Molecular Phylogenetics and Evolution 120 (2018) 183–195

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Discordance between genomic divergence and phenotypic variation in a rapidly evolving avian genus (*Motacilla*)

Rebecca B. Harris[a,b,*], Per Alström[c,d,e], Anders Ödeen[c], Adam D. Leaché[a,b]

[a] Department of Biology, University of Washington, Seattle, WA 98195, United States
[b] Burke Museum of Natural History and Culture, Seattle, WA 98195, United States
[c] Department of Ecology and Genetics, Animal Ecology, Evolutionary Biology Centre, Uppsala University, Norbyvägen 18D, Uppsala SE-752 36, Sweden
[d] Swedish Species Information Centre, Swedish University of Agricultural Sciences, Box 7007, Uppsala SE-750 07, Sweden
[e] Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China

ARTICLE INFO

We dedicate this work to Anders Ödeen, a dear friend, colleague, and pioneer in wagtail genetics, who passed away during the preparation of this manuscript.

Keywords:
Phylogenetics
Genetic differentiation
Plumage divergence
Mito-nuclear discordance

ABSTRACT

Generally, genotypes and phenotypes are expected to be spatially congruent; however, in widespread species complexes with few barriers to dispersal, multiple contact zones, and limited reproductive isolation, discordance between phenotypes and phylogeographic groups is more probable. Wagtails (*Motacilla*) are a genus of birds with striking plumage pattern variation across the Old World. Up to 13 subspecies are recognized within a single species, yet previous studies using mitochondrial DNA have supported polyphyletic phylogeographic groups that are inconsistent with subspecies plumage characteristics. In this study, we investigate the link between phenotypes and genotype by taking a phylogenetic approach. We use genome-wide SNPs, nuclear introns, and mitochondrial DNA to estimate population structure, isolation by distance, and species relationships. Together, our genetic sampling includes complete species-level sampling and comprehensive coverage of the three most phenotypically diverse Palearctic species. Our study provides strong evidence for species-level patterns of differentiation, however population-level differentiation is less pronounced. SNPs provide a robust estimate of species-level relationships, which are mostly corroborated by a combined analysis of mtDNA and nuclear introns (the first time-calibrated species tree for the genus). However, the mtDNA tree is strongly incongruent and is considered to misrepresent the species phylogeny. The extant wagtail lineages originated during the Pliocene and the Eurasian lineage underwent rapid diversification during the Pleistocene. Three of four widespread Eurasian species exhibit an east-west divide that contradicts both subspecies taxonomy and phenotypic variation. Indeed, SNPs fail to distinguish between phenotypically distinct subspecies within the *M. alba* and *M. flava* complexes, and instead support geographical regions, each of which is home to two or more different looking subspecies. This is a major step towards our understanding of wagtail phylogeny compared to previous analyses of fewer species and considerably less sequence data.

1. Introduction

In birds, male plumage differences among closely related taxa are often believed to be the result of sexual selection, and to play an important role in reproductive isolation (Price, 2008). Plumage differences can evolve rapidly (Olsson et al., 2010; Omland and Lanyon, 2000; Milá et al., 2007), and when populations are geographically structured, may result from spatial variation in selection regimes (Price, 2008). Recent studies have demonstrated that a small number of genes can cause dramatic plumage differences despite limited genetic differentiation throughout the remainder of the genome (Poelstra et al., 2015; Toews et al., 2016; Vijay et al., 2016; Mason and Taylor, 2015). Lack of overall genetic differentiation in taxa with distinct phenotypic differences is likely due to either (1) recent divergence, with strong selection on phenotype, or (2) large-scale introgression, except on pre-existing adaptive genetic differences. In such cases, it is unlikely that phylogenetic relationships gleaned from few loci accurately reflect true species trajectories.

Genera that contain widespread species complexes are useful systems for investigating geographic variation in phenotypes because they offer comparisons between populations and species at different stages of the speciation continuum. Species complexes are often characterized by high frequencies of hybridization and poorly developed isolation barriers, despite being structured geographically (Price, 2008).

* Corresponding author at: Department of Biology, University of Washington, Seattle, WA 98195, United States.
E-mail address: rbharris@uw.edu (R.B. Harris).

https://doi.org/10.1016/j.ympev.2017.11.020
Received 10 August 2017; Received in revised form 10 November 2017; Accepted 29 November 2017
Available online 12 December 2017
1055-7903/

Traditionally, it was thought that sympatric species would respond similarly to environmental factors influencing divergence, and therefore the study of sympatric complexes might reveal important biogeographic barriers. However, a growing body of literature suggests that species specific differences have a direct effect on demography, and spatially concordant genetic breaks should not be the expectation (Zamudio et al., 2016). For example, sexually selected traits, such as plumage, can directly affect genetic diversity via assortative mating or species recognition (Price, 1998). The complex interactions of gene flow, drift, and selection play a large role in determining the outcome of speciation and the rate at which species move along the speciation continuum. While the placement of recently diverged and introgressed lineages may be difficult, these can be contrasted with older, reproductively isolated groups within the same genus.

One bird system that is particularly well suited for such studies is the passerine genus *Motacilla* in the family Motacillidae. *Motacilla* consists of 12 species distributed throughout the Old World (Alström and Mild, 2003; del Hoyo et al., 2004) that have earned the common name wagtails due to their propensity to pump their long tails up and down. Within wagtails, there are multiple examples of taxa at different stages in the speciation process: from barely differentiated parapatric populations, to subspecies/species with distinct plumages that meet in hybrid zones, to fully reproductively isolated species (Alström and Mild, 2003). Previous phylogenetic and phylogeographic studies of *Motacilla* report mitochondrial relationships incongruent with both taxonomy (Pavlova et al., 2005; Voelker, 2002; Pavlova et al., 2003; Li et al., 2016; Alström and Ödeen, 2002; Ödeen and Björklund, 2003; Ödeen & Alström, 2001) and nuclear relationships (Alström and Ödeen, 2002; Ödeen and Björklund, 2003; Ödeen & Alström, 2001), with suggestions that mitochondrial DNA (mtDNA) poorly reflects the true phylogeny (Alström and Mild, 2003; Alström and Ödeen, 2002; Ödeen and Björklund, 2003; Ödeen & Alström, 2001). Several of these studies focused on aspects of wagtails' plumage diversity, some proposing cases of remarkable parallel plumage evolution (Alström and Mild, 2003; Alström and Ödeen, 2002; Ödeen & Alström, 2001) and others implicating the role of selection in rapid plumage evolution (Pavlova et al., 2005).

Of particular interest have been the four sympatric, migratory wagtail species White Wagtail *M. alba*, Grey Wagtail *M. cinerea*, Citrine Wagtail *M. citreola*, and Yellow Wagtail *M. flava* which are widely distributed across the Palearctic during the breeding season. These species represent a striking contrast in spatial variation in male breeding plumage (cf. Fig. 1). Currently, subspecies are defined by differences in both color and pattern of head plumage in the *M. flava* complex (13 subspecies) and by head, back, and wing-covert plumage in *M. alba* (9 subspecies) (Alström and Mild, 2003). On the basis of both genetic and plumage data, many of these subspecies have been treated as separate species (reviewed in Alström and Mild, 2003). Plumage differences are thought to have evolved rapidly and in conflict with phylogeographic structure (Pavlova et al., 2005; Pavlova et al., 2003; Li et al., 2016; Ödeen and Björklund, 2003; Ödeen & Alström, 2001). In contrast, the other two Palearctic breeding species, *M. cinerea* and *M. citreola*, lack this extreme plumage variation (3 subtly different and 2 distinct subspecies, respectively (Alström and Mild, 2003).

