Downloaded from genome.cshlp.org on November 3, 2026 . Published by Cold Spring Harbor Laboratory Press



Fast-X on the Z: Rapid evolution of sex-linked genes in birds

Judith E. Mank, Erik Axelsson and Hans Ellegren

Genome Res. published online April 6, 2007
Access the most recent version at doi:10.1101/gr.6031907

P<P Published online April 6, 2007 in advance of the print journal.

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**



To subscribe to *Genome Research* go to:
https://genome.cshlp.org/subscriptions

Copyright © 2007, Cold Spring Harbor Laboratory Press

Downloaded from genome.cshlp.org on November 3, 2026 . Published by Cold Spring Harbor Laboratory Press

Letter

Fast-X on the Z: Rapid evolution of sex-linked genes in birds

Judith E. Mank, Erik Axelsson, and Hans Ellegren[1]

Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18 D, SE 752 36 Uppsala, Sweden

Theoretical work predicts natural selection to be more efficient in the fixation of beneficial mutations in X-linked genes than in autosomal genes. This "fast-X effect" should be evident by an increased ratio of nonsynonymous to synonymous substitutions (d_N/d_S) for sex-linked genes; however, recent studies have produced mixed support for this expectation. To make an independent test of the idea of fast-X evolution, we focused on birds, which have female heterogamety (males ZZ, females ZW), where analogous arguments would predict a fast-Z effect. We aligned 2.8 Mb of orthologous protein-coding sequence of zebra finch and chicken from 172 Z-linked and 4848 autosomal genes. Zebra finch data were in the form of EST sequences from brain cDNA libraries, while chicken genes were from the draft genome sequence. The d_N/d_S ratio was significantly higher for Z-linked (0.110) than for all autosomal genes (0.085; $P = 0.002$), as well as for genes linked to similarly sized autosomes 1–10 (0.0948; $P = 0.04$). This pattern of fast-Z was evident even after we accounted for the nonrandom distribution of male-biased genes. We also examined the nature of standing variation in the chicken protein-coding regions. The ratio of nonsynonymous to synonymous polymorphism (p_N/p_S) did not differ significantly between genes on the Z chromosome (0.104) and on the autosomes (0.0908). In conjunction, these results suggest that evolution proceeds more quickly on the Z chromosome, where hemizygous exposure of beneficial nondominant mutations increases the rate of fixation.

[Supplemental material is available online at www.genome.org and http://www.egs.uu.se/evbiol/Research/Data/fast-Z/.]

Sex chromosomes can exhibit several unusual properties, including inheritance pattern, reduced recombination, and hemizygosity, which influence the mechanisms of natural selection (Rice 1984; Vicoso and Charlesworth 2006). These differences often make evolutionary comparisons between sex chromosomes and autosomes particularly revealing, as they help answer fundamental questions regarding the signature and pattern of mutation and natural selection, as well as how these forces vary across the genome. For instance, consider new autosomal mutations, which are initially low in frequency and primarily present in heterozygotes. If these mutations are either wholly or partially recessive, they will be obscured by the ancestral allele, and will only rarely be exposed directly to selective forces. In contrast, novel recessive and otherwise nondominant mutations on sex chromosomes are directly exposed to selection in the hemizygous sex. Selection would therefore be expected to act faster on sex chromosomes to fix some types of beneficial mutations (Charlesworth et al. 1987), a situation referred to as fast-X evolution.

The fast-X (or fast-Z in the case of female heterogamety) effect could potentially explain several evolutionary phenomena. For instance, it has been invoked to explain Haldane's Rule (Haldane 1922), which states that the heterogametic sex suffers higher fitness consequences in interspecific crosses, which could be explained if the X chromosome accumulates epistatic mutations at a faster rate than the autosomes (Charlesworth et al. 1987). Additionally, many studies have found that a disproportionately large percentage of genes relating to speciation, reproductive isolation, and mate choice map to the sex chromosomes (Dobzhansky 1974; Templeton 1977; Coyne 1985; Coyne and Orr 1989; Prowell 1998; Reinhold 1998; Ritchie and Phillips 1998; Presgraves 2002). This may be because sex-linked incompatibility loci evolve at a faster rate than autosomal incompatability sites (Tao and Hartl 2003; Coyne et al. 2004).

However, the theory behind fast-X evolution does not hold for all types of mutations under all forms of selection. If selection acts on standing genetic variation, rather than on novel mutations, the sex-linked genomic regions will have a slower rate of evolution than autosomal regions (Orr and Betancourt 2001). Moreover, if most protein sequence evolution is due to fixation of weakly deleterious alleles through genetic drift, the autosomal genes will show higher rates of evolutionary change than sex-linked genes, where hemizygous exposure would result in faster purging (Charlesworth et al. 1987). Finally, the nature of mutations is also important to consider, as those effects of completely dominant DNA mutations would result in genetic drift fixing mildly deleterious alleles more often on the X.

It is possible to analyze the rate and pattern of nucleotide substitution for the molecular signature of fast-X evolution. Evidence of fast-X would originate from the accelerated accumulation of differences in the coding regions of sex-linked genes, and would manifest in a higher ratio of nonsynonymous (d_N) to synonymous (d_S) divergence. The ratio of these metrics (d_N/d_S) can be used as a genomic beacon for fast-X evolution when averaged across large genomic regions. However, several recent tests for fast-X evolution using this or similar methodology have produced mixed results (Betancourt et al. 2002; Thornton and Long 2002, 2005; Torgerson and Singh 2003, 2006; Lu and Wu 2005; Richards et al. 2005), even when overlapping data sets were analyzed (Counterman et al. 2004; Musters et al. 2006; Thornton et al. 2006). This may be because the theoretical requirements for fast-X (beneficial mutations that are novel and nondominant)

[1]Corresponding author.
E-mail Hans.Ellegren@ebc.uu.se; fax 46-18-4716310.
Article published online before print. Article and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.6031907.

Downloaded from genome.cshlp.org on November 3, 2026 . Published by Cold Spring Harbor Laboratory Press

are too rare to result in a detectable genomic signature, or because other phenomena act to obscure the signal in some of the analyzed taxa.

