J Mol Evol (1999) 49:750–759

JOURNAL OF MOLECULAR EVOLUTION

Male-Driven Evolution Among Eoaves? A Test of the Replicative Division Hypothesis in a Heterogametic Female (ZW) System

Nate W. Kahn,* Tom W. Quinn

Department of Biological Sciences, University of Denver, Denver, CO 80208, USA

Received: 19 January 1999 / Accepted: 8 June 1999

Abstract. Because avian females are heterogametic, the reverse of mammals, avian sex chromosomes undergo significantly different patterns and numbers of DNA replications than do those in mammals. This makes the W (female-specific) and the Z chromosomes an excellent model system for the study of the replicative division hypothesis, which purports that DNA substitution rate is determined by the number of germline replications. The sex-specific chromosome in birds (the W) is predicted to change at the slowest rate of all avian chromosomes because it undergoes the fewest rounds of replication per unit of evolutionary time. Using published data on gametogenesis from a variety of sources, we estimated the ratio of male-to-female germline replications (c) in galliforms and anseriforms to be approximately 4.4. The value of c should predict the value of the ratio of male-to-female mutation rates (α_m) if the replicative division hypothesis is true. Homologous DNA sequences including an intron and parts of two exons of the CHD gene were obtained from the W and the Z chromosomes in ostrich, sage grouse, canvasback duck, tundra swan, and snow goose. The exons show significantly different nucleotide composition from the introns, and the W-linked exons show evidence of relaxed constraint. The Z-linked intron is diverging ≈ 3.1 times faster than the W-linked intron. From this, α_m was calculated to be approximately 4.1, with a confidence interval of 3.1 to 5.1. The data support the idea that the number of replicative divisions is a major determinant of substitution rate in the Eoavian genome.

Key words: Male-driven evolution — ZW chromosomes — male/female mutation rate (α_m) — replication error — molecular clock

Introduction

A Molecular Clock and the Replicative Division Hypothesis

Because of its importance in phylogenetic inference, understanding the rates at which DNA substitutions occur is among the most important recent issues in evolutionary biology. Although it has been proposed that sequence change is constant enough that it can be used as a "molecular clock" (Zuckerkandl and Pauling 1965; Wilson et al. 1977), several recent studies have challenged that idea. For example, the metabolic rate hypothesis proposes that higher metabolic rates cause higher levels of oxidative damage to DNA, leading ultimately to different substitution rates (Martin et al. 1992; Martin and Palumbi 1993; Rand 1994; Mindell et al. 1996). Others have proposed that DNA repair mechanisms in some taxa are significantly more or less efficient than in others, and this determines the substitution rate (Drake and Baltz 1976; Goodman et al. 1984; Britten 1986). The focus of this study, the replicative division hypothesis (RDH), asserts that if point mutations occur predominantly during DNA replication (Topal and Fresco 1976; Wilson et

* *Present address:* University of Colorado Health Sciences Center, Department of Biochemistry and Molecular Genetics, Campus Box B121, 4200 East Ninth Ave., Denver, CO 80262-0295, USA
Correspondence to: Nate W. Kahn; *e-mail:* Nate.Kahn@UCHSC.edu

al. 1977; Miyata et al. 1987b), then the rate of mutation (and ultimately the substitution rate) should be determined by the number of replicative divisions that occur in those cell lineages that span generations, i.e., the germline cells. If the number of germline replications per generation were constant among different taxa, then the RDH predicts that the molecular clock should run faster in organisms with a short generation time because their germline cells will undergo more rounds of replication per unit chronological time. This aspect of the RDH has been termed the generation time effect, and it was used to explain why hominoids seem to be evolving more slowly than rodents at the sequence level (Kohne 1970; Kohne et al. 1972; Li et al. 1996).

If, as proposed by the RDH, replicative divisions in germline cells drive mutation (and substitution) rates, some specific predictions can be made and used to test its validity. Because the number of germline replications that occur during sperm production in male vertebrates far exceeds those required for egg production in females, the RDH predicts that chromosomes passed only through male lineages, such as the mammalian Y chromosome, should undergo sequence evolution more rapidly than autosomes (which are shared equally between male and female lineages). Similarly, chromosomes passed only through female lineages, such as the avian W chromosome, should evolve more slowly. Because the RDH predicts that more sequence change occurs during the numerous replicative divisions that occur in males, it has also been termed "male-driven evolution" (Miyata et al. 1987b).

Previous Studies

With one exception, previous studies of the RDH have focused on mammalian chromosomes. Two general methods for estimating substitution rates have been used. One approach is to sample many unrelated sites or genes on each chromosome and to average those values to arrive at a chromosome-specific rate (Miyata et al. 1987a, 1987b). This approach is difficult because there is a large (fivefold) range of synonymous substitution rates among genes located on autosomes (or the X chromosome) making it necessary to sample high numbers of genes/sites (Shimmin et al. 1993b). A second approach has been to make comparisons using related but genetically isolated (paralogous) sequences that are located on those chromosomes being compared (Hayashida et al. 1992; Shimmin et al. 1993a; Chang et al. 1994; Chang and Li 1995). As an example, Shimmin et al. (1993a) made orthologous comparisons among primates using an intron of a Y-linked gene (ZFY) to calculate substitution rates of that chromosome and then made similar comparisons within primates using a genetically isolated X-linked paralogue (ZFX) to calculate a comparable X chromosomal rate. In this case, the Y chromosomal substitution rate was shown to be 2.3 times that of the X chromosome. Overall, if gene conversion can be ruled out, or at least chronologically identified, then comparison among paralogous sequences located on all chromosomes of interest provides a powerful method for calculating the male-to-female mutation ratio (α_m) (Shimmin et al. 1993b; Li et al. 1996).

The observed value of the male-to-female mutation ratio ($\mu_m / \mu_f = \alpha_m$) in higher primates is approximately four to six (Miyata et al. 1987a, 1987b; Hayashida et al. 1992; Wolfe and Sharp 1993; Pamilo and Bianchi 1993; Shimmin et al. 1993a; Li et al. 1996), whereas the value of α_m in rodents is approximately two (Lanfear and Holland 1991; Chang et al. 1994; Change and Li 1995). These observed α_m ratios are in close agreement with the corresponding male-to-female ratios of germline replications ($N_m / N_f = c$), estimated independently using information about mammalian reproductive biology (Chang et al. 1994; Li et al. 1996).

