Open access • Posted Content • DOI:10.1101/692194

Female-specific and dosage selections restore genes through transpositions onto the degenerated songbird W chromosomes — Source link

Luohao Xu, Luohao Xu, Qi Zhou, Qi Zhou

Institutions: University of Vienna, Life Sciences Institute

Published on: 04 Jul 2019 - bioRxiv (Cold Spring Harbor Laboratory)

Topics: Dosage compensation, Housekeeping gene and Genome

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

Luohao Xu[1,2], 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

[*]Correspondence should be addressed to Q. Z. (zhouqi1982@zju.edu.cn)

Abstract

Sex chromosomes are usually suppressed for homologous recombination, which leads to the loss of functional genes on the Y or W chromosomes. It remains unclear how species like birds with a ZW sex system cope with the consequential gene expression imbalance, usually in the absence of global dosage compensation mechanism. Here we tackle this conundrum by reporting 14 genes recently transposed from the Z to the W chromosomes of three songbird lineages, after analyzing a total of 12 songbird species' genomes. These 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, the other retained genes after transposition are enriched for 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 provide direct evidence that sex-specific and dosage selections may have recurrently driven the restoration of genes on the Y or W chromosomes, and suggest their evolutionary processes are more dynamic than simply becoming completely degenerated.

Y chromosome embarks on an irreversible trajectory of functional degeneration, at regions where its homologous recombination with the X chromosome was suppressed. This is demonstrated by the widely-observed difference of genomic and epigenomic compositions between X and Y chromosomes: while the X is euchromatic and gene rich, the Y chromosome usually has lost most of its functional genes and become highly heterochromatic (Bachtrog 2013; Hughes, et al. 2015). Similar patterns have been found on the W chromosomes in species like birds and butterflies, which have a pair of ZZ chromosomes in males, and ZW chromosomes in females. The divergent evolutionary trajectories between sex chromosome pair are 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 (Charlesworth and Charlesworth 2000; Ponnikas, et al. 2018). The consequential cost of maintaining sex is essentially a much compromised level of natural selection on the Y/W chromosome due to the lack of recombination(Charlesworth and Charlesworth 2000). This creates a conundrum that when recombination was initially suppressed, the affected regions must contain many other genes with important functions besides the SD/SA genes.

A direct resolution to such 'collateral damage' is evolution of dosage compensation on the X/Z chromosome, so that the balance of expression level can be restored. On the other hand, studies showed that the Y/W chromosomes also come up with various strategies to 'rescue' functions of certain genes during their complex and dynamic evolutionary course. 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 avian W chromosomes (Smeds, et al. 2015; Bellott, et al. 2017; Xu, et al. 2019) due to a much higher level of selective constraints. The human Y chromosome has evolved palindromic sequence structures to 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 through transposition(Hughes, et al. 2015), or recruiting novel genes onto the Y chromosomes from various resources. Emerging cases of such gene movements on the Y chromosome have been reported since the characterization of 'X-transposed, XTR' region 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). This region was transposed from the X 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 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 how the avian W chromosomes resolve the conundrum of losing dosage sensitive genes, particularly without global dosage compensation ever evolved on the homologous Z chromosome (Itoh, et al. 2007; Graves 2014; Gu and Walters 2017). Little genomic information is available except for the euchromatic parts of W chromosomes of chicken (Bellott, et al. 2017), and tens of other bird species (Zhou, et al. 2014; Smeds, et al. 2015; Xu, et al. 2019), although a previous study suggested palindrome structure also exists on the W chromosomes of sparrows and blackbirds (Davis, et al. 2010). One might expect transposition or 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, relative to over 15,000 cases in mammals(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 a songbird W chromosome (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 looked into this question by studying 12 songbird genomes which both male and female sequencing data is available. We reasoned that these Illumina-based genomes do not contain complete information of complex sequence structures (e.g., palindromes) or traces of ancient transposition events on the W chromosome. We therefore focused on identifying the recent transpositions, if any from the Z onto the W chromosome that are manifested as female-specific elevations of both read coverage and heterozygosity level. While other regions with such an elevation pattern in both sexes are inferred as Z-linked duplications, those at the end of the chromosome with elevation of female coverage to the rest Z-linked regions, but without sex-biased patterns of heterozygosity are inferred as pseudoautosomal regions (PAR) (Figure 1,

Supplementary Figure 1). Intriguingly, we identified four Z-to-W transposition events involving 14 genes among great tit (*Parus major*), medium ground finch (*Geospiza fortis*), red bird-of-paradise (*Paradisaea rubra*) and Raggiana bird-of-paradise (*P. raggiana*). The two birds-of-paradise species share the same transposition, 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 red bird-of-paradise. We dated the transposition of medium ground finch about 8.3 million years (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 Figure 2). Similarly, we dated the transpositions of red bird-of-paradise within 4MY (Supplementary Figure 3) and that of great tit about 7 MY ago, after examining their sister species.

