J Mol Evol (2006) 62:226–233
DOI: 10.1007/s00239-005-0103-6

JOURNAL OF MOLECULAR EVOLUTION

Substitution Rate Heterogeneity and the Male Mutation Bias

Sofia Berlin,[1,*] Mikael Brandström,[1] Niclas Backström,[1] Erik Axelsson,[1] Nick G.C. Smith,[1,2†] Hans Ellegren[1]

[1] Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18 D, SE-752 36 Uppsala, Sweden
[2] Department of Mathematics and Statistics, Lancaster University, Lancaster LA1 4YF, UK

Received: 28 April 2005 / Accepted: 18 October 2005 [*Reviewing Editor:* Dr. Martin Kreitman]

Abstract. Germline mutation rates have been found to be higher in males than in females in many organisms, a likely consequence of cell division being more frequent in spermatogenesis than in oogenesis. If the majority of mutations are due to DNA replication error, the male-to-female mutation rate ratio (α_m) is expected to be similar to the ratio of the number of germ line cell divisions in males and females (c), an assumption that can be tested with proper estimates of α_m and c. α_m is usually estimated by comparing substitution rates in putatively neutral sequences on the sex chromosomes. However, substantial regional variation in substitution rates across chromosomes may bias estimates of α_m based on the substitution rates of short sequences. To investigate regional substitution rate variation, we estimated sequence divergence in 16 gametologous introns located on the Z and W chromosomes of five bird species of the order Galliformes. Intron ends and potentially conserved blocks were excluded to reduce the effect of using sequences subject to negative selection. We found significant substitution rate variation within Z chromosome ($G_{15} = 37.6$, $p = 0.0010$) as well as within W chromosome introns ($G_{15} = 44.0$, $p = 0.0001$). This heterogeneity also affected the estimates of α_m, which varied significantly, from 1.53 to 3.51, among the introns (ANOVA: $F_{13,14} = 2.68$, $p = 0.04$). Our results suggest the importance of using extensive data sets from several genomic regions to avoid the effects of regional mutation rate variation and to ensure accurate estimates of α_m.

Key words: Male mutation bias — Sex chromosomes — Birds — Substitution rate variation

Introduction

Studies of a variety of organisms have shown that the germline mutation rate is male-biased, i.e., that more mutations are produced in males than in females (mammals [Shimmin et al. 1993a; Chang et al. 1994; Bohossian et al. 2000; Makova and Li 2002], birds [Ellegren and Fridolfsson 1997; Kahn and Quinn 1999; Bartosch-Härlid et al. 2003; Axelsson et al. 2004], fish [Ellegren and Fridolfsson 2003], and plants [Filatov and Charlesworth 2002; Whittle and Johnston 2002]). The male mutation bias was first suggested by Haldane (1935) under the assumption that if mutations are predominantly introduced by errors during cell division (DNA replication), then the germline mutation rate should correlate positively with the number of germline cell divisions, which in many animal species are in vast excess in spermatogenesis compared to oogenesis. It has subsequently been recognized that comparisons of estimates of the male mutation bias and the sex bias in germline cell division actually provide a means for testing the relative importance of faulty replication in germline mutation (Chang et al. 1994; Chang and Li 1995; Kahn and Quinn 1999; Li et al. 2002). Specifically, if the male-to-female mutation rate ratio (α_m) is significantly different from the ratio of the numbers of cell division in spermatogenesis and oogenesis (c), then other mechanisms than replication errors are

†Deceased
Current address: Department of Genetics and Genomics, Roslin Institute (Edinburgh), Roslin, Midlothian EH25 9PS, UK
Correspondence to: Hans Ellegren; *email:* hans.ellegren@ebc.uu.se

important for causing mutations too. Alternatively, this could indicate that the per cell generation mutation rate differs between spermatogenesis and oogenesis. Analyses of these basic aspects of germline biology require appropriate estimates of α_m and c.

The extent of male-biased mutation is also important in other contexts. Molecular clock rates may vary in response to life history characteristics such as generation time and sexual selection if these characteristics affect the number of germ cell divisions. For instance, when sexual selection is intense and leads to increased sperm production, the relative excess of male mutation is expected to increase (Bartosch-Härlid et al. 2003). Moreover, the relative rate of adaptive evolution in different parts of the genome is affected by α_m (Kirkpatrick and Hall 2004), as is the chromosome-specific levels of nucleotide diversity (Sachidanandam et al. 2001). Furthermore, a significant sex bias in mutation rates will have practical implications for paternity testing (Ibarguchi et al. 2004) and is also important in health risk assessment (Crow 1997).

Mutation rates prove difficult to estimate by direct counts from one generation to another, however, Miyata et al. (1987) offered an evolutionary approach for studying α_m. By comparing neutral substitution rates in chromosomes that spend different amounts of time in male and female germline, the relative mutation rates of the two sexes can be indirectly inferred. For instance, since the mammalian Y chromosome is only transmitted through the male germline, the rate of Y sequence evolution is solely governed by the male mutation rate. Knowing this rate from divergence data of neutral Y chromosome sequences, the female mutation rate can be obtained from divergence data of neutral X chromosomes sequences by taking into account the fact that, during evolution, X chromosomes are in the female germline two-thirds of the time and in the male germline one-third of the time. Correspondingly, α_m can be estimated in organisms with female heterogamety, such as birds, where the W chromosome evolves under the influence of female mutation rate only, while the Z chromosome is in the male germline two-thirds of the time.

A potential pitfall in α_m estimation using the evolutionary approach derives from the fact that substitution rates in regions traditionally thought to be free of selective constraint are not uniform along chromosomes (Lercher et al. 2001). Such variation may be due either to mutation rate heterogeneity or to negative selection on, e.g., noncoding DNA or synonymous sites. Indeed, recent observations from different organisms point at significant substitution rate heterogeneity among genomic regions (Lercher et al. 2001; Ebersberger et al. 2002; Smith et al. 2002; Waterston et al. 2002; Hardison et al. 2003), a variation that may occur at various scales, from sequence context effects of neighboring nucleotides up to regional variation at the Mb-scale level (Ellegren et al. 2003). Potentially, the effect of such variation on α_m estimation could be significant. For example, consider estimates of α_m in the human-chimpanzee comparison. Typical sequence divergence for the X and Y chromosomes in this case are 0.010 and 0.019, respectively (Ebersberger et al. 2002), which gives $\alpha_m = 3.45$ according to the formula of Miyata et al. (1987). However, if data from a region of unusually high substitution rate on the Y chromosome (say 1.5 times the average rate) are used for α_m estimation, then $\alpha_m = 38$.

