Evolutionary Strata on the Chicken Z Chromosome: Implications for Sex Chromosome Evolution

Lori-Jayne Lawson Handley,1 Helene Ceplitis and Hans Ellegren2

Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, SE-752 36 Uppsala, Sweden

Manuscript received May 20, 2003
Accepted for publication July 13, 2003

ABSTRACT

The human X chromosome exhibits four “evolutionary strata,” interpreted to represent distinct steps in the process whereby recombination became arrested between the proto X and proto Y. To test if this is a general feature of sex chromosome evolution, we studied the Z-W sex chromosomes of birds, which have female rather than male heterogamety and evolved from a different autosome pair than the mammalian X and Y. Here we analyze all five known gametologous Z-W gene pairs to investigate the “strata” hypothesis in birds. Comparisons of the rates of synonymous substitution and intronic divergence between Z and W gametologs reveal the presence of at least two evolutionary strata spread over the p and q arms of the chicken Z chromosome. A phylogenetic analysis of intronic sequence data from different avian lineages indicates that Z-W recombination ceased in the oldest stratum (on Zq; CHD1Z, HINTZ, and SPINZ) 102–170 million years ago (MYA), before the split of the Neoaves and Eoaves. However, recombination continued in the second stratum (on Zp; UBAP2Z and ATP5A1Z) until after the divergence of extant avian orders, with Z and W diverging 58–85 MYA. Our data suggest that progressive and stepwise cessation of recombination is a general feature behind sex chromosome evolution.

Sex chromosomes are found in such diverse groups as mammals, birds, fish, insects, and plants and have probably evolved many times in different lineages. Despite this, empirical data indicate that many parallels can be drawn from sex chromosome evolution in different taxa (Ohno 1967; Charlesworth and Charlesworth 2000; Graves and Shetty 2001; Lahn et al. 2001). Sex chromosome evolution from an autosomal pair has been suggested to be initiated by selection for linkage between sexually antagonistic alleles and a sex-determining locus, being favored in one sex but not in the other. Selective pressure to maintain a favorable genotype in one sex should lead to the creation of “cold spots” where recombination between the two proto-sex chromosomes is reduced (Rice 1996). Following cessation of recombination, degeneration of the hemizygous sex chromosome may proceed by any of the evolutionary forces associated with clonal inheritance, for example, Muller’s ratchet, background selection, the Hill-Robertson effect with weak selection, and the hitchhiking of deleterious alleles to advantageous mutations (reviewed in Charlesworth and Charlesworth 2000).

Cessation of recombination is therefore a critical step in sex chromosome evolution since differentiation of the proto-sex chromosomes can begin only once recombination between them has ceased. By comparing relic genes shared between homologous sex chromosomes (“gametologs”; Garcia Moreno and Mindell 2000; Ellegren 2002), insight can be gained into when and how recombination ceased between them, assuming they have been evolving independently since recombination was disrupted. For instance, Lahn and Page (1999) compared a number of gametologous gene pairs shared between the human X and Y chromosomes and found their divergence to be highly correlated to their location on the X chromosome. The gene pairs clustered into four distinct groups, with decreasing synonymous substitution rates ($K_s$) from the distal long arm of X (Xq) to the distal short arm (Xp). Since the accumulation of synonymous substitutions can be assumed to be linear with time, this suggests that the first group on Xq forms the oldest evolutionary “stratum” and group 4 on Xp forms the youngest. From these data, Lahn and Page (1999) hypothesize that sex chromosome evolution in the human lineage was punctuated by at least four disruptive events (e.g., inversion mutations on the Y chromosome), which suppressed in turn the recombination between X and Y in each stratum.

While sex chromosome evolution appears to be a stepwise process in the human lineage, whether this holds true in a wider taxonomic perspective is poorly understood. So far, no evidence suggests that such clearly defined strata are present on the X chromosomes of other mammals, and the situation may be different...

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under accession nos. AY188782–AY188763, AY189754–AY189777, and AY194125–AY194149.

1Present address: Institute of Ecology, Biology Bldg., University of Lausanne, CH-1015 Lausanne, Switzerland.

2Corresponding author: Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Norbyvagen 18D, SE-752 36 Uppsala, Sweden. E-mail hans.ellegren@ebc.uu.se

in mice, for example, where gene order of the X is different from that in primates (Searle et al. 1989; Lyon 1994). Birds provide an ideal model to test the ubiquity of such a hypothesis since they have female rather than male heterogamety (males ZZ, females ZW), and the Z and W sex chromosomes evolved from a different autosomal pair from X and Y (Fredolfsen et al. 1998). Despite this, however, the mammalian and avian sex chromosomes are similar in organization and structure, and this may be suggestive of a common mechanism driving sex chromosome evolution. The X and the Z chromosomes are large, gene rich, and highly conserved between mammalian and avian lineages, respectively [the Z constitutes ~7–10% of the total genome (Bloom et al. 1993; Schmid et al. 2000)—this is a slightly larger genomic proportion compared to mammalian X]. Similarly the W chromosome, like the Y, is generally small, heterochromatin rich, and gene poor, and since most of the W does not recombine it should be subjected to a degeneration similar to that of the Y chromosome (Ellegren 2000, 2001). The chicken W, for example, represents only 1.5% of the total genome and ~65% is composed of repeat-sequence families (Clint and Haines 1999).

To study the process of avian sex chromosome evolution we here investigate the degree of divergence between gametologous gene pairs shared between the chicken Z and W chromosomes. We also perform a phylogenetic analysis to study the process of divergence of these gene pairs relative to ordinal divergence among birds. These combined analyses indicate that at least two evolutionary strata are present on the chicken Z chromosome.

