

# Variation in promiscuity and sexual selection drives avian rate of Faster-Z evolution

ALISON E. WRIGHT,\*† PETER W. HARRISON,† FABIAN ZIMMER,†  
STEPHEN H. MONTGOMERY,† MARIE A. POINTER\* and JUDITH E. MANK†  
\*Department of Zoology, Edward Grey Institute, University of Oxford, Oxford OX1 3PS, UK, †Department of Genetics,  
Evolution and Environment, University College London, London WC1E 6BT, UK

## Abstract

Higher rates of coding sequence evolution have been observed on the Z chromosome relative to the autosomes across a wide range of species. However, despite a considerable body of theory, we lack empirical evidence explaining variation in the strength of the Faster-Z Effect. To assess the magnitude and drivers of Faster-Z Evolution, we assembled six *de novo* transcriptomes, spanning 90 million years of avian evolution. Our analysis combines expression, sequence and polymorphism data with measures of sperm competition and promiscuity. In doing so, we present the first empirical evidence demonstrating the positive relationship between Faster-Z Effect and measures of promiscuity, and therefore variance in male mating success. Our results from multiple lines of evidence indicate that selection is less effective on the Z chromosome, particularly in promiscuous species, and that Faster-Z Evolution in birds is due primarily to genetic drift. Our results reveal the power of mating system and sexual selection in shaping broad patterns in genome evolution.

**Keywords:** effective population size, Faster-Z evolution, genetic drift, sexual selection

Received 25 August 2014; revision received 10 February 2015; accepted 11 February 2015

## Introduction

Sex chromosomes are subject to unique evolutionary forces as a result of their unusual pattern of inheritance (Charlesworth *et al.* 1987; Vicoso & Charlesworth 2009; Connallon *et al.* 2012). The magnitude of selection, genetic drift and recombination are all predicted to differ between the sex chromosomes and autosomes (Rice 1984; Kirkpatrick & Hall 2004a; Mank *et al.* 2010a; Meisel & Connallon 2013) and studies contrasting the evolution of sex-linked to autosomal genes can shed light on the fundamental evolutionary forces acting across the genome as a whole.

Faster rates of coding sequence evolution have been observed on the Z and X chromosomes relative to the autosomes across a wide range of species (recently reviewed by Meisel & Connallon 2013), and Faster-X and Faster-Z Effects appear to be a common feature of sex chromosome evolution. However, despite elevated rates of evolution for both X-linked and Z-linked genes,

the underlying causes of Faster-X and Faster-Z Evolution are predicted to differ (Vicoso & Charlesworth 2009; Meisel & Connallon 2013).

The effective population size of X and Z chromosomes ( $N_{EX}$  and  $N_{EZ}$ ) is  $\frac{3}{4}$  that of the autosomes ( $N_{EA}$ ) when there is no difference in the variance of male and female reproductive success, such as in strictly monogamous breeding systems (Charlesworth *et al.* 1987). However, many forms of sexual selection cause elevated variance in male reproductive success (Andersson 1994), which reduces  $N_{EZ}/N_{EA}$ , and in extreme cases where a single male monopolizes the reproductive output of many females,  $N_{EZ}$  approaches  $\frac{1}{2} N_{EA}$  (Vicoso & Charlesworth 2009; Wright & Mank 2013) (Fig. 1). Correspondingly, genetic drift and fixation of weakly deleterious mutations is greater on the Z chromosome (Charlesworth 2009), and we predict a Faster-Z Effect largely due to neutral, nonadaptive processes. Empirical evidence in birds and snakes is consistent with this nonadaptive and neutral explanation of Faster-Z (Mank *et al.* 2010b; Corl & Ellegren 2012; Vicoso *et al.* 2013a); however, silk moths may present a recent exception

Correspondence: Alison E. Wright, E-mail: alison.e.wright@ucl.ac.uk

(Sackton *et al.* 2014). It is worth noting that a major factor determining the relative contribution of nonadaptive and adaptive drivers of Faster-Z is overall effective population size (Meisel & Connallon 2013). Overall  $N_E$  mediates the distribution of fitness effects, and specifically, we expect the efficacy of selection and adaptive component of Faster-Z to be weaker in populations with smaller  $N_E$  (Kimura & Ohta 1971).

The opposite relationship between male mating success and relative  $N_{EX}$  is predicted in male heterogametic systems (Laporte & Charlesworth 2002; Vicoso & Charlesworth 2009; Wright & Mank 2013). Increasing variance in male reproductive success results in  $N_{EX}/N_{EA} > 3/4$ , and  $N_{EX}/N_{EA}$  may approach 1 in extreme cases (Fig. 1). Correspondingly, the higher ratio of  $N_{EX}/N_{EA}$  is expected to decrease the effect of genetic drift in Faster-X Evolution. Elevated rates of evolution on X chromosomes are therefore more often thought to be the product of increased efficacy of selection acting on recessive X-linked alleles in the heterogametic sex, thereby increasing the rate of fixation of beneficial alleles relative to the autosomes. Consistent with adaptive Faster-X Evolution, signatures of positive selection have been uncovered on the X chromosome of mammals and *Drosophila* (Thornton & Long 2005; Baines *et al.* 2008; Hvilsom *et al.* 2012; Langley *et al.* 2012).

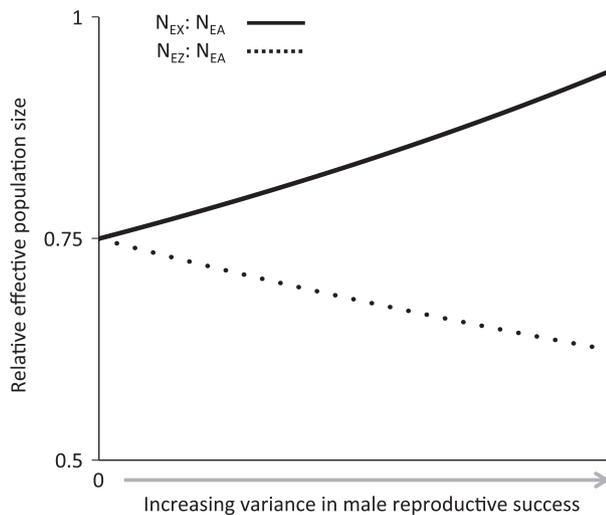
A key prediction is that the magnitude of Faster-Z Evolution can be explained by variation in the effective population size of the sex chromosomes relative to the

autosomes driven by sexual selection (Vicoso & Charlesworth 2009). Here, we explicitly test this prediction in the Galloanserae, a clade of birds spanning 90 million years (Fig. 2), for which there is extensive variation in mating system (Moller 1988, 1991; Birkhead & Petrie 1995). Using *de novo* transcriptomes for six Galloanserae species, we measured sequence divergence, polymorphism and expression and combined these molecular data with phenotypic measures of mating system to explore the nature of Faster-Z Evolution. Our results build on previous findings to reveal the dominant role nonadaptive processes play in Faster-Z. Furthermore, we uncover a positive association between Faster-Z and measures of sperm competition, a widely used indicator of the strength of postcopulatory sexual selection (Birkhead & Moller 1998). Our results suggest that variation in male mating success drives Z-linked divergence, and present the first empirical evidence in support of the considerable body of theory (Charlesworth *et al.* 1987; Vicoso & Charlesworth 2009) outlining the relationship between sexual selection and sex chromosome evolution.

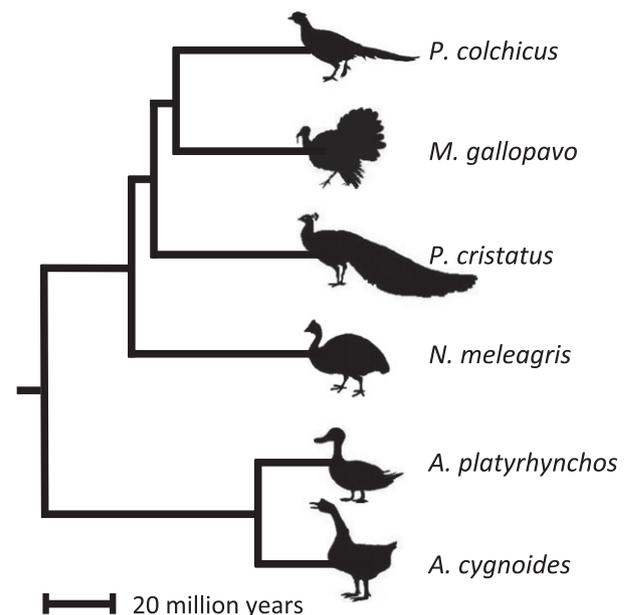
## Materials and methods

### *De novo* transcriptome assembly

RNA-Seq data were obtained from captive populations of the following Galloanserae species at the start of their first breeding season; *Anas platyrhynchos* (mallard



**Fig. 1** Relationship between effective population size ( $N_E$ ) and variance in male reproductive success. Schematic outlining the predicted relationship between variance in male reproductive success and relative  $N_{EZ}$  and  $N_{EX}$ . When variance in reproductive success is the same in males and females, under monogamy, both  $N_{EZ}$  and  $N_{EX} = 3/4 N_{EA}$ . As variance in male mating success increases,  $N_{EZ} < 3/4 N_{EA}$  and  $N_{EX} > 3/4 N_{EA}$ .



**Fig. 2** Phylogenetic relationship of the Galloanserae species in this study.

duck), *Meleagris gallopavo* (wild turkey), *Phasianus colchicus* (common pheasant), *Numida meleagris* (helmeted guinea fowl), *Pavo cristatus* (Indian peafowl) and *Anser cygnoides* (swan goose) (Fig. 2). Samples were collected with permission from institutional ethical review committees and in accordance with national guidelines. The left gonad and spleen were dissected separately from five males and five females of each species. The exceptions were *P. colchicus*, where six male gonad and spleen samples were collected, and *M. gallopavo*, where four male and two female spleens were collected. Samples were homogenized and stored in RNA later until preparation. We used the Animal Tissue RNA Kit (Qiagen) to extract RNA, and the samples were prepared and barcoded at The Wellcome Trust Centre for Human Genetics, University of Oxford using Illumina's Multiplexing Sample Preparation Oligonucleotide Kit with an insert size of 280 bp. RNA was sequenced on an Illumina HiSeq 2000 resulting in on average 26 million 100 bp paired-end reads per sample (Tables S1 and S2, Supporting Information).

The data were quality assessed using FastQC v0.10.1 ([www.bioinformatics.babraham.ac.uk/projects/fastqc](http://www.bioinformatics.babraham.ac.uk/projects/fastqc)) and filtered using Trimmomatic v0.22 (Lohse *et al.* 2012). Specifically, we removed reads containing adaptor sequences and trimmed reads if the sliding window average Phred score over four bases was <15 or if the leading/trailing bases had a Phred score <4. Reads were removed post filtering if either read pair was <25 bases in length. We constructed *de novo* transcriptome assemblies for each species using TRINITY with default parameters (Grabherr *et al.* 2011). We separately mapped back all of the reads from each sample to the Trinity contigs using RSEM v1.1.21 with default parameters (Li & Dewey 2011) to obtain expression levels. We applied a minimum expression filter of 2 reads per kilobase per million mapped reads (RPKM) requiring that each contig has expression above unlogged 2 RPKM in at least half of any of the tissues from either sex. For each Trinity contig cluster, the isoform with the highest expression level was selected for further analysis. We removed rRNA transcripts using *G. gallus* known sequences. This generated 37453 contigs for *A. platyrhynchos*, 50817 for *M. gallopavo*, 56090 for *P. colchicus*, 45535 for *N. meleagris*, 56604 for *P. cristatus* and 44144 for *A. cygnoides*.

