MOLECULAR ECOLOGY

Molecular Ecology (2011) **20**, 3812–3822 doi: 10.1111/j.1365-294X.2011.05183.x

Hybrid speciation in sparrows I: phenotypic intermediacy, genetic admixture and barriers to gene flow

JO S. HERMANSEN, STEIN A. SÆTHER, TORE O. ELGVIN, THOMAS BORGE, ELIN HJELLE and GLENN-PETER SÆTRE

Centre for Ecological and Evolutionary Synthesis, Department of Biology, University of Oslo, PO Box 1066, Blindern, N-0316 Oslo, Norway

Abstract

Homoploid hybrid speciation is thought to require unusual circumstances to yield reproductive isolation from the parental species, and few examples are known from nature. Here, we present genetic evidence for this mode of speciation in birds. Using Bayesian assignment analyses of 751 individuals genotyped for 14 unlinked, nuclear microsatellite loci, we show that the phenotypically intermediate Italian sparrow (*Passer italiae*) does not form a cluster of its own, but instead exhibits clear admixture (over its entire breeding range) between its putative parental species, the house sparrow (*P. domesticus*) and the Spanish sparrow (*P. hispaniolensis*). Further, the Italian sparrow possesses mitochondrial (mt) DNA haplotypes identical to both putative parental species (although mostly of house sparrow type), indicating a recent hybrid origin. Today, the Italian sparrow has a largely allopatric distribution on the Italian peninsula and some Mediterranean islands separated from its suggested parental species by the Alps and the Mediterranean Sea, but co-occurs with the Spanish sparrow on the Gargano peninsula in southeast Italy. No evidence of interbreeding was found in this sympatric population. However, the Italian sparrow hybridizes with the house sparrow in a sparsely populated contact zone in the Alps. Yet, the contact zone is characterized by steep clines in species-specific male plumage traits, suggesting that partial reproductive isolation may also have developed between these two taxa. Thus, geographic and reproductive barriers restrict gene flow into the nascent hybrid species. We propose that an origin of hybrid species where the hybrid lineage gets geographically isolated from its parental species, as seems to have happened in this system, might be more common in nature than previously assumed.

Keywords: birds, homoploid hybrid speciation, hybridization, reproductive isolation

Received 29 October 2010; revision received 25 May 2011; accepted 1 June 2011

Introduction

Two hundred and fifty years after its introduction to the scientific literature (Linnaeus 1760 as cited in Rieseberg 1997), the role of natural hybridization in the evolution remains controversial (Arnold 1997; Seehausen 2004; Mallet 2005, 2008). Depending on the circumstances, this process can lead to outcomes ranging from despeciation at one end of a continuum to hybrid speciation at the other (Arnold 1997). Traditionally, botanists have emphasized the potentially creative roles of hybridization to a much larger extent than zoologists (Anderson 1949; Stebbins 1959), and hence a number of studies, largely descriptive ones, have been conducted on various aspects of the process in plants (Arnold 1997). Zoologists, on the other hand, have seen hybridization mainly as a 'reproductive mistake' (Fisher 1930; Mayr 1942; Arnold 1997; Mallet 2005, 2007, 2008) and have thus, until recently, largely neglected the creative

Correspondence: Glenn-Peter Sætre, Fax: (47) 228 54001;
E-mail: g.p.satre@bio.uio.no

potential of the phenomenon except for evaluating its importance in the reinforcement scenario of speciation (Dobzhansky 1937). Although seeing hybridization as of limited direct importance in speciation, zoologists have studied it as a means of understanding the processes leading to reproductive isolation, and hence speciation (Arnold 1997; Coyne & Orr 2004; Seehausen 2004). However, today most zoologists also view natural hybridization as a potentially creative evolutionary process (Arnold 1997, 2006).

Homoploid hybrid speciation (HHS) is the least studied among the creative outcomes of natural hybridization. In this mode of speciation, interbreeding between diverged lineages leads to the formation of a third, novel and stable lineage, without a change in chromosome number (Mallet 2007; Mavárez & Linares 2008; Nolte & Tautz 2010). However, there exists no real consensus on what exact criteria must be met for a lineage to be considered a hybrid species, e.g., with respect to the relative contribution of the parental genomes to the mixed lineage. In our opinion, the crucial distinction between hybrid speciation and introgression is whether or not the hybrid lineage has developed reproductive barriers against its parental taxa or not. Hence, even a contribution of only a few genes from one of the parental taxa would be sufficient for a lineage to be considered a hybrid species if these genes are causally linked to the build-up of reproductive isolation.

Few examples of HHS are known from nature, both in plants and animals (Mallet 2007; Mavárez & Linares 2008). For instance in birds, one of the most intensively studied taxa in speciation research where approximately 10% of all species are known to hybridize (Grant & Grant 1992), convincing evidence of HHS has until recently been lacking (Brelsford *et al.* 2011). This lack of documented examples of HHS may result from it being a genuinely rare phenomenon in nature; a hybrid lineage has to escape both gene flow and competition from its initially sympatric parental species, seemingly making it inherently less likely than most modes of bifurcating speciation. However, the lack of documented examples may also partly be an effect of the historical skepticism towards the process as alluded to above or an effect of logistic problems involved in detecting homoploid hybrid species (Jiggins *et al.* 2008; Mavárez & Linares 2008). Here, we test whether the Italian sparrow, for a long time proposed as a potential example of HHS in birds, really is of hybrid origin.

The Italian sparrow (*Passer italiae*) is a human-associated, seed-eating bird replacing the house sparrow (*P. domesticus*) on the Italian peninsula and a few Mediterranean islands (Summers-Smith 1988). It has been suggested to be of hybrid origin because the plumage of male individuals is intermediate to males of the house sparrow and the Spanish sparrow (*P. hispaniolensis*) (Zedlitz 1913; Meise 1936; Johnston 1969; Summers-Smith 1988; Töpfer 2006) (Fig. 1). Male Italian sparrows have a chestnut-coloured crown and nape, and white cheeks similar to the Spanish sparrow (house sparrows have a broad, grey band on the crown and nape, and grey cheeks), but a small bib and a brown-streaked back similar to the house sparrow (Spanish sparrows have a large black bib that extends all along the body flanks and a black- and yellow-streaked back) (Summers-Smith 1988). Interestingly, male *F*1-hybrids between house sparrows and Spanish sparrows resemble Italian sparrows (Macke 1965; Alonso 1984). Although the Italian sparrow is phenotypically uniform over most of its breeding range (from the southern slopes of the Alps in northern Italy and southwards excluding the Calabria peninsula), males in southwest Italy (the Calabria peninsula and especially Sicily) are approaching Spanish sparrows phenotypically, possessing larger bibs than those further north, and partially black-streaked flanks (Fig. 1) (Summers-Smith 1988; Töpfer 2006).