MATERIALS AND METHODS

Locus information: We analyzed all five gametologous gene pairs so far characterized on the avian sex chromosomes (Figure 1). The genes are as follows, using the standardized nomenclature for avian gametologists as suggested by Ellegren (2002): CHD1Z/CHD1W (Ellegren 1996; Griffiths and Korn 1996), HINTZ/HINTW (chicken HINTZ was referred to as chPKCI, and HINTW as WPKCI, by Hori et al. 2000, and HINTW as ASW, by O’Neill et al. 2000; see Cepilis and Ellegren 2004 for the nomenclature of this gene), SPINZ/SPINW (Itoh et al. 2001; the original report referred to these as chSpin-Z and chSpin-W), UBAP2Z/UBAP2W (our unpublished observations), and ATP5A1Z/ATP5A1W (Dvorak et al. 1992; Fredolfsen et al. 1998). CHD1Z, HINTZ, and SPINZ are situated on Zq, with HINTZ situated in close proximity to SPINZ (both at Zq1.1; Schmid et al. 2000; Figure 1). UBAP2Z and ATP5A1Z are both near the telomere on Zp (Figure 1). It should be noted that HINTZ and SPINZ have been mapped to the Z chromosome by only fluorescence in situ hybridization (FISH), and their exact orientation has not been confirmed by linkage mapping. The gene order on the W chromosome is unknown and only the positions of HINTW and SPINW have been assigned by FISH mapping (to Wp14; Schmid et al. 2000). The positions of CHD1W and ATP5A1W have not yet been resolved although they are also provisionally assigned to Wp14 given the heterochromatic nature of large parts of Wq (Schmid et al. 2000). HINTW is present in ~40 tandemly repeated copies in chicken although it has not been established whether all of these copies are functional (Hori et al. 2000).

Coding region analysis: Coding region sequences were analyzed from chicken only, because avian mapping work has focused on this species. cDNA sequences were taken from GenBank as follows: CHD1Z, AF004397; CHD1W, AF181826; HINTZ, AB026675; HINTW, AB026677; SPINZ, nucleotides 28–816 in AB047853; SPINW, nucleotides 50–838 in AB04-7852; UBAP2Z, AF188762; UBAP2W, AF188763; ATP5A1Z, AF301567 and AF301579; and finally ATP5A1W, AF301566 and AF301578. To estimate the Z-W divergence in coding regions, synonymous substitution rates (Ks) were calculated using the Nei-Gojobori method (Nei and Gojobori 1989), with Jukes-Cantor correction and excluding gaps in MEGA2 (Kumar et al. 2000). Confidence intervals for Ks were calculated from estimates of standard error obtained after 1000 bootstrap replicates.

Intron analysis: DNA samples: DNA was extracted from avian tissue or blood using a standard extraction procedure (Hoelzel and Green 1998). We performed phylogenetic analyses of introns from species from four avian orders that are representative of the two infraclasses within Aves (Sibley and Ahlquist 1990). From the infraclass Neoaves we sampled two Galliformes (chicken, Gallus gallus, sample code Gga and turkey, Meleagris gallopavo, Mga) and six Anseriformes (Canada goose, Branta canadensis, Bca; goldeneye, Bucephala clangula, Bck; barnacle goose, Branta leucopsis, Bln; long-tailed duck, Clangula hyemalis, Chl; trumpeter swan, Cygnus buccinator, Cbu; and eider, Somateria mollissima, Smo). From the Infraclass Eoaves, we analyzed two Passeriformes (blue tit, Parus caeruleus, Pca; and great tit, Parus major, Pma) and two Strigiformes (eagle owl, Bubo bubo, Bbu; and Tengmalm’s owl, Aegolius funereus, Afu).

Amplification of UBAP2Z and UBAP2W: We PCR amplified the first two introns of both UBAP2Z and UBAP2W, using the primers UBAP2-1F (5’-TGA GTG TTA CCA GAC TGT CT) and UBAP2-1R (5’-GCA GGC AGA AGT CCT CCA GG) and UBAP2-2F (5’-AAA GGT CCT CCA AAC GTG CG) and UBAP2-2R (5’-CTG CAT TGA AGC ATC TGG AG). Reactions contained 0.2 µM of each primer, 0.2 mM dNTPs (Pharmacia, Piscataway, NJ), 1 × AmpliTaq Gold PCR buffer, 2 mM MgCl2, and 1.25 units of AmpliTaq Gold (Applied Biosystems, Foster City, CA), with ~100 ng template DNA, made up to 50 µl with water. Amplification conditions were 95° for 5 min; followed by 30 cycles of 95° for 30 sec, 60° for 40 sec, and 72° for 1 min; and a final extension step for 7 min at 72°.

Amplification of SPINZ and SPINW: The exon-intron structure of SPINZ/SPINW was inferred by comparing the published Z-linked chicken cDNA sequence (AB047853) to the genome sequence of human SPIN (NT_029395.13). We identified three potential intron sites in the avian genes and these were confirmed by PCR using flanking exon primers. Estimated locations of the introns are at positions 128, 382, and 614 in the SPINZ chicken cDNA sequence. Intron 2 was amplified with the primers SPIN319F (5’-TAT GGA CTA GAA CTG GAC AAA G) and SPIN472R (5’-AGA CCA TCC CCC TCC ATT CAT C), using the PCR conditions described above.