Thus, the RDH is supported from the mammalian data sets. However, there is still a major issue that mammalian studies have not clarified: Is the observation of a greater substitution rate on the Y chromosome the result of an asymmetrical number of germline replications between the sexes, or is it the result of the unusual nature of the Y itself (i.e., a small, heterochromatic chromosome, the bulk of which is haploid, repetitive, and noncoding)? Because the Y chromosome is unique in several ways, it becomes difficult to attribute its increased substitution rate solely to the effects of number of rounds of replication in male lineages (Wu and Li 1985; Miyata et al. 1987a; Easteal et al. 1995).

Avian Model System

Perhaps the best way to address these ambiguities is to use information from other organisms in which both the magnitude and direction of rate predictions made by the RDH differ radically from those predicted for mammals. Such altered predictions are generated in taxa with female (ZW), rather than male (XY), heterogamety—a condition found in birds (see Miyata et al. 1987b; Crow 1997). The Z chromosome spends (through many generations) two-thirds of its time in males and one-third in females, thus undergoing spermatogenesis two-thirds and oogenesis one-third of the time. The W chromosome spends 100% of its time in females and thus always undergoes oogenesis. According to the RDH, because of the greater numbers of replications involved in spermatogenesis, the chromosome predicted to be the fastest evolving one in birds is the shared one (Z), rather than the sex-specific one as in mammals (Y). In contrast, the female-specific chromosome in birds (W) is predicted to change more slowly than all other avian chromosomes

because it undergoes the fewest rounds of replication per unit time. If the predictions of the RDH are true, then the W chromosome should, in general, be the slowest evolving vertebrate chromosome ever studied. Therefore, examination of substitution rates in an avian model system should further clarify interpretations of the mammalian studies.

The discovery that the chromo-ATPase/helicase-DNA–binding domain (CHD) gene (Delmas et al. 1993; Funahashi et al. 1993) exists on both the W and Z chromosomes in almost all avian species (Ellegren 1996; Griffiths et al. 1996) has provided an excellent opportunity to test the RDH in birds. Ellegren and Fridolfsson (1997) have completed such a study in passerines (Neoaves). We have used a different portion of the CHD gene in Galliformes and Anseriformes (Eoaves) to compare mutation rates on the W and Z chromosomes, and thus calculate the value of α_m in large, longer-lived birds. This study (1) uses DNA sequence data from both exon and intron regions that are homologous to both the W and the Z chromosomes to examine divergence rates and to test the hypothesis that DNA in males will accumulate substitutions faster than DNA in females; (2) contributes to the overall understanding of molecular evolution and the rate of evolution in vertebrates; and (3) adds information to the growing data set about the sex chromosomes of birds.

Materials and Methods

Species Choices and Tissue Extraction

Ostrich (*Struthio camelus;* blood), sage grouse (*Centrocercus urophasianus;* muscle), canvasback duck (*Aythya valisineria;* blood), tundra swan (*Olor columbianus;* heart), and snow goose (*Chen caerulescens;* blood) were specifically chosen to sample phylogenetic distances at different depths within the Eoaves clade of the avian phylogenetic tree (Bellrose 1976; Monroe and Sibley 1993). By making comparisons within the subfamily Anserinae, within the order Anseriformes, between two orders thought to be sister taxa (Galliformes and Anseriformes), and including the order Struthioniformes for a distant outgroup, our intention was to obtain a set of distance measurements that span a wide variety of divergences such that final rate calculations could be based on subsets of sequence comparisons that were not so divergent that saturation effects were extreme, nor so similar that too few observed differences made sampling error excessive. Our choices were guided by calculating mitochondrial sequence distances among various waterfowl (Shields and Wilson 1987; Van Wagner and Baker 1990; Zimmer et al. 1994) and Galliformes (Desjardins and Morais 1990) and using these to estimate the expected autosomal sequence difference using the approximation that the nuclear genome undergoes substitutions at roughly 10 times the rate of the nuclear genome (Brown et al. 1982). Total genomic DNA was extracted from male and female tissue samples as described in Quinn and White (1987). The resulting aqueous solution of genomic DNA was ethanol-precipitated and resuspended in TE (Tris-HCl 10 m*M*, EDTA 1 m*M* pH 8.0) to a concentration of 300 μg/ml.

Primer Design and PCR

Five primers were designed at locations both upstream and downstream of Griffiths et al.'s (1996) previously published P2 and P3 primer sites on the CHD gene, by aligning mouse, chicken, and *Drosophila* CHD cDNA sequences (Delmas et al. 1993; Funahashi et al. 1993; Stokes et al. 1996) and then targeting conserved regions. All possible pairwise combinations of primers were characterized by PCR; the two that provided the most consistent amplifications across all taxa were chosen for continued analysis (1092F 5′-TCCAATATGGATGAAGATGA-3′ [numbered according to Fuhahashi et al.'s (1993) chicken cDNA] and P2 [Griffiths et al. 1996]). PCR amplifications were performed and products visualized as described in Quinn (1992) using 30 cycles in a Perkin-Elmer DNA thermal cycler, each cycle including the following steps: 94°C for 1 min, 56°C for 1 min, 72°C for 2 min.

Cloning

Each 1092L-P2 PCR product was purified using QIAQuick PCR Purification minicolumns (Qiagen) following manufacturer's recommendations. Each was then ligated into pGEM-T (Promega) and used to transform JM109 Supercompetant cells (Promega), following the manufacturer's recommended procedures. A minimum of eight white colonies from male and female targets respectively were chosen for sequence determination.

Amplification and Sequencing

"Colony touch" asymmetrical PCR (CTAPCR) was performed directly on bacterial colonies as described by Hofmann and Brian (1991). Each CTAPCR product was cleaned and sequenced following Quinn (1992) and using Sequenase 2.0 kits (Amersham). Sequencing products were electrophoresed through 6% salt-gradient polyacrylamide gels and through 5% polyacrylamide gels for resolution of larger fragments following Sambrook et al. (1989).

Sequence and Phylogenetic Analyses

The resulting sequences (Fig. 1) were aligned using the Higgins algorithm of MacDNASIS v3.0 (Hitachi Software Eng. Co. Ltd.) and by eye, starting with most closely related taxa (snow goose/tundra swan) and successively working out to the most distant (canvasback duck, sage grouse, and then ostrich) (cf. Mindell 1991). Maximum parsimony analyses were performed using the exhaustive search algorithm of PAUP v3.1.1 (Swofford 1993). Bootstrap values were calculated by branch-and-bound search using 1,000 repetitions. For introns, calculation of rates and their standard errors was performed on MEGA v1.0 (Sudhir et al. 1993), using the complete-deletion option for gaps; the RDH is concerned with base pair substitutions, as it is believed that the mechanism causing insertions and deletions is different (Shimmin et al. 1993a; Li et al. 1996). Two portions of sequence, one among W and one among Z introns, border deletions and were difficult to align; these portions (shown boxed in Fig. 1B) were removed from the analysis. Divergence values were corrected to account for saturation effects using the Tajima-Nei correction (Tajima and Nei 1984). For exons, calculation of synonymous (Ks) and nonsynonymous (Ka) rates and their standard errors was performed according to the method of Li (1993) implemented in FENS pre-release vβ1.2 (de Koning et al. 1999). The neighbor-joining algorithm in MEGA v1.0 was used to construct trees and calculate branch lengths within each chromosomal clade.