In practice, we do not know to what extent this represents a real problem in molecular evolutionary analyses of male-biased mutation. A common approach in studies of this kind is to target one or a few regions shared (yet nonrecombining, i.e., "gametologous") between the X and the Y chromosomes, or the Z and the W chromosomes, and estimate α_m based on substitution rate estimates from these regions. Hence, it is considered to be an advantage to use gametologous genes to correct for direct sequence effects. However, the use of gametologues would not correct for regional effects on mutation rates. To empirically test whether substitution rate heterogeneity along the sex chromosomes affects α_m estimates, we study here the male mutation bias in five birds using sequence data from 16 different Z- and W-linked gametologous introns.

Materials and Methods

Sequence Data

Sixteen introns from three gametologous gene pairs (*CHD1Z*/*CHD1W*, *SPINZ*/*SPINW*, and *UBAP2Z*/*UBAP2W*) on the avian Z and W sex chromosomes were sequenced in five species from the order Galliformes: chicken (*Gallus gallus*), red-legged partridge (*Alectoris rufa*), quail (*Coturnix coturnix*), turkey (*Meleagris gallopavo*), and black grouse (*Tetrao tetrix*) (Table 1). PCR amplification was achieved by the use of conserved Z-specific and W-specific exon primers (Supplementary Material). PCR reactions for all introns except *UBAP2* introns 1 and 4 were performed in 20-μl volumes in a Perkin Elmer 9600 Thermal Cycler using 0.5 U AmpliTaq Gold (Applied Biosystems), 2.5–3 m*M* $MgCl_2$, 0.08 m*M* dNTPs, 1 × PCR Gold Buffer, 5 pmol of each primer, and 50 ng of template DNA. *UBAP2* introns 1 and 4 were amplified with *Pfu* DNA polymerase (Promega) according to the manufacturer's instructions. The error rate of *Pfu* DNA polymerase is approximately 1.6×10^{-6} errors per base (Lundberg et al. 1991), while the error rate of Taq DNA polymerase is approximately 1.1×10^{-4} (e.g., Barnes 1992). We assume that this difference has had no effect on our study. Prior to sequencing, PCR products were purified with ExoSAP-IT (Amersham Biosciences), followed by direct sequencing in forward and reverse directions using the original PCR primers and the DYEnamic ET DyeTerminator Kit (Amersham Biosciences). PCR products from *UBAP2* introns, *CHD1* introns 12 and 22, and *SPINZ* intron 3 were ligated into pGEM-T Easy Vectors, transformed into *JM109* competent cells (Promega), and thereafter sequenced using modified M13 primers OMNI

Table 1. Data on intron length, sequence divergence (K_Z and K_W), and male-to-female mutation rate ratio (α_m) based on untrimmed and trimmed data sets, respectively

	Untrimmed data set				Trimmed data set			
Intron	Length (Z/W ungapped)	$K_Z \pm$ SE	$K_W \pm$ SE	$\alpha_m \pm$ SE	Length (Z/W ungapped)	$K_Z \pm$ SE	$K_W \pm$ SE	$\alpha_m \pm$ SE
UBAP2-1	1118/1048	0.208 ± 0.015	0.126 ± 0.011	1.97 ± 0.29	886/840	0.234 ± 0.018	0.132 ± 0.013	2.16 ± 0.36
UBAP2-2	374/764	0.258 ± 0.033	0.106 ± 0.012	3.16 ± 0.62	332/733	0.287 ± 0.036	0.107 ± 0.012	3.51 ± 0.70
UBAP2-4	1370/862[a]	0.244 ± 0.015	0.133 ± 0.013	2.25 ± 0.32	1302/796[a]	0.238 ± 0.016	0.138 ± 0.014	2.09 ± 0.31
CHD1-6	580/619[a]	0.266 ± 0.023	0.195 ± 0.021	1.55 ± 0.29	544/583[a]	0.265 ± 0.024	0.186 ± 0.021	1.64 ± 0.32
CHD1-7	221/216	0.302 ± 0.041	0.214 ± 0.034	1.62 ± 0.48	182/180	0.324 ± 0.049	0.198 ± 0.038	1.96 ± 0.60
CHD1-10	360/458	0.278 ± 0.032	0.193 ± 0.023	1.66 ± 0.36	321/436	0.297 ± 0.035	0.201 ± 0.023	1.72 ± 0.37
CHD1-11	606/338	0.321 ± 0.028	0.187 ± 0.026	2.07 ± 0.45	570/306	0.324 ± 0.030	0.166 ± 0.026	2.44 ± 0.56
CHD1-12	567[a]/882[a]	0.299 ± 0.026	0.181 ± 0.015	1.97 ± 0.31	535[a]/857[a]	0.298 ± 0.026	0.183 ± 0.016	1.95 ± 0.21
CHD1-13	226/163	0.264 ± 0.040	0.183 ± 0.047	1.67 ± 0.63	192/120	0.265 ± 0.040	0.179 ± 0.043	1.72 ± 0.77
CHD1-15	402/238	0.390 ± 0.037	0.190 ± 0.029	2.57 ± 0.56	364/200	0.384 ± 0.039	0.200 ± 0.033	2.38 ± 0.57
CHD1-17	489/607	0.288 ± 0.027	0.165 ± 0.019	2.11 ± 0.41	451/576	0.293 ± 0.029	0.159 ± 0.018	2.26 ± 0.43
CHD1-21	119/127	0.251 ± 0.045	0.199 ± 0.094	1.39 ± 0.59	86/89	0.237 ± 0.054	0.175 ± 0.081	1.53 ± 0.95
CHD1-22	1000/492[a]	0.297 ± 0.019	0.117 ± 0.016	3.29 ± 0.60	980/454[a]	0.298 ± 0.020	0.115 ± 0.017	3.38 ± 0.64
CHD1-24	347/353	0.264 ± 0.033	0.139 ± 0.020	2.35 ± 0.59	310/314	0.273 ± 0.036	0.129 ± 0.021	2.67 ± 0.69
CHD1-25	621/490	0.286 ± 0.026	0.130 ± 0.019	2.80 ± 0.59	588/462	0.295 ± 0.026	0.135 ± 0.019	2.77 ± 0.53
SPIN-3	669/544	0.269 ± 0.024	0.132 ± 0.016	2.56 ± 0.46	639/515	0.272 ± 0.027	0.129 ± 0.017	2.66 ± 0.55

[a] The complete intron was not sequenced.

(5′ ACAGGAAACAGCTATGACCATGAT) and UNI (5′ CGAC GTTGTAAAACGAGGCCAGT), as well as with internal sequencing primers to cover the full length of the introns. *Pfu* DNA polymerase generates blunt-end fragments during PCR amplification and these products were subsequently A-tailed, ligated into pGEM-T Easy Vector, and transformed into *JM109* competent cells (Promega). Positive clones were amplified with Templiphi DNA sequencing template amplification kit (Amersham Biosciences) and sequenced in both directions using the OMNI/UNI primers. Reactions were electrophoresed on a MegaBACE 1000 sequencing instrument (Amersham Biosciences). All sequences have been deposited in GenBank (accession numbers in Supplementary Material).