These very recent Z-to-W transpositions provide a unique window for us to examine the evolution of W-linked genes at their early stages. They show clear signatures 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. The most prominent case of gene loss after transposition is found in red bird-of-paradise. Almost half of the originally transposed sequences, involving a large 583kb region and 4 encompassing genes and 2 partial genes, and a nearby smaller 2kb region (Supplementary Figure 4) have become deleted, with the deleted regions showing similar levels of coverage and heterozygosity with other non-transposed Z-linked regions in the female (Figure 2, Supplementary Note). Based on analyses of insert size of mate-pair libraries, we have not identified any large-scale insertions into the transposed regions.

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 red bird-of-paradise (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 Figure 5), and also their outgroup species green anole lizard (Figure 3). Although *ALDH1A1* has a relatively lower

126 expression in ovary than in testis, it has been recently shown in mice that the disruption of this
127 gene delays the onset of meiosis in ovary (Bowles, et al. 2016). Besides, *ANXA1* and *CDK7*
128 probably have been restored by strong dosage selection, indicated by their much higher levels
129 of predicted haploinsufficiency (HP score) than most other genes on the Z chromosome
130 (Supplementary Figure 5) (Huang, et al. 2010), as well as a lack of any nonsynonymous
131 changes compared to their Z-linked homologs (Supplementary Table 1). Several medium
132 ground finch genes, for example, *SERINC5* and *MTX3*, have a low HP score, but a very broad
133 expression pattern across tissues measured by tissue-specificity matrix *tau*, thus are likely
134 restored as housekeeping genes (Figure 3). In fact, the restored genes have a generally higher
135 HP score (P=0.051, Wilcoxon test) than the genes that have become lost after the
136 transpositions.
137 These results together provide clear evidence for the female-specific and dosage
138 selections have driven the frequent restoration of W-linked genes through transpositions among
139 songbird species. Because similar X-to-Y transpositions have been reported in insects and
140 mammals, we propose that restoration of once-lost genes onto the non-recombining sex
141 chromosomes is probably a general feature in sex chromosomes evolution. Although such
142 restoration is not expected to alter the evolutionary trajectories of W or Y chromosomes toward
143 complete functional degeneration, in fact, we found some transposed genes have already
144 become lost or show signatures of functional degeneration (e.g., *THBS4*). Such loss-and-
145 restoration cycles may recurrently occur throughout the evolution of sex chromosomes,
146 particularly in ZW systems that usually do not have global dosage compensation to cope with
147 the imbalance of gene expression. We have to point out that our method can only identify recent
148 transpositions, and probably has missed ancient transpositions that have become too divergent
149 in sequence between Z and W chromosomes. The genes involved in the such cases
150 nevertheless have probably already become pseudogenes. Our results are in line with the
151 reported cases in avian W or mammalian Y chromosomes that dosage-sensitive genes are
152 retarded for their functional degeneration due to the strong selective constraints (Bellott, et al.
153 2014; Smeds, et al. 2015; Bellott, et al. 2017; Xu, et al. 2019). We also provided new evidence
154 that sex-specific selection is shaping the evolution of the W chromosome, which was assumed
155 to be less frequent than that shaping the Y chromosome, due to the more frequent and intensive
156 male-targeted sexual selection.
157
158
159 **Materials and Methods**

The genomic, transcriptomic and resequencing data used in this study are listed in Supplementary Table 2-4. For the 12 songbird genomes, genomic data are available for both sexes except for three species. 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 (Supplementary Figure 6, Supplementary Note). 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", "ReadPosamplenkSum < -8.0". We further required the variants showing an allele frequency between 0.3 and 07 (the expected heterozygosity should be 0.5 for one individual). 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 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)

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. 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.

194

195 **Figure legend**

196 **Fig. 1 Transpositions from the Z to W chromosomes in songbirds.** Genomic regions on the
197 Z chromosome showing female-specific elevations of SNP density and coverage were inferred
198 as recent transposition event. Pseudoautosomal regions (PARs) and Z-linked duplications do
199 not show elevated SNP density. We show seven representative species out of the 12 studied
200 songbirds, including three species with signatures of transpositions. The red asterisks indicate
201 the origination branch of the transpositions.

202

203 **Fig. 2 The Z-to-W transposition in red bird-of-paradise**. a) the loci of transposition (at ~60
204 Mb) shows an elevated heterozygosity and coverage in females. b)-c) a zoom-in view of the
205 transposed region. The 1.3 Mb transposed sequence contains 9 genes, but 5 compete and 2
206 partial genes probably have become lost through a 583 kb sequence deletion. Only *ANXA1* and
207 *ALDH1A1* are retained on the W.

208

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

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

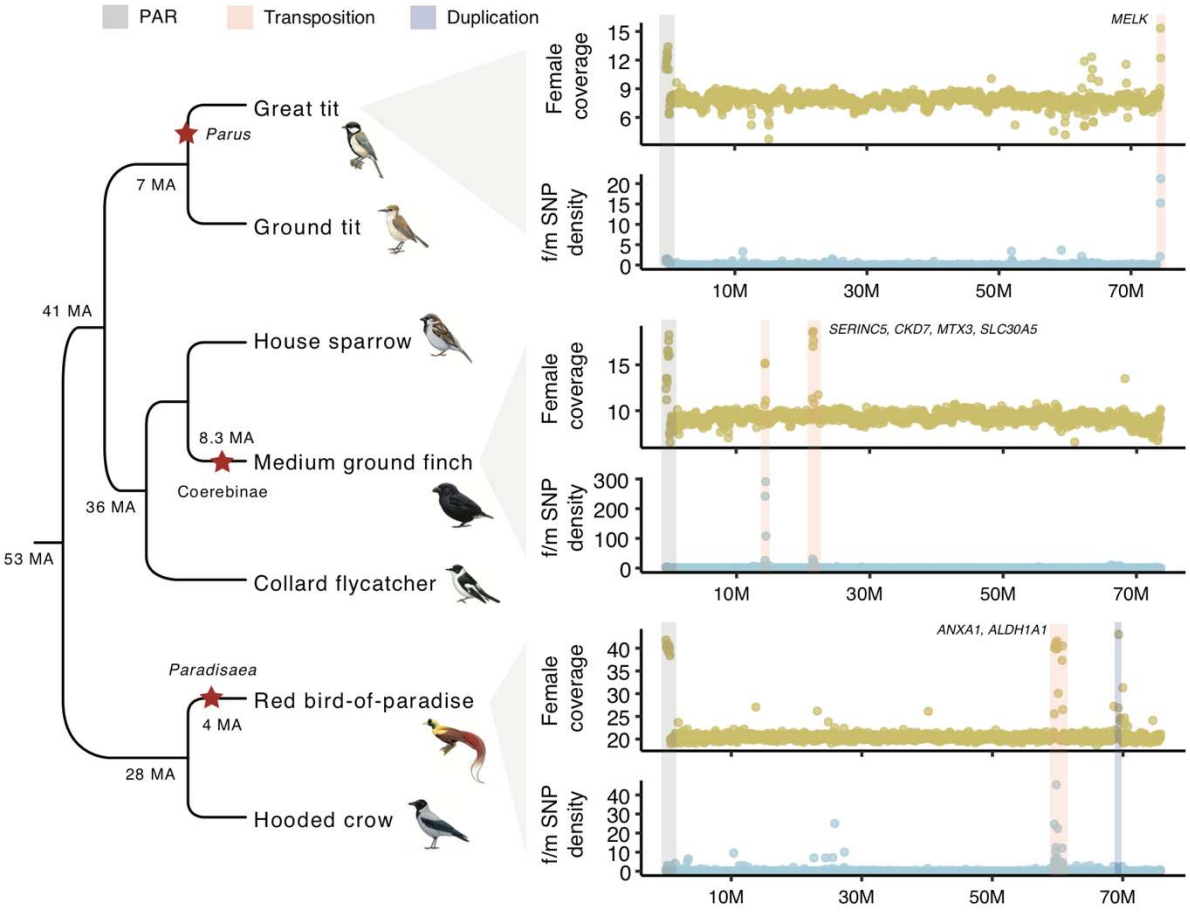


Fig. 2 The Z-to-W transposition in red bird-of-paradise.