A novel approach to the question of fast-X evolution is, in contrast to previous molecular evolutionary studies on this topic, to study a system with female heterogamety (males ZZ, females ZW). One obvious advantage with analyzing a ZW system is that it offers an independent test of the theoretical predictions since this type of sex chromosome inheritance allows for the parsing of the effects of maleness from the effects of heterogamety. The theoretical predictions are analogous to those for the X chromosome in male heterogamety; new beneficial and recessive mutations would be more easily fixed on the Z chromosome than on autosomes, manifesting in a higher d_N/d_S ratio on the Z. Contrasting data from XY and ZW systems have previously proved useful for addressing many aspects of sex chromosome evolution, such as distinguishing between sex- and chromosome-specific effects in the context of male-biased mutation (Ellegren and Fridolfsson 1997), and distinguishing between sexual selection and restricted recombination in explaining the low genetic variability typically seen in the sex-limited chromosome (Berlin and Ellegren 2004).

Female heterogamety is seen in a variety of animals, including birds, butterflies, and some fish, lizards, and snakes. The chicken (*Gallus gallus*) was the first bird for which a draft genome sequence was presented (International Chicken Genome Sequencing Consortium 2004). Large-scale sequence data for a second bird species, the zebra finch (*Taeniopygia guttata*), are available in the form of expressed sequence tags (ESTs) from brain cDNA libraries (Clayton 2004). Here we align >5000 zebra finch–chicken 1:1 orthologs, two highly divergent birds that account for roughly 100 million years of avian evolution (van Tuinen et al. 2000), in order to assess the rate of nonsynonymous (d_N) and synonymous (d_S) substitution over a wide window of avian evolution. We also analyze large-scale polymorphism data from domestic breeds of chicken to be able to examine the pattern of standing genetic diversity in a bird, both the number of nonsynonymous polymorphisms per nonsynonymous site (p_N), as well as the number of synonymous polymorphisms per synonymous site (p_S). We use these data to examine the nature of coding-region mutations on sex chromosomes and autosomes, and the mechanism by which natural selection acts on those mutations. We then use these metrics to look for evidence of fast-Z evolution in birds and to gain insight into the specific selective processes that produce it.

Results

Divergence data

We found reciprocal best hits and aligned 172 Z-linked and 4848 autosomal zebra finch (unique EST contigs from brain cDNA libraries) and chicken (annotated draft genome sequence) orthologs, which comprised roughly 2.8 Mb of coding sequence (90.6 kb Z-linked and 2.72 Mb autosomal). Data on chromosomal location were obtained from the chicken genome, and there is strong indirect evidence that genes located on the chicken Z chromosome are Z-linked in zebra finch as well (Supplemental Material). It can thus be assumed that sex linkage has been retained in the lineages leading to chicken and zebra finch since they shared a common ancestor.

Across all 5020 genes, mean $d_N = 0.0396$ and $d_S = 0.462$. Z-linked genes had a higher rate of nonsynonymous substitutions (0.0451), compared to autosomal (0.0394) regions (permutation test, 1000 repetitions; $P = 0.01$) (Table 1). The synonymous substitution rate showed the opposite pattern, with the Z-linked average lower (0.410) than the autosomal average (0.464), which was also significantly different ($P = 0.03$). The d_N/d_S for Z-linked (0.110) and autosomal (0.0849) genes is significantly different ($P = 0.002$), with the ratio roughly 30% larger for coding sequences that mapped to the Z chromosome.

The avian karyotype differs from that of many other organisms by showing significant heterogeneity in chromosome size, including a large number of very small (<20 Mb) chromosomes, the "microchromosomes." The chicken diploid karyotype has 78 chromosomes, with chromosomes 11–38 defined as microchromosomes by the International Chicken Genome Sequencing Consortium (2004), while the zebra finch has $2n = 80$ (Pigozzi and Solari 1999). As judged from the analysis of the chicken genome sequence, the microchromosomes differ from the larger chromosomes by showing higher gene density, fewer repeats, shorter introns, higher GC content, and much higher recombination rates (International Chicken Genome Sequencing Consortium 2004). As several of these factors, in particular recombination, could influence the character and efficacy of selection, it is potentially more relevant to compare evolutionary rates in the Z chromosome with those in similarly sized autosomes. This could pose a problem if chromosomal rearrangements have been frequent and differentiated the chicken and zebra finch karyotypes. However, the rates of translocation, fusion, and fission have generally been low during avian genome evolution (Burt et al. 1999; Bourque et al. 2005), and, on the whole (see Supplemental Material), chicken and zebra finch chromosomes 1–10 are syntenic (Griffin et al. 2007). As the Z chromosome is similar in size to autosomes 5 and 6 in chicken, and to obtain a reasonably large set of genes for comparison, we extracted data from genes located on chicken chromosomes 1–10. For 3185 autosomal transcripts (1.79 Mb) from chicken chromosomes 1–10, d_N was still smaller (0.0392; $P = 0.04$) than for the Z chromosome, while d_S was no longer significantly different (0.413; $P = 0.43$) from Z-linked genes. Importantly, the d_N/d_S ratio was still significantly higher ($P = 0.04$) in the Z-linked genes (0.110) than in genes that mapped to chicken chromosomes 1–10 (0.0948). The reduction in d_S for chromosomes 1–10 compared to all autosomes is consistent with previous observations of a lower mutation rate on macro- than on microchromosomes (International Chicken Genome Sequencing Consortium 2004; Axelsson et al. 2005). This seems largely because of the higher GC content of microchromosomes and the positive correlation between GC and substitution rate in birds (Webster et al. 2006).