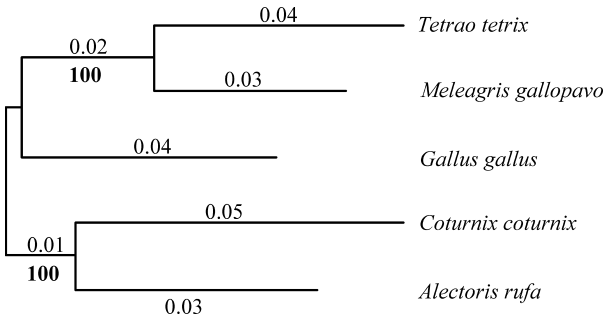

Fig. 1. Neighbor-joining tree with branch lengths (Tamura-Nei) and bootstrap supports in boldface.

Sequence Analyses

Sequences were edited in Autoassembler (Applied Biosystems) and forward and reverse sequences compared to construct consensus sequences. Regions of simple repeats, which may be difficult to align, were identified with Sputnik (Abajian 1994) and masked before further analyses. Z- and W-linked sequences from each intron were aligned separately, using Clustal W (Thompson et al. 1994) in Sequence Navigator 1.0 (Applied Biosystems). Phylogenetic reconstruction was performed with a concatenation of all Z- and all W-linked alignments using the neighbor joining method and Tamura-Nei distance method implemented in MEGA 2.1 (Kumar et al. 2001) (Fig. 1). The bootstrap analysis consisted of 1000 replicates. Individual trees constructed for each intron did not give any indication of gene conversion between gametologous sequences (data not shown). Pairwise comparisons between all five species for K_Z and K_W were estimated using the Tamura-Nei distance method in MEGA 2.1 (Supplementary Material).

Sequence divergence (K_Z and K_W) for each intron was estimated using the baseml program in PAML 3.13d (Yang 1997), with the Tamura-Nei (1993) model of sequence evolution. The estimation of intronic standard errors and hypothesis testing was carried out using nonparametric bootstrapping, which involved random sampling with replacement of sites from an alignment to generate alignments of the same length as the original. The bootstrapping process was repeated 1000 times for each alignment, thereby giving 1000 sets of K_Z and K_W estimates. Following the method of Miyata et al. (1987), the male-to-female mutation rate ratio, α_m, was estimated using K_Z and K_W. The expected ratio of K_Z to K_W, taking into account the time the chromosomes spend in male and female germlines, is $(1/3 + 2/3\ \alpha_m)$:1, which gives $\alpha_m = (3 \times K_Z/K_W - 1)/2$. Standard errors for K_Z, K_W, and hence α_m were given by the standard deviation of the 1000 bootstrap replicates. The K_Z/K_W ratio can be obtained by matching any divergence estimates, including pairwise divergence between two species, divergence down one lineage, or total divergence down all lineages in a tree.

Statistical Analyses

In order to study variation in substitution rates among introns we used estimates of divergence down all lineages, i.e., the total branch lengths for each intron and chromosomal class. *G* tests (Sokal and Rohlf 2000) were performed, comparing the observed numbers of substitutions ($O(D)$) with the expected ($E(D)$), given by the following formula:

Table 2. Relative rate tests for K_Z, K_W, and α_m, assuming a star phylogeny

Pairwise comparison	K_Z	p value	K_W	p value	α_m	p value
G .gallus–M. gallopavo	0.051–0.037	$p < 0.001$	0.022–0.020	n.s.	2.95–2.30	n.s.
G. gallus–T. tetrix	0.051–0.046	n.s.	0.022–0.026	n.s.	2.95–2.13	$p = 0.015$
G. gallus–A. rufa	0.051–0.042	$p = 0.010$	0.022–0.028	$p = 0.014$	2.95–1.79	$p = 0.002$
G. gallus–C. coturnix	0.051–0.064	$p = 0.001$	0.022–0.030	$p = 0.004$	2.95–2.72	n.s
M. gallopavo–T. tetrix	0.037–0.046	$p = 0.015$	0.020–0.026	$p = 0.014$	2.30–2.13	n.s.
M. gallopavo–A. rufa	0.037–0.042	n.s.	0.020–0.028	$p = 0.002$	2.30–1.79	n.s.
M. gallopavo–C. coturnix	0.037–0.064	$p < 0.001$	0.020–0.030	$p = <0.001$	2.30–2.72	n.s
T. tetrix–A. rufa	0.046–0.042	n.s.	0.026–0.028	n.s.	2.13–1.79	n.s.
T. tetrix–C. coturnix	0.046–0.064	$p < 0.001$	0.026–0.030	n.s.	2.13–2.72	n.s.
A. rufa–C. coturnix	0.042–0.064	$p < 0.001$	0.028–0.030	n.s.	1.79–2.72	$p = 0.003$

Table 3. Relative rate tests for K_Z, K_W, and α_m using the whole phylogeny (i.e., not assuming a star phylogeny)

Pairwise comparison	K_Z	p value	K_W	p value	α_m	p value
G. gallus–M. gallopavo	0.051–0.062	$p = 0.002$	0.022–0.033	$p < 0.001$	2.95–2.29	n.s.
G. gallus–T. tetrix	0.051–0.070	$p < 0.001$	0.022–0.039	$p < 0.001$	2.95–2.18	$p = 0.016$
G. gallus–A. rufa	0.051–0.057	n.s.	0.022–0.038	$p < 0.001$	2.95–1.73	$p < 0.001$
G. gallus–C. coturnix	0.051–0.079	$p < 0.001$	0.022–0.041	$p < 0.001$	2.95–2.42	n.s.
M. gallopavo–T. tetrix	0.037–0.046	$p = 0.015$	0.020–0.026	$p = 0.014$	2.30–2.13	n.s.
M. gallopavo–A. rufa	0.062–0.057	n.s.	0.033–0.038	p = 0.05	2.29–1.73	n.s.
M. gallopavo–C. coturnix	0.062–0.079	$p < 0.001$	0.033–0.041	$p = 0.014$	2.29–2.42	n.s.
T. tetrix–A. rufa	0.070–0.057	$p = 0.003$	0.039–0.038	n.s.	2.18–1.73	n.s.
T. tetrix–C. coturnix	0.070–0.079	$p = 0.028$	0.039–0.041	n.s.	2.18–2.42	n.s.
A. rufa–C. coturnix	0.042–0.064	$p < 0.001$	0.028–0.030	n.s.	1.79–2.72	$p = 0.003$