#### Identification of Galloanserae orthogroups

*G. gallus* (Galgal4/GCA\_000002315.2) cDNA sequences were obtained from ENSEMBL v73 (Flicek *et al.* 2013), and the longest transcript for each gene was identified. We determined orthology using reciprocal BLASTN v2.2.27+ (Altschul *et al.* 1990) with an *E*-value cut-off of

$1 \times 10^{-10}$  and minimum percentage identity of 30%. Reciprocal 1-1 orthologs across all seven species (orthogroups) were identified using the highest BLAST score.

Avian chromosome structure is unusually stable, potentially due to a lack of active transposons (Toups *et al.* 2011), and major genomic rearrangements are infrequent (Stiglec *et al.* 2007). Synteny of the Z chromosome has previously been shown to be highly conserved across both extant birds (Vicoso *et al.* 2013b), as well as within the Galloanserae (Skinner *et al.* 2009). Chromosomal location was therefore assigned from *G. gallus* reciprocal orthologs.

#### Estimating sequence divergence across orthogroups

To extract Galloanserae protein-coding sequences, *G. gallus* (Galgal4/GCA\_000002315.2) protein sequences were obtained from ENSEMBL v73 (Flicek *et al.* 2013). For each orthogroup, each contig was translated into all potential reading frames and BLASTED against the orthologous *G. gallus* protein sequence using BLASTX. BLASTX outputs were used to determine coding frame, and protein-coding sequences for each species were extracted. Protein-coding sequences were defined as sequences starting with the amino acid M and terminating with a stop codon or end of the contig. Orthogroups with no BLASTX hits or a valid protein-coding sequence were excluded.

Orthogroups were aligned with PRANK v121218 using the orthologous *Taeniopygia guttata* cDNA (taeGut3.2.4.75) as an outgroup and specifying the following guidetree (((*A. cygnoides*, *A. platyrhynchos*), (*N. meleagris*, (*P. cristatus*, (*M. gallopavo*, *P. colchicus*)))), *T. guttata*). Retrotransposons were removed with REPEATMASKER (v open-4.0.3), and sequences with internal stop codons were also removed. SWAMP v0.9 (Harrison *et al.* 2014) with a cut-off of 4 and window size of 15, and a minimum length of 75 bp was used to preprocess the data.

To obtain divergence estimates for each orthogroup, we used the branch model (model=2, nssites=0) in the CODEML package in PAML v4.7a (Yang 2007), using the specified phylogeny; ((*A. cygnoides*, *A. platyrhynchos*), (*N. meleagris*, (*P. cristatus*, (*M. gallopavo*, *P. colchicus*)))), *T. guttata*). The branch model was used to calculate mean  $d_N/d_S$  across all Galloanserae branches, excluding the *T. guttata* outgroup. We will refer to this as the Galloanserae analysis. We also used the branch model to calculate mean  $d_N/d_S$  for each of the six Galloanserae species separately. Specifically, for each species, we calculated mean  $d_N/d_S$  from the terminal tip to the Galloanserae common ancestor. We will refer to this as the species-specific analysis. This approach ensures that the branch length over which  $d_N/d_S$  is calculated is

identical for each species and therefore prevents interspecific variation in branch length biasing our conclusions (Montgomery *et al.* 2011). As mutational saturation and double hits can lead to inaccurate divergence estimates (Axelsson *et al.* 2008), orthogroups were excluded if tree length  $d_S > 2$  across all branches.

#### Using sequence divergence to estimate the Faster-Z Effect

The avian genome exhibits considerable karyotypic variation in chromosome size. Therefore, mean  $d_N$ ,  $d_S$  and  $d_N/d_S$  were calculated separately for all autosomes, autosomes 1–10, microchromosomes and the Z chromosome. Microchromosomes exhibit an elevated recombination rate, greater gene density and GC content, all of which have been shown to impact the nature and efficacy of selection (Burt 2002; Ellegren 2013). The fairest measure of Faster-Z Evolution is therefore to contrast divergence between the Z chromosome and similar-sized autosomes 1–10 (Mank *et al.* 2010b).

For each genomic category, mean  $d_N$  and mean  $d_S$  were calculated as the sum of the number of substitutions across all contigs in a given category divided by the number of sites ( $d_N = \text{sum } D_N / \text{sum } N$ ,  $d_S = \text{sum } D_S / \text{sum } S$ , where  $D_{N/S}$  is an estimate of the number of nonsynonymous/synonymous substitutions and  $N/S$  is the number of nonsynonymous/synonymous sites). This approach avoids the problems of infinitely high  $d_N/d_S$  estimates arising from contigs with extremely low  $d_S$  (Mank *et al.* 2007a, 2010b) and prevents disproportionate weighting of shorter contigs.

Bootstrapping with 1000 repetitions was used to generate 95% confidence intervals, and significant differences between genomic categories were determined from 1000 permutation tests. One-tailed  $P$ -values are reported because we specifically test whether  $d_N$ ,  $d_S$  and  $d_N/d_S$  are significantly higher for Z-linked contigs vs. autosomal contigs. Mean Z-linked and autosomal  $d_N$ ,  $d_S$  and  $d_N/d_S$  values were calculated for the whole Galloanserinae (Galloanserinae analysis) and for each of the six species (species-specific analysis). Faster-Z Effect was calculated as  $d_{NZ}/d_{SZ}$ :  $d_{NA}/d_{SA}$ .

#### Testing the relationship between sexual selection and Faster-Z Effect

To test the hypothesis that the magnitude of Faster-Z increases with increased variance in male reproductive success, we performed phylogenetically controlled regression analyses between Faster-Z ( $d_{NZ}/d_{SZ}$ :  $d_{NA}/d_{SA}$ ) and relative  $N_{EZ}$  for each Galloanserinae species and two measures of female promiscuity. The intensity of sperm competition, a widely used proxy for the

magnitude of postcopulatory sexual selection and therefore variance in male reproductive success, is strongly predicted by relative testes weight and sperm number (Moller 1991; Moller & Briskie 1995; Birkhead & Moller 1998). These measures are also frequently used to test genotype–phenotype hypotheses (e.g. Dorus *et al.* 2004; Ramm *et al.* 2008). Residual testes weight was calculated using the following equation describing the linear relationship between log testes weight and body weight across a large number of birds (Pitcher *et al.* 2005):  $\log_2[\text{testes mass(g)}] = -1.56 + 0.61 \log_2 [\text{body mass(g)}]$  (Moller 1988, 1991; Birkhead & Petrie 1995). For all six species in this study, relative testes weight was less than expected given body weight. Log sperm number ( $10^6$ ) has been measured in previous studies (Moller 1988, 1991; Birkhead & Petrie 1995). Estimates for body weight and sperm number were not available for *A. cygnoides* and therefore *A. anser* estimates were used instead, as these species are closely related (Ruokonen *et al.* 2000) and both exhibit strictly monogamous mating systems.

These analyses were performed using phylogenetic generalized least squares models (PGLS) in BAYESTRAITS v2-beta (Pagel 1999; Pagel *et al.* 2004) with maximum likelihood and 1000 runs for each analysis. PGLS corrects for phylogenetic nonindependence. Phylogenies were obtained from birdtree.org using the Ericson data set. For each regression analysis, mean  $r^2$  and mean  $t$ -value (mean regression coefficient/mean standard error) were calculated. A one-tailed  $t$ -test with four degrees of freedom was used to determine whether the slope was significantly  $> 0$ .

Differences in the rate of male-biased mutation across the six species could contribute to variation in Faster-Z Effect because the Z chromosome is more often present in males than the autosomes (Kirkpatrick & Hall 2004a). We explicitly tested for significant differences in mean Z-linked  $d_S$  across the six species using permutation tests with 1000 replicates to verify that there were no underlying differences in mutation rate.

#### Tests of positive selection using sequence data

To test for signatures of positive selection acting at a subset of sites, we used the site models in the CODEML package in PAML v4.7a (Yang 2007). These models allow  $d_N/d_S$  to vary among sites but not across lineages. To test for positive selection, we compared likelihoods from two models; M1a (Nearly neutral, model=0, nssites=1) and M2a (Positive selection, model=0, nssites=2). Under model M1a, sites can fall into one of two categories (purifying selection  $d_N/d_S < 1$  and neutral evolution  $d_N/d_S = 1$ ), whereas there is an additional category under model M2a (positive selection  $d_N/d_S > 1$ ).

The following phylogeny was specified; ((*A. cygnoides*, *A. platyrhynchos*), (*N. meleagris*, (*P. cristatus*, (*M. gallopavo*, *P. colchicus*))), *T. guttata*).

#### Tests of positive selection using polymorphism data

We tested for deviations from neutrality using polymorphism data. Polymorphism data was obtained by first mapping RNA-seq reads to orthogroups using the two-pass alignment method of the STAR aligner with default parameters (Dobin *et al.* 2013). SNPs were called using VARSCAN v2.3.6 (Koboldt *et al.* 2009, 2012) and SAMTOOLS (Li *et al.* 2009) following the recommendations of Quinn *et al.* 2013 (Quinn *et al.* 2013). Only uniquely mapping reads were used to call SNPs. SAMTOOLS was run with probabilistic alignment disabled and a maximum read depth of 10 000 000. VARSCAN mpileup2snp was run with a minimum coverage of 2, a minimum average quality of 20, with the strand filter, *P*-value of 1, a minimum variant allele frequency threshold of 1E-1 and a minimum frequency to call homozygote of 0.85. SNPs were required to have a minor allele frequency >0.15 and to be from regions where at least 4 samples had a read depth >20 and have a Phred quality >20. Valid SNPs were matched to the reading frame to determine whether they were synonymous or nonsynonymous. Fixed sites were identified using the same quality and coverage thresholds used to call SNPs.

We explicitly tested whether our power to identify SNPs is equal across the Z and autosomes, despite differences in sequencing coverage. We generated random diploid populations of individuals with varying minor allele frequencies. From these populations, we sampled 20 (autosomal) and 15 (Z-linked) alleles separately 1000 times without replacement and for each sample determined the presence or absence of polymorphism. At a minor allele frequency of 0.15%, the false-negative rate for both the autosomes and Z chromosome was very low (autosomes = 0.023, Z chromosome = 0.068), although marginally lower for the autosomes. We also repeated analyses using a minor allele frequency threshold of 25% (false-negative rate autosomes = 0.001, Z chromosome = 0.009); however, our power is limited at this threshold due to a large reduction in detectable SNPs (Tables S3 and S4, Supporting Information). Our conclusions were broadly comparable across both minor allele frequency thresholds.

For each species, mean nonsynonymous polymorphism ( $p_N$ ), synonymous polymorphism ( $p_S$ ) and  $p_N/p_S$  were calculated separately for Z-linked and autosomal 1–10 orthogroups. Specifically, mean polymorphism was calculated as the sum of the number of polymorphic sites across all contigs in a given genomic category divided by the number of sites ( $p_N = \text{sum } P_N / \text{sum } N$ ,

$p_S = \text{sum } P_S / \text{sum } S$  where  $P_{N/S}$  is the number of non-synonymous/synonymous polymorphic sites and  $N/S$  is the number of nonsynonymous/synonymous sites). Faster-Z was calculated as  $p_{NZ}/p_{SZ}$ :  $p_{NA}/p_{SA}$ . Bootstrapping with 1000 repetitions was used to generate 95% confidence intervals, and significance differences between genomic categories were determined from 1000 permutation tests.

