Mitochondrial introgression and replacement between yellowhammers (*Emberiza citrinella*) and pine buntings (*Emberiza leucocephalos*) (Aves: Passeriformes)

DARREN E. IRWIN[1]*, ALEXANDER S. RUBTSOV[1,2] and EUGENE N. PANOV[3]

[1]*Biodiversity Research Centre, and Department of Zoology, University of British Columbia, 6270 University Boulevard, Vancouver, BC, V6T 1Z4, Canada*
[2]*State Darwin Museum, Moscow 117292, Russia*
[3]*Severtsov Institute of Ecology and Evolution, Moscow 117071, Russia*

In studies of phylogeography and taxonomy, strong emphasis is usually placed on the study of mitochondrial (mt)DNA. In the present study, we present a remarkable case in which highly phenotypically divergent species have almost no divergence in mtDNA. Yellowhammers (*Emberiza citrinella* Linnaeus) and pine buntings (*Emberiza leucocephalos* S. G. Gmelin) differ noticeably in appearance and song but hybridize in some areas of contact. They share a variety of closely-related mtDNA haplotypes, with little divergence in frequencies, indicating a mitochondrial divergence time sometime during or after the last major glacial period. By contrast, nuclear DNA (amplified fragment length polymorphism markers and *CHD1Z* gene sequences) differs more strongly between the species, and these differences can be used to identify intermediate genetic signatures of hybrids. The combined amount of mitochondrial diversity within yellowhammers and pine buntings is very low compared to other *Emberiza* species pairs, whereas the level of variation at the nuclear gene *CHD1Z* is comparable to that within other species pairs. Although it is difficult to completely reject the possibility that the two species split extremely recently and experienced rapid nuclear and phenotypic differentiation, we argue that the evidence better supports another possibility: the two species are older and mtDNA has recently introgressed between them, most likely as a result of a selective sweep. Mismatches between mitochondrial and nuclear phylogeographic patterns may occur more commonly than previously considered, and could have important implications for the fields of phylogeography and taxonomy. © 2009 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2009, **98**, 422–438.

ADDITIONAL KEYWORDS: AFLP – CHD1Z – hybridization – mtDNA – phylogeography – Siberia – speciation.

INTRODUCTION

The study of variation in mitochondrial (mt)DNA has played a central role in studies of taxonomy and phylogeography (Avise, 2004; Ballard & Whitlock, 2004; Zink & Barrowclough, 2008). Mitochondrial DNA is unusual in that, in most taxa (such as birds), it is inherited solely through the maternal line. As a result, mtDNA generally has a lower effective population size than nuclear DNA, causing mtDNA within a single population to have a lower expected coalescence time than most nuclear genes, assuming that no selection is involved. This characteristic potentially makes mtDNA more useful than a single nuclear gene in inferring phylogeographic history (Avise, 2004; Zink & Barrowclough, 2008).

Despite these advantages, the sole use of mtDNA in inferring history has recently come under criticism (Irwin, 2002; Ballard & Whitlock, 2004; Edwards *et al.*, 2005; Rubinoff & Holland, 2005; Bazin, Glémin & Galtier, 2006), primarily because the entire mitochondrial genome is inherited as a single unit and captures only a portion of phylogeographic history. Because of stochastic effects, sex-biased gene flow, lack of recombination between distinct mitochondrial

*Corresponding author. E-mail: irwin@zoology.ubc.ca

lineages, and direct selection on mitochondrial variation (Dowling, Friburg & Lindell, 2008), patterns of geographic variation in mtDNA might be highly discordant with patterns of variation throughout the nuclear genome. In response to these criticisms, Zink & Barrowclough (2008) have defended the primary use of mtDNA, emphasizing its greater ability to capture patterns of population division because of its lower effective population size and arguing that the use of mtDNA leads to accurate conclusions regarding taxonomic relationships the great majority of the time. Zink (2004) even argued that mtDNA is of greater importance than phenotypic traits in determining correct subspecies boundaries. In the present study, we explore these issues by studying patterns of genetic differentiation in mitochondrial and nuclear DNA between two phenotypically distinct but hybridizing taxa that have been considered distinct species ever since their scientific description in the late 1700s.

The debate regarding the importance of mtDNA as an indicator of population history has important implications for the study of speciation, in which genetic patterns are compared between relatively young taxa. Past variation in climate and habitat distribution (e.g. during the Pleistocene glaciations) has frequently divided species ranges into geographically separated areas, and then allowed those areas to expand and come into contact, producing contact zones between divergent groups. Such contact zones enable an examination of interactions between closely-related taxa and an assessment of whether they are reproductively isolated. One region where such contact zones are common is central Siberia (Haffer, 1989; Rogacheva, 1992; Newton, 2003; Irwin & Irwin, 2005), where many pairs of related western and eastern forms meet. Siberia was treeless (although not highly glaciated) for long periods during the Pleistocene glaciations (Frenzel, 1968; Adams & Faure, 1997); hence, any forest-dependent species found there now must have expanded from other regions. In birds, patterns of migratory behavior (Irwin & Irwin, 2005) and molecular biogeography (Irwin, Bensch & Price, 2001; Irwin *et al.*, 2005) suggest that many species expanded into Siberia from a forested refugium in central Asia (i.e. the region west and north-west of the Tibetan Plateau and Taklamakan Desert), whereas others expanded from a refugium in eastern or north-eastern China (or nearby regions such as the Korean Peninsula; Nazarenko, 1990). Some species had related forms in both of these regions.

Thus, central Siberia is a meeting place of distinct avian faunas, making it a prime location for the study of how related forms interact when they come into secondary contact. Yet little research has focused on the genetic relationships between western and eastern relatives in central Siberia. In perhaps the only previous such study, genetic variation indicated that western and eastern forms of greenish warblers [*Phylloscopus trochiloides viridanus* and *Phylloscopus (trochiloides) plumbeitarsus*] started diverging long before the last major period of Pleistocene glaciation and then moved into central Siberia from different directions: one from central Asia and the other from eastern China (Irwin *et al.*, 2001, 2005).

In the present study, we investigate interactions between the yellowhammer (*Emberiza citrinella* Linnaeus 1758) and the pine bunting (*Emberiza leucocephalos* S. G. Gmelin 1771), two sister taxa in the Emberizinae family (Alström *et al.*, 2007). Both species breed across a wide region of western and central Siberia, with yellowhammers extending westward to western Europe and pine buntings extending eastward to the Russian Far East (Fig. 1). The two species differ noticeably in plumage patterns, although a variety of phenotypic intermediates across the sympatric area suggest that they hybridize extensively (Panov, 1989; Panov, Rubtsov & Monzikov, 2003; Panov, Rubtsov & Mordkovich, 2007). The two forms are similar morphometrically but apparently differ to some degree in habitat preference, with yellowhammers being more likely to inhabit shrubby habitat along forest edges and in mountain steppe vegetation, and pine buntings being more likely to inhabit open forests where coniferous trees predominate (Ravkin, 1973; Panov *et al.*, 2003).

The present study was motivated by the desire to use yellowhammers and pine buntings as a model system for the study of speciation between western and eastern Palearctic forms. Our primary goal was to use patterns of genetic differentiation to: (1) reconstruct the time of population splitting between yellowhammers and pine buntings and (2) clarify patterns of current introgression due to hybridization. We hypothesized that genetic variation would indicate that the two forms started diverging sometime before the last of a long series of major Pleistocene glaciations; this last glacial period began approximately 110 000 years ago (Adams & Faure, 1997). We intended to use mtDNA as the primary genetic marker in this effort, as is often done in the field of phylogeography (Avise, 2004; Zink & Barrowclough, 2008). Our surprising initial results indicating virtually no mitochondrial divergence between the species prompted us to investigate patterns of variation in the nuclear genome as well, using amplified fragment length polymorphism (AFLP) analysis and the sequencing of a sex-linked gene (*CHD1Z*) on the Z-chromosome. Given the large phenotypic differences between yellowhammers and pine buntings, we hypothesized that the two groups are genetically
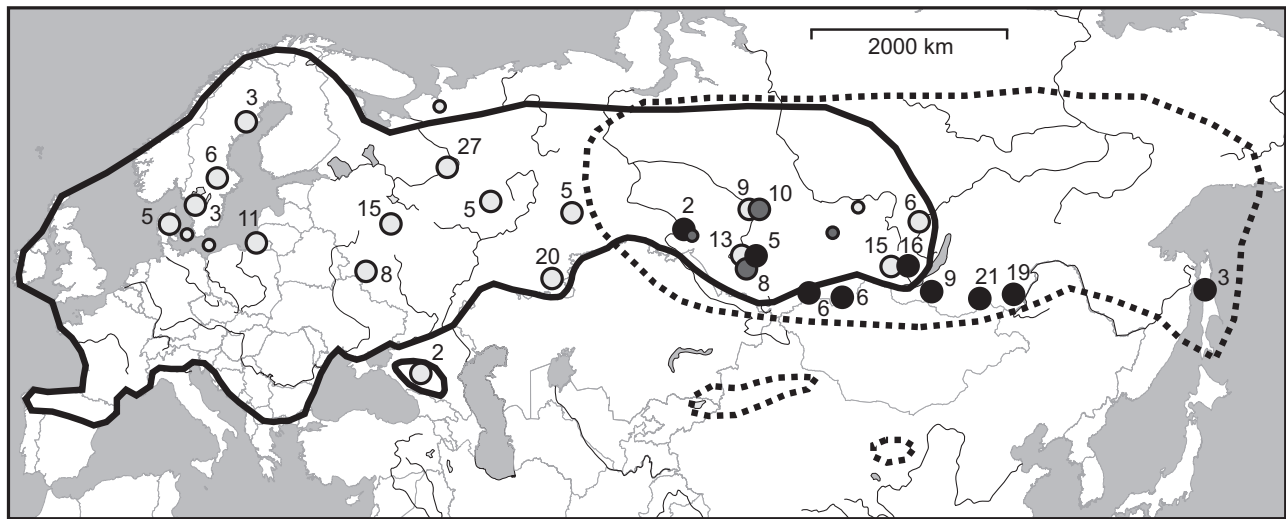
Figure 1. Geographic distribution of yellowhammers (*Emberiza citrinella*; solid lines) and pine buntings (*Emberiza leucocephalos*; dashed lines), which hybridize extensively in their area of overlap in western and central Siberia. Sampling sites are indicated by small circles (one sample) or large circles (multiple samples, with numbers indicating sample sizes). Phenotypic yellowhammers are indicated by light grey circles, phenotypic hybrids by dark grey, and phenotypic pine buntings by black.

divergent in nuclear DNA, even though they are so similar in mtDNA. We also aimed to use nuclear DNA markers that differed between yellowhammers and pine buntings to test whether phenotypically intermediate individuals (i.e. apparent hybrids) have intermediate patterns of genetic variation. The results reveal a variety of nuclear markers that differ between the taxa and are useful in identifying hybrids. More importantly, our findings indicate rapid mitochondrial introgression and fixation, a phenomenon that is often not considered in the fields of phylogeography and taxonomy.

MATERIAL AND METHODS

SAMPLING

We obtained blood and/or tissue (muscle or liver) samples from field-caught birds and from museum specimens. In total, we examined 156 yellowhammers, 87 pine buntings, and 20 phenotypic hybrids (see Supporting information, Table S1). Each individual from which a sample was obtained (both live birds and museum specimens) was photographed and described phenotypically according to the standard scoring system used by Panov *et al.* (2003). We also obtained samples of eight other *Emberiza* species in order to estimate the genetic relationships of the yellowhammer/pine bunting complex to other species within the genus. Sample sizes and localities are shown in Figure 1 (see also Supporting information, Table S1).

From birds captured in the field, we took approximately 50 μL of blood and diluted it in 500 μL of 'Queen's Lysis Buffer' (Seutin, White & Boag, 1991). Blood samples were stored in a cool place during field work and frozen just after returning from the field. Tissue samples from the museum collections were stored in 96% ethanol. DNA was extracted using a standard phenol–chloroform protocol.

DNA SEQUENCE ANALYSIS

To test whether yellowhammers and pine buntings differ in mtDNA we sequenced a subset of samples of each species from allopatric populations: yellowhammers from the Baltic Sea, Kursk, and Krasnodar regions (41 samples), and pine buntings from Chita and Sakhalin regions (33 samples; for locations, see Supporting information, Table S1). We amplified 1032 bp of the *ND2* gene using the primers L5215: 5′-TATCGGGCCCATACCCCGAAAAT-3′ and 1064: 5′-CTTTGAAGGCCTTCGGTTTA-3′ (Drovetski *et al.*, 2004). Polymerase chain reaction (PCR) amplification was conducted using 1 μM each of primers L5215 and 1064, 1 × PCR buffer and 1.5 U *Taq* DNA polymerase (New England Biolabs), 2.5 mM MgCl₂ (Invitrogen), and 0.25 mM dNTP-mix (New England Biolabs) in a 25-μL total volume. PCR included an initial denaturing step of 94 °C for 3 min, followed by 35 cycles of 94 °C for 30 s, 55 °C for 30 s, and 72 °C for 1 min, and a final elongation step at 72 °C for 10 min. To compare levels of *ND2* variation in the

yellowhammer/pine bunting species complex with that throughout the *Emberiza* genus, we sequenced the same *ND2* fragment from eight other species: *Emberiza aureola*, *Emberiza calandra*, *Emberiza cioides*, *Emberiza cirlus*, *Emberiza hortulana*, *Emberiza spodocephala*, *Emberiza stewarti*, and *Emberiza tristrami* (see Supporting information, Table S1).

We are reasonably confident that our *ND2* sequences are from the mitochondrial genome rather than from a nuclear pseudogene (Sorenson & Quinn, 1998) for several reasons. First, we used a combination of tissue samples and blood samples, and the two sources of DNA gave compatible results (pseudogenes are more likely to be amplified from blood than from tissue). Second, our resulting *ND2* phylogeny of ten *Emberiza* species (see Results) is highly consistent with the cytochrome *b* (cyt*b*) phylogeny presented by Alström *et al.* (2007), differing only slightly in nodes that received low bootstrap support in the cyt*b* phylogeny (e.g. the placement of *E. cirlus*). Third, our *ND2* phylogeny had much greater depth than our *CHD1Z* phylogeny, consistent with the expectation that mitochondrial sequences evolve more quickly than nuclear sequences (Sorenson & Quinn, 1998).

We sequenced 612 bp of an intron of a sex-linked gene, *CHD1Z* on the Z-chromosome (Fridolfsson & Ellegren, 1999), from 173 samples representing both parental species and their hybrids. PCR amplification was performed using 1 µM each of primers 2550F (5′-GTTACTGATTCGTCTACGAGA-3′) and 2718R (5′-ATTGAAATGATCCAGTGCTTG-3′; Fridolfsson & Ellegren, 1999) and other PCR reagents as described above for *ND2*, in a 25-µL total volume. PCR included an initial denaturing step of 94 °C for 3 min, followed by 35 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 40 s, and a final elongation step at 72 °C for 10 min. To compare *CHD1Z* variation within the yellowhammer/pine bunting species complex with that throughout the genus, we sequenced the same *CHD1Z* fragment from four other species: *E. calandra*, *E. hortulana*, *E. cirlus*, and *E. stewarti* (see Supporting information, Table S1). Purification and sequencing of PCR products for both the *ND2* gene and the *CHD1Z* intron were performed by Macrogen (Seoul, South Korea).

Sequences were aligned and edited using BIOEDIT (Hall, 1999). Both avian sex chromosomes (Z and W) have a copy of the *CHD1Z* gene, but the intron used differs in length between these chromosomes (Fridolfsson & Ellegren, 1999). For sequencing, we used the samples obtained from the homogametic sex (males) only, to avoid sequencing complications from the presence of the W-chromosome copy. To reconstruct haplotypes from unphased genotypes we used the software FASTPHASE under default settings (Scheet & Stephens, 2006).

For both *ND2* and *CHD1Z*, we produced minimum spanning haplotype networks with the assistance of the software TCS (Clement, Posada & Crandall, 2000) and ARLEQUIN, version 3.11 (Excoffier, Laval & Schneider, 2005), which produced identical results. ARLEQUIN was also used to calculate Φ_{ST}, defined as the fraction of variance in pairwise sequence differences (i.e. the number of mismatches) that is explained by the difference between two groups (Excoffier *et al.*, 2006; Φ_{ST} takes into account both genetic distance between and frequency of haplotypes, and is therefore not equivalent to some formulations of F_{ST} that use only haplotype frequency).

To estimate genetic distances and phylogenetic trees among *Emberiza* species based on *ND2* and *CHD1Z*, we used the software TREE-PUZZLE (Schmidt *et al.*, 2002), which uses a maximum likelihood approach. We used an HKY + Γ model of sequence evolution (Price, 2008: 38–40), eight gamma-distributed rate categories, and accurate parameter estimation using neighbour-joining trees. A likelihood ratio test as implemented in TREE-PUZZLE revealed that, for both the *ND2* and the *CHD1Z* trees, the assumption of a constant molecular clock does not significantly decrease the likelihood of the tree; hence, we use the clock-like trees for illustration.

Although there is support for a reasonably constant molecular clock for mtDNA in birds (Weir & Schluter, 2008), we did not assume the same clock would apply to *CHD1Z*. To calculate a range of reasonable estimates for the *CHD1Z* molecular clock that was independent of the yellowhammer/pine bunting data, we compared *ND2* distances with *CHD1Z* distances for the other species pairs that were in both datasets (*E. calandra*, *E. hortulana*, *E. cirlus*, and *E. stewarti*).

DNA sequences have been deposited in GenBank (accession numbers GQ370010–GQ370064).

DEMOGRAPHIC MODELLING BASED ON MTDNA

Many studies of phylogeography, historical demography, and conservation genetics rely heavily on inference from mtDNA-based demographic modelling, using a variety of methods. We used such methods to explore the range of historical scenarios that are consistent with the mitochondrial data under certain assumptions. We note that our goal was not to infer the precise history of the two species; rather, we wished to determine whether the historical scenarios inferred using mtDNA were consistent with variation in nuclear DNA and phenotypic traits.

To test whether the *ND2* data were consistent with neutral evolution under constant population size, we used ARLEQUIN to perform three tests of selective

neutrality, each tested for significance using 10 000 simulations: Tajima's *D*-test (Tajima, 1989), Fu's F_S test (Fu, 1997), and Ewens–Watterson–Slatkin's exact test (Slatkin, 1996).