Amplification of HINTZ and HINTW: Intron 2 of HINTW was amplified using the primers HINTW-F1 (5’-TGG CTG GTC CTT CAT GAT A) and HINTW-R1 (5’-GTA AAA CCC ATT CCG TGG C), whereas HINTZ was amplified with HINTZ-F1 (5’-TGC CTT GGC TTC CAT ATC) and HINTZ-R1 (5’-CAA ATC TAT TGT CTA GTG AT). PCR conditions were the same as those described for UBAP2Z/W, but with an annealing temperature of 52° for HINTW and 54° for HINTZ, and an MgCl2 concentration of 2.5 mM. Intron sequences for chicken were taken from GenBank (AB026676 and AB026678).
Cloning and DNA sequencing: For UBA2PZ/UBAP2W and SPINZ/SPINW introns, “heterozygous” PCR products from females were cloned to separate the Z and W copies, and Z-linked sequences were verified by amplification and direct sequencing from males. In addition, HINTW products were cloned to separate the different W-linked copies. PCR products were purified using the Qiagen PCR purification kit (Qiagen, Valencia, CA), ligated into pGEM-T Easy Vector (Promega, Madison, WI), and transformed into JM109 competent cells (Promega) following the manufacturer’s guidelines. Inserts were subsequently amplified using the modified M13 primers OMNI (5'-ACA GGA AAC AGC TAT GAC CAT GAT) and UNI (5'-AGA GTT TAA AAA GCG AGG CCA GT) with the same reagent concentrations as described for UBA2P. PCR conditions were 95°C for 5 min; followed by 35 cycles of 95°C for 40 sec, 50°C for 30 sec, and 72°C for 1 min; and a subsequent hold of 72°C for 10 min. PCR products were purified and sequenced directly using the BigDye Terminator cycle sequencing kit (Applied Biosystems). Reactions were electrophoresed on an ABI 377 automated sequencer and sequences were edited using AutoAssembler 2.1 software (both from Applied Biosystems).

Sequence analysis: Intron sequences from CHDI and ATP5A1 were taken from GenBank as follows: chicken CHDI/CHDIW intron E, AFS26055–AF526056; and chicken, turkey, eider, and goldeneye ATP5A1/ATP5A1W intron 3, AF165968–AF165975. Sequences were aligned using CLUSTALW (Thompson et al. 1994), with gap opening penalties of 10 and gap extension penalties of 0.5 (alignments are available from the authors upon request). In all our analyses (divergence estimates and phylogenetic), indels were removed. Nucleotide divergence between Z and W introns in the chicken was calculated with Jukes-Cantor correction, and 95% confidence intervals were obtained from standard errors after 1000 bootstrap replicates. For the multicopy gene HINTW, we obtained very similar results when different copies were compared to the gametologous Z sequence; hence we present the results of just one comparison here. A detailed analysis of the evolution of multiple copies of HINTW is to be presented elsewhere (our unpublished results). All intron sequences generated in this study have been submitted to the GenBank database with the following accession numbers: UBA2PZ/UBAP2W intron 1, AY189754–AY189766; UBA2PZ/UBAP2W intron 2, AY189767–AY189777; HINTZ/HINTW intron 2, AY194143–AY194149; and SPINZ/SPINW intron 2, AY194125–AY194143.

Estimating Z/W divergence times: We estimated the minimum divergence time of SPINZ and SPINW, HINTZ and HINTW, UBA2PZ and UBAP2W, and ATP5A1Z and ATP5A1W, using a maximum-likelihood quartet method that incorporates independently derived divergence dates as calibration points (Rambaut and Bromham 1998; implemented in “Qdate” and available from http://evolve.zoo.ox.ac.uk/software/QDate/main.html). The same approach has previously been used to estimate a divergence between CHDI and CHDIW (Garcia Moreno and Mindell 2000) and is therefore not repeated here. The HKY85 model of nucleotide substitution (Hasegawa et al. 1985) was used for all introns, with correction for transition/transversion ratio and for gamma distribution rate heterogeneity after verification using Modeltest v3.06 (Posada and Crandall 1998). Pairwise species divergences used for calibration were obtained from fossil and molecular data (Table 3); however, these calibration dates should be considered with caution since they are very approximate. For example, the estimated divergence between chicken and turkey based on DNA-DNA hybridization data ranges between 20 and 40 million years ago (Mya; Sibley and Ahlquist 1990). The rates of nucleotide substitution on the Z and W chromosomes are expected to be different due to differences in sex-specific mutation rates and effective population size. We therefore calculated log-likelihood values for the constrained one-rate and two-rate substitution models and used a χ² test to compare the divergence in log-likelihood between them and the unconstrained five-rate model (Rambaut and Bromham 1998). In all comparisons, the log-likelihood values of the one-rate model were significantly different from the five-rate model, which vindicates using a two-rate model in the analysis, in which one pair in the quartet (i.e., corresponding to the two Z-linked sequences) has a different rate from the other (the W-linked). The use of the two-rate model also makes intuitive sense since we expect the Z and W sequences to be evolving at different rates due to relative effective population sizes and male-biased mutation rates. Where the log-likelihood values differ significantly from those of the five-rate model (in other words indicating rate differences other than those due to Z or W linkage) is indicated in Table 3.

