

# FEMALE HETEROGAMETY AND SPECIATION: REDUCED INTROGRESSION OF THE Z CHROMOSOME BETWEEN TWO SPECIES OF NIGHTINGALES

Radka Storchová,<sup>1,2,3</sup> Jiří Reif,<sup>4</sup> and Michael W. Nachman<sup>1</sup>

<sup>1</sup>Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona 85721

<sup>2</sup>E-mail: radkas@natur.cuni.cz

<sup>3</sup>Department of Zoology, Faculty of Science, Charles University in Prague, Czech Republic

<sup>4</sup>Institute for Environmental Studies, Faculty of Science, Charles University in Prague, Czech Republic

Received July 1, 2009

Accepted September 3, 2009

Several lines of evidence suggest that the X chromosome plays a large role in intrinsic postzygotic isolation. The role of the Z chromosome in speciation is much less understood. To explore the role of the Z chromosome in reproductive isolation, we studied nucleotide variation in two closely related bird species, the Thrush Nightingale (*Luscinia luscinia*) and the Common Nightingale (*L. megarhynchos*). These species are isolated by incomplete prezygotic isolation and female hybrid sterility. We sequenced introns of four Z-linked and eight autosomal loci and analyzed patterns of polymorphism and divergence using a divergence-with-gene flow framework. Our results suggest that the nightingale species diverged approximately 1.8 Mya. We found strong evidence of gene flow after divergence in both directions, although more introgression occurred from *L. megarhynchos* into *L. luscinia*. Gene flow was significantly higher on the autosomes than on the Z chromosome. Our results support the idea that the Z chromosome plays an important role in intrinsic postzygotic isolation in birds, although it may also contribute to the evolution of prezygotic isolation through sexual selection. This highlights the similarities in the genetic basis of reproductive isolation between organisms with heterogametic males and organisms with heterogametic females during the early stages of speciation.

**KEY WORDS:** Birds, hybrid zone, reproductive isolation, sex chromosomes, speciation.

Understanding the genetic basis of reproductive isolation may provide important insights into the mechanisms responsible for the origin of new species. Previous studies have suggested that the X chromosome contributes disproportionately to the origin of intrinsic postzygotic isolation, including both hybrid sterility and hybrid inviability. This has been demonstrated in laboratory crosses (e.g., True et al. 1996; Turelli and Begun 1997; Tao et al. 2003; Oka et al. 2004; Storchová et al. 2004; Masly and Presgraves

2007; Good et al. 2008) as well as in studies of hybrid zones (Payseur et al. 2004; Macholán et al. 2007; Teeter et al. 2008). The disproportionately large role of the X chromosome in intrinsic postzygotic isolation may have several distinct causes including exposure of recessive mutations in hemizygous individuals, or a higher density of genes contributing to postzygotic isolation on the X (Masly and Presgraves 2007). A higher density of X-linked genes contributing to postzygotic isolation may, in turn, be caused by one of several mechanisms including faster X evolution, the accumulation of sex-chromosome meiotic drive mechanisms, or disruption of epigenetic regulation of the X chromosome in

Sequence data from this article have been deposited with the GenBank Data Libraries under accession nos. FJ695620-FJ696173.

the male germline (reviewed in Coyne and Orr 2004; Presgraves 2008).

Despite the extensive literature on the role of the X chromosome in reproductive isolation, relatively few studies have focused on the role of the Z chromosome in speciation (e.g., Jiggins et al. 2001; Sætre et al. 2003, 2007; Carling and Brumfield 2008, 2009). However, several lines of evidence suggest that the Z chromosome might be even more important in reproductive isolation than the X chromosome. The first is that faster evolution of the Z compared to the autosomes is clear in birds (Mank et al. 2007; Ellegren 2009), whereas evidence for “faster X evolution” is equivocal (e.g., Betancourt et al. 2002; Thornton and Long 2002; Torgerson and Singh 2003; Counterman et al. 2004; Connallon 2007; Baines et al. 2008). A higher rate of adaptive evolution on the Z chromosome relative to autosomes could underlie either intrinsic or extrinsic reproductive isolation. Second, the Z chromosome may promote speciation through sexual selection (Ritchie 2007). Several studies suggest that sexual selection is facilitated in species with female heterogamety, because sons inherit the Z chromosome directly from their father (Hastings 1994; Reeve and Pfennig 2003; Kirkpatrick and Hall 2004; Albert and Otto 2005). An important prediction of these studies is that male traits and female preferences will often be Z-linked, and indeed, there is some evidence for such Z-linkage in birds (Sætre et al. 2003; Sæther et al. 2007) and Lepidoptera (Iyengar et al. 2002). Finally, genetic association between genes involved in pre mating and postzygotic isolation, which can arise if both are on the Z chromosome, is expected to facilitate the evolution of reinforcement (Servedio and Sætre 2003; Hall and Kirkpatrick 2006). In *Ficedula* flycatchers, genes underlying female preference and low hybrid fitness are both Z-linked (Sætre et al. 2003). Thus, there are a variety of reasons for expecting that the Z chromosome might be involved in reproductive isolation, including prezygotic isolation and either intrinsic or extrinsic postzygotic isolation.

Two closely related bird species, the Thrush Nightingale (*Luscinia luscinia*) and the Common Nightingale (*L. megarhynchos*), are well suited for studying the role of the Z chromosome in speciation. These species are believed to have diverged in geographical isolation during the Pleistocene and have subsequently come into secondary contact in central and Eastern Europe where they form a narrow hybrid zone (Sorjonen 1986) (Fig. 1). Despite their overall morphological and ecological similarity, the two species can be distinguished by wing feather characteristics and slightly different songs (Cramp 1988). However, males with a mixed song are frequently observed in areas in which both species co-occur (Becker 2007). Although there is strong assortative mating between the species, mixed pairs as well as interspecific hybrids are occasionally observed in the region of sympatry (Becker 2007; Kverek et al. 2008). Morphological as well as genetic studies indicate that approximately 5% of birds in



**Figure 1.** Geographic distribution of the Common Nightingale (dark gray) and the Thrush Nightingale (light gray) with indicated sampled populations. (1) Distal allopatry for the Common Nightingale (Spain), (2) close allopatry for the Common Nightingale (Czech Republic, SW Poland), (3) close allopatry for the Thrush Nightingale (NE Poland), and (4) distal allopatry for the Thrush Nightingale (Finland). The zone of sympatry is depicted only schematically owing to uncertainty about the exact position and width of the hybrid zone.

sympatry are hybrids (Becker 2007; R. Storchová et al. unpubl. ms.). Interestingly, interspecific pairs show slightly lower hatching performance compared to conspecific pairs, indicating that hybrids might be partially inviable (Becker 2007). Crosses between the two species in captivity have shown that, in accordance with Haldane’s rule,  $F_1$  hybrid females are sterile, whereas  $F_1$  hybrid males are fertile (Stadie 1991). Overall, these observations indicate that prezygotic isolation exists but is not complete, and that intrinsic postzygotic isolation also contributes to the reproductive barrier between the Common Nightingale and the Thrush Nightingale.

To investigate the role of the Z chromosome in reproductive isolation between nightingales, we examined patterns of polymorphisms and divergence at multiple autosomal and Z-linked loci in allopatric populations of both species. Two other pairs of recently diverged species of birds have been studied for patterns of differentiation at Z-linked and autosomal loci. Carling

and Brumfield (2008) studied gene flow across a hybrid zone between Lazuli and Indigo Buntings in North America using four autosomal and two Z-linked markers, and showed that the two Z-linked markers had an average cline width of 309 km whereas the four autosomal markers had an average cline width of 466 km. Sætre and colleagues have studied Pied and Collared Flycatchers in Europe in both sympatric and allopatric populations and found that the Z chromosome shows greater differentiation than the autosomes, likely as a consequence of reduced introgression of the Z chromosome (Sætre et al. 2003) as well as positive selection at some Z-linked loci (Borge et al. 2005). Here, we expand on these studies by using DNA sequence data and an explicit divergence-with-gene-flow model to assess the contribution of gene flow and selection to observed patterns of differentiation at autosomal and Z-linked loci in nightingales.

## Materials and Methods

### ANIMALS AND SAMPLES

*Luscinia megarhynchos* and *L. luscinia* are sister species (Jönsson and Fjeldså, 2006) and they are not known to hybridize with other species (see the database of bird hybrids: www.bird-hybrids.com). Samples from 25 males of *L. megarhynchos* and 25 males of *L. luscinia* were analyzed in this study. Birds were caught during the breeding season (from late April until mid June) in 2006 and 2007. Two different populations from each species were sampled. According to their distance from the hybrid zone, we refer to these populations as close allopatry (Czech Republic and SW Poland for *L. megarhynchos*, NE Poland for *L. luscinia*) and distal allopatry (Spain for *L. megarhynchos*, Finland for *L. luscinia*) (Fig. 1). Within each population, birds were captured at different localities several kilometers apart. The exact positions of the localities are provided in Table A1. Additionally, one male Bluethroat (*Luscinia svecica*) was used as an outgroup. *Luscinia svecica* is the sister taxon to the clade containing *L. megarhynchos* and *L. luscinia* (Jönsson and Fjeldså, 2006). Genetic divergence at cytochrome b between *L. svecica* and either *L. megarhynchos* or *L. luscinia* is approximately 14% (Wink et al. 2002), which corresponds to a divergence time of 7 million years assuming approximately 2% sequence divergence per million years (Wilson et al. 1985). From each individual, a blood sample was collected by brachial vein puncture and stored in pure ethanol for further extraction of genomic DNA.

### DNA EXTRACTION, PCR, AND SEQUENCING

Genomic DNA was purified by DNeasy Tissue Kit (Qiagen) and used for PCR amplification and sequencing of eight autosomal and four Z-linked loci. Primer sequences were obtained from published studies (Borge et al. 2005; Backström et al. 2006, 2008). These primers were designed in conserved exonic regions of the

chicken genome to amplify intronic sequences. PCR conditions as well as the length of PCR products are given in Table S1. In a few cases, nightingale-specific primers were designed from initial sequence to reduce nonspecific PCR amplification. Primer sequences are provided in Table S2. All PCR products were sequenced in both directions. The chromosomal location of each locus (i.e., whether autosomal or Z-linked) was determined on the basis of the conserved synteny between the chicken genome and the genome of passerines (Backström et al. 2006; Itoh et al. 2006; Nanda et al. 2008). We were also interested in knowing whether any of these loci were closely linked to each other. In the absence of a physical or genetic map for nightingales, we determined the chromosomal position of each locus in two bird species with known genome sequences, *Gallus gallus* and *Taeniopygia guttata*, using the Ensembl database (Table S3). We reasoned that if linkage relationships were similar in chickens and in zebra finches, they might be conserved in nightingales as well given the generally conserved karyotypes of birds (Price 2007). Most of the autosomal genes lie on different chromosomes both in the *G. gallus* and *T. guttata* genomes. Of the four Z-linked genes, *PPWD1* and *ADAMTS6* are close physically in both species (and thus probably also in nightingales), and therefore may not have independent evolutionary histories. However, the other two loci on the Z chromosome are separated from each other and from *PPWD1* and *ADAMTS6* by > 5 Mb in both species. The recombination rate on the Z chromosome is 1.43 cM/Mb in *T. guttata* (Pigozzi 2008) and 2.7 cM/Mb in *G. gallus* (Wahlberg et al. 2007). Thus, these loci are likely separated by at least 7 cM in *Luscinia* and are unlikely to be in linkage disequilibrium.