Wagtails can be broadly categorized by breeding distribution (i.e. Palearctic, Afrotropical). Whereas only some of the Palearctic species are migratory, all of the Afrotropical species are resident (Alström and Mild, 2003; del Hoyo et al., 2004). Species with Palearctic breeding distributions can be further categorized by plumage color (i.e. "black-and-white" and yellow). Past phylogenetic reconstructions have not fully supported these groupings. *M. cinerea*, *M. citreola*, and the *M. flava* complex all have yellow plumage, but the latter two species have repeatedly been found to be polyphyletic (Voelker, 2002; Pavlova et al., 2003; Alström and Ödeen, 2002; Ödeen & Alström, 2001). Genetic data places the polytypic *M. alba* within the "black-and-white" plumage group, along with three monotypic species with rather restricted allopatric distributions in the Indian subcontinent (White-browed Wagtail *M. maderaspatensis*), Cambodia (Mekong Wagtail *M. samveasnae*), and Japan (Japanese Wagtail *M. grandis*) (Alström and Mild, 2003; Alström and Ödeen, 2002). The black-and-white Afrotropical African Pied Wagtail *M. aguimp* has also been placed within this group (Alström and Ödeen, 2002; Ödeen & Alström, 2001). The other two Afrotropical (Cape Wagtail *M. capensis* and Mountain Wagtail *M. clara*) and the Malagasy (Madagascar Wagtail *M. flaviventris*) species are closely related (Voelker, 2002; Alström and Ödeen, 2002; Alström et al., 2015) and have slight or no geographical variation in plumage (del Hoyo et al., 2004). A recent phylogenetic exploration of the family found the São Tomé endemic São Tomé Shorttail *Amaurocichla bocagii* nested within *Motacilla*, and proposed its inclusion within this genus (Alström et al., 2015). Overall, relationships among species are unclear and are in need of reexamination.

In this study, we utilize genome-wide SNPs, nuclear introns, and mtDNA to analyze phylogenetic relationships and divergence patterns in *Motacilla*, with complete species-level sampling and comprehensive coverage of the three most diverse Palearctic species. We (1) estimate the first complete time-calibrated species tree for this group; (2) use genome-wide SNPs to reconstruct the phylogeny and investigate the agreement between genotype and phenotype in the three most variable wagtail species; and (3) demonstrate conclusively that mtDNA alone is inappropriate for phylogenetic studies of *Motacilla*.

2. Materials and methods

2.1. Sanger data

Throughout the manuscript, we follow the taxonomy of Alström and Mild (2003) and Alström et al. (2015).

To resolve species-level relationships and infer divergence times, we utilize previously published and unpublished sequences from (1) three nuclear introns (CHD1Z, ODC, Mb) for 42 individuals across all 12 *Motacilla* species (Alström and Ödeen, 2002; Ödeen and Björklund, 2003), and (2) two mitochondrial regions (ND2, CR) for 103 individuals across all species, including all subspecies of *M. alba*, *M. flava*, and *M. citreola* (Table S1) (Pavlova et al., 2005; Voelker, 2002; Li et al., 2016; Ödeen and Björklund, 2003). *Dendronanthus indicus*; *Anthus pratensis*, and *Anthus trivialis* were used as outgroups (Alström et al., 2015).

2.2. ddRADseq data

2.2.1. Sampling

If wagtail divergence was recent or shaped by rapid ancestral radiations, then the timing between divergence events may have been too short for the emergence of phylogenetically informative mutations (Gohli et al., 2015; Rokas and Carroll, 2006), potentially leading to the mito-nuclear discordance shown in previous studies. We therefore enhanced our inferential power by collecting thousands of genome-wide SNPs.

We obtained extensive geographic sampling and near-complete taxonomic coverage across Eurasia from samples at the Burke Museum of Natural History and Culture. We augmented these specimens with samples from other natural history museums to provide complete sampling for the genus (Fig. S1, Table S1). A total of 246 birds were sampled from 11 of the 12 recognized *Motacilla* species (Alström and Mild, 2003; Alström et al., 2015) (*M. maderaspatensis* not included due to lack of tissue samples). Our sampling focused on the widespread migratory Eurasian wagtail species (*M. alba, M. flava, M. citreola, M. cinerea*) with multiple described subspecies. We examined museum skins and assigned individuals to subspecies using morphological criteria outlined in Alström and Mild (2003). For rooting phylogenetic trees, we sampled three individuals from the monotypic sister genus (*Dendronanthus* Alström et al., 2015).

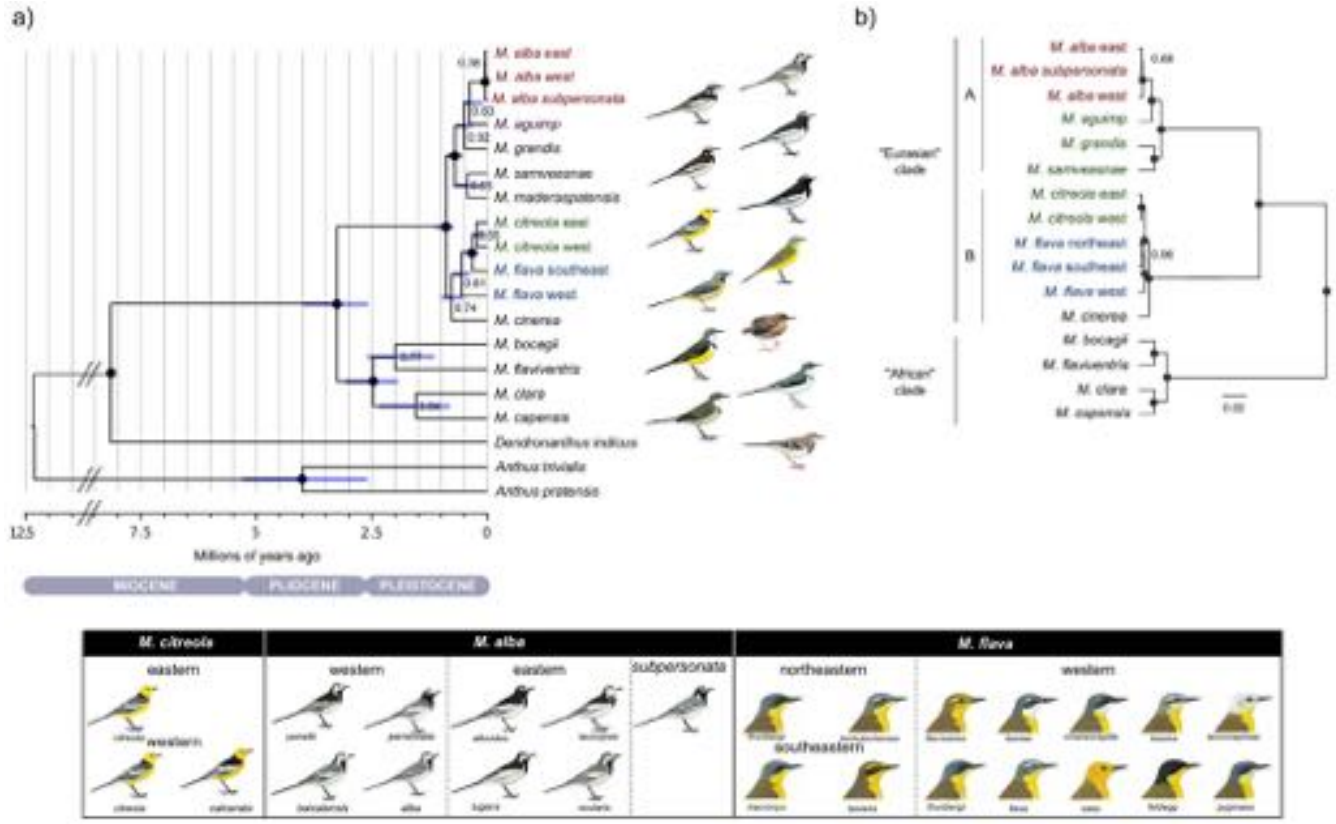

Fig. 1. Species trees. (a) *BEAST tree inferred from mtDNA and nuclear introns, calibrated using a 2.7% ND2 rate. Node bars indicate the 95% HPD of height (not shown for outgroups). (b) SNAPP maximum clade credibility species tree (1467 biallelic SNPs with 8.5% missing data) rooted with *Dendronanthus* (not shown). In both trees, posterior probability is indicated at the nodes, where a circle denotes PP ≥ 0.95. Intron sequences were only generated for one sample of *M. flava* northeast, therefore it was excluded from the *BEAST analysis. Paintings by Bill Zetterström (from Alström and Mild, 2003), Ren Hathway (from del Hoyo et al., 2004), and Jon Fjeldså (from Alström et al., 2015).