There is evidence that the avian Z chromosome has an excess of coding regions that exhibit a male-biased expression pattern (Kaiser and Ellegren 2006; Storchova and Divinia 2006), which has been shown to be positively correlated with the rate of evolution (Meiklejohn et al. 2003; Zhang et al. 2004; Zhang and Parsch 2005). Therefore, we wanted to investigate the relationship between gene expression pattern and d_N/d_S values in order to determine whether the pattern of fast-Z described above is simply the result of the hyperaccumulation of male-biased genes on the avian Z chromosome. Using data from microarray studies of gene expression in chicken brain (H. Ellengren, L. Hultin-Rosenberg, B. Brunström, L. Dencker, K. Kultima, and B. Scholz, unpubl.), we identified and then removed 64 coding regions with male-biased expression (32 Z-linked, 32 autosomal) from our

Downloaded from genome.cshlp.org on November 3, 2026 . Published by Cold Spring Harbor Laboratory Press

Table 1. Mean d_N, d_S, and d_N/d_S for Z-linked and autosomal chicken–zebra finch orthologs

	Z chromosome (90.6 kb)	Chromosomes 1–10 (1.79 Mb)	All autosomes (2.72 Mb)
d_N (95% CI)	0.0451 (0.0378–0.0508)	0.0392 (0.0375–0.0407) $P = 0.04$	0.0394 (0.0381–0.0404) $P = 0.01$
d_S (95% CI)	0.410 (0.389–0.444)	0.413 (0.406–0.423) $P = 0.43$	0.464 (0.456–0.472) $P = 0.03$
d_N/d_S (95% CI)	0.110 (0.0938–0.1234)	0.0948 (0.0910–0.0985) $P = 0.04$	0.0849 (0.0828–0.881) $P = 0.002$

Significance values were determined from permutation tests of autosomes versus Z chromosomes, and 95% confidence intervals were derived from a bootstrap test, both with 1000 repetitions.

analysis; the pattern of fast-Z evolution still remains evident after this ($P = 0.02$ for Z vs. all autosomal genes; $P = 0.005$ for Z vs. autosomes 1–10) (data not shown).

Polymorphism data

To get a broader picture of the nature of selective forces on autosomes and the Z chromosome, we next turned to the intraspecific diversity data of the domestic chicken. Coding-sequence polymorphisms have been identified through large-scale EST sequencing of multiple cDNA libraries of chicken (>485,000 sequences) (Hubbard et al. 2005) and are available in the BBSRC ChickEST database (http://www.chick.umist.ac.uk/). This database includes high-confidence SNPs predicted from stringent filtering criteria, which included the removal of singleton polymorphisms. To be able to make a direct comparison to chicken–zebra finch divergence data, we focused on the same genes as used for divergence estimates; 1724 nonsynonymous and synonymous substitutions were identified in this gene set.

While the p_N/p_S ratio for Z-linked genes was higher (0.104) than that for autosomal genes (0.0925), this difference was not statistically significant (randomization test, 1000 permutations; $P > 0.3$ in all cases) (Table 2). The p_N/p_S ratios for different genomic regions are similar as well (0.0934 vs. 0.104 for chromosomes 1–10 and the Z chromosome, respectively; $P > 0.3$ in all cases). The observation of similar p_N/p_S ratios but a higher d_N/d_S ratio for sex-linked genes indicates that the rate of fixation of nonsynonymous substitutions is higher on the Z chromosome than on autosomes, as the pattern in the underlying pool of genetic variability does not differ among chromosomal classes.

Another source of chicken polymorphism data is from the International Chicken Polymorphism Map Consortium (2004), which identified 2.8 million single nucleotide polymorphisms (SNPs) in light shotgun (0.25× coverage) sequencing of three individuals from different breeds of domestic chicken. This is the most extensive polymorphism data set available for a bird, however, since the majority of SNPs identified in that study have not been validated and typically originate from nonoverlapping reads aligned to the reference genome sequence; a certain fraction likely represent sequencing errors. Such errors would have the largest effect on relative frequency estimates for rare classes of polymorphism, primarily nonsynonymous mutations. Indeed, small-scale validation by resequencing indicates that while 94% of genome-wide SNPs are confirmed, this drops to 80% for nonsynonymous polymorphisms (International Chicken Polymorphism Map Consortium 2004). Additionally, since the Z chromosome harbors less variation in general (Sundström et al. 2004), it would be expected that sequencing errors would disproportionately influence our estimates of Z-linked polymorphisms. We therefore sought to reduce the impact of potential sequencing errors by limiting the International Chicken Polymorphism Map Consortium (ICPMC) data set to only include polymorphisms observed in at least two chicken strains. In this resulting data set, p_N/p_S ratios of sex-linked and autosomal genes are again very similar (0.185 for the Z chromosome, 0.195 for autosomes, and 0.199 for autosomes 1–10) (Table 3). The overall higher p_N/p_S ratios in the BBSRC compared to the ICPMC data sets are likely explained by the fact that we could not limit the ICPMC data to only include those genes used for divergence estimates, as this would have resulted in extremely limited sample sizes. This means that the ICPMC polymorphism data were not limited to genes expressed in brain, and include genes predicted from the chicken genome sequence that are not confirmed by zebra finch EST sequencing.

Discussion

Our results show that Z-linked protein-coding genes have evolved at a significantly faster average nonsynonymous-to-synonymous rate than autosomal genes in the lineages leading to zebra finch and chicken. While our analysis is comprised of genes expressed in the brain, there are several reasons to think that these results are more broadly applicable. First, the 5050 coding sequences analyzed here represent roughly a quarter of anno-

Table 2. Absolute numbers and p_N/p_S for polymorphisms in chicken protein-coding regions from the BBSRC EST data set

	Z chromosome (38.6 kb)	Chromosomes 1–10 (1.19 Mb)	All autosomes (1.73 Mb)
Nonsynonymous polymorphisms	7	208	297
Synonymous polymorphisms	27	952	1377
p_N/p_S (95% CI)	0.104 (0.0419–0.192)	0.0925 (0.0804–0.106) $P = 0.39$	0.0908 (0.0776–0.103) $P = 0.36$

Significance values were determined from permutation tests of autosomes versus Z chromosomes, and 95% confidence intervals were derived from a bootstrap test, both with 1000 repetitions.