Sequences were edited using phred/phrap/consed/polyphred (Nickerson et al. 1997; Ewing and Green 1998; Ewing et al. 1998; Gordon et al. 1998) together with auxiliary shell scripts and Perl programs kindly provided by August Woerner. Alignments were generated by ClustalW (Thompson et al. 1994) as implemented in the program BioEdit (Hall 1999). All alignments were visually checked and manually adjusted. Exonic sequences as well as indel polymorphisms were excluded from the analyses. Individuals and/or positions with missing data were eliminated from the dataset. Sequences have been deposited in GenBank under accession numbers FJ695620-FJ696173.

### DATA ANALYSES

Diploid sequences at each locus were separated into two haplotypes computationally using the program PHASE, version 2.1 (Stephens et al. 2001; Stephens and Donnelly 2003) with the following parameters: number of iterations = 10,000, thinning interval = 1, burnin = 1000 (see PHASE 2.1 documentation for details). We used the default recombination model, which is the general model for varying recombination rate from Li and Stephens (2003). For each dataset, we applied the algorithm

five times with different random seeds, and we checked for consistency of results across independent runs. We obtained identical haplotypes across all runs for 11 of 12 loci. The only locus that showed slightly different haplotypes in different runs was 25149. We note that haplotype reconstruction can be affected by several demographic and evolutionary factors, including population subdivision, recent population expansion, or selection (Andrés et al. 2007). However, we find no evidence for population size changes or selection in these data (see Results).

Basic population genetic analyses of polymorphism, divergence, recombination, and tests of neutrality based on the allele frequency spectrum were performed with the programs DnaSP (Rozas et al. 2003) and SITES (Hey and Wakeley 1997). Genealogical relationships among haplotypes for each locus were reconstructed with the Neighbor-Joining method (Saitou and Nei 1987) using the program MEGA4 (Tamura et al. 2007). Hudson-Kreitman Aguadé (HKA) tests (Hudson et al. 1987) were performed using the HKA program (<http://lifesci.rutgers.edu/~heylab/>). The  $K_{ST}$  test for detecting geographic subdivision (Hudson et al. 1992) was performed using DnaSP.

We looked for evidence of gene flow by calculating the amount of linkage disequilibrium (LD) among specific classes of segregating sites as described in Machado et al. (2002) using the software SITES.  $D'$  (Lewontin 1964) was used as a measure of LD. The test of Machado et al. (2002) is based on comparison of LD among pairs of shared polymorphisms (DSS), LD among pairs of exclusive polymorphisms (DXX), and LD among pairs of sites where one member is a shared and the other an exclusive polymorphism (DSX). If gene flow occurs after divergence, the difference,  $DSS - DXX$  or  $DSS - DSX$ , which is called  $x$ , is expected to be positive because polymorphisms that are shared due to common ancestry are expected to be older, on average, than polymorphisms that are shared due to postdivergence gene flow. Less LD is expected between older polymorphisms because recombination will have had more time to break down associations. Statistical significance of the observed  $x$  values was obtained by simulating datasets using the program WH (Wang et al. 1997; <http://lifesci.rutgers.edu/~heylab/>). This program simulates a model of isolation without gene flow. Because our dataset exceeded the maximum input for WH (which is 32 chromosomes per locus), we divided the dataset into three subsamples with the appropriate sample size and performed three independent simulations. Parameter estimates for all three subsamples are given in Table S4. The results of the LD test are shown only for  $x = DSS - DXX$ ; however, the results based on  $x = DSS - DSX$  were similar.

The data were fit to the isolation-with-migration model (IM) using the programs IM and IMA (Nielsen and Wakeley 2001; Hey and Nielsen 2004, 2007). These programs estimate six param-

eters by Markov chain Monte Carlo simulations of genealogies. These parameters include the population-split time, the effective population size for the ancestral population and for the current populations, and the migration rates in both directions. Because IM and IMA assume no recombination within loci, we determined the longest region without observed recombination for each locus using the program IMgc (Woerner et al. 2007). This program removes either sites or haplotypes to produce the most information-rich contiguous DNA sequence segment that passes the four-gamete test. Nonrecombinant regions represented 79% of the length of each locus, on average, and were used as an input for the IM and IMA programs. We ran the programs separately for (1) all 12 loci, (2) the eight autosomal loci and (3) the four Z-linked loci. For each dataset, the programs were run three or four times with identical starting conditions, with the exception of the random number seed, to assess convergence. To facilitate mixing of the Markov chains we used Metropolis coupling with 15–20 chains and a geometric heating scheme. All runs began with a burn-in period of 100,000 steps and were allowed to continue for 1–10 millions steps. We were able to achieve adequate mixing of the Markov chains as indicated by trend line plots and effective sample size (ESS) values (for most runs, all ESS values were higher than 500 and no ESS value was lower than 50). Independent runs converged to the same result (e.g., maximum-likelihood estimates and marginal posterior probability distributions for the demographic parameters were essentially the same for all three runs; see Fig. S1). All estimated parameters of the IM model are scaled to the mutation rate. To convert these parameters to biologically more meaningful quantities (i.e.,  $N_e$ , effective population size in number of individuals;  $m$ , migration rate per year;  $2Nm$ , population migration rate;  $t$ , divergence time in years) we calculated the neutral mutation rate for each locus ( $\mu$ ) using divergence to the outgroup. The neutral mutation rate per year was calculated from the formula  $D = 2\mu t_1$  where  $D$  is the estimated  $D_{xy}$  (Nei 1987) between *L. megarhynchos/luscinia* and *L. svecica* for each locus and  $t_1$  is the divergence time. We assumed that the divergence time between *L. megarhynchos* or *L. luscinia* and *L. svecica* is 7 million years as estimated on the basis of cytochrome *b* sequence divergence (Wink et al. 2002). The geometric mean of the locus-specific mutation rates was then used for the parameter conversion. In analyses involving only Z-linked loci or only autosomal loci, the geometric mean of locus-specific mutation rates was calculated using only Z-linked or autosomal loci, respectively. To estimate mutation rates per generation, we need to know the generation time. Nightingales, like most passerines, first reproduce one year after hatching and thus a generation time of one year was assumed. This may be an underestimate because some individuals may reproduce over multiple years. Uncertainty in generation time will affect estimates of  $N_e$ , but will not affect estimates of  $2Nm$  or  $t$ .

## Results

We sequenced eight autosomal and four Z-linked loci in close allopatric and distal allopatric populations of both species. For each locus, sequence from 38 to 50 chromosomes in *L. megarhynchos* and 34 to 50 chromosomes in *L. luscinia* was obtained (Table 1). The same loci were also sequenced in one male *L. svecica*, representing an outgroup. No significant genetic differentiation was revealed between close and distal allopatric populations within species for the vast majority of loci ( $K_{ST}$  ranged from  $-0.1964$  to  $0.0335$ ,  $P > 0.05$ ). The only significant differentiation was detected for locus 25149 in *L. megarhynchos* ( $K_{ST} = 0.0454$ ,  $P = 0.007$ ) and for locus 25374 in *L. luscinia* ( $K_{ST} = 0.0580$ ,  $P = 0.022$ ). However, after applying a Bonferroni correction for multiple testing, population subdivision was not significant for any locus. Thus, we pooled data from close and distal allopatric populations in all subsequent analyses. The absence of strong geographic subdivision may be related to high dispersal rates due to migration.

### LEVELS OF INTRASPECIFIC POLYMORPHISM

Levels of nucleotide variation were quite high for both species, although slightly higher for *L. luscinia* than for *L. megarhynchos* (on autosomes, average  $\pi = 0.850\%$  for *L. luscinia* and  $0.796\%$  for *L. megarhynchos*; average  $\theta = 0.994\%$  for *L. luscinia* and  $0.679\%$  for *L. megarhynchos*) (Table 1). This suggests a large effective population size for both species, consistent with their high abundances and large geographic ranges. Roughly similar values of nucleotide diversity have been observed, for example, in *Drosophila* (Moriyama and Powell 1996) and rabbits (Carneiro et al. 2009).

Nucleotide variation was lower on the Z chromosome than on the autosomes, even after we corrected for the different effective population size of the Z chromosome, which is three-fourth that of the autosomes assuming an equal sex ratio (after correction, the Z/A ratio for  $\pi$  was  $0.33$  in *L. megarhynchos* and  $0.54$  in *L. luscinia*; the Z/A ratio for  $\theta$  was  $0.56$  in *L. megarhynchos* and  $0.65$  in *L. luscinia*). To test whether differences in mutation rate could account for reduced Z chromosome variation within species, we divided  $\theta$  by divergence (Dxy) to *L. svecica* (which can be used as a surrogate for the mutation rate). The ratio  $\theta/Dxy$  was still slightly lower for Z-linked than for autosomal loci (Z/A ratio was  $0.58$  in *L. megarhynchos*,  $0.72$  in *L. luscinia*). We then applied a 12-locus HKA test as a more formal approach to asking if the Z chromosome shows reduced diversity relative to the autosomes. This test takes into account differences in mutation rate (by using an outgroup) as well as differences in the effective population size of the Z and autosomes. The test was applied separately to each species and to both species together, in each case using a single sequence from *L. svecica* as an outgroup. None of these three tests rejected a neutral model ( $P > 0.10$ ; 10,000 permuta-

tions). We also performed two-locus HKA tests with the pooled data from the autosomes as one locus and the pooled data from the Z chromosome as a second locus to test for differences in nucleotide variation between the autosomes and the Z chromosome, without testing for variation among individual loci. This was done for each species separately as well as both species together. None of these three tests were significant ( $P > 0.10$ ; 10,000 permutations) (Table S5). Together, these results suggest that there is only a modest (nonsignificant) reduction of variation on the Z chromosome compared to the autosomes.

