sex chromosome differentiation.

In recent years, long-read sequencing technology has made it possible to sequence through the heterochromatic part of the genomes (Jain et al. 2018; Chang and Larracuente 2019; Khost, Eickbush, and Larracuente 2017), including the avian W chromosome (Weissensteiner and Suh 2019; Peona, Weissensteiner, and Suh 2018). This will allow for the study of chromosomal rearrangements and the proliferation of TEs of the W chromosomes, and their impacts on the recombination suppression and the evolution of avian sex chromosomes.

Reference

Bachtrog, Doris. 2008. “The Temporal Dynamics of Processes Underlying Y Chromosome Degeneration.” Genetics 179 (3): 1513–25.

Bellott, Daniel W., Jennifer F. Hughes, Helen Skaletsky, Laura G. Brown, Tatyana Pyntikova, Ting-Jan Cho, Natalia Koutseva, et al. 2014. “Mammalian Y Chromosomes Retain Widely Expressed Dosage-Sensitive Regulators.” Nature 508 (7497): 494–99.

Bellott, Daniel W., Helen Skaletsky, Ting-Jan Cho, Laura Brown, Devin Locke, Nancy Chen, Svetlana Galkina, et al. 2017. “Avian W and Mammalian Y Chromosomes Convergently Retained Dosage-Sensitive Regulators.” Nature Genetics 49 (3): 387–94.

Chang, Ching-Ho, and Amanda M. Larracuente. 2019. “Heterochromatin-Enriched Assemblies Reveal the Sequence and Organization of the Y Chromosome.” Genetics 211 (1): 333–48.

Cooke, Thomas F., Curt R. Fischer, Ping Wu, Ting-Xin Jiang, Kathleen T. Xie, James Kuo, Elizabeth Doctorov, et al. 2017. “Genetic Mapping and Biochemical Basis of Yellow Feather Pigmentation in Budgerigars.” Cell 171 (2): 427–39.e21.

Damas, Joana, Jaebum Kim, Marta Farré, Darren K. Griffin, and Denis M. Larkin. 2018. “Reconstruction of Avian Ancestral Karyotypes Reveals Differences in the Evolutionary History of Macro- and Microchromosomes.” Genome Biology 19 (1): 155.

Dierickx, Elisa, Simon Sin, Pieter van Veelen, Michael de L. Brooke, Yang Liu, Scott Edwards, and Simon Martin. 2019. “Neo-Sex Chromosomes and Demography Shape Genetic Diversity in the Critically Endangered Raso Lark.” bioRxiv. https://doi.org/10.1101/617563.

Gan, Han Ming, Stephanie Falk, Hernán E. Morales, Christopher M. Austin, Paul Sunnucks, and Alexandra Pavlova. 2019. “Genomic Evidence of Neo-Sex Chromosomes in the Eastern Yellow Robin.” GigaScience 8 (9). https://doi.org/10.1093/gigascience/giz111.

Jain, Miten, Hugh E. Olsen, Daniel J. Turner, David Stoddart, Kira V. Bulazel, Benedict Paten, David Haussler, Huntington F. Willard, Mark Akeson, and Karen H. Miga. 2018. “Linear Assembly of a Human Centromere on the Y Chromosome.” Nature Biotechnology 36 (4): 321–23.

Khost, Daniel E., Danna G. Eickbush, and Amanda M. Larracuente. 2017. “Single-Molecule Sequencing Resolves the Detailed Structure of Complex Satellite DNA Loci in Drosophila Melanogaster.” Genome Research 27 (5): 709–21.

Leroy, Thibault, Yoann Anselmetti, Marie-Ka Tilak, Sèverine Bérard, Laura Csukonyi, Maëva Gabrielli, Céline Scornavacca, Borja Milá, Christophe Thébaud, and Benoit Nabholz. 2019. “A Bird’s White-Eye View on Neosex Chromosome Evolution.” bioRxiv. https://doi.org/10.1101/505610.

Moran, J. V., R. J. DeBerardinis, and H. H. Kazazian Jr. 1999. “Exon Shuffling by L1 Retrotransposition.” Science 283 (5407): 1530–34.

Pala, I., S. Naurin, M. Stervander, D. Hasselquist, S. Bensch, and B. Hansson. 2012. “Evidence of a Neo-Sex Chromosome in Birds.” Heredity 108 (3): 264–72.

Peona, Valentina, Matthias H. Weissensteiner, and Alexander Suh. 2018. “How Complete Are ‘Complete’ Genome Assemblies?-An Avian Perspective.” Molecular Ecology Resources 18 (6): 1188–95.