To estimate the time of divergence and demographic history of allopatric yellowhammers and allopatric pine buntings based on mitochondrial *ND2* data and an assumption of neutrality, we used the software IMa (Hey, 2007; Hey & Nielsen, 2007), which used a Bayesian Markov chain Monte Carlo approach to fit the *ND2* sequence data to an 'Isolation with Migration' model (Nielsen & Wakeley, 2001). This model assumes that two populations (in this case, allopatric yellowhammers and allopatric pine buntings) split some time ago from a single population, and that there has been some constant rate of gene exchange subsequent to the populations splitting. Given an estimated mutation rate, IMa attempts to simultaneously estimate six parameters: the effective population sizes of both current populations and the ancestral population, the time splitting of the populations, and the migration rates of each population to the other. The model assumes selective neutrality and no recombination within a locus.

We ran IMa under two scenarios, using *ND2* sequence data from allopatric yellowhammers and allopatric pine buntings. The purpose of these model runs was not to determine precisely the true history of the two species; the model assumptions are too restrictive and simple for that purpose. Rather, our goal was to estimate the range of population splitting times that are consistent with the mitochondrial *ND2* data, in the hope that this information would help us to determine whether mitochondrial patterns were consistent with patterns in nuclear DNA. First, we ran IMa to estimate all six parameters simultaneously. Second, we ran the model with the maximum migration rate in each direction set to zero, effectively making the model into a two-island isolation model. Several preliminary runs were conducted under each scenario to determine proper upper bounds on parameters, as well as the run times necessary to achieve convergence and have sufficient sampling (Hey, 2007). Each model was run with a burn period of 10^7 steps followed by a recording period of 10^8 steps. Each scenario was run twice using different random number seeds, to check repeatability of results. In the runs that included post-split migration, we used IMa to record the mean time of migration events (Won & Hey, 2005).

Conversion of the IMa parameters t and θ_A to divergence time in years and effective population size requires an estimate of generation time and mutation rate (Hey, 2007). We used an estimated generation time of 1.7 years (i.e. the average time for seven passerine species; Kondo *et al.*, 2008; after Sæther

et al., 2005; note that using a different estimate of generation time would affect estimates of population size but would not affect the resulting estimate of time in years since population expansion). We based our estimate of the *ND2* mutation rate on the well-supported mitochondrial molecular clock estimate of 1% mutation per million years along a single lineage (2% between lineages; Lovette, 2004; Weir, 2006; Weir & Schluter, 2008). This rate corresponds to 1.03×10^{-5} mutations per entire *ND2* sequence (1032 bp) per year, or an estimate for μ of 1.75×10^{-5} per generation. To account for uncertainty in the rate of mutation, we also calculated parameter values for both 0.5% and 1.5% per million year mutation rates, which appeared to be appropriate given the apparent variation in calibrated rates reported by Weir & Schluter (2008).

AFLP ANALYSIS

Whereas mitochondrial DNA is inherited matrilineally as a single unit, AFLP markers are spread throughout the nuclear genome and represent a variety of genealogical histories. To generate AFLP markers, we used the protocol provided by LI-COR Biosciences (2003), based on the method developed by Vos *et al.* (1995). We used the restriction enzymes *EcoR*I and *Mse*I to digest genomic DNA, and then synthetic oligonucleotides ('adaptors') were ligated to the fragments. We performed two rounds of PCR using primers corresponding to the adaptor plus one arbitrary base pair in the first round and the adaptor plus three arbitrary base pairs in the second round. Fluorescein-labelled primers were used in the second round of PCR (the 'selective amplification'). The products were separated in 6.5% KB^{Plus} gels and visualized using a LI-COR 4300 DNA Analyzer.

We generated two sets of AFLP data for subsequent analysis; each allowed particular questions to be addressed most effectively.

AFLP dataset 1: genetic differentiation between taxa
Here, our goal was to obtain an unbiased estimate of genetic differentiation between allopatric yellowhammers and allopatric pine buntings. We scored all AFLP bands, regardless of their patterns of variation, from five primer combinations. This was performed for a subset of samples from allopatric yellowhammers ($N = 13$) and allopatric pine buntings ($N = 15$). We did not include bands that appeared in three or fewer individuals. We summarized variation in the resulting presence/absence matrix using principal components analysis (PCA), using R (R Development Core Team, 2006).

We used AFLP-SURV (Vekemans, 2002) to calculate F_{ST}, the proportion of total variance in allele frequencies that is explained by differences between the two

species. Estimation of allele frequencies was performed using a Bayesian approach (with non-uniform prior distributions of allele frequencies) assuming Hardy–Weinberg equilibrium within each species. This approach is necessary to avoid biases in the estimation of allele frequencies from dominant markers (Lynch & Milligan, 1994; Zhivotovsky, 1999). A locus-specific F_{ST} was generated using these allele frequencies, using the equation:

$$F_{ST} = \frac{(p_1 - \bar{p})^2 + (p_2 - \bar{p})^2}{2\bar{p}(1 - \bar{p})},$$

where p_x is the frequency of the allele for band presence in population x, and $\bar{p}$ is the mean of p_1 and p_2.

Some previous AFLP studies of population differentiation have reported an alternative F_{ST} statistic that is based on AFLP band frequencies rather than allele frequencies. This statistic is based simply on two phenotypes, these being presence or absence of each AFLP band. Because a number of other studies report this statistic (Svensson *et al.*, 2004; Bensch & Åkesson, 2005; Helbig *et al.*, 2005; Irwin *et al.*, 2005; Parchman, Benkman & Britch, 2006; Toews & Irwin, 2008), for purposes of comparison we also calculated it for our data, using ARLEQUIN, version 3.11 (Excoffier *et al.*, 2005).

AFLP dataset 2: selected markers for distinguishing taxa and hybrids

Here, our goal was to select only the most useful markers in distinguishing the taxa and presumably in identifying hybrids as well. We examined AFLP profiles from ten primer combinations and determined AFLP markers that differed in presence/absence frequencies between 26 allopatric yellowhammers (group 1 below) and 15 allopatric pine buntings (group 7 below) according to two criteria: we selected all markers that were (1) at least 1.5-fold more common in one of the groups than in the other and (2) present in at least 50% of individuals in the group with the highest frequency of that marker. These criteria ensured that selected markers were picked objectively and had a large difference in frequency between the two groups.

This procedure resulted in 20 selected markers, which were then also scored in individuals from within or close to the contact zone. In total, this analysis included 65 samples divided into seven phenotypic groups (for detailed descriptions of each group, see Panov *et al.*, 2003): (1) allopatric yellowhammers (from Baltic Sea, Kursk, Moscow, and Orenburg) ($N = 26$); (2) phenotypic yellowhammers from the contact zone ($N = 4$); (3) yellow hybrids ($N = 8$); (4) white hybrids ($N = 7$); (5) pine buntings with slight hybrid phenotypes ($N = 2$); (6) pine bun-

tings in or near the contact zone ($N = 3$); and (7) allopatric pine buntings (from East Transbaikalia and Sakhalin) ($N = 15$). We again summarized variation in the resulting presence/absence matrix using PCA, using R (R Development Core Team, 2006). We also used STRUCTURE, version 2.2 (Pritchard, Stephens & Donnelly, 2000; Falush, Stephens & Pritchard, 2007) to calculate assignment probabilities of individuals to two genetic clusters based on their AFLP signatures.

RESULTS

MITOCHONDRIAL DNA

Allopatric yellowhammers and pine buntings differ remarkably little in mtDNA (Fig. 2A), with many haplotypes shared between the two species. Φ_{ST} is only 0.025, which does not differ significantly from zero ($P = 0.073$). Out of 1032 bp sequenced, the average pairwise difference between species, after correction for within-species polymorphism, was only 0.056 bp (uncorrected pairwise differences between species: 2.231 bp; within yellowhammers: 2.285 bp; within pine buntings: 2.064 bp). This amounts to a percentage divergence of 0.0054%, which is an extraordinarily low amount for such phenotypically divergent taxa.

A variety of analyses indicate that mtDNA of both yellowhammers and pine buntings have experienced a selective sweep and/or population growth. Tajima's test of selective neutrality (Tajima, 1989) was significant for yellowhammers ($D = -1.49$, $P = 0.049$) and pine buntings ($D = -1.81$, $P = 0.018$), as well as for all samples combined ($D = -2.06$, $P = 0.004$), where significance indicates rejection of the neutral model. Fu's (1997) test also rejected neutrality in each case (yellowhammer: $F_S = -9.87$, $P < 0.00001$; pine bunting: $F_S = -9.16$, $P = 0.00010$; all samples combined: $F_S = -24.37$, $P < 0.00001$). Finally, Ewens–Watterson–Slatkin's exact test (Slatkin, 1996) also indicated a selective sweep and/or population growth (yellowhammer: $P = 0.0182$; pine bunting: $P = 0.0007$; all samples combined: $P < 0.00001$).

The IMa analyses provided fairly precise estimates of the time of population splitting of yellowhammer and pine bunting mtDNA. All four runs (two allowed migration after the split and two did not) produced similar posterior probability distributions for the time of the split (see Supporting information, Fig. S1), with the peak time being 32 000 years ago [95% confidence interval (CI) = 20 000–48 000]. The analysis also produced a narrow range of estimates of the ancestral effective population size (see Supporting information, Fig. S1), which was relatively small, with the peak estimate being 110 000 (95% CI = 47 000–290 000).
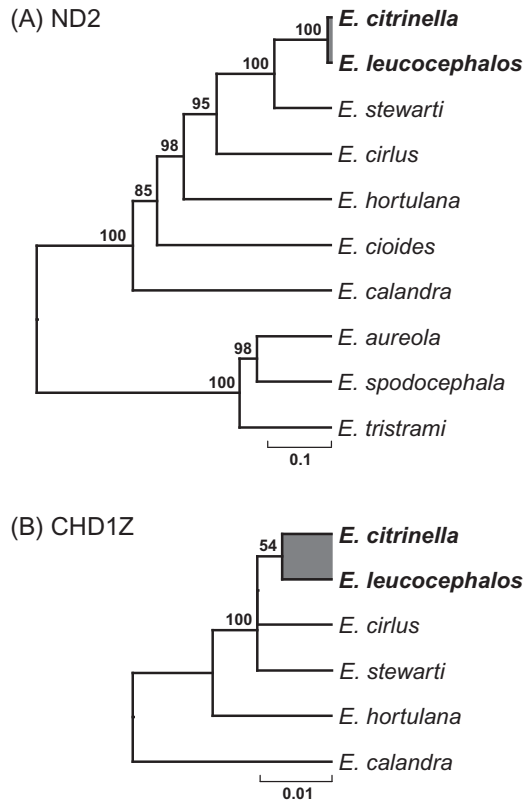
(A) ND2
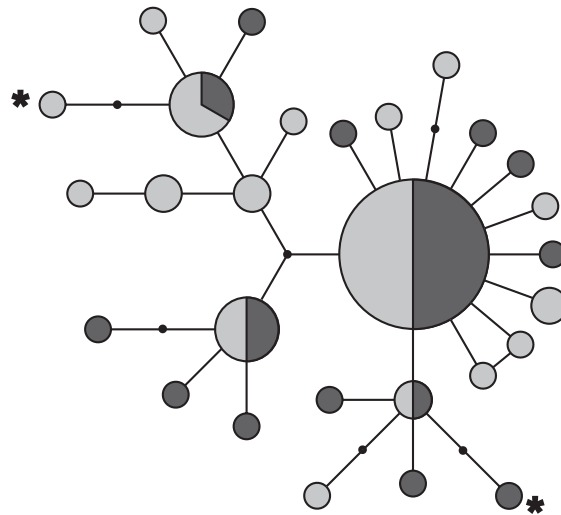
(B) CHD1Z
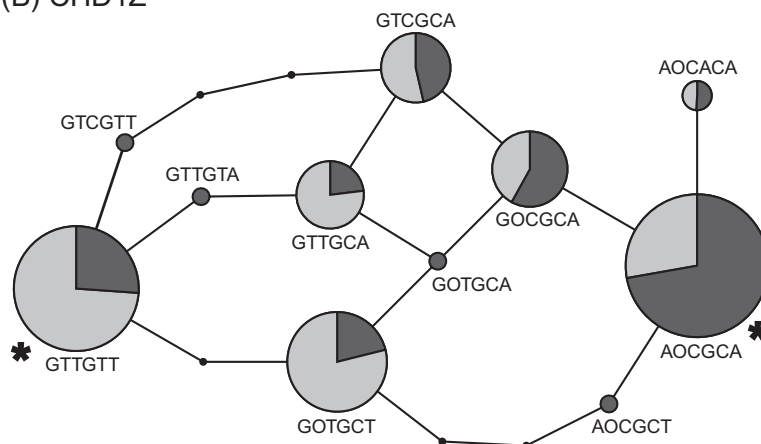
Figure 2. Minimum spanning network showing the relationships among (A) *ND2* haplotypes and (B) *CHD1Z* haplotypes of allopatric yellowhammers (light gray) and pine buntings (dark gray). In (A), circle area is proportional to the number of samples with that haplotype, with the smallest size corresponding to a single sample, and small dots representing missing haplotypes. Asterisks indicate the two most distantly-related haplotypes in the two species; these haplotypes were used in the interspecific phylogeny (Fig. 3A) and to calculate maximum genetic distance (Table 1). In (B), the circle area is proportional to the number of haplotypes of that sequence (with a single individual represented by GTCGTT). The light and dark areas of the circles represent the relative frequencies of particular haplotypes in yellowhammer and pine buntings, respectively. Missing haplotypes are indicated by small dots. Most haplotypes could have been produced by recombination of the two most common haplotypes (GTTGTT, which is more common in yellowhammers; and AOCGCA, which is more common in pine buntings). These two most common haplotypes (marked with asterisks) were used in the phylogenetic reconstruction in Fig. 3B.

Incorporating uncertainty in the rate of mutation in the range 0.5–1.5% per million years along a lineage, these 95% CIs expand to 14 000–97 000 years ago for the time of splitting and the ancestral effective population size of 31 000–580 000. The IMa runs were not able to produce stable distributions for the estimates of present population sizes, nor migration rates after

the time of the split (see Supporting information, Fig. S1).

The most divergent mitochondrial haplotypes within the yellowhammer/pine bunting complex differ by only eight substitutions, or 0.8% sequence divergence. This maximal amount of sequence divergence within the yellowhammer/pine bunting complex is

Figure 3. Phylogenetic trees of *Emberiza* species based on (A) *ND2* and (B) *CHD1Z*. Numbers to the left of nodes represent quartet puzzling support values, with numbers above 90 indicating very strong confidence in the clade joined at that node. Scale bars indicate expected rates of nucleotide divergence between two lineages, under the HKY + Γ substitution model (see Material and methods). The grey areas between *Emberiza citrinella* and *Emberiza leucocephalos* denote haplotype sharing between the two species, presumably as a result of ongoing gene flow. For these two species, highly divergent haplotypes were used to represent the two species in the phylogeny (indicated with asterisks in Fig. 2; see Results).

very low compared to that between other species of *Emberiza* buntings (Fig. 3A, Table 1; to represent yellowhammers and pine buntings in the phylogeny, we used the most divergent sequences, denoted with asterisks, in Fig. 2A). The genetic distance between the most divergent haplotypes in the pine bunting/yellowhammer complex is 16-fold less than that between the pine bunting and another closely-related species, the chestnut-breasted bunting (*E. stewarti*). These results closely parallel the patterns seen at the mitochondrial cyt*b* gene, as reported by Alström *et al.* (2007).

CHD1Z SEQUENCES

Sequences from 173 samples revealed five single nucleotide polymorphisms and one insertion/deletion ('indel') within 613 bp of the *CHD1Z* intron. Positions of these polymorphisms in the sequence, and the alternative characters at those positions, are: 132:A/G, 160:T/O, 241:T/C, 379:G/A, 495:T/C, 501:T/A, where ATGC represent mononucleotides and O represents a deletion. The software FASTPHASE (Scheet & Stephens, 2006) revealed that the genotypes most likely consist of 11 haplotypes, five of which are quite rare (Fig. 2B). Common haplotypes are shared between species, indicating that *CHD1Z* would be of little use in characterizing the genetic origin of individuals in the hybrid zone. We therefore focused our analysis on allopatric yellowhammers (75 samples) and pine buntings (36 samples), addressing the question of how genetically divergent the two groups are.

Φ_{ST} between the allopatric groups based on *CHD1Z* is 0.176, which is significantly different than zero ($P < 10^{-5}$) and much larger than Φ_{ST} based on *ND2* (0.025, see above). This is a surprising result because, under neutral theory, estimates of Φ_{ST} based on mitochondrial sequences are expected to exceed estimates of Φ_{ST} based on nuclear genes (Zink & Barrowclough, 2008).

The haplotype network shows multiple possible pathways by which haplotypes might have arisen, an apparent pattern of reticulate evolution suggesting a role for recombination in generating new haplotypes (Fig. 2B). Explaining the pattern without recombination would require postulating that several cases of point mutations arose multiple times, which seems unlikely. The two most highly diverged of the common haplotypes differ substantially in frequency between the species, with GTTGTT being more common in yellowhammers (chi-squared contingency test: $\chi^2 = 10.7$, $P = 0.002$) and AOCGCA more common in pine buntings ($\chi^2 = 26.0$, $P < 10^{-6}$; Fig. 2B). Taken together, a possible explanation for these patterns is that GTTGTT is an ancestral yellowhammer allele, whereas AOCGCA is an ancestral pine bunting allele. All of the other haplotypes, with the sole exception of the low-frequency haplotype AOCACA, could have arisen through hybridization and recombination without any novel point mutation.

For the purpose of estimating the *CHD1Z* relationships of the yellowhammer and pine bunting with four other bunting species, we took the common haplotypes in the two species (GTTGTT for yellowhammer and AOCGCA for pine bunting) as representing these species, respectively. In contrast with the mtDNA result, *CHD1Z* shows a genetic distance between the yellowhammer and pine bunting comparable with other closely-related species, *E. cirlus* and *E. stewarti* (Fig. 3B, Table 1). A comparison of the *CHD1Z* distances and *ND2* distances for the other species pairs that were in both datasets (*E. calandra*, *E. hortulana*, *E. cirlus*, and *E. stewarti*), and assum-

Table 1. Pairwise genetic distances (see Material and Methods for details) between the ten species of *Emberiza* genus for mtDNA (*ND2* gene, below the diagonal) and Z-chromosome (*CHD1Z* intron, above the diagonal)

	1	2	3	4	5	6	7	8	9	10
1. *Emberiza calandra*	–					0.023	0.031	0.028	0.032	0.028
2. *Emberiza tristrami*	0.394	–								
3. *Emberiza aureola*	0.419	0.140	–							
4. *Emberiza spodocephala*	0.463	0.126	0.117	–						
5. *Emberiza cioides*	0.249	0.404	0.456	0.466	–					
6. *Emberiza hortulana*	0.313	0.505	0.545	0.545	0.234	–	0.016	0.012	0.015	0.012
7. *Emberiza cirlus*	0.346	0.524	0.510	0.545	0.302	0.308	–	0.010	0.010	0.010
8. *Emberiza stewarti*	0.265	0.466	0.471	0.490	0.243	0.213	0.179	–	0.010	0.007
9. *Emberiza citrinella*	0.307	0.490	0.562	0.558	0.275	0.243	0.216	0.084	–	0.007
10. *Emberiza leucocephalos*	0.307	0.494	0.566	0.563	0.282	0.239	0.222	0.082	0.005	–

For yellowhammers (*E. citrinella*) and pine buntings (*E. leucocephalos*), haplotypes indicated with asterisks in Fig. 2 were used to calculate genetic distances.

ing an *ND2* molecular clock of 2% per million years, results in *CHD1Z* clock estimates ranging from 0.1% (*E. hortulana* – *E. stewarti*) to 0.2% (*E. calandra* – *E. stewarti*) per million years. Applying this range of estimates to the divergence between the two most common haplotypes leads to an estimated time of divergence of these two haplotypes of approximately 3–6 million years ago.
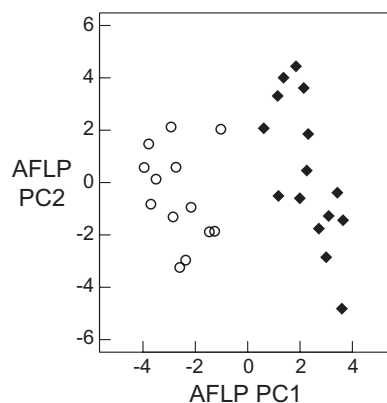
AFLP ANALYSIS

Genomic differentiation between taxa

AFLP analysis of allopatric yellowhammers and allopatric pine buntings reveals a moderately strong signal of divergence in their nuclear genomes. Our analysis of five AFLP primer pairs (AFLP dataset 1) revealed 367 AFLP markers in 13 allopatric yellowhammers and 15 allopatric pine buntings. Of these, 63 markers (17.2%) were variable. A PCA (Fig. 4) reveals that the two taxa are clearly separated along the first principal component (PC1), which explains 11.4% of the variance in the dataset. This difference in PC1 is highly significant (*t*-test: $t_{26} = 13.57$, $P < 10^{-12}$). F_{ST}, defined as the fraction of variance in AFLP signatures that is explained by the difference between two groups, is 0.078 based on estimated allele frequencies (using AFLP-SURV) and 0.140 based on band frequencies (using ARLEQUIN), both of which are significantly different from zero ($P < 10^{-5}$ and $P < 10^{-4}$, respectively). Although the overall signal of differentiation between the species is clear, it is due to a relatively small number of markers (Fig. 5), a pattern that is generally expected for multilocus genetic data (Whitlock, 2008).

AFLP analysis of hybridization

By examining bands produced by ten AFLP primer combinations, we identified 20 markers that showed

Figure 4. Variation in amplified fragment length polymorphism (AFLP) signatures (based on AFLP dataset 1) among allopatric yellowhammers (open circles) and pine buntings (filled diamonds), illustrated using principal components analysis. All variable markers (63 total, not chosen based on their pattern of variation but simply included if they were variable) from five primer combinations were used (these were only determined for allopatric samples; see Material and Methods). PC1 explains 11.4% of the variation, while PC2 explains 8.7%. There is no overlap between the two species in their AFLP signatures, and the difference in PC1 is highly significant (*t*-test: $t_{26} = 13.57$, $P < 10^{-12}$).

strong frequency differences between 26 allopatric yellowhammers and 15 allopatric pine buntings (Table 2; see Material and Methods). A PCA of these 20 markers on all phenotypic groups (AFLP dataset 2) reveals that phenotypic hybrids tend to have intermediate AFLP signatures (Fig. 6), and the STRUCTURE analysis produced similar results (see Supporting information, Fig. S2). Many of the phenotypic hybrids are generally more similar to allopatric yellowhammers than to allopatric pine buntings.
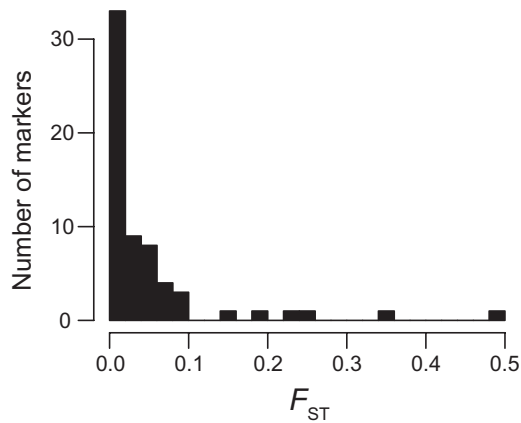
Figure 5. Distribution of F_{ST} values between allopatric yellowhammers and allopatric pine buntings for 63 variable amplified fragment length polymorphism (AFLP) markers from five primer combinations (AFLP dataset 1). F_{ST} was calculated based on allele frequencies estimated using AFLP-SURV (Vekemans, 2002). Most markers differ little in frequency between the species, but a small percentage show a strong difference (seven out of 63, or 11%, show an F_{ST} larger than 0.1).

DISCUSSION

GENETIC DIFFERENTIATION BETWEEN THE SPECIES

Yellowhammers and pine buntings differ noticeably in plumage, song, and geographic distributions (Panov *et al.*, 2003, 2007; Rubtsov, 2007), suggesting a long period of allopatric divergence. Thus, we expected to find sizeable genetic divergence between them. Surprisingly, patterns of mtDNA variation were extraordinarily similar between the species, with many shared haplotypes, a nonsignificant Φ_{ST} of 0.025, and a corrected sequence divergence of only 0.00054%. The estimated mitochondrial divergence date between species based on both the isolation model and the isolation with migration model is only approximately 30 000 years ago, which is an extraordinarily short time compared to most estimated divergence times between avian sister species, the bulk of which are between one and six million years ago (Weir & Schluter, 2007). A variety of analyses (Tajima's D, Fu's F_S, Ewens–Watterson–Slatkin's exact test, and IMa results) all indicate that mtDNA has undergone substantial population growth and/or selective

Table 2. Identities of the 20 informative AFLP fragments that differ in frequency (see Material and Methods for criteria) between allopatric yellowhammers and pine buntings

*Eco*RI primer* (NNN-3′)	*Mse*I primer† (NNN-3′)	Approximate fragment length	Frequency in yellowhammers	Frequency in pine buntings
AAC	CTT	112	0.19	0.73
AAC	CTT	135	0.62	0.07
ACA	CTA	231	0.08	0.53
ACA	CTA	406	0.27	0.53
ACC	CAA	65	0.31	0.53
ACC	CAA	223	0.38	0.67
ACC	CAA	323	0.00	1.00
ACC	CAT	468	0.19	0.80
ACT	CAA	64	0.46	1.00
ACT	CAA	225	0.31	0.60
ACT	CAA	226	0.27	0.80
ACT	CAA	236	0.23	0.60
ACT	CAA	251	0.23	0.67
ACT	CAT	176	0.38	1.00
ACT	CAT	182	0.15	0.60
ACT	CAT	279	0.31	0.53
ACT	CAT	303	0.31	0.60
AGC	CTT	245	0.58	0.20
AGG	CTA	187	0.69	0.13
AGG	CTA	552	0.15	0.67

*EcoRI primer: 5′-GACTGCGTACCAATTCNNN-3′.
†*Mse*I primer: 5′-GATGAGTCCTGAGTAANNN-3′.
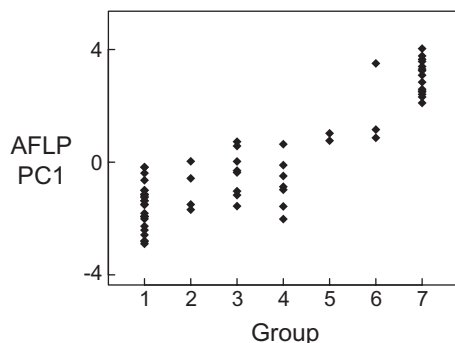
Figure 6. Genetic variation among yellowhammers, pine buntings, and phenotypic hybrids. The primary axis of variation (PC1) in a principal components analysis of 20 informative amplified fragment length polymorphism (AFLP) markers (AFLP dataset 2) is shown for seven phenotypic categories (Panov *et al.*, 2003), with each diamond representing a single individual. Group numbers along the horizontal axis indicate (1) allopatric yellowhammers, (2) yellowhammers from the contact zone, (3) yellow hybrids, (4) white hybrids, (5) pine buntings with slight hybrid phenotypes, (6) pine buntings in or near the contact zone, and (7) allopatric pine buntings. PC1 explains 20.1% of the variance, and varies significantly among phenotypic groups (analysis of variance: $F = 65.98$, d.f. = 6 and 58, $P < 10^{-15}$). AFLP variation clearly distinguishes the two species (*t*-test between allopatric groups: $t_{39} = -19.56$, $P < 10^{-15}$), and hybrids have a range of intermediate values.

sweeps, and that the pattern in yellowhammers and pine buntings are remarkably similar.

In contrast to mtDNA, nuclear markers do show moderately strong divergence between yellowhammers and pine buntings. Significant divergence was observed at the nuclear *CHD1Z* intron in frequencies of alleles, although alleles were shared between the species. In AFLP markers, there is a clear signal of divergence, although it is relatively modest compared to the amounts of AFLP divergence between other avian species pairs. Our F_{ST} estimate based on AFLP band frequencies is 0.14, whereas this measurement of divergence is 0.18 between greater and lesser spotted eagles (*Aquila clanga* and *Aquila pomarina*; Helbig *et al.*, 2005), 0.38 between white-winged crossbills and Hispaniolan crossbills (*Loxia leucoptera* and *Loxia megaplaga*; Parchman *et al.*, 2006), 0.4 between two reproductively isolated taxa of greenish warbler [*P. t. viridanus* and *P. (t.) plumbeitarsus*; Irwin *et al.*, 2005), and 0.42 between two cryptic species of winter wrens (*Troglodytes troglodytes* and *Troglodytes (Troglodytes) pacificus*; Toews & Irwin, 2008]. In three of these other cases, divergence in mtDNA has been assessed and is quite substantial (eagles: 1.75% in cyt*b*; warblers: 5% in control region; wrens: 6.2% in *ND2*).

These findings lead to two strong conclusions. First, the two species do differ genetically, suggesting that the two species are evolutionarily significant units that have experienced some period of relatively independent evolution. Second, this genetic difference can be clearly seen only in approximately 10% of the genome (Fig. 5). These results demonstrate that estimates of genetic divergence can differ dramatically between different molecular markers, and that a small but important subset of the genome can differ markedly even when most of the genome does not. It is possible that the few AFLP markers that show high divergence between the species are closely linked to genes that are under divergent selection in the two species.

These patterns can be explained in two ways. First, if we were to assume that the mitochondrial relationships are representative of true history, we would have to conclude that yellowhammers and pine buntings shared a common ancestor approximately 30 000 years ago (95% CI = 14 000–97 000 years). Under this scenario, the differentiation in phenotypes and in some AFLP markers would have occurred extremely rapidly, most likely as a result of strong selection. This would qualify as one of the fastest known cases of bird speciation, which, in most cases, takes more than a million years (Price, 2008). This scenario is extremely difficult to reconcile with the noticeable divergence in plumage colour, songs, ecology, and nuclear DNA.

Second, a much more parsiminoious explanation of these patterns is recent introgression of mtDNA between divergent forms. Recent hybridization might have introduced mtDNA from one species into the other, and that mitochondrial clade might have then become fixed in both species. This process can in theory occur rather easily, even when there is fairly strong selection against hybrids (Takahata & Slatkin, 1984). The smaller effective population size of mtDNA compared to nuclear DNA makes mtDNA particularly susceptible to fixation of foreign haplotypes (Funk & Omland, 2003). A selective advantage of one type of mitochondria over another could have hastened this process (Grant, Spies & Canino, 2006; Dowling *et al.*, 2008). This process of introgression, drift, and possible selection would also apply to nuclear genes. Over a long period of hybridization, much of the genome could have become homogenized between the species, whereas parts that were strongly selected in different directions continued to diverge between the species. We suggest that this second scenario, of introgression between two highly distinctive species, is easier to reconcile with the observed patterns.

The estimated time at which this mitochondrial introgression occurred (14 000–97 000 years ago) is difficult to reconcile with paleoclimatological history

because this span of time occurred during the last major glacial period, which began approximately 110 000 years ago and ended approximately 12 000 years ago (Adams & Faure, 1997). If yellowhammers and pine buntings were confined to separate refugia (e.g. the former in Europe or central Asia, the latter in south-east Asia), it might be unlikely that hybridization would have occurred during this major glacial period. However, there is evidence that the period from 55 000–25 000 years ago was relatively mild, with evidence for some wooded vegetation across Siberia (Adams & Faure, 1997). Hence, it is possible that yellowhammers and pine buntings did hybridize during that period, which would be consistent with the estimates from IMa. This is the first study that we of aware of to provide evidence of such genetic contact between western and eastern Palearctic refugial forms during that more mild phase of the last glacial period. An important caveat is that the IMa analysis assumes selective neutrality; it is possible that introgression of mtDNA may have occurred much more recently through the actions of selection.

CHD1Z patterns are remarkably supportive of this introgression scenario rather than a scenario of recent speciation. The *CHD1Z* haplotype network strongly suggests a recombinant origin of many haplotypes; the alternative explanation of so many identical mutations occurring multiple times is implausible. The two most common haplotypes (GTTGTT, the most common haplotype in yellowhammers; and AOCGCA, the most common haplotype in pine buntings) are also the most divergent of all haplotypes (except for one rare haplotype, AOCACA), suggesting that they represent the haplotypes ancestral in the two species. These two haplotypes are roughly as divergent as we might expect sister species of *Emberiza* to be, given the genetic distances throughout the *Emberiza* phylogeny (e.g. the GTTGTT and AOCGCA haplotypes are as divergent as the GTTGTT haplotype is from *E. stewarti*). All of the other haplotypes, with the sole exception of AOCACA, are simple combinations of these major haplotypes. If mutation rates were high, we would expect to see many haplotypes that differ from a common one by a single mutation, a pattern not seen. Overall, the patterns are suggestive of recombination between highly divergent forms of *CHD1Z*. These forms likely correspond to distinct yellowhammer and pine bunting forms of *CHD1Z* before hybridization, recombination, and introgression led to the current mixed pattern. The estimated time of divergence between these haplotypes, three to six million years ago, provides a very approximate maximum estimate for when the two species started diverging. It should be noted that the Z chromosome may play an especially large role in speciation in female-heterogametic groups such as birds, and may be less subject to genetic mixing between incipient species compared to other parts of the genome (Qvarnström & Bailey, 2009).

MOLECULAR IDENTIFICATION OF HYBRIDIZATION

Our analysis strongly supports the hypothesis that hybrids can be identified based on appearance, as individuals with apparent hybrid phenotypes usually had intermediate AFLP signatures. This result confirms the utility of screening large numbers of AFLP markers between two allopatric samples, and then using a subset of markers that are most divergent in frequency to test whether there is hybridization in sympatry and to compare the genetic signatures of birds in the hybrid zone to those in allopatry. It would require much further analysis, as well as larger sample sizes, to accurately assign birds in the contact zone to hybrid categories, such as F1, F2, backcross, etc.; this goal is beyond the scope of the present study.

Most phenotypic hybrids are more similar to allopatric yellowhammers than they are to allopatric pine buntings. This pattern suggests that gene flow from the contact zone has affected allopatric European populations of yellowhammers to a greater extent than allopatric populations of pine buntings. Perhaps backcrosses with yellowhammers have higher fitness than backcrosses with pine buntings, or perhaps yellowhammers have larger dispersal distances, leading to greater gene flow across their range. Whatever the cause of this pattern, it is in accordance with the distribution of various colour phenotypes within the yellowhammer breeding range. Populations of yellowhammers in the eastern part of their range tend to have some chestnut coloration on the throat; this and other plumage traits have led some authors to treat them as a distinct subspecies (*Emberiza citrinella erythrogenys*; Cramp & Perrins, 1994; Byers, Curson & Olsson, 1995), although this variation is subtle and varies clinally from the traits of the European *Emberiza citrinella citrinella*. The suggestion of Panov *et al.* (2003) that *E. c. erythrogenys* could be a product of ancient hybridization between yellowhammers and pine buntings appears to be consistent with our molecular data. These results indicate the possibility of hybridization being a creative force, rearranging gene combinations to create novel phenotypes that are relatively successful over time (Arnold, 1997).

MITOCHONDRIAL INTROGRESSION

There is growing recognition that introgression of mtDNA between species might be quite common

(Good *et al.*, 2008), as is suggested by the high rate of species paraphyly and polyphyly in interspecific mitochondrial gene trees (Funk & Omland, 2003). Cases of partial introgression of heterospecific mtDNA has been observed in a variety of species (Rohwer, Bermingham & Wood, 2001; Weckstein *et al.*, 2001; Good *et al.*, 2003, 2008; Deffontaine *et al.*, 2005; Melo-Ferreira *et al.*, 2005; Plötner *et al.*, 2008), although the case of the yellowhammer and pine bunting is unusual in the magnitude of phenotypic differentiation between the species and the extent of mitochondrial blending, with the type of one species apparently completely replacing that of the other. Such complete replacement is remarkable between species that differ so noticeably in appearance. Cases of complete replacement might be much more common than presently thought because they are difficult to detect. When mtDNA has only partially introgressed, it is detectable because members of one species have two very divergent forms of mtDNA, one of which is similar to the other species (Plötner *et al.*, 2008). When complete replacement has occurred, there are no surviving examples of the extinct haplotype group to reveal the presence of introgression; the remaining pattern is simply one of mtDNA similarity between the two species, which could be mistakenly interpreted as recent population splitting.

It is increasingly recognised that mitochondrial DNA could often be under selection, challenging the assumption of neutrality that is common to many analytical methods used in studies of phylogeography, speciation, and conservation genetics (Ballard & Whitlock, 2004; Bazin *et al.*, 2006; Dowling *et al.*, 2008). Selection can occur in many ways, including local adaptation of different mtDNA variants to different environmental conditions, coevolution between mitochondrial and nuclear genes (Dowling *et al.*, 2008), and selective sweeps of universally favoured mitochondrial mutations. We suggest that a selective sweep is the likely explanation for the mitochondrial introgression between yellowhammers and pine buntings, as it is implausible that a complete replacement of one species' mtDNA by another occurred by drift alone. A possible scenario consistent with the data is as follows: a favourable mutation arose in mtDNA of one of the species and rose in frequency as a result of selection. A short time later, hybridization introduced this variant to the other species. The favourable variant then continued to grow in frequency within both species, and hybridization continued to transfer variants of this favourable mtDNA between the species. Eventually, all other variants of mtDNA vanished from both populations. Plötner *et al.* (2008) suggested that a variant of mtDNA that arose in one species of water frog has spread to another species because it is selectively advantageous to both species in the more northerly parts of their ranges. Parallel shifts in climate or other environmental variables may be quite common for sister species; during such shifts, the mitochondria in one of the species may become better adapted to the new conditions and, assisted by hybridization, sweep to fixation in both species.

These findings highlight the challenges inherent in the use of molecular variation to reconstruct biogeographic history and identify evolutionarily significant units. Because of both shared ancestral polymorphism and introgressive hybridization, two groups that have experienced much independent evolution can be similar in most of their genome while still differing in those parts that are under divergent selection. These differences can be maintained even when two groups hybridize extensively, given strong enough selection for alternative alleles in the two groups (Barton & Hewitt, 1985, 1989; Wu, 2001). Hybridization is being increasingly recognized for its potentially creative role in evolution because differential introgression can lead to novel gene combinations (Arnold, 1997; Mallet, 2005; Price, 2008). In the case of yellowhammers and pine buntings, mtDNA appears to have introgressed between two phenotypically distinct forms that differ in parts of their nuclear genomes. Thus, mtDNA, as well as many parts of the nuclear genome, would provide a misleading picture of the history of the species complex.

Our findings should be considered in light of the traditional emphasis on the use of mtDNA in taxonomy at the species and subspecies level. For example, Zink (2004) argued that phenotypically-defined subspecies should not be considered evolutionarily significant units if they do not correspond to distinct mitochondrial clades. Such reasoning, when applied to our study group, would lead to the mistaken conclusion that yellowhammers and pine buntings are a single genetically undifferentiated group. Plumage patterns (Panov *et al.*, 2003, 2007), song (Rubtsov, 2007), AFLP, and *CHD1Z* lead to a different conclusion. Introgression is just one of the many reasons that patterns of relationships in mtDNA may not accurately represent relationships in nuclear DNA (Irwin, 2002; Ballard & Whitlock, 2004; Chan & Levin, 2005; Edwards *et al.*, 2005; Jennings & Edwards, 2005; Shaffer & Thomson, 2007). We predict that distinct phenotypic groupings will increasingly be supported by multilocus nuclear-DNA surveys based on AFLP (Bensch, Åkesson & Irwin, 2002), single nucleotide polymorphisms (Shaffer & Thomson, 2007), or multiple gene sequences (Jennings & Edwards, 2005), even when mtDNA does not differ between groups.

REFERENCES

Adams JM, Faure H, eds. 1997. *Review and atlas of palaeovegetation: preliminary land ecosystem maps of the world since the Last Glacial Maximum*. Oak Ridge, TN: Oak Ridge National Laboratory. Accessed online 7 August 2008. Available at http://www.esd.ornl.gov/projects/qen/adams1.html

Alström P, Olsson U, Lei F, Wang H-t, Gao W, Sundberg P. 2007. Phylogeny and classification of the Old World Emberizini (Aves, Passeriformes). *Molecular Phylogenetics and Evolution* **47:** 960–973.

Arnold ML. 1997. *Natural hybridization and evolution*. Oxford: Oxford University Press.

Avise JC. 2004. *Molecular markers, natural history and evolution*. Sunderland, MA: Sinauer Associates.

Ballard JWO, Whitlock MC. 2004. The incomplete natural history of mitochondria. *Molecular Ecology* **13:** 729–744.

Barton NH, Hewitt GM. 1985. Analysis of hybrid zones. *Annual Review of Ecology and Systematics* **16:** 113–148.

Barton NH, Hewitt GM. 1989. Adaptation, speciation and hybrid zones. *Nature* **341:** 497–503.

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

Bensch S, Åkesson M. 2005. Ten years of AFLP in ecology and evolution: why so few animals? *Molecular Ecology* **14:** 2899–2914.

Bensch S, Åkesson S, Irwin DE. 2002. The use of AFLP to find an informative SNP: genetic differences across a migratory divide in willow warblers. *Molecular Ecology* **11:** 2359–2366.

Byers C, Curson J, Olsson U. 1995. *Sparrows and buntings: a guide to the sparrows and buntings of North America and the world*. New York, NY: Houghton Mifflin.

Chan KMA, Levin SA. 2005. Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. *Evolution* **59:** 720–729.

Clement M, Posada D, Crandall K. 2000. TCS: a computer program to estimate gene genealogies. *Molecular Ecology* **9:** 1657–1660.

Cramp S, Perrins CM, eds. 1994. *The birds of the western Palearctic*, Vol. 9. Oxford: Oxford University Press.

Deffontaine V, Libois R, Kotlik P, Sommer R, Nieberding C, Paradis E, Searle JB, Michaux JR. 2005. Beyond the Mediterranean peninsulas: evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (*Clethrionomys glareolus*). *Molecular Ecology* **14:** 1727–1739.

Dowling DK, Friburg U, Lindell J. 2008. Evolutionary implications of non-neutral mitochondrial genetic variation. *Trends in Ecology and Evolution* **23:** 546–554.

Drovetski SV, Zink RM, Fadeev IV, Nesterov EV, Koblik EA, Red'kin YA, Rohwer S. 2004. Mitochondrial phylogeny of Locustella and related genera. *Journal of Avian Biology* **35:** 105–110.

Edwards SV, Kingan SB, Calkins JD, Balakrishnan CN, Jennings WB, Swanson WJ, Sorensen MD. 2005. Speciation in birds: genes, geography, and sexual selection. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 6550–6557.

Excoffier L, Laval G, Schneider S. 2005. Arlequin ver. 3.0: an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online* **1:** 47–50.

Excoffier L, Laval G, Schneider S. 2006. Arlequin ver 3.1 user manual. Available at http://cmpg.unibe.ch/software/arlequin3

Falush D, Stephens M, Pritchard JK. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology Notes* **7:** 574–578.

Frenzel B. 1968. The Pleistocene vegetation of northern Eurasia. *Science* **161:** 637–649.

Fridolfsson A-K, Ellegren H. 1999. A simple and universal method for molecular sexing of non-ratite birds. *Journal of Avian Biology* **30:** 116–121.

Fu Y-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* **147:** 915–925.

Funk DJ, Omland KE. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* **34:** 397–423.

Good JM, Demboski JR, Nagorsen DW, Sullivan J. 2003. Phylogeography and introgressive hybridization: chipmunks (genus *Tamias*) in the northern Rocky Mountains. *Evolution* **57:** 1900–1916.

Good JM, Hird S, Reid N, Demboski JR, Steppan SJ, Martin-Nims TR, Sullivan J. 2008. Ancient hybridization and mitochondrial capture between two species of chipmunks. *Molecular Ecology* **17:** 1313–1327.

Grant WS, Spies IB, Canino MF. 2006. Biogeographic evidence for selection on mitochondrial DNA in North Pacific walleye pollock *Theragra chalcogramma*. *Journal of Heredity* **97:** 571–580.

Haffer J. 1989. Parapatrische Vogelarten der paläarktischen Region. *Journal für Ornithologie* **130:** 475–512.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41:** 95–98.

Helbig AJ, Seibold I, Kocum A, Liebers D, Irwin J, Bergmanis U, Meyburg BU, Scheller W, Stubbe M, Bensch S. 2005. Genetic differentiation and hybridization between greater and lesser spotted eagles (Accipitriformes: *Aquila clanga*, *A. pomarina*). *Journal of Ornithology* **146:** 226–234.

Hey J. 2007. Using the IMa program. Distributed by the author. Available at http://lifesci.rutgers.edu/~heylab/heylabsoftware.htm

Hey J, Nielsen R. 2007. Integration with the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proceedings of the National Academy of Sciences of the United States of America* **104:** 2785–2790.

Irwin DE. 2002. Phylogeographic breaks without geographic barriers to gene flow. *Evolution* **56:** 2383–2394.

Irwin DE, Bensch S, Irwin JH, Price TD. 2005. Speciation by distance in a ring species. *Science* **307:** 414–416.

Irwin DE, Bensch S, Price TD. 2001. Speciation in a ring. *Nature* **409:** 333–337.

Irwin DE, Irwin JH. 2005. Siberian migratory divides: the role of seasonal migration in speciation. In: Greenberg R, Marra PP, eds. *Birds of two worlds: the ecology and evolution of migratory birds*. Baltimore, MD: Johns Hopkins University Press, 27–40.

Jennings WB, Edwards SV. 2005. Speciational history of Australian grass finches (*Poephila*) inferred from thirty gene trees. *Evolution* **59:** 2033–2047.

Kondo B, Peters JL, Rosensteel BB, Omland KE. 2008. Coalescent analyses of multiple loci support a new route to speciation in birds. *Evolution* **62:** 1182–1190.

LI-COR Biosciences. 2003. *Applications manual: model 4300 DNA analyzer*. Lincoln, NE: LI-COR Biosciences.

Lovette IJ. 2004. Mitochondrial dating and mixed-support for the '2% rule' in birds. *Auk* **121:** 1–6.

Lynch M, Milligan B. 1994. Analysis of population genetic structure with RAPD markers. *Molecular Ecology* **3:** 91–99.

Mallet J. 2005. Hybridization as an invasion of the genome. *Trends in Ecology and Evolution* **20:** 229–237.

Melo-Ferreira J, Boursot P, Suchentrunk F, Ferrand N, Alves PC. 2005. Invasion from the cold past: extensive introgression of mountain hare (*Lepus timidus*) mitochondrial DNA into three other hare species in northern Iberia. *Molecular Ecology* **14:** 2459–2464.

Nazarenko AA. 1990. [Avifaunal interchange between south and north Asia at the eastern periphery of the continent: the last glacial-interglacial cycle.]. *Zhurnal Obshchei Biologii* **51:** 89–106 (in Russian).

Newton I. 2003. *The speciation and biogeography of birds*. London: Academic Press.

Nielsen R, Wakeley J. 2001. Distinguishing migration from isolation. A Markov chain Monte Carlo approach. *Genetics* **158:** 885–896.

Panov EN 1989. *[Hybridization and ethological isolation in birds.]* Moscow: Nauka (in Russian).

Panov EN, Rubtsov AS, Monzikov DG. 2003. Hybridization between yellowhammer and pine bunting in Russia. *Dutch Birding* **25:** 17–31.

Panov EN, Rubtsov AS, Mordkovich MV. 2007. [New data on interrelationships of two bunting species (*Emberiza citrinella*, *E. leucocephala*) interbreeding in zone of their ranges overlap.]. *Zoologicheskii Zhurnal* **86:** 1362–1378 (in Russian).

Parchman TL, Benkman CW, Britch SC. 2006. Patterns of genetic variation in the adaptive radiation of New World crossbills (Aves: *Loxia*). *Molecular Ecology* **15:** 1873–1887.

Plötner J, Uzzell T, Beerli P, Spolsky C, Ohst T, Litvinchuk SN, Guex G-D, Reyer H-U, Hotz H. 2008. Widespread unidirectional transfer of mitochondrial DNA: a case in western Palaearctic water frogs. *Journal of Evolutionary Biology* **21:** 668–681.

Price T. 2008. *Speciation in birds*. Greenwood Village, CO: Roberts and Company.

Pritchard JK, Stephens M, Donnelly P. 2000. Inference of population structure using multilocus genotype data. *Genetics* **155:** 945–959.

Qvarnström A, Bailey RI. 2009. Speciation through evolution of sex-linked genes. *Heredity* **102:** 4–15.

R Development Core Team. 2006. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available at http://www.R-project.org

Ravkin YS. 1973. *[The birds of north-eastern Altai.]* Novosibirsk, Russia: Nauka (in Russian).

Rogacheva H. 1992. *The birds of central Siberia*. Husum, Germany: Husum Druck- und Verlagsgesellschaft.

Rohwer S, Bermingham E, Wood C. 2001. Plumage and mitochondrial DNA haplotype variation across a moving hybrid zone. *Evolution* **55:** 405–422.

Rubinoff D, Holland BS. 2005. Between two extremes: mitochondrial DNA is neither the panacea nor the nemesis of phylogenetic and taxonomic inference. *Systematic Biology* **54:** 952–961.

Rubtsov AS. 2007. [Variation in songs of yellowhammer and pine bunting (*Emberiza citrinella*, *E. leucocephala*) as an evidence for the population structure dynamics and evolutionary history of species.] *Zoologicheskiĭ. Zhurnal* **86:** 863–876 (in Russian).

Sæther B-E, Lande R, Engen S, Weimerskirch H, Lillegård M, Altwegg R, Becker PH, Bregnballe T,

Brommer JE, McCleery RH, Merilä J, Nyholm E, Rendell W, Robertson RR, Tryjanowski P, Visser ME. 2005. Generation time and temporal scaling of bird population dynamics. *Nature* **436:** 99–102.

Scheet P, Stephens M. 2006. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *American Journal Human Genetics* **78:** 629–644.

Schmidt HA, Strimmer K, Vingron M, von Haeseler A. 2002. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **18:** 502–504.

Seutin G, White BN, Boag PT. 1991. Preservation of avian blood and tissue samples for DNA analyses. *Canadian Journal of Zoology* **69:** 82–90.

Shaffer HB, Thomson RC. 2007. Delimiting species in recent radiations. *Systematic Biology* **56:** 896–906.

Slatkin M. 1996. A correction to the exact test based on the Ewens sampling distribution. *Genetical Research* **68:** 259–260.

Sorenson MD, Quinn TW. 1998. Numts: a challenge for avian systematics and population biology. *Auk* **115:** 214–221.

Svensson EI, Kristoffersen L, Oskarsson K, Bensch S. 2004. Molecular population divergence and sexual selection on morphology in the banded demoiselle (*Calopteryx splendens*). *Heredity* **93:** 423–433.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123:** 585–595.

Takahata N, Slatkin M. 1984. Mitochondrial gene flow. *Proceedings of the National Academy of Sciences of the United States of America* **81:** 1764–1767.

Toews DPL, Irwin DE. 2008. Cryptic speciation in a Holarctic passerine revealed by genetic and bioacoustic analyses. *Molecular Ecology* **17:** 2691–2705.

Vekemans X. 2002. AFLP-SURV, *Version 1.0*. Distributed by the author. Belgium: Laboratoire de Génétique et Ecologie Végétale, Université Libre de Bruxelles.

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M. 1995. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research* **23:** 4405–4414.

Weckstein JD, Zink RM, Blackwell-Rago RC, Nelson DA. 2001. Anomalous variation in mitochondrial genomes of White-crowned (*Zonotrichia leucophrys*) and Golden-crowned (*Z. atricapilla*) Sparrows: pseudogenes, hybridization, or incomplete lineage sorting? *Auk* **118:** 231–236.

Weir JT. 2006. Divergent patterns of species accumulation in lowland and highland neotropical birds. *Evolution* **60:** 842–855.

Weir JT, Schluter D. 2007. The latitudinal gradient in recent speciation and extinction rates in birds and mammals. *Science* **315:** 1928–1933.

Weir JT, Schluter D. 2008. Calibrating the avian molecular clock. *Molecular Ecology* **17:** 2321–2328.

Whitlock MC. 2008. Evolutionary inference from QST. *Molecular Ecology* **17:** 1885–1896.

Won YJ, Hey J. 2005. Divergence population genetics of chimpanzees. *Molecular Biology and Evolution* **22:** 297–307.

Wu C-I. 2001. The genic view of the process of speciation. *Journal of Evolutionary Biology* **14:** 851–865.

Zhivotovsky LA. 1999. Estimating population structure in diploids with multilocus dominant DNA markers. *Molecular Ecology* **8:** 907–913.

Zink RM. 2004. The role of subspecies in obscuring avian biological diversity and misleading conservation policy. *Proceedings of the Royal Society of London Series B, Biological Sciences* **271:** 561–564.

Zink RM, Barrowclough GF. 2008. Mitochondrial DNA under siege in avian phylogeography. *Molecular Ecology* **17:** 2107–2121.

SUPPORTING INFORMATION

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

Figure S1. Output of demographic modelling based on mitochondrial *ND2* haplotypes and the isolation with migration model as implemented in IMa. Shown are posterior probability distributions of (a) time of population splitting, (b) mean time of migration events, (c) post-split migration rates, and (d) ancestral population size of the common ancestor of yellowhammers and pine buntings. The runs included two with migration rate set to zero (solid and dotted black lines) and two that allowed migration after splitting (solid and dotted grey lines). In (b) and (c), migration from yellowhammers to pine buntings (going forward in time) is represented by light grey, and the reverse is represented by dark grey. Peaks of the probability distributions as well as 95% confidence intervals (CIs) were similar between runs for both t [peak and 95% CIs: no migration, run 1: 0.33 (0.21–0.50); run 2: 0.33 (0.21–0.50); with migration, run 1: 0.32 (0.21–0.48); run 2: 0.33 (0.21–0.52)] and θ_A [no migration, run 1: 7.4 (3.3–19.9); run 2: 8.0 (3.3–20.5); with migration, run 1: 8.0 (3.3–20.5); run 2: 8.0 (3.9–20.5)]. Assuming a mtDNA mutation rate of 1% per million years and a generation time of 1.7 years (see Material and Methods), these estimates convert to a splitting time of approximately 32 000 years ago (95% CI = 20 000–48 000 years ago) and an ancestral effective population size of 110 000 (47 000–290 000). Migration events (b) had a mean time that was approximately halfway between the original split and the present, suggesting that

migration events were spread throughout the period of time subsequent to the split. The migration rate (c) could not be estimated confidently – values of migration rates ranging from $m = 0$ to 100 had similar probability.

Figure S2. Probabilities of population assignment, based on analysis of 20 informative amplified fragment length polymorphism (AFLP) markers (AFLP dataset 2) by STRUCTURE 2.2. Phenotypic categories are arranged along the horizontal axis (Fig. 6), and vertical axis represents the probability of assignment to the pine bunting cluster. STRUCTURE was run with an admixture model, a burn-in time of 10 000 steps and a run length of 10 000 steps. Because the markers were specifically chosen to best distinguish between allopatric yellowhammers and allopatric pine buntings, we expected STRUCTURE to reveal that the optimal number of genetic groups (k) was 2. To check, we ran STRUCTURE at a range of k values (1–7), and compared the likelihood of the data among the values of k. The highest likelihood was for $k = 2$ (log likelihoods for $k = 2$: –1068; $k = 1$: –1138; $k = 3$–7: –1074 to –2950). Ten separate runs of STRUCTURE at $k = 2$ produced almost identical results; hence, only the results from the first run are shown.

Table S1. Origin of samples used in the molecular analysis of yellowhammers, pine buntings, their hybrids, and related species.

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