Downloaded from genome.cshlp.org on November 3, 2026 . Published by Cold Spring Harbor Laboratory Press

Table 3. Absolute numbers and p_N/p_S for polymorphisms in chicken protein-coding regions from the ICPMC data set with polymorphisms that were present in at least two sequenced chicken strains

	Z chromosome (172 kb)	Chromosomes 1–10 (4.44 Mb)	All autosomes (6.41 Mb)
Nonsynonymous polymorphisms	91	2325	3408
Synonymous polymorphisms	200	4956	7331
p_N/p_S (95% CI)	0.185 (0.138–0.275)	0.199 (0.186–0.213) $P = 0.33$	0.195 (0.815–0.206) $P = 0.39$

Significance values were determined from permutation tests of autosomes versus Z chromosomes, and 95% confidence intervals were derived from a bootstrap test, both with 1000 repetitions.

tated avian genes. Additionally, while our data set is comprised of ESTs expressed in the brain, most of them are by no means brain-specific, and many are expressed broadly throughout the organism. Both these lines of evidence suggest that these results would apply to the majority of the avian transcriptome.

These birds represent two highly divergent clades in the avian phylogenetic tree. Within Neognathae (all extant birds except the mostly flightless ratites), the first split occurred between Galloanserae (including chicken) and Neoaves (remaining extant orders including song birds to which zebra finch belongs) ~100 million years ago (van Tuinen et al. 2000). The vast span of evolutionary time separating our study organisms suggests that the results obtained in this analysis may be representative for birds in general.

The increased d_N/d_S for Z-linked coding regions during avian evolution is compatible with the notion that many novel beneficial mutations are nondominant, and that they are fixed faster through hemizygous exposure to natural selection on the sex chromosomes than they would be in heterozygous exposure on an autosome (Charlesworth et al. 1987). This is the first comparative genomic support for fast-X evolution in a female heterogametic system, an observation that, by analogy, we denote fast-Z evolution.

Two questions arise from this observation. First, are there other explanations to fast-Z than a higher fixation rate of advantageous mutations? Second, are there reasons to expect fast-Z to be more evident than fast-X? The latter question arises from the conflicting evidence for the existence of a fast-X effect in recent studies of male heterogametic systems (Betancourt et al. 2002; Torgerson and Singh 2003; Counterman et al. 2004; Thornton et al. 2006).

It is theoretically possible that the lower effective population size (N_e) of the Z chromosome compared to autosomes could create a fast-Z effect through the fixation of some weakly deleterious nonsynonymous mutations. Under random mating in an ideal population, N_e of the Z is three-quarters as large as that for autosomes. With reproductive skews, common in birds where sexual selection often reduces the number of males contributing to the next generation (Andersson 1994), Z chromosome N_e can be expected to be even less relative to the autosomes. Although Charlesworth et al. (1987) accounted for a difference in N_e due to chromosome counts in their prediction for relative rates of evolution on sex chromosomes and autosomes, these predictions require some very specific assumptions about the nature of mutations, namely, that they are nondominant and in many cases beneficial. When fully dominant, slightly deleterious mutations would more easily drift to fixation on the Z chromosome, leading to an increased d_N/d_S ratio. While the limiting sample size of our polymorphism data set prevents the drawing of definitive conclusions, it does suggest that the pattern of fast-Z we observe is not due to N_e, as the presence of a significant proportion of weakly deleterious mutations segregating at Z-linked loci would have manifested in a higher p_N/p_S than for autosomal loci, and we do not observe this pattern in our data set.

The difference in p_N and p_S levels between the BBSRC and ICPMC polymorphism estimates is likely due to the nature of the underlying data sets. While the genes in our study are by no means brain-specific, the library that forms the basis of the data set is enriched for brain genes, which, on average, seem to be subject to stronger constraint than genes expressed in other tissues (Nielsen et al. 2005). This library is also likely enriched for housekeeping genes, which are more slowly evolving than the transcriptome when taken as a whole (e.g., Duret and Mouchiroud 2000). Because we limited our analysis of the BBSRC data to those coding sequences identified from brain EST libraries, it is likely that this data set is also biased toward slowly evolving housekeeping genes. This would suggest that the reduced level of polymorphism in the BBSRC data set is due to stronger evolutionary constraints acting on housekeeping genes, and the higher level of polymorphism observed in the ICPMC data is derived from the fact that this data set is comprised of all coding regions with sufficient sequencing coverage and does not differentiate according to expression locus. This enrichment in the underlying data set likely results in an underestimate of the true fast-Z effect, but would not be expected to create a spurious correlation. It is therefore doubtful that the nature of the brain EST library affects the overall conclusions of our analysis.

Theoretical predictions on the rate of evolution of sex-linked genes could be violated if genes are nonrandomly distributed across the genome, producing a fast-Z pattern for reasons other than those predicted by Charlesworth et al. (1987). An important conclusion from recent genome-wide gene expression analysis is that a significant proportion of the transcriptome shows sex-biased expression (Reinke et al. 2000; Jin et al. 2001; Ranz et al. 2003; Oliver and Parisi 2004; Yang et al. 2006). Male- and female-biased genes, especially if sex-biased gene expression has arisen from sexual antagonism, are predicted to have nonrandom genomic distributions depending on the heterogametic system (Rice 1984) or incomplete dosage compensation. This could influence studies of fast-Z as male-biased genes evolve more quickly (Zhang et al. 2004; Zhang and Parsch 2005), possibly because of their involvement with sexual selection or male–male competition. If these genes are over- or under-represented on the sex chromosome, this will impact on the mean rates of divergence of sex-linked genes. For male heterogametic organisms such as worms (Reinke et al. 2000; Jiang et al. 2001; Kelly et al. 2002), flies (Meiklejohn et al. 2003; Parisi et al. 2003; Ranz et al. 2003), and mammals (Betran et al. 2004; Khil et al. 2004; Yang et al. 2006), male-biased genes are under-represented on the X chromosome (with the exception of pre-meiotic genes expressed

Downloaded from genome.cshlp.org on November 3, 2026 . Published by Cold Spring Harbor Laboratory Press

in testis (Wang et al. 2001; Lercher et al. 2003) because of X inactivation during male meiosis. Preliminary studies in birds show the opposite pattern, with male-biased genes overrepresented on the chicken Z chromosome (Kaiser and Ellegren 2006; Storchova and Divinia 2006), possibly because of incomplete dosage compensation in birds (Itoh et al. 2007). However, the removal of genes with male-biased expression does not statistically alter our results, suggesting that the nonrandom distribution of male-biased genes is not responsible for the observed pattern of fast-Z, and offering further indirect support for the causative role of positive selection in fast-Z evolution.