For each species, we used the McDonald–Kreitman test (McDonald & Kreitman 1991) to estimate the number of contigs evolving under adaptive and neutral evolution. The McDonald–Kreitman test contrasts the number of nonsynonymous and synonymous substitutions ( $D_N$  and  $D_S$ ) with polymorphisms ( $P_N$  and  $P_S$ ).  $D_N$  and  $D_S$  for each species were obtained from the species-specific PAML analysis, where divergence was calculated from the terminal tip to the Galloanserae common ancestor, excluding the *T. guttata* outgroup. A deficit of nonsynonymous polymorphisms relative to substitutions is indicative of positive selection [ $(D_N/D_S) > (P_N/P_S)$ ], and an excess of nonsynonymous polymorphisms relative to substitutions is indicative of relaxed purifying selection [ $(D_N/D_S) < (P_N/P_S)$ ]. For each contig, we tested for departures from neutrality using a  $2 \times 2$  contingency table and Pearson's chi-squared test (Hope 1968; Patefield 1981) in R v3.1.0 (R Core Team 2014). Contigs were only included in the analysis if the sum of each marginal row and column of the  $2 \times 2$  contingency table was greater or equal than 6 (Begun *et al.* 2007; Andolfatto 2008). We used the qvalue function in R with a false discovery rate = 0.05 and lambda = 0 to correct for multiple testing. After identifying contigs with signatures of positive selection, we tested for significant differences in the proportion of these contigs on the Z chromosome vs. the autosomes using Pearson's chi-squared test in R.

Lastly, we used polymorphism data to test for an excess or under-representation of Z-linked nonsynonymous polymorphisms relative to the autosomes. Excess or underrepresentation is indicative of relaxed purifying selection or positive selection, respectively. For this analysis, we separately concatenated  $P_N$  and  $P_S$  for each species and used Pearson's chi-squared test to test for significant differences in  $P_N/P_S$  between the Z chromosome and autosomes (Mank *et al.* 2007a).

#### Calculating relative effective population size of the Z chromosome

We calculated the effective population size ( $N_E$ ) of the Z chromosome and autosomes 1–10 for each species using two separate approaches based on  $\pi$  and  $\theta$ .

For each contig, the number of fourfold degenerate sites (4D) and polymorphic fourfold degenerate sites

( $P_{4D}$ ) was calculated. Nucleotide diversity was calculated for each genomic category as  $\pi = \text{sum } P_{4D} / \text{sum } 4D$ . Watterson's estimator of theta ( $\theta$ ) (Watterson 1975) was also calculated as  $\theta = \text{sum } 4D / (\text{sum}[i = 1 \dots n - 1] 1/i)$  where  $n$  is the number of chromosomes in the sample.  $\theta$  per site was then calculated. Finally, we recalculated  $\pi$  and  $\theta$  using all polymorphic synonymous sites.

Effective population size was calculated separately for the Z and autosomes as  $N_E = (\pi \text{ or } \theta) / [4 * (U * \text{generation time})]$ . The mutation rate per site per year ( $U$ ) was calculated separately for the Z chromosome ( $1.45E-09$ ) and autosomes ( $1.33E-09$ ) to account for male-mutation bias, using previous Galliform estimates of Z-linked and autosomal divergence (Dimcheff *et al.* 2002; Axelson *et al.* 2004; van Tuinen & Dyke 2004; Mank *et al.* 2010a).  $U = K/2T$ , where  $K$  is the no of substitutions per site between homologous sequences and  $T$  is divergence time. Bootstrapping with 1000 repetitions was used to generate 95% confidence intervals for effective population size estimates.

#### Tests of positive selection using gene expression

The relative role of selection vs. drift in driving Faster-Z Evolution can be disentangled using gene expression (Baines *et al.* 2008; Mank *et al.* 2010b; Sackton *et al.* 2014). Gene expression was quantified using only adult gonad samples, because this tissue exhibits the greatest magnitude of sex-biased transcription (Mank *et al.* 2007b; Pointer *et al.* 2013) and therefore maximizes the number of female-biased contigs used in the analysis. Expression was estimated as reads per kilobase per million mappable reads (RPKM) and normalized to control for differences in sequencing depth across samples (Brawand *et al.* 2011).

Mean male and female RPKM of each orthogroup were calculated separately for each species, together with fold change [a measure of sex-bias:  $\log_2(\text{male RPKM}) - \log_2(\text{female RPKM})$ ]. A  $t$ -test was used to identify significantly sex-biased contigs, and the Benjamini-Hochberg method (FDR of 5%) (Benjamini & Hochberg 1995) used to correct for multiple testing (Mank *et al.* 2010c; Pointer *et al.* 2013; Perry *et al.* 2014). Female-biased and male-biased contigs were classified as significantly sex-biased ( $P < 0.05$ ) or sex-limited with a  $\log_2$  fold change of  $< -1$  and  $> 1$ , respectively. Unbiased contigs had a  $\log_2$  fold change between  $< 1$  and  $> -1$ .

To verify that our method of defining sex bias was consistent with other approaches, we also used EDGER to categorize sex bias and compared the overlap between both approaches. Briefly, for each species, we extracted raw read counts for 2 RPKM filtered contigs from RSEM (Li & Dewey 2011), normalized to control for differences in sequencing depth across samples using TMM

in EDGER and tested for sex-biased gene expression using the exactTest function in EDGER (Robinson & Oshlack 2010; Robinson *et al.* 2010; McCarthy *et al.* 2012). Female-biased and male-biased contigs were classified as above using a significant  $P$ -value and  $\log_2$  fold change of  $< -1$  and  $> 1$ , respectively. Our approach of categorizing sex bias was consistent with the results from EDGER, and we observe an overlap of 89–96% between expression categories as defined by both approaches.

We used three approaches to test the predictions of the selection and drift hypotheses. First, we calculated Faster-Z for orthogroups where expression category was conserved across all six species. This was to avoid diluting significant signals of selection or drift by including orthogroups where exposure to the dominant evolutionary force has not been consistent over time due to rapid expression turnover. Mean  $d_N$ ,  $d_S$  and  $d_N/d_S$  were calculated separately for each expression category for Z-linked and autosomal contigs using divergence estimates from the Galloanserae analysis in CODEML (Yang 2007). Bootstrapping with 1000 repetitions was used to generate 95% confidence intervals. Significant differences between genomic categories were determined using permutation tests with 1000 repetitions.

We then repeated this analysis with relaxed criteria to maximize the number of orthogroups in each expression category. Specifically, we compared the Faster-Z Effect between putatively female-biased contigs (defined as contigs where at least half of the species had female-limited or significantly female-biased expression, and the fold change was  $< 0$  across all species) and male-biased contigs (where at least half of the species had male-limited or significantly male-biased expression, and the fold change was  $> 0$  across all species).

Finally, we assessed the relationship between species-specific Faster-Z Evolution and gene expression. For each species, we separately calculated  $d_{NZ}/d_{SZ}$ :  $d_{NA}/d_{SA}$  for female-, male- and unbiased contigs for each species as defined with  $t$ -tests and fold change thresholds. Significance was assessed using permutation tests with 1000 repetitions.

#### Gene ontology analysis

We used GORILLA (Eden *et al.* 2007, 2009) to perform a Gene Ontology enrichment analysis to test for enriched gene function terms for Z-linked contigs compared with the autosomes. Mouse reciprocal orthologs were identified using BIOMART (ENSEMBL v.77) for Z-linked and autosomal 1–10 orthologs. The target list contained Z-linked orthologs and the background list contained autosomal orthologs.  $P$ -values were corrected for multiple testing

using the Benjamini–Hochberg method (Benjamini & Hochberg 1995).

## Results

### Faster–Z Evolution

We assembled *de novo* transcriptomes for six Galloanserae species, spanning approximately 90 million years of avian evolution van Tuinen and Hedges (2001) (Fig. 2), and identified 160 Z-linked and 2431 autosomal orthogroups. Across the Galloanserae, mean  $d_N/d_S$  of the Z chromosome is significantly higher than that of the autosomes, due to significantly elevated  $d_{NZ}$  (Table 1, Fig. 3). There is no difference in  $d_S$  between the Z chromosomes and all autosomes ( $P = 0.865$ ).

Seven-hundred and forty-one autosomal orthogroups are located on microchromosomes in the chicken genome, and microchromosomes exhibit different genomic properties to the rest of the autosomes. These properties impact the nature and efficacy of selection (Burt 2002; Ellegren 2013); therefore, the fairest measure of Faster-Z Evolution is to contrast divergence between the Z chromosome and similar-sized autosomes 1–10 (Mank *et al.* 2010b). We identified 1690 orthogroups located on autosomes 1–10. Mean  $d_{NZ}/d_{SZ}$  and  $d_{NZ}$  are both significantly higher than mean  $d_N/d_S$  and  $d_N$  of autosomal 1–10 orthogroups (Table 1, Fig. 3). This pattern is

consistent with the results of the previous analysis using all autosomes, and with previous estimates of Faster-Z Evolution in birds (Mank *et al.* 2007a, 2010b; Dalloul *et al.* 2010; Ellegren *et al.* 2012; Wang *et al.* 2014a). For the rest of the manuscript, autosomal will refer to autosomal 1–10 orthogroups and Faster-Z will refer to the comparison between Z-linked and autosomal 1–10 orthogroups  $d_{NZ}/d_{SZ}$ :  $d_{NA}/d_{SA}$ .

In each of the six Galloanserae species,  $d_{NZ}/d_{SZ}$  is higher than  $d_{NA}/d_{SA}$  based on the species-specific analysis, and there is interspecific variation in the magnitude of this difference (Table 2). We find no significant difference in  $d_S$  between the Z chromosome and autosomes for any species, consistent with previous findings that male-biased mutation rate is weak across the Galloanserae (Bartosch-Harlid *et al.* 2003; Axelsson *et al.* 2004). This suggests that Z-linked mutation rate does not vary significantly across the six species (addressed further in the Discussion).

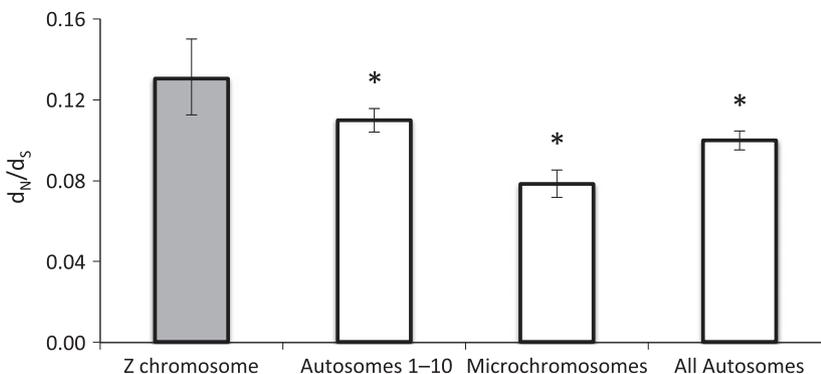
### Variation in sperm competition drives Faster-Z Evolution

The intensity of sperm competition, a widely used indicator of postcopulatory sexual selection and therefore one measure of variance in male mating success, is strongly predicted by relative testes weight and sperm number in birds (Moller 1991; Birkhead & Moller 1998;

**Table 1**  $d_N$ ,  $d_S$  and  $d_N/d_S$  for Z-linked and autosomal genes across Galloanserae clade

	Z chromosome (160 contigs)	Autosomes 1–10 (1690 contigs)	Microchromosomes (741 contigs)	All autosomes (2431 contigs)
$d_S$ 95% CI	0.432 (0.413–0.454)	0.424 (0.417–0.432) $P = 0.229$	0.510 (0.493–0.528) $P = 1.000$	0.447 (0.440–0.454) $P = 0.865$
$d_N$ 95% CI	0.056 (0.049–0.065)	<b>0.047 (0.044–0.049)</b> $P = 0.007$	<b>0.040 (0.037–0.043)</b> $P < 0.001$	<b>0.045 (0.042–0.047)</b> $P = 0.005$

Significance values were determined from 1000 permutation tests, and bootstrapping with 1000 repetitions was used to generate 95% confidence intervals. Significant differences between autosomal and Z-linked orthogroups are in bold.