We also looked for evidence of selection by comparing the distribution of allele frequencies with the expectations under a neutral equilibrium model using Tajima's D (TD) (Table 1). None of the 24 values of TD within species were significantly different from the neutral expectation of 0. There was a very slight tendency toward negative values of TD (17 of these 24 values were negative), and this may reflect mild population expansions.

### INTERSPECIFIC DIFFERENTIATION

Levels of differentiation between the two species of nightingales are summarized in Table 2. We observed a striking contrast in the level of differentiation between the Z chromosome and the autosomes. The mean  $F_{ST}$  as well as net divergence (Da) were consistently and significantly higher for Z-linked loci ( $F_{ST} = 0.783$ ,  $Da = 0.976\%$ ) than for autosomal loci ( $F_{ST} = 0.298$ ,  $Da = 0.327\%$ ) (Mann-Whitney U-test,  $P < 0.01$ ). A similar contrast was found in the number of fixed differences and shared polymorphisms. In total, the Z-linked loci showed 14 fixed differences and only one shared polymorphism. On the other hand, the autosomal loci showed 0 fixed differences and 52 shared polymorphisms (Fisher exact test,  $P < 0.0001$ ).

Another way of examining differentiation is to look at gene genealogies. Figure 2 shows neighbor-joining trees for each locus. Despite generally low bootstrap values, the genealogies suggest that lineage sorting has been completed on the Z chromosome, with the exception of the *PPWD1* locus. A closer inspection of the sequences of *PPWD1* gene revealed that one haplotype of *L. luscinia* seems to be introgressed into *L. megarhynchos*. The length of this introgressed haplotype is 344bp and includes six polymorphic sites. Because the length of this introgression is a little bit shorter than the length of the whole sequence, the neighbor-joining tree shows this sequence somewhere between *L. luscinia* and *L. megarhynchos*. In contrast, the autosomal loci did not show clear phylogenetic differentiation between species. The lack of concordance between genealogical relationships and species identity for autosomal loci might indicate that lineage sorting is not yet complete or that interspecific gene flow has occurred.

**Table 1.** Polymorphism statistics for eight autosomal and four Z-linked loci.

Locus	Chr	L <sup>1</sup>	Species <sup>2</sup>	N <sup>3</sup>	S <sup>4</sup>	$\pi$ (%) <sup>5</sup>	$\theta$ (%) <sup>6</sup>	TD <sup>7</sup>	D (%) <sup>8</sup>	$\gamma$ (%) <sup>9</sup>
<i>CKB-2</i>	A	453	LM	48	7	0.369	0.348	0.156	0.814	0.000
			LL	50	14	0.624	0.690	-0.289	1.307	4.278
			LM+LL	98	18	0.783	0.770	0.048	1.065	1.132
<i>ACTB-3</i>	A	291	LM	40	15	1.671	1.212	1.207	3.505	5.374
			LL	42	23	1.471	1.837	-0.664	3.068	2.750
			LM+LL	82	29	1.645	2.002	-0.553	3.281	5.414
<i>GAPD-3</i>	A	133	LM	44	3	0.828	0.519	1.272	4.033	0.000
			LL	46	1	0.064	0.171	-0.860	3.792	-
			LM+LL	90	4	0.557	0.593	-0.123	3.910	-
<i>ALAS1-8</i>	A	377	LM	38	11	0.634	0.694	-0.267	4.761	1.188
			LL	34	17	0.818	1.103	-0.864	4.595	3.093
			LM+LL	72	24	0.794	1.313	-1.223	4.682	1.632
<i>GHI-2</i>	A	547	LM	40	18	0.667	0.774	-0.450	5.495	3.613
			LL	48	26	0.936	1.071	-0.417	5.891	7.865
			LM+LL	88	32	0.934	1.159	-0.602	5.711	4.711
25374	A	710	LM	48	16	0.681	0.508	1.065	3.744	0.286
			LL	46	35	1.160	1.122	0.117	4.057	0.727
			LM+LL	94	47	1.258	1.294	-0.089	3.897	0.791
1152	A	568	LM	50	13	0.461	0.511	-0.291	4.092	0.247
			LL	42	12	0.399	0.491	-0.571	3.999	0.375
			LM+LL	92	23	0.561	0.795	-0.878	4.049	0.309
25149	A	747	LM	44	28	1.056	0.862	0.759	4.974	5.367
			LL	46	48	1.327	1.464	-0.326	4.840	5.987
			LM+LL	90	59	1.370	1.559	-0.394	4.909	5.254
<b>Average</b>	<b>A</b>	<b>487</b>	<b>LM</b>	<b>44.0</b>	<b>13.9</b>	<b>0.796</b>	<b>0.679</b>	<b>0.431</b>	<b>3.927</b>	<b>2.009</b>
			<b>LL</b>	<b>44.3</b>	<b>22.0</b>	<b>0.850</b>	<b>0.994</b>	<b>-0.484</b>	<b>3.944</b>	<b>3.582</b>
			<b>LM+LL</b>	<b>88.3</b>	<b>29.5</b>	<b>0.988</b>	<b>1.186</b>	<b>-0.477</b>	<b>3.938</b>	<b>2.749</b>
<i>GHR-5.2</i>	Z	404	LM	50	8	0.302	0.442	-0.863	4.772	0.000
			LL	50	13	0.451	0.718	-1.117	4.144	2.843
			LM+LL	100	24	1.033	1.147	-0.296	4.458	1.243
<i>PPWD1</i>	Z	574	LM	44	10	0.189	0.400	-1.545	4.328	0.000
			LL	42	8	0.265	0.324	-0.516	3.824	0.000
			LM+LL	86	18	0.710	0.624	0.399	4.082	0.000
<i>ADAMTS6</i>	Z	483	LM	50	3	0.110	0.139	-0.432	2.133	-
			LL	50	8	0.307	0.370	-0.463	2.075	2.754
			LM+LL	100	13	0.555	0.520	0.182	2.104	0.006
<i>SPINZ-2</i>	Z	822	LM	50	6	0.191	0.163	0.439	4.063	0.000
			LL	50	19	0.350	0.516	-1.019	4.195	0.000
			LM+LL	100	32	0.756	0.752	0.016	4.129	0.000
<b>Average</b>	<b>Z</b>	<b>571</b>	<b>LM</b>	<b>48.5</b>	<b>6.8</b>	<b>0.198</b>	<b>0.286</b>	<b>-0.600</b>	<b>3.824</b>	<b>0.000</b>
			<b>LL</b>	<b>48.0</b>	<b>12.0</b>	<b>0.343</b>	<b>0.482</b>	<b>-0.779</b>	<b>3.560</b>	<b>1.399</b>
			<b>LM+LL</b>	<b>96.5</b>	<b>21.8</b>	<b>0.764</b>	<b>0.761</b>	<b>0.075</b>	<b>3.693</b>	<b>0.312</b>

<sup>1</sup>Length of sequence (bp).<sup>2</sup>LM, *L. megarhychos*; LL, *L. Luscinia*.<sup>3</sup>Number of haploid sequences.<sup>4</sup>Number of segregating sites.<sup>5</sup>Average number of nucleotide differences (Nei 1987).<sup>6</sup>Proportion of polymorphic sites (Watterson 1975).<sup>7</sup>Tajima's D (Tajima 1989).<sup>8</sup>Divergence to outgroup measured as average pairwise divergence, Dxy (Nei 1987).<sup>9</sup>Population recombination parameter (Hey and Wakeley 1997).

**Table 2.** Genetic differentiation, net divergence, and numbers of shared and fixed polymorphisms between the Common Nightingale and the Thrush Nightingale.

Locus	Chr	$F_{ST}^1$	$D_a$ (%) <sup>2</sup>	$S_s^3$	$S_f^4$
<i>CKB-2</i>	A	0.532	0.563	3	0
<i>ACTB-3</i>	A	0.089	0.153	10	0
<i>GAPD-3</i>	A	0.346	0.236	0	0
<i>ALAS1-8</i>	A	0.166	0.145	4	0
<i>GH1-2</i>	A	0.231	0.240	12	0
<i>25374</i>	A	0.424	0.678	4	0
<i>1152</i>	A	0.372	0.255	2	0
<i>25149</i>	A	0.225	0.346	17	0
<b>All autosomal loci<sup>5</sup></b>	<b>A</b>	<b>0.298</b>	<b>0.327</b>	<b>52</b>	<b>0</b>
<i>GHR-5.2</i>	Z	0.775	1.299	1	5
<i>PPWD1</i>	Z	0.808	0.957	0	0
<i>ADAMTS6</i>	Z	0.767	0.686	0	2
<i>SPINZ-2</i>	Z	0.780	0.961	0	7
<b>All Z-linked loci<sup>5</sup></b>	<b>Z</b>	<b>0.783</b>	<b>0.976</b>	<b>1</b>	<b>14</b>

<sup>1</sup>Fixation index.<sup>2</sup>Net nucleotide divergence (Nei 1987).<sup>3</sup>Number of shared polymorphisms.<sup>4</sup>Number of fixed polymorphisms.<sup>5</sup>The average is shown for  $F_{ST}$  and  $D_a$ . The sum is shown for  $S_s$  and  $S_f$ .

## TESTS OF INTERSPECIFIC INTROGRESSION AND ESTIMATION OF DEMOGRAPHIC PARAMETERS

To reveal whether interspecific gene flow occurred after the divergence of the two species, we performed two tests. The first test is based on patterns of LD within genes (Machado et al. 2002). Under a scenario of postdivergence gene flow, observed values of the  $x$  statistic are expected to be more positive than simulated values (see Materials and Methods). The observed values of  $x$  were consistently higher than the simulated values in both species, and in four of five genes,  $x$  was significant in at least one run (Table 3). These results suggest that gene flow has occurred after the divergence of both species.