2.2.2. Data collection

For detailed information on DNA extraction, library construction, and sequencing, see Appendix A. We generated SNPs using the double-digest restriction site-associated DNA sequencing (ddRADseq) protocol following methodology described in Peterson et al. (2012). A total of five lanes were sequenced (single-end reads: four 50 bp and one 100 bp). We then constructed a reference genome for a single male individual of *M. alba* to improve the accuracy of our ddRADseq locus assembly (see Appendix A). We constructed three long-insert libraries (2 kbp, 4 kbp, and 9 kbp) using the Nextera mate-pair library kit and one short-insert (350 bp) library using Illumina's TruSeq kit. These libraries were sequenced on two lanes (paired-end reads: 125 bp) of an Illumina HiSeq 2500.

2.2.3. Data filtering

We used Trimmomatic (Bolger et al., 2014) and NxTrim (O'Connell et al., 2015) to trim adapter sequences from our short-read and mate-pair libraries, respectively. Quality was checked with FastQC prior to alignment to the zebra finch genome in CLCGenomics Workbench v6. Raw ddRAD Illumina reads were processed into candidate RAD loci using Stacks v.1.35 (Catchen et al., 2011). Reads were separated by barcode and index, and low-quality reads were removed. Each individual was aligned to the *M. alba* reference genome in bowtie2 (Langmead and Salzberg, 2012). We then used the ref_map.pl pipeline to generate a catalogue for each species using a conservative estimate of 10% for the upper bound of the error rate (Masretta-Yanes et al., 2014). We explored a range of mismatch thresholds (2–4) but report only the results using two allowed mismatches, as higher values did not markedly affect the number of loci obtained or downstream clustering analyses. Finally, loci were clustered across individuals using a minimum allele frequency of 0.05. To preclude paralogous sequences, a maximum observed heterozygosity of 50%. We used custom R scripts to further filter our data according to linkage and missing data (described below).

The vast majority of RAD sequencing studies are conducted without a reference genome and assume that choosing one SNP per locus will result in an unlinked dataset. However, this may violate the requirements of downstream population genetic structure and phylogenetic analyses. To explore the impact of this assumption, we compiled a "pseudo-unlinked" and a "truly" unlinked dataset by filtering out all but one locus within a 100 kb range.

As missing data can affect downstream PCA results (Dray and Josse, 2014), phylogenetic inference (Huang and Knowles, 2014; Wagner et al., 2013), and other analyses as well (Nakagawa and Freckleton, 2008), we further filtered datasets according to different levels of missing data (all possible combinations of 25%, 50%, and 75% missing loci and individuals). Together with our linkage filtering, we compiled 18 datasets per species (two linkage treatments × three missing loci treatments × three missing individual treatments). We used all 18 datasets independently when conducting population structure analyses, Mantel tests, and RAxML phylogenetic trees. For all other analyses, we used a 50% threshold for both missing loci and individuals.

2.3. Phylogenetic analyses

2.3.1. Time-calibrated species tree and gene trees

Nuclear DNA and mtDNA can support contradicting phylogenetic relationships and methods that ignore incomplete lineage sorting may fail to accurately estimate species tree relationships (Maddison, 1997). Because *Motacilla* lacks a species tree, we estimated the first time-calibrated species tree for this group using combined mtDNA and nuclear intron data. We implemented molecular rate calibration in *BEAST v1.8.4 (Drummond and Rambaut, 2007; Heled and Drummond, 2012)

using a published *Motacilla*-specific rate of 2.7% per million years for ND2 (Li et al., 2016). We also placed a broad prior on the *Dendronanthus-Motacilla* split (Gamma (9;2)) to reflect the oldest reliable *Motacilla* fossil (5.5–7 mya Jánossy, 1991).

Recent simulation studies demonstrate that tree priors can have a large impact on divergence time dating when using datasets with mixed intra- and interspecies sampling (Ritchie et al., 2016). Therefore, we conducted model selection on converged runs. As mtDNA and nuclear DNA may support different topologies for reasons other than incomplete lineage sorting, such as introgression and sex-biased dispersal, we also conducted analyses on each data type independently. Using nuclear introns, we estimated a species tree in *BEAST. We also estimated separate nuclear and mtDNA gene trees in BEAST. See Appendix A for details.

2.3.2. Concatenated SNP tree

To place subspecies in a wider phylogenetic context, we conducted ML phylogenetic analyses using concatenated ddRAD loci using RAxML v8.2 (Stamatakis, 2014). Due to their large population sizes and assumed recent divergences, wagtails contain a high level of heterozygosity with few fixed SNPs: over 99% of all variable sites contain at least one individual with a heterozygous SNP. Most methods count heterozygous sites as missing data and researchers have typically excluded these sites. We implemented the method of Lischer et al. (2014) to generate 500 random haplotype samples from sequences with multiple heterozygous sites. We then inferred ML phylogenies using RAxML for each of these datasets. To account for potential SNP ascertainment bias, we implemented the Felsenstein correction (Leaché et al., 2015). See Appendix A for details.

2.3.3. SNP species tree

To estimate a species tree, we implemented SNAPP (Bryant et al., 2012). SNAPP uses biallelic loci and requires at least one representative SNP from each species at each locus. Species assignments were based on population structure estimates (see Section 3.2), and individuals with admixture were excluded to avoid model violations and branch length underestimation (Leaché et al., 2013). Because SNAPP is computationally intensive, convergence issues prevented us from using our full sampling scheme. We therefore ranked individuals by missing data and admixture proportions, and only included the top four from each population. All individual assignments were made with > 95% posterior probability. Mutation rates (u, v) were both fixed at 1 and default parameters were used for the gamma prior (alpha 11.75, beta 109.73). Altering parameter values for the prior distributions on population size and tree length to better reflect wagtail population history caused convergence failure. For further information on running SNAPP, see Appendix A.

2.4. Population genetic structure

Grouping individuals based on phenotype may be misleading, especially given that current wagtail taxonomy does not fully reflect the mitochondrial or nuclear trees (Alström and Mild, 2003; Voelker, 2002; Pavlova et al., 2003; Alström and Ödeen, 2002; Ödeen and Björklund, 2003; Ödeen & Alström, 2001). Objective, genetic-based methods are preferable for inferring the number of genetically distinct populations and the assignment of individuals to those populations. As published mtDNA-based studies suggest *M. flava* and *M. citreola* are polyphyletic (Voelker, 2002; Pavlova et al., 2003; Alström et al., 2002; Ödeen and Björklund, 2003; Ödeen & Alström, 2001), we initially ran all population structure analyses on these two species combined. A similar approach was taken for the "African" clade (*M. clara*, *M. capensis*, *M. flaviventris*, *M. bocagii*) and the "black-and-white wagtails" (*M. alba*, *M. aguimp*, *M. samveasnae*, and *M. grandis*). Focal Eurasian wagtail species were further analyzed independently.

To *de novo* identify the optimal number of clusters in our data, we implemented two methods: the model free Discriminant Analyses of Principal Components (DAPC) in adegenet (Jombart et al., 2010, 2008) and the model-based maximum-likelihood (ML) method, ADMIXTURE v1.3 (Alexander et al., 2009). One caveat of ADMIXTURE is that it assumes discrete ancestral or parental populations. When organisms exhibit continuous spatial population structure, recent studies tend to use PCA methods like adegenet (Royal et al., 2010). However, adegenet has the undesirable behavior of assigning individuals to populations with unrealistically high probability. Therefore, to reduce the impact of their respective biases on downstream analyses, we employed these methods in concert to estimate K and individual assignment probabilities.

First, we ran ADMIXTURE with variable numbers of clusters K = 1–10. We then plotted 10-fold cross-validation values terminated with default criteria, to choose the optimum value of K. Second, we ran the k-means clustering to assess groups using both AIC and BIC. We implemented DAPC to maximize differences between groups while minimizing variation within groups. To assess how many PCs to retain, we used cross-validation (xvalDapc) with 100 replicates and retained the number of PCs with the lowest mean squared error.

To assess genetic differentiation among phenotypes, we used DAPC to find the largest distance between subspecies defined a priori. For each species, we ensured that a minimum of one individual per subspecies per locus was present. On average, this additional filtering step reduced our dataset by 40%. To determine the diagnosability of subspecies, we then compared our DAPC results to analyses with randomized subspecies definitions. Because rare alleles may be younger than common alleles, and may track more recent demographic or selective events (Gompert et al., 2014), we explored their effect on population genetic inference by alternatively (1) pruning of all sites with minor allele frequency < 10% or (2) building a matrix with only low-frequency alleles.