RESULTS

Divergence of avian Z/W gene pairs: Intrinsic nucleotide divergence was estimated for all five known gametologous Z-W gene pairs from chicken (Table 1). There is a clear correlation between divergence and the Z chromosome position (Figure 1), ranging from 0.47 between CHD1Z (which is distal on Zq) and CHD1W to 0.23–0.25 between ATP5A1Z (distal on Zp) and ATP5A1W and UBAP2Z and UBAP2W. Since divergence

TABLE 1

Intron divergence in gametologous gene pairs shared between the chicken Z and W chromosomes

<table>
<thead>
<tr>
<th>Gene pair</th>
<th>Intron</th>
<th>Lengtha</th>
<th>Divergenceb</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP5A1Z/ATP5A1W</td>
<td>3</td>
<td>769 (838)</td>
<td>0.249</td>
<td>0.225–0.275</td>
</tr>
<tr>
<td>UBA2PZ/UBAP2W</td>
<td>1 and 2</td>
<td>1570 (2113)</td>
<td>0.229</td>
<td>0.211–0.247</td>
</tr>
<tr>
<td>SPINZ/SPINW</td>
<td>2</td>
<td>612 (804)</td>
<td>0.395</td>
<td>0.355–0.437</td>
</tr>
<tr>
<td>HINTZ/HINTW</td>
<td>2</td>
<td>701 (892)</td>
<td>0.429</td>
<td>0.387–0.471</td>
</tr>
<tr>
<td>CHDI1Z/CHDI1W</td>
<td>“E”</td>
<td>384 (466)</td>
<td>0.470</td>
<td>0.409–0.531</td>
</tr>
</tbody>
</table>

a Alignment length in base pairs excluding indels, with total length in parentheses.
b Nucleotide divergence (calculated with indels excluded) with Jukes-Cantor correction. The 95% confidence intervals were calculated after standard errors obtained after 1000 bootstrap replicates.
Figure 1.—Physical location of gametologous gene pairs shared between the Z (GgaZ) and W (GgaW) chromosomes of chicken. G-banded ideograms are adapted from Schmid et al. (2000). No asterisk, FISH-mapped only; *, linkage mapped on Z; **, assigned by linkage and FISH mapping. Parentheses refer to W-linked genes whose exact position has yet to be defined. The pseudoautosomal region, not indicated here, is at Zpter.

should be approximately linear with time, this indicates that Z and W have diverged in a progressive fashion, starting with the long arm of the Z and proceeding to the short arm. Overlapping confidence intervals between divergence estimates of CHD1Z/CHD1W [95% confidence interval (C.I.), 0.409–0.581], HINTZ/HINTW (0.387–0.471), and SPINZ/SPINW (0.353–0.437) suggest the presence of an evolutionary stratum (group 1, the oldest group) defined as CHD1Z + HINTZ + SPINZ. The much lower divergence estimates obtained for UBAP2Z/UBAP2W (95% C.I., 0.211–0.247) and ATP5A1Z/ATP5A1W (0.223–0.275) indicate a second evolutionary stratum (group 2, the youngest group) consisting of ATP5A1Z + UBAP2Z. Confidence intervals do not overlap between groups.

A pattern similar to that found with intronic divergence is expected from comparison of Ks and we again found strong correlation between Ks and Z position (Table 2). However, estimates of Ks were lower than those for intronic divergence for all Z-W pairs (Table 2) and divergence for SPINZ/SPINW (0.187) was more similar to that of UBAP2Z/UBAP2W (0.170) than to that of HINTZ/HINTW (0.321). It should be noted, though, that the confidence intervals associated with Ks are large and overlap considerably between genes, so the definition of the boundary between the two strata is less evident from Ks than from the estimates based on intronic divergence. Because of the ambiguity in the placement of SPINZ/SPINW based on coding and noncoding sequence analysis, the distribution of Ks in the SPINZ/SPINW alignment was investigated further. We dissected SPINZ/SPINW into its respective exons and analyzed each one separately. Ks was found to be similar in exons 1 and 3 (exon 1, $K_s = 0.126 \pm 0.087$ SE; exon 3, $0.128 \pm 0.060$) and in exons 2 and 4 (exon 2, $0.264 \pm 0.081$; exon 4, $0.200 \pm 0.077$). The estimates for exons 1 and 3 are similar to those for ATP5A1Z/ATP5A1W + UBAP2Z/UBAP2W, whereas estimates for exons 2 and 4 are intermediate between ATP5A1Z/ATP5A1W + UBAP2Z/UBAP2W and HINTZ/HINTW.

**Phylogenetic analysis of Z-W introns:** The divergence of Z- and W-linked copies of SPINZ/SPINW, HINTZ/HINTW, and UBAP2Z/UBAP2W, relative to ordinal divergence, was investigated using a phylogenetic ap-

---

**TABLE 2**

Synonymous substitution rates ($K_s$) for gametologous gene pairs shared between the chicken Z and W chromosomes

<table>
<thead>
<tr>
<th>Gene pair</th>
<th>Exons</th>
<th>Length*</th>
<th>$K_s$</th>
<th>95% C.I.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP5A1Z/ATP5A1W</td>
<td>4–5</td>
<td>336</td>
<td>0.140</td>
<td>0.082–0.198</td>
</tr>
<tr>
<td>UBAP2Z/UBAP2W</td>
<td>2–5</td>
<td>858</td>
<td>0.165</td>
<td>0.125–0.205</td>
</tr>
<tr>
<td>SPINZ/SPINW</td>
<td>All</td>
<td>789</td>
<td>0.187</td>
<td>0.136–0.238</td>
</tr>
<tr>
<td>HINTZ/HINTW</td>
<td>All</td>
<td>321</td>
<td>0.321</td>
<td>0.212–0.430</td>
</tr>
<tr>
<td>CHD1Z/CHD1W</td>
<td>All</td>
<td>2943</td>
<td>0.423</td>
<td>0.376–0.470</td>
</tr>
</tbody>
</table>

*Alignment length in base pairs. No indels were found in any of the alignments of coding regions.
Evolutionary Strata on the Avian Z Chromosome

Figure 3.—Phylogenetic analysis of HINTZ and HINTW intron 2 sequences. An unrooted tree was constructed using parsimony based on 357 bp. See also the legend to Figure 2.