**Fig. 3** Estimates of mean  $d_N/d_S$  for loci on autosomes and the Z chromosome across the Galloanserae. Synonymous and nonsynonymous divergence estimates were calculated using the branch model in PAML (Galloanserae analysis). 95% confidence intervals were calculated by bootstrapping with 1000 replicates, and significant differences in  $d_N/d_S$  between autosomal and Z-linked orthogroups (permutation test, 1000 replicates) are indicated (\*).

**Table 2**  $d_N$ ,  $d_S$  and  $d_N/d_S$  for Z-linked and autosomal genes across Galloanserae species

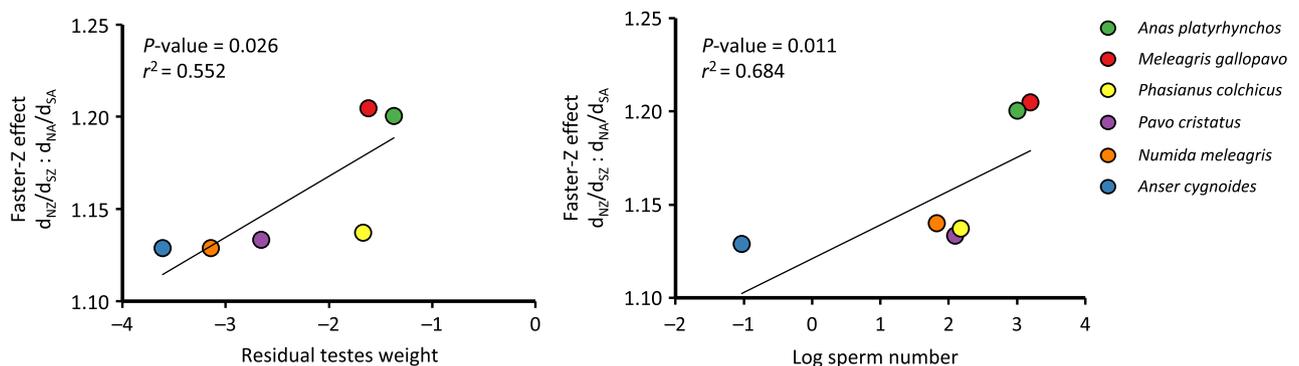
Species	Z chromosome			Autosomes 1–10			Faster-Z Effect $d_{NZ}/d_{SZ}$ : $d_{NA}/d_{SA}$ (95% CI)
	$d_N$ (95% CI)	$d_S$ (95% CI)	$d_N/d_S$ (95% CI)	$d_N$ (95% CI)	$d_S$ (95% CI)	$d_N/d_S$ (95% CI)	
<i>Meleagris gallopavo</i>	0.023 (0.020–0.027)	0.163 (0.155–0.170)	0.144 (0.123–0.165)	<b>0.019</b> <b>(0.018–0.020)</b> $P = 0.005$	0.158 (0.154–0.161) $P = 0.168$	<b>0.120</b> <b>(0.113–0.127)</b> $P = 0.011$	1.205 (1.035–1.390)
<i>Phasianus colchicus</i>	0.021 (0.018–0.025)	0.161 (0.153–0.168)	0.134 (0.114–0.154)	<b>0.018</b> <b>(0.017–0.020)</b> $P = 0.035$	0.157 (0.153–0.160) $P = 0.215$	0.118 (0.111–0.125) $P = 0.061$	1.137 (0.961–1.331)
<i>Numida meleagris</i>	0.019 (0.016–0.022)	0.133 (0.127–0.140)	0.140 (0.119–0.162)	<b>0.016</b> <b>(0.015–0.017)</b> $P = 0.041$	0.132 (0.129–0.135) $P = 0.393$	<b>0.123</b> <b>(0.116–0.130)</b> $P = 0.049$	1.140 (0.965–1.332)
<i>Anas platyrhynchos</i>	0.015 (0.012–0.018)	0.116 (0.108–0.126)	0.131 (0.107–0.155)	<b>0.013</b> <b>(0.012–0.014)</b> $P = 0.030$	0.116 (0.113–0.119) $P = 0.518$	<b>0.109</b> <b>(0.103–0.116)</b> $P = 0.024$	1.200 (0.974–1.457)
<i>Anser cygnoides</i>	0.012 (0.010–0.015)	0.100 (0.093–0.107)	0.125 (0.103–0.148)	0.011 (0.010–0.012) $P = 0.083$	0.099 (0.097–0.101) $P = 0.378$	0.111 (0.105–0.118) $P = 0.083$	1.129 (0.939–1.360)
<i>Pavo cristatus</i>	0.020 (0.017–0.023)	0.147 (0.139–0.154)	0.134 (0.114–0.157)	0.017 (0.016–0.018) $P = 0.068$	0.147 (0.144–0.150) $P = 0.502$	0.118 (0.112–0.125) $P = 0.056$	1.133 (0.951–1.303)

Significance values were determined from 1000 permutation tests and bootstrapping with 1000 repetitions was used to generate 95% confidence intervals. Significant differences between autosomal and Z-linked orthologs are shown in bold.

Pitcher *et al.* 2005). We recovered a significant positive association between magnitude of Faster-Z Evolution and both log sperm number ( $r^2 = 0.684$ ,  $P = 0.011$ ,  $t_4 = 3.629$ ) and residual testes weight ( $r^2 = 0.552$ ,  $P = 0.026$ ,  $t_4 = 2.744$ ) after correcting for phylogeny (Fig. 4). To test the strength of these associations, we sequentially removed each species and repeated the analyses (Table S5). Despite the reduction in sample size and therefore statistical power, there was no change to the significance or direction of the slope for log sperm number. For residual testes weight, there

was no change to the direction of the slope but when either *A. cygnoides* or *A. platyrhynchos* was excluded, the relationship was nonsignificant (Table S5).

There are two plausible explanations for our finding that the magnitude of Z-linked divergence increases with increasing female promiscuity. A recent study in silk moths has shown that Faster-Z Evolution is adaptive, and results from increased efficacy of selection acting on recessive advantageous mutations in the hemizygous sex (Sackton *et al.* 2014). Conversely, a study in birds suggested that avian Faster-Z Evolution



**Fig. 4** Phylogenetically controlled regression between proxies of sperm competition and Faster-Z Effect. Data points are raw species values but  $P$ -values and  $r^2$  estimates were calculated using phylogenetic generalized least squares regression with maximum likelihood and 1000 runs for each analysis. Autosomes refers to macrochromosomes (autosomes 1–10).

**Table 3** Effective population size estimates of the Z chromosome and autosomes

Species	$N_{EZ}$ (E + 05) (95% CI)	$N_{EA1-10}$ (E + 05) (95% CI)	$N_{EZ}/N_{EA1-10}$ (95% CI)
<i>Meleagris gallopavo</i>	1.761 (1.087–2.702)	6.047 (5.656–6.469)	0.291 (0.179–0.426)
<i>Phasianus colchicus</i>	3.188 (2.308–4.210)	9.481 (8.948–10.054)	0.336 (0.234–0.460)
<i>Numida meleagris</i>	1.695 (0.773–3.213)	7.233 (6.682–7.848)	0.234 (0.103–0.423)
<i>Anas platyrhynchos</i>	6.150 (3.927–8.758)	18.427 (17.447–19.544)	0.334 (0.209–0.470)
<i>Anser cygnoides</i>	4.045 (2.774–5.591)	10.894 (10.233–11.570)	0.371 (0.250–0.529)
<i>Pavo cristatus</i>	1.088 (0.167–2.811)	2.393 (2.095–2.697)	0.455 (0.057–1.227)

$N_E$  was calculated using the same method as Mank *et al.* 2010b; Mutation rate estimates are from Axelsson *et al.* 2004; Dimcheff *et al.* 2002 and van Tuinen & Dyke 2004.

Minor allele frequency threshold of 0.15.

Nucleotide diversity ( $\pi$ ) was calculating using fourfold degenerate sites.

is a neutral process, driven by relaxed efficacy of purifying selection as a consequence of relative differences in  $N_{EZ}/N_{EA}$  (Mank *et al.* 2010b). Under the latter hypothesis, variation in male reproductive success, associated with sexual selection, is predicted to alter the relationship between  $N_{EZ}$  and  $N_{EA}$ , and therefore the relative magnitude of drift acting on the Z chromosome (Charlesworth *et al.* 1993; Vicoso & Charlesworth 2009). Specifically, with increasing variance in male reproductive success, relative  $N_{EZ}$  decreases, resulting in greater magnitude of drift and therefore Faster-Z Effect (Wright & Mank 2013).

We use sequence divergence, polymorphism and expression data to test whether the relationship between female promiscuity and Faster-Z Evolution is adaptive or neutral.

#### Estimates of relative $N_{EZ}$

After filtering for quality and read depth, across Z-linked and autosomal 1–10 contigs, we identified 12 436 SNPs in *A. platyrhynchos*, 4584 in *M. gallopavo*, 6850 in *P. colchicus*, 5205 in *N. meleagris*, 2012 in *P. cristatus* and 8128 in *A. cygnoides* (Table S3).

For each species, we calculated the effective population size of the Z chromosome ( $N_{EZ}$ ) and autosomes 1–10 ( $N_{EA}$ ) using a number of approaches. We accounted for male-biased mutation rate and generation time using previous Galliform estimates (Dimcheff *et al.* 2002; Axelsson *et al.* 2004; van Tuinen & Dyke 2004; Mank *et al.* 2010a) (Table 3, Tables S6, S7 and S8, Supporting Information) (Vicoso & Charlesworth 2009). Under strict monogamy,  $N_{EZ}$  is predicted to equal  $\frac{3}{4} N_{EA}$ . For all species with the exception of *P. cristatus*,  $N_{EZ}$  was significantly  $<\frac{3}{4} N_{EA}$ . However, the 95% CI for this species was unusually wide, probably as a result of the low frequency of SNPs detected (Table S3).

The relationship between  $N_{EZ}/N_{EA}$  and sperm number, residual testes weight or Faster-Z was not statisti-

cally significant (sperm number:  $r^2 = 0.083$ ,  $P = 0.252$ ,  $t_4 = 0.735$ ; residual testes weight:  $r^2 = 0.068$ ,  $P = 0.275$ ,  $t_4 = 0.656$ ; Faster-Z:  $r^2 = 0.220$ ,  $P = 0.132$ ,  $t_4 = 1.300$ ; Table S9, Supporting Information). Additionally, the autosomal effective population size of *P. cristatus* is significantly smaller than the other six species, indicating either a very recent bottleneck or variation in family structure across the individuals sampled in this study. This finding hints at the sensitivity of  $N_E$  calculations to many factors (Hartl & Clark 2007), including recombination rate and recent demographic perturbations (Pool & Nielsen 2007). This may explain both the unusually low  $N_E$  estimates in *P. cristatus* as well as the lack of significant association between  $N_{EZ}/N_{EA}$  and measures of sperm competition (addressed further in the Discussion).