$$E(D)_j = \frac{\sum_j O(D)_j}{\sum_j L_j} L_j \quad (1)$$

where L is the ungapped length of each intron alignment, $O(D)$ is given by the product of L and the intronic divergence, and summation is over all introns. Variation in estimates of the male mutation bias among gametologous introns was studied with a one-way analysis of variance (ANOVA) using the logarithm of α_m for each intron to ensure normality. The error variance was estimated by dividing each intron alignment in odd-and-even numbered nucleotides, thereby giving two estimates of total branch lengths for each intron (Smith and Eyre-Walker 2003). Variations in substitution rate and in the estimates of the male mutation bias with introns and lineages as independent factors were studied in the same way with two-way ANOVAs. The two shortest introns, *CHD1-13* and *CHD1-21*, displayed excessively high error variances, so we omitted these from the analyses of variance to attain homogeneity of variances among the introns (Sokal and Rohlf 2000). Moreover, *CHD1-15*, *CHD1-24*, and *CHD1-25* showed zero branch lengths and were excluded from the two-way ANOVAs.

Search for Conserved Elements

We made alignments of Z- and W-linked gametologues from chicken alone to search for conserved blocks, and performed two types of analyses. First, alignments were examined to see if the distribution of lengths of perfectly conserved blocks deviated from the distribution that would be expected by chance. To obtain the expected distribution of block lengths, each alignment was permuted 1000 times and the length distribution of conserved blocks was noted for each permutation. We compared the observed distribution of conserved blocks with the expected distributions with a Kolmogorv-Smirnov test. Second, we used Pipmaker (Schwartz et al. 2000), which is based on the local alignment method BLASTZ and outputs a list of conserved regions and their respective conservation level. Each intron alignment was search for conserved blocks with a sequence similarity of at least 90%.

In order to examine sequence conservation in intron edges, we estimated sequence divergence in nonoverlapping sequence blocks at various distances from the intron-exon boundaries as well as in the bulk middle part of the introns (excluding 30 base pairs in each end). The blocks were positioned at nucleotides 1–10, 11–20, and 21–30 from the splice sites at both the ends (running $5' \rightarrow 3'$) and the $3'$ ends (running $3' \rightarrow 5'$), respectively. Blocks comprising nucleotides 1–20 and 1–30 were also analyzed. Sequence divergence within the end blocks, as well as the middle block of the introns was compared using a chi-square test of a 7×2 contingency table. The contingency table was constructed by assigning the first row as the estimated number of nucleotide changes for each block using the baseml program of PAML, and the second row as the number of bases in each block.

Relative Rate Tests

In order to investigate whether different lineages have evolved at different rates, relative rate tests were performed for all possible species pairs and chromosomal classes using concatenated alignments of all introns using trimmed data sets. Substitutions in terminal branches from the split of any two lineages were compared, and the bootstrapped branch lengths were used for evaluating the significance of the relative rate tests (Table 2 and 3). A similar approach was used to study variation in α_m among lineages. In this

case α_m was estimated using terminal branch lengths from the concatenated alignments of Z and W sequences, respectively. In addition, substitution rates and α_m were estimated using the whole phylogeny (Table 3) followed by relative rate tests.

Results

Five galliform species of birds (chicken, red-legged partridge, quail, turkey, and black grouse) were sequenced for 16 introns of the Z- and the W-linked copy of three gametolgous genes (Table 1). A total of about 9000 bp Z-linked and 8200 bp W-linked sequence was thereby obtained for each species. We examined variation in substitution rates (K_Z and K_W; total branch lengths of the respective phylogenetic tree) and subsequently also estimates of the male mutation bias (α_m) for each intron (untrimmed data set). K_Z varied between 0.21 and 0.39, with statistically significant variation among introns ($G_{15} = 51.4, p < 0.0001$). K_W varied between 0.11 and 0.21 and the variation among introns was significant also in this case ($G_{15} = 56.1$, $p < 0.0001$). Somewhat surprisingly, K_w estimates fell broadly within two bimodal ranges; *CHD1-6, 7, 10, 11, 12, 13, 15, 17*, and *21* show divergences of 0.165–0.214, while *CHD1-22, 24*, and *25*, *SPIN-3*, and *UBAP-1, 2*, and *4* show divergences of 0.106–0.139, a pattern that is unaccounted for. Estimates of α_m based on each gametologous intron ranged between 1.39 and 3.29, with a mean of 2.19. There was evidence for significant variation in $\log\alpha_m$ among the introns (ANOVA: $F_{13, 14} = 3.56, p = 0.01$).

There was no significant correlation between substitution rates of gametologous introns ($\rho = 0.318$, $p = 0.228$, Spearman rank test). Moreover, the variance in α_m estimates obtained from the 16 gametologous introns (var = 0.35) was not significantly lower than the variance in a permuted data set (95% confidence limit = 0.29, single sided) where introns from Z and W were randomly selected to estimate α_m ($p = 0.13$)

A detectable fraction of the noncoding DNA of vertebrate genomes, including chicken (Chicken Genome Consortium 2004), consists of conserved blocks that most likely evolve under purifying selection (Ludwig 2002). The presence of sequences subject to selective constraint in our data set may potentially contribute to the heterogeneity in substitution rates seen among introns. Comparative genomics offers a means for identification of conserved blocks and here we can use data from gametologous Z-linked and W-linked sequences for this purpose. The avian Z and W sex chromosomes started to differentiate from an autosomal state 100–170 Mya (Handley et al. 2004). Given this long time of independent evolution, any remaining sign of conservation between paralogous Z-linked and W-linked noncoding DNA would suggest functional constraints. However, the incidence of perfectly conserved sequence blocks in alignments of the chicken Z- and W-copy of each intron was not higher than in a permuted data set (Kolmogorov-Smirnov test, $p > 0.2$).

It has recently been observed in mammalian genomes (Hare and Palumbi 2003; Chamary and Hurst 2004) and in the chicken genome (Chicken Genome Consortium 2004) that sequence conservation in intron ends may extend beyond the splice acceptor and donor sites. However, in our untrimmed data set of the respective alignments of Z- and W-linked sequences from all species, we were not able to detect lower divergence in blocks 10, 20, and 30 bp from exon-intron boundaries compared to the bulk interior parts of introns (χ^2 test; $p > 0.05$). Although the analyses described above do not present evidence for a large effect of negative selection on intron sequences, we cannot exclude that some sequences have been evolving under such constraint. To be conservative in the analysis of mutation rate variation among introns, we therefore trimmed the data set by removing 20 bp in the 5′ and 3′ end of each intron and conserved blocks with a similarity of at least 90% according to Pipmaker analysis. The block lengths varied between 3 and 57 nucleotides; a total of 576 bp was in this way removed from the complete data set (17,270 bp). However, using the trimmed data set, the earlier conclusions basically remained unchanged. There was significant variation in K_Z ($G_{15} = 37.6$, $p = 0.0010$) and in K_W ($G_{15} = 44.0$, $p = 0.0001$) as well as in $\log\alpha_m$ among introns (ANOVA: $F_{13, 14} = 2.61$, $p = 0.04$) (Table 1).