Figure 4.—Phylogenetic analysis of UBA2Z and UBA2W intron 1. An unrooted tree was constructed using parsimony based on 687 bp. Identical topology was obtained for intron 2 (data not shown). See also the legend to Figure 2.

Estimates of Z-W divergence times: Minimum divergence times of SPINZ/SPINW, HINTZ/HINTW, UBA2Z/UBA2W, and ATP5A1Z/ATP5A1W were estimated on the basis of intron data to date the time of cessation of recombination between the Z- and W-linked genes. Nucleotide substitution rate under the two-rate model varies from 0.475–0.656 substitutions per site per million years.

In contrast to the situation for CHD1Z/CHD1W, HINTZ/HINTW, and SPINZ/SPINW, a phylogenetic analysis of UBA2Z and UBA2W provided strong support for the grouping together of Z and W copies within each avian order (Figure 4). This indicates that UBA2Z and UBA2W ceased to recombine independently in the different orders analyzed here, subsequent to ordinal radiation. This pattern is similar to that seen for ATP5A1Z and ATP5A1W (Ellegren and Carmichael 2001). The contrasting phylogenetic patterns of CHD1Z/CHD1W, HINTZ/HINTW, and SPINZ/SPINW on the one hand and UBA2Z/UBA2W and ATP5A1Z/ATP5A1W on the other hand support the definition of two evolutionary strata on the avian Z chromosome as indicated by data from intronic divergence.

The contrasting phylogenetic patterns of CHD1Z/CHD1W, HINTZ/HINTW, and SPINZ/SPINW on the one hand and UBA2Z/UBA2W and ATP5A1Z/ATP5A1W on the other hand support the definition of two evolutionary strata on the avian Z chromosome as indicated by data from intronic divergence.

Similar analyses have previously been reported for CHD1Z/CHD1W (Kahn and Quinn 1999; Fridolfsson and Ellegren 2000; Garcia Moreno and Mindell 2000) and ATP5A1Z/ATP5A1W (Ellegren and Carmichael 2001). For SPINZ/SPINW we found very strong support for the monophyletic grouping of all W sequences from four avian orders from the infraclasses Neoaves and Eoaves in one clade and of all Z sequences in the other (Figure 2). This pattern is identical to that seen in CHD1Z/CHD1W (Kahn and Quinn 1999; Fridolfsson and Ellegren 2000; Garcia Moreno and Mindell 2000) and indicates that recombination ceased between SPINZ and SPINW prior to the Neoave-Eoave split. A similar pattern was found for HINTZ/HINTW, although it was possible to amplify this gene only in Galliformes and Anseriformes (Figure 3).

In contrast to the situation for CHD1Z/CHD1W, HINTZ/HINTW, and SPINZ/SPINW, a phylogenetic analysis of UBA2Z and UBA2W provided strong support for the grouping together of Z and W copies within each avian order (Figure 4). This indicates that UBA2Z and UBA2W ceased to recombine independently in the different orders analyzed here, subsequent to ordinal radiation. This pattern is similar to that seen for ATP5A1Z and ATP5A1W (Ellegren and Carmichael 2001). The contrasting phylogenetic patterns of CHD1Z/CHD1W, HINTZ/HINTW, and SPINZ/SPINW on the one hand and UBA2Z/UBA2W and ATP5A1Z/ATP5A1W on the other hand support the definition of two evolutionary strata on the avian Z chromosome as indicated by data from intronic divergence.
TABLE 3

Divergence dates of gametologous gene pairs shared between Z and W chromosomes based on quartet analysis