If the intermediate plumage of the Italian sparrow is a result of hybridization in the past, the Italian sparrow is consequently of homoploid hybrid origin because all three taxa share the same chromosome number ($2n = 76$) (Bulatova *et al.* 1972; Fulgione *et al.* 2000). However, not all intermediate-looking organisms are hybrids, and an intermediate plumage is not in itself evidence of hybrid origin. Consequently, the taxonomic status of the Italian sparrow has been much debated (Töpfer 2006).

Using genetic analyses of an extensive sample of birds, we test the homoploid hybrid speciation hypothesis against alternatives. First, the Italian sparrow may be the result of bifurcating speciation, and hence have a most recent common ancestor with either the house sparrow or the Spanish sparrow. If so, we expect to find (i) unique haplotypes in the Italian sparrow (amount depending on the time since divergence) and (ii) that it forms a monophyletic clade with either the house sparrow or the Spanish sparrow. If on the other hand, the Italian sparrow is of hybrid origin we expect to find (iii) relatively larger divergence between the house sparrow and Spanish sparrow than between the Italian sparrow and its putative parental species, (iv) the genome of the Italian sparrow to be a mixture of those found in the house sparrow and the Spanish sparrow and (v) few or no unique haplotypes in the Italian sparrow. However, if there is much introgression into the Italian sparrow under a bifurcation scenario, or long time has passed since the hybridization event(s) in the HHS scenario, our ability to draw conclusive inferences may be impeded. Incomplete lineage sorting may also

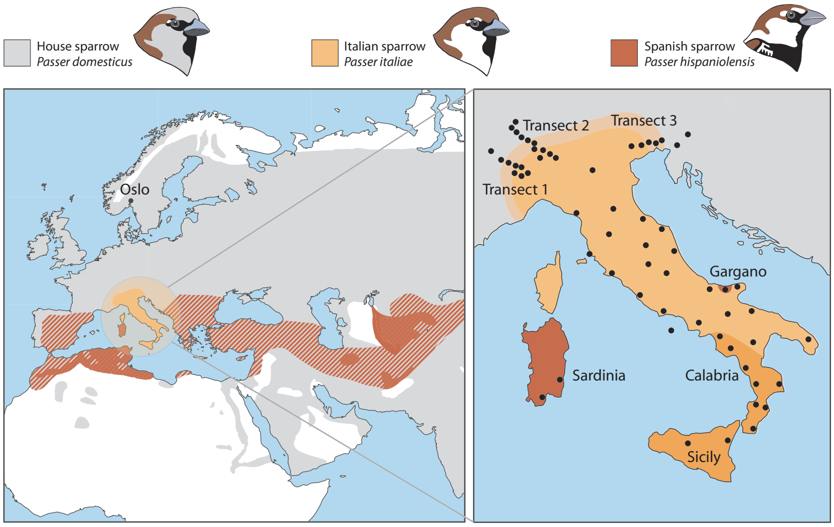

Fig. 1 Geographic distribution of the studied *Passer* taxa (left) and a zoom in on the Italian distribution (right). Black dots indicate sampling localities. Grey indicates the distribution of the house sparrow, orange indicates the distribution of the Italian sparrow and red indicates the distribution of the Spanish sparrow. Hatched areas indicate regions where the house sparrow and Spanish sparrow distribution overlap. Grey-orange indicates a hybrid zone in the Alps between the house sparrow and Italian sparrow. Transect 1, 2 and 3 refers to three sampling transects through this hybrid zone. The red colour on the Gargano peninsula refers to a sympatric pocket of Spanish sparrows within the Italian sparrow distribution. The darker orange colour on the Calabria peninsula and Sicily indicate an area where Italian sparrows look more similar to Spanish sparrows. Black dots refer to sampling localities (n = 62). One sampling locality is shown in the left panel (house sparrows from Oslo, Norway). Bird drawings indicate species-specific male plumage characteristics of the house sparrow, Italian sparrow and Spanish sparrow.

obscure the actual evolutionary relationships between the taxa because of retained ancestral polymorphisms.

Our sampling scheme includes (i) transects across an Italian sparrow–house sparrow contact zone in the Alps, (ii) Italian sparrow populations across the Italian peninsula and Sicily, (iii) an Italian sparrow–Spanish sparrow contact zone on the Gargano peninsula and finally (iv) allopatric populations of the putative parental species separated from any Italian sparrow populations by distance (house sparrows from Norway) or open bodies of water (Spanish sparrows from Sardinia) to avoid interpretational problems resulting from recent introgression. We think Norway is an adequate allopatric reference population for house sparrows as a recent article by Schrey *et al.* (2011) showed that house sparrows sampled in Norway were approximately equally diverged from the other house sparrow populations included in their study. Likewise, Sardinia is one of very few areas where the Spanish sparrow occurs allopatrically and which is also separated from other sparrow populations by a significant geographic barrier (open water).

In an accompanying article, we look further into the system by analysing variation at 15 genes (nine autosomal, five sex-linked and one mitochondrial) for individuals sampled from allopatric, phenotypically pure populations of the three taxa studied here (Elgvin *et al.* 2011).

Materials and methods

Study material

A total of 751 adult *Passer* sparrows were caught using mist nets at 62 localities in Italy, Switzerland, France and Slovenia (Fig. 1) as well as in Oslo, Norway during spring 2007–2008. A blood sample (20–50 μL) for use in genetic analyses was taken from each individual, by puncturing a brachial vein. DNA was extracted from blood samples stored in standard buffer using the E-Z 96® Blood DNA Kit (Omega Bio-Tek, Inc., Doraville, GA, USA) according to the manufacturer's instructions. Authorization to catch birds and take blood samples was obtained from the appropriate authorities in the respective countries.

Microsatellites and phenotypic data

All 751 sparrows were genotyped for 14 unlinked, nuclear microsatellite loci (Table S1, Supporting information). All polymerase chain reactions were carried out in 10 μL reactions containing 5 μL Qiagen Multiplex PCR Kit Master Mix (Qiagen, Valencia, CA, USA), 0.125 μL of each primer, 3 μL template DNA (concentrations ranging from 10 to 30 ng/μL). Samples were subjected to initial denaturation at 94 °C for 15 min; followed by 12 cycles with -1 °C per cycle; 94 °C for

30 s, 62–50 °C 90 s, 72 °C 60 s; then 20 cycles of 94 °C for 30 s, 50 °C for 90 s and 72 °C for 60 s. The final extension phase was at 60 °C for 5 min. Fluorescently labelled PCR fragments were analysed by electrophoresis in an ABI 3730 high-throughput capillary electrophoresis instrument (Applied Biosystems, Foster City, CA, USA). During electrophoresis, each reaction contained 10.3 μL Hi-Di Formamide (Applied Biosystems), 0.2 μL GeneScan 600 LIZ Size Standard (Applied Biosystems) and 1.2 μL 10× diluted PCR product. Alleles were scored using GeneMapper 3.7 (Applied Biosystems).