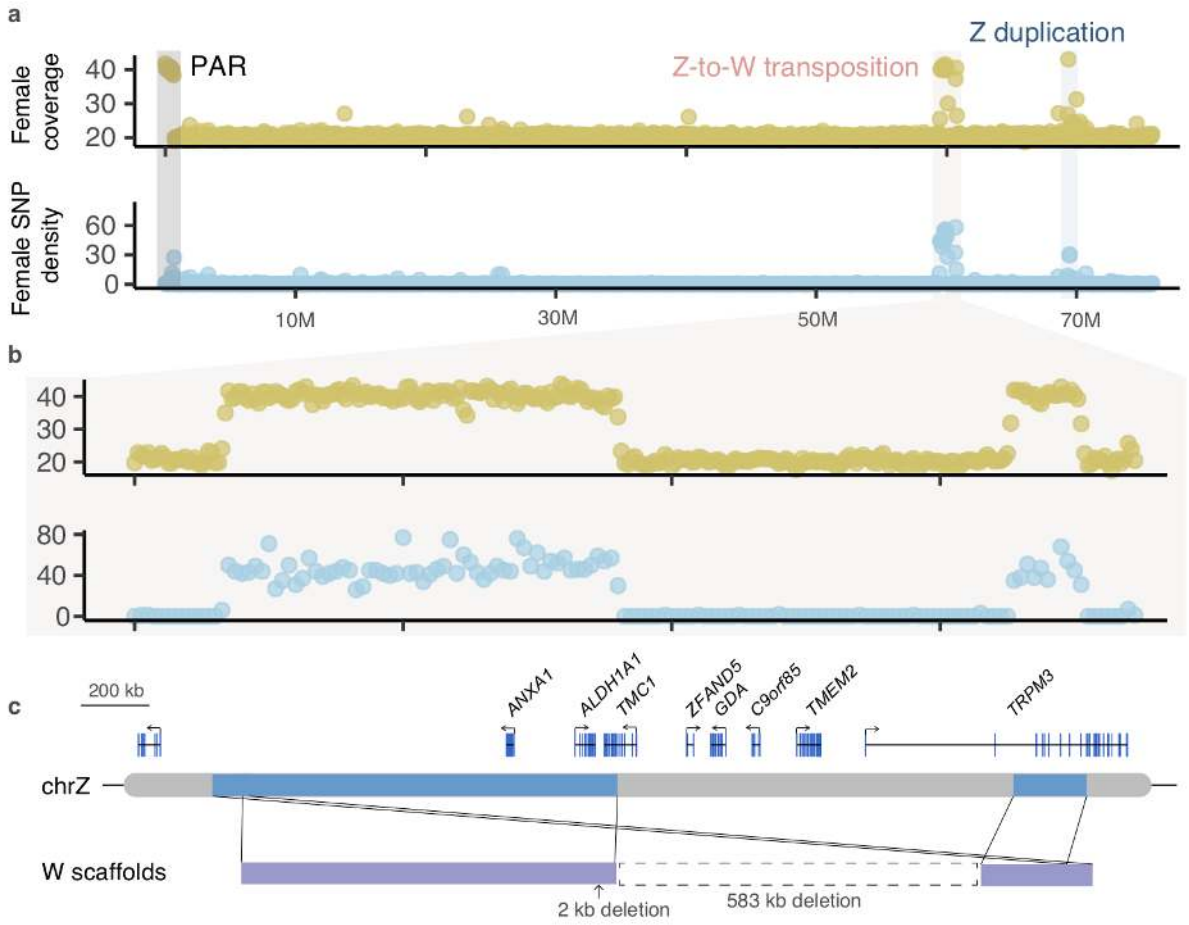
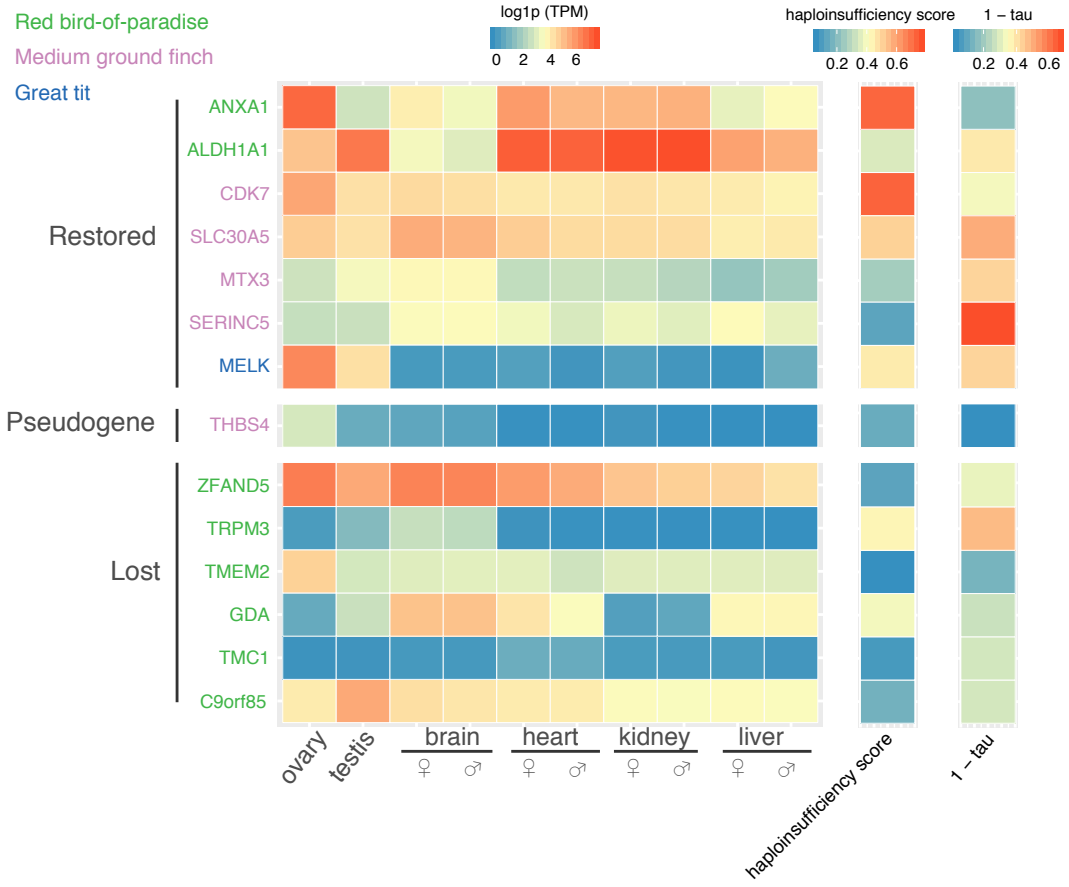