<table>
<thead>
<tr>
<th>Gene/species pairs</th>
<th>Families</th>
<th>L′ (bp)</th>
<th>Calibration</th>
<th>Reference</th>
<th>Rate</th>
<th>Divergence (95% C.I.)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ATP5A1Z/ATP5A1W</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gga/Mga</td>
<td>GAL</td>
<td>764</td>
<td>20</td>
<td>A, B</td>
<td>0.656*a</td>
<td>58 (49–70)</td>
</tr>
<tr>
<td>Smo/Bcl</td>
<td>ANS</td>
<td>770</td>
<td>10</td>
<td>A, B</td>
<td>0.475</td>
<td>60 (44–86)</td>
</tr>
<tr>
<td><strong>UBAP2Z/UBAP2W</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intron 1</td>
<td>Gga/Mga</td>
<td>GAL</td>
<td>1149</td>
<td>20</td>
<td>A, B</td>
<td>0.521 85 (71–103)</td>
</tr>
<tr>
<td>Intron 2</td>
<td>Gga/Mga</td>
<td>GAL</td>
<td>363</td>
<td>20</td>
<td>A, B</td>
<td>0.505 79 (59–104)</td>
</tr>
<tr>
<td><strong>SPINZ/SPINW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gga/Mga</td>
<td>GAL</td>
<td>584</td>
<td>20</td>
<td>A, B</td>
<td>3.387</td>
<td>111 (88–142)</td>
</tr>
<tr>
<td>Gga/Chy</td>
<td>GAL/ANS</td>
<td>524</td>
<td>68</td>
<td>C, D</td>
<td>1.951</td>
<td>142 (122–168)</td>
</tr>
<tr>
<td>Gga/Ble</td>
<td>523</td>
<td></td>
<td></td>
<td></td>
<td>2.062</td>
<td>144 (124–171)</td>
</tr>
<tr>
<td>Gga/Smo</td>
<td>526</td>
<td></td>
<td></td>
<td></td>
<td>1.991</td>
<td>145 (125–172)</td>
</tr>
<tr>
<td>Gga/Cbu</td>
<td>529</td>
<td></td>
<td></td>
<td></td>
<td>1.717</td>
<td>146 (125–172)</td>
</tr>
<tr>
<td>Gga/Bbu</td>
<td>GAL/STR</td>
<td>529</td>
<td>90</td>
<td>E</td>
<td>1.430</td>
<td>153 (135–176)</td>
</tr>
<tr>
<td>Gga/Pma</td>
<td>GAL/PAS</td>
<td>508</td>
<td>90</td>
<td>E</td>
<td>1.583*a 145 (128–160)</td>
<td></td>
</tr>
<tr>
<td>Mga/Chy</td>
<td>GAL/ANS</td>
<td>519</td>
<td>68</td>
<td>C, D</td>
<td>1.641</td>
<td>132 (114–154)</td>
</tr>
<tr>
<td>Mga/Ble</td>
<td>518</td>
<td></td>
<td></td>
<td></td>
<td>1.617</td>
<td>134 (116–157)</td>
</tr>
<tr>
<td>Mga/Smo</td>
<td>521</td>
<td></td>
<td></td>
<td></td>
<td>1.673</td>
<td>132 (114–154)</td>
</tr>
<tr>
<td>Mga/Cbu</td>
<td>523</td>
<td></td>
<td></td>
<td></td>
<td>1.502</td>
<td>133 (116–156)</td>
</tr>
<tr>
<td>Mga/Bbu</td>
<td>GAL/STR</td>
<td>525</td>
<td>90</td>
<td>E</td>
<td>1.449</td>
<td>140 (125–160)</td>
</tr>
<tr>
<td>Mga/Pma</td>
<td>GAL/PAS</td>
<td>502</td>
<td>90</td>
<td>E</td>
<td>1.675</td>
<td>136 (121–155)</td>
</tr>
<tr>
<td>Bbu/Chy</td>
<td>STR/ANS</td>
<td>551</td>
<td>90</td>
<td>E</td>
<td>1.522</td>
<td>164 (144–189)</td>
</tr>
<tr>
<td>Bbu/Ble</td>
<td>548</td>
<td></td>
<td></td>
<td></td>
<td>1.588</td>
<td>158 (138–182)</td>
</tr>
<tr>
<td>Bbu/Smo</td>
<td>551</td>
<td></td>
<td></td>
<td></td>
<td>1.493</td>
<td>170 (149–198)</td>
</tr>
<tr>
<td>Bbu/Cbu</td>
<td>551</td>
<td></td>
<td></td>
<td></td>
<td>1.472</td>
<td>166 (147–196)</td>
</tr>
<tr>
<td>Pma/Chy</td>
<td>PAS/ANS</td>
<td>528</td>
<td>90</td>
<td>E</td>
<td>1.614** 146 (129–169)</td>
<td></td>
</tr>
<tr>
<td>Pma/Ble</td>
<td>525</td>
<td></td>
<td></td>
<td></td>
<td>1.690** 144 (128–165)</td>
<td></td>
</tr>
<tr>
<td>Pma/Smo</td>
<td>528</td>
<td></td>
<td></td>
<td></td>
<td>1.661** 147 (128–168)</td>
<td></td>
</tr>
<tr>
<td>Pma/Cbu</td>
<td>528</td>
<td></td>
<td></td>
<td></td>
<td>1.651** 144 (128–165)</td>
<td></td>
</tr>
<tr>
<td><strong>HINTZ/HINTW</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gga/Mga</td>
<td>GAL</td>
<td>688</td>
<td>20</td>
<td>A, B</td>
<td>1.811</td>
<td>102 (84–124)</td>
</tr>
<tr>
<td>Gga/Cbu</td>
<td>GAL/ANS</td>
<td>361</td>
<td>68</td>
<td>C, D</td>
<td>1.034</td>
<td>113 (97–133)</td>
</tr>
<tr>
<td>Gga/Bca</td>
<td>554</td>
<td></td>
<td></td>
<td></td>
<td>1.094</td>
<td>120 (106–138)</td>
</tr>
<tr>
<td>Mga/Cbu</td>
<td>363</td>
<td></td>
<td></td>
<td></td>
<td>1.018</td>
<td>112 (97–132)</td>
</tr>
<tr>
<td>Mga/Bca</td>
<td>548</td>
<td></td>
<td></td>
<td></td>
<td>1.242</td>
<td>124 (109–142)</td>
</tr>
</tbody>
</table>

*a For species codes, please refer to materials and methods.
*b Family abbreviations are GAL, Galliformes; ANS, Anseriformes; STR, Strigiformes; and PAS, Passeriformes.
*c Number of base pairs in alignment (without gaps).
*d Divergence dates (in MYA) used as calibration points to estimate Z-W divergence times.
*e References associated with estimated Z-W divergence times are A, Sibley and Ahlquist (1990); B, Ellegren and Carmichael (2001); C, Waddell et al. (1999); D, Garcia Moreno and Mindell (2000); and E, van Tuinen and Hedges (2001).
*f Nucleotide substitution rate (in substitutions per site per million years) under the two-rate model, where asterisks refer to significance of the likelihood-ratio test using χ² approximation to compare the two-rate vs. five-rate model, where *p < 0.05 and **p < 0.01.
*g Calculated Z-W divergence date in MYA and corresponding 95% confidence intervals.