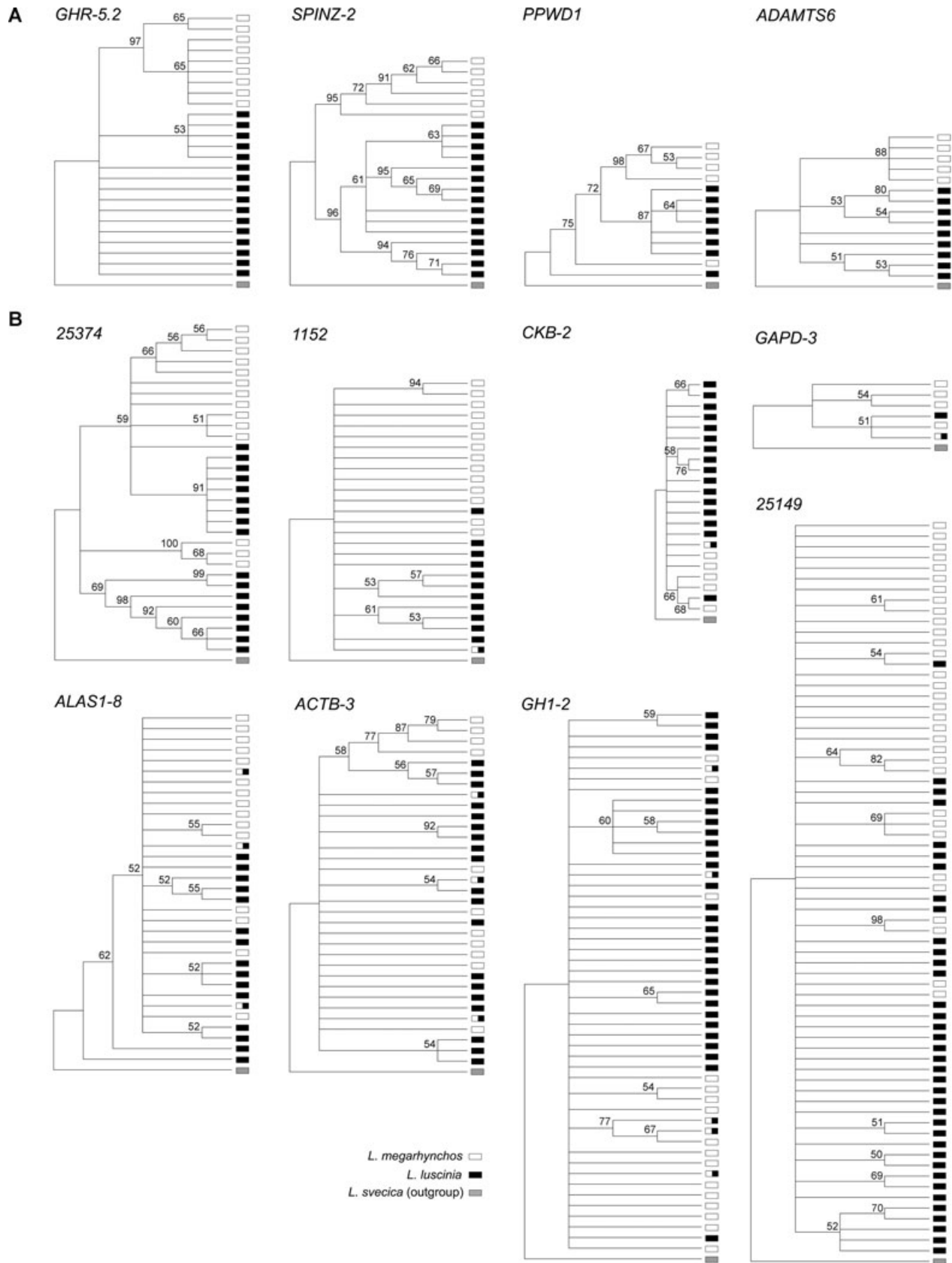
To further explore patterns of interspecific gene flow and to estimate divergence time as well as effective population sizes for both species, we fitted an IM model to our data using the IMA program. The program was run separately for (1) all 12 loci, (2) the eight autosomal loci and (3) the four Z-linked loci. Maximum-likelihood estimates (MLE) and 90% highest posterior density (HPD) intervals of all model parameters are given in Table 4, and posterior distributions are shown in Figure 3. For the whole dataset, the estimated  $N_e$  was 448,584 for *L. megarhynchos*, 853,268 for *L. luscinia*, and 220,888 for the ancestral population. Similar estimates were obtained for autosomal and Z-linked loci, although estimates of  $N_e$  for Z-linked genes were slightly lower.  $N_e$  of the ancestral population could not be estimated accurately for the Z-linked loci due to the absence of shared polymorphisms.

We estimated the time since divergence between the two species to be 1.8 Mya for the whole dataset. Interestingly, the estimated divergence time was two times higher for Z-linked loci (2.6 Mya) than for autosomal loci (1.3 Mya). This discrepancy might be partly explained by an underestimation of  $N_e$  of the ancestral population for Z-linked loci due to the absence of shared polymorphisms. However, because estimates of  $N_e$  of the ancestral population from autosomal loci are still quite low, the interesting possibility remains that reproductive isolation and thus species divergence occurred earlier on the Z chromosome than on the autosomes.

Both migration rate parameters revealed evidence of gene flow between the species. Gene flow was generally higher from *L. megarhynchos* to *L. luscinia* ( $2N_2m_2 = 0.326$ ), than in the opposite direction ( $2N_1m_1 = 0.119$ ). The estimates of gene flow for autosomal loci were quite high in both directions. In contrast, no gene flow was detected on the Z chromosome from *L. megarhynchos* to *L. luscinia* and limited gene flow was detected in the opposite direction (corresponding to the introgression of only one haplotype at *PPWD1*; Fig. 2).

To test whether models with gene flow are a significantly better fit to the data than models without gene flow, we performed log-likelihood ratio tests of nested models (Hey and Nielsen 2007) as implemented in the IMA program, where migration parameters,  $m_1$  and/or  $m_2$ , were set to zero (Table 5). For the whole dataset, a model with gene flow was a significantly better fit to the data than all models without gene flow. For autosomal loci, a model with gene flow was significantly better than a model without gene flow in both directions. It was also significantly better than a model in which gene flow occurs only from *L. luscinia* into *L. megarhynchos*, however, it was not significantly better than a model with gene flow only in the opposite direction. In contrast, for Z-linked loci, a model with gene flow was not a significantly better fit to the data than any model without gene flow. Finally, we tested whether a model with asymmetrical introgression is better than models where gene flow is the same in both directions. This test, however, did not yield significant results in any of the comparisons (Table 5).

We also estimated migration rates separately for each locus using the IM program (Table 6). Not all of these estimates were recovered with high confidence, either because the tails of the posterior probability densities did not reach zero or because the 90% HPD intervals were quite large (the posterior probability distributions for locus-specific migration rates are shown in Fig. S2). Nonetheless, the estimates were remarkably consistent with classic estimators of gene flow ( $F_{ST}$ -based estimates of  $Nm$ ), suggesting that introgression is the major cause of the contrasting patterns of differentiation between the Z chromosome and the autosomes. Average migration rates for Z-linked



**Figure 2.** Neighbor-joining trees for four Z-linked (A) and eight autosomal (B) loci. The nodes that do not have majority bootstrap support have been compressed. Because recombination events were detected for most loci, the trees represent overall similarity between the haplotypes; portions of each gene may have slightly different genealogies.



**Table 3.** Linkage disequilibrium test of gene introgression. Observed and simulated values of the  $x$  statistics counted as difference between LD among pairs of shared polymorphism (DSS) and LD among pairs of exclusive polymorphism (DXX). Only genes with more than four shared and four exclusive polymorphisms in at least one subsample are presented. No Z-linked loci had sufficient segregating sites for the analysis. NA, not applicable, because observed values are shown only if both DSS and DXX were calculated from at least four pairs of sites in a particular subsample (see Methods). \* $P < 0.05$ , \*\* $P < 0.01$ .

Locus	Chr	Subsample	<i>L. megarhynchos</i>				<i>L. luscinia</i>			
			Observed			Simulated	Observed			Simulated
			DSS	DXX	$x = \text{DSS} - \text{DXX}$	$x = \text{DSS} - \text{DXX}$	DSS	DXX	$x = \text{DSS} - \text{DXX}$	$x = \text{DSS} - \text{DXX}$
25149	A	1	0.044	-0.438	0.482	0.103	0.171	-0.489	0.660	0.124*
		2	0.131	-0.470	0.601	0.028*	0.022	-0.232	0.254	0.055
		3	0.284	-0.425	0.709	0.103*	-0.091	-0.607	0.516	0.116*
25374	A	1	NA	NA	NA	NA	NA	NA	NA	NA
		2	0.074	-0.556	0.630	0.040*	0.083	-0.308	0.391	0.082
		3	NA	NA	NA	NA	NA	NA	NA	NA
ACTB-3	A	1	-0.160	-0.048	-0.112	0.133	-0.067	-0.294	0.227	0.154
		2	0.104	0.200	-0.096	0.033	-0.333	-0.714	0.381	0.077
		3	0.504	-0.048	0.552	0.105	0.644	-0.556	1.200	0.141**
ALASI-8	A	1	NA	NA	NA	NA	NA	NA	NA	NA
		2	NA	NA	NA	NA	NA	NA	NA	NA
		3	-0.190	-0.666	0.476	0.095	-0.810	-0.464	-0.346	0.128
GHI-2	A	1	-0.015	-0.600	0.585	0.098	-0.010	-0.273	0.263	0.146
		2	-0.095	NA	NA	NA	-0.051	-0.302	0.251	0.069
		3	0.054	-0.200	0.254	0.095	0.401	-0.530	0.931	0.126*

loci were significantly lower than for autosomal loci (Mann-Whitney U test,  $P = 0.021857$ ). This difference was caused mainly by introgression from *L. megarhynchos* into *L. luscinia* (Table 6).

### Discussion

To study the role of the Z chromosome in speciation, we analyzed patterns of polymorphism and divergence at eight autosomal and four Z-linked loci in allopatric populations of two nightingale

**Table 4.** Maximum-likelihood estimates (MLE) and 90% highest posterior density (HPD) intervals of demographic parameters.

Comparison	$N_1^1$	$N_2^1$	$N_A^1$	$m_1^2$	$m_2^3$	$2N_1m_1^4$	$2N_2m_2^5$	$t^6$
All genes								
MLE	448,584	853,268	220,888	1.32E-07	1.91E-07	0.119	0.326	1,844,605
Lower 90% HPD	338,611	671,354	4,318	1.45E-08	9.27E-08			1,034,130
Higher 90% HPD	581,698	1,069,838	524,096	3.27E-07	3.37E-07			2,420,352
Autosomes								
MLE	540,786	889,687	218,616	7.52E-08	4.29E-07	0.081	0.763	1,254,893
Lower 90% HPD	381,078	645,670	5,074	8.08E-10	2.09E-07			720,264
Higher 90% HPD	729,515	1,204,895	503,319	3.84E-07	7.57E-07			1,918,229
Z chromosome								
MLE	303,003	820,407	3,124	5.53E-08	1.13E-09	0.034	0.002	2,591,343
Lower 90% HPD	190,917	589,566	3,124	1.13E-09	1.13E-09			1,407,337
Higher 90% HPD	471,143	1,126,113	2,473,718	2.65E-07	7.33E-08			4,292,909

<sup>1</sup>Effective population size of *L. megarhynchos* ( $N_1$ ), *L. luscinia* ( $N_2$ ) and the ancestral population ( $N_A$ ) in numbers of individuals.

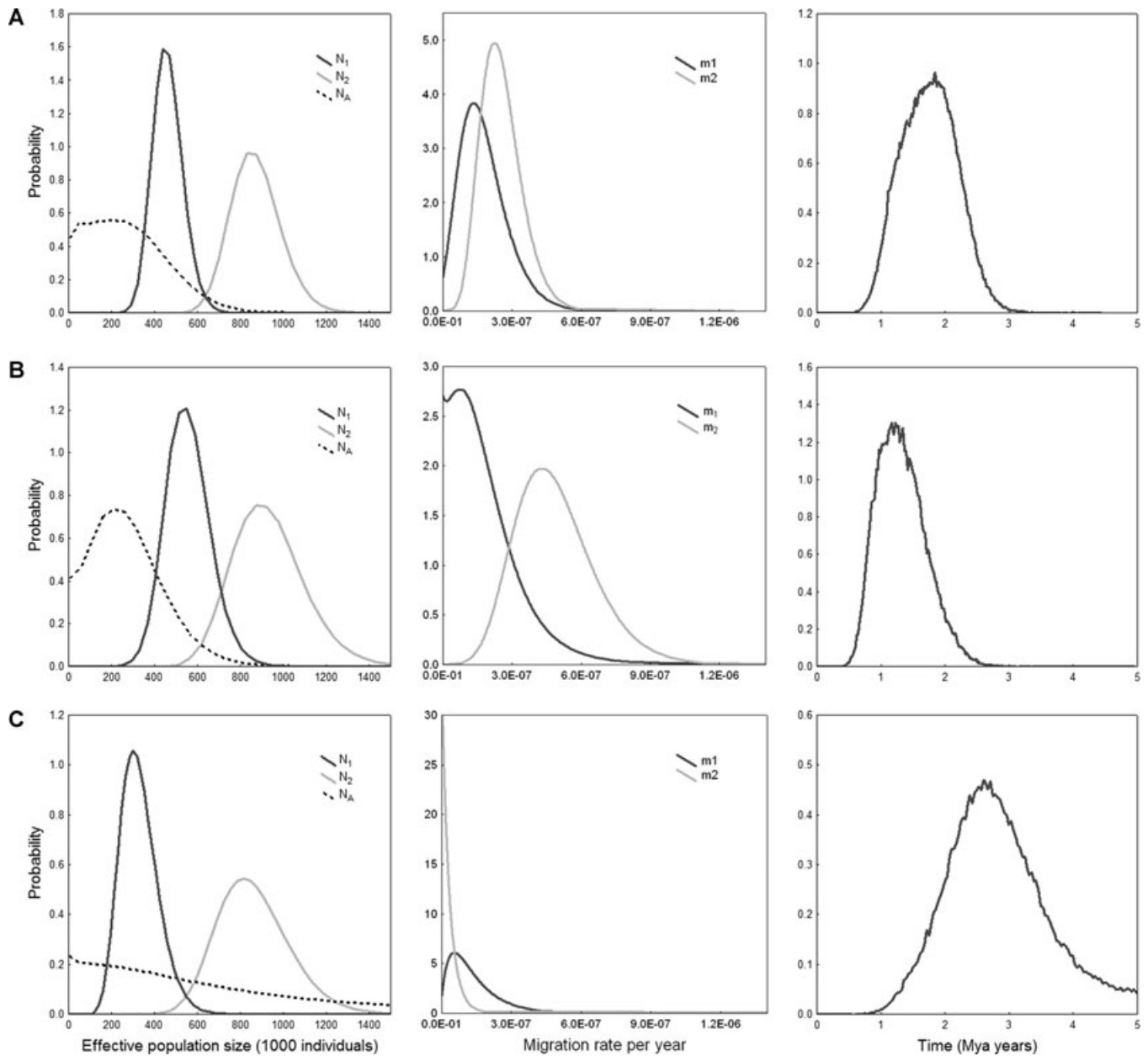
<sup>2</sup>Migration rate per year from *L. luscinia* into *L. megarhynchos*.

<sup>3</sup>Migration rate per year from *L. megarhynchos* into *L. luscinia*.

<sup>4</sup>The population migration rate from *L. luscinia* into *L. megarhynchos*.

<sup>5</sup>The population migration rate from *L. megarhynchos* into *L. luscinia*.

<sup>6</sup>The time since population split in number of years.



**Figure 3.** The marginal posterior probability distributions for the demographic parameters of the IM model. Effective population size is shown for *L. megarhynchos* ( $N_1$ ), *L. luscinia* ( $N_2$ ), and ancestral population ( $N_A$ ). Migration rates are from *L. luscinia* to *L. megarhynchos* ( $m_1$ ), and from *L. megarhynchos* to *L. luscinia* ( $m_2$ ). Distributions are shown for all genes (A), autosomal genes (B), and Z-linked genes (C).

species. Four main results stem from our study. (1) We estimated that the two species diverged approximately 1.8 Mya and have large effective population sizes (*L. megarhynchos*  $N_e = 448,584$ , *L. luscinia*  $N_e = 853,268$ ). (2) We observed strikingly different patterns of polymorphism and divergence between the Z chromosome and the autosomes. Z-linked loci showed slightly, but not significantly, lower levels of variation within species compared to autosomes. On the other hand, the Z chromosome was characterized by significantly higher levels of differentiation between

species than were the autosomes. This was seen in levels of  $F_{ST}$ , net nucleotide divergence, and in the ratio of fixed differences to shared polymorphisms, all of which were higher for Z-linked loci than for autosomal loci. (3) We found strong evidence of gene flow following divergence of the two species. This gene flow occurred in both directions, although more introgression was from *L. megarhynchos* into *L. luscinia*. (4) Gene flow was significantly higher on the autosomes than on the Z chromosome. Below, we discuss our results in light of recent theories of speciation, with

**Table 5.** Log-likelihood ratio tests of nested models.

Models <sup>1</sup>	log(P) <sup>2</sup>	2LLR <sup>3</sup>	df <sup>4</sup>	P <sup>5</sup>
All genes				
m1, m2=0	-21.502	52.3115	1	<0.0000**
m1=0, m2	1.4405	6.4266	1	0.0112*
m1=0, m2=0	-72.2962	153.8999	2	<0.0000**
m1=m2	4.4378	0.4321	1	0.5110
Autosomes				
m1, m2=0	-18.4623	43.8713	1	<0.0000**
m1=0, m2	1.6533	3.64	1	0.0564
m1=0, m2=0	-98.1575	203.2617	2	<0.0000**
m1=m2	1.9105	3.1257	1	0.0771
Z chromosome				
m1, m2=0	4.2341	0.0099	1	0.9207
m1=0, m2	2.6578	3.1624	1	0.0754
m1=0, m2=0	2.657	3.164	2	0.2056
m1=m2	3.2204	2.0372	1	0.1535

<sup>1</sup>Four nested models with different patterns of gene flow were compared to the estimated full model: (1) model in which migration from *L. megarhynchos* to *L. luscinia* is equal to 0 (m1, m2=0), (2) model in which migration from *L. luscinia* to *L. megarhynchos* is equal to 0 (m1=0, m2), (3) model without migration in both directions (m1=0, m2=0), and (4) model in which migration rates are equal in both directions (m1=m2). The effective population sizes were same as for the full model.

<sup>2</sup>Estimates of the posterior density function under the full model.

<sup>3</sup>Log-likelihood ratio statistics calculated as the difference between the highest posterior probability for the full model and the highest posterior probability for the nested model.

<sup>4</sup>The degrees of freedom. Models in which migration at least in one direction is equal to 0 have distributions of 2LLR that are a mixture and  $\chi^2$  mixture distribution was thus used to calculate the *P*-value (see Hey and Nielsen 2007).

<sup>5</sup>The probability of achieving the test statistics by chance under the null model (\**P*<0.05, \*\**P*<0.01).

respect to the large effect of the sex chromosomes in reproductive isolation and adaptive evolution.

## HISTORY OF SPECIATION IN NIGHTINGALES

The Common Nightingale and the Thrush Nightingale are thought to have diverged in allopatry during the Pleistocene climatic oscillations (Sorjonen 1986) which started approximately 2.4–2.6 Mya (Hewitt 2000; Provan and Bennett 2008). Our estimated divergence time of 1.8 Mya is consistent with this idea. However, it is important to bear in mind that estimates of divergence time depend on estimates of mutation rate, which should be considered approximate owing to the uncertainty of divergence time to the outgroup.

Interestingly, previous estimates based on mitochondrial cytochrome *b* sequences were considerably higher (3.3 Mya; Wink et al. 2002). The high estimated divergence times from mitochondrial sequences in nightingales and other species were used to question the importance of Pleistocene climatic oscillations in the

**Table 6.** Maximum-likelihood estimates of population migration rates for eight autosomal and four Z-linked loci.

Locus	Chr	$2N_1m_1^1$	$2N_2m_2^2$	Av. $2Nm^3$	$Nm^4$
<i>CKB-2</i>	A	0.003	0.529	0.266	0.220
<i>ACTB-3</i>	A	0.025	2.737	1.381	2.559
<i>GAPD-3</i>	A	0.523	0.007	0.265	0.473
<i>ALAS1-8</i>	A	0.003	0.007	0.005	1.256
<i>GHI-2</i>	A	0.115	3.574	1.844	0.832
<i>25374</i>	A	0.003	0.268	0.135	0.340
<i>1152</i>	A	0.545	0.124	0.335	0.422
<i>25149</i>	A	0.691	0.346	0.519	0.861
<b>Average</b>	<b>A</b>	<b>0.238</b>	<b>0.949</b>	<b>0.594</b>	<b>0.589</b>
<i>GHR-5.2</i>	Z	0.003	0.007	0.005	0.097
<i>PPWD1</i>	Z	0.143	0.007	0.075	0.079
<i>ADAMTS6</i>	Z	0.003	0.007	0.005	0.101
<i>SPINZ-2</i>	Z	0.003	0.007	0.005	0.094
<b>Average</b>	<b>Z</b>	<b>0.038</b>	<b>0.007</b>	<b>0.022</b>	<b>0.092</b>

<sup>1</sup>Estimates of population migration rate from *L. luscinia* to *L. megarhynchos* using the IM software.

<sup>2</sup>Estimates of population migration rate from *L. megarhynchos* to *L. luscinia* using the IM software.