Using the trimmed data set (similar results were obtained with the untrimmed data set), two-way ANOVA analyses showed significant intron and lineage effects for substitution rates among Z-linked (lineage: $F_{4,86} = 11.0, p < 0.001$; intron: $F_{9,86} = 1.98$, $p = 0.05$) as well as W-linked introns (lineage: $F_{4,86} = 2.90$, $p = 0.03$; intron: $F_{9,86} = 4.26$, $p = 0.0001$). Moreover, there were significant intron and lineage effects on estimates of the male mutation bias (lineage: $F_{4,85} = 2.76, p = 0.03$; intron: $F_{9,85} = 2.80$, $p = 0.0006$). The phylogenetic relationship among the species included in this study approximately follows a star-like phylogeny (Dimcheff et al. 2002), although our tree (Fig. 1) gives significant support for positive internal branches. Relative rate tests of divergence estimates were therefore conducted both between all pairs of lineages in the phylogeny (Table 2) and using the phylogeny shown in Fig. 1 (Table 3). In both cases, quail displayed the highest substitution rates among the five lineages for Z as well as for W-linked sequences using the trimmed data set for estimating divergences. This may be consistent with a generation time effect on the molecular clock, quail being the smallest species of the five galliform birds analyzed. Also in both ap-

proaches, the male mutation bias was significantly higher in chicken (≈3.0) than in black grouse (≈2.1) and red-legged partridge (≈1.7) and significantly higher in quail (≈2.5) than in red-legged partridge (Tables 2 and 3).

Discussion

Previous studies of male-biased mutation have revealed variable and sometimes inconsistent estimates of α_m for the same group of organisms. In humans, for instance, estimates have ranged between 1.7 and 6 (Shimmin et al. 1993a; Bohossian et al. 2000; Makova and Li 2002). In birds, point estimates of α_m of between 1.7 and 6.5 have been reported (Ellegren and Fridolfsson 1997; Kahn and Quinn 1999; Carmichael et al. 2000; Fridolfsson and Ellegren 2000; Bartosch-Harlid 2003; Axelsson et al. 2004). There are at least three possible methodological causes to such variation. One obvious explanation is that most estimates are associated with rather large confidence intervals (note that α_m is obtained from the ratio of two divergence estimates) so the wide range of the point estimates is at least in part due to stochastic variation in substitution rate estimation. It follows that, for comparative purposes, meaningful estimates of α_m need to be based on divergences estimated with low variance.

Second, ancestral polymorphism affects the analysis of male-biased mutation if the evolutionary distance between the taxa for which divergence is estimated is short (Makova and Li 2002). Specifically, if the preexisting levels of polymorphism differ between the chromosomal classes being analyzed and these levels are appreciable compared to levels of divergence, then α_m estimates will be biased since the coalescence time (and thus divergence) for two sequences from one chromosomal class will tend to differ from that of two sequences from another class. This was noted in studies based on Y chromosome data in human and great apes (Makova and Li 2002) and W chromosome data in terminal branches of an extensive bird phylogeny (Bartosch-Härlid et al. 2003); both Y and W are typically low in genetic diversity (Sachidanandam et al. 2001; Berlin and Ellegren 2004). In the present study we consider the effect of ancestral polymorphism negligible since the pairwise distance between all taxa is of the order of 5–10%, which is high compared to diversity levels.

Third, data presented herein demonstrate that heterogeneity in substitution rates among genomic regions on the sex chromosomes can lead to significant variation in α_m estimates. It has previously been noted that the male mutation bias is insufficient to explain the variation in substitution rates seen among and within individual chromosomes (McVean and Hurst 1997; Lercher et al. 2001). We extend these observations by showing that even with the analysis of gametologous sequences shared between sex chromosomes, an approach commonly used in α_m estimation, significant variation in α_m estimates is obtained. With the exception of large-scale analysis of genomic sequences from human-chimpanzee (Ebersberger et al. 2002; Lu and Wu 2005) and mouse-rat (Lercher et al. 2001; Malcom et al. 2003; Makova et al. 2004), most studies of male bias have been based on data from one or just a few gametologous regions shared between the sex chromosomes (e.g., Shimmin et al. 1993a; Chang et al. 1994; Agulnik et al. 1997). It is thus possible that part of the variation in α_m seen among these studies reflects mutation rate heterogeneity inherent to different genomic regions rather than underlying differences in the mutational sex bias.

Given the observed heterogeneity in estimates of the male mutation bias, is there a single meaningful evolutionary-based α_m for each species (lineage)? On the one hand, large-scale approaches involving sequence data from many different genomic regions should reveal a genomic average for the male mutation bias. On the other hand, however, just as sex-specific recombination rates vary over the genome (Nachman 2002), so may sex-specific mutation rates. Clearly, this is true for hot spot sites where differences in methylation levels between sexes can lead to distinct differences in their rates of mutation (Wilkin et al. 1998; Girard et al. 2001; Trappe et al. 2001). For comparative purposes, e.g., in tests of the correlation between the male mutation bias and various life history parameters, this suggests that either of two approaches should be taken. If it is possible to retrieve large numbers of sequence data (for instance, if the genome sequence is available for one of the species), it would be preferable to include data from many genomic regions rather than doing more extensive sampling from one or just a few regions. Alternatively, for less well-characterized genomes, comparative studies should be based on data from the same genomic region(s) analyzed in all species (cf. Bartosch-Härlid et al. 2003).

One important observation in this study was that gametologous introns did not seem to perform better in α_m estimation than random combinations of introns from Z and W. Furthermore, the absence of a correlation between the substitution rates of gametologous introns suggests that the pattern of mutation rate heterogeneity is not conserved in these paralogous sequences. Deterministic mutation rate variation has been observed over much shorter evolutionary distances, e.g., for orthologous sequences along the human and chimpanzee lineages (Smith et al. 2002). Whatever factors govern repeatability in patterns of regional mutation rate variation, avian

sex chromosome evolution may represent too long a time of divergence (>100 MYA) for such effects to remain or, at least, to be strong enough to be detectable in our analysis. Based on the assumption that gametologous noncoding sequences are subject to similar patterns of mutation, it has been argued that they are preferably used for α_m estimation (Shimmin et al. 1993, b; Chang et al. 1994; Chang and Li 1995; Ellegren and Fridolfsson 1997; Kahn and Quinn 1999; Fridolfsson and Ellegren 2000). However, our analysis thus does not support this assumption, something, which may make it easier to gather larger sets of sequence data from sex chromosomes of less well-characterized genomes.