Sigeman, Hanna, Suvi Ponnikas, Pallavi Chauhan, Elisa Dierickx, M. de L. Brooke, and Bengt Hansson. 2019. “Genomics of Expanded Avian Sex Chromosomes Shows That Certain Chromosomes Are Predisposed towards Sex-Linkage in Vertebrates.” bioRxiv. https://doi.org/10.1101/585059.

Tan, Shengjun, Margarida Cardoso-Moreira, Wenwen Shi, Dan Zhang, Jiawei Huang, Yanan Mao, Hangxing Jia, et al. 2016. “LTR-Mediated Retroposition as a Mechanism of RNA-Based Duplication in Metazoans.” Genome Research 26 (12): 1663–75.

Veerappa, Avinash M., Prakash Padakannaya, and Nallur B. Ramachandra. 2013. “Copy Number Variation-Based Polymorphism in a New Pseudoautosomal Region 3 (PAR3) of a Human X-Chromosome-Transposed Region (XTR) in the Y Chromosome.” Functional & Integrative Genomics 13 (3): 285–93.

Vicoso, B., V. B. Kaiser, and D. Bachtrog. 2013. “Sex-Biased Gene Expression at Homomorphic Sex Chromosomes in Emus and Its Implication for Sex Chromosome Evolution.” Proceedings of the National Academy of Sciences 110 (16): 6453–58.

Wang, Zongji, Jilin Zhang, Xiaoman Xu, Christopher Witt, Yuan Deng, Guangji Chen, Guanliang Meng, et al. 2019. “Phylogeny, Transposable Element and Sex Chromosome Evolution of the Basal Lineage of Birds.” bioRxiv. https://doi.org/10.1101/750109.

Wang, Zongji, Jilin Zhang, Wei Yang, Na An, Pei Zhang, Guojie Zhang, and Qi Zhou. 2014. “Temporal Genomic Evolution of Bird Sex Chromosomes.” BMC Evolutionary Biology 14 (December): 250.

Weissensteiner, Matthias H., and Alexander Suh. 2019. “Repetitive DNA: The Dark Matter of Avian Genomics.” In Avian Genomics in Ecology and Evolution, edited by Robert H. S. Kraus, 93–150.

Xu, Luohao, Gabriel Auer, Valentina Peona, Alexander Suh, Yuan Deng, Shaohong Feng, Guojie Zhang, et al. 2019. “Dynamic Evolutionary History and Gene Content of Sex Chromosomes across Diverse Songbirds.” Nature Ecology & Evolution 3 (5): 834–44.

Xu, Luohao, Simon Yung Wa Sin, Phil Grayson, Scott V. Edwards, and Timothy B. Sackton. 2019. “Evolutionary Dynamics of Sex Chromosomes of Paleognathous Birds.” Genome Biology and Evolution 11 (8): 2376–90.

Xu, Luohao, and Qi Zhou. 2019. “Female-Specific and Dosage Selections Restore Genes through Transpositions onto the Degenerated Songbird W Chromosomes.” bioRxiv. https://doi.org/10.1101/692194.

Yazdi, Homa Papoli. 2019. “The Evolution of Sex Chromosomes and Sex-Linked Sequences in Birds.” Acta Universitatis Upsaliensis. http://www.diva-portal.org/smash/record.jsf?pid=diva2:1298940.

Yazdi, Homa Papoli, and Hans Ellegren. 2014. “Old but Not (so) Degenerated--Slow Evolution of Largely Homomorphic Sex Chromosomes in Ratites.” Molecular Biology and Evolution 31 (6): 1444–53.

Yonezawa, Takahiro, Takahiro Segawa, Hiroshi Mori, Paula F. Campos, Yuichi Hongoh, Hideki Endo, Ayumi Akiyoshi, et al. 2017. “Phylogenomics and Morphology of Extinct Paleognaths Reveal the Origin and Evolution of the Ratites.” Current Biology: CB 27 (1): 68–77.

Zhang, Guojie. 2018. “The Bird’s-Eye View on Chromosome Evolution.” Genome Biology 19 (1). https://doi.org/10.1186/s13059-018-1585-z.

Zhang, Guojie, Cai Li, Qiye Li, Bo Li, Denis M. Larkin, Chul Lee, Jay F. Storz, et al. 2014. “Comparative Genomics Reveals Insights into Avian Genome Evolution and Adaptation.” Science 346 (6215): 1311–20.

Zhou, Qi, Jilin Zhang, Doris Bachtrog, Na An, Quanfei Huang, Erich D. Jarvis, M. Thomas P. Gilbert, and Guojie Zhang. 2014. “Complex Evolutionary Trajectories of Sex Chromosomes across Bird Taxa.” Science 346 (6215): 1246338.