<sup>3</sup>Average population migration rate in both directions using estimates from the IM software.

<sup>4</sup>Estimates of population migration rate calculated from the expression  $F_{ST}=1/(1+4Nm)$  for autosomal and  $F_{ST}=1/(1+3Nm)$  for Z-linked loci.

origin of many sister species of birds (Klicka and Zink 1997). However, single-locus estimates of age can be unreliable, and if mitochondrial divergence times have generally been overestimated, as seems to have been the case in nightingales, we may have to re-evaluate the importance of ice ages in bird speciation (Weir and Schluter 2004).

Interestingly, our estimate of divergence time was much higher for the Z chromosome (2.6 Mya) than for the autosomes (1.3 Mya). This might indicate that the two species did not diverge in strict allopatry, but exchanged genes during the speciation process either in sympatry or in parapatry. Under such a scenario, divergence times might vary across the genome, because interspecific gene flow is expected to be restricted earlier in genomic regions harboring loci causing hybrid incompatibilities than in genomic regions without such loci (Wu 2001). Indeed, we found strong evidence of post divergence gene flow between the nightingale species. This gene flow was significantly higher on the autosomes compared to the Z chromosome, which is consistent with lower estimates of divergence time for the autosomes. Glacial fluctuations were a major factor in shaping bird distributions during the Quaternary. It is possible that nightingales underwent repeated cycles of allopatry followed by periods of gene flow during interglacial events.

Although we do not know the exact position of the four analyzed Z-linked genes in nightingales, *PPWD1* and *ADAMTS6*

are physically close to each other in both the chicken and in the zebra finch, and thus they may also be physically close in nightingales. The other two Z-linked loci are farther apart in both chicken and zebra finch, and thus may lie in different parts of the Z chromosome in nightingales (Table S3). Loci that are closely linked may have correlated histories, whereas loci that are far apart can provide independent information about the levels of divergence and gene flow. Genetic linkage can be caused not only by close physical proximity of genes on the chromosome, but also by chromosomal rearrangements, such as inversions. We cannot exclude the possibility that some chromosomal rearrangements differ between the nightingale species and might result in linkage of multiple Z-linked loci across a large part of the Z chromosome. However, this seems unlikely given the low frequency of karyotypic changes in birds (Price 2007) and the young age of these nightingale species.

### LARGE EFFECT OF THE Z CHROMOSOME IN SPECIATION

The sex chromosomes are known to play a prominent role in reproductive isolation in species with heterogametic males (e.g., Coyne and Orr 1989; Masly and Presgraves 2007). Here, we show that in species with heterogametic females, the Z chromosome is also likely to be important in reproductive isolation. We found that Z-linked loci systematically showed reduced interspecific introgression compared to autosomal loci. Similar results have been demonstrated in two other bird systems. Genetic analysis of a sympatric population of two naturally hybridizing *Ficedula* flycatcher species revealed that interspecific gene flow occurs at autosomal, but not at Z-linked loci (Sætre et al. 2003). Recently, reduced introgression of the Z chromosome has also been observed across the hybrid zone between Indigo and Lazuli Bunting (Carling and Brumfield 2008, 2009). Involvement of the Z chromosome in reproductive isolation has also been found between different species of *Heliconius* butterflies (Jiggins et al. 2001; Naisbit et al. 2002). Together, these results indicate that the importance of the Z chromosome in speciation might represent a general phenomenon.

The Z chromosome could play a large role in postzygotic as well as prezygotic isolation (Qvarnström and Bailey 2008). The Z chromosome could underlie postzygotic barriers, such as hybrid sterility and inviability, due to the exposure of recessive mutations in the hemizygous individuals or due to a higher density of genes contributing to reproductive isolation on the Z. The higher density of isolation genes on Z may be caused by faster Z evolution, accumulation of meiotic drive elements on the Z or disruption of meiotic Z chromosome inactivation that has been recently reported in birds (Schoenmakers et al. 2009). The meiotic drive hypothesis for the large-Z effect may be unlikely because sex ratio distorters are less likely to evolve on the Z chromosome

compared to the X chromosome because male-biased populations are especially prone to extinction (Tao et al. 2007). All these mechanisms lead to preferential occurrence of hybrid sterility or inviability in the heterogametic sex (Coyne and Orr 2004). Nightingales (Stadie 1991), *Ficedula* flycatchers (Gelter et al. 1992) as well as *Heliconius* butterflies (Jiggins et al. 2001) all show female-limited hybrid sterility, supporting the idea that the Z chromosome is involved in the evolution of intrinsic postzygotic isolation.

On the other hand, the Z chromosome could promote the evolution of prezygotic isolation through sexual selection and/or reinforcement. Sexual selection may facilitate speciation because it can cause rapid divergence of male-mating signals and female preferences (Ritchie 2007). According to the theory of reinforcement, natural selection against unfit hybrids could enhance prezygotic isolation in areas in which two species coexist (Noor 1999). Both theoretical and empirical studies suggest that sexual selection as well as reinforcement are facilitated in species with female heterogamety if genes for male traits and female preferences are Z-linked (Hastings 1994; Reeve and Pfennig 2003; Servedio and Sætre 2003; Kirkpatrick and Hall 2004; Albert and Otto 2005; Hall and Kirkpatrick 2006). Prezygotic isolation plays an important role in nightingale speciation as shown by strong assortative mating between species and the low frequency of interspecific hybrids. Both the Common Nightingale and the Thrush Nightingale have dull brown plumage without any differences between the sexes. Sexual selection based on plumage color is thus unlikely to contribute to speciation. However, nightingales are known for their outstanding song, which differs slightly between the species, and might be under sexual selection (Cramp 1988). If reinforcement is responsible for the evolution of prezygotic isolation, we would expect more pronounced interspecific differences in male-mating signals in sympatry than in allopatry. This has been shown for example in *Ficedula* flycatchers, where sympatric males show increased differences in plumage compared to allopatric males (Sætre et al. 1997). Empirical observations in nightingales, however, show just the opposite pattern; songs of both nightingale species tend to converge in sympatry (Sorjonen 1986; Becker 2007). This indicates that, contrary to the situation in *Ficedula* flycatchers, reinforcement is unlikely to contribute to the evolution of prezygotic isolation in nightingales. It remains unclear, however, whether divergence of male song and female preferences in allopatry might contribute to reduced introgression of the Z chromosome.

The greater differentiation observed on the Z chromosome relative to the autosomes could, in principle, be explained by reduced introgression on the Z (e.g., Sætre et al. 2003) or positive selection that increases interspecific divergence by eliminating shared variants (e.g., Borge et al. 2005). We sought to disentangle these two effects by using tests for positive selection and by using

several models that explicitly test for postdivergence gene flow. Neither the HKA test (Hudson et al. 1987) nor tests based on the allele frequency spectrum (Tajima 1989) provided evidence for positive selection on the Z. In contrast, substantial gene flow on the autosomes but not on the Z was revealed by divergence-with-gene-flow models (Nielsen and Wakeley 2001; Hey and Nielsen 2004, 2007). Thus, these data provide no evidence for recent strong selection on the Z, although we cannot rule out weak or older selection. On the other hand, the absence of gene flow on the Z in the face of substantial gene flow on the autosomes is consistent with the presence of Z-linked genes contributing to reproductive isolation. The fact that these species are separated by hybrid female sterility and show Z-linked genes contributing to isolation is analogous to X-linked male sterility in species with heterogametic males. This highlights the similarities in the genetic basis of reproductive isolation between organisms with heterogametic males and organisms with heterogametic females during the early stages of speciation.

#### ACKNOWLEDGMENTS

We are grateful to M. Ahola, M. Antczak, D. B. Campas, A. Cabrera, P. A. Del Baño, P. Dołata, K. L. Evans, P. Kverek, T. Polo Aparisi, and J. Vokurková for help with the sample collection. V. Pavel kindly provided the sample from the Bluethroat. We thank members of the Nachman lab, especially A. Geraldes and M. Carneiro, for many useful discussions. J. Good, A. Geraldes, M. Sans, D. Storch, J. Lindell, and two anonymous reviewers provided critical comments on early versions of the manuscript. The research was supported by junior grant GAAV (KJB501110701) and post-doctoral grant GACR (206/08/P160) to RS, by the Ministry of Education, Youth and Sport of the Czech Republic (MSM0021620828), by a Fulbright fellowship to RS, and by NSF and NIH grants to MWN.

#### LITERATURE CITED

- Albert, A. Y., and S. P. Otto. 2005. Sexual selection can resolve sex-linked sexual antagonism. *Science* 310:119–121.
- Andrés, A. M., A. G. Clark, L. Shimmin, E. Boerwinkle, C. F. Sing, and J. E. Hixson. 2007. Understanding the accuracy of statistical haplotype inference with sequence data of known phase. *Genet. Epidemiol.* 31:659–671.
- Backström, N., M. Brandström, L. Gustafsson, A. Qvarnström, H. Cheng, and H. Ellegren. 2006. Genetic mapping in a natural population of collared flycatchers (*Ficedula albicollis*): conserved synteny but gene order rearrangements on the avian Z chromosome. *Genetics* 174:377–386.
- Backström, N., S. Fagerberg, and H. Ellegren. 2008. Genomics of natural bird populations: a gene-based set of reference markers evenly spread across the avian genome. *Mol. Ecol.* 17:964–980.
- Baines, J. F., S. A. Sawyer, D. L. Hartl, and J. Parsch. 2008. Effects of X-linkage and sex-biased gene expression on the rate of adaptive protein evolution in *Drosophila*. *Mol. Biol. Evol.* 25:1639–1650.
- Becker, J. 2007. About Nightingales (*Luscinia megarhynchos*), Trush Nightingales (*Luscinia luscinia*) and their hybrids—further results of an investigation via bird ringing in the Frankfurt (Oder) area. *Vogelwarte* 45:15–26.
- Betancourt, A. J., D. C. Presgraves, and W. J. Swanson. 2002. A test for faster X evolution in *Drosophila*. *Mol. Biol. Evol.* 19:1816–1819.
- Borge, T., M. T. Webster, G. Andersson, and G. P. Sætre. 2005. Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes in two *Ficedula* flycatcher species. *Genetics* 171:1861–1873.
- Carling, M. D., and R. T. Brumfield. 2008. Haldane's rule in an avian system: using cline theory and divergence population genetics to test for differential introgression of mitochondrial, autosomal, and sex-linked loci across the *Passerina* bunting hybrid zone. *Evolution* 62:2600–2615.
- . 2009. Speciation in *Passerina* buntings: introgression patterns of sex-linked loci identify a candidate gene region for reproductive isolation. *Mol. Ecol.* 18:834–847.
- Carneiro, M., N. Ferrand, and M. W. Nachman. 2009. Recombination and speciation: loci near centromeres are more differentiated than loci near telomeres between subspecies of the European rabbit (*Oryctolagus cuniculus*). *Genetics* 181:593–606.
- Connallon, T. 2007. Adaptive protein evolution of X-linked and autosomal genes in *Drosophila*: implications for faster-X hypotheses. *Mol. Biol. Evol.* 24:2566–2572.
- Counterman, B. A., D. Ortiz-Barrientos, and M. A. Noor. 2004. Using comparative genomic data to test for fast-X evolution. *Evolution* 58:656–660.
- Coyne, J. A., and H. A. Orr. 1989. Two rules of speciation. Pp. 180–207 in D. Otte and J. Endler, eds. *Speciation and its consequences*. Sinauer Associates, Sunderland, MA.
- . 2004. *Speciation*. Sinauer Associates, Sunderland, MA.
- Cramp, S. 1988. *The Birds of the Western Palearctic*. Oxford Univ. Press, Oxford, UK.
- Ellegren, H. 2009. Genomic evidence for a large-Z effect. *Proc. R. Soc. Lond. B* 276:361–366.
- Ewing, B., and P. Green. 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. *Genome Res.* 8:186–194.
- Ewing, B., L. Hillier, M. C. Wendl, and P. Green. 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* 8:175–185.
- Gelter, H. P., H. Tegelström, and L. Gustafsson. 1992. Evidence from hatching success and DNA fingerprinting for the fertility of hybrid Pied x Collared Flycatchers *Ficedula hypoleuca* x *albicollis*. *Ibis* 134:62–68.
- Good, J. M., M. D. Dean, and M. W. Nachman. 2008. A complex genetic basis to X-linked hybrid male sterility between two species of house mice. *Genetics* 179:2213–2228.
- Gordon, D., C. Abajian, and P. Green. 1998. Consed: a graphical tool for sequence finishing. *Genome Res.* 8:195–202.
- Hall, T. A. 1999. BioEdit: a user friendly biological sequence alignment editor and analyses program for Windows 95/98/NT. *Nucleic Acids Symp. Series*:95–98.
- Hall, D. W., and M. Kirkpatrick. 2006. Reinforcement and sex linkage. *Evolution* 60:908–921.
- Hastings, I. M. 1994. Manifestations of sexual selection may depend on the genetic bases of sex determination. *Proc. R. Soc. Lond. B* 258:83–87.
- Hewitt, G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405:907–913.
- Hey, J., and R. Nielsen. 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167:747–760.
- . 2007. Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc. Natl. Acad. Sci. USA* 104:2785–2790.
- Hey, J., and J. Wakeley. 1997. A coalescent estimator of the population recombination rate. *Genetics* 145:833–846.
- Hudson, R. R., M. Kreitman, and M. Aguade. 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics* 116:153–159.

- Hudson, R. R., D. D. Boos, and N. L. Kaplan. 1992. A statistical test for detecting geographic subdivision. *Mol. Biol. Evol.* 9:138–151.
- Itoh, Y., K. Kampf, and A. P. Arnold. 2006. Comparison of the chicken and zebra finch Z chromosomes shows evolutionary rearrangements. *Chromosome Res.* 14:805–815.
- Iyengar, V. K., H. K. Reeve, and T. Eisner. 2002. Paternal inheritance of a female moth's mating preference. *Nature* 419:830–832.
- Jiggins, C. D., M. Linares, R. E. Naisbit, C. Salazar, Z. H. Yang, and J. Mallet. 2001. Sex-linked hybrid sterility in a butterfly. *Evolution* 55:1631–1638.
- Jönsson, K. A., and J. Fjeldså. 2006. A phylogenetic supertree of oscine passerine birds (Aves: Passeri). *Zool. Scripta* 35:149–186.
- Kirkpatrick, M., and D. W. Hall. 2004. Sexual selection and sex linkage. *Evolution* 58:683–691.
- Klicka, J., and R. M. Zink. 1997. The importance of recent ice ages in speciation: a failed paradigm. *Science* 277:1666–1669.
- Kverek, P., R. Storchová, J. Reif, and M. W. Nachman. 2008. Occurrence of a hybrid between the Common Nightingale (*Luscinia megarhynchos*) and the Thrush Nightingale (*Luscinia luscinia*) in the Czech Republic confirmed by genetic analysis. *Sylvia* 44:17–26.
- Lewontin, R. C. 1964. The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics* 49:49–67.
- Li, N., and M. Stephens. 2003. Modelling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics* 165:2213–2233.
- Machado, C. A., R. M. Kliman, J. A. Markert, and J. Hey. 2002. Inferring the history of speciation from multilocus DNA sequence data: the case of *Drosophila pseudoobscura* and close relatives. *Mol. Biol. Evol.* 19:472–488.
- Macholán, M., P. Munclinger, M. Šugerková, P. Dufková, B. Břimová, E. Božíková, J. Zima, and J. Piálek. 2007. Genetic analysis of autosomal and X-linked markers across a mouse hybrid zone. *Evolution* 61:746–771.
- Mank, J. E., E. Axelsson, and H. Ellegren. 2007. Fast-X on the Z: rapid evolution of sex-linked genes in birds. *Genome Res.* 17:618–624.
- Masly, J. P., and D. C. Presgraves. 2007. High-resolution genome-wide dissection of the two rules of speciation in *Drosophila*. *PLoS Biol.* 5:e243.
- Moriyama, E. N., and J. R. Powell. 1996. Intraspecific nuclear DNA variation in *Drosophila*. *Mol. Biol. Evol.* 13:261–277.
- Naisbit, R. E., C. D. Jiggins, M. Linares, C. Salazar, and J. Mallet. 2002. Hybrid sterility, Haldane's rule and speciation in *Heliconius cydno* and *H. melpomene*. *Genetics* 161:1517–1526.
- Nanda, I., K. Schlegelmilch, T. Haaf, M. Schartl, and M. Schmid. 2008. Synteny conservation of the Z chromosome in 14 avian species (11 families) supports a role for Z dosage in avian sex determination. *Cytogenet. Genome Res.* 122:150–156.
- Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia Univ. Press, New York, NY.
- Nickerson, D. A., V. O. Tobe, and S. L. Taylor. 1997. PolyPhred: automating the detection and genotyping of single nucleotide substitutions using fluorescence-based resequencing. *Nucleic Acids Res.* 25:2745–2751.
- Nielsen, R., and J. Wakeley. 2001. Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics* 158:885–896.
- Noor, M. A. 1999. Reinforcement and other consequences of sympatry. *Heredity* 83:503–508.
- Oka, A., A. Mita, N. Sakurai-Yamatani, H. Yamamoto, N. Takagi, T. Takano-Shimizu, K. Toshimori, K. Moriwaki, and T. Shiroishi. 2004. Hybrid breakdown caused by substitution of the X chromosome between two mouse subspecies. *Genetics* 166:913–924.
- Payseur, B. A., J. G. Krenz, and M. W. Nachman. 2004. Differential patterns of introgression across the X chromosome in a hybrid zone between two species of house mice. *Evolution* 58:2064–2078.
- Pigozzi, M. I. 2008. Relationship between physical and genetic distances along the zebra finch Z chromosome. *Chromosome Res.* 2008:839–849.
- Presgraves, D. C. 2008. Sex chromosomes and speciation in *Drosophila*. *Trends Genet.* 24:336–343.
- Price, T. 2007. *Speciation in birds*. Roberts & Company Publishers, Greenwood Village, CO.
- Provan, J., and K. D. Bennett. 2008. Phylogeographic insights into cryptic glacial refugia. *Trends Ecol. Evol.* 23:564–571.
- Qvarnström, A., and R. I. Bailey. 2008. Speciation through evolution of sex-linked genes. *Heredity* 102:4–15.
- Reeve, H. K., and D. W. Pfennig. 2003. Genetic biases for showy males: are some genetic systems especially conducive to sexual selection? *Proc. Natl. Acad. Sci. USA* 100:1089–1094.
- Ritchie, M. G. 2007. Sexual selection and speciation. *Annu. Rev. Ecol. Evol. Syst.* 38:79–102.
- Rozas, J., J. C. Sánchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496–2497.
- Sæther, S. A., G. P. Sætre, T. Borge, C. Wiley, N. Svedin, G. Andersson, T. Veen, J. Haavie, M. R. Servedio, S. Bureš, et al. 2007. Sex chromosome-linked species recognition and evolution of reproductive isolation in flycatchers. *Science* 318:95–97.
- Sætre, G. P., T. Moum, S. Bureš, M. Král, M. Adamjan, and J. Moreno. 1997. A sexually selected character displacement in flycatchers reinforces premating isolation. *Nature* 387:589–592.
- Sætre, G. P., T. Borge, K. Lindroos, J. Haavie, B. C. Sheldon, C. Primmer, and A. C. Syvänen. 2003. Sex chromosome evolution and speciation in *Ficedula* flycatchers. *Proc. R. Soc. Lond. B* 270:53–59.
- Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406–425.
- Schoenmakers, S., E. Wassenaar, J. W. Hoogerbrugge, J. S. Laven, J. A. Grootegeod, and W. M. Baarends. 2009. Female meiotic sex chromosome inactivation in chicken. *PLoS Genet.* 5:e1000466.
- Servedio, M. R., and G. P. Sætre. 2003. Speciation as a positive feedback loop between postzygotic and prezygotic barriers to gene flow. *Proc. R. Soc. Lond. B* 270:1473–1479.
- Sorjonen, J. 1986. Mixed singing and interspecific territoriality—consequences of secondary contact of 2 ecologically and morphologically similar nightingale species in Europe. *Ornis Scand.* 17:53–67.
- Stadie, C. 1991. Erdsanger I; Nachtigall und Sprosser. *Europäische Vogelwelt* 3, Sonderheft:130–189.
- Stephens, M., and P. Donnelly. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *Am. J. Hum. Genet.* 73:1162–1169.
- Stephens, M., N. J. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68:978–989.
- Storchová, R., S. Gregorová, D. Buckiová, V. Kyselová, P. Divina, and J. Forejt. 2004. Genetic analysis of X-linked hybrid sterility in the house mouse. *Mamm. Genome* 15:515–524.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.
- Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol. Biol. Evol.* 24:1596–1599.
- Tao, Y., S. Chen, D. L. Hartl, and C. C. Laurie. 2003. Genetic dissection of hybrid incompatibilities between *Drosophila simulans* and *D. mauritiana*. I. Differential accumulation of hybrid male sterility effects on the X and autosomes. *Genetics* 164:1383–1397.

- Tao, Y., L. Araripe, S. B. Kingan, Y. Ke, H. Xiao, and D. L. Hartl. 2007. A sex-ratio meiotic drive system in *Drosophila simulans*. II: an X-linked distorter. *PLoS Biol.* 5:e303.
- Teeter, K. C., B. A. Payseur, L. W. Harris, M. A. Bakewell, L. M. Thibodeau, J. E. O'Brien, J. G. Krenz, M. A. Sans-Fuentes, M. W. Nachman, and P. K. Tucker. 2008. Genome-wide patterns of gene flow across a house mouse hybrid zone. *Genome Res.* 18:67–76.
- Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673–4680.
- Thornton, K., and M. Long. 2002. Rapid divergence of gene duplicates on the *Drosophila melanogaster* X chromosome. *Mol. Biol. Evol.* 19:918–925.
- Torgerson, D. G., and R. S. Singh. 2003. Sex-linked mammalian sperm proteins evolve faster than autosomal ones. *Mol. Biol. Evol.* 20:1705–1709.
- True, J. R., B. S. Weir, and C. C. Laurie. 1996. A genome-wide survey of hybrid incompatibility factors by the introgression of marked segments of *Drosophila mauritiana* chromosomes into *Drosophila simulans*. *Genetics* 142:819–837.
- Turelli, M., and D. J. Begun. 1997. Haldane's rule and X-chromosome size in *Drosophila*. *Genetics* 147:1799–1815.
- Wahlberg, P., L. Strömstedt, X. Tordoir, M. Foglio, S. Heath, D. Lechner, A. R. Hellström, M. Tixier-Boichard, M. Lathrop, I. G. Gut, et al. 2007. A high-resolution linkage map for the Z chromosome in chicken reveals hot spots for recombination. *Cytogenet. Genome Res.* 117:22–29.
- Wang, R. L., J. Wakeley, and J. Hey. 1997. Gene flow and natural selection in the origin of *Drosophila pseudoobscura* and close relatives. *Genetics* 147:1091–1106.
- Watterson, G. A. 1975. On the number of segregating sites in genetical models without recombination. *Theor. Popul. Biol.* 7:256–276.
- Weir, J. T., and D. Schluter. 2004. Ice sheets promote speciation in boreal birds. *Proc. R. Soc. Lond. B* 271:1881–1887.
- Wilson, A. C., R. L. Cann, S. M. Carr, M. George, U. B. Gyllensten, K. M. Helm-Bychowski, R. G. Higuchi, S. R. Palumbi, E. M. Prager, R. D. Sage, et al. 1985. Mitochondrial DNA and two perspectives on evolutionary genetics. *Biol. J. Linn. Soc.* 26:375–400.
- Wink, M., H. Sauer-Gurth, and E. Gwinner. 2002. Evolutionary relationships of stonechats and related species inferred from mitochondrial-DNA sequences and genomic fingerprinting. *British Birds* 95:349–355.
- Woerner, A. E., M. P. Cox, and M. F. Hammer. 2007. Recombination-filtered genomic datasets by information maximization. *Bioinformatics* 23:1851–1853.
- Wu, C. I. 2001. The genic view of the process of speciation. *J. Evol. Biol.* 14:851–865.

Associate Editor: D. Presgraves

## Appendix

**Table A1. Samples.**

Sample ID	Species <sup>1</sup>	Population <sup>2</sup>	Date <sup>3</sup>	Locality	Country	Sex	Age <sup>4</sup>	GPS position
S0014	LM	DA	4/21/2007	Benaguasil	Spain	M	4	39°35.099'N 0°37.689'W
S0019	LM	DA	4/21/2007	Marines	Spain	M	4	39°44.431'N 0°32.008'W
S0022	LM	DA	4/22/2007	Vilamarxant	Spain	M	4	39°33.560'N 0°36.491'W
S0027	LM	DA	4/23/2007	Manises	Spain	M	4	39°29.122'N 0°26.387'W
S0029	LM	DA	4/23/2007	Eslida	Spain	M	4	39°53.369'N 0°19.007'W
S0032	LM	DA	4/24/2007	Almenara	Spain	M	5	39°45.063'N 0°11.697'W
S0035	LM	DA	4/26/2007	Amposta	Spain	M	5	40°40.268'N 0°35.799'W
S0037	LM	DA	4/26/2007	Amposta	Spain	M	4	40°41.760'N 0°33.341'W
S0038	LM	DA	4/27/2007	Campredo	Spain	M	4	40°44.359'N 0°34.268'W
S0041	LM	DA	4/27/2007	Campredo	Spain	M	5	40°45.949'N 0°32.361'W
S0042	LM	DA	4/27/2007	Tortosa	Spain	M	5	40°50.858'N 0°31.578'W
S0043	LM	DA	4/27/2007	Tortosa	Spain	M	5	40°52.526'N 0°30.683'W
S0007	LM	CA	5/28/2006	Mělník	Czech Republic	M	4	50°21.674'N 14°27.758'E
S0008	LM	CA	5/28/2006	Smiřice	Czech Republic	M	4	50°17.828'N 15°26.473'E
S0044	LM	CA	5/2/2007	Litoměřice	Czech Republic	M	6	50°32.374'N 14°06.263'E
S0049	LM	CA	5/3/2007	Roudnice nad Labem	Czech Republic	M	4	50°26.197'N 14°14.568'E
S0057	LM	CA	5/4/2007	Káraný	Czech Republic	M	4	50°10.509'N 14°44.943'E
S0060	LM	CA	5/5/2007	Lysá nad Labem	Czech Republic	M	5	50°11.213'N 14°51.681'E
S0069	LM	CA	5/5/2007	Nový Bydžov	Czech Republic	M	5	50°13.393'N 15°29.930'E
S0070	LM	CA	5/5/2007	Kopidlno	Czech Republic	M	4	50°20.294'N 15°14.903'E
S0073	LM	CA	5/6/2007	Březno	Czech Republic	M	6	50°24.616'N 15°00.277'E
S0076	LM	CA	5/6/2007	Kolomuty	Czech Republic	M	5	50°24.484'N 14°58.642'E
S0078	LM	CA	5/7/2007	Kilianów	SW Poland	M	5	51°00.798'N 16°45.027'E
S0081	LM	CA	5/7/2007	Brzeg Dolny	SW Poland	M	5	51°15.366'N 16°46.392'E
S0083	LM	CA	5/8/2007	Šestajovice	Czech Republic	M	5	50°20.332'N 16°00.870'E

**Table A1. Samples.**

Sample ID	Species <sup>1</sup>	Population <sup>2</sup>	Date <sup>3</sup>	Locality	Country	Sex	Age <sup>4</sup>	GPS position	
S0195	LL	CA	5/31/2007	Lipa	NE Poland	M	5	52°44.974'N	21°06.621'E
S0197	LL	CA	5/31/2007	Gnojno	NE Poland	M	4	52°46.789'N	21°08.524'E
S0205	LL	CA	6/1/2007	Nozewo	NE Poland	M	4	53°01.823'N	21°30.214'E
S0208	LL	CA	6/1/2007	Teodorowo	NE Poland	M	4	53°07.031'N	21°38.518'E
S0211	LL	CA	6/1/2007	Laskowiec	NE Poland	M	4	53°07.915'N	21°40.806'E
S0214	LL	CA	6/2/2007	Wizna	NE Poland	M	4	53°11.990'N	22°24.588'E
S0218	LL	CA	6/2/2007	Wizna	NE Poland	M	4	53°09.272'N	22°24.531'E
S0221	LL	CA	6/3/2007	Grajewo	NE Poland	M	5	53°38.712'N	22°29.157'E
S0222	LL	CA	6/3/2007	Sojczyn Gradowy	NE Poland	M	4	53°35.038'N	22°36.211'E
S0227	LL	CA	6/4/2007	Szczebra	NE Poland	M	4	53°54.987'N	22°58.666'E
S0228	LL	CA	6/4/2007	Smolniki	NE Poland	M	4	54°16.503'N	22°54.669'E
S0229	LL	CA	6/4/2007	Juryzdyka	NE Poland	M	5	53°58.727'N	22°57.442'E
S0009	LL	DA	8/7/2006	Ruissalo	Finland	M	3	60°25.089'N	22°12.184'E
S0230	LL	DA	6/7/2007	Ruissalo	Finland	M	6	60°26.215'N	22°12.326'E
S0233	LL	DA	6/7/2007	Turku	Finland	M	4	60°27.212'N	22°11.701'E
S0238	LL	DA	6/8/2007	Kustavi	Finland	M	5	60°32.494'N	21°25.762'E
S0241	LL	DA	6/8/2007	Laitila	Finland	M	6	60°53.498'N	21°41.560'E
S0245	LL	DA	6/9/2007	Poytya	Finland	M	4	60°46.258'N	22°38.926'E
S0247	LL	DA	6/9/2007	Suksela	Finland	M	4	60°30.360'N	22°41.531'E
S0250	LL	DA	6/9/2007	Paimio	Finland	M	6	60°27.404'N	22°42.078'E
S0251	LL	DA	6/10/2007	Salo	Finland	M	6	60°22.546'N	23°03.563'E
S0254	LL	DA	6/10/2007	Kaarina	Finland	M	4	60°23.801'N	22°23.195'E
S0255	LL	DA	6/10/2007	Kuusisto	Finland	M	4	60°24.425'N	22°28.482'E
S0257	LL	DA	6/11/2007	Karstu	Finland	M	6	60°19.138'N	23°53.224'E
S0258	LL	DA	6/11/2007	Lohja	Finland	M	4	60°16.255'N	24°02.331'E

<sup>1</sup>LM=*L. megarhynchos*; LL=*L. luscinia*.

<sup>2</sup>DA=distal allopatry; CA=close allopatry.

<sup>3</sup>Date of capturing.

<sup>4</sup>Age according to the Euring exchange-code.

## Supporting Information

The following supporting information is available for this article:

**Figure S1:** Marginal posterior probability of model parameters (scaled by the neutral mutation rate) for all loci (All), autosomal loci (Chr A), and Z-linked loci (Chr Z).

**Figure S2:** Marginal posterior probability distributions for the locus-specific migration rate estimates (scaled by the neutral mutation rate).

**Table S1:** Primers and PCR conditions.

**Table S2:** Nightingale-specific primers.

**Table S3:** Ensembl chromosomal position of best blast hit for eight autosomal and four Z-linked genes.

**Table S4:** Isolation model fitting.

**Table S5:** HKA test of positive selection.

Supporting Information may be found in the online version of this article.

(This link will take you to the article abstract).

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting informations supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.