Acknowledgments. Financial support was obtained from the Swedish Research Council to H.E and to S.B. H.E. is a Royal Academy of Sciences Research Fellow supported by a grant from the Knut and Alice Wallenberg Foundation. Cecilia Berg, Jacob Höglund, Anna Mattsson, and Ettore Randi are acknowledged for help with DNA samples.

References

Abajian (1994) Sputnik. Available at: http://abaijan.net/sputnik/index.html

Agulnik AI, Bishop CE, Lerner JL, Agulnik SI, Solovyev VV (1997) Analysis of mutation rates in the SMCY/SMCX genes shows that mammalian evolution is male driven. Mammal Genome 8:134–138

Axelsson E, Smith NG, 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. Mol Biol Evol 21:1538–1547

Barnes WM (1992) The fidelity of Taq polymerase catalyzing PCR is improved by an N-terminal deletion. Gene 112:29–35

Bartosch-Härlid A, Berlin S, Smith NG, Møller AP, Ellegren H (2003) Life history and the male mutation bias. Evolution 57:2398–2406

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

Bohossian HB, Skaletsky H, Page DC (2000) Unexpectedly similar rates of nucleotide substitution found in male and female hominids. Nature 406:622–625

Carmichael AN, Fridolfsson AK, Halverson J, Ellegren H (2000) Male-biased mutation rates revealed from Z and W chromosome-linked ATP synthase alpha-subunit (*ATP5A1*) sequences in birds. J Mol Evol 50:443–447

Chamary JV, Hurst LD (2004) Similar rates but different modes of sequence evolution in introns and at exonic silent sites in rodents: evidence for selectively driven codon usage. Mol Biol Evol 21:1014–1023

Chang BH, Li WH (1995) Estimating the intensity of male-driven evolution in rodents by using X-linked and Y-linked *Ube 1* genes and pseudogenes. J Mol Evol 40:70–77

Chang BH, Shimmin LC, Shyue SK, Hewett-Emmett D, Li WH (1994) Weak male-driven molecular evolution in rodents. Proc Natl Acad Sci USA 91:827–831

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

Crow JF (1997) The high spontaneous mutation rate: is it a health risk? Proc Natl Acad Sci USA 94:8380–8386

Dimcheff DE, Drovetski SV, Mindell DP (2002) Phylogeny of *Tetraoninae* and other galliform birds using mitochondrial *12S* and *ND2* genes. Mol Phylogenet Evol 24:203–215

Ebersberger I, Metzler D, Schwarz C, Paabo S (2002) Genomewide Comparison of DNA Sequences between Humans and Chimpanzees. Am J Hum Genet 70:1490–1497

Ellegren H, Fridolfsson AK (1997) Male-driven evolution of DNA sequences in birds. Nat Genet 17:182–184

Ellegren H, Fridolfsson AK (2003) Sex–specific mutation rates in salmonid fish. J Mol Evol 56:458–463

Ellegren H, Smith NGC, Webster MT (2003) Mutation rate variation in the mammalian genome. Curr Opin Genet Dev 13:562–568

Filatov DA, Charlesworth D (2002) Substitution rates in the X- and Y-linked genes of the plants, *Silene latifolia* and *S. dioica*. Mol Biol Evol 19:898–907

Filatov DA (2005) Substitution rates in a new *Silene latifolia* sex-linked gene, *SlssX/Y*. 22:402–408

Fridolfsson AK, Ellegren H (2000) Molecular evolution of the avian *CHD1* genes on the Z and W sex chromosomes. Genetics 155:1903–1912

Girard M, Couvert P, Carrie A, Tardieu M, Chelly J, Beldjord C, Bienvenu T (2001) Parental origin of de novo MECP2 mutations in Rett syndrome. Eur J Hum Genet 9:231–236

Haldane JBS (1935) The rate of spontaneous mutation of a human gene. J Genet 31:317–326

Handley LJ, Ceplitis H, Ellegren H (2004) Evolutionary strata on the chicken Z chromosome: implications for sex chromosome evolution. Genetics 167:367–76

Hardison RC, Roskin KM, Yang S, Diekhans M, Kent WJ, Weber R, Elnitski L, Li J, O'Connor M, Kolbe D, Schwartz S, Furey TS, Whelan S, Goldman N, Smit A, Miller W, Chiaromonte F, Haussler D (2003) Covariation in frequencies of substitution, deletion, transposition, and recombination during eutherian evolution. Genome Res 13:13–26

Hare MP, Palumbi SR (2003) High intron sequence conservation across three mammalian orders suggests functional constraints. Mol Biol Evol 20:969–978

Ibarguchi G, Gissing GJ, Gaston AJ, Boag PT, Friesen VL (2004) Male-biased mutation rates and the overestimation of extrapair paternity: problem, solution, and illustration using thick-billed murres (*Uria lomvia*, Alcidae). J Hered 95:209–216

Kahn NW, Quinn TW (1999) Male-driven evolution among Eoaves? A test of the replicative division hypothesis in a heterogametic female (ZW) system. J Mol Evol 49:750–759

Kirkpatrick M, Hall DW (2004) Male-biased mutation, sex linkage, and the rate of adaptive evolution. Evol Int J Org Evol 58:437–440

Kumar S, Tamura K, Jakobsen IB, Nei M (2001) MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17:1244–1245

Lercher MJ, Williams EJ, Hurst LD (2001) Local similarity in evolutionary rates extends over whole chromosomes in human-rodent and mouse-rat comparisons: implications for understanding the mechanistic basis of the male mutation bias. Mol Biol Evol 18:2032–2039

Li WH, Yi S, Makova K (2002) Male-driven evolution. Curr Opin Genet Dev 12:650–656

Lu J, Wu CI (2005) Weak selection revealed by the whole–genome comparison of the X chromosome and autosomes of human and chimpanzee. Proc Natl Acad Sci USA 102:4063–4067

Ludwig MZ (2002) Functional evolution of noncoding DNA. Curr Opin Genet Dev 12:634–639

Lundberg KS, Shoemaker DD, Adams MW, Short JM, Sorge JA, Mathur EJ (1991) High–fidelity amplification using a thermostable DNA polymerase isolated from Pyrococcus furiosus. Gene 108:1–6