Without using prior population information, the genetic admixture of individuals and number of genetic clusters were evaluated using the Bayesian algorithm implemented in STRUCTURE 2.2 (Pritchard *et al.* 2000, 2002; Falush *et al.* 2003), using all 751 individuals. Number of genetic clusters was assayed using the DeltaK-method proposed by Evanno *et al.* (2005) as well as the method proposed in the original article by Pritchard *et al.* (2000). We tested for number of clusters (*K*) ranging from 1 through 9 (2 through 8 with the DeltaK-method), with ten replicates per *K*. All analyses were run using the admixture model and correlated allele frequencies. Burn-in period was set to 500 000 and Markov Chain Monte Carlo iterations set to 1 000 000. As preliminary runs with the original STRUCTURE model showed limited assignment power (Fig. S1, Supporting information), we analysed the dataset (for $K = 2$ and $K = 3$) with the LOCPRIOR model implemented in STRUCTURE 2.3 that allows for the use of sample group information in the clustering process (Hubisz *et al.* 2009). This model has been shown to detect structure at lower levels of divergence, or with less data, than previous STRUCTURE models. The model uses sample group information as a prior when the data indicate that this information will be helpful in the clustering process, e.g., in settings where individuals can be classified into discrete groups on the basis of a phenotypic characteristic. The LOCPRIOR model does not tend to find structure when none is present and it is able to ignore the sample group information when the ancestry of individuals does not correlate with sampling locations (Hubisz *et al.* 2009). Hence, the model-developers recommend using the LOCPRIOR model in most situations where the amount of available data is limited, especially when the standard STRUCTURE model does not provide a clear signal of structure.

For the LOCPRIOR STRUCTURE run, we divided the sparrow populations into six groups based on phenotype (Summers-Smith 1988; Töpfer 2006): (i) phenotypically pure house sparrows (Norway), (ii) individuals from the contact zone between house sparrows and Italian sparrows in the Alps, (iii) typical Italian sparrows (Italian peninsula excluding the Calabria peninsula), (iv) Southern Italian sparrows (Calabria peninsula from Eboli and southwards), (v) Sicilian Italian sparrows (Sicily) and (vi) phenotypically pure Spanish sparrows (Sardinia and Gargano peninsula). Males in category (iv) and especially (v) are phenotypically closer to Spanish sparrows than are typical Italian sparrows in having larger bibs and some black streaking on the flanks (Summers-Smith 1988; Töpfer 2006).

In addition to the main clusters used in the LOCPRIOR STRUCTURE run, males from the contact zones between the Italian sparrow and its putative parental species were scored phenotypically on an individual basis for separate analysis. Intermediates between house sparrows and Italian sparrows in the Alpine contact zone were defined according to percentage of grey feathers on the crown, with phenotypically pure house sparrows and Italian sparrow designated as 100 and 0, respectively. In the sympatric population of Italian and Spanish sparrows on Gargano peninsula, we classified males as typical Italian sparrow (category iii) and Spanish sparrow (category vi) based on the colour on the back, bib size and presence/absence of black streaking on the flanks. We also specifically looked for males of intermediate phenotype to be classified in the categories corresponding to those in southern Italy (category iv) or Sicily (category v).

Females of the three taxa cannot be reliably differentiated between based on phenotype and were therefore included in the genetic analyses based on geographical location and corresponding male phenotype. Females from the mixed population of Italian and Spanish sparrows on Gargano were hence placed in a separate group.

Mitochondrial DNA

To further test the hybrid origin hypotheses, we analysed variation in mitochondrial DNA. Seventy-five individuals (eight house sparrows, 41 Italian sparrows and 26 Spanish sparrows) were sequenced for the mitochondrial *ND2* gene using primers published in Sorenson *et al.* (1999) as modified in http://people.bu.edu/msoren/Bird.mt.Primers.pdf (see Table S2, Supporting information). We obtained high-quality sequence in all individuals across a continuous stretch of 594 bp's of the gene, which were then aligned and analysed. The sequences are available in GenBank (accession no's JN090513–JN090587).

Polymerase chain reactions were performed in 10 μL volumes containing 5.1 μL mqH_2O, 1 μL 10× optimized DyNAzyme buffer (Finnzymes Oy, Espoo, Finland) 0.5 μL of each primer (10 μM), 0.4 μL 2 mM dNTPs, 0.3 μL 100% DMSO, 0.2 μL 2 U/μL DyNAzyme II Polymerase (Finnzymes) and 2 μL template DNA (concentration ranging from 10 to 30 ng/μL). The following PCR program was used: Initial denaturation at 94 °C for 1 min, followed by 35 cycles of denaturation at 94 °C for 30 s and annealing at 56 °C for 30 s, followed by extension at 72 °C for 45 s. The final extension phase was at 72 °C for 7 min. PCR products were screened on a 2% agarose gel stained with ethidium bromide.

PCR products were cleaned for excessive primers and nucleotides using ExoSAP-IT (USB Corporation, Cleveland, OH, USA). The manufacturer's recommendations were followed except for the amount of ExoSAP-IT used per μL PCR product (ExoSAP-IT was diluted 10 times) and the incubation time for degradation of primers and nucleotides (37 °C for 45 min instead of 15 min). The sequencing was performed on an ABI 3730 high-throughput capillary electrophoresis instrument and was conducted using the same primers as used in the PCR. Sequences were edited by eye and aligned using SEQUENCHER 4.8 (Gene Codes Corp., Ann Arbor, MI, USA).

Phylogenetic analyses were conducted in MEGA 4 (Tamura *et al.* 2007). Tree sparrow (*Passer montanus*) was used as outgroup in the analysis (sequence obtained from GenBank, accession no: AY030144). Phylogenetic trees were reconstructed using both a distance method (neighbour-joining (Saitou & Nei 1987) (NJ)) and methods based on optimality criteria [maximum parsimony (Fitch 1971) (MP) and minimum evolution (Rzhetsky & Nei 1993) (ME)]. All positions containing missing data were eliminated from the dataset, and number of bootstraps was set to 2000. NJ, MP and ME gave identical topologies using both model based (Kimura 2-parameter, Kimura 1980; Tamura & Nei 1993) and p-distance calculations of the divergence). House sparrows used in the analysis of mtDNA were sampled in Norway (Oslo), Italian sparrows were sampled on the Italian peninsula (Accettura, Acquaviva-Picena, Assisi, Barletta, L'Aquila, Lecce, Lesina, Mondolfo, Scisciano and Terni) and on Sicily (Montemaggiore Belsito and Giardini-Naxos) and Spanish sparrows were sampled on Sardinia (Pula and San Priamo) and the Gargano peninsula (Lesina).