years for the genes on Zp to 1.018–3.387 for the genes on Zq, but is generally similar within introns between species (Table 3). For SPINZ/SPINW we constructed 21 quartets and the divergence between Z and W ranges from 111 to 170 MYA (95% C.I., 88–198). Divergence estimates between HINTZ and HINTW from five quartets are 102–124 MYA (95% C.I., 84–142). The estimates for SPINZ/SPINW and HINTZ/HINTW are similar to an estimate recently obtained for CHD1Z/CHD1W (123 MYA; Garcia Moreno and Mindell 2000), and, although they should be treated as rough approximates only, these combined results are consistent with the phylogenetic analyses, suggesting diversification of these gene pairs prior to the Neoave-Eoave split, 90–130 MYA (Sibley and Ahlquist 1990; Hedges et al. 1996; Cooper and Penny 1997; Kumar and Hedges 1998; Waddell et al. 1999; Garcia...
The time of this split coincides with the time of the major avian radiation. To implement the quartet-dating method, each pair of taxa must be monophyletic with respect to the other pair (Rambaut and Bromham 1998). We were therefore not able to build as many quartets for UBAP2Z/UBAP2W and ATP5A1Z/ATP5A1W as for SPINZ/SPINW since Z and W diverged within each order. However, we estimate that UBAP2Z and UBAP2W diverged within the Galliformes ~79–85 MYA (95% C.I., 59–104). For ATP5A1Z we estimated divergence times of 58 MYA (95% C.I., 49–70) within Galliformes, and 60 MYA (95% C.I., 44–86) within Anseriformes (Table 3). In summary, genes from stratum 1 on chicken Zq (CHD1Z, HINTZ, and SPINZ) seem to have diverged 102–170 MYA while genes from stratum 2 on Zp (UBAP2Z and ATP5A1Z) probably diverged 58–85 MYA.

**DISCUSSION**

Molecular evolutionary and phylogenetic data from Z- and W-linked gametologs of the five gene pairs so far characterized on the chicken sex chromosomes are compatible with a scenario where cessation of recombination between Z and W was progressive, starting with a point on the long arm of Z (Zq) and continuing to the distal region of the Z short arm (Zp), resulting in the formation of at least two evolutionary strata on the Z chromosome. This seems to have been punctuated by at least one mutational event, for example, an inversion on W that disrupted the homology with Zq, 102–170 MYA (Table 3), and led to the creation of stratum 1 prior to the major avian radiations. Moreover, our data suggest that recombination continued in the younger stratum 2 (Zp) until after the divergence of extant avian orders, when it was subsequently disrupted many times, independently in the different lineages between 58 and 85 MYA (summarized in Figure 5). The latter illustrates that fully differentiated sex chromosomes of birds have evolved by convergence rather than by reflecting a common ancestral state.

Although our locus sampling is based on data from all known gametologs so far found on avian sex chromosomes, it should be acknowledged that it is not extensive and that we lack divergence estimates for parts of the chromosome. However, some factors suggest thenumber of evolutionary strata on the chicken Z chromosome, they do not exclude the possibility of additional chromosomal segments with separate evolutionary history. Lahn and Page (1999) identified four evolutionary strata on the human X chromosome, something that may partly relate to the fact that a larger number of human gametologous gene pairs were analyzed than is possible for birds. However, some factors suggest the number of evolutionary strata on the avian Z chromosome to be rather limited. First, the segment we refer to as stratum 1 covers a major part of Zq, with HINTZ and SPINZ close to the centromere, and the distance between CHD1Z and Zqter represents only some 10% of the whole chromosome according to in situ hybridization data (Figure 1). If existing, a potential additional stratum telomeric to CHD1Z on Zq would therefore have to be small. Second, the same argument applies to a potential additional stratum between ATP5A1Z...
and the pseudoautosomal region at Zpter, since this distance is even shorter than that between CHD1Z and Zpter.

If additional strata do exist on chicken Z, a more likely candidate region is between the centromere and UBAP2Z, comprising ≈25% of the chromosome, which is (so far) not covered by any gene shared with the W chromosome. This region contains the suggested boundary between strata 1 and 2 as defined by intron divergence. However, if one (or more) additional stratum is present in this region it must have evolved during a relatively narrow time interval, possibly between 85 and 102 MYA according to our estimates. The four strata identified on the human X chromosome evolved during 240–320 million years of sex chromosome evolution (Lahn and Page 1999), indicating that stratification is a slow process. Overall, the combined data from mammals and birds suggest that the number of strata formed during vertebrate sex chromosome evolution is limited. Moreover, the events leading to cessation of recombination (possibly chromosomal inversions occurring on the nonrecombining chromosome; Bengtsson and Goodfellow 1987; Yen et al. 1998; Lahn and Page 1999) seem to involve large chromosomal segments.

It is of interest to compare the chromosomal distribution of gametologous genes on the human X chromosome and chicken Z chromosome. On human X an increasing density of genes is still shared with Y with decreasing age of strata. For instance, while only three genes shared with Y are present on the long arm of X (stratum 1), 16 (strata 2–4) shared genes are on the short arm (Lahn and Page 1999). The observation of three gametologous genes on the long arm of chicken Z (stratum 1) but only two on the short arm (stratum 2) is significantly different from the distribution of gametologous genes on the long and short arms of human X (χ² = 4.1, P < 0.05). One possible explanation for this difference in chromosomal distribution is that chicken stratum 1 is younger (102–170 MYA) than human stratum 1 (240–320 MYA). It is generally believed that once recombination between sex chromosomes ceases, most genes on the nonrecombining chromosome will decay and eventually become obliterated. The density of genes on X or Z with “surviving” gametologs on Y or W should thus be negatively correlated with the time since recombination between the two (proto-) sex chromosomes was suppressed. One interesting consequence of this idea is that the proportion of gametologous gene pairs on the chicken sex chromosomes may be higher than that in human. Only a few percent of genes on the human X chromosome have detectable gametologs on Y, whereas in birds 5 out of ~25 genes mapped to the chicken Z chromosome represent gametologous genes shared with W (Ellegren 2000; Schmid et al. 2000).

Although the latter ratio is probably biased by the active search for genes on the avian W chromosome, this may reflect a difference between mammalian and avian sex chromosomes.

Although the density of genes on Z (or X) with surviving gametologs on W (Y) is expected to be highest in the youngest stratum, we may at the same time expect genes with gametologs currently being degenerated on W (Y) to be concentrated mainly in this stratum. The majority of human X-linked genes that have nonfunctional Y copies (pseudogenes) are concentrated in the youngest stratum on X (Lahn and Page 1999). Consistent with this, ATP5AJW in chicken stratum 2 seems to have been lost from the W chromosome in some avian lineages and is thought to be a pseudogene in others (de Kloet 2001).

While human sex chromosome evolution was apparently initiated shortly after the split of mammalian and bird lineages, dated ~300 MYA (Kumar and Hedges 1998), differentiated sex chromosomes among birds are of more recent origin. Birds and crocodiles diverged 225–245 MYA (Benton 1990) and our dating of the formation of chicken stratum 1 to 102–170 MYA is compatible with the observation that crocodiles typically have homomorphic (undifferentiated) sex chromosomes. The timing of avian sex chromosome evolution is relevant to the disputed question of basal lineages in the phylogenetic tree of extant birds. A traditional view (Sibley and Ahlquist 1990) holds that the deepest split is that between the Paleognathae [Struthiformes (ratites, e.g., ostriches) and Tinamiformes (tinnamous)] and Neognathae (all other birds). However, mitochondrial DNA (mtDNA) sequence data suggest that Passeriformes represent the basal offshoot (Minnell et al. 1997, 1999; Härlid et al. 1998), and additional alternatives have also been proposed (see Härlid et al. 1998). Unlike other birds, ratites have homomorphic sex chromosomes (ostrich, cassowary, and kiwi) or only marginally differentiated ones (emu and rhea). Genes that are W-linked in other birds are not female-specific in ratites. Moreover, three genes found on the Z, but not on W, in other birds are present on both Z and (the very similar) W of emus (Ogawa et al. 1998). Now, critical in this context is our observation that the differentiation of all three genes from chicken stratum 1 predates the radiation of major avian lineages, such as Galliformes, Anseriformes, Strigiformes, and Passeriformes. The most parsimonious scenario compatible with these observations is that ratites branched off the lineage leading to all other extant birds before Z and W sex chromosome evolution was initiated, which in turn took place before the radiation of extant lineages. In fact, these three events may have happened within a rather limited period of time, as the difficulty in resolving early branch orders with phylogenetic analysis probably indicates that the time between the Paleognathae-Neognathae split and major avian radiation was limited. This reasoning relies on the assumption that homomorphic sex chromosomes of ratites are not a derived character state.
Indeed, the suppression of recombination between differentiated sex chromosomes is likely to be an irreversible process (Ogawa et al. 1998).

Related to the above it should finally be noted that there is no evidence for large rearrangements involving the Z chromosome among extant bird lineages. This contrasts with the situation for the eutherian X chromosome to which additional material was transferred between the time of divergence of the eutherian and metatherian lineages and the divergence of eutherian orders (Watson et al. 1991; Graves 1996). Available gene mapping data indicate that the gene content of Z has been conserved during avian evolution (Orno 1967; Bavestock et al. 1982; Schmid et al. 1989, 2000; Saitoh et al. 1993; Nanda et al. 1998; Ogawa et al. 1998; Shetty et al. 1999; Ellegren 2000), although more detailed genetic or physical maps of species other than chicken will be required before this can be studied at the scale of gene order. As a consequence, we do not yet know whether the two evolutionary strata on chicken Z identified in this study are similarly organized in other bird lineages.

In conclusion, our data suggest that similar mechanisms appear to be driving the evolution of sex chromosomes in mammals and birds, leading to the formation of evolutionary strata. Moreover, the observation of independent disruption of recombination in stratum 2, in different avian lineages, strengthens the idea that progressive cessation of recombination is a general feature of sex chromosome evolution. In fact, this may extend also to plants since different estimates of silent site divergence have been obtained for two gene pairs on the X and Y sex chromosomes of the dioecious plant Silene latifolia (Atanassov et al. 2001; Charlesworth 2002).

We thank Hannah Sundström for unpublished data and members of our research group for useful discussions throughout this project. Financial support was obtained from the Swedish Natural Sciences Research Council. H.E. is a Royal Swedish Academy of Sciences Research Fellow supported by a grant from the Knut and Alice Wallenberg Foundation.

LITERATURE CITED