#### Tests of positive selection

We used sequence and polymorphism data from our six species to test whether selection is more effective for Z-linked vs. autosomal loci. Using the site-model test in CODEML, we found significant evidence for positive selection acting on 5/160 Z-linked loci (1/160 after sequential Bonferroni's correction) and 51/1690 autosomal loci (5/1690 after sequential Bonferroni's correction) (Table 4, Table S10, Supporting Information). There was no significant difference in the proportion of positively selected loci on the Z chromosome or autosomes 1–10 either before or after multiple testing correction ( $\chi^2$ , d.f. = 1,  $P > 0.400$  in both comparisons). This indicates that selection is not more effective on the Z chromosome; however, the power of this analysis is limited by the low number of total contigs under positive selection.

We next used polymorphism data to test for deviations from neutrality. With the exception of *N. meleagris* and *P. cristatus*,  $p_{NZ}/p_{SZ}$  is significantly greater than  $p_{NA}/p_{SA}$  (Table 5, Table S11, Supporting Information). This finding of excess nonsynonymous polymorphism

**Table 4** Site-model test results for contigs under positive selection

<i>G. gallus</i> ortholog*	Chromosome	$\omega$	Proportion of sites	M1a likelihood ratio	M2a likelihood ratio	LRT	<i>P</i> -value	<i>P</i> -fdr value <sup>†</sup>
22552	1	2.897	0.122	-6535.857	-6522.227	27.259	<0.001	0.003
21101	1	4.155	0.033	-14063.297	-14050.286	26.023	<0.001	0.006
31776	3	4.608	0.130	-1270.098	-1256.430	27.337	<0.001	0.003
39919	6	4.226	0.310	-1630.735	-1611.278	38.915	<0.001	<0.001
03831	8	4.817	0.080	-9607.226	-9560.287	93.878	<0.001	<0.001
10504	15	3.343	0.072	-5389.616	-5375.473	28.287	<0.001	0.002
01868	20	9.422	0.013	-4192.958	-4179.195	27.526	<0.001	0.003
02022	28	4.914	0.068	-2768.690	-2753.634	30.110	<0.001	0.001

\*ENSGALT000000.

†Sequential Bonferroni's correction (Holm 1979).

on the Z chromosome relative to the autosomes suggests that selection is less effective at removing mildly deleterious mutations from the Z chromosome. This finding is consistent with the drift hypothesis of Faster-Z, rather than the adaptive hypothesis. Interestingly, *N. meleagris* Z chromosome exhibits a nonsignificant deficit of  $p_N$ , potentially as a consequence of monogamy, which would maximize  $N_{EZ}/N_{EA}$  and therefore the potential of selection to act on the Z chromosome in this species.

For each species, we estimated the number of contigs evolving under adaptive evolution using the McDonald–Kreitman test (McDonald & Kreitman 1991). This test contrasts the number of nonsynonymous and synonymous substitutions ( $D_N$  and  $D_S$ ) with polymorphisms ( $P_N$  and  $P_S$ ) for each contig. An excess of nonsynonymous substitutions relative to polymorphism is indicative of positive selection [ $(D_N/D_S) > (P_N/P_S)$ ], and under-representation of nonsynonymous substitutions relative to polymorphism is indicative of relaxed purifying selection [ $(D_N/D_S) < (P_N/P_S)$ ]. We detected no Z-linked contigs with signatures of positive selection, and there was no difference between the Z chromosome and autosomes 1–10 in the proportion of loci under positive selection in any species ( $\chi^2$ , d.f. = 1,  $P > 0.500$  in all cases) (Table S12, Supporting Information). However, only contigs with sufficient numbers of substitutions and polymorphisms were included in the analysis (Begun *et al.* 2007; Andolfatto 2008), and therefore, our ability to draw species-specific conclusions is limited by low sample sizes.

Lastly, for each species, we concatenated the number of  $P_N$  and  $P_S$  across all Z-linked and all autosomal 1–10 contigs separately (Table 6, Table S13, Supporting Information) and tested for significant differences between Z-linked and autosomal  $P_N/P_S$ . For each species, there is a significant excess of Z-linked nonsynonymous polymorphism relative to the autosomes for all species with the exceptions of *P. cristatus* and *N. meleagris*. This is

again consistent with a reduction in the power of selection to remove mildly deleterious alleles from this chromosome.

The lack of difference in Z-linked and autosomal nonsynonymous polymorphism in *P. cristatus* and *N. meleagris* could be attributed to a number of factors. It could reflect biological differences in sexual selection and therefore the magnitude of drift acting on the Z chromosome. However, although this explanation is consistent with the monogamous mating system of *N. meleagris*, it is not consistent with the *P. cristatus*, which exhibits a lek mating system (Petrie *et al.* 1999). More likely, this pattern reflects the limitations of polymorphism data and the difficulty in controlling for family structure and demographic effects (Hartl & Clark 2007). For example, the number of SNPs in *P. cristatus* is much lower than the other five species, and therefore, the statistical power of this analysis is limited (Table 6).

Differences in gene content between the sex chromosomes and autosomes can contribute to observed patterns of Faster-Z/X (Meisel & Connallon 2013) by biasing the potential for positive selection in different genomic categories. However, the results of our GORILLA functional enrichment test reveal no significantly enriched gene ontology terms for Z-linked orthogroups compared with autosomes 1–10 after correcting for multiple tests.

### Gene expression

We used gene expression data from gonads of our six avian species to identify the dominant force driving Faster-Z Evolution across the Galloanserae clade. If Faster-Z Evolution is adaptive and driven by increased efficacy of selection acting on recessive mutations in the hemizygous sex, we predict the Faster-Z Effect to be largest for female-biased, followed by unbiased and then male-biased genes. If it is due to neutral causes, there will be no difference in the rate of Faster-Z

**Table 5**  $p_N$ ,  $p_S$  and  $p_N/p_S$  for Z-linked and autosomal genes across Galloanserae species

Species	Z chromosome			Autosomes 1–10			Faster-Z Effect $p_{NZ}/p_{SZ}$ : $p_{NA}/p_{SA}$ (95% CI)
	$p_N$ (95% CI)	$p_S$ (95% CI)	$p_N/p_S$ (95% CI)	$p_N$ (95% CI)	$p_S$ (95% CI)	$p_N/p_S$ (95% CI)	
<i>Melaneris gallopazo</i>	0.000 (0.000–0.000)	0.001 (0.001–0.002)	0.176 (0.112–0.256)	0.000 (0.000–0.001) $P = 1.000$	0.005 (0.005–0.005) $P = 1.000$	<b>0.102 (0.093–0.111)</b> $P < 0.001$	1.721 (1.039–2.716)
<i>Phasianus colchicus</i>	0.000 (0.000–0.001)	0.002 (0.002–0.003)	0.162 (0.109–0.236)	0.001 (0.001–0.001) $P = 1.000$	0.007 (0.007–0.008) $P = 1.000$	<b>0.095 (0.088–0.103)</b> $P < 0.001$	1.704 (1.122–2.546)
<i>Numida meleagris</i>	0.000 (0.000–0.000)	0.002 (0.001–0.003)	0.083 (0.050–0.150)	0.001 (0.001–0.001) $P = 1.000$	0.005 (0.005–0.006) $P = 1.000$	0.102 (0.093–0.112) $P = 0.897$	0.813 (0.483–1.370)
<i>Anas platyrhynchos</i>	0.001 (0.000–0.001)	0.005 (0.004–0.008)	0.103 (0.064–0.156)	0.001 (0.001–0.001) $P = 1.000$	0.014 (0.013–0.015) $P = 1.000$	<b>0.072 (0.066–0.078)</b> $P = 0.002$	1.426 (0.863–2.120)
<i>Anser cygnoides</i>	0.001 (0.000–0.001)	0.003 (0.002–0.004)	0.177 (0.109–0.262)	0.001 (0.001–0.001) $P = 0.999$	0.008 (0.008–0.009) $P = 1.000$	<b>0.108 (0.099–0.117)</b> $P < 0.001$	1.642 (1.009–2.427)
<i>Pavo cristatus</i>	0.000 (0.000–0.000)	0.001 (0.000–0.002)	0.173 (0.096–0.541)	0.000 (0.000–0.000) $P = 0.930$	0.002 (0.002–0.002) $P = 1.000$	0.134 (0.116–0.156) $P = 0.137$	1.293 (0.681–4.002)

Significance values were determined from 1000 permutation tests, and bootstrapping with 1000 repetitions was used to generate 95% confidence intervals. Significant differences between autosomal and Z-linked orthologs are shown in bold. Minor allele frequency threshold of 0.15.

Evolution among expression classes (Baines *et al.* 2008; Mank *et al.* 2010b; Sackton *et al.* 2014). We tested this prediction at three levels in our data.

First, we identified orthogroups with consistent male-, female- and unbiased expression across all six species, thereby excluding any orthogroups where the nature of sex-bias, and therefore exposure to the dominant evolutionary force, has varied over Galloanserae evolutionary history. The rapid change in sex bias across this clade (Harrison *et al.* in press) means that relatively few genes are consistently sex-biased in our data set, resulting in 17 male-biased, 9 female-biased and 7 unbiased Z-linked orthogroups alongside 104 male-biased, 116 female-biased and 205 unbiased autosomal orthogroups. Among these gene sets, there was no significant difference in Faster-Z Effect (male-biased vs. female-biased  $P = 0.542$ , female-biased vs. unbiased  $P = 1.000$ , male-biased vs. unbiased  $P = 0.616$ , all two-tailed pairwise permutation tests with 1000 repetitions), shown in Fig. 5.

To exclude the possibility that we lack statistical power to distinguish between drift and selection due to low sample sizes, we next repeated the analysis and relaxed the definition of sex bias (see Materials and Methods). In doing so, we nearly doubled the number of orthogroups in each expression category; identifying 54 male-biased and 15 female-biased Z-linked orthogroups, together with 347 male-biased and 319 female-biased autosomal orthogroups. Again, there was no significant difference in Faster-Z Effect between these gene sets ( $P = 0.916$ , permutation test, 1000 repetitions), with female-biased  $d_{NZ}/d_{SZ}$ :  $d_{NA}/d_{SA} = 1.491$  (95% CI = 0.997–2.137) and male-biased  $d_{NZ}/d_{SZ}$ :  $d_{NA}/d_{SA} = 1.456$  (95% CI = 1.112–1.869).

Finally, we assessed whether there was any species-specific pattern in Faster-Z Evolution across male-, female- and unbiased contigs. There is no significant difference between Faster-Z of any expression category in any species after correction for multiple testing, with the exception of *N. meleagris* where we found a significantly larger Faster-Z Effect for male-biased compared with unbiased contigs (Tables S14 and S15, Supporting Information). At all three levels of analysis, our expression data are consistent with Faster-Z Evolution resulting predominantly from neutral forces.