Results

Genetic analysis

The assignment analyses recognized only two genetic clusters (Fig. 2a, see Fig S2, Supporting information for analyses of $K = 3$), corresponding to the house sparrow and the Spanish sparrow (Fig. 2b). Italian sparrows showed clear admixture, i.e., were assigned partially to both clusters (Fig. 2b, 95% assignment probability intervals (PI) hardly overlap between any of the

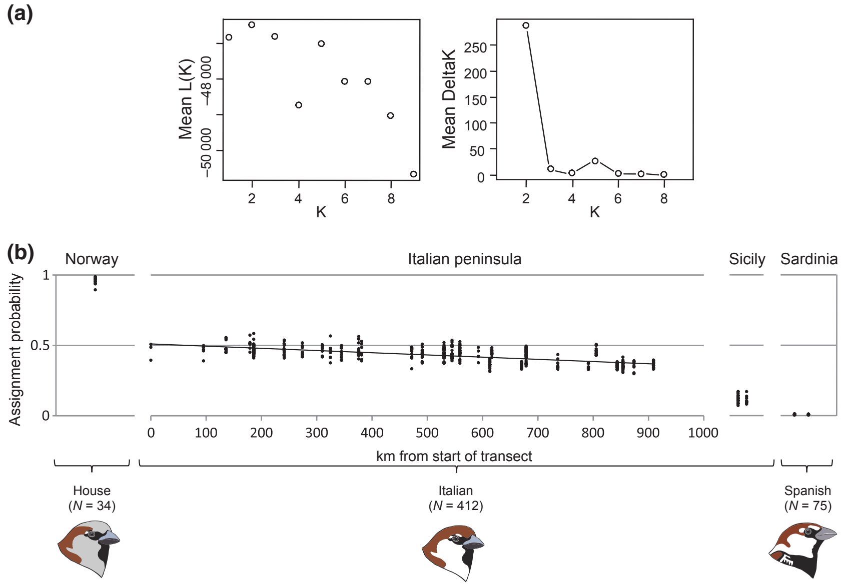

Fig. 2 Genetic structure of the studied *Passer* taxa. (a) Summary of runs for values of *K* ranging from 1 to 9 in STRUCTURE v. 2.2. Left panel shows mean log likelihood for each *K* plotted against *K*, right panel shows DeltaK plotted against *K*. Most likely number of clusters is two using both methods. (b) Each point represents an individual's probability of belonging to the house sparrow relative to the Spanish sparrow cluster. Results shown when assignment probabilities were calculated using the LOCPRIOR model (and admixture model with correlated allele frequencies). The trend line is from a linear regression of assignment probability on geographical distance from the northernmost population of phenotypically pure Italian sparrows. When indicated on the *x*-axis, numbers refer to distance in kilometres from the northernmost population of phenotypically pure Italian sparrows.

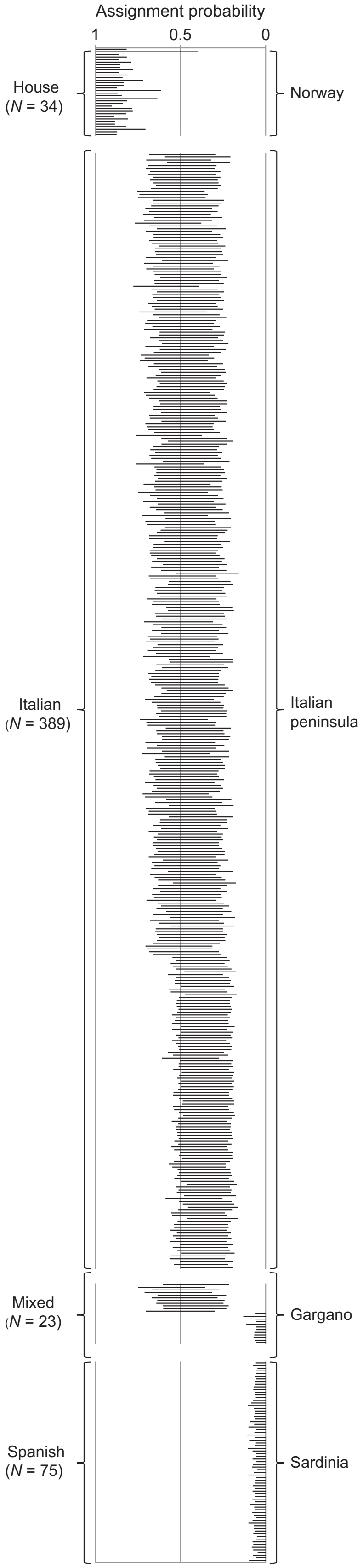

Fig. 3 95% probability intervals (PI) from STRUCTURE run with the LOCPRIOR model, admixture model and correlated allele frequencies for $K = 2$. Each line represents an individual's 95% PI of belonging to the house sparrow relative to the Spanish sparrow cluster.

individuals in the three taxa; Fig. 3, see also Table S3, Supporting information).

The assignment analyses further indicated clinal variation in nuclear genetic markers along the Italian peninsula; birds assigning on average somewhat closer to the Spanish sparrow towards the south (Fig. 2b; linear regression of assignment probability on geographical distance from the northernmost population of phenotypically pure Italian sparrows: $r = -0.64$, $N = 389$, $P < 0.001$). The sparrows on Sicily assigned much closer to Spanish sparrows than the rest of the Italian sparrows (t-test: Sicily: mean = 0.12, $N = 23$; mainland Italian sparrows: mean = 0.43, $N = 389$, $t = 25.64$, $P < 0.0001$, t-test: Sicily: mean = 0.12, $N = 23$; Spanish sparrows: mean = 0.01, $N = 75$, $t = 32.17$, $P < 0.0001$), but still showed clear evidence of admixture. Moreover, all the birds from Sicily that were analysed for mtDNA had house sparrow haplotypes.

Our phylogenetic analyses of the mtDNA data showed that the putative parental species differ in approximately four out of every 100 bases on average. According to a standard *ND2* molecular clock, the genetic distance between the House sparrow and Spanish sparrow translates into a time to the most recent common ancestor of about 2 million years (Johnson & Sorenson 1998; Norman *et al.* 2007; Price 2008). The data further indicated a recent origin of the Italian sparrow as it shares identical haplotypes with both parental species (although 40 of 41 haplotypes were of the house sparrow variant), and that it consequently does not form a haplogroup of its own (Fig. 4).

Reproductive barriers

Our data suggest reproductive barriers between the Italian sparrow and both its putative parental species where they now co-occur. No birds of intermediate phenotype were observed in a sympatric pocket of Spanish sparrows within the distribution of Italian sparrows on the Gargano peninsula (Fig. 5a). The birds on Gargano classified as Spanish and Italian sparrows, respectively, based on phenotype, had significantly different mean assignment probabilities that do not differ significantly from conspecifics sampled elsewhere (95% PI do not overlap for any of the individuals in the two different taxa; Fig. 3). These results also hold without using the LOCPRIOR model: The two taxa still have significantly different mean assignment probabilities (t-test, Italian;

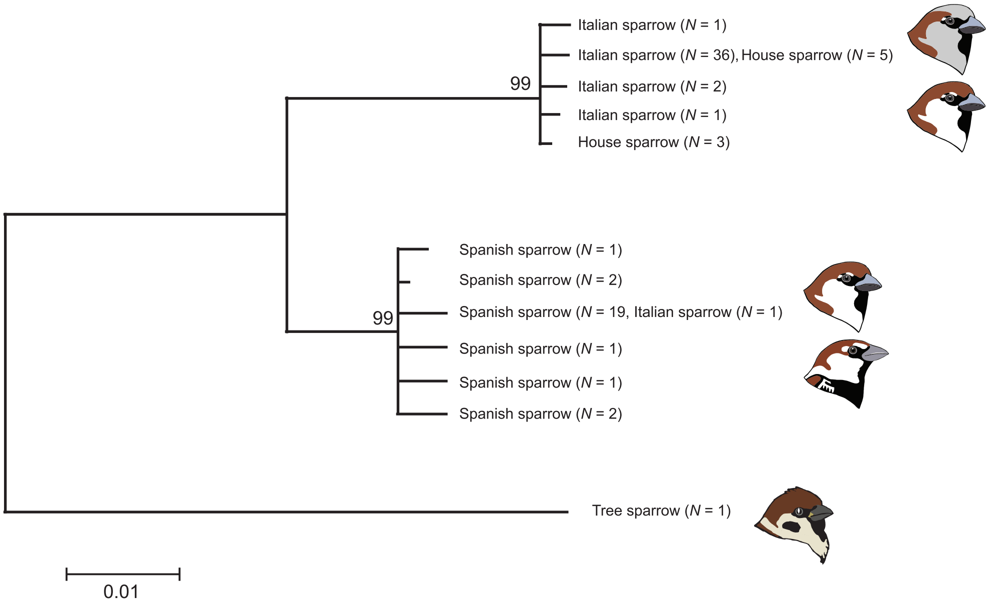

Fig. 4 Neighbour-joining tree (2000 bootstraps) of 594 bp's of the mtDNA gene *ND2* calculated using p-distances. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances are in the units of number of base substitutions per site. Each node represents a haplotype with number of individuals indicated. Tree sparrow *P. montanus* was used as outgroup in the analysis.

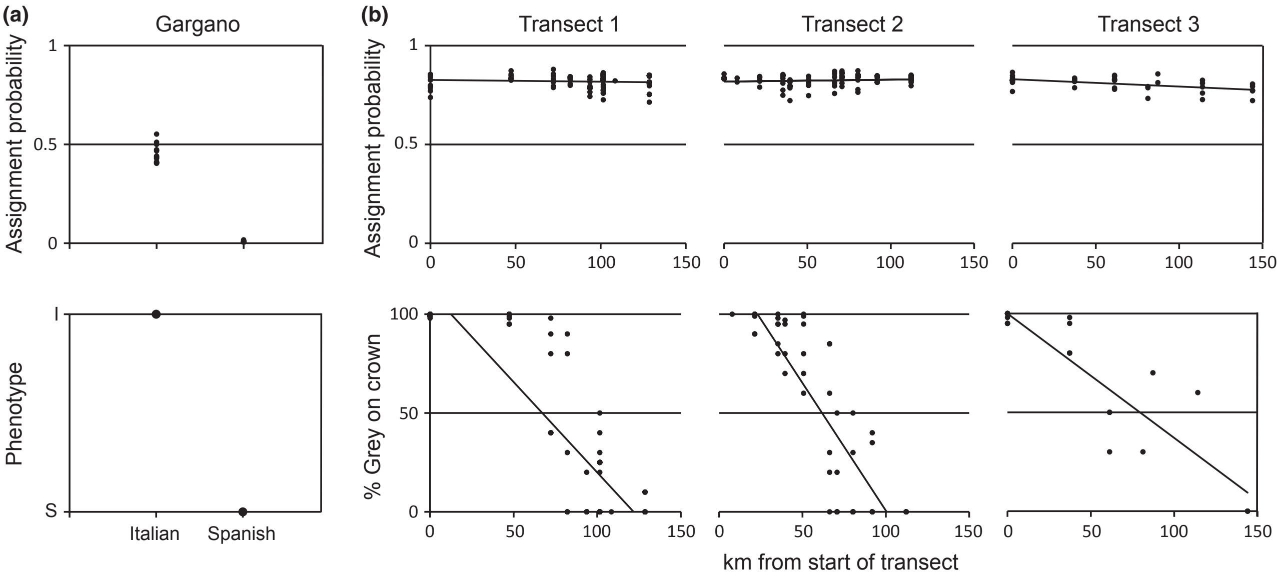

Fig. 5 Upper panels display results from STRUCTURE run with LOCPRIOR model on 751 individuals genotyped for 14 unlinked nuclear microsatellites. Assignment probabilities refer to an individual's probability of belonging to the house sparrow relative to the Spanish sparrow cluster. Lower panels show the phenotypic distribution of males at species-specific plumage traits. (a) Results for the mixed population of Italian and Spanish sparrows on the Gargano peninsula. No individuals with intermediate phenotypes were observed. (b) Results from three transects through the hybrid zone between the Italian sparrow and the house sparrow in the Alps. The trend lines in the upper panel refer to linear regressions of assignment probability on geographical distance from the first population on the house sparrow-end of the hybrid zone. In the lower panel of each transect, a point refers to the amount of grey feathers on the crown of male individuals in each sampling location and trend lines from linear regressions are fitted. The *x*-axis refers to number of kilometres from the first sampling location at the house sparrow-end of the hybrid zone.

$N = 11$, mean = 0.47; Spanish: $N = 12$, mean = 0.18, $t = 4.82$, $P < 0.0001$) that do not differ significantly from those of conspecifics sampled elsewhere (Italian sparrows: *t*-test, Gargano: $N = 11$, mean = 0.47, mainland Italian sparrows: $N = 389$, mean = 0.45, $t = 0.43$, $P = 0.67$; Spanish sparrows: *t*-test, Gargano: $N = 12$, mean = 0.18, Sardinia: $N = 75$, mean = 0.20, $t = 1.05$, $P = 0.30$). Further, of the birds from Gargano analysed for mtDNA, all phenotypic Italian sparrows had house sparrow haplotypes, and all phenotypic Spanish sparrows had Spanish sparrow haplotypes.

The Italian sparrow and the house sparrow do, however, hybridize where they come into contact in a narrow, stable and sparsely populated hybrid zone in the Alps (Summers-Smith 1988) (Fig. 1). Prezygotic barriers between these two taxa are therefore not absolute. Mean assignment probabilities in the Alpine contact zone were intermediate to those of pure house sparrows (*t*-test: contact zone: mean = 0.82, $N = 203$; house sparrows: mean = 0.97, $N = 34$, $t = 26.03$, $P < 0.0001$) and Italian sparrows (*t*-test: contact zone: mean = 0.82, $N = 203$; Italian sparrows: mean = 0.43, $N = 389$, $t = 88.73$, $P < 0.0001$). However, we did not find evidence of any prominent genetic cline over this hybrid zone (Fig. 5b; linear regressions of assignment probability on geographical distance from the first population on the house sparrow-end of the hybrid zone; Transect 1: $r = -0.10$, $N = 88$, $P = 0.36$, Transect 2: $r = 0.09$, $N = 77$, $P = 0.41$, Transect 3: $r = -0.53$, $N = 38$, $P < 0.001$). Yet, the hybrid zone is characterized by a steep cline in species-specific male plumage traits (Fig. 5b; linear regressions: Transect 1: $r = -0.86$, $N = 61$, $P < 0.001$, Transect 2: $r = -0.84$, $N = 49$, $P < 0.001$, Transect 3: $r = -0.89$, $N = 19$, $P < 0.001$).

Discussion

Hybrid origin of the Italian sparrow

The assignment analyses based on the nuclear microsatellites yielded two groups corresponding to the house sparrow, and the Spanish sparrow with the Italian sparrow showing admixture between them. Further, the results from the analyses of the microsatellite markers are in concordance with the phenotypic data: Italian sparrows get more similar to Spanish sparrows towards the south-western end of their distribution, both genetically and morphologically, especially so on Sicily. Finally, the results of the phylogenetic analyses of the mtDNA show a deep split between the house sparrow and the Spanish sparrow similar to other congeneric bird species (Price & Bouvier 2002; Price 2008), and that the Italian sparrow shares identical haplotypes with both species (98% of the house sparrow type) and thus does not form a clade of its own.

As the Italian sparrow does not form a cluster/clade of its own using neither microsatellites nor mtDNA, and neither forms a monophyletic clade with either of the putative parental species, the obtained results seem to contradict a scenario of bifurcating speciation. Although only one (from Barletta south of the Gargano peninsula) of the sequenced Italian sparrows possessed Spanish sparrow mtDNA, the Italian sparrow is fixed for the Spanish sparrow variant at one Z-linked marker (CHD1Z) whereas both parental variants segregate at another Z-linked marker (PLAA) (Elgvin *et al.* 2011). This demonstrates that the Italian sparrow does not form a monophyletic clade with the house sparrow as could be suggested by the mtDNA data alone. Importantly, the haplotypes possessed by the Italian sparrows do not form separate haplogroups at any of the markers analysed here or in the study by Elgvin *et al.* (2011), but consistently clusters with one or the other, or both, of the putative parental species. This further suggests that the Italian sparrow is of recent origin, as it has not diverged from its putative parental species.

On the other hand, all the data are consistent with a hybrid origin of the Italian sparrow. As expected under this hypothesis, we found (i) greater divergence between the putative parental species than between the Italian sparrow and any of them, (ii) evidence for a mixed nuclear genome both in the microsatellite data presented here and the nuclear sequences presented by Elgvin *et al.* (2011) and (iii) no unique haplogroups in the Italian sparrow at any of the studied loci. Such genealogical discordance and intermediate allele frequencies compared with the putative parental species have been found in other suggested homoploid hybrid taxa (Mavárez & Linares 2008). Further, (iv) the Italian sparrow has an intermediate morphology which has been reproduced in laboratory *F*1-crosses between house sparrows and Spanish sparrows (Macke 1965; Alonso 1984). Although an intermediate plumage is only circumstantial evidence of hybrid origin, such a pattern would not necessarily follow from any bifurcation scenario. Moreover, (v) the Italian sparrow has a restricted geographic distribution relative to those of its putative parental species. Such a pattern has also been found in most other putative examples of HHS and is consistent with hybridization as hybridization usually occurs on a local scale (Mavárez & Linares 2008). Taken together, the most parsimonious explanation of the genetic results and the more indirect evidence from morphology and distribution patterns seem to be that the Italian sparrow really is of hybrid origin.

Reproductive barriers against parental species

The major obstacle facing homoploid hybrid speciation is to produce a lineage that can escape both gene flow and competition from its parental species (e.g., Rieseberg 1997; Buerkle *et al.* 2000; Coyne & Orr 2004; Schwarz *et al.* 2005; Mallet 2007). Ecologically, the Italian sparrow is similar to the house sparrow, i.e., closely associated with human settlements (agricultural and urban environments) whereas the Spanish sparrow is affiliated with more mesic habitat, often breeding in natural or semi-natural habitat, e.g., along riverbeds and in wetlands (Summers-Smith 1988). Considering this close ecological similarity with especially one of the parental species, what has prevented the Italian sparrow from getting swamped by gene flow from, or outcompeted by, one or both of them? Although ecological divergence from parental forms have been shown to be of importance in the formation of some hybrid taxa (Gross & Rieseberg 2005), this is not a general pattern and behavioural or genetic factors have been invoked to explain reproductive isolation in other proposed hybrid taxa (Mavárez & Linares 2008).

Our data indicate current reproductive barriers between the Italian sparrow and both its putative parental species. We found no signs of interbreeding between the Italian sparrow and the Spanish sparrow, i.e., we observed no individuals of intermediate phenotype, in an area of distributional overlap on the Gargano peninsula, and all individuals sampled here were genetically similar to pure Italian and Spanish sparrows sampled elsewhere. In birds, mate choice and species recognition are usually strongly affected by sexual imprinting on parental phenotypes early in life (ten Cate & Vos 1999; Hansen *et al.* 2008; Price 2008) and such early learning can facilitate rapid development of premating isolation, because no genetic changes are required (Servedio *et al.* 2009). However, as the observed pattern might also result from differences in breeding habitats or breeding phenology between the two taxa or from hybrid inviability, it remains to be tested whether sexual imprinting explains the observed pattern between the Italian and Spanish sparrows.

In contrast, the Italian sparrow does hybridize with the house sparrow where they meet in the Alps. Hence, the Italian sparrow is evidently not fully reproductively isolated from its other putative parental species. The contact zone is characterized by a steep cline in species-specific plumage traits, but we found only weak evidence of any coinciding genetic cline using the neutral microsatellite markers. Although we only found weak evidence of a genetic cline in the contact zone, the mean assignment probabilities of the birds sampled in that area were intermediate relative to phenotypically pure house sparrows and Italian sparrows. This indicates that the birds in the contact zone are genetically intermediate between pure house sparrows and Italian sparrows, and that the lack of a cline may be because of the relatively low power of the applied markers and of all individuals being lumped in one group. Selection against intermediate genotypes and/or positive assortative mating is predicted to yield a sharp transition zone as observed for the phenotypic trait (Barton & Hewitt 1985). On the other hand, the width of a contact zone depends on a number of factors, including dispersal distances, population densities and topography. The Italian sparrow/house sparrow contact zone occurs at high altitudes in the Alps where population densities are low, factors that are likely to reduce gene flow across the Alps. Further studies are therefore needed to investigate the relative importance of reproductive and geographic barriers in reducing gene flow between the Italian sparrow and the house sparrow.

A scenario for the origin of the Italian sparrow

It has been debated whether reproductive isolation must arise as a direct consequence of hybridization for a lineage of mixed ancestry to qualify as a homoploid hybrid species (e.g., Mallet 2007; Jiggins *et al.* 2008). However, as Mavárez & Linares (2008) make explicit in their review, theory has shown that reproductive isolation might evolve after the actual hybridization event(s) when the evolution of the homoploid hybrid lineage occurs in geographical isolation (Buerkle *et al.* 2000; Duenez-Guzman *et al.* 2009), and that such geographical isolation might generally facilitate HHS. In concordance with this, the Italian sparrow is today sheltered from gene flow from its putative parental species through its largely allopatric distribution on the Italian Peninsula and adjacent islands as it is surrounded by the massive geographical barriers of the Alps and the Mediterranean Sea. During the last approximately 10 000 years, *Passer* sparrows have taken advantage of artificial habitat and have expanded greatly with the spread of agriculture and human civilizations (Summers-Smith 1988; Ericson *et al.* 1997, G.P. Sætre, S. Riyahi, M. Aliabadian, T.O. Elgvin, J.S. Hermansen, U. Olsson and SA Sæther unpublished data). Moreover, agriculture developed earlier in Italy than on the European continent (Zeder 2008) and this probably facilitated hybridization in this system as house sparrows and Spanish sparrows do not interbreed when both co-occur in large numbers, but may interbreed extensively in areas where one or both species are rare (e.g., in areas in North Africa and Cape Verdes) (Summers-Smith 1988), presumably because of limited availability of conspecific mates. Hence, interbreeding between once rare house sparrows and Spanish sparrows in Italy would have led to the formation of a hybrid swarm that would have increased in numbers with the expansion of agriculture. When agriculture and numbers of sparrows eventually increased also elsewhere in Europe, extensive secondary gene flow into the Italian sparrow would have been hindered by the amount of birds already present on the Italian peninsula, the barriers of the Alps and the Mediterranean Sea and by reproductive barriers having developed, and thus allowed the Italian sparrow to be maintained as an independent lineage. Geographic isolation seems to have played a similar role in the origin of another suggested hybrid bird, the Audubon's warbler, which may have been isolated from its parental lineages in a glacial refugium (Brelsford *et al.* 2011).

Conclusions

We suggest that an origin of homoploid hybrid taxa as proposed for the Italian sparrow, where the hybrid lineage gets geographically isolated from its parental species, might be more common in nature than previously assumed. First, there is often ample time for hybridization to occur between diverging lineages from the point that reproductive barriers starts to develop until they are complete (Mallet 2005, 2008). Second, geographic isolation inevitably leads to divergence over time (Coyne & Orr 2004), also in traits that may promote reproductive isolation. Accordingly, if a hybrid lineage becomes geographically isolated from its parental species, allopatric divergence will increase the likelihood of speciation (Buerkle *et al.* 2000; Duenez-Guzman *et al.* 2009). Together this implies that homoploid hybrid taxa may originate relatively frequently, but also that the traces—-both molecular and morphological—of such an origin may eventually disappear. It is therefore possible that many animals have a hybrid origin now eroded by time.

Acknowledgments

We thank H Jensen for access to unpublished primer sequences, A Andreotti, J Cecere, D Fulgione, A Herland, S Hogner, GA Lien, E Mallia, V Maselli, M Sacchi, R Tinarelli, WWF Italia and Istituto Nazionale per la Fauna Selvatica for help during fieldwork, NW Steen for technical assistance in the lab, KKH Clausen for graphical assistance, and AØ Bollvåg, A Brelsford, F Haas, KL Voje and anonymous referees for helpful comments on previous drafts. The study was financially supported by the Centre for Ecological and Evolutionary Synthesis (CEES) and Molecular Life Sciences (MLS) at the University of Oslo, and the Research Council of Norway.

References

Alonso JC (1984) Kreuzung spanischer haus- (*Passer domesticus*) und weidensperlinge (*Passer hispaniolensis*) in gefangenschaft. *Journal für Ornithologie*, **125**, 339–340.

Anderson E (1949) *Introgressive Hybridization*. John Wiley & Sons, New York, NY.

Arnold ML (1997) *Natural Hybridization and Evolution*. Oxford University Press, Oxford.

Arnold ML (2006) *Evolution Through Genetic Exchange*. Oxford University Press, New York.

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

Brelsford A, Milá B, Irwin DE (2011) Hybrid origin of Audubon's warbler. *Molecular Ecology*, **20**, 2380–2389.

Buerkle CA, Morris RJ, Asmussen MA, Rieseberg LH (2000) The likelihood of homoploid hybrid speciation. *Heredity*, **84**, 441–451.

Bulatova NS, Panov EN, Radjabli SI (1972) Karyological description of three species of the genus *Passer*. *Experientia*, **28**, 1369–1371.

ten Cate C, Vos DR (1999) Sexual imprinting and evolutionary processes in birds: a reassessment. *Advances in the Study of Behavior*, **28**, 1–31.

Coyne JA, Orr HA (2004) *Speciation*. Sinauer Associates, Sunderland, MA.

Dobzhansky T (1937) *Genetics and the Origin of Species*. Columbia University Press, New York, NY.

Duenez-Guzman EA, Mavarez J, Vose MD, Gavrilets S (2009) Case studies and mathematical models of ecological speciation. 4. Hybrid speciation in butterflies in a jungle. *Evolution*, **26**, 2611–2626.

Elgvin TO, Hermansen JS, Fijarczyk A *et al.* (2011) Hybrid speciation in sparrows II: a role for sex chromosomes? *Molecular Ecology*, doi: 10.1111/j.1365-294X.2011.05182.x

Ericson PGP, Tyrberg T, Kjellberg AS, Jonsson L, Ullén I (1997) The earliest record of house sparrows (*Passer domesticus*) in northern europe. *Journal of Archaeological Science*, **24**, 183–190.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Fisher RA (1930) *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford.

Fitch WM (1971) Toward defining the course of evolution: minimum change for a specific tree topology. *Systematic Zoology*, **20**, 406–416.

Fulgione D, Aprea G, Milone M, Odierna G (2000) Chromosomes and heterochromatin in the Italian sparrow, *Passer italiae*, a taxon of presumed hybrid origins. *Folia Zoologica*, **49**, 199–204.

Grant PR, Grant BR (1992) Hybridization of bird species. *Science*, **256**, 193–197.

Gross BL, Rieseberg LH (2005) The ecological genetics of homoploid hybrid speciation. *Journal of Heredity*, **96**, 241–251.

Hansen BT, Johannessen LE, Slagsvold T (2008) Imprinted species recognition lasts for life in free-living great tits and blue tits. *Animal Behaviour*, **75**, 921–927.

Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, **9**, 1322–1332.

Jiggins CD, Salazar C, Linares M, Mavarez J (2008) Hybrid trait speciation and *Heliconius* butterflies. *Philosophical Transactions of the Royal Society Series B*, **363**, 3047–3054.

Johnson KP, Sorenson MD (1998) Comparing molecular evolution in two mitochondrial protein coding genes (Cytochrome *b* and ND2) in the dabbling ducks (Tribe: Anatini). *Molecular Phylogenetics and Evolution*, **10**, 82–94.

Johnston RF (1969) Taxonomy of house sparrows and their allies in the Mediterranean basin. *Condor*, **71**, 129–139.

Kimura M (1980) A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, **16**, 111–120.

Linnaeus C (1760) Disquisitio de sexu plantarum, ab Academia Imperiali Scientiarum Petropolitana praemio ornata. *Amoenitates Academicae*, **10**, 100–131.

Macke T (1965) Gelungene Kreuzung von *Passer domesticus* und *hispaniolensis*. *Journal für Ornithologie*, **10**, 461–462.

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

Mallet J (2007) Hybrid speciation. *Nature*, **446**, 279–283.

Mallet J (2008) Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society Series B*, **363**, 2971–2986.

Mavárez J, Linares M (2008) Homoploid hybrid speciation in animals. *Molecular Ecology*, **17**, 4181–4185.

Mayr E (1942) *Systematics and the Origin of Species*. Columbia University Press, New York.

Meise W (1936) Zur Systematik und Verbreitungsgeschichte der Haus- und Weidensperlinge, *Passer domesticus* (L.) und *hispaniolensis* (T.). *Journal für Ornithologie*, **84**, 631–672.

Nolte AW, Tautz D (2010) Understanding the onset of hybrid speciation. *Trends in Genetetics*, **26**, 54–58.

Norman JA, Rheindt FE, Rowe DL, Christidis L (2007) Speciation dynamics in the Australo-Papuan Meliphaga honeyeaters. *Molecular Phylogenetics and Evolution*, **42**, 80–91.

Price T (2008) *Speciation in Birds*. Ben Roberts, Greenwood, CO.

Price T, Bouvier M (2002) The evolution of F-1 postzygotic incompatibilities in birds. *Evolution*, **56**, 2083–2089.

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

Pritchard JK, Falush D, Stephens M (2002) Inference of population structure in recently admixed populations. *American Journal of Human Genetetics*, **71**, 177–177.

Rieseberg LH (1997) Hybrid origins of plant species. *Annual Review in Ecology and Systematics*, **28**, 359–389.

Rzhetsky A, Nei M (1993) Theoretical foundation of the Minimum-Evolution method of phylogenetic inference. *Molecular Biology and Evolution*, **10**, 1073–1095.

Saitou N, Nei M (1987) The neighbour-joining method - a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 406–425.

Schrey AW, Grispo M, Awad M *et al.* (2011) Broad-scale latitudinal patterns of genetic diversity among native European and introduced house sparrow (*Passer domesticus*) populations. *Molecular Ecology*, **20**, 1133–1143.

Schwarz D, Matta BM, Shakir-Botteri NL, McPheron BA (2005) Host shift to an invasive plant triggers rapid animal hybrid speciation. *Nature*, **436**, 546–549.

Seehausen O (2004) Hybridization and adaptive radiation. *Trends in Ecology and Evolution*, **19**, 198–207.

Servedio MR, Sæther SA, Sætre G-P (2009) Reinforcement and learning. *Evolutionary Ecology*, **23**, 109–123.

Sorenson MD, Ast JC, Dimcheff DE, Yuri T, Mindell DP (1999) Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. *Molecular Phylogenetics and Evolution*, **12**, 105–114.

Stebbins GL (1959) The role of hybridization in evolution. *Proceedings of the American Philosophical Society*, **103**, 231–251.

Summers-Smith JD (1988) *The Sparrows: A Study of the Genus Passer*. T & D Poyser, Calton.

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial-DNA in humans and chimpanzees. *Molecular Biology and Evolution*, **10**, 512–526.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–1599.

Töpfer T (2006) The taxonomic status of the Italian sparrow *Passer italiae* (Vieillot 1817): speciation by stabilised hybridisation? A critical analysis. *Zootaxa*, **1325**, 117–145.

Zeder MA (2008) Domestication and early agriculture in the Mediterranean basin: origins, diffusion, and impact. *Proceedings of the National Academy of Sciences of the USA*, **105**, 11597–11604.

Zedlitz OG (1913) Ornithologische Ergebnisse der Reise von Paul Spatz in die algerische Sahara im Sommer 1912. *Novitates Zoologicae*, **20**, 164–186.

The authors share a common interest in evolutionary biology with a special focus on the role of hybridization in evolution. Topics of interest include hybrid speciation, reinforcement, sexual selection and phylogeography.

Data accessibility

mtDNA sequences: Genbank accessions JN090513–JN090587.

Sample locations and microsatellite data: Data deposited at Dryad: doi:10.5061/dryad.k7vh9.

Supporting information

Additional supporting information may be found in the online version of this article.

Table S1 Information on the microsatellite loci used in the study.

Table S2 Information on the mtDNA primer pair used in the study.

Table S3 Pairwise F_{ST}-estimates between phenotypic classes.

Fig. S1 95% probability intervals from STRUCTURE run with admixture model and correlated allele frequencies for $K = 2$ (with and without LOCPRIOR model).

Fig. S2 Comparisons of STRUCTURE runs for $K = 2$ and $K = 3$.