2.5. Isolation by distance

Populations in close proximity are expected to be more genetically similar than those located farther apart (Wright, 1943). To explore whether our estimates are the result of the clustering of individuals with distinct allele frequencies or structure due to separation in space, we conducted Mantel tests on each species using the mantel.randtest function in ade4 (Dray and Dufour, 2007) and ran these for 1 million permutations. To account for Earth's curvature, geographic distance was calculated using the Great Circle distance in the R package sp (Bivand et al., 2013).

Given that the ability of Mantel tests to detect isolation by distance (IBD) has been a recent area of debate (Legendre et al., 2015; Guillot and Rousset, 2013), we also implemented the Estimated Effective Migration Surfaces (EEMS) method (Petkova et al., 2016) to model the relationship between geography and genetics. EEMS allows visualization of variation in effective migration across each species' breeding distribution and the identification of corridors and barriers to gene flow by implementing a stepping-stone model over a dense grid. We used species' breeding ranges (BirdLife International and NatureServe, 2015) and a dense grid of equally sized demes. To show that our results were independent of grid size, we ran each analysis using 250, 500, and 750 demes from three independent chains for 20 million MCMC iterations with a 10 million iteration burn-in. We first ran a series of short preliminary runs to choose parameter values that gave acceptance ratios between 20% and 30%. Graphs were constructed using rEEMSplots (Petkova et al., 2016). We visualized each run separately and checked convergence of MCMC runs (log posterior plots and Gewke diagnostic tests, Heidelberger and Welch test in coda (Plummer et al., 2006) before combining across runs and grid numbers to construct final consensus graphs.

3. Results

We successfully constructed a 956 Mbp reference genome for *M. alba* which was used to assemble ddRADseq loci. A total of 219 million quality filtered reads were aligned with 5x average coverage to 73% of the zebra finch genome. Reflecting this low coverage, the N50 of the alignment was only ~3 kbp. Reference mapping of 442.5 million ddRADseq reads resulted in 8.2 × 104 unique loci across 246 individuals with an average coverage of 29.7 (BioProject PRJNA356768). See Table S1 for individual-level details.

For each species group, we analyzed 18 different dataset combinations consisting of either "pseudo-unlinked" or unlinked SNPs, and varying levels of missing data at the locus and individual level. On average, controlling for linkage reduced each dataset by 70%. This SNP pruning had potential to either remove biased SNPs or, if linkage was not a problem, to remove relevant information from the analysis. However, linkage did not alter our population structure estimates (Table S2) or Mantel tests (Table S3). A common rule-of-thumb is to consider a 1 Mbp window, however, a window of this size drastically reduced our dataset. Therefore, we used a window of 100 kbp. While our study suggests that the use of "pseudo-unlinked" loci may not be an issue for estimation of population genetic structure, a rigorous investigation of this question is needed.

Overall, filtering data based on missing data at the individual level had a larger impact than filtering data based at the locus level (Table S2). As individuals with excess missing data were removed, population structure estimators were less likely to find a consistent result than when loci were removed. We present the majority-rule analyses (50% missing data) and "pseudo-unlinked" SNPs.

3.1. Phylogenetic analyses

The *Motacilla* relationships inferred in our time-calibrated species tree Fig. 1a were confirmed by genome-wide SNPs Figs. 1b and 2b. Relationships were, for the most part, consistent with current wagtail taxonomy, but incongruent with the mtDNA Fig. 2a and nuclear (Fig. S2) gene trees.

3.1.1. Mitochondrial relationships

The mtDNA tree Fig. 2a did not recover monophyletic clades consistent with geographic distribution. Instead, mtDNA recovered an "African clade" consisting of four out of the five Afrotropical/Malagasy species and an "Eurasian clade", which includes the Afrotropical *M. aguimp*. Consistent with previous mtDNA studies, both *M. flava* and *M. citreola* were polyphyletic within the Eurasian clade. *M. flava* was separated into a western clade (Z) and an eastern clade, which was further divided into two sub-clades (X, southeast; Y, northeast). Each *M. flava* clade was sister to a corresponding *M. citreola* clade, which was not geographically consistent: *M. c. citreola* was split between clades X (western *citreola* samples) and Y (eastern *citreola* samples), and clade Z included the southern *M. c. calcarata*. There was no structure consistent with subspecies designations in *M. alba, M. citreola,* or *M. flava*. Similar to previous mtDNA trees, we found no support for a monophyletic "yellow" plumage clade. However, the black-and-white Palearctic species did form a monophyletic group, although this clade excludes the Afrotropical *M. aguimp*.

3.1.2. Nuclear relationships

Both the SNAPP species tree analysis Fig. 1b and the RAxML analysis of the concatenated SNP data Fig. 2b supported African and Eurasian clades, with the former divided into two primary clades notable for their plumage coloration: clade A includes only "black-and-white" (melanin-based) species, whereas clade B contains only species with green/yellow (carotenoid-based) plumages. These nuclear relationships are strikingly incongruent with mtDNA in (1) placing *M. aguimp* within the "black-and-white" clade, (2) supporting the monophyly of both *M. flava* and *M. citreola*, and (3) placing *M. bocagii* within the African clade. In the SNAPP tree, the Afrotropical mainland species (*M. capensis* and *M. clara*) formed a clade and the insular species (*M. flaviventris* and *M. bocagii*) formed another Fig. 1b. RAxML found high support for splits

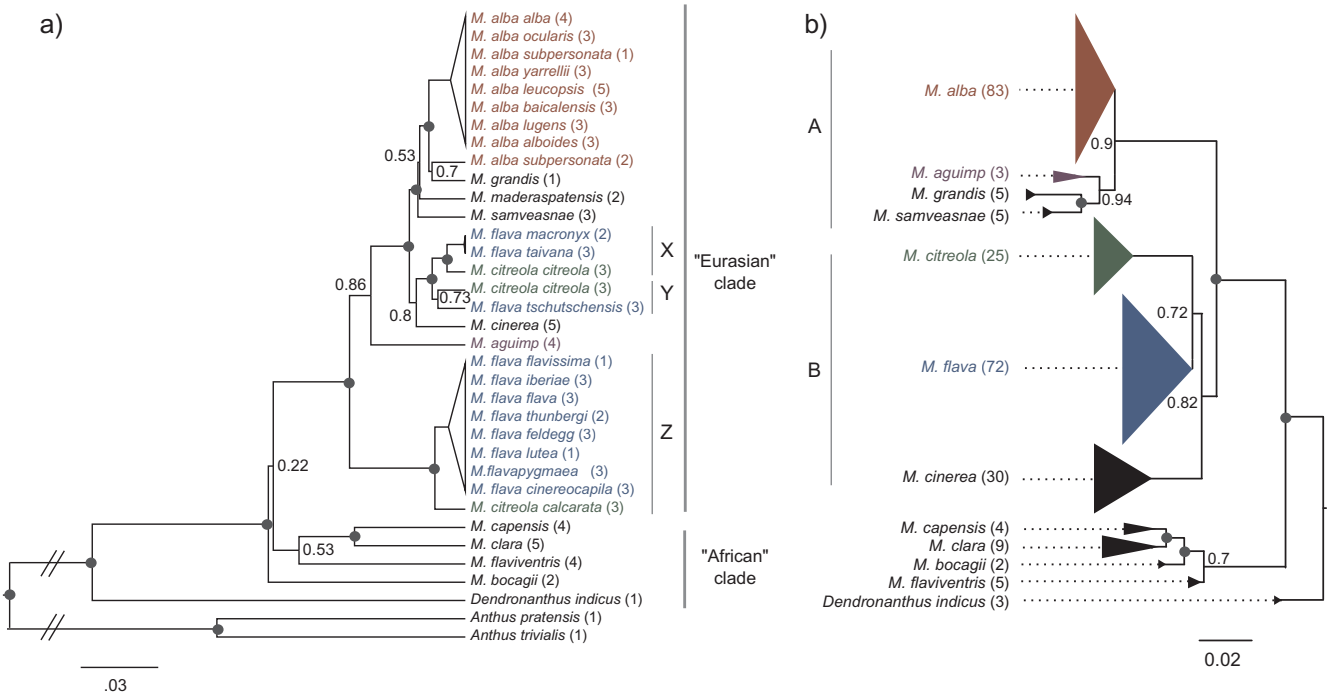

Fig. 2. Gene trees. (a) BEAST mitochondrial gene tree, with bars indicating polyphyletic groupings of *M. flava* and *M. citreola* (X, Y, Z). (b) RAxML consensus tree of 500 random haplotype datasets (Lischer et al. (2014) method) estimated from genome-wide SNPs. In both trees, nodes of monophyletic, single-species clades are collapsed for ease of viewing. All collapsed nodes have bootstrap values > 95%. Numbers in parentheses indicate sample size (a and b). Bars indicate the two Eurasian-breeding clades: A = "black-and-white" plumage and B = "yellow" plumage clade. In both trees, posterior probability indicated at nodes, where a circle denotes PP ≥ 0.95.

among species but no support for within-species relationships (see Dryad). Accordingly, within *M. alba*, *M. flava,* and *M. citreola*, no subspecies were monophyletic.