There are at least two reasons to suggest that fast-Z is more easily discerned than fast-X. First, the nonrandom genomic distribution of sex-biased genes described above could theoretically mask fast-X in XY organisms because of the observed shortage of male-biased genes on X (Reinke et al. 2000; Meiklejohn et al. 2003; Parisi et al. 2003; Ranz et al. 2003). While we observed no effect of male-biased gene expression in our analysis of fast-Z, it is possible that the dearth of sex-linked male-biased genes has a stronger effect in some XY organisms because of the inactivation of the X chromosome during male meiosis (Lifschytz and Lindsley 1972; Oliver and Parisi 2004). If the under-representation of X-linked male-biased genes is sufficiently large, it could make it very difficult to detect any fast-X signal, as these forces are contradictory.

Additionally, when favorable mutations have additive effects, male-biased mutation will cause Z-linked loci to evolve faster and X-linked loci slower than autosomal genes (Kirkpatrick and Hall 2004a). Moreover, when mutations are partly dominant, male-biased mutation implies that Z-linked genes evolve faster than autosomal loci over a broader range of dominance values than is the case for X-linked genes (Kirkpatrick and Hall 2004a). Admittedly, the excess of paternally derived mutation in birds is relatively weak, with recent estimates of the male-to-female mutation rate ratio (α_m) of ~2.5 (Bartosch-Härlid et al. 2003; Axelsson et al. 2004), suggesting that the role of male-biased mutation is weak in our analysis. However, male-biased mutation is much more pronounced in some mammals, in particular, primates (Makova and Li 2002). Under this premise, the effects of male-biased mutation might also mask evidence of fast-X.

Finally, it is of interest to make the link between fast-Z and large-Z. There are observations that ZW organisms have more conspicuous male secondary sexual traits, and are thus more prone to extreme sexual selection, than XY organisms (Hastings 1994; Prowell 1998; Reinhold 1998). This idea is lent support from theoretical work comparing the likelihood of Fisher's runaway and the good-genes mechanisms of sexual selection in ZW and XY systems (Reeve and Pfennig 2003; Kirkpatrick and Hall 2004b). Moreover, there is observation of Z-linkage of plumage traits involved in species recognition in birds (Saetre et al. 2003). If fast-Z is more prevalent than fast-X, it could contribute to more rapid evolution of sex-linked genes involved with sexual selection and reproductive isolation in female heterogametic systems, making the large-Z effect stronger than the large-X.

Methods

Alignments of zebra finch–chicken orthologs

Zebra finch (*T. guttata*) EST contigs (n = 22,638), each thought to represent a unique transcript and generated by the "Songbird Neurogenomics Initiative," were available at http://titan.biotec.uiuc.edu/songbird/. The transcripts are all derived from multiple cDNA libraries made from telencephalon of embryonic, juvenile, and adult birds. BlastN was used to search for chicken orthologs to these transcripts among all known and ab initio predicted protein-coding chicken genes identified by Ensembl in the WASHUC 1 assembly (Genebuild Ensembl, December 2005; http://www.ensembl.org/Gallus_gallus/index.html). Orthology was established using the principle of best reciprocal hit, given a minimum *E*-value of 10^{-30} for a match. A total of 5658 chicken–zebra finch orthologs were thereby identified. Because *e*-scores are influenced by sequence length, we excluded all alignments <100 bp from further analysis in order to reduce the possibility of false homology between short alignments.

We screened all zebra finch sequences for high complexity contamination using RepeatMasker (A.F.A. Smit, R. Hubley, and P. Green. 1996–2004. RepeatMasker Open-3.0; http://www.repeatmasker.org) using the *Gallus*-derived repeat library. As the Ensembl gene annotation methodology automatically removes repetitive DNA elements via RepeatMasker (Curwen et al. 2004), it was not necessary to screen the orthologous chicken coding regions for retroviral contamination.

DNA sequences were subsequently translated into amino acid sequences and aligned using DIALIGN2 (Morgenstern 1999), with a minimum fragment weight (thr) of 10. DIALIGN2 was considered suitable for the alignment of EST sequences with full-length transcripts since it has been shown to perform well on local alignments (Morgenstern 1999; Edgar and Batzoglou 2006) and since it tests all possible reading frames. Gaps introduced in the aligned chicken sequence were removed as they are likely to represent erroneous base calls in EST sequencing. Gap-free regions longer than 7 bp interpreted as nonhomologous by DIALIGN2 were also excluded.

Divergence estimates

From the complete set of 5658 chicken–zebra finch orthologs, 92 were excluded as they contained premature stop codons in the zebra finch sequence; these transcripts could represent pseudogenes but could also contain sequencing errors. Furthermore, as too short sequences make parameter estimations unreliable, 63 alignments shorter than 100 bp were excluded from analysis.

Codeml (PAML package version 3.15) (Yang 1997) was used to estimate the pairwise (run mode = −2) nonsynonymous (d_N) and synonymous (d_S) divergence for each set of orthologs with settings seqtype = 1 and CodonFreq = 2. Since divergence estimates are not reliable for saturated sites, we excluded 73 orthologs with $d_S > 3$. This left 5430 genes with a mean ungapped alignment length of 559 bp for molecular evolutionary analyses.

The inference of chromosomal location was taken from the May 2006 galGal3 assembly of the chicken genome (http://www.ensembl.org/Gallus_gallus/index.html), and 5050 of the aligned ESTs were mapped to specific chromosomes and were used in further analysis. Mean values of d_N (defined as the number of nonsynonymous substitutions per nonsynonymous site) and d_S (the number of synonymous substitutions per synonymous site) for autosomes and the Z chromosome, respectively, were calculated by dividing the sum of the number of substitutions over genes by the sum of the number of sites over genes. This means that the problem of infinitely high d_N/d_S values arising from genes with no synonymous substitutions is circumvented and also that data for individual genes are weighted by alignment length.