## Discussion

Faster rates of coding sequence evolution on the Z chromosome relative to the autosomes have been observed across a wide range of species (Mank *et al.* 2007a, 2010b; Dalloul *et al.* 2010; Ellegren *et al.* 2012; Sackton *et al.* 2014; Wang *et al.* 2014a,b); however, the underlying cause is unclear. Indirect evidence from an

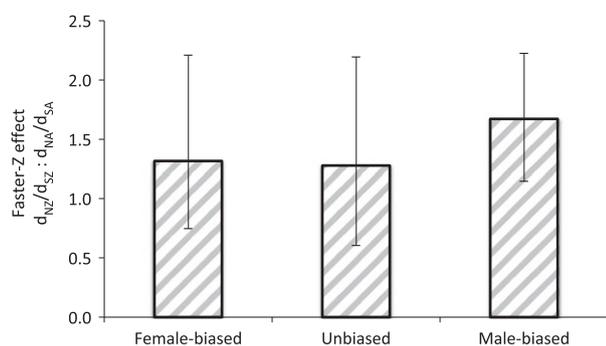
**Table 6** Significant differences between nonsynonymous and synonymous polymorphism on the Z chromosome and autosomes

Species	Z chromosome		Autosomes 1–10		Faster-Z Effect $P_{NZ}/P_{SZ}$ : $P_{NA}/P_{SA}$ P-value
	$P_N$	$P_S$	$P_N$	$P_S$	
<i>Meleagris gallopavo</i>	51	83	1174	3276	<b>1.715</b> <b><math>P = 0.004</math></b>
<i>Phasianus colchicus</i>	89	157	1654	4950	<b>1.700</b> <b><math>P &lt; 0.001</math></b>
<i>Numida meleagris</i>	29	100	1339	3737	0.809 $P = 0.372$
<i>Anas platyrhynchos</i>	126	351	2417	9542	<b>1.417</b> <b><math>P = 0.001</math></b>
<i>Anser cygnoides</i>	127	206	2138	5657	<b>1.631</b> <b><math>P &lt; 0.001</math></b>
<i>Pavo cristatus</i>	38	63	610	1301	1.286 $P = 0.277$

Significant differences were determined using Pearson's chi-squared test in R.

Significant differences between autosomal and Z-linked orthologs are shown in bold.

Minor allele frequency threshold of 0.15.



**Fig. 5** Estimates of mean Faster-Z across sex-biased gene expression categories. Sex bias was defined using fold change thresholds and *t*-tests. 95% confidence intervals were calculated by bootstrapping with 1000 replicates. Autosomal orthologs were limited to chromosomes 1–10.

expression-based approach suggests that avian Faster-Z Evolution is driven by genetic drift (Mank *et al.* 2010b), but a recent study in silk moths postulated an adaptive explanation (Sackton *et al.* 2014). To determine the cause of Faster-Z Evolution in birds, we assembled *de novo* transcriptomes for six Galloanserae species, spanning 90 million years of avian evolution and combined expression, sequence and polymorphism data with measures of sperm competition and promiscuity. We present the first empirical evidence demonstrating the positive relationship between the Faster-Z Effect and measures of postcopulatory sexual selection and variance in male reproductive success.

This pattern is consistent with a considerable body of theory predicting that Faster-Z Evolution in birds is

driven by changes in the relative strength of genetic drift as a result of increased variance in male reproductive success (Vicoso & Charlesworth 2009). In support of the predominant role of genetic drift in shaping rates of Z chromosome evolution, we used multiple sequence-, polymorphism- and expression-based approaches. Our expression analysis is consistent with previous work that found no difference in Faster-Z Evolution among sex-biased expression categories (Mank *et al.* 2010b). However, our analysis significantly extends this previous work by incorporating tests of positive selection based on divergence and polymorphism. The results from these multiple lines of evidence are broadly convergent, indicating that selection is not more effective on the Z chromosome. We conclude that Faster-Z Evolution in birds is due primarily to relaxed power of purifying selection and that the magnitude of this effect is dependent on the nature of sexual selection.

#### *Promiscuity and sperm competition are drivers of Faster-Z Evolution*

Changes in the skew of male reproductive success are commonly associated with promiscuity and the intensity of postcopulatory sexual selection (Andersson 1994), both of which decrease the  $N_{EZ}/N_{EA}$  ratio. If Faster-Z is neutral and nonadaptive, we predict that the magnitude of Faster-Z Evolution should increase as  $N_{EZ}/N_{EA}$  decreases (Vicoso & Charlesworth 2009), and therefore, we should expect both lower  $N_{EZ}/N_{EA}$  and increased rates of Faster-Z Evolution in promiscuous compared with monogamous populations (Fig. 1).

We uncovered a significant and positive association between the magnitude of Faster-Z and relative testes weight and sperm number, both reliable predictors of the intensity of sperm competition in birds (Fig. 4) (Moller 1991; Birkhead & Moller 1998; Pitcher *et al.* 2005). Sperm competition is a widely used indicator of the strength of postcopulatory sexual selection and therefore a good proxy for variance in male mating success and the magnitude of drift acting on the Z chromosome (Moller 1991; Birkhead & Moller 1998; Dorus *et al.* 2004). It is even possible we have underestimated the role of male mating success in driving Z chromosome divergence, as the birds sampled in this study have a lower testes weight than expected given their body weight (Pitcher *et al.* 2005).

Although the relationship between  $N_{EZ}/N_{EA}$  and sperm number or residual testes weight was not significant,  $N_{EZ}/N_{EA}$  across the Galloanserae is consistent with the nonadaptive hypothesis of Faster-Z Evolution (Vicoso & Charlesworth 2009) and is significantly less than the 0.75 predicted under strict monogamy, with the exception of *P. cristatus* (Table 3). We calculated effective population size using parameters estimated from previous Galliform studies (Dimcheff *et al.* 2002; Axelsson *et al.* 2004; van Tuinen & Dyke 2004; Mank *et al.* 2010a), and although mutation rate, male-biased mutation and generation time are not expected to vary substantially across the Galloanserae, we might expect slight differences. Overall  $N_E$  is also predicted to have a large effect on the magnitude of Faster-Z and relative contribution of nonadaptive and adaptive evolutionary forces. However, patterns of autosomal  $N_E$  do not reflect differences in Faster-Z across species.

Polymorphism estimates are sensitive to recent demographic perturbations, bottlenecks and recombination rate (Hartl & Clark 2007). Changes in population size have been shown to differentially impact  $N_{EZ}$  relative to  $N_{EA}$  and variation in population history across the Galloanserae may contribute to the lack of a significant relationship between  $N_{EZ}/N_{EA}$  and measures of promiscuity and sperm competition (Pool & Nielsen 2007). Previous attempts to estimate  $N_{EZ}/N_{EA}$  in birds (Corl & Ellegren 2012) showed sizable variation from what would be predicted by mating system, suggesting that  $N_{EZ}/N_{EA}$  estimates may simply be too inaccurate for the types of analyses used here. Because divergence data are not as sensitive to recent demographic perturbations, it can be argued that it is a fairer test for the role of male mating success and sperm competition in Faster-Z Evolution.

#### Tests of positive selection

We used sequence and polymorphism data to test the relative strength of selection on the Z chromosome vs.

autosomes. In both the site-model tests in PAML as well as species-specific McDonald–Kreitman tests, there was no difference in the proportion of positively selected loci on the Z chromosome compared with the autosomes. The McDonald–Kreitman test is limited to sequences with sufficient numbers of substitutions and polymorphisms (McDonald & Kreitman 1991; Andolfatto 2008), and this restricted our analysis to a handful of Z-linked contigs. Therefore, to maximize the power of our data set, we concatenated polymorphism data across all Z-linked and autosomal contigs (Mank *et al.* 2007a). For the majority of species, an excess of Z-linked nonsynonymous polymorphism relative to the autosomes was observed, suggesting that selection is less able to purge mildly deleterious alleles from the Z chromosome. This pattern is consistent with the theoretical expectations of elevated levels of genetic drift. We would expect the opposite pattern, a deficit of Z-linked nonsynonymous polymorphism, under both positive and purifying selection.

Differences in gene content between the sex chromosomes and autosomes can bias the potential for positive selection to act on different genomic categories, and therefore may contribute to our observed patterns of Faster-Z (Meisel & Connallon 2013). The avian Z chromosome is enriched in male-biased genes (Mank & Ellegren 2009), which typically exhibit rapid rates of evolution (Meisel 2011; Parsch & Ellegren 2013). However, we do not find an elevated Faster-Z Effect for male-biased genes, and the results of our GORILLA functional enrichment analysis reinforce that differences in gene content are not likely to drive the pattern of Faster-Z we observe.

Overall, we failed to detect any indication that selection is more effective for Z-linked loci, consistent with the nonadaptive explanations for Faster-Z Evolution. However, it is important to note that our analyses are limited to orthologs conserved across 90 million years, and conservation across this span of time suggests that purifying selection is a dominant force acting on these genes. The important role of purifying selection in this gene set may bias our ability to detect positive selection using this data set. Nevertheless, our neutral explanation of Faster-Z is consistent with previous work indicating that sex chromosome dosage compensation status mediates the contribution of positive selection to Faster-Z Effect (Charlesworth *et al.* 1987; Mank 2009). Theory predicts that the adaptive component of Faster-Z is weaker in species with incomplete dosage compensation, such as birds (Ellegren *et al.* 2007; Mank 2009; Itoh *et al.* 2010; Uebbing *et al.* 2013), compared to those with complete dosage compensation.

Theory predicts that the magnitude of Faster-Z Effect should increase as  $N_{EZ}/N_{EA}$  decreases (Vicoso &

Charlesworth 2009), and therefore, we should expect increased rates of Faster-Z Evolution in promiscuous compared with monogamous populations. This prediction is consistent with our finding that Faster-Z is positively correlated with the intensity of sperm competition, and therefore variance in male reproductive success.

#### *Faster-Z vs. Faster-X Evolution*

Faster rates of coding sequence divergence have repeatedly been documented on the X and Z chromosomes relative to the autosomes, and there is considerable variation in the magnitude of this difference across species (Meisel & Connallon 2013). Moreover, there is a stark contrast between our results and those of Faster-X Evolution in *Drosophila* and mammals, where X-linked male-biased genes evolve more rapidly than unbiased and female-biased genes (Khaitovich *et al.* 2005; Baines *et al.* 2008; Grath & Parsch 2012). This pattern is consistent with an adaptive explanation of Faster-X Evolution driven by increased efficacy of selection acting on recessive mutations in the heterogametic sex. In addition, there is considerable evidence for signatures of adaptation on the X chromosome across many species (Thornton & Long 2005; Baines *et al.* 2008; Hvilsom *et al.* 2012; Langley *et al.* 2012).

The empirical evidence for neutral vs. adaptive explanations of Faster-Z and Faster-X Evolution, respectively, is supported by theoretical predictions (Vicoso & Charlesworth 2009). As variance in male reproductive fitness increases,  $N_{EZ} < \frac{3}{4} N_{EA}$ , reducing the ability of selection to purge mildly deleterious alleles. In contrast,  $N_{EX} > \frac{3}{4} N_{EA}$  under increased variance in male reproductive success, indicating that Faster-X is more often due to positive selection acting on recessive mutations exposed in the heterogametic sex. However, a recent study in silk moths (Sackton *et al.* 2014) indicates that this prediction may not hold for all female heterogametic species and is dependent on numerous other factors, including overall population size and sex-specific recombination rates (Connallon *et al.* 2012).