Fig. 3 Female-specific and dosage selections restore avian W-linked genes.



References

Bachtrog D. 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. 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. 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. 2016. ALDH1A1 provides a source of meiosis-inducing retinoic acid in mouse fetal ovaries. Nat. Commun. 7:10845.

Brashear WA, Raudsepp T, Murphy WJ. 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. 2015. Birth of a new gene on the Y chromosome of Drosophila melanogaster. Proc. Natl. Acad. Sci. 112:12450-12455.

Charlesworth B, Charlesworth D. 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. 2014. Origins and functional evolution of Y chromosomes across mammals. Nature 508:488-493.

Crow TJ. 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. 2010. A W-linked palindrome and gene conversion in New World sparrows and blackbirds. Chromosome Res. 18:543-553.

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

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. 9:2461-2476.

Huang N, Lee I, Marcotte EM, Hurles ME. 2010. Characterising and predicting haploinsufficiency in the human genome. PLoS Genet. 6:e1001154.

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

International Chicken Genome Sequencing C. 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. 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. 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. 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. 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. 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. 2013. Comparative analysis of mammalian Y chromosomes illuminates ancestral structure and lineage-specific evolution. Genome Res. 23:1486-1495.

Mahajan S, Bachtrog D. 2017. Convergent evolution of Y chromosome gene content in flies. Nat. Commun. 8.

Moghadam HK, Pointer MA, Wright AE, Berlin S, Mank JE. 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. 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. 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. 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. 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. 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. 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. 2015. Evolutionary analysis of the female-specific avian W chromosome. Nat. Commun. 6:7330.

Speevak MD, Farrell SA. 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. 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. 2019. Dynamic evolutionary history and gene content of sex chromosomes across diverse songbirds. Nat Ecol Evol 3:834-844.

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