DNA sequence from ostrich was used as an outgroup for parsimony analysis of all sequences together in one total data set, because ratites are considered to be an outgroup to galliforms and waterfowl based on morphological traits (Wetmore 1960) and DNA–DNA hybridization

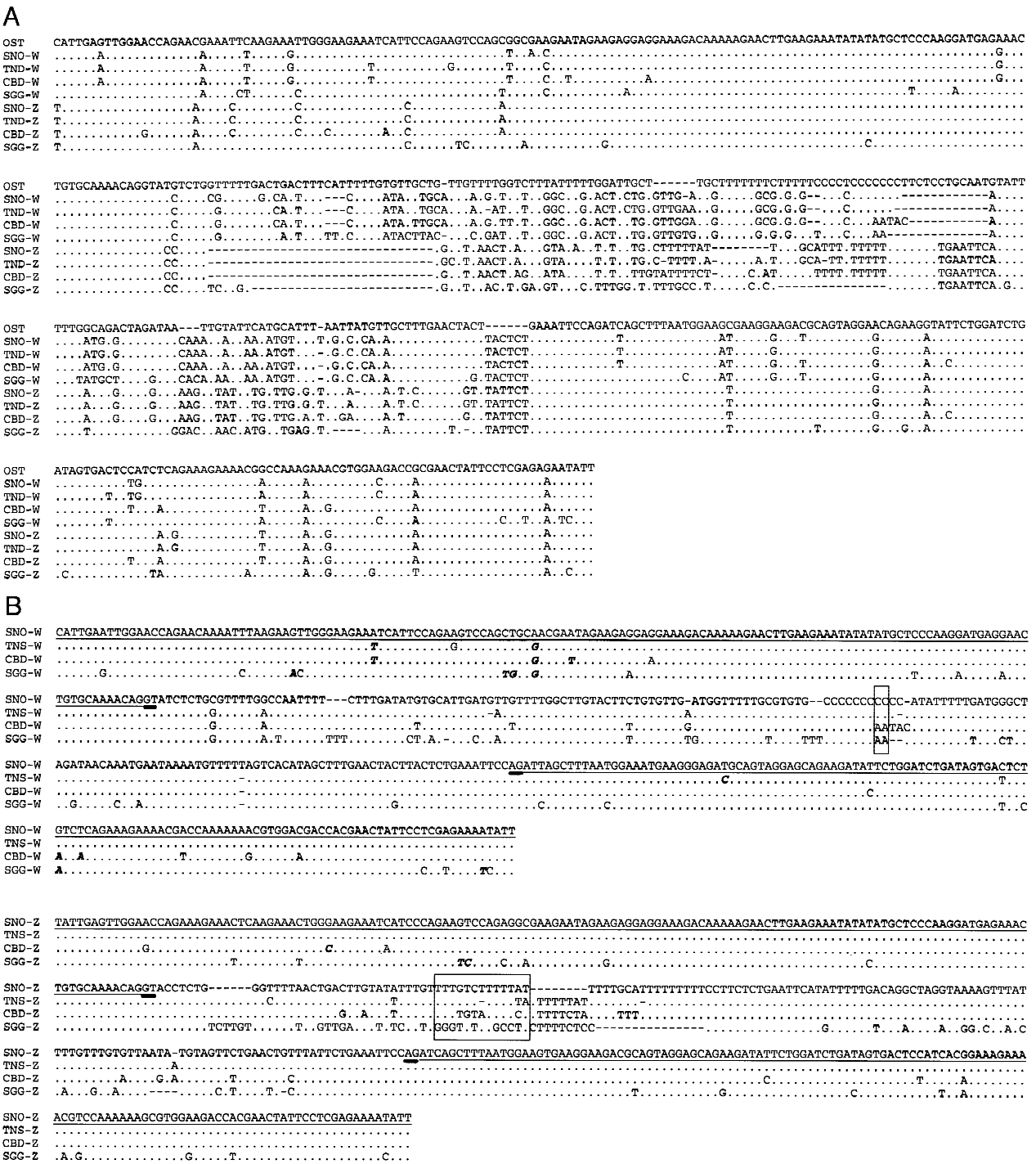

Fig. 1. **(A)** Alignment of all paralogous sequences to ostrich CHD. Abbreviations are: OST = ostrich, SNO = snow goose, TNS = tundra swan, CBD = canvasback duck, SGG = sage grouse, W = W chromosome–linked, Z = Z chromosome–linked. As described in the text, all parsimony analyses from this data set, and its subsets, resulted in the same overall tree that separates W and Z chromosome–linked sequences into two separate clades, with high bootstrap values (see Fig. 2). **(B)** Alignment of orthologous W-CHD and Z-CHD sequence sets. Abbreviations are the same as above. Indels are shown marked with dashes. Exon regions are underlined. Intron splice sites are bold-underlined. Nonsynonymous substitutions are bold-italicized. Boxed regions were removed from neighbor-joining analysis because of difficulty in alignment. As described in the text, neighbor-joining analysis was performed within each of these orthologous clades to estimate branch lengths (see Fig. 2).

data (Sibley et al. 1988). Additionally, parsimony and distance analyses were performed on W- and Z-linked sequences, respectively, as separate groups. Within the "W clade," sage grouse W sequence was used as outgroup; within the "Z clade," sage grouse Z sequence was used as outgroup. Galliformes are the closest sister group to Anseriformes, based on DNA–DNA hybridization (Sibley et al. 1988), mitochondrial DNA (Mindell et al. 1997), and traditional classifications (Johnsgard 1965; Bellrose 1976; Peters 1986).

Results

Discovery of an Intron Between the Helicase and DNA-Binding Regions

The PCR primers were designed to flank the boundary between the helicase and the DNA-binding regions of the CHD gene, and all products were approximately 190 bp larger than expected from cDNA information. By comparison with published cDNA sequence, the observation of 5′-GT and AT-3′ splice sites, and the lack of an open reading frame consistent with other parts of the sequenced fragments, it was apparent that the size discrepancy was caused by an intron located roughly in the middle of the amplified fragments. With the exception of ostrich, female samples generated two sizes of product, and males one size; further research has shown that this results from intron size differences between the Z- and the W-linked copies of the CHD gene. This size difference was used by Kahn et al. (1998) to develop a primer pair that can be used to determine gender in almost all avian taxa.

Determination of W and Z Sequence from Each Species

A minimum of eight clones from each sex were sequenced in all species. With the exception of ostrich, female-derived clones revealed two sequences: one matching the single sequence found in all male clone sets of each species, and a different homologous sequence. Overall, the two sequences found in female-derived clones from sage grouse, tundra swan, canvasback duck, and snow goose appeared at a ratio of ≈ 1:1 (26 presumed W sequences: 24 presumed Z sequences). In cases where the eight sequences from females did not include a minimum of three consensus sequences of each type, additional clones were sequenced to reach that total. Verification that CHD exists on the W and the Z chromosome in waterfowl was confirmed by Southern blot analysis, using multiple digests of duck and goose (unpublished data), which showed same-size bands of double intensity in males and single intensity in females, as well as some bands unique to females. From these combined findings we concluded that the sequence shared in males and females was amplified from the Z chromosome, and the sequence unique to females was amplified from the W chromosome.

With a few exceptions (n = 4 sequences, each with one nucleotide difference), sequences from male-derived clones proved to be identical within a species. We assume those exceptions to result from *Taq* polymerase errors or single base polymorphisms. Because CHD has been reported to be a single-copy gene on both W and Z chromosomes in birds except for ratites (Ellegren and Fridolfsson 1997; Griffiths and Korn 1997) and because sequencing revealed just two related sequences in females and one in males, we concluded that these sequences were representative of a gene present in single copy on each sex chromosome, except in ostrich. In ostrich, only one sequence was found, identical between sexes, despite extensive sampling (males n = 8, females n = 16). This finding adds support to the observations by others (Ellegren 1996; Ellegren and Fridolfsson 1997; Fridolfsson et al. 1998) that the CHD gene is not sex-linked or does not exist on the W chromosomes in ratites. Another plausible explanation is that an identical CHD gene exists on both sex chromosomes of ratites (cf. Ogawa et al. 1998).

Altogether, Z-chromosomal and W-chromosomal sequences were obtained from sage grouse, Z = 428 bp, W = 457 bp; canvasback duck, Z = 441 bp, W = 456 bp; tundra swan, Z = 439 bp, W = 453 bp; and snow goose, Z = 433 bp, W = 455 bp (Fig. 1); only one type of sequence was found in ostrich, 457 bp, from both males and females (Fig. 1A). Each included one complete intron flanked by two partial exons. The two exon sequences provided highly conserved boundaries on both sides of the intron (Fig. 1B, underlined); the GT-AG intron splice sites were found immediately inside of the homologous exon regions (Fig. 1B, bold-underlined), and were considered the outer boundaries for the intron. An area of difficult alignment was found in each sex-linked clade of sequences (Fig. 1B, boxed sequences); each area was next to a large deletion. We excluded these areas from distance analysis as a conservative precaution against possible misalignment. However, making these exclusions probably underestimates the true distance between Z and W sequences slightly, perhaps more so among Z sequences. All sequences have been deposited in the GenBank Sequence Database under accession numbers AF138961 to AF138969.

Sequence Analysis and Diversity

Pairwise distances between intron sequences (Table 1), and between exon sequences (Table 2) for synonymous sites (Ks) and for nonsynonymous sites (Ka), were calculated for W-linked and Z-linked sequences. For introns, there is less divergence among W-intron sequences than among Z-intron sequences, and the greatest distances are found between the W- and the Z-intron sequences. In the introns, based on branch-length differences overall, the Z sequences are diverging approximately 3.1 times faster than W sequences. In the exons, overall, W-linked nonsynonymous (replacement) sites have distance measures approximately 3.3 times greater than Z-linked nonsynonymous sites (Ka, Table 2), whereas W- and Z-linked synonymous (silent) positions have approximately the same distance measures (Ks, Table 2). Z-linked exon synonymous sites have approximately the same distance measures as their corresponding Z-linked introns; W-linked exon synonymous sites

Table 1. Number of nucleotide substitutions per site

For CHD-W Intron sequences				
	SGG-W	CBD-W	TNS-W	SNO-W
SGG-W		0.1017	0.1018	0.1229
CBD-W	0.0263		0.0239	0.0300
TNS-W	0.0264	0.0121		0.0179
SNO-W	0.0295	0.0136	0.0104	
For CHD-Z Intron sequences				
	SGG-Z	CBD-Z	TNS-Z	SNO-Z
SGG-Z		0.2432	0.2633	0.2631
CBD-Z	0.0553		0.0524	0.0522
TNS-Z	0.0569	0.0221		0.0168
SNO-Z	0.0568	0.0219	0.0120	

Distances above diagonal, standard errors below diagonal. Distances adjusted using Tajima-Nei correction. Abbreviations are the same as Fig. 1.

Table 2. Number of nucleotide substitutions per site

For CHD-W Exon sequences				
	SGG-W	CBD-W	TNS-W	SNO-W
SGG-W		0.0271	0.0292	0.0269
CBD-W	0.2952		0.0182	0.0243
TNS-W	0.1915	0.0989		0.0141
SNO-W	0.1937	0.0790	0.0193	
For CHD-Z Exon sequences				
	SGG-Z	CBD-Z	TNS-Z	SNO-Z
SGG-Z		0.0141	0.0100	0.0100
CBD-Z	0.2815		0.0040	0.0040
TNS-Z	0.2407	0.0499		0.0000
SNO-Z	0.2407	0.0499	0.000	

Ka (nonsynonymous) values above diagonal. Ks (synonymous) values below diagonal. Distances calculated by the method of Li (1993). Abbreviations are the same as Fig. 1.

have overall slightly greater distance measures than their corresponding W-linked introns, but are not significantly different ($p > 0.5$) based on a *t* test that corrects for multiple pairwise comparisons.

Phylogenetic Relationships of Sequences

The possibility of gene conversion was examined by parsimony analyses. The entire data set generated a single tree that contains two monophyletic groups of sequences, W and Z chromosome derived (Fig. 2). Searches were also conducted with sequence data from the introns only and from the exons only. Each analysis produced a single most parsimonious tree, each with the same branching order, except the exon data, which combined W-derived sequences from canvasback duck, tundra swan, and snow goose into a polytomy, rather than placing tundra swan and snow goose as sister taxa. We attribute this result to fewer informative sites in the exons. Parsimony analysis using midpoint rooting without ostrich generates the same two sex-chromosome-specific clades, i.e., W and Z. Within each clade, phylogenetic relationships within these taxa are congruent with the conventional placements based on morphology, fossils, and DNA–DNA hybridization. Bootstrap analysis shows that the separation of the W and the Z chromosomal sequences into distinct clades is robust (see Fig. 2). Thus, there is no evidence of gene conversion on a large or moderate scale within our data set. The possibility remains that very small tracts of recombination may have occurred; no tests were conducted to rule this out. If small portions of information exchange occurred, then it would cause the divergence between W- and Z-linked copies to be slightly underestimated. Because recombination between the W and Z clades is absent on the larger scale, each orthologous tree (W or Z) could also be reconstructed using paralogous sequences as an outgroup.

Discussion

Analysis of Germline Replication Number in Birds

The RDH predicts that the ratio of male-to-female germline replications (*c*) should match the ratio of male-to-female substitution rates (α_m) calculated from sequence data. Estimates of *c* in mammals have generally been made using a few well-studied species and those results extrapolated to less well-studied species using the rationale that the processes involved with oogenesis and spermatogenesis are very similar among all mammalian taxa (Roosen-Runge 1977). In birds, however, *c* has never been estimated specifically (Crow 1997; Ellegren and Fridolfsson 1997). Here, we have used information about Japanese quail and chickens (Galliformes) to estimate the number of germline replications in male and female Eoaves. In the following calculation of *c*, information about reproduction in females is presented first, followed by information about males.

Female chickens generate a maximum of approximately 400,000 primordial oogonia during embryogenesis in their active ovary by mitotic divisions (Hughes 1963; Mendez-Herrera et al. 1993). Some cell loss (up to 40%) occurs during the earlier migration of primordial germ cells (PGCs) from the epiblast and subsequent incorporation into the germinal epithelium (Swift 1914; Meyer 1964; Gilbert 1979; Kuwana 1993). With 50% PGC loss during migration, 5.2×10^5 primordial oogonia would be generated during embryogenesis by approximately 20 rounds of replication. With no loss during migration, 19 rounds of replication would be required. We have used an average 25% loss, which requires 19 replications to produce $\approx$ 400,000 primordial oogonia. One additional replicative round occurs in preparation for meiosis (with two subsequent delayed divisions). The total of 20 replicative divisions per generation (N_f) is independent of age of reproduction. The number of primary oocytes that the female carries into adulthood is less than the maximum oogonia created because of cell

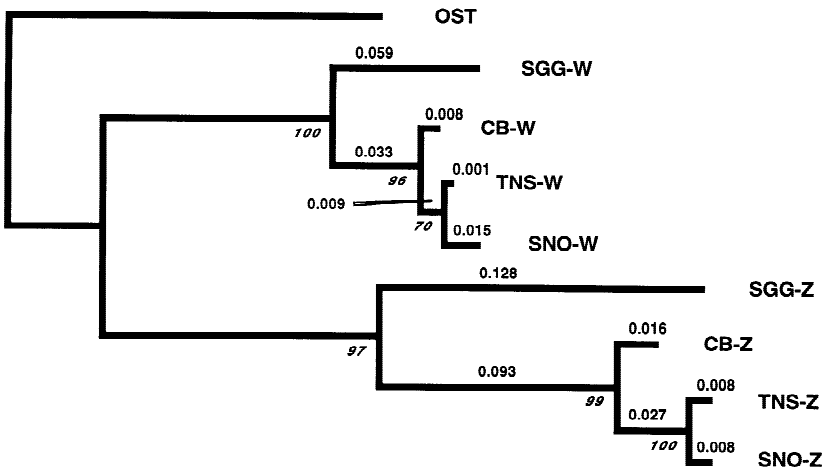

Fig. 2. Phylogenetic tree of Z chromosome and W chromosome sequences. The two separate clades of sequences (Z and W) resolved by parsimony are shown, with bootstrap values for the resolution of the two clades from the total data set shown in italicized numbers below their respective branch nodes. The branches within (but not between) the W and Z clades are drawn to show their relative lengths from neighbor-joining analysis, with the estimated branch-length values shown in bold above each branch. The Z-linked sequences are diverging approximately three times faster than the W-linked sequences.

degeneration that occurs in later stages of embryogenesis (Hughes 1963).

Spermatogenesis proceeds via a set of organs and cellular processes similar to the way it does in mammals (Roosen-Runge 1977; Lake 1981; Froman 1994; Kirby and Froman 1995). In Japanese quail, 3.1×10^8 spermatozoa exit the seminiferous tubules per day (Clulow and Jones 1982). Following the method of Vogel and Rathenberg (1975), we estimate that 6×10^8 PGCs inhabit each testicle, based on average testicle weight, number of spermatogenic divisions as derived from sperm output, and estimates of sperm loss (Vogel et al. 1976; Lake 1981; Clulow and Jones 1982). Assuming that this represents all PGCs produced, approximately 29 mitotic cell divisions are required in embryogenesis. PGCs undergo further replicative divisions in the seminiferous epithelial tissues to produce spermatogonia. Epithelial cycle time in the seminiferous tubule is 2.7 days (Lin et al. 1990). The average yearly period of fertility in sage grouse and snow geese is approximately 2.5 months (75 days), and it occurs once during the spring breeding season when the testis swell (Eng 1963; Ricklefs 1972; Bellrose 1976; Lake 1981; Gill 1995). Thus, there are approximately 75/2.7 = 27.8 spermatogonial stem cell divisions per year. In Japanese quail, 32 spermatids are produced from each spermatogonial stem cell (Lin and Jones 1992); thus, four replications occur (and five divisions) in the production of haploid cells. Average age of reproductive maturation and average age of reproduction, respectively, are approximately 1 year old and 3 years old in sage grouse (Clait Braun, personal communication; Eng 1963; Ricklefs 1972), and 3 years old and 5 years old in snow geese (Fred Cooke, personal communication; Ricklefs 1972; Cooke et al. 1995). Thus, for sage grouse and snow geese, we can estimate the total number of germline divisions that have occurred in the average male at the average time of reproduction (N_m) to be: 29 embryogenic div. + (27.8 spermatogonial div. / year × 2 years) + 4 spermatogenic div. = 88.6 total germline divisions.

Combining the estimates of replicative numbers from males and females in a ratio, we obtain the value of $c \approx$ 88.6 male-replications / 20 female-replications ≈ 4.4. Based on this estimate, the RDH predicts that the ratio of male-to-female mutation rates (α_m) should equal 4.4.

Comparison of the Male and Female Mutation Rates

The equation $\alpha_m = ((3\mu_z / \mu_w) - 1) / 2$ allows the male-to-female mutation ratio (α_m) to be calculated from our sequence data (Miyata et al. 1987b; Crow 1997). Here, mutation rates (μ) were estimated separately for introns and exons using corrected substitution values (Tables 1 and 2).

Our data contain sequence information from two introns; one is W-linked and the other is Z-linked. Neighbor-joining trees and branch lengths were calculated internally for each clade (W and Z, respectively) of intron sequence. Using the same statistical procedures that have been used for comparisons of mammalian sex chromosomal sequence (Shimmin et al. 1993a; Chang et al. 1994), we calculated the ratio E(Z) / E(W) ≈ 3.06 with SE = 0.35, from intron Z- / W-branch lengths. As in previous studies of mammals, the error in this calculation is assumed to be normally distributed. From this, the mean α_m value for the intron sequence is 4.1, with a confidence interval (CI) of 3.1 to 5.1. Thus, introns show the male substitution rate to be significantly greater than the female substitution rate ($p < 0.01$; also see branch lengths in Fig. 2), and the confidence interval for α_m covers a range of values that includes the one predicted (4.4) by the RDH from estimates of germline replications.

Our data also include W-linked and Z-linked sequence information from two exon regions. Exons are generally more constrained than introns, in part because nonsynonymous substitutions lead to changes in the amino acid composition of the gene's protein product. Thus, exon substitutions were categorized as nonsynonymous (replacement) or synonymous (silent). Whereas

nonsynonymous substitutions most often reflect levels of selective constraint, synonymous substitutions have often been considered to be selectively "neutral" and thus equivalent in nature to intron sequence (Kimura 1983). However, evidence that the synonymous rate varies widely (Bernardi et al. 1993; Wolfe and Sharp 1993; Li 1997), together with evidence that the synonymous rate varies in correlation with the nonsynonymous rate and in correlation to the degree of codon bias (Sharp and Li 1987; Caccio et al. 1995; Zoubak et al. 1995; Alvarez-Valin et al. 1998) all argue against complete neutrality and indicate that synonymous and nonsynonymous rates, as well as codon usage, may be under some common selective constraints.

With these caveats in mind, it appears, based on nonsynonymous substitutions, that the CHD W exon is evolving more rapidly than the CHD Z exon. The W exon of the CHD gene in this region has a significantly greater rate of nonsynonymous substitution (Ka) than that of the Z exon ($p \leq 0.005$ for most pairwise comparisons, and $p \leq 0.05$ in all cases), as determined using the *t* test method of Hughes and Nei (1988) implemented in FENS pre-release vβ1.2 (de Koning et al. 1999); the approximately 3.3 times greater rate is readily seen in the comparison of W-Ka versus Z-Ka (Table 2). Whereas the markedly higher number of nonsynonymous site substitutions (Fig. 1.B, bold-italicized bases) in W-exons is as high as 5 W compared to 1 Z in some cases, there is an absence of statistical support for positive selection ($p > 0.05$). Thus, we speculate that this represents relaxed selective constraint on the W exons.

Because intragenic synonymous and nonsynonymous rates of some genes are positively correlated (Alvarez-Valin et al. 1998), the possibility exists that there is also relaxed constraint on synonymous sites of the W exon and that they are reaching a new equilibrium between regional mutational bias (the condition presumably reflected by the introns) and relaxed functional constraint (cf. Li 1997). While it appears that selective constraints at work on the W exons of CHD are different than those on Z exons, there is no evidence yet to indicate what specific roles the respective W- and Z-linked genes play in the biology of birds, so this must be investigated further. Because the measurement of rates (and thus calculation of α_m) in these exon regions is complicated by the effects of different constraints, and by the possibility of a coupled rate increase in W-nonsynonymous and W-synonymous sites, we have excluded the exon data from further consideration as a measure of α_m.

Comparisons to Other Studies

For these data, α_m (4.1) is approximately equal to *c* (4.4). In studies of primates and rodents, values of α_m have also been found to be close to the estimates of *c* ($c = 6.2$ in humans, $c = 2.0$ in rodents [Chang et al. 1994; Li et al. 1996]). The correlation between α_m and *c* from observations in such a wide variety of taxa suggests that errors made during germline replication are a primary source of mutation and that this mechanism exists throughout mammalian and avian orders. Ellegren and Fridolfsson (1997) found, within Neoaves, an α_m (6.5) that is higher than the value predicted by the estimate of *c* for Eoaves (4.4). Overall, Eoaves are larger, longer lived, and have a longer period of reproduction than Neoaves. Passerines like those used in Ellegren and Fridolfsson's (1997) study have an average age of reproduction of 2 years, with sexual maturity occurring at 1.5 years old (Ricklefs 1972; Sternberg 1989; Potti and Montalvo 1991). Most of these short-lived birds will breed only once in their lifetimes (Sternberg 1989). These factors would lead to a prediction that α_m should be lower, not higher, in those passerines than in the Eoaves included in our study. However, there is evidence that collared flycatchers (*Ficedula albicollis*) engage in extra-pair matings (Sheldon and Ellegren 1996); if high sperm production/output results from sperm competition in these birds, it would increase the predicted magnitude of *c*. Tuttle et al. (1996) have measured very large sperm counts and sperm production in fairy-wrens (Aves: Maluridae), which they attribute to sperm competition in these birds; the frequency of extra-pair matings is higher in fairy-wrens than in any other socially monogamous bird species (Tuttle et al.1996). More information about the reproductive biology of passerines may bring this analysis into better perspective. Possible future studies might examine birds with sperm competition for a specific correlation of faster substitution rates with higher sperm production and measured sperm counts.

Male-Driven Evolution?

Because the intron value for $\alpha_m \approx 4.1$, with CI = 3.1 to 5.1, is significantly larger than unity ($p < 0.01$), we conclude that this data set supports the idea that sex-specific asymmetrical number of germline replications has a marked effect on mutation rate in waterfowl and galliforms. It is worth noting that while the observed α_m value is slightly less than the predicted value of $c = 4.4$, this is one of the stronger correlations found between the observed ratio of male-to-female mutations and the predicted ratio of male-to-female germline replications. It can be concluded that a greater rate of substitution occurs in males, regardless of whether males are heterogametic (mammals) or homogametic (birds). These findings suggest that the faster rate found by others for the mammalian Y chromosome is related to number of replications and is not an artifact of the unusual nature of the Y. Although other mechanisms of DNA mutation cannot be ruled out, this data set indicates that the force of male-driven evolution is one of the primary influences at work in the vertebrate genome.

Acknowledgments. We wish to thank Bill Payne, Mike Monahan, Phil Danielson, Jim Hagler, Walter Messier, Caro-Beth Stewart, and Wen-Hsiung Li for their help during the preparation of this manuscript. Funding for this study was provided by NSF grant DEB9629462 to TWQ.

References

Alvarez-Valin F, Jabbari K, Bernardi G (1998) Synonymous and nonsynonymous substitutions in mammalian genes: intragenic correlations. J Mol Evol 46:37–44

Bellrose FC (1976) Ducks, geese, and swans of North America. Wildlife Management Institute, Stackpole Books, Harrisburg, PA

Bernardi G, Mouchiroud D, Gautier C (1993) Silent substitutions in mammalian genomes and their evolutionary implications. J Mol Evol 37:583–589

Brown WM, Prager EM, Wang A, Wilson AC (1982) Mitochondrial DNA sequences of primates: Tempo and mode of evolution. J Mol Evol 18:225–239

Britten RJ (1986) Rates of DNA sequence evolution differ between taxonomic groups. Science 231:1393–1398

Caccio S, Zoubak S, D'Onofrio G, Bernardi G (1995) Nonrandom frequency patterns of synonymous substitutions in homologous mammalian genes. J Mol Evol 40:280–292

Chang BH-J, Li W-H (1995) Estimating the intensity of male-driven evolution in rodents by using X-linked and Y-linked *Ube1* genes and pseudogenes. J Mol Evol 40:70–77

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

Clulow J, Jones RC (1982) Production, transport, maturation, storage and survival of spermatozoa in the male Japanese quail, *Coturnix coturnix*. J Reprod Fertil 64:259–266

Cooke F, Rockwell RF, Lank DB (1995) The snow geese of La Perouse Bay: Natural selection in the wild. Oxford University Press, New York, NY

Crow JF (1997) Molecular evolution—who is in the driver's seat? Nat Genet 17:129–130

de Koning J, Palumbo M, Messier W, Stewart CB (1999) FENS: facilitated estimates of nucleotide substitutions, pre-release version β1.2. University of Albany, SUNY, Albany, NY

Delmas V, Stokes DG, Perry RP (1993) A mammalian DNA-binding protein that contains a chromo-domain and an SNF2/SW12-like helicase domain. Proc Natl Acad Sci USA 90:2414–2418

Desjardins P, Morais R (1990) Sequence and gene organization of the chicken mitochondrial genome. A novel gene order in higher vertebrates. J Mol Biol 212:599–634

Drake JW, Baltz RH (1976) The biochemistry of mutagenesis. Ann Rev Biochem 45:11–37

Easteal S, Collet C, Betty D (1995) The mammalian molecular clock. R.G. Landes Co., Austin, TX

Ellegren H (1996). First gene on the avian W chromosome (CHD) provides a tag for universal sexing of non-ratite birds. Proc R Soc Lond B 263:1251–1256

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

Eng RL (1963) Observations on the breeding biology of male sage grouse. J Wildl Manage 27:841–846

Fridolfsson A-K, Cheng H, Copeland NG, Jenkins NA, Liu H-C, Raudsepp T, Woodage T, Chowdhary B, Halverson J, Ellegren H (1998) Evolution of the avian sex chromosomes from an ancestral pair of autosomes. Proc Natl Acad Sci USA 95:8147–8152

Froman DP (1994) Biology of semen production and ejaculation. In: Bakst MR, Wishart GJ (eds) Proceedings First International Symposium on the Artificial Insemination of Poultry. The Poultry Science Association, Inc., Savoy, IL, pp 21–38

Funahashi J-I, Sekido R, Murai K, Kamachi Y, Kondoh H (1993) δ-Crystallin enhancer binding protein δEF1 is a zinc-finger homeodomain protein implicated in gastrulation embryogenesis. Development 119:433–446

Gilbert AB (1979) Female genital organs. In: King AS, McLelland J (eds) Form and function in birds, vol. 1. Academic Press, New York, NY, pp 237–360

Gill FB (1995) The annual cycles of birds. In: Ornithology, 2d ed. W.H. Freeman and Co., New York, NY, pp 263–285

Goodman M, Koop BF, Czelusniak J, Weiss ML, Slightom JL (1984) The η-globin gene: Its long history in the β-globin gene family of mammals. J Mol Biol 180:803–823

Griffiths R, Korn RM (1997) A CHD1 gene is Z chromosome linked in the chicken *Gallus domesticus*. Gene 197:225–229

Griffiths R, Daan S, Dijkstra C (1996) Sex identification using two CHD genes. Proc R Soc Lond B 263:1251–1256

Hayashida H, Kuma K, Miyata T (1992) Interchromosomal gene conversion as a possible mechanism for explaining divergence patterns of ZFY-related genes. J Mol Evol 35:181–183

Hofmann MA, Brian DA (1991) Sequencing PCR DNA directly from a bacterial colony. Biotechniques 11:30–31

Hughes GC (1963) The population of germ cells in the developing female chick. J Embryol Exp Morph 11:513–536

Hughes AL, Nei M (1988) Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature 335:167–170

Johnsgard PA (1965) Handbook of waterfowl behaviour. Cornell University Press, New York, NY

Kahn NW, John JS, Quinn TW (1998) Sex identification in birds using an intron from CHD. Auk 115:1074–1078

Kimura M (1983) The neutral theory of evolution. Cambridge University Press, Cambridge

Kirby JD, Froman DP (1995) Reproduction in the male bird. In: Whittow GC (ed) Sturkie's avian physiology. Academic Press, Orlando, FL

Kohne DE (1970) Evolution of higher-organism DNA. Rev Biophys 33:1–48

Kohne DE, Chiscon JA, Hoyer BH (1972) Evolution of primate DNA sequences. J Hum Evol 1:627–644

Kuwana T (1993) Migration of avian primordial germ cells toward the gonadal anlage. Develop Growth Differ 35:237–243

Lake PE (1981) Male genital organs. In: King AS, McLelland J (eds) Form and function in birds, vol. 2. Academic Press, New York, NY, pp 1–61

Lanfear J, Holland PWH (1991) The molecular evolution of ZFY-related genes in birds and mammals. J Mol Evol 32:310–315

Li W-H (1993) Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J Mol Evol 36:96–99

Li W-H (1997) Rates and patterns of nucleotide substitution; Genome organization and evolution. In: Molecular evolution. Sinauer Associates, Sunderland, MA, pp 196–202; 379–418

Li W-H, Ellsworth DL, Krushkal J, Chang BH-J, Hewett-Emmett D (1996) Rates of nucleotide substitution in primates and rodents and the generation-time effect hypothesis. Mol Phylogenet Evol 5:182–187

Lin M, Jones RC (1992) Renewal and proliferation of spermatogonia during spermatogenesis in the Japanese quail, *Coturnix coturnix japonica*. Cell Tissue Res 267:591–601

Lin M, Jones RC, Blackshaw AW (1990) The cycle of the seminiferous epithelium in the Japanese quail (*Coturnix coturnix japonica*) and estimation of its duration. J Reprod Fertil 88:481–490

Martin AP, Palumbi SR (1993) Body size, metabolic rate, generation time, and the molecular clock. Proc Natl Acad Sci USA 90:4087–4091

Martin AP, Naylor GJP, Palumbi SR (1992) Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. Nature 357:153–155

Mendez-Herrera C, Hofmann P, Pedernera E (1993) Effect of 17β-

estradiol on somatic and germ cell subpopulations in the ovary of the newly hatched chick. Gen Comp Endocrinol 89:182–188

Meyer DB (1964) The migration of primordial germ cells in the chick embryo. Develop Biol 10:154–190

Mindell DP (1991) Phylogenetic analysis of DNA sequences. In: Miyamoto M, Cracraft J (eds) Oxford University Press, New York, NY

Mindell DP, Knight A, Baer C, Huddleston CJ (1996) Slow rates of molecular evolution in birds and the metabolic rate and body temperature hypothesis. Mol Biol Evol 13:422–426

Mindell DP, Sorenson MD, Huddleston CJ, Miranda HC Jr, Knight A, Sawchuk SJ, Yuri T (1997) Phylogenetic relationships among and within select avian orders based on mitochondrial DNA. In: Avian molecular evolution and systematics. Academic Press, San Diego, CA, pp 213–247

Miyata T, Hayashida H, Kuma K, Yasunaga T (1987a) Male-driven molecular evolution demonstrated by different rates of silent substitutions between autosome- and sex chromosome–linked genes. Proc Japan Acad Ser B 63:327–331

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

Monroe BL, Sibley CG (1993) A world checklist of birds. Yale University Press, New Haven, CT

Ogawa A, Murata K, Mizuno S (1998) The location of Z- and W-linked marker genes and sequence on the homomorphic sex chromosomes of the ostrich and the emu. Proc Natl Acad Sci USA 95:4415–4418

Pamilo P, Bianchi NO (1993) Evolution of the ZFX and ZFY genes: rates and interdependence between the genes. Mol Biol Evol 10: 271–281

Peters JL (1986) Check-list of the birds of the world. Mayr E, Greenway JC, Paynter RA Jr (eds) Harvard University Press and Museum of Comparative Zoology, Cambridge, MA

Potti J, Montalvo S (1991) Return rate, age at first breeding and natal dispersal of pied flycatchers *Ficedula hypoleuca* in Central Spain. Ardea 79:419–428

Quinn TW (1992) The genetic legacy of Mother Goose—phylogeographic patterns of lesser snow goose *Chen caerulescens caerulescens* maternal lineages. Mol Ecol 1:105–117

Quinn TW, White BN (1987) Identification of restriction-fragment-length polymorphisms in genomic DNA of the lesser snow goose (*Anser caerulescens caerulescens*). Mol Biol Evol 4:126–143

Rand DM (1994) Thermal habit, metabolic rate and the evolution of mitochondrial DNA. Trends Ecol Evol 9:125–131

Ricklefs RE (1972) Fecundity, mortality, and avian demography. In: Farner DS (ed) Breeding biology of birds. National Academy of Sciences, Washington DC, pp 366–435

Roosen-Runge EC (1977) Vertebrata. In: The process of spermatogenesis in animals. Cambridge University Press, New York, NY, pp 102–132

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2d ed. Cold Spring Harbor Laboratory Press, New York, NY

Sharp PM, Li W-H (1987) The rate of synonymous substitution in enterobacterial genes is inversely related to codon usage bias. Mol Biol Evol 4:222–230

Sheldon BC, Ellegren H (1996) Offspring sex and paternity in the collared flycatcher. Proc R Soc Lond B 263:1017–1021

Shields GF, Wilson AC (1987) Calibration of mitochondrial DNA evolution in geese. J Mol Evol 24:212–217

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

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

Sibley GC, Ahlquist JE, Monroe BL Jr (1988) A classification of the living birds of the world based on DNA–DNA hybridization studies. Auk 105:409–423

Sternberg H (1989) Part I. Short-lived hole nesters, pied flycatcher. In: Newton I (ed) Lifetime reproduction in birds. Academic Press, London, pp 55–74

Stokes DG, Tartof KD, Perry RP (1996) CHD1 is concentrated in interbands and puffed regions of *Drosophila polytene* chromosomes. Proc Natl Acad Sci USA 93:7137–7142

Sudhir K, Tamara K, Nei M (1993) MEGA: molecular evolutionary genetic analysis, version 1.0. Pennsylvania State University, University Park, PA

Swift CH (1914) Origin and early history of the primordial germ cells in the chick. Am J Anat 15:483–516

Swofford DL (1993) PAUP: phylogenetic analysis using parsimony, version 3.1.1. Illinois Natural History Survey, Champaign, IL

Tajima F, Nei M (1984) Estimation of the evolutionary distance between nucleotide sequences. Mol Biol Evol 1:269–285

Topal MD, Fresco JR (1976) Complementary base pairing and the origin of substitution mutations. Nature 263:285–289

Tuttle EM, Pruett-Jones S, Webster MS (1996) Cloacal protuberances and extreme sperm production in Australian fairy-wrens. Proc R Soc Lond B 263:1359–1364

Van Wagner CE, Baker AJ (1990) Association between mitochondrial DNA and morphological evolution in Canada geese. J Mol Evol 31:373–382

Vogel F, Rathenberg R (1975) Spontaneous mutations in man. In: Advances in human genetics, vol. 5. Plenum Press, New York, NY

Vogel F, Kopun M, Rathenberg R (1976) Mutation and molecular evolution. In: Goodman M, Tashian RE, Tashian JH (eds) Molecular anthropology. Plenum Press, New York, NY, pp 13–33

Wetmore A (1960) A classification of birds of the world. Smithson Misc Collect 139:1–37

Wilson AC, Carlson SS, White TJ (1977) Biochemical evolution. Ann Rev Biochem 46:573–639

Wolfe KH, Sharp PM (1993) Mammalian gene evolution: nucleotide sequence divergence between mouse and rat. J Mol Evol 37:441–456

Wu C-I, Li W-H (1985) Evidence for higher rates of nucleotide substitution in rodents than in man. Proc Natl Acad Sci USA 82:1741–1745

Zimmer R, Erdtmann B, Thomas WK, Quinn TW (1994) Phylogenetic analysis of the *Coscoroba coscoroba* using mitochondrial srRNA gene sequences. Mol Phylogenet Evol 3:85–91

Zoubak S, D'Onofrio G, Caccio S, Bernardi G (1995) Specific compositional patterns of synonymous positions in homologous mammalian genes. J Mol Evol 40:293–307

Zuckerkandl E, Pauling L (1965) Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ (eds) Evolving genes and proteins. Academic Press, London, pp 97–166