#### *Male-biased mutation*

The relative rate of Z-linked divergence is thought to be influenced by multiple factors, not only variance in male reproductive success (Kirkpatrick & Hall 2004a; Connallon *et al.* 2012). The number of cell divisions, and therefore potential for mutations, is inherently higher in spermatogenesis compared with oogenesis. This male-biased mutation has been documented across a number of species (Bartosch-Harlid *et al.* 2003; Axelsson *et al.* 2004; Xu *et al.* 2012), and as the Z chromosome is present more often in males than females, it could con-

tribute to the observed differences in relative Z-linked divergence (Kirkpatrick & Hall 2004a; Xu *et al.* 2012). However, previous estimates indicate the magnitude of male-biased mutation may be relatively weak across the Galloanserae (Bartosch-Harlid *et al.* 2003), ranging from 1.6 to 3.8 in Anseriformes (Wang *et al.* 2014b) and 1.7 to 2.52 in Galliformes (Axelsson *et al.* 2004). We failed to find a significant difference between  $d_{SZ}$  and  $d_{SA}$  in any species indicating that male-mutation bias does not vary significantly across this clade. This is consistent with the observation that the wild species in this study are seasonal breeders where spermatogenesis ceases in the nonbreeding season. Consequentially, the difference in number of meiotic cell divisions between males and females is reduced, and therefore, the potential for male-biased mutation is lower. In contrast, many previous estimates of male-biased mutation were based on domesticated species with continuous breeding cycles and spermatogenesis (Bartosch-Harlid *et al.* 2003; Axelsson *et al.* 2004). However, it is possible there is also a confounding effect of Z-linked codon usage bias, an excess of which has been observed on the *Drosophila* X chromosome (Singh *et al.* 2008).

#### *Sexual selection and the Z chromosome*

The sex chromosomes are predicted to play a disproportionate role in encoding sex-specific fitness due to their unequal inheritance pattern (Rice 1984). The Z chromosome in particular is thought to foster tight linkage between female preference genes and flashy male traits, and promote rapid evolution of some types of sexually selected traits (Rice 1984; Reeve & Pfennig 2003; Kirkpatrick & Hall 2004b). However, evidence that the Z chromosome harbours genes encoding sexually dimorphic phenotypes is mixed (Dean & Mank 2014). Z-linked male plumage genes have been documented in flycatchers (Saetre *et al.* 2003; Saether *et al.* 2007), but other studies have failed to find an association between sexually dimorphic traits and sex linkage (Knief *et al.* 2012; Schielzeth *et al.* 2012; Pointer *et al.* 2013). Our findings may help explain this discrepancy between theoretical and empirical data. The low effective population size of the Z chromosome relative to the autosomes may weaken the efficacy of sex-specific selection, particularly in the species under the strongest sexual selection regimes. This may limit the adaptive role of the Z chromosome in general, and in particular its role in encoding sexually selected traits. Given this, it is important to note that our results do not exclude the potential for selection acting on the Z chromosome, but suggests that relaxed purifying selection is more dominant on the Z chromosome relative to the autosomes.

## Conclusions

We assessed the magnitude and drivers of Faster-Z Evolution across a clade of birds spanning 90 million years of evolution. Our analysis combines expression, sequence and polymorphism data with measures of sperm competition and promiscuity. The results from these multiple lines of evidence are broadly convergent, indicating that selection is less effective on the Z chromosome, and suggesting that Faster-Z Evolution in birds is due primarily to genetic drift. Moreover, we present the first empirical evidence demonstrating the positive relationship between the Faster-Z Effect and measures of promiscuity and sperm competition, and therefore variance in male mating success.

## Acknowledgements

We thank Victoria Roberts and John Hall for collection assistance, and Geoff Parker for assistance in securing samples. We also thank Jen Perry, Rebecca Dean and Natasha Bloch for helpful suggestions and comments on the manuscript, Mario dos Reis for help with the PAML analyses and four reviewers for their extremely helpful comments. Sequencing was performed by the Wellcome Trust Centre for Human Genetics Sequencing Hub (funded by Wellcome Trust Grant reference 090532/Z09/Z and MRC Hub Grant G0900747 91070). The authors acknowledge the use of the UCL Legion High Performance Computing Facility (Legion@UCL) and UCL Unity SMP Facility, and associated support services, in the completion of this work. The work described here was funded by the European Research Council under the Framework 7 agreement (grant agreement 260233). A.E.W. is funded by a National Environment Research Council Ph.D. studentship. SHM is funded by a research fellowship from the Royal Commission for the Exhibition of 1851.

## References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *Journal of Molecular Biology*, **215**, 403–410.
- Andersson M (1994) *Sexual Selection*. Princeton University Press, New Jersey, USA.
- Andolfatto P (2008) Controlling type-I error of the McDonald–Kreitman Test in genomewide scans for selection on noncoding DNA. *Genetics*, **180**, 1767–1771.
- Axelsson E, Smith NGC, Sundstrom H, Berlin S, Ellegren H (2004) Male-biased mutation rate and divergence in autosomal, Z-linked and W-linked introns of chicken and turkey. *Molecular Biology and Evolution*, **21**, 1538–1547.
- Axelsson E, Hultin-Rosenberg L, Brandström M, Zwahlen M, Clayton DF, Ellegren H (2008) Natural selection in avian protein-coding genes expressed in brain. *Molecular Ecology*, **17**, 3008–3017.
- 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*. *Molecular Biology and Evolution*, **25**, 1639–1650.
- Bartosch-Harlid A, Berlin S, Smith NGC, Moller AP, Ellegren H (2003) Life history and the male mutation bias. *Evolution*, **57**, 2398–2406.
- Begun DJ, Holloway AK, Stevens K *et al.* (2007) Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. *PLoS Biology*, **5**, 2534–2559.
- Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B*, **57**, 289–300.
- Birkhead T, Moller AP (1998) *Sperm Competition and Sexual Selection*. Academic Press Inc., San Diego, California.
- Birkhead TR, Petrie M (1995) Ejaculate features and sperm utilization in Peafowl *Pavo Cristatus*. *Proceedings of the Royal Society B-Biological Sciences*, **261**, 153–158.
- Brawand D, Soumillon M, Necsulea A *et al.* (2011) The evolution of gene expression levels in mammalian organs. *Nature*, **478**, 343–348.
- Burt DW (2002) Origin and evolution of avian microchromosomes. *Cytogenetic and Genome Research*, **96**, 97–112.
- Charlesworth B (2009) Effective population size and patterns of molecular evolution and variation. *Nature Reviews Genetics*, **10**, 195–205.
- Charlesworth B, Coyne JA, Barton NH (1987) The relative rates of evolution of sex chromosomes and autosomes. *The American Naturalist*, **130**, 113–146.
- Charlesworth B, Morgan MT, Charlesworth D (1993) The effect of deleterious mutations on neutral molecular variation. *Genetics*, **134**, 1289–1303.
- Connallon T, Singh ND, Clark AG (2012) Impact of genetic architecture on the relative rates of X versus autosomal adaptive substitution. *Molecular Biology and Evolution*, **29**, 1933–1942.
- Corl A, Ellegren H (2012) The genomic signature of sexual selection in the genetic diversity of the sex chromosomes and autosomes. *Evolution*, **66**, 2138–2149.
- Dalloul RA, Long JA, Zimin AV *et al.* (2010) Multi-platform next-generation sequencing of the domestic turkey (*Meleagris gallopavo*): genome assembly and analysis. *PLoS Biology*, **8**, e1000475.
- Dean R, Mank JE (2014) The role of sex chromosomes in sexual dimorphism: discordance between molecular and phenotypic data. *Journal of Evolutionary Biology*, **27**, 1443–1453.
- Dimcheff DE, Drovetski SV, Mindell DP (2002) Phylogeny of *Tetraoninae* and other galliform birds using mitochondrial 12S and ND2 genes. *Molecular Phylogenetics and Evolution*, **24**, 203–215.
- Doabin A, Davis CA, Schlesinger F *et al.* (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, **29**, 15–21.
- Dorus S, Evans PD, Wyckoff GJ, Choi SS, Lahn BT (2004) Rate of molecular evolution of the seminal protein gene SEMG2 correlates with levels of female promiscuity. *Nature Genetics*, **36**, 1326–1329.
- Eden E, Lipson D, Yogev S, Yakhini Z (2007) Discovering motifs in ranked lists of DNA sequences. *PLoS Computational Biology*, **3**, 508–522.
- Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z (2009) GORILLA: a tool for discovery and visualization of enriched GO terms in ranked gene lists. *BMC Bioinformatics*, **10**, 48.
- Ellegren H (2013) The evolutionary genomics of birds. *Annual Review of Ecology, Evolution, and Systematics*, **44**, 239–259.
- Ellegren H, Hultin-Rosenberg L, Brunstrom B, Dencker L, Kulitima K, Scholz B (2007) Faced with inequality: chicken do