Malcom CM, Wyckoff GJ, Lahn BT (2003) Genic mutation rates in mammals: local similarity, chromosomal heterogeneity, and X-versus-autosome disparity. Mol Biol Evol 20:1633–1641

Makova KD, Li WH (2002) Strong male-driven evolution of DNA sequences in humans and apes. Nature 416:624–626

Makova KD, Yang S, Chiaromonte F (2004) Insertions and deletions are male biased too: a whole-genome analysis in rodents. Genome Res 14:567–573

McVean GT, Hurst LD (1997) Evidence for a selectively favourable reduction in the mutation rate of the X chromosome. Nature 386:388–392

Miyata T, Hayashida H, Kuma K, Mitsuyasu K, Yasunaga T (1987) Male-driven molecular evolution: a model and nucleotide sequence analysis. Cold Spring Harbor Symp Quant Biol 52:863–867

Nachman MW (2002) Variation in recombination rate across the genome: evidence and implications. Curr Opin Genet Dev 12:657–663

Sachidanandam R, Weissman D, Schmidt SC, et al. (2001) A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. Nature 409:928–933

Schwartz S, Zhang Z, Frazer KA, Smit A, Riemer C, Bouck J, Gibbs R, Hardison R, Miller W (2000) PipMaker—a web server for aligning two genomic DNA sequences. Genome Res 10:577–586

Shimmin LC, Chang BH, Li WH (1993a) Male-driven evolution of DNA sequences. Nature 362:745–747

Shimmin LC, Chang BH, Hewett-Emmett D, Li WH (1993b) Potential problems in estimating the male-to-female mutation rate ratio from DNA sequence data. J Mol Evol 37:160–166

Smith NGC, Eyre-Walker A (2003) Partitioning the variation in mammalian substitution rates. Mol Biol Evol 20:10–17

Smith NGC, Webster MT, Ellegren H (2002) Deterministic mutation rate variation in the human genome. Genome Res 12:1350–1356

Sokal RR, Rohlf FJ (2000) Biometry—the principles and practice of statistics ion biological research. W.H. Freeman, New York,

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol 10:512–526

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position–specific gap penalties and weight matrix choice. Nucleic Acids Res 22:4673–4680

Trappe R, Laccone F, Cobilanschi J, Meins M, Huppke P, Hanefeld F, Engel W (2001) MECP2 mutations in sporadic cases of Rett syndrome are almost exclusively of paternal origin. Am J Hum Genet 68:1093–1101

Waterston RH, Lindblad-Toh K, Birney E et al. (2002) Initial sequencing and comparative analysis of the mouse genome. Nature 420:520–562

Whittle CA, Johnston MO (2002) Male-driven evolution of mitochondrial and chloroplastidial DNA sequences in plants. Mol Biol Evol 19:938–949

Wilkin DJ, Szabo JK, Cameron R, Henderson S, Bellus GA, Mack ML, Kaitila I, Loughlin J, Munnich A, Sykes B, Bonaventure J, Francomano CA (1998) Mutations in fibroblast growth-factor receptor 3 in sporadic cases of achondroplasia occur exclusively on the paternally derived chromosome. Am J Hum Genet 63:711–716

Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci 13: 555–556

Supplementary Table 1. Primer sequences and PCR conditions

Intron	F primer name and sequence 5′ → 3′	R primer name and sequence 5′ → 3′	Ta °C	$MgCl_2$ mM	Cloned
UBAP2-1Z	ADO12-1F (Handley et al. 2004)	Adex2RZ: CCAGGTCCCACAATATACTGG	54	2.5	X
UBAP2-1W	Adex1FW: GTGTCTATACCAGCTGCTACTG	Adex2RW: CCAGGTCCCACAATATACTGA	54	2.5	X
UBAP2-2Z	Adex2FZ: GTACCACCCTTGTTGCATAAC	UBAP2-2R (Handley et al. 2004)	54	2	X
UBAP2-2W	Adex2FW: GTACCACCCTTGTTGCATAAT	UBAP2-2R (Handley et al. 2004)	54	2	X
UBAP2-4Z	Adex4FZ: CTATGGAATTACATTCCCTG	UBAP2-4R GTAGAATCCCCACGGCCAAA	52	2.5	X
UBAP2-4W	Adex4FW: CTATGGAATTACATTCCCCA	UBAP2-4R GTAGAATCCCCACGGCCAAA	54	2.5	X
UBAP2-4W (Aru)	Ado4WSF1: GCAAAATTGTTAAAAATG	UBAP2-4R GTAGAATCCCCACGGCCAAA	54	2.5	
UBAP2-4W (Tte)	Adex4FWb: CACTCCTGCAACCTTGACAG	Ado4WSR1: CATGACCAGTTAGATGTCC	51	2.5	
CHD1-6Z	Ex6FZ: ATGAACAAACTGGACAACTAC	Ex7RZ: CTCCTGCTGGCAGTTATAA	52	2.5	
CHD1-6W	Ex6FW: ATGAAGAAACTAGACAATTAT	Ex7RW: TTCCTGTTGGCAGTTATAG	50	3	
CHD1-7Z	Ex7FZ: CAATATCAAATAGTGGAAAGA	Ex8RZ: CCCTGCCATTTGCAATAGTAG	52	2.5	
CHD1-7W	Ex7FW: CAGTATCAAGTAGTGGAAAGG	Ex8RW: CCCTGCCATTTGCAATAGTAA	50	3	
CHD1-10Z	Ex10FZ: CTCCTCAGATGAATGCTGTA	Ex11RZ: CATATGTCGTCAGAAGTATG	52	2.5	
CHD1-10W	Ex10FW: CTCCTCAGATGAATGCTGTG	Ex11RW: CATAAGTTGTTAGAAGTATA	50	3	
CHD1-11Z	Ex11FZ: CACAGACTAAACGATTAAAGTTTAAC	Ex12RZ: GCTTCATCAACTCCTATGAATG	52	2.5	
CHD1-11W	Ex11FW: CACAGACCAAGCGATTAAAATTTAAT	Ex12RW: GCTTCATCAATCCCTATAAACA	50	3	
CHD1-12Z	Ex12FZ: CTCAAAGAGCTGTGGTCTTTGT	Ex13RZ: TTTGTGAAGACTTGCATAACCAT	52	2.5	X
CHD1-12W	Ex12FW: CTCAAAGAGCTTTGGTCTTTGC	Ex13RW: TTTGTGAAGACTTGCATAACCAA	50	3	X
CHD1-13Z	Ex13FZ: CCTGCTAAGGTTGAACAAATTC	Ex14RZ: CTGCCTTTTGAACCTTTACTG	52	2.5	
CHD1-13W	Ex13FW: CCTGCTAAGGTTGAGCAAATTT	Ex14RW: CTGCCTTTAGAACCTTTGCTC	50	3	
CHD1-15Z	2669FZ: CTCAGATGGTGAGGATGCTG	2718R: ATTGAAATGATCCAGTGCTTG	52	3	
CHD1-15W	2669FW: CTCAGATGGTGAGGATGCTA	2718R: ATTGAAATGATCCAGTGCTTG	TD 60–50	3	
CHD1-17Z	Ex17FZb: GATTCTGACTGGAATCCACAC	Ex17RZb: GAATATCTTCTTCTACTGATCC	50	2.5	
CHD1-17Z (Aru)	Ex17FZc: CACAGGCGAGAGCTCATAG	Ex18RZc: CTTCTACTGATCCTTTTGTGAC	50	2.5	
CHD1-17W	Ex17FW: GGAATCCACAAAATGATCTGCAA	Ex18RW: CTAAATGATCTAACACCATCTTT	53	3	
CHD1-21Z	Ex21FZ: CAAAAAGAACTTGAAGAAATATAC	Ex22RZ: TCGAGGAATGGTTCGAGGTCT	52	2.5	
CHD1-21W	Ex21FW: CAAAAAGAACTTGAAGAAATATAT	Ex22RW: ACGGGGAATAGTTCGTGGTCG	50	2.5	
CHD1-22Z	Ex22FZ: AGACCTCGAACCATTCCTC	Ex23RZ: AGGGCCACCAAATTTCTTGTAAC	52	2.5	X
CHD1-22W	Ex22FW: CGACCACGAACTATTCCCC	Ex23RW: TGGGCCACCAAATTTCTTGTAAC	50	3	X
CHD1-24Z	3770FZ: GGCTTTAAAGGACARTTCA	3830RZ: CATTCACCTGCACTCCTGAG	52	3	
CHD1-24W	3720FW: TGAAACAGACCTTAGACGTC	3830RW: CATTCACCTGCACTCCTGCT	50	3	
CHD1-25Z	Ex25FZ: GGCCCAACGTTTCGAATCT	Ex26RZ: CGAAGTGAGCAGCCTTGGTGTGGC	52	2.5	
CHD1-25W	Ex25FWb: AAAGGCCCAACATTCCGAATA	Ex26RWb: CCAATCTATATCAAAATGAGCT	53	3	
SPIN-3Z	SP2WF: GGGAGTGGACCTGTAACACAG	SP2WR: CATATGTTCCACAGCTTTGCCA	52	2.25	X
SPIN-3W	SP2ZF: GGAAGCGGACCCATAACACAA	SP2ZR: CATATGTTCCACAGCTTTGCCT	TD 60–50	2.5	