A permutation test with 1000 repetitions was used to assess significant differences for each metric (d_N, d_S, and d_N/d_S) between

Downloaded from genome.cshlp.org on November 3, 2026 . Published by Cold Spring Harbor Laboratory Press

the Z and autosomal coding regions. The permutation test calculates the probability that the observed difference in d_N, d_S, and d_N/d_S is due to chance association alone. Additionally, 95% confidence intervals were calculated based on bootstrap simulations with 1000 repetitions.

We identified male-biased genes from an Affymetrix microarray study (H. Ellegren et al., unpubl.) of gene expression data in chicken brain tissue. Any coding region that exhibited twofold or greater bias in males compared to females was designated as male biased, according to previous studies of the relationship between male-biased gene expression and rates of evolution (Zhang et al. 2004; Pröschel et al. 2005; Zhang and Parsch 2005). We repeated the above described statistical treatments for the data set excluding these identified as coding regions from both the Z and autosomal categories.

Polymorphism analysis

The BBSRC ChickEST Database at http://www.chick.umist.ac.uk/ harbors 11,000 high-quality polymorphisms found in at least two EST sequences. We first placed these polymorphisms in the chicken genome by using the 100-bp sequence surrounding each SNP in BlastN searches against all known and predicted Ensembl chicken protein-coding transcripts, applying an *E*-value threshold of 1×10^{-30}. To make polymorphism data directly comparable to data on divergence, we restricted the data set to only include SNPs from those genes for which a zebra finch–chicken alignment was available. Although this significantly reduced the number of SNPs available for analysis, it was motivated by the fact that while divergence data were limited to genes found expressed in brain, the total BBSRC SNPs data set was derived from several different tissues. This left a set of 1674 autosomal and 34 Z-linked polymorphisms. The number of synonymous and nonsynonymous polymorphisms was obtained using self-written perl code. In case of multiple mutations per codon, the shortest evolutionary pathway was chosen. Codeml (runmode = −2, CodonFreq = 2) was used to calculate genewise *N* and *S* values. Mean values of p_N and p_S for autosomes and the Z chromosome, respectively, were calculated by dividing the sum of the number of polymorphisms over genes by the sum of the number of sites over genes. This was also done for the autosomal subset of genes that mapped to chromosomes 1–10. Significant differences in p_N, p_S, and p_N/p_S were assessed with a permutation test of 1000 repetitions, and 95% confidence intervals determined via bootstrap simulations with 1000 repetitions.

Polymorphism data from the International Chicken Polymorphism Map Consortium (2004) (available at http://genome.ucsc.edu/) were analyzed in a similar way, with codeml estimating the *N* and *S* values genewise for sequence covered in shotgun reads. Restricting the data set to polymorphisms seen in at least two chicken strains left 11,036 polymorphisms.

All PERL scripts and alignments used in this analysis are available at http://www.egs.uu.se/evbiol/Research/Data/fast-Z/.

Acknowledgments

Financial support was obtained from the Swedish Research Council and the Wenner-Gren Foundation. The useful comments of three anonymous reviewers are acknowledged.

References

Andersson, M. 1994. *Sexual selection*. Princeton University Press, Princeton, NJ.

Axelsson, E., Smith, N.G.C., Sundstrom, H., Berlin, S., and Ellegren, H. 2004. Male-biased mutation rate and divergence in autosomal, Z-linked and W-linked introns of chicken and turkey. *Mol. Biol. Evol.* **21:** 1538–1547.

Axelsson, E., Webster, M.T., Smith, N.G.C., Burt, D.W., and Ellegren, H. 2005. Comparison of the chicken and turkey genomes reveals a higher rate of nucleotide divergence on microchromosomes than macrochromosomes. *Genome Res.* **15:** 120–125.

Bartosch-Härlid, A., Berlin, S., Smith, N.G.C., Møller, A.P., and Ellegren, H. 2003. Life history and the male mutation bias. *Evolution Int. J. Org. Evolution* **57:** 2398–2406.

Berlin, S. and Ellegren, H. 2004. Chicken W: A genetically uniform chromosome in a highly variable genome. *Proc. Natl. Acad. Sci.* **101:** 15967–15969.

Betancourt, A.J., Presgraves, D.C., and Swanson, W.J. 2002. A test for faster *X* evolution in *Drosophila*. *Mol. Biol. Evol.* **19:** 1816–1819.

Betran, E., Emerson, J.J., Kaessmann, H., and Long, M. 2004. Sex chromosomes and male functions—Where do new genes go? *Cell Cycle* **3:** 873–875.

Bourque, G., Zdobnov, E.M., Bork, P., Pevzner, P.A., and Tesler, G. 2005. Comparative architecture of mammalian and chicken genomes reveal highly variable rates of genomic rearrangements across different lineages. *Genome Res.* **15:** 98–110.

Burt, D.W., Bruley, C., Dunn, I.C., Jones, C.T., Ramage, A., Law, A.S., Morrice, D.R., Paton, I.R., Smith, J., Windsor, D., et al. 1999. The dynamics of chromosome evolution in birds and mammals. *Nature* **402:** 411–413.

Charlesworth, B., Coyne, J.A., and Barton, N.H. 1987. The relative rates of evolution of sex chromosomes and autosomes. *Am. Nat.* **130:** 113–146.

Clayton, D.F. 2004. Songbird genomics—Methods, mechanisms, opportunities, and pitfalls. *Ann. N. Y. Acad. Sci.* **1016:** 45–60.

Counterman, B.A., Ortiz-Barrientos, D., and Noor, M.A.F. 2004. Using comparative genomic data to test for fast-X evolution. *Evolution Int. J. Org. Evolution* **58:** 656–660.

Coyne, J.A. 1985. The genetic basis of Haldane's rule. *Nature* **314:** 736–738.

Coyne, J.A. and Orr, H.A. 1989. Two rules of speciation. In *Speciation and its consequences* (eds. D. Otte and J. Endler), pp. 180–207. Sinauer, Sunderland, MA.

Coyne, J.A., Elwyn, S., Kim, S.Y., and Llopart, A. 2004. Genetic studies of two species in the *Drosophila melanogaster* subgroup, *D. yakuba* and *D. santomea*. *Genet. Res.* **84:** 11–26.

Curwen, V., Eyras, E., Andrews, T.D., Clarke, L., Mongin, E., Searle, S.M.J., and Clamp, M. 2004. The Ensembl automatic gene annotation system. *Genome Res.* **14:** 942–950.

Dobzhansky, T. 1974. Genetic analysis of hybrid sterility within the species *Drosophila pseudoobscura*. *Hereditas* **77:** 81–88.

Duret, L. and Mouchiroud, D. 2000. Determinants of substitution rates in mammalian genes: Expression pattern affects selection intensity but not mutation rate. *Mol. Biol. Evol.* **17:** 68–74.

Edgar, R.C. and Batzoglou, S. 2006. Multiple sequence alignment. *Curr. Opin. Struct. Biol.* **16:** 368–373.

Ellegren, H. and Fridolfsson, A.K. 1997. Male-driven evolution of DNA sequences in birds. *Nat. Genet.* **17:** 182–184.

Griffin, D.K., Robertson, L.B.W., Tempest, H.G., and Skinner, B.M. 2007. The evolution of the avian genome as revealed by comparative molecular cytogenetics. *Cytogenet. Genome Res.* (in press).

Haldane, J.B.S. 1922. Sex-ratio and unisexual sterility in hybrid animals. *J. Genet.* **12:** 101–109.

Hastings, I.M. 1994. Manifestations of sexual selection may depend on the genetic basis of sex determination. *Proc. R. Soc. Lond. B Biol. Sci.* **258:** 83–87.

Hubbard, S.J., Grafham, D.V., Beattie, K.J., Overton, I.M., McLaren, S.R., Croning, M.D.R., Boardman, P.E., Bonfield, J.K., Burnside, J., Davies, R.M., et al. 2005. Transcriptome analysis for the chicken based on 19,626 finished cDNA sequences and 485,337 expressed sequence tags. *Genome Res.* **15:** 174–183.

International Chicken Genome Sequencing Consortium. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432:** 695–716.

International Chicken Polymorphism Map Consortium. 2004. A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature* **432:** 717–722.

Itoh, Y., Melamed, E., Yang, X., Kampf, K., Wang, S., Yehya, N., Van Nas, A., Replogle, K., Band, M.R., Clayton, D.F., et al. 2007. Dosage compensation is less effective in birds than in mammals. *J. Biol.* **6:** 2.

Jiang, M., Ryu, J., Kiraly, M., Duke, K., Reinke, V., and Kim, S.K. 2001. Genome-wide analysis of developmental and sex-regulated gene expression profiles in *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci.* **98:** 218–223.

Jin, W., Riley, R.M., Wolfinger, R.D., White, K.P., Passador-Gurgel, G., and Gibson, G. 2001. The contributions of sex, genotype and age to

Downloaded from genome.cshlp.org on November 3, 2026 . Published by Cold Spring Harbor Laboratory Press

transcriptional variance in *Drosophila melanogaster*. *Nat. Genet.* **29:** 389–395.
Kaiser, V.B. and Ellegren, H. 2006. Nonrandom distribution of genes with sex-biased expression in the chicken genome. *Evolution Int. J. Org. Evolution* **60:** 1945–1951.
Kelly, W.G., Schaner, C.E., Dernburg, A.F., Lee, M.H., Kim, S.K., Villeneuve, A.M., and Reinke, V. 2002. X-chromosome silencing in the germline of *C. elegans*. *Development* **129:** 479–492.
Khil, P.P., Smirnova, N.A., Romanienko, P.J., and Camerini-Otero, R.D. 2004. The mouse X chromosome is enriched for sex-biased genes not subject to selection by meiotic sex chromosome inactivation. *Nat. Genet.* **36:** 642–646.
Kirkpatrick, M. and Hall, D.W. 2004a. Male-biased mutation, sex linkage, and the rate of adaptive evolution. *Evolution Int. J. Org. Evolution* **58:** 437–440.
Kirkpatrick, M. and Hall, D.W. 2004b. Sexual selection and sex linkage. *Evolution Int. J. Org. Evolution* **58:** 683–691.
Lercher, M.J., Urrutia, A.O., and Hurst, L.D. 2003. Evidence that the human X chromosome is enriched for male-specific but not female-specific genes. *Mol. Biol. Evol.* **20:** 1113–1116.
Lifschytz, E. and Lindsley, D.L. 1972. The role of X-chromosome inactivation during spermatogenesis. *Proc. Natl. Acad. Sci.* **69:** 182–186.
Lu, J. and 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.* **102:** 4063–4067.
Makova, K.D. and Li, W.H. 2002. Strong male-driven evolution of DNA sequences in humans and apes. *Nature* **416:** 624–626.
Meiklejohn, C.D., Parsch, J., Ranz, J.M., and Hartl, D.L. 2003. Rapid evolution of male-biased gene expression in *Drosophila*. *Proc. Natl. Acad. Sci.* **100:** 9894–9899.
Morgenstern, B. 1999. DIALIGN 2: Improvements of the segment-to-segment approach to multiple sequence alignment. *Bioinformatics* **15:** 211–218.
Musters, H., Huntley, M.A., and Singh, R.S. 2006. A genomic comparison of faster-sex, faster-X, and faster-male evolution between *Drosophila melanogaster* and *Drosophila pseudoobscura*. *J. Mol. Evol.* **62:** 693–700.
Nielsen, R., Bustamante, C., Clark, A.G., Glanowski, S., Sackton, T.B., Hubisz, M.J., Fledel-Alon, A., Tanenbaum, D.M., Civello, D., White, T.J., et al. 2005. A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biol.* **3:** 976–985.
Oliver, B. and Parisi, M. 2004. Battle of the Xs. *Bioessays* **26:** 543–548.
Orr, H.A. and Betancourt, A.J. 2001. Haldane's sieve and adaptation from the standing genetic variation. *Genetics* **157:** 875–884.
Parisi, M., Nuttall, R., Naiman, D., Bouffard, G., Malley, J., Andrews, J., Eastman, S., and Oliver, B. 2003. Paucity of genes on the *Drosophila* X chromosome showing male-biased expression. *Science* **299:** 697–700.
Pigozzi, M.I. and Solari, A.J. 1999. The ZW pairs of two paleognath birds from two orders show transitional stages of sex chromosome differentiation. *Chromosome Res.* **7:** 541–551.
Presgraves, D.C. 2002. Patterns of postzygotic isolation in Lepidoptera. *Evolution Int. J. Org. Evolution* **56:** 1168–1183.
Pröschel, M., Zhang, Z., and Parsch, J. 2005. Widespread adaptive evolution of *Drosophila* genes with sex-biased expression. *Genetics* **174:** 893–900.
Prowell, D. 1998. Sex linkage and speciation in Lepidoptera. In *Endless forms: Species and speciation* (eds. S. Berlocher and D. Howard), pp. 309–319. Oxford University Press, New York.
Ranz, J.M., Castillo-Davis, C.I., Meiklejohn, C.D., and Hartl, D.L. 2003. Sex-dependent gene expression and evolution of the *Drosophila* transcriptome. *Science* **300:** 1742–1745.
Reeve, H.K. and Pfennig, D.W. 2003. Genetic biases for showy males: Are some genetic systems especially conducive to sexual selection? *Proc. Natl. Acad. Sci.* **100:** 1089–1094.
Reinhold, K. 1998. Sex linkage among genes controlling sexually selected traits. *Behav. Ecol. Sociobiol.* **44:** 1–7.
Reinke, V., Smith, H.E., Nance, J., Wang, J., Van Doren, C., Begley, R., Jones, S.J.M., Davis, E.B., Scherer, S., Ward, S., et al. 2000. A global profile of germline gene expression in *C. elegans*. *Mol. Cell* **6:** 605–616.
Rice, W.R. 1984. Sex chromosomes and the evolution of sexual dimorphism. *Evolution Int. J. Org. Evolution* **38:** 735–742.
Richards, S., Liu, Y., Bettencourt, B.R., Hradecky, P., Letovsky, S., Nielsen, R., Thornton, K., Hubisz, M.J., Chen, R., Meisel, R.P., et al. 2005. Comparative genome sequencing of *Drosophila pseudoobscura*: Chromosomal, gene, and *cis*-element evolution. *Genome Res.* **15:** 1–18.
Ritchie, M.G. and Phillips, S.D.F. 1998. The genetics of sexual isolation. In *Endless forms: Species and speciation* (eds. S. Berlocher and D. Howard), pp. 291–308. Oxford University Press, New York.
Saetre, G.P., Borge, T., Lindroos, K., Haavie, J., Sheldon, B.C., Primmer, C., and Syvanen, A.C. 2003. Sex chromosome evolution in *Ficedula* flycatchers. *Proc. Roy. Soc. Lond. B Biol. Sci.* **270:** 53–59.
Storchova, R. and Divinia, P. 2006. Nonrandom representation of sex-biased genes on the chicken Z chromosome. *J. Mol. Evol.* **63:** 676–681.
Sundström, H., Webster, M.T., and Ellegren, H. 2004. Reduced variation on the chicken Z chromosome. *Genetics* **167:** 377–385.
Tao, Y. and Hartl, D.L. 2003. Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. III. Heterogeneous accumulation of hybrid incompatibilities, degree of dominance, and implications for Haldane's rule. *Evolution Int. J. Org. Evolution* **57:** 2580–2598.
Templeton, A.R. 1977. Analysis of head shape differences between two infertile species of Hawaiian *Drosophila*. *Evolution Int. J. Org. Evolution* **31:** 630–641.
Thornton, K. and Long, M. 2002. Rapid divergence of gene duplicates on the *Drosophila melanogaster X* chromosome. *Mol. Biol. Evol.* **19:** 918–925.
Thornton, K. and Long, M. 2005. Excess of amino acid substitutions relative to polymorphism between *X*-linked duplications in *Drosophila melanogaster*. *Mol. Biol. Evol.* **22:** 273–284.
Thornton, K., Bachtrog, D., and Andolfatto, P. 2006. *X* chromosomes and autosomes evolve at similar rates in *Drosophila*: No evidence for faster-X protein evolution. *Genome Res.* **16:** 498–504.
Torgerson, D.G. and Singh, R.S. 2003. Sex-linked mammalian sperm proteins evolve faster than autosomal ones. *Mol. Biol. Evol.* **20:** 1705–1709.
Torgerson, D.G. and Singh, R.S. 2006. Enhanced adaptive evolution of sperm-expressed genes on the mammalian X chromosome. *Heredity* **96:** 39–44.
van Tuinen, M., Sibley, C.G., and Hedges, S.B. 2000. The early history of modern birds inferred from DNA sequences of nuclear and mitochondrial ribosomal genes. *Mol. Biol. Evol.* **17:** 451–457.
Vicoso, B. and Charlesworth, B. 2006. Evolution on the X chromosome: Unusual patterns and processes. *Nat. Rev. Genet.* **7:** 645–653.
Wang, P.J., McCarrey, J.R., Yang, F., and Page, D.C. 2001. An abundance of X-linked genes expressed in spermatogonia. *Nat. Genet.* **27:** 422–426.
Webster, M.T., Axelsson, E., and Ellegren, H. 2006. Strong regional biases in nucleotide substitution in the chicken genome. *Mol. Biol. Evol.* **23:** 1203–1216.
Yang, Z. 1997. PAML: A program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13:** 555–556.
Yang, X., Schadt, E.E., Wang, S., Wang, H., Arnold, A.P., Ingram-Drake, L., Drake, T.A., and Lusis, A.J. 2006. Tissue-specific expression and regulation of sexually dimorphic genes in mice. *Genome Res.* **16:** 995–1004.
Zhang, Z. and Parsch, J. 2005. Positive correlation between evolutionary rate and recombination rate in *Drosophila* genes with male-biased expression. *Mol. Biol. Evol.* **22:** 1945–1947.
Zhang, Z., Hambuch, T.M., and Parsch, J. 2004. Molecular evolution of sex-biased genes in *Drosophila*. *Mol. Biol. Evol.* **21:** 2130–2139.

Received October 12, 2006; accepted in revised form March 6, 2007.