- not have a general dosage compensation of sex-linked genes. *BMC Biology*, **5**, 40.
- Ellegren H, Smeds L, Burri R *et al.* (2012) The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature*, **491**, 756–760.
- Flicek P, Ahmed I, Amode MR *et al.* (2013) ENSEMBL 2013. *Nucleic Acids Research*, **41**, D48–D55.
- Grabherr MG, Haas BJ, Yassour M *et al.* (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, **29**, 644–652.
- Grath S, Parsch J (2012) Rate of amino acid substitution is influenced by the degree and conservation of male-biased transcription over 50 Myr of *Drosophila* evolution. *Genome Biology and Evolution*, **4**, 346–359.
- Harrison PW, Jordan GE, Montgomery SH (2014) SWAMP: sliding window alignment masker for PAML. *Evolutionary Bioinformatics Online*, **10**, 1–8.
- Harrison PW, Wright AE, Zimmer F *et al.* (in press) Sexual selection drives evolution and rapid turnover of male-biased genes. PNAS.
- Hartl DL, Clark AG (2007) *Principles of Population Genetics*. Sinauer Associates, Inc., Sunderland, Massachusetts.
- Holm S (1979) A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, **6**, 65–70.
- Hope ACA (1968) A simplified Monte Carlo significance test procedure. *Journal of the Royal Statistical Society Series B, Statistical Methodology*, **30**, 582–598.
- Hvilsom C, Qian Y, Bataillon T *et al.* (2012) Extensive X-linked adaptive evolution in central chimpanzees. *Proceedings of the National Academy of Sciences of the USA*, **109**, 2054–2059.
- Itoh Y, Replogle K, Kim Y-H, Wade J, Clayton DF, Arnold AP (2010) Sex bias and dosage compensation in the zebra finch versus chicken genomes: general and specialized patterns among birds. *Genome Research*, **20**, 512–518.
- Khaitovich P, Hellmann I, Enard W *et al.* (2005) Parallel patterns of evolution in the genomes and transcriptomes of humans and chimpanzees. *Science*, **309**, 1850–1854.
- Kimura M, Ohta T (1971) On the rate of molecular evolution. *Journal of Molecular Evolution*, **1**, 1–17.
- Kirkpatrick M, Hall DW (2004a) Male-biased mutation, sex linkage, and the rate of adaptive evolution. *Evolution*, **58**, 437–440.
- Kirkpatrick M, Hall DW (2004b) Sexual selection and sex linkage. *Evolution*, **58**, 683–691.
- Knief U, Schielzeth H, Kempnaers B, Ellegren H, Forstmeier W (2012) QTL and quantitative genetic analysis of beak morphology reveals patterns of standing genetic variation in an Estrildid finch. *Molecular Ecology*, **21**, 3704–3717.
- Koboldt DC, Chen K, Wylie T *et al.* (2009) VARSCAN: variant detection in massively parallel sequencing of individual and pooled samples. *Bioinformatics*, **25**, 2283–2285.
- Koboldt DC, Zhang Q, Larson DE *et al.* (2012) VARSCAN 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Research*, **22**, 568–576.
- Langley CH, Stevens K, Cardeno C *et al.* (2012) Genomic variation in natural populations of *Drosophila melanogaster*. *Genetics*, **192**, 533–598.
- Laporte V, Charlesworth B (2002) Effective population size and population subdivision in demographically structured populations. *Genetics*, **162**, 501–519.
- 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, Handsaker B, Wysoker A *et al.* (2009) The sequence alignment/map format and SAMTOOLS. *Bioinformatics*, **25**, 2078–2079.
- Lohse M, Bolger AM, Nagel A *et al.* (2012) ROBINA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Research*, **40**, W622–W627.
- Mank JE (2009) The W, X, Y and Z of sex chromosome dosage compensation. *Trends in Genetics*, **25**, 226–233.
- Mank JE, Ellegren H (2009) All dosage compensation is local: Gene-by-gene regulation of sex-biased expression on the chicken Z chromosome. *Hereditas*, **102**, 312–320.
- Mank JE, Axelsson E, Ellegren H (2007a) Fast-X on the Z: rapid evolution of sex-linked genes in birds. *Genome Research*, **17**, 618–624.
- Mank JE, Hultin-Rosenberg L, Axelsson E, Ellegren H (2007b) Rapid evolution of female-biased, but not male-biased, genes expressed in the avian brain. *Molecular Biology and Evolution*, **24**, 2698–2706.
- Mank JE, Vicoso B, Berlin S, Charlesworth B (2010a) Effective population size and the Faster-X Effect: empirical results and their interpretation. *Evolution*, **64**, 663–674.
- Mank JE, Nam K, Ellegren H (2010b) Faster-Z evolution is predominantly due to genetic drift. *Molecular Biology and Evolution*, **27**, 661–670.
- Mank JE, Nam K, Brunström B, Ellegren H (2010c) Ontogenetic complexity of sexual dimorphism and sex-specific selection. *Molecular Biology and Evolution*, **27**, 1570–1578.
- McCarthy DJ, Chen Y, Smyth GK (2012) Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research*, **40**, 4288–4297.
- McDonald JH, Kreitman M (1991) Adaptive protein evolution at the ADH locus in *Drosophila*. *Nature*, **351**, 652–654.
- Meisel RP (2011) Towards a more nuanced understanding of the relationship between sex-biased gene expression and rates of protein coding sequence evolution. *Molecular Biology and Evolution*, **28**, 1893–1900.
- Meisel RP, Connallon T (2013) The faster-X effect: integrating theory and data. *Trends in Genetics*, **29**, 537–544.
- Moller AP (1988) Testes size, ejaculate quality and sperm competition in birds. *Biological Journal of the Linnean Society*, **33**, 273–283.
- Moller AP (1991) Sperm competition, sperm depletion, paternal care, and relative testis size in birds. *The American Naturalist*, **137**, 882–906.
- Moller AP, Briskie JV (1995) Extra-pair paternity, sperm competition and the evolution of testis size in birds. *Behavioral Ecology and Sociobiology*, **36**, 357–365.
- Montgomery SH, Capellini I, Venditti C, Barton RA, Mundy NI (2011) Adaptive evolution of four microcephaly genes and the evolution of brain size in anthropoid primates. *Molecular Biology and Evolution*, **28**, 625–638.
- Pagel M (1999) Inferring the historical patterns of biological evolution. *Nature*, **401**, 877–884.
- Pagel M, Meade A, Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. *Systematic Biology*, **53**, 673–684.

- Parsch J, Ellegren H (2013) The evolutionary causes and consequences of sex-biased gene expression. *Nature Reviews. Genetics*, **14**, 83–87.
- Patefield WM (1981) ALGORITHM AS159. An efficient method of generating  $r \times c$  tables with given row and column totals. *Applied Statistics*, **30**, 91–97.
- Perry JC, Harrison PW, Mank JE (2014) The ontogeny and evolution of sex-biased gene expression in *Drosophila melanogaster*. *Molecular Biology and Evolution*, **31**, 1206–1219.
- Petrie M, Krupa A, Burke T (1999) Peacocks lek with relatives even in the absence of social and environmental cues. *Nature*, **401**, 155–157.
- Pitcher TE, Dunn PO, Whittingham LA (2005) Sperm competition and the evolution of testes size in birds. *Journal of Evolutionary Biology*, **18**, 557–567.
- Pointer MA, Harrison PW, Wright AE, Mank JE (2013) Masculinization of gene expression is associated with exaggeration of male sexual dimorphism. *PLoS Genetics*, **9**, 1–9.
- Pool JE, Nielsen R (2007) Population size changes reshape genomic patterns of diversity. *Evolution*, **63**, 3001–3006.
- Quinn EM, Cormican P, Kenny EM *et al.* (2013) Development of strategies for SNP detection in RNA-Seq data: Application to lymphoblastoid cell lines and evaluation using 1000 genomes data. *PLoS One*, **8**, e58815.
- R Core Team (2014) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Ramm SA, Oliver PL, Ponting CP, Stockley P, Emes RD (2008) Sexual selection and the adaptive evolution of Mammalian ejaculate proteins. *Molecular Biology and Evolution*, **25**, 207–219.
- Reeve HK, Pfennig DW (2003) Genetic biases for showy males: are some genetic systems especially conducive to sexual selection? *Proceedings of the National Academy of Sciences of the USA*, **100**, 1089–1094.
- Rice WR (1984) Sex chromosomes and the evolution of sexual dimorphism. *Evolution*, **38**, 735–742.
- Robinson MD, Oshlack A (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology*, **11**, R25.
- Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, **26**, 139–140.
- Ruokonen M, Kvist L, Lumme J (2000) Close relatedness between mitochondrial DNA from seven Anser goose species. *Journal of Evolutionary Biology*, **13**, 532–540.
- Sackton TB, Corbett-Detig RB, Nagaraju J, Vaishna L, Arunkumar KP, Hartl DL (2014) Positive selection drives Faster-Z evolution in silkworms. *Evolution*, **68**, 2331–2342.
- Saether SA, Saetre G-P, Borge T *et al.* (2007) Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. *Science*, **318**, 95–97.
- Saetre GP, Borge T, Lindroos K *et al.* (2003) Sex chromosome evolution and speciation in *Ficedula* flycatchers. *Proceedings of the Royal Society B-Biological Sciences*, **270**, 53–59.
- Schielzeth H, Kempnaers B, Ellegren H, Forstmeier W (2012) QTL linkage mapping of zebra finch beak color shows an oligogenic control of a sexually selected trait. *Evolution*, **66**, 18–30.
- Singh N, Larracuente A, Clark A (2008) Constrasting the efficacy of selection on the X and autosomes in *Drosophila*. *Molecular Biology and Evolution*, **25**, 454–467.
- Skinner BM, Robertson LBW, Tempest HG *et al.* (2009) Comparative genomics in chicken and Pekin duck using FISH mapping and microarray analysis. *BMC Genomics*, **10**, 357.
- Stiglec R, Ezaz T, Graves JAM (2007) A new look at the evolution of avian sex chromosomes. *Cytogenetic and Genome Research*, **117**, 103–109.
- Thornton K, Long M (2005) Excess of amino acid substitutions relative to polymorphism between X-linked duplications in *Drosophila melanogaster*. *Molecular Biology and Evolution*, **22**, 273–284.
- Toups MA, Pease JB, Hahn MW (2011) No excess gene movement is detected off the avian or lepidopteran Z chromosome. *Genome Biology and Evolution*, **3**, 1381–1390.
- van Tuinen M, Dyke GJ (2004) Calibration of galliform molecular clocks using multiple fossils and genetic partitions. *Molecular Phylogenetics and Evolution*, **30**, 74–86.
- van Tuinen M, Hedges SB (2001) Calibration of avian molecular clocks. *Molecular Biology and Evolution*, **18**, 206–213.
- Uebbing S, Künstner A, Mäkinen H, Ellegren H (2013) Transcriptome sequencing reveals the character of incomplete dosage compensation across multiple tissues in flycatchers. *Genome Biology and Evolution*, **5**, 1555–1566.
- Vicoso B, Charlesworth B (2009) Effective population size and the Faster-X effect: an extended model. *Evolution*, **63**, 2413–2426.
- Vicoso B, Emerson JJ, Zektser Y, Mahajan S, Bachtrog D (2013a) Comparative sex chromosome genomics in snakes: differentiation, evolutionary strata, and lack of global dosage compensation. *PLoS Biology*, **11**, e1001643.
- Vicoso B, Kaiser VB, Bachtrog D (2013b) Sex-biased gene expression at homomorphic sex chromosomes in emus and its implication for sex chromosome evolution. *Proceedings of the National Academy of Sciences of the USA*, **110**, 6453–6458.
- Wang B, Ekblom R, Bunikis I, Siitari H, Høglund J (2014a) Whole genome sequencing of the black grouse (*Tetrao tetrix*): reference guided assembly suggests faster-Z and MHC evolution. *BMC Genomics*, **15**, 180.
- Wang Z, Zhang J, Yang W *et al.* (2014b) Temporal genomic evolution of bird sex chromosomes. *BMC Evolutionary Biology*, **14**, 250.
- Watterson GA (1975) On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, **7**, 256–276.
- Wright AE, Mank JE (2013) The scope and strength of sex-specific selection in genome evolution. *Journal of Evolutionary Biology*, **26**, 1841–1853.
- Xu K, Oh S, Park T, Presgraves DC, Yi SV (2012) Lineage-specific variation in slow- and fast-X evolution in primates. *Evolution*, **66**, 1751–1761.
- Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, **24**, 1586–1591.

---

A.E.W., M.A.P. and J.E.M. were involved in data collection. A.E.W. and J.E.M. designed the research. A.E.W., P.W.H., F.Z., S.H.M. and J.E.M. analysed the data and wrote the paper.

---

## Data accessibility

Illumina raw reads: SRA: PRJNA271731.

Trinity assembly, RPKM data, sequence alignments and SNP data: Dryad: doi:10.5061/dryad.4gv50.

## Supporting information

Additional supporting information may be found in the online version of this article.

**Table S1** Assembly statistics for each Galloanserae species.

**Table S2** Assembly statistics for each sample.

**Table S3** Number of nonsynonymous and synonymous polymorphic and fixed sites at a minor allele frequency threshold of 0.15.

**Table S4** Number of nonsynonymous and synonymous polymorphic and fixed sites at a minor allele frequency threshold of 0.25.

**Table S5** Phylogenetically controlled regression analyses between Faster-Z and log sperm number, and residual testes weight.

**Table S6** Nucleotide diversity estimates of the Z chromosome and autosomes.

**Table S7.** Effective population size estimates of the Z chromosome and autosomes.

**Table S8** Effective population size estimates of the Z chromosome and autosomes calculated using Watterson's estimation of theta.

**Table S9** Phylogenetically controlled regression analyses between different measures of NEZ/NEA and residual testes weight, log sperm number, and Faster-Z.

**Table S10** Site model test results for contigs with signatures of positive selection.

**Table S11** pN, Ps and pN/pS for Z-linked and autosomal genes across Galloanserae species with a minor allele frequency of 0.25.

**Table S12** McDonald Kreitman test results.

**Table S13** Significant differences between nonsynonymous and synonymous polymorphism on the Z chromosome and autosomes with a minor allele frequency threshold of 0.25.

**Table S14** Faster-Z Effect across expression classes.

**Table S15** Differences in Faster-Z Effect between expression classes.