3.1.3. Species tree relationships

The topology of the time-calibrated species tree Fig. 1a was similar, but not identical, to the SNAPP tree Fig. 1b. First, it included *M. maderaspatensis*, for which no SNP data were available, and therefore found different relationships within the "black-and-white" plumage clade. Second, in both trees there was strong support for a clade comprising *M. citreola* and *M. flava*, however these trees differed in whether they supported a monophyletic *M. flava* (SNAPP) or not (*BEAST). Overall, the major splits were consistent across species trees. The African and Eurasian clades split towards the end of the Pliocene (~3.25 mya; 2.55–3.97 95% HPD), with early and relatively widely spaced divergences from the early Pleistocene in the former clade, and an explosive radiation within < 0.88 million years (my) during the mid- to late Pleistocene.

3.2. Population structure

At the species level, *de novo* population estimation methods were able to distinguish recognized species, except *M. aguimp*, which was grouped with western *M. alba* by both DAPC (Table S2) and ADMIXTURE. Narrowing our focus to phenotypically distinct subspecies, the full dataset showed genetic structure discordant with phenotype in all analyzed species. Instead, eastern and western populations (K = 2) were resolved by both ADMIXTURE and *de novo* DAPC in *M. citreola* (2 subspecies, Fig. 3) and *M. flava* (13 subspecies, Fig. 4, Table S2). Depending on which *M. alba* (9 subspecies) + *M. aguimp* dataset was analyzed in DAPC, there was support for either K = 2 or K = 3 (Table S2).

ADMIXTURE consistently supported K = 2 Fig. 5c). For all three datasets, individual assignments were consistent across methods when K = 2. Finally, we found no difference in population structure when rare variants were either considered separately or ignored.

Ordination plots from DAPC analyses with subspecies defined a priori provide additional resolution of population structure in *M. flava* Fig. 4b and *M. alba* Fig. 5b. In both species, the x-axis of the ordination plots showed a correlation between diagnosability and longitude. In *M. flava*, all subspecies clusters overlapped, suggesting genetic distinctiveness is limited (Fig. S3), especially since the DAPC method acts to maximize whatever differences are present. We note our sampling of *M. f. tschutschensis* only included individuals from the easternmost portion of its range. However, *M. f. tschutschensis* is distributed from Central Eurasia through the Kamchatka Peninsula and into Western Alaska with its geographic center close to that of *M. f. macronyx* and *M. f. taivana* (Alström and Mild, 2003). Therefore, its position in the center of the DAPC ordination plot conforms to the idea that this pattern tracks longitude, and it is unclear whether these clusters correspond to subspecies distinctions or merely IBD.

The ordination plot generated by assigning *M. alba* samples to subspecies was influenced by missing data. The less stringent the threshold, the more distinct *M. a. subpersonata* became. This trend was found in both pseudo-unlinked and unlinked datasets, with the unlinked presenting the most extreme cases (Fig. S4). However, it is remarkable that in three out of four cases, *M. a. subpersonata* was more distinct than *M. aguimp* (Fig. S4) and all other *M. alba* subspecies Fig. 5b. As no other subspecies displayed this behavior, we removed *M. a. subpersonata* and recompiled datasets. This resulted in *M. aguimp* samples being tightly clustered and diagnosable in all analyses (Fig. S5). The remaining *M. alba* subspecies clusters were overlapping and did not show clustering consistent with subspecies, but did show clustering on either side of the vertical axis that was consistent with our *de novo* K = 2 clusters. Taken together with evidence that clustering methods may not work well with small sample sizes (Fogelqvist et al., 2010), we considered both K = 2 (without *M. a. subpersonata*) and K = 3 (*M. a subpersonata* included as own population) population histories in the following analyses.

Population structure in *M. cinerea* was not consistent across methods. Whereas ADMIXTURE supported a single panmictic population, DAPC supported an east-west split (K = 2) consistent with the other Eurasian wagtails. These contrasting findings may be explained by simulation and other empirical studies, which have demonstrated that Bayesian clustering methods fail to detect structure when genetic divergence is very low (Waples and Gaggiotti, 2006; Benestan et al., 2015).

3.3. Isolation by distance

When dealing with genetic data from evenly spaced samples from a spatially structured organism, the expected behavior of PCA is to return clusters that are related to the geographic origin of each individual sample (Leaché et al., 2013). However, spatial autocorrelation may bias interpretation of PC analyses. While Mantel tests generally support a history of IBD in all species groups (p-value < .05), p-values vary

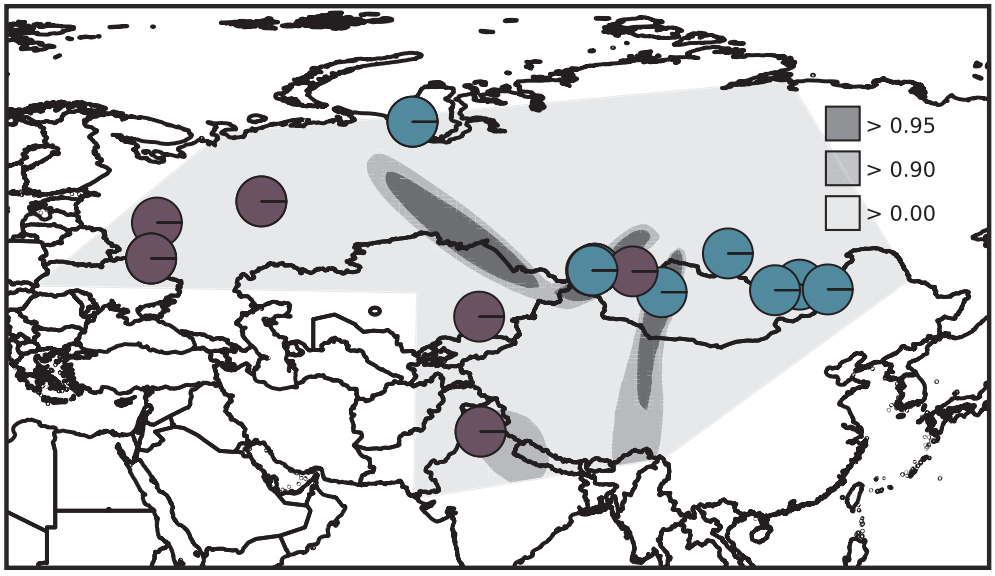

Fig. 3. Posterior probabilities of effective migration rates of *M. citreola* estimated by EEMS. Birds from India belong to *M. c. calcarata*, all others to *M. c. citreola*. Pie charts are located at sampling sites and denote the posterior probability of ADMIXTURE assignments (here, K = 2).

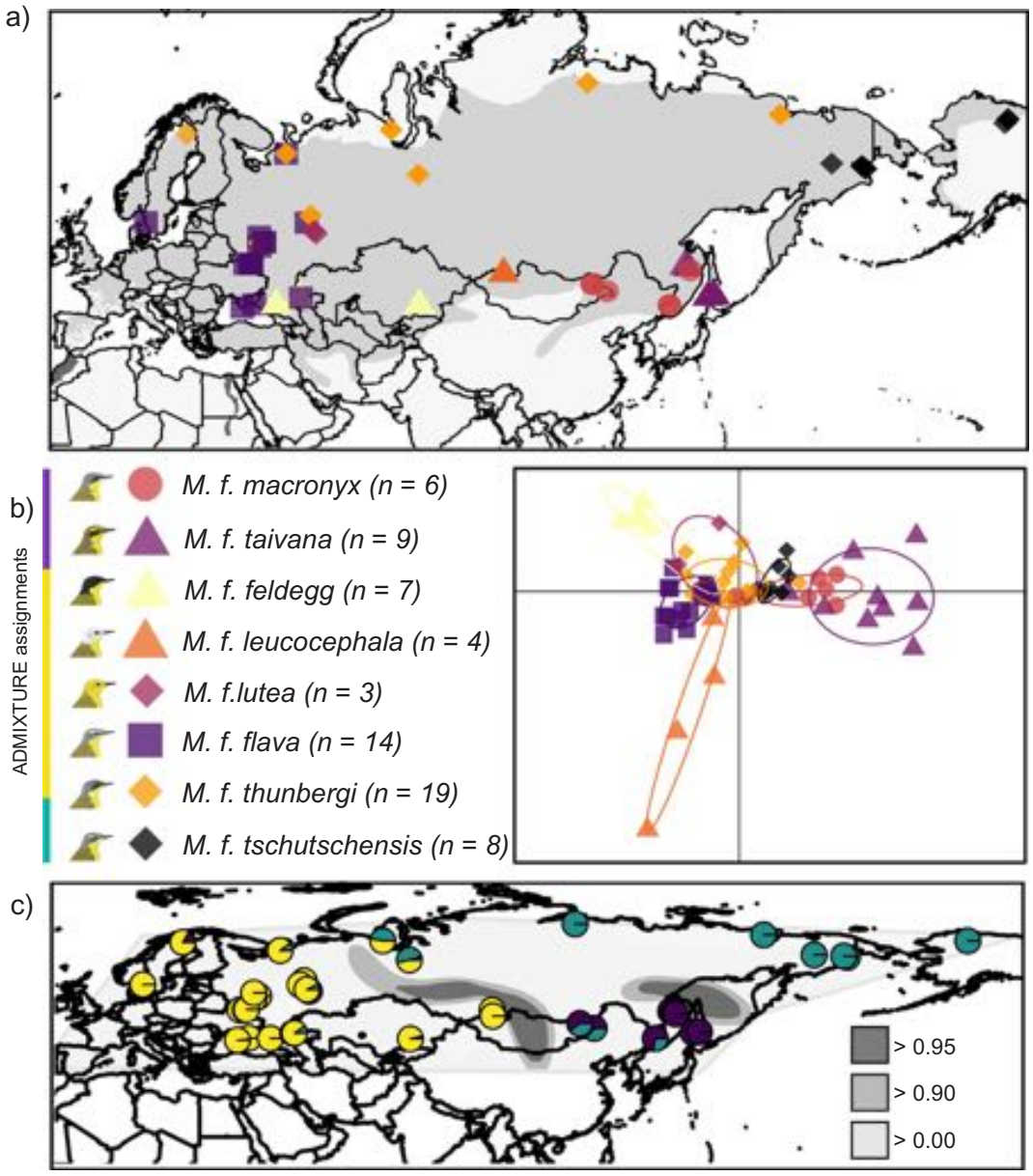

Fig. 4. (a) *M. flava* breeding distribution and sampling localities. (b) DAPC plot of genetic clustering by subspecies (c) Posterior probability of effective migration rates estimated by EEMS show two barriers and K = 3. Pie charts are located at sampling sites and denote the posterior probability of ADMIXTURE assignments (here, K = 3). *M. f. thunbergi* is the only subspecies split between two populations.

according to missing data (Table S3).

As recent studies demonstrate that sampling design can strongly bias interpretations of Mantel tests (Guillot and Rousset, 2013), we further explored IBD using EEMS, a program that can distinguish whether support for IBD is the result of either geographically distant and differentiated populations or a continuous cline in genetic differentiation. Strong linear relationships between predicted and observed genetic dissimilarities confirm that the EEMS model fit our data (Petkova et al., 2016). To assess support for true barriers, we examined plots of dissimilarity between pairs of sampled demes for non-linearity and resulting effective migration map for a singular uniform barrier. EEMS strongly supported the existence of a barrier between eastern and western/southern *M. citreola* Fig. 3, consistent with the findings of population structure estimators.

In *M. cinerea*, EEMS found no support for a barrier, consistent with population structure estimates from ADMIXTURE (K = 1) but not DAPC (K = 2). We were unable to include samples from two of the three subspecies (small populations restricted to either the Azores and Madeira Islands) and can only conclude that there is support for a single panmictic mainland Eurasian *M. cinerea*.

We found conflicting support for IBD in *M. flava*. Despite a linear trend in predicted and observed genetic dissimilarities, supporting IBD, there was also strong support for the existence of two discontinuous barriers Fig. 4c. We interpreted this to mean that while IBD exists across the whole *M. flava* range, there are areas where gene flow is not significant. For assignment of subspecies into these populations, see Figs. 1b and 4. Since these results differ from those of *de novo* DAPC and ADMIXTURE, we tested whether DAPC can distinguish between these groups when given individual assignments. First, we fixed K = 3 and ran *de novo* DAPC to see if we could find the same populations resolved by EEMS. Second, we defined EEMS populations and ran DAPC. As both methods overwhelmingly supported the three populations found by EEMS, we used these definitions in species tree analysis (see Section 2.3).

M. alba also showed a linear trend and a patchy effective migration map, lending support for IBD. However, a closer look at posterior mean

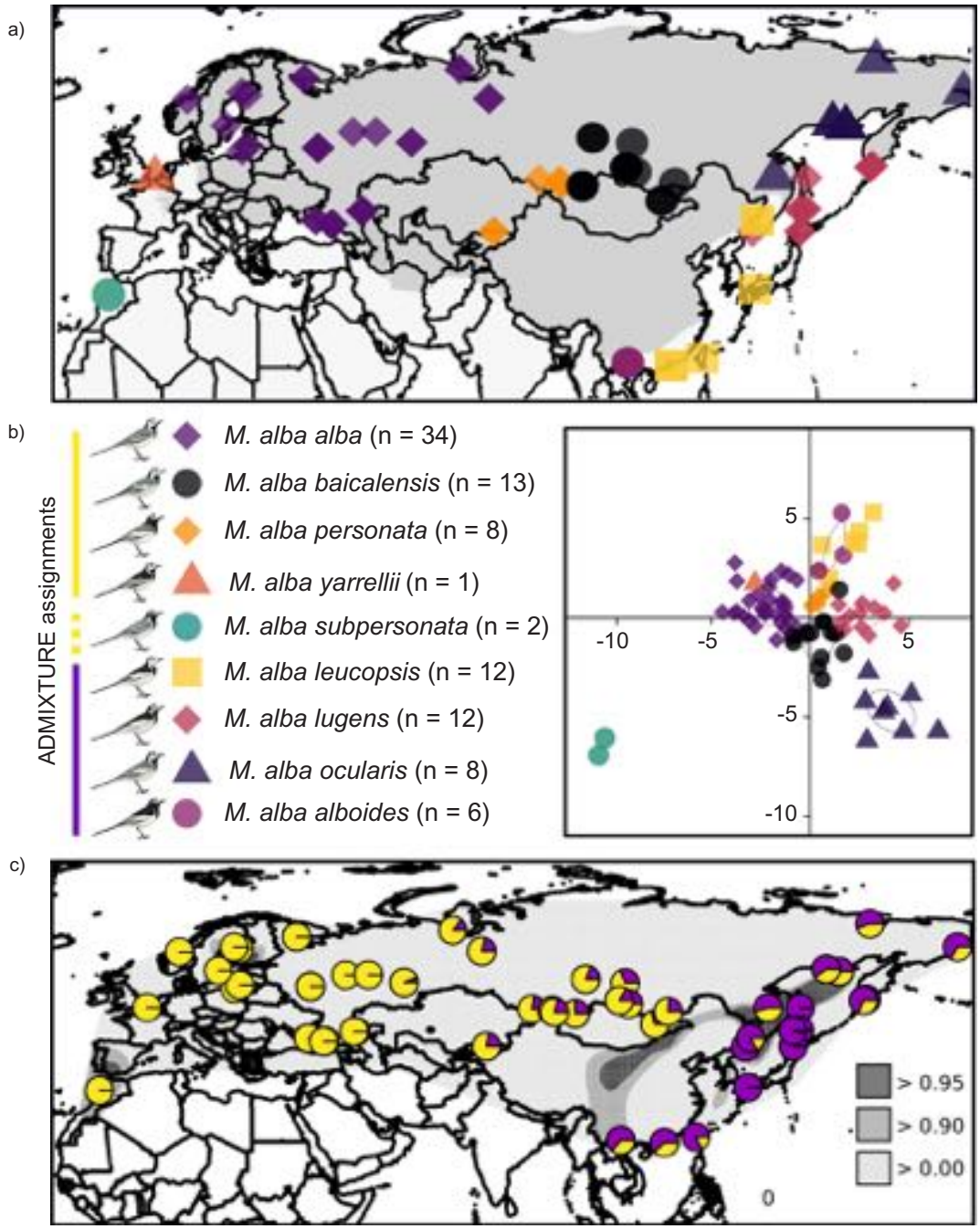

Fig. 5. (a) *M. alba* distribution and sampling localities. (b) DAPC plot of genetic clustering by subspecies. Colored bar denotes ADMIXTURE assignments, where the dotted yellow bar indicates conflicting evidence for either including *M. a. subpersonata* in the eastern *M. alba* population or delimiting a third monotypic population. (c) Posterior probabilities of estimated effective migration rates show 2–4 barriers and K = 2. Pie charts are located at sampling sites and denote the posterior probability of ADMIXTURE assignments (here, K = 2). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of effective migration rates demonstrated that some areas have a markedly lower rate of migration than average Fig. 5c. EEMS largely supported the east-west division of *M. alba* and the distinctness of *M. a. subpersonata*. However, an inverted Y-shaped migration surface split eastern *M. alba* into northern (congruent with *M. a. lugens* and *M. a. ocularis*) and southern (congruent with *M. a. leucopsis* and *M. a. alboides*) populations Fig. 5c. To explore the north-south division further, we implemented DAPC, but it was unable to distinguish between these populations. These results may stem from low sample sizes, as EEMS is more robust to biased sampling than PCA methods (Petkova et al., 2016).

4. Discussion

Modern phylogeographic studies hope that the population divergence history of widespread and variable species can be disentangled with sufficient numbers of variable, neutrally evolving SNPs (Brumfield et al., 2003; McCormack et al., 2013). In the present study, genome-wide SNPs provide a robust estimate of species-level relationships, most of which are corroborated by the combined analysis of mtDNA and nuclear introns, but disagree stronger with the mtDNA gene tree. This is a major step towards our understanding of wagtail phylogeny compared to previous analyses of fewer species and considerably less sequence data (Pavlova et al., 2005; Voelker, 2002; Pavlova et al., 2003; Li et al.,

2016; Alström and Ödeen, 2002; Ödeen and Björklund, 2003; Ödeen & Alström, 2001; Alström et al., 2015). In contrast, SNPs fail to distinguish between recently diverged, phenotypically distinct taxa that are usually treated as subspecies. Instead, three of the most widespread Eurasian wagtails are structured along an east-west axis. Recent divergence, ongoing or past gene flow, and demographic processes may have obscured patterns of differentiation at neutral loci (cf. Campagna et al., 2011; André et al., 2010).

4.1. Phylogeny

The topologies estimated from the analyses of SNP data and combined mtDNA and nuclear intron data are largely congruent and, for reasons discussed below, provide a better estimate of the phylogeny than that given by mtDNA alone. The SNP-based species tree (SNAPP) differs from the concatenated SNP tree (RAxML) regarding the positions of *M. aguimp* and *M. flaviventris*. Simulations have shown that concatenation can be statistically inconsistent and may favour the wrong tree (e.g. Degnan and Rosenberg, 2006; Edwards et al., 2007; Kubatko and Degnan, 2007). Both SNAPP and *BEAST are based upon the multispecies coalescent and, using different data, are topologically identical with respect to *M. aguimp* and *M. flaviventris*. Because no SNP data was available for *M. maderaspatensis*, the topological incongruence between the two species trees with respect to the "black-and-white" clade (A) cannot be evaluated.

Consistent with previous studies, we find strong support for an initial split into an "African" and a "Eurasian" clade (Voelker, 2002; Alström and Ödeen, 2002). Within the African clade, we recover a sister relationship between the two widely sympatric mainland species, *M. capensis* and *M. clara*. The other two "African" species are insular: *M. bocagii* is endemic to São Tomé Island off the west coast of Africa, while *M. flaviventris* is endemic to Madagascar. Originally thought to belong to the sylvioid superfamily, *M. bocagii* bears little resemblance to other *Motacilla* species in both plumage, structural morphology, habitat, and behavior (Alström et al., 2015) (cf. Fig. 1). Using previously unsampled individuals of *M. bocagii*, we confirm its placement within *Motacilla*, and the species trees suggest that it is sister to the widely disjunct *M. flaviventris*.

The "Eurasian" clade is subdivided into monophyletic "black-and-white" (A) and "yellow" (B) plumage clades. The latter comprises of *M. cinerea*, *M. flava,* and *M. citreola*, which are broadly sympatric across large parts of the Palearctic (Alström and Mild, 2003). Contrary to the mtDNA tree, we find strong support for monophyletic *M. flava* and *M. citreola* in our SNAPP tree. In contrast, the *BEAST species tree recovers a monophyletic *M. citreola* but a paraphyletic *M. flava,* but with weaker support. Taken together, the short branch lengths and conflicting results suggest that these species underwent a rapid, complex divergence process. However, past introgression may have been an integral part of this process and may have affected branch lengths (Leaché et al., 2013). Indeed, *M. flava* and *M. citreola* hybridize occasionally, predominantly on the expanding western edge of the latter's range (reviewed in Alström and Mild, 2003). Future studies would benefit from more explicit tests of introgression.

Species belonging to the "black-and-white" clade (A) have mostly allopatric distributions, together covering nearly all of Europe, Asia, and Africa. While our *de novo* population structure analyses are unable to differentiate between *M. aguimp* and *M. alba*, *M. aguimp* is identifiable using *a priori* assignment. This is because our SNAPP analyses were conducted using three *M. alba* tips (defined by population structure estimates) and one *M. aguimp* tip. By accounting for the stochastic nature of lineage sorting, we are able to show that *M. alba* is indeed a monophyletic species and is sister to *M. aguimp*. The distinctness of *M. aguimp* is further substantiated by its multiple differences (size, plumage, song, call) from all subspecies of *M. alba* (Alström and Mild, 2003).

The remaining relationships among the species within clade A are not fully resolved, partly because of the lack of the Indian endemic *M. maderaspatensis* in the SNP dataset. However, we confirm monophyly of two of its members, *M. grandis* and *M. samveasnae*, which had been poorly sampled in previous studies. Endemic to Japan, the monotypic *M. grandis* is easily distinguishable from its partially sympatric relative *M. alba* by its plumage, larger body size, song, call, and habitat preference (Alström and Mild, 2003). The only previous analyses including *M. grandis* were inconclusive with respect to the position of *M. grandis* relative to *M. alba* (Li et al., 2016; Alström and Ödeen, 2002; Ödeen & Alström, 2001), and indicated possible incomplete lineage sorting or introgression. Using four newly sampled *M. grandis* individuals, our results confirm the genetic distinctness of this species, and find no evidence of introgression despite documented cases of hybridization (Alström and Mild, 2003).

Within *M. alba*, we only find evidence for 2–4 genetically distinct populations. Previous mtDNA-based studies on *M. alba* delineated three (Pavlova et al., 2005) or four (Li et al., 2016) populations. The only population fully consistent with our SNP results was the rare Moroccan endemic *M. a. subpersonata*. Li et al. (2016) included denser sampling of the southeastern part of the *M. alba* range (*M. a. leucopsis*; *M. a. alboides*), but this sampling difference is unlikely to explain the incongruent population structure and instead likely reflects differences inherent to mitochondrial and nuclear data, such as mutational rate and inheritance patterns.

4.2. Mito-nuclear discordance

Mito-nuclear discordance is prevalent in many animal systems (Toews and Brelsford, 2012) and leads to conflicting biogeographic patterns inferred from these two marker systems (reviews in Toews and Brelsford, 2012; Funk and Omland, 2003; McKay and Zink, 2010). There are relatively few documented cases where mitochondrial and nuclear trees infer strongly supported, yet conflicting relationships among avian species (Dong et al., 2014; Kearns et al., 2014; Lavretsky et al., 2014; Shipham et al., 2016), although this has been shown in several other animal groups (e.g. Bryson et al., 2010; Hailer et al., 2012; Tang et al., 2012).

The genus *Motacilla* represents a compelling avian example of mito-nuclear discordance. Here, we show that nuclear data better reflect true species relationships. Non-genetic data lend support to our findings. For example, the two subspecies of *M. citreola* are nearly identical in all plumages except male dorsal breeding plumage color, but differ markedly from all *M. flava* subspecies by multiple plumage characters (Alström and Mild, 2003). Likewise, *M. aguimp* resembles the other black-and-white species much more than any other species in multiple plumage traits and vocalizations (Alström and Mild, 2003). Parallel evolution in unrelated lineages is highly unlikely to produce such similarity in multiple, independent traits. However, this pattern could be the result of lineage sorting of genetic polymorphisms retained across successive nodes in the tree (Avise and Robinson, 2008), introgression and successive fixation of foreign mtDNA (Rheindt and Edwards, 2011), or other processes.

Several, non-exclusive, explanations for mito-nuclear discord have been proposed, e.g. adaptive introgression of mtDNA, lineage sorting effects, demographic disparities, asymmetric mate choice, and sex-biased asymmetries in dispersal and/or hybrid fitness (Toews and Brelsford, 2012; Funk and Omland, 2003; McKay and Zink, 2010; Rheindt and Edwards, 2011). For example, Pons et al. (2014) invoked drift in small populations and asymmetric mate choice in a colonization context as the most likely explanation for mitochondrial paraphyly in two North American gull species, while noting that selection could not be ruled out. Wang et al. (2014) suggested that historical introgression and selective sweep and/or genetic drift might be the main causes of unexpectedly low mtDNA differentiation between two long-tailed tit species. In contrast, McKay and Zink (2010) concluded based on a survey of the literature that the most common cause of gene tree

paraphyly was due to “imperfect taxonomy”.

We have not tried to distinguish between the alternative hypotheses for mito-nuclear discordance in wagtails, but we suggest that ancient introgression and subsequent fixation of foreign mtDNA is the most likely explanation for discordance at the species-level. It has often been suggested that because females are the heterogametic sex in birds (ZW), introgression of mtDNA should be less likely than introgression of autosomal loci in accordance with “Haldane’s rule” (Funk and Omland, 2003; McKay and Zink, 2010; Moore, 1995; Carling and Brumfield, 2008). However, there was limited support for this prediction in a recent review (Toews and Brelsford, 2012).

In general, hybridization is more prevalent during the early stages of the speciation process, whereas hybrid infertility and/or inviability are more likely to evolve later. (e.g., Price and Bouvier (2002) concluded in a review of postzygotic incompatibilities in birds that complete loss of F1 hybrid fertility usually takes millions of years). Accordingly, we suggest that mtDNA introgression was equally probable as nuclear introgression at the early phases of wagtail speciation. Introgressed mtDNA is more likely to reach fixation than introgressed nuclear DNA because, among other things, it is non-recombining and has a smaller effective population size (Moore, 1995). Because of the latter, mtDNA is more likely than autosomal DNA to accurately represent the species phylogeny under a lineage sorting scenario. However, when times between successive splitting events are short, stochastic lineage sorting may become a problem for mtDNA (Moore, 1995).

We also see mito-nuclear discordance at the population level, with mtDNA producing paraphyletic clades of *M. flava, M. citreola,* and *M. alba*. If populations were isolated during the Pleistocene with foreign, introgressed mtDNA segregating in the population, then reduced population sizes would have increased the likelihood of these foreign mtDNA haplotypes being fixed in the population. However, selection may favor different mtDNA haplotypes in different environments (Toews and Brelsford, 2012) and it is unclear from this study alone whether this may be playing a role in wagtail mito-nuclear discordance.

4.3. Divergence times

The time-calibrated species tree finds support for wagtail diversification beginning in the Pliocene (~3.25 mya). This is in conflict with previous studies which found an older split between clades (4.5–5 mya Voelker, 2002; Alström et al., 2015). While there is variation across studies in topology, all have resolved more recent divergence times for the Eurasian clade compared to the African clade. However, the divergence times recovered in this study are nearly half those previously estimated. This inconsistency is likely due to different sampling schemes. Whereas Voelker (2002) used only mtDNA to construct a time-calibrated gene tree, Alström et al. (2015) used both mtDNA and nuclear introns but fixed the topology and did not include comprehensive sampling of the genus. Instead, the latter study prioritized sampling of the entire Motacillidae family. We also note that the *M. alba* specific mtDNA rate used to calibrate our species tree was derived from the generic avian cyt-b rate, the same rate used in previous studies. It is well established that molecular rates vary across lineages (Britten, 1986) and this rate may be inaccurate. Furthermore, gene flow leads to underestimation of divergence times (Leaché et al., 2013). While infrequent hybrids between parapatric Eurasian species pairs have been documented (reviewed in Alström and Mild, 2003), ancient hybridization may also impact divergence time estimates. Future studies should conduct a comprehensive exploration of introgression between wagtail species.

4.4. Discordance between phenotype and genotype

Discordance between populations identified by genetic markers and plumage phenotypes have been noted in several young radiations (Olsson et al., 2010; Milá et al., 2007; Mason and Taylor, 2015; Poelstra et al., 2014; Lamichhaney et al., 2015). Wagtails provide a striking example of complex plumage evolution, which has apparently taken place very rapidly and recently in some lineages, and with multiple examples of discordance between genetic population structure and phenotype. The most dramatic, rapid and recent divergences are found in the *M. alba* and *M. flava* complexes, where the pronounced plumage differences among the many subspecies are not at all reflected in any of our genetic datasets. Instead, the only indications of genetic divergence are between broad geographical regions, each of which is home to two or more different-looking subspecies.

For example, the *M. alba* complex consists of nine subspecies with markedly different plumages. However, we only find evidence for 2–4 genetically distinct populations in broad agreement with previous mtDNA-based studies (Pavlova et al., 2005; Li et al., 2016). Only the rare endemic Moroccan *M. a. subpersonata* stands out as distinct in both plumage, mtDNA, and SNPs. Similarly, none of the phenotypically distinct *M. flava* or *M. citreola* subspecies align with the population structure boundaries. The “*thunbergi* phenotype” (*thunbergi*, *macronyx*) is found in all three *M. flava* populations, while the “*flava*” phenotype is found in both the western and northeastern populations. Furthermore, taxa with bright yellow supercilium are found in two of these clades, as well as in widely disjunct geographical areas within the western clade.

Lack of strong genetic differentiation is characteristic of the other Eurasian species. The *citreola* phenotype is split across both *M. citreola* populations, while *M. cinerea* has no population structure. We consider the Eurasian *M. cinerea* a single panmictic population with a strong signal of IBD. This is consistent with a previously described pattern of shallow clinal plumage variation across much of Eurasia (Alström and Mild, 2003).

Most of the plumage similarities and differences among wagtail taxa concern differently coloured plumage patches. It seems possible that these can be switched on and off in different combinations through a rather simple system, which might explain the many cases of seemingly divergent and parallel plumage evolution in wagtails. Recent studies have demonstrated that strong selection can occur at few genes, and that plumage differences can evolve rapidly without corresponding divergence in the rest of the genome (Toews et al., 2016; Mason and Taylor, 2015; Poelstra et al., 2014). For example, within a Eurasian crow (*Corvus*) species complex, there is evidence of assortative mating based on plumage despite no genetic differentiation across most of the genome. Instead, there appears to be simple genetic control of phenotype, but multiple loci are leading to the same phenotype in different populations, thus implicating a more complex multi-genic pathway than previously thought (Vijay et al., 2016; Poelstra et al., 2014).

While our SNP data was sampled from across the genome, it only represents 0.05% of the entire 1.1 Gb wagtail genome. If few genes are responsible for plumage differences and populations experienced recent divergence and selection, then it is not surprising that our data fails to find a genetic signal congruent with phenotype. It is also clear that hybridization is prevalent among parapatrically distributed subspecies of *M. flava*, *M. citreola,* and *M. alba*. Strong selection on plumage loci, