ORIGINAL ARTICLE

Distinguishing the effects of selection from demographic history in the genetic variation of two sister passerines based on mitochondrial–nuclear comparison

C-M Hung¹ and RM Zink²

Determining the mechanisms responsible for the distribution of genetic diversity in natural populations has occupied a central role in molecular evolution. Our study was motivated by the unprecedented observation that a widespread Eurasian flycatcher, *Ficedula albicilla*, exhibited no variation at the mitochondrial DNA (mtDNA) ND2 gene in 75 individuals sampled over a 5000-km distance. In contrast, its sister species, *F. parva*, had low but considerably higher levels of mtDNA variation. We assessed whether natural selection or demographic factors could explain the absence of mtDNA variation in *F. albicilla*. Eighteen nuclear genes were sequenced to estimate the two species' phylogeographic histories, and for comparison to the mtDNA data. Multilocus coalescence analyses suggested that *F. albicilla* experienced a population expansion perhaps following a population bottleneck. Simulations based on this demographic history, however, did not replicate the extremely low level of mtDNA variation. Historical range changes based on ecological niche models also failed to explain the observed mtDNA patterns. Neutrality tests (DHEW and ML-HKA) suggested a non-neutral pattern in the mtDNA of *F. albicilla*. We found a transmembrane-skewed distribution of nonsynonymous substitutions between the two species, three of which caused functional change; the results implied that positive selection could have targeted mtDNA. Several lines of evidence support selection rather than demographic history as the main force influencing the patterns of mtDNA variation. Despite the influence of natural selection, many of the phylogeographic inferences derived from mtDNA were robust, including species limits and a high level of gene flow among populations within species.

Heredity (2014) **113**, 42–51; doi:10.1038/hdy.2014.9; published online 12 March 2014

INTRODUCTION

Whether the distribution of genetic variation is best explained by neutral or selective forces has fueled an extended debate in evolutionary biology (Hughes, 2007; Hahn, 2008). The neutral theory predicts that most genetic variants have few or no effects on fitness, and thus genetic variation is a function of mutation rate and population size (Kimura, 1983). Some authors argue that positive (or adaptive) selection is widespread and that the predicted relationship between population size and genetic variation is rarely observed (for example, Hahn, 2008). Many authors have inferred that positive selection influences many genes, particularly mitochondrial DNA (mtDNA; for example, Bazin *et al.*, 2006; Bensch *et al.*, 2006). If true, estimates of phylogeographical history based solely on mtDNA could be biased (Edwards and Bensch, 2009).

Genetic variation at the population level can be influenced by demographic factors, selective factors or both. For example, population bottlenecks can reduce genetic variation through reducing population sizes and facilitating genetic drift (Nei *et al.*, 1975). Similarly, selective fixation of a beneficial mutation can abruptly reduce genetic variation in a gene and genes linked to the targeted gene via hitchhiking (selective sweeps; Maynard-Smith and Haigh, 1974). Although theoretical work for each of these factors is well

developed, disentangling the two confounding factors from empirical data remains challenging.

The influence of demographic history should be observed genome-wide, whereas selection is more likely to affect different regions of a genome independently (Hudson *et al.*, 1987). Therefore, comparisons of genetic variation across unlinked genes have been used to test for departures from neutrality. However, stochastic coalescence or mutation processes can result in seemingly inconsistent variation among genes even under neutrality (Rosenberg and Nordborg, 2002). Thus, empirical comparisons should take into account such stochasticity to prevent biased inferences. In addition, demographic and selective factors could work in concert to shape the genetic variation of populations or species. Therefore, understanding the confounding effect of demographic history is essential to assess the effect of selection on observed patterns of genetic variation in nature.

mtDNA variation in two sister species of Old World flycatchers, *Ficedula albicilla* and *F. parva*, was studied to evaluate the potentially confounding influences of demography and selection in shaping observed genetic variation. Zink *et al.* (2008) found that 75 out of 75 *F. albicilla* individuals, sampled across its breeding range spanning around 5000 km, shared the same ND2 haplotype based on entire gene sequences (1041 bp). On the other hand, they found five ND2

¹Department of Life Science, National Taiwan Normal University, Taipei, Taiwan and ²Bell Museum and Department of Ecology, Evolution and Behavior, University of Minnesota, St Paul, MN, USA

Correspondence: Dr C-M Hung, Department of Life Science, National Taiwan Normal University, 88 Ting-Chow Road, Sec 4, Taipei 116, Taiwan.

E-mail: ymwur1@gmail.com

Received 9 September 2013; revised 1 December 2013; accepted 9 January 2014; published online 12 March 2014

haplotypes among 16 individuals of its sister species, *F. parva*. This discordant pattern raises the question of whether the different patterns of mtDNA variation are caused by different demographic histories or natural selection. If selection is the driving force, another question is whether selection influences mtDNA directly or indirectly via hitchhiking. In this study, we estimated the demographic histories of the two study species based on 18 nuclear genes, and then used the estimated histories to simulate neutral expectations to test mtDNA diversity for departures from neutrality. Furthermore, we applied neutrality tests, protein structure modeling and ecological niche modeling (ENM) to explore other influences on observed mtDNA variation. Three questions were addressed. First, we examined whether the two sister species experienced different population dynamics. Second, we tested whether the pattern of mtDNA variation departed from neutral expectation by accounting for species' demographic history. Finally, we determined whether positive selection might have eliminated most of the mtDNA variation.

MATERIALS AND METHODS

Sequence data collection and analyses

ND2 sequence data (1041 bp) for 16 *F. parva* and 75 *F. albicilla* (NCBI accession numbers EU326696–326786) were used from the study by Zink *et al.* (2008) study, in which *F. albicilla* samples were referred to as *F. parva* (The two species were considered the same species). The ND2 sequences of one *F. parva* from Sweden were downloaded from GenBank (NCBI accession number GU358799). Zink *et al.* (2008) used two primer pairs to amplify and sequence ND2 to insure the authenticity of the mitochondrial sequences. We further

checked mitochondrial authenticity in two ways: (1) using long PCR to amplify about half the mtDNA genome from ND6 to COI (around 8000 bp) or a fragment from 16S to COI (around 4000 bp) and then sequencing ND2, and (2) sequencing two other mtDNA genes, *Cytb* and COI (see Supplementary Materials for the PCR and sequencing protocols for mtDNA). The *Cytb* (1026 bp) gene was sequenced for 15 *F. parva* and 25 *F. albicilla* (Table 1). The *Cytb* gene of one *F. parva* from the Czech Republic was downloaded from GenBank (NCBI accession number AJ299689). The COI (1551 bp) gene was sequenced for 16 *F. parva* and 26 *F. albicilla* (Table 1).

We sequenced 15 autosomal and two Z-linked introns (Friesen *et al.*, 1999; Shapiro and Dumbacher, 2001; Primmer *et al.*, 2002; Backström *et al.*, 2006, 2008; Kimball *et al.*, 2009) and one exon (*Mc1r*; MacDougall-Shackleton *et al.*, 2003; Table 1) for variable numbers of individuals. Fourteen allele copies (two allele copies in one individual) of *Mc1r* were collected for *F. parva* and 22 allele copies for *F. albicilla* (Table 1). We collected 13–24 allele copies in the 17 introns for each species (mean = 18.5 for *F. parva* and 22.2 for *F. albicilla*; Table 1). Sample localities are given in Figure 1 and Supplementary Table 1.

The phases of sequences containing indels were sorted manually by subtracting chromatogram peaks upstream of the indel in the reverse primer sequences from the double peaks downstream of the indel in the forward primer sequences; this process was repeated in the alternative direction (Sousa-Santos *et al.*, 2005). The lengths of indels could also be determined by this approach. Multibase indels were collapsed to single-base polymorphisms for further analyses.

The allelic states of individuals with multiple heterozygous sites but no indels were resolved using PHASE 2.1.1 (Stephens *et al.*, 2001). Homozygous genotypes and genotypes with single heterozygous sites or with indels were set as known alleles to improve the performance of PHASE analyses. The phases of some sites could not be estimated confidently at 90% posterior probability.

Table 1 Characteristics of the 19 genes (the three mtDNA regions belong to the same linkage group) in this study

Gene	Chr	L	N _{FIPA}	N _{FIAL}	π_{FIPA}	π_{FIAL}	Dxy	Indel length
ND2	Mt ^a	1041	17	75	0.00138	0	0.06854	
Cytb	Mt ^a	1026	16	25	0.00193	0.00016	0.06069	
COI	Mt ^a	1551	16	26	0.00092	0.00005	0.05027	
Mean					0.00141	0.00007	0.05983	
00132	26	324	20	24	0.00332	0	0.00549	6
06635	6	540	20	24	0.00541	0.00749	0.01767	2, 4, 5
08352	5	348	20	24	0.00166	0.00378	0.01468	1
09385	20	283	20	22	0.00802	0.01859	0.01397	1, 1, 4, 7, 7, 7
11074	20	376	18	22	0.00688	0.00656	0.01013	15
12021	24	633	20	24	0.00125	0.00026	0.00891	1
13380	7	686	20	24	0.00814	0.00159	0.01882	1, 1, 1
14765	3	817	20	24	0.00462	0.00565	0.01245	1, 1
15743	1	530	16	24	0.01372	0.00931	0.02686	1, 10, 10
17483	4	491	20	24	0.00412	0.00750	0.01211	1, 5
ACL	27	372	16	24	0.00690	0.00657	0.01384	1
AK	17	707	18	24	0.00960	0.00602	0.01231	2, 61, 61, 61
MPP	4	216	18	22	0.00732	0.00459	0.00657	1
RHO	12	779	16	22	0.00090	0.00023	0.01214	1
TGFB2-I5	3	518	18	24	0.00703	0.00032	0.01099	
ABCA1	Z	281	17	13	0	0.01058	0.01122	
ADAMTS6	Z	529	17	13	0.00390	0.00210	0.01408	1, 1
Mean					0.00546	0.00536	0.01307	
<i>Mc1r</i>	11	590	14	22	0.0027	0.00164	0.00581	

Chr indicates the chromosome where the gene is located. L indicates the length of the gene used in the analyses. N_{FIPA} and N_{FIAL} show the sample size (in numbers of alleles) of *Ficedula parva* and *F. albicilla*, respectively, in each gene. π indicates the nucleotide diversity of each species. Dxy indicates the nucleotide divergence between the two species. Indel length shows the length of individual indel in the corresponding gene.

^aMitochondrial DNA

Figure 1 Breeding ranges of *F. parva* (dark gray) and *F. albicilla* (light gray) and distribution of samples sequenced for mtDNA and/or nuclear genes. Black circles indicate the localities where sample(s) were sequenced for mtDNA. Circles with black and white colors indicate the localities where sample(s) were sequenced for both mtDNA and nuclear genes. ARH indicates Arhangay, BUR indicates Buryatiya, CZE indicates Czech Republic, DOR indicates Dornod, IRK indicates Irkutskaya, KAM indicates Kamchatka, KHA indicates Khabarov, KRD indicates Krasnodar, MAG indicates Magadan, MAR indicates Markovo, MED indicates Medvedevo, MOS indicates Moscow, OMN indicates Omnogovi, ROS indicates Rostov, SWE indicates Sweden, TYV indicates Tyva, TÖV indicates Töv, YAK indicates Yakutiya and YEK indicates Yekaterinburg. Information on the numbers of samples sequenced for each gene at each locality is given in Supplementary Table 1.

The phases of these ambiguous sites were resolved using Clark's (1990) parsimony algorithm.

Intra-locus recombination rates were estimated using a Bayesian method (Li and Stephens, 2003) implemented in PHASE 2.1.1. A background recombination rate and the recombination rates between two polymorphic sites were estimated and portrayed as 5000 sampled points from the posterior distribution. The factors by which the estimated recombination rates exceed the background rate were calculated, and the upper and lower bounds of 95% credible intervals of these values were recorded. If one or more estimated values were significantly larger than one, it meant that there was significant recombination occurring between the corresponding sites (Hung *et al.*, 2012). Otherwise, we assumed no intra-locus recombination.

DnaSP 5 (Librado and Rozas, 2009) was used to compute nucleotide diversity (π ; Nei, 1987) for each species and nucleotide distance (D; Nei, 1987) between the two species for each gene.

Allele networks

NETWORK 4.5.1.6 (fluxus-engineering.com) was used to generate a minimum spanning allele network for each nuclear gene to reveal the pattern of divergence between *F. parva* and *F. albicilla*.

Demographic and divergence history

The program IMA (Hey and Nielsen, 2007) was used to estimate the effective population sizes of two species and their common ancestor, migration rates and divergence times between the two species. A minimum of 2×10^7 steps after a burn-in of 10^6 steps were performed for each analysis. Plots of trend lines and effective sample size values ($ESS > 250$) were examined to assess convergence in parameter estimates. In addition, at least two independent analyses were performed to assure convergence. The IMA analyses were applied to all 17 introns combined. To convert the scaled demographic parameters of IMA, we calculated the geometric mean of the substitution rates of these 17 loci by multiplying the sequence lengths by 1.35×10^{-9} substitutions per site per year for autosomal introns (Ellegren, 2007), and 1.6×10^{-9} substitutions per site per year for Z-linked introns (given the mean divergence of Z-linked introns between chicken and turkey is 1.2 times higher than that of autosomal introns (Ellegren 2007). The generation times of the two flycatcher species were set at 2 years because most male birds do not breed until the red throat or breast patch is attained after the second year (Mitrus, 2006). Additional IMA analyses were applied to two other data sets, one contained the 17 introns plus the exon and the other contained 15 introns excluding two introns (00132 and ABCA1), each of which was fixed in one species (see Results).

Although IMA can reconstruct divergence history between sister species by considering multiple demographic parameters, it cannot reveal chronological population dynamics owing to the assumption of no population size change since divergence. To complement the IMA results, we applied the Extended Bayesian Skyline Plot (EBSP) method (Heled and Drummond, 2008) implemented in BEAST v1.6.1 (Drummond and Rambaut, 2007) to estimate

population size changes through time. However, EBSP can only trace the history of one species per analysis, not coalescence between sister species. Thus, integrating IMA and EBSP results can provide a better picture of species' divergence and demographic history. The data sets used for the EBSP analyses were the same as those for the IMA analyses. Unlinked substitution estimates based on jModelTest (Posada, 2008) were assigned to each nuclear gene. The strict molecular clock model was used as it generally considered a better fit with analyses at the intraspecific level (Yang, 2006; Bisconti *et al.*, 2011) and it helps to avoid over-parameterization (Thalmann *et al.*, 2013) and thus facilitates convergence of the analyses. A locus that evolves at the fastest rate or is the most divergent was set as the reference locus to calibrate time and population sizes. ADAMTS6, a Z-linked intron, with a substitution rate of 1.6×10^{-9} substitutions per site per year and 15743, which had high π values in both species (see Table 1) with a substitution rate of 1.35×10^{-9} substitutions per site per year, were used as the reference loci for separate analyses to test for the effects of different reference loci. Each EBSP analysis included a MCMC chain of 500 million steps, which were sampled every 50 000 steps, and the first 25% of the steps were discarded as burn-in. Trace plots were checked using TRACER v 1.5 (Rambaut and Drummond, 2007) to assess convergence in MCMC analyses and three independent runs were performed to assure convergence in estimates. If the 95% confidence interval (CI) of the number of size-change steps excluded zero, we concluded that significant change(s) in the population size of the focal species had occurred.

Modeling current and historical distributions of the two species

ENMs can predict historical range change of species independent from genetic data. Locality records for breeding birds from each species were obtained from specimen records and ORNIS2 (<http://ornis2.ornisnet.org>); duplicate records were eliminated. ENMs were constructed using Maxent v3.2.2 (Phillips *et al.*, 2006) for both species at the present, the Last Glacial Maximum (LGM, 21 000 ybp, CCSM model) and the Last Interglacial (LIG, 120 000 ybp). Climatic data (19 layers) were obtained from the Worldclim bioclimatic database (Hijmans *et al.*, 2005). On the basis of an initial Maxent run of three replicates, climate layers contributing 5% or more to the model were selected, and the analysis was re-run using these layers. The final niche model was based on the average of five Maxent runs and plotted using DIVA-GIS v7.1.7.2 (Hijmans *et al.*, 2004). The goal was to determine whether the range of *F. albicilla* was severely contracted at the LGM, which could have led to a bottleneck.

Compound neutrality tests

The DHEW neutrality test (Zeng *et al.*, 2007), which is robust to demographic history, is a compound test of Tajima's *D* (Tajima, 1989), Fay and Wu's *H* (Fay and Wu, 2000) and Ewens-Watterson test (Watterson, 1978) and was used to detect positive selection on mtDNA. We applied DHEW tests implemented in the DH program (<http://zeng-lab.group.shef.ac.uk/wordpress>) to concatenated

sequences of ND2, Cytb and COI because the DHEW test relies on segregating sites (S) but *F. albicilla*'s ND2 was invariant and its Cytb ($S=2$) and COI ($S=1$) provided little variability (see Results). One sequence from each of the two species was used as the outgroup for the other. The significance levels of tests were determined through 50 000 simulations.

Multilocus neutrality tests

Maximum likelihood multilocus Hudson–Kreitman–Aguade (ML-HKA) tests implemented in the MLHKA program (Wright and Charlesworth, 2004) were used to test the neutrality of the mtDNA, the exon and two introns 00132 and ABCA1 that were fixed in *F. albicilla* and *F. parva*, respectively. This test is revised from the standard HKA test (Hudson *et al.*, 1987) and provides a test for selection at specific loci (Wright and Charlesworth, 2004). We performed likelihood ratio tests by assigning the remaining 15 introns to a neutral gene set, which was supported by the standard HKA test ($P=0.415$), against each individual candidate gene (that is, ND2, Cytb, COI, a concatenation of the three mtDNA genes, *Mcl1*, 00132 and *ABCA1*). The likelihood ratio tests compared a model considering the candidate gene under selection with a null model assuming that all 16 genes were neutral (degrees of freedom = 1). Each of the two species was used as the outgroup for the other. All runs included 500 000 cycles of Markov chains and two independent runs were performed for each model to check for the convergence of likelihood scores.

Testing mtDNA neutrality using simulations

To test whether the pattern of mtDNA variation deviated from neutral models, while considering demographic history and coalescence variation, we performed neutral coalescence simulations of mtDNA evolution using the program Bayesian Serial SimCoal (Excoffier *et al.*, 2000; Anderson *et al.*, 2005). As Bayesian Serial SimCoal simulates genes not individuals, the ploidy of genes was taken into account. That is, the effective population size of mtDNA (N_{ef}) was $0.5 \times$ the N_e of diploid individuals ($= 0.25 \times$ the N_e of autosomal genes) assuming a balanced sex ratio, in which the N_e of diploid individuals was the effective population size estimated from IMA (that is, $\theta/4\mu$; J. Hey, pers. comm.). A generation time of 2 years was used for the simulations. The following demographic model incorporating the IMA and EBSP estimates (see Results and Discussion) was used to simulate neutral mtDNA diversity: two species diverged from each other 1.6 million generations ago when the N_{ef} of their common ancestor was 370 000, the N_{ef} of *F. parva* remained at a similar and constant size, and *F. albicilla* experienced a recent population increase ($\sim 6X$) around 210 000 ($\sim 0.5 N_{ef}$) generations ago to a N_{ef} of 410 000. We assumed an $\sim 5X$ population bottleneck in the N_{ef} of *F. albicilla* when it split from *F. parva*, which we suggest provides a reasonable test for detecting departure from a neutral pattern. That is, if neutrality is

rejected, factor(s) other than a population bottleneck were the cause of the observed low mtDNA polymorphism of *F. albicilla*. To address uncertainty in the IMA estimates of divergence times and population sizes, we specified simulation parameters as normal distributions, rather than fixed values, with means equal to the values with the highest posterior probability (that is, those described above), covering at least 95% of the posterior distributions. To maintain a model with a population bottleneck followed by a population expansion in *F. albicilla*, we used the current N_{ef} of *F. albicilla* to scale all other population sizes and the timing of the population expansion in simulations (Figure 4a). The simulations used parameters based on the 15 introns, instead of all 17 introns, considering the potentially non-neutral patterns of 00132 and ABCA1, although the results of the two data sets were similar (see Results).

The substitution rates of mtDNA used for the simulations were estimated based on the results of this study. As the nucleotide difference between the two species were 0.06854, 0.06069 and 0.05027 in ND2, Cytb and COI, respectively (see Table 1), and the divergence time was about 3.327 million years (IMA estimate based on the 15 introns; see Table 2), the substitution rates of the three mtDNA regions were 0.0103, 0.0091 and 0.0076 per site per million years, respectively. In addition, we used the substitution rates estimated from Hawaiian honeycreepers, which are possibly more applicable to recent divergence (Lerner *et al.*, 2011), for simulation. Those were 0.029, 0.014 and 0.016 per site per million years for ND2, Cytb and COI, respectively.

We simulated 100 000 data sets with the same sequence lengths and sample sizes as the empirical data for *F. parva* and *F. albicilla*. The π values of the simulated data, calculated by Bayesian Serial SimCoal, were compared with those of the empirical data to test whether the mtDNA genetic diversity of the two species departed from neutral expectations.

We also applied the coalescence simulation test to 00132 and ABCA1 to examine whether they departed from neutral expectations and whether the patterns of low variation in mtDNA were more extreme than those in the two introns. The N_e of (autosomal) 00132 and (Z-linked) ABCA1 were assumed as $2 \times$ and $1.5 \times$ the N_e of diploid individuals, respectively. The substitution rates of 00132 and ABCA1 used for the simulations were estimated using the same approach as that for mtDNA and were 0.0008251 and 0.001686 per site per million years, respectively. Other simulation settings were the same as those for mtDNA.

Test for direct positive selection on mtDNA

The widely used M-K test (McDonald and Kreitman, 1991) for direct positive selection could not be used in this study owing to the absence or low numbers of nonsynonymous polymorphisms in the mtDNA genes of the two study species. Thus, we devised a test based on the distribution of nonsynonymous and synonymous interspecific differences between transmembrane (TM) and

Table 2 Effective population sizes, migration frequencies and divergence times estimated by IMA

	$N_{e(FIAL)}$	$N_{e(FIPA)}$	$N_{e(anc)}$	M_1	M_2	t
15 Introns						
HiSmth	820 666	742 987	749 649	0	0	3 327 422
95Lo	662 270	594 115	451 984	0	0	2 645 477
95Hi	1 039 143	954 096	1 467 906	0.0399	0.0372	4 432 643
17 Introns						
HiSmth	754 922	654 876	663 658	0	0	3 112 058
95Lo	610 042	524 777	389 416	0	0	2 492 660
95Hi	952 169	837 710	1 259 636	0.0354	0.0330	4 169 719
17 Introns + 1 exon						
HiSmth	758 884	660 159	653 003	0	0	3 229 241
95Lo	616 531	531 170	375 327	0	0	2 569 250
95Hi	951 031	841 109	1 215 532	0.0334	0.0306	4 232 697

Effective population sizes of *Ficedula albicilla*, *F. parva* and their common ancestor are denoted by $N_{e(FIAL)}$, $N_{e(FIPA)}$ and $N_{e(anc)}$, respectively. M_1 indicates the number of individual per generation migrating from *F. parva* to *F. albicilla* and M_2 indicates the migration in the other direction. t indicates the divergence time between the two species in the unit of years. These analyses are based on 15 introns, 17 introns and 17 introns plus one exon, respectively. HiSmth indicates the value of the highest posterior probability after being smoothed using surrounding points. 95Lo indicates the value to which 2.5% of the total distribution lies to the left. 95Hi indicates the value to which 2.5% of the total distribution lies to the right.

surface (SF) segments of mtDNA genes (see Wise *et al.*, 1998). If positive selection directly targets mtDNA gene function, nonsynonymous substitutions should be concentrated in TM segments that are subject to more stringent structural and functional constraints (Tourasee and Li, 2000).

The crystallized chicken Cytb protein (3170C.pdb) and bovine heart COI protein (2ocCA.pdb) were used as templates to model the three-dimensional structures of the two proteins for *F. parva* and *F. albicilla* using the SWISS-MODEL online server (Arnold *et al.*, 2006; online server at <http://swissmodel.expasy.org/>). The protein structures and the locations of amino acid replacements were visualized using the program Swiss-PdbViewer v4.02 (Guex and Peitsch, 1997). The TM and SF segments of the Cytb and COI proteins were predicted using MEMSAT (Jones *et al.*, 1994) through the SWISS-MODEL server. We were unable to model the protein structure of ND2 because a crystallized template was not available. Thus, the TM and SF segments of the ND2 protein were predicted using the program HMMTOP v2 (Tusnady and Simon, 2001; the online server at <http://www.enzim.hu/hmmtop/index.html>). We also used HMMTOP to predict TM and SF segments for the Cytb protein and found that the HMMTOP and MEMSAT predictions were very similar and the results of Fisher's exact tests based on the two predictions were the same. Thus, predictions of the two programs were compatible.

We separated the fixed substitutions between the two species into four categories, TM nonsynonymous, SF nonsynonymous, TM synonymous and SF synonymous substitutions, to conduct Fisher's exact tests. The ratio of TM synonymous to SF synonymous substitutions revealed relative mutation rates between TM and SF and was used to test whether the distribution of nonsynonymous substitutions was skewed. We applied this test to the ND2 and Cytb but not to the COI because there was only one fixed nonsynonymous substitution between the two species.

We further used PROVEAN v.1.1.2 online server (Choi *et al.*, 2012; online server at <http://provean.jcvi.org/index.php>) to examine whether the nonsynonymous substitutions between the two species' mtDNA have significant effect in protein function. PROVEAN assesses the effect of a substitution based on its resulting change in sequence similarity of a query protein to homologous proteins. The protein database was set as 'NCBI nr, September 2012' and the default delta score threshold (-2.5) was used. The PROVEAN prediction was applied to the ND2, Cytb and COI data.

RESULTS

Genetic diversity of mitochondrial and nuclear genes

The π values for *F. albicilla* (average of ND2, Cytb and COI = 0.00007 ± 0.00005 (s.e.)) were consistently lower than those of *F. parva* (average = 0.00141 ± 0.00029) in the three mtDNA genes, although the π values of *F. albicilla* were not zero for Cytb (0.00016) or COI (0.00005) as was found for ND2 (Table 1 and Figure 2). There were three Cytb haplotypes in 25 *F. albicilla* individuals; however, two singleton haplotypes differed only by one bp from the most frequent one. There were two COI haplotypes in 26 *F. albicilla* individuals, one of which was one singleton differing by one bp from the most frequent one. In contrast, π for the nuclear loci was not significantly different between the two species (0.00536 ± 0.00116 for *F. albicilla*, and 0.00546 ± 0.00086 for *F. parva* at introns, $P = 0.936$, paired Student's *t*-test; 0.00164 for *F. albicilla*, and 0.0027 for *F. parva* at *Mc1r*; Figure 2). Two introns had no genetic variation among the samples of one species. Those were 00132 in *F. albicilla* and ABCA1 in *F. parva* (Table 1). Most introns (15/17) contained one to six indels ranging from one to 61 bp (Table 1). No intra-locus recombination was detected in the 18 nuclear loci.

Allele networks of nuclear genes

Most nuclear networks (13/18) supported the split of the two species except for 09385, 11074, MPP, TGFB2-I5 and ABCA1, which were unresolved and therefore not conflicting (Supplementary Figure 1).

Figure 2 Nucleotide diversity (π) of three mtDNA genes, 17 introns and one exon of *F. parva* (gray bars) and *F. albicilla* (white bars). Shown are the means and SE for mtDNA and introns.

Demographic and divergence history

The IMA analyses based on the 17 introns suggested similar effective population sizes for *F. parva* (654 876; 95% CI = 524 777–837 710) and *F. albicilla* (754 922; 95% CI = 610 042–952 169; Table 2). The effective population size of their common ancestor (663 658; 95% CI = 389 416–1 259 636) was similar to those of the two species. The divergence time between the two species was estimated at 3 112 058 (95% CI = 2 492 660–4 169 719) years ago. There has been no gene flow between the two species (Table 2). The analyses based on the 17 introns plus one exon and 15 introns excluding 00132 and ABCA1 suggested similar histories as those based on the 17 introns (Table 2).

The EBS analyses, however, suggested that *F. albicilla* experienced a recent population expansion, whereas *F. parva* has had a stable effective population size based on the 17 introns (Figure 3; the 95% CIs of the number of size-change steps were 1–2 for the former and 0–1 for the latter). The EBS results based on the three data sets suggested similar history of a population expansion in *F. albicilla* and a stable population for *F. parva* (data not shown).

Modeling current and historical distributions of the two species

The ENMs for both species suggested that their ranges were reduced at the LGM compared with the LIG and present (Supplementary Figure 2). However, there was no evidence for a dramatic range contraction during the past 120 000 years that would explain the extremely low mtDNA variability in *F. albicilla*.

Neutrality tests

The DHEW tests suggested that the pattern of concatenated mtDNA sequences for *F. albicilla* was consistent with positive selection or selective sweeps ($P = 0.00034$) but not for that of *F. parva* ($P = 0.12$).

The ML-HKA tests rejected the neutral model in which the three mtDNA genes were added individually or in concert for both *F. albicilla* ($P \leq 0.000015$) and *F. parva* ($P = 0.0012$ – 0.0222), indicating a non-neutral pattern of variation in mtDNA. However, the ML-HKA tests could not reject the neutral models for *Mc1r*, 00132 and ABCA1 in either species ($P = 0.161$ – 0.599) except for ABCA1 in

Figure 3 Extended Bayesian Skyline Plots of population sizes of the two study species. Black dashed lines indicate the median values. Gray lines indicate 95% highest posterior density (HPD). ADAMTS6 is set as the reference locus to calibrate time and population size in this figure.

Figure 4 (a) The outline of the coalescence model incorporating the IMA and EBP estimates based on the 15 introns excluding 00132 and ABCA1. This model is used to simulate neutral mtDNA sequences of *F. parva* and *F. albicilla*. N_a , N_{FIPA} , N_{FIAL} and N_{FIALa} indicate the female effective population sizes of their common ancestor, *F. parva*, *F. albicilla*, and the latter before a population expansion, respectively. t_1 and t_2 indicate the divergence time between the two species and the initial time of the population expansion of *F. albicilla*, respectively, in a unit of generations from present. The parameters are set as normal distributions instead of fixed values. The symbol $\{N: 410\,000; 35\,000\}$ indicates a normal distribution with a mean of 410 000 and a s.d. of 35 000, and $\{N: 1\,664\,000; 165\,000\}$ indicates one with a mean of 1 664 000 and a s.d. of 165 000. (b) The distributions of nucleotide diversity (π) values for 100 000 neutrally simulated ND2 data sets in *F. parva* (grey bars) and *F. albicilla* (white bars). The grey triangle represents the empirical π value of ND2 in *F. parva* and the white one for *F. albicilla*.

F. parva, for which the neutral model was marginally rejected ($P = 0.0494$).

The results of DHEW and ML-HKA tests collectively suggested that the mtDNA variation of *F. albicilla* departed from neutral expectation, whereas selection was less likely an influence on *F. parva*.

Testing mtDNA neutrality using simulations

The comparison of the simulated and empirical data rejected a neutral pattern for the mtDNA of both *F. parva* and *F. albicilla*. The

empirical π value of ND2 in *F. albicilla* was lower than the range of π values of the simulated neutral data and that in *F. parva* was lower than the 99.99% range of the simulated π values based on the substitution rates estimated in this study (Figure 4b); the empirical π values of ND2 in both species were also lower than the range of the simulated π values based on the substitution rate of Lerner *et al.* (2011) (data not shown). The same result was found for the Cytb and COI data for both species (data not shown).

Table 3 Numbers of fixed nonsynonymous and synonymous substitutions between *Ficedula parva* and *F. albicilla* at the transmembrane and surface segments of ND2 and Cytb.

Fixed substitution	ND2		Cytb	
	Transmembrane	Surface	Transmembrane	Surface
Nonsynonymous	12	1	8	1
Synonymous	30	24	30	21
FET probability	0.012		0.083	

Abbreviation: FET, Fisher's exact one-tailed test.

FET probability indicates the probability of Fisher's exact one-tailed test.

The simulation results rejected a neutral pattern for ABCA1 in both species with lower confidence than those for mtDNA but could not reject a neutral pattern for 00132 in either species. The empirical π values of *F. albicilla* and *F. parva* for ABCA1 were out of the 95% range but were within the 99% range of the simulated ones (Supplementary Figure 3). Noticeably, the former was located in the upper tail of the range (that is, non-neutrally high polymorphism) and the latter was in the lower tail of the range (that is, non-neutrally low polymorphism). In contrast, the empirical π values of both species for 00132 were within the 90% range of the simulated ones (Supplementary Figure 3).

Test for direct positive selection on mtDNA

Nonsynonymous substitutions tended to occur in the TM regions (12 TM:1 SF) of the ND2 protein more than synonymous ones between the two study species (30 TM:24 SF; $P_{\text{one-tail}} = 0.012$, Fisher's exact test; Table 3). We also found more TM nonsynonymous substitutions ($n = 8$) than SF nonsynonymous substitutions ($n = 1$) in the Cytb protein (Table 3; Supplementary Figure 4), although the TM:SF ratios of nonsynonymous and synonymous substitutions were not significantly different ($P_{\text{one-tail}} = 0.083$; Table 3).

PROVEAN predicted that three TM nonsynonymous substitutions in ND2 between the two species, Threonine (*F. albicilla*) to Isoleucine (*F. parva*) at the 18th codon, Isoleucine (*F. albicilla*) to Threonine (*F. parva*) at the 19th codon and Threonine (*F. parva*) to Alanine (*F. albicilla*) at the 152th codon could change protein function. No nonsynonymous substitutions in the Cytb and COI data were predicted to have a functional effect.

DISCUSSION

Contrasting genetic diversity between mtDNA and nuclear DNA (ncDNA)

One explanation for the lack of polymorphism in the ND2 of *F. albicilla* is that Zink *et al.* (2008) inadvertently sequenced pseudogenes. However, mtDNA polymorphism in *F. albicilla* was consistently lower than that in *F. parva* across ND2, Cytb and COI. Therefore, it is unlikely that nuclear copies of mtDNA were sequenced in all three regions because they were well separated throughout the mtDNA genome. Furthermore, (1) the PCR products for ND2 derived from four primer pairs returned the same sequencing results, (2) there were no indels or stop codons found in ND2 and (3) correct sister relationships between the two species and between their clade and the other clade containing *F. hypoleuca*, *F. albicollis*, *F. speculigera* and *F. semitorquata* were obtained based on the ND2 data (results not shown). Together, these observations support the mitochondrial authenticity of the ND2 data.

In contrast to the mtDNA data, the ncDNA polymorphism levels and the estimated N_e based on multiple nuclear genes were similar between *F. albicilla* and *F. parva*. Therefore, difference in N_e is not an explanation for the lower level of mtDNA polymorphism in *F. albicilla*. However, the EBSF revealed a recent population expansion in *F. albicilla* but not *F. parva*. Thus, the effect of historical population changes on mtDNA patterns should be considered. In addition, this result can be confounded by large standard errors in mtDNA diversity due to stochastic coalescence processes (Edwards and Bensch, 2009). Thus, repeated coalescence simulations based on population history, which we did in this study, are needed to confirm whether these striking patterns depart from neutral expectation.

Departure from neutral patterns in mtDNA

The seeming disagreement between the EBSF and IMA estimated histories could be caused by the simplified model of IMA. On the other hand, the EBSF estimate of population size and expansion timing varied with the choice of a reference locus. For example, the current effective population size of *F. albicilla* was estimated as 23 500 when ADAMTS6 was assigned as the reference locus, whereas the estimated size was 867 000 when 15743 was used. The initial time of the *F. albicilla* population expansion was around 12 000 and 390 000 years ago when ADAMTS6 and 15743 were the reference loci, respectively. Nevertheless, the extent of population increase ($\sim 6X$) and the ratios of the initial time of the population expansion to the current population size (~ 0.5) were consistent across different reference loci. Thus, we incorporated the IMA and EBSF results for coalescence simulations to take the advantage of precise IMA estimates, which are based on the geometrical mean of substitution rates for all loci, and explicit demographic fluctuation revealed by EBSF (see Materials and Methods).

Neutral coalescence simulations predicted slightly lower π values in *F. albicilla* than for *F. parva*; however, the difference was not significant (Figure 4b). Thus, different demographic histories are unable to explain the different levels of mtDNA diversity between the two species.

The low mtDNA diversity in *F. albicilla* departed from the neutral expectation based on our simulations. Therefore, the modeled bottleneck cannot explain such low mtDNA variation observed in *F. albicilla*. However, the bottleneck (with a magnitude of $\sim 5X$) we modeled is much less severe than typically expected for great genetic effect (Nei *et al.*, 1975), although the magnitude of population reduction we assumed was consistent with our multilocus estimates. It explains the observed mtDNA results of the DHWE and ML-HKA tests, which suggested that selection is likely responsible for the low mtDNA diversity of *F. albicilla*.

The observed mtDNA diversity of *F. parva* was also significantly lower than the neutral expectation, which is supported by the ML-HKA test but not the DHEW test. The inconsistent results between the DHEW test and the ML-HKA test or the coalescence simulation approach could be due to the strengths and sensitivities of the tests. In particular, the DHEW test can be more conserved because it integrates multiple tests. The mtDNA diversity of *F. parva* (π_S (synonymous nucleotide polymorphism) = 0.0047 ± 0.0006 , π_N (nonsynonymous nucleotide polymorphism) = 0.0003 ± 0.0002) is lower than other avian species ($\pi_S = 0.02\text{--}0.06$, $\pi_N = 0.0015\text{--}0.005$ for 72 avian species; see Hughes and Hughes, 2007). Of course, the mtDNA diversity of *F. albicilla* is even lower. The mtDNA variation patterns of the two flycatchers were anomalously low for birds.

The fact that the two introns ABCA1 and 00132 were fixed in *F. parva* and *F. albicilla*, respectively, points to the possibility that the

low mtDNA variation in the two species could happen by chance. The coalescence simulation approach and the ML-HKA test could marginally reject a neutral model for ABCA1 but could not for 00132. Although the results are mixed, they suggest that the magnitude of departure from neutral expectation for mtDNA is more extreme than that for the two introns, and unlikely to happen by chance.

Tests for positive selection targeting mtDNA

The TM segments of ND2 accumulated more nonsynonymous substitutions than the SF segments between the two flycatchers. This result differs from some studies. Kerr (2011) found that nonsynonymous substitutions were equally distributed between helix (primarily TM) and loop (primarily SF) sites based on the COI data of 43 avian species across 12 orders; he found no evidence for a selective sweep in the avian COI data. Tourasse and Li (2000) reported that the TM segments of proteins accumulated fewer nonsynonymous substitutions than the SF segments in mammals across six orders and suggested that the TM segments were subject to more stringent structural constraint. The contrasting patterns between this study and others imply that the ND2 protein of either *F. parva* or *F. albicilla* or both may be targeted by positive selection. Although relaxed purifying selection can also lead to elevated nonsynonymous substitutions in one or both types of segments, it should not lead to extremely low levels of genetic variation. In addition, coalescence analyses suggested a stable population size for *F. parva* and a recent or ongoing population expansion in *F. albicilla*, which impede relaxation of purifying selection (Ohta, 1973). Thus, positive selection is more likely to be the explanation for the observed pattern rather than relaxed purifying selection.

Furthermore, three TM nonsynonymous substitutions in the ND2 between the two study species were predicted to cause functional change in the protein. Two of the substitutions (Thr¹⁸ to Ile¹⁸; Ile¹⁹ to Thr¹⁹) were from *F. albicilla* to *F. parva* and the third one (Thr¹⁵² to Ala¹⁵²) was from *F. parva* to *F. albicilla*. All three substitutions change the polarity of amino acids. Ile¹⁸, Thr¹⁹ and Ala¹⁵² are unique derived states, whereas Thr¹⁸, Ile¹⁹ and Thr¹⁵² are ancestral states found in most of the *Ficedula* species whose ND2 sequences are available in Genbank (Supplementary Table 2). The results imply positive selection on novel advantageous mutations in each of the two study species. Additional studies on mechanistic links between phenotypic or physiological changes and genetic substitutions and consequent ecological interactions are warranted. Despite the plausible evidence found here supporting positive selection directly targeting the mtDNA ND2 gene, alternative hypotheses such as mitochondrial–nuclear coadaptation (Gershnoi *et al.*, 2009) and selective sweeps caused by linkage to either other mtDNA genes (Oliveira *et al.*, 2007) or the W chromosome due to shared segregation process (Berlin *et al.*, 2007) are worthy of further investigation.

The phylogeography of the two flycatchers based on selected mtDNA, ncDNA and ENMs

The multilocus analyses suggested that *F. albicilla* and *F. parva*, which had long been considered conspecific, diverged from each other around 3 million years ago. The well-sorted nuclear genes (Supplementary Figure 1) supported their taxonomic split, which was based on distinguishable plumage characters (Cramp and Perrins, 1993) and highly diverged mtDNA (Zink *et al.*, 2008).

The ENMs for the two species exhibit characteristic features of northern hemisphere species that experienced southward range retractions during the LGM, followed by northward advance, despite the lack of major ice sheets occurring in northern Asia (Adams, 2002).

The magnitude of range contraction at the LGM relative to the LIG and the present is not commensurate with the near absence of mtDNA variation in *F. albicilla*. Although the timing of the population fluctuations suggested by the genetic analyses apparently predate the range fluctuations since the LIG, the ncDNA results suggest no distinct differences in population history between the two species, which the ENMs confirm for at least the past 120 000 years.

The ENMs help to reconstruct the selection-driven history of the mtDNA. The time from introduction to fixation of a favored haplotype in a population is $\sim 2 \ln(2N)/s$, where N is the population size of the focal gene and s is the selection coefficient (Stephan *et al.*, 1992). Assuming that s acting on *F. albicilla*'s mtDNA is within the range of those on color polymorphisms across diverse taxa, 0.013–0.67 (Hoekstra *et al.*, 2004), the fixation time is 41–2095 generations if N is set as the current effective population size (410 000), or 35–1822 generations if the N is set as that before the recent population expansion (70 000; Figure 4a). Thus, the dominant haplotype in the mtDNA of *F. albicilla* would have arisen and spread when this species was experiencing a post-LGM population expansion. It is consistent with the theoretical prediction that the fixation frequencies of advantageous mutations increase with population sizes (Meiklejohn *et al.*, 2007).

Therefore, demographic and selective forces might have worked in concert to lead to the unique pattern in which there is almost no mtDNA variation in *F. albicilla*, across an area of 5000 km. One has to assume that if a highly advantageous mutant arose, presumably in a single female, a high level of gene flow was needed to lead to its rapid spread over a maximum of 2095 generations (*ca.* 4190 years). Thus to explain the current distribution of mtDNA variation in *F. albicilla*, both a high selection coefficient and a high level of gene flow are inferred.

CONCLUSION

The simulations and ML-HKA and DHEW tests indicate that the unusually low mtDNA variation of *F. albicilla* has mainly been driven by selection and that demographic history alone has had little impact, although high levels of gene flow are required to explain the distribution of the common mtDNA haplotype. The mtDNA variation of *F. parva* has also been reduced by selection, but to a lesser extent compared with *F. albicilla*. It is possible that selection strength is lower or selection operates too recently to reach fixation in the mtDNA variation of *F. parva*. The TM-skewed nonsynonymous substitution distribution and the functional change caused by TM amino-acid substitutions in ND2 imply that positive selection has drastically reduced mtDNA variation, although the possibility that selection on linked (such as other mtDNA genes) or functionally associated genes causes their low mtDNA variation cannot be completely ruled out. This study presents a clear case that demonstrates how natural selection can shape genetic variation within and among populations.

A number of published studies (for example, Edwards and Bensch, 2009; Galtier *et al.*, 2009) have claimed that natural selection on mtDNA is sufficiently pervasive to prevent phylogeographic inference. We found strong evidence of selection on the mtDNA of these flycatchers, especially *F. albicilla*. However, we suggest that three important inferences can be obtained from the mtDNA pattern, even if it was influenced by natural selection. First, the mtDNA data corroborate the species limits, which are revealed in morphology and nuclear genes. Second, the existence of the same haplotype across a wide area means that there are no barriers to dispersal, because even if positive selection was the cause of the lack of variability, the favored

haplotype cannot occur where individuals do not disperse. Third, the lack of variation suggests that the event leading to the homogenization of the mitochondrial genome was recent and rapid (Przeworski, 2002), as otherwise one would expect a series of random point mutations found idiosyncratically across the range. Thus, although the level of variability observed in the mtDNA of *F. albicilla* is anomalously low, several important evolutionary inferences could be obtained. If we only had nuclear loci, this complex history would not be apparent. Several recent studies compared mtDNA and nuclear genes, discovering intriguing processes of population differentiation, local adaptation or unexpected mtDNA dynamics (for example, Cheviron and Brumfield, 2009; Ribeiro *et al.*, 2011; Pavlova *et al.*, 2013). Accumulating multilocus studies have shown that when mtDNA is under selection, adding this marker will gain a more comprehensive phylogeographic picture. Lastly, a few persuasive exceptions should not discount the fact that mtDNA often reveals geographically and taxonomically important patterns, especially among recently isolated taxa (Zink and Barrowclough, 2008; McKay and Zink, 2010).

DATA ARCHIVING

Data deposited in GenBank: accession numbers KJ362224-KJ363064 and in the Dryad repository: doi:10.5061/dryad.m4sn8.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We are grateful to S Rowher, S Drovetski and A Pavlova for their contribution in specimen collection and preparation. M Westberg provided laboratory assistance, T Rodrigues aided in map-making and S Weller helped to improve the manuscript. We thank L Dong for advising *F. albicilla* occurrence data collection. Support came from NSF (DEB 0919494). The University of Minnesota Supercomputing Institute provided critical assistance with the computations.

Adams JM (2002). Global land environments since the last interglacial. Available at: <http://www.esd.ornl.gov/projects/gen/merc/html>.

Anderson CNK, Ramakrishnan U, Chan YL, Hadly EA (2005). Serial Simcoal: a population genetic model for data from multiple populations and points in time. *Bioinformatics* **21**: 1733–1734.

Arnold K, Bordoli L, Kopp J, Schwede T (2006). The SWISS-MODEL Workspace: a web-based environment for protein structure homology modeling. *Bioinformatics* **22**: 195–201.

Backström N, Brandström M, Gustafsson L, Qvarnström A, Cheng H, Ellegren H (2006). Genetic mapping in a natural population of collared flycatchers (*Ficedula albicollis*): conserved synteny but gene order rearrangements on the avian z chromosome. *Genetics* **17**: 377–386.

Backström N, Fagerberg S, Ellegren H (2008). Genomics of natural bird populations: a gene-based set of reference markers evenly spread across the avian genome. *Mol Ecol* **17**: 964–980.

Bazin E, Glémin S, Galtier N (2006). Population size does not influence mitochondrial genetic diversity in animals. *Science* **312**: 570–572.

Bensch S, Irwin DE, Irwin JH, Kvist L, Åkesson S (2006). Conflicting patterns of mitochondrial and nuclear DNA diversity in *Phylloscopus* warblers. *Mol Ecol* **15**: 161–171.

Berlin S, Tomaras D, Charlesworth B (2007). Low mitochondrial variability in birds may indicate Hill-Robertson effects on the W chromosome. *Heredity* **99**: 389–396.

Bisconti R, Canestrelli D, Colangelo P, Nascetti G (2011). Multiple lines of evidence for demographic and range expansion of a temperate species (*Hyla sarda*) during the last glaciation. *Mol Ecol* **20**: 5313–5327.

Cheviron ZA, Brumfield RT (2009). Migration-selection balance and local adaptation of mitochondrial haplotypes in rufous-collared sparrows (*Zonotrichia capensis*) along an elevational gradient. *Evolution* **63**: 1593–1605.

Choi Y, Sims GE, Murphy S, Miller JR, Chan AP (2012). Predicting the functional effect of amino acid substitutions and indels. *PLoS One* **7**: e46688.

Clark AG (1990). Inference of haplotypes from PCR-amplified samples of diploid populations. *Mol Biol Evol* **7**: 111–122.

Cramp S, Perrin C (1993). *Handbook of the Birds of Europe, the Middle East and North Africa: the Birds of the Western Palearctic: Flycatchers to Shrikes*, Vol. 8. Oxford University Press: Oxford, UK.

Drummond AJ, Rambaut A (2007). BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**: 214.

Edwards S, Bensch S (2009). Looking forwards or looking backwards in avian phylogeography? A comment on Zink and Barrowclough 2008. *Mol Ecol* **18**: 2930–2933.

Ellegren H (2007). Molecular evolutionary genomics of birds. *Cytogenet Genome Res* **117**: 120–130.

Excoffier L, Novembre J, Schneider S (2000). SIMCOAL: a general coalescent program for simulation of molecular data in interconnected populations with arbitrary demography. *J Hered* **91**: 506–509.

Fay JC, Wu C-I (2000). Hitchhiking under positive Darwinian selection. *Genetics* **155**: 1405–1413.

Friesen VL, Congdon BC, Kidd MG, Birt TP (1999). Polymerase chain reaction (PCR) primers for the amplification of five nuclear introns in vertebrate. *Mol Ecol* **8**: 2147–2149.

Galtier N, Nabholz B, Glémin S, Hurst GDD (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Mol Ecol* **18**: 4541–4550.

Gershoni M, Templeton AR, Mishmar D (2009). Mitochondrial bioenergetics as a major motive force of speciation. *BioEssays* **31**: 642–650.

Guex N, Peitsch MC (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* **18**: 2714–2723.

Hahn MW (2008). Toward a selection theory of molecular evolution. *Evolution* **62**: 255–265.

Heled J, Drummond AJ (2008). Bayesian inference of population size history from multiple loci. *BMC Evol Biol* **8**: 289.

Hey J, Nielsen R (2007). Integration within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc Natl Acad Sci USA* **104**: 2785–2790.

Hijmans RJ, Guarino L, Bussink C, Mathur P, Cruz M, Barrientes I *et al.* (2004). DIVA-GIS Vsn. 7.1.7. A geographic information system for the analysis of species distribution data. Available at: <http://www.diva-gis.org>

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005). Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* **25**: 1965–1978.

Hoekstra HE, Drumm KE, Nachman MW (2004). Ecological genetics of adaptive color polymorphism in pocket mice: geographic variation in selected and neutral genes. *Evolution* **58**: 1329–1341.

Hudson RR, Kreitman M, Aguadé M (1987). A test of neutral molecular evolution based on nucleotide data. *Genetics* **116**: 153–159.

Hughes AL (2007). Looking for Darwin in all the wrong places: the misguided quest for positive selection at the nucleotide sequence level. *Heredity* **99**: 364–373.

Hughes AL, Hughes AK (2007). Coding sequence polymorphism in avian mitochondrial genomes reflects population histories. *Mol Ecol* **16**: 1369–1376.

Hung C-M, Drovetski SV, Zink RM (2012). Multilocus coalescence analyses support a mtDNA-based phylogeographic history for widespread Palearctic passerine bird, *Sitta europaea*. *Evolution* **66**: 2850–2864.

Jones DT, Taylor WR, Thornton JM (1994). A model recognition approach to the prediction of all-helical membrane protein structure and topology. *Biochemistry* **33**: 3038–3049.

Kerr KCR (2011). Searching for evidence of selection in avian DNA barcodes. *Mol Ecol Resour* **11**: 1045–1055.

Kimball RT, Braun EL, Barker FK, Bowie RCK, Braun MJ, Chojnowski JL *et al.* (2009). A well-tested set of primers to amplify regions spread across the avian genome. *Mol Phylog Evol* **50**: 654–660.

Kimura M (1983). *The Neutral Theory of Molecular Evolution*. Cambridge University Press: Cambridge, UK.

Lerner HRL, Meyer M, James HF, Hofreiter M, Fleischer RC (2011). Multilocus resolution of phylogeny and timescale in the extant adaptive radiation of Hawaiian Honeycreepers. *Curr Biol* **21**: 1–7.

Li N, Stephens M (2003). Modeling linkage disequilibrium, and identifying recombination hotspots using SNP data. *Genetics* **165**: 2213–2233.

Librado P, Rozas J (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* **25**: 1451–1452.

MacDougall-Shackleton EA, Blanchard L, Gibbs HL (2003). Unmelanized plumage patterns in old world leaf warblers do not correspond to sequence variation at the Melanocortin-1 Receptor locus (MC1R). *Mol Biol Evol* **20**: 1675–1681.

Maynard-Smith J, Haigh J (1974). The hitch-hiking effect of a favourable gene. *Genet Res* **23**: 23–35.

McDonald JH, Kreitman M (1991). Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**: 652–654.

McKay BD, Zink RM (2010). The causes of mitochondrial DNA gene tree paraphyly in birds. *Mol Phylog Evol* **54**: 647–650.

Meiklejohn CD, Montooth KL, Rand DM (2007). Positive and negative selection on the mitochondrial genome. *Trends Genet* **23**: 259–263.

Mitrus C (2006). The influence of male age and phenology on reproductive success of the red-breasted flycatcher (*Ficedula parva* Bechst.). *Ann Zool Fennici* **43**: 358–365.

Nei M (1987). *Molecular Evolutionary Genetics*. Columbia University Press: New York, NY, USA.

Nei M, Maruyama T, Chakraborty R (1975). The bottleneck effect and genetic variability in populations. *Evolution* **29**: 1–10.

Ohta T (1973). Slightly deleterious mutant substitutions in evolution. *Nature* **246**: 96–98.

- Oliveira DCS, Raychoudhury R, Lavrov DV, Werren JH (2007). Rapidly evolving mitochondrial genome and directional selection in mitochondrial genes in the parasitic wasp *Nasonia* (Hymenoptera: Pteromalidae). *Mol Biol Evol* **25**: 2167–2180.
- Pavlova A, Amos JN, Joseph L, Loynes K, Austin JJ, Keogh J *et al.* (2013). Perched at the mito-nuclear crossroads: divergent mitochondrial lineages correlate with environment in the face of ongoing nuclear gene flow in an Australian bird. *Evolution* **67**: 3412–3428.
- Phillips SJ, Anderson RP, Schapire RE (2006). Maximum entropy modeling of species geographic distribution. *Ecol Model* **19**: 231–259.
- Posada D (2008). jModelTest: phylogenetic model averaging. *Mol Biol Evol* **25**: 1253–1256.
- Primmer CR, Borge T, Lindell J, Sætre G-P (2002). Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. *Mol Ecol* **11**: 603–612.
- Przeworski M (2002). The signature of positive selection at randomly chosen loci. *Genetics* **160**: 1179–1189.
- Rambaut A, Drummond AJ (2007). Tracer v1.4, Available from: <http://beast.bio.ed.ac.uk/Tracer>.
- Ribeiro AM, Lloyd R, Bowie RCK (2011). A tight balance between natural selection and gene flow in a southern African arid-zone endemic bird. *Evolution* **65**: 3499–3514.
- Rosenberg N, Nordborg M (2002). Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. *Nat Rev Genet* **3**: 380–390.
- Shapiro LH, Dumbacher JP (2001). Adenylate Kinase Intron 5: a new nuclear locus for avian systematics. *Auk* **118**: 248–255.
- Stephan W, Wiehe THE, Lenz MW (1992). The effect of strongly selected substitutions on neutral polymorphism: analytical results based on diffusion theory. *Theor Popul Biol* **41**: 237–254.
- Stephens M, Smith N, Donnelly P (2001). A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* **68**: 978–989.
- Sousa-Santos C, Robalo JI, Collars-Pereira M-J, Almada VC (2005). Heterozygous indels as useful tools in the reconstruction of DNA sequences and in the assessment of ploidy level and genomic constitution of hybrid organisms. *DNA Seq* **16**: 462–467.
- Tajima F (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**: 585–595.
- Thalmann O, Shapiro B, Cui P, Schuenemann VJ, Sawyer SK, Greenfield DL *et al.* (2013). Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. *Science* **342**: 871–874.
- Tourasse NJ, Li W-H (2000). Selective constraints, amino acid composition, and the rate of protein evolution. *Mol Biol Evol* **17**: 656–664.
- Tusnady GE, Simon I (2001). The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**: 849–850.
- Watterson GA (1978). The homozygosity test of neutrality. *Genetics* **88**: 405–417.
- Wise CA, Sraml M, Eastal S (1998). Departure from neutrality at the mitochondrial NADH dehydrogenase subunit 2 gene in humans, but not in chimpanzees. *Genetics* **148**: 409–421.
- Wright SI, Charlesworth B (2004). The HKA test revisited: A Maximum-Likelihood-Ratio Test of the standard neutral model. *Genetics* **168**: 1071–1076.
- Yang Z (2006). *Computational Molecular Evolution*. Oxford University Press: Oxford, UK.
- Zeng K, Suhua S, Wu C-I (2007). Compound tests for the detection of hitchhiking under positive selection. *Mol Biol Evol* **24**: 1898–1908.
- Zink RM, Barrowclough GF (2008). Mitochondrial DNA under siege in avian phylogeography. *Mol Ecol* **17**: 2107–2121.
- Zink RM, Pavlova A, Drovetski SV, Rohwer S (2008). Mitochondrial phylogeographies of five widespread Eurasian bird species. *J Ornithol* **149**: 399–413.

Supplementary Information accompanies this paper on Heredity website (<http://www.nature.com/hdy>)