Supplementary Table 2. Genbank accessions numbers

Locus	*G. gallus*	*M. gallopavo*	*T. tetrix*	*A. rufa*	*C. coturnix*
UBAP2-1Z	AY189761	AY189760	AY628466	AY628465	AY628464
UBAP2-2Z	AY189777	AY189776	AY628471	AY628470	AY628469
UBAP2-4Z	AY426726	AY426732	AY628476	AY628475	AY628474
CHD1-6Z	AY298997	AY299005	AY628601	AY628600	AY628599
CHD1-7Z	AY298998	AY299006	AY628606	AY628605	AY628604
CHD1-10Z	AY298993	AY299000	AY628551	AY628550	AY628549
CHD1-11Z	AY298994	AY299001	AY628556	AY628555	AY628554
CHD1-12Z	AY426729–30	AY426733	AY628561	AY628560	AY628559
CHD1-13Z	AY298995	AY299002	AY628566	AY628565	AY628564
CHD1-15Z	AF525980	AF526014	AY628571	AY628570	AY628569
CHD1-17Z	AY628573	AY628577	AY628576	AY628575	AY628574
CHD1-21Z	AY628578	AY628582	AY628581	AY628580	AY628579
CHD1-22Z	AY426728	AY426734	AY628586	AY628585	AY628584
CHD1-24Z	AF526056	AY299008	AY628591	AY628590	AY628589
CHD1-25Z	AY298996	AY299004	AY628596	AY628595	AY628594
SPIN-3Z	AY194142	AY194143	AY628546	AY628545	AY628544
UBAP2-1W	AY189754	AY189755	AY628451	AY628450	AY628449
UBAP2-2W	AY189767	AY189768	AY628456	AY628455	AY628454
UBAP2-4W	AY426725	AY426731	AY628461	AY628460	AY628459
CHD1-6W	AY628528	AY628532	AY628531	AY628530	AY628529
CHD1-7W	AY298971	AY426735	AY628536	AY628535	AY628534
CHD1-10W	AY298965	AY298959	AY628481	AY628480	AY628479
CHD1-11W	AY298966	AY298960	AY628486	AY628485	AY628484
CHD1-12W	AY426727	AY426736-7	AY628491	AY628490	AY628489
CHD1-13W	AY628493	AY628497	AY628496	AY628495	AY628494
CHD1-15W	AY298972	AF526013	AY628501	AY628500	AY628499
CHD1-17W	AY298970	AY299013	AY628506	AY628505	AY628504
CHD1-21W	AY628508	AY628512	AY628511	AY628510	AY628509
CHD1-22W	AY298968	AY298963	AY628516	AY628515	AY628514
CHD1-24W	AF526055	AF526077	AY628521	AY628520	AY628519
CHD1-25W	AY298969	AY298964	AY628526	AY628525	AY628524
SPIN-3W	AY194125	AY194126	AY628541	AY628540	AY628539

Supplementary Table 3. Pairwise K_z with standard errors estimated by 1000 bootstrap replicates

	Gallus gallus	*Alectoris rufa*	*Coturnix coturnix*	*Meleagris gallopavo*
Gallus gallus				
Alectoris rufa	0.107 ± 0.004			
Coturnix coturnix	0.129 ± 0.004	0.106 ± 0.003		
Meleagris gallopavo	0.110 ± 0.004	0.116 ± 0.004	0.138 ± 0.005	
Tetrao tetrix	0.121 ± 0.005	0.126 ± 0.004	0.147 ± 0.005	0.081 ± 0.003

Supplementary Table 4. Pairwise K_w with standard errors estimated by 1000 bootstrap replicates

	Gallus gallus	*Alectoris rufa*	*Coturnix coturnix*	*Meleagris gallopavo*
Gallus gallus				
Alectoris rufa	0.062 ± 0.003			
Coturnix coturnix	0.065 ± 0.002	0.059 ± 0.002		
Meleagris gallopavo	0.057 ± 0.002	0.071 ± 0.003	0.077 ± 0.003	
Tetrao tetrix	0.064 ± 0.003	0.080 ± 0.003	0.083 ± 0.004	0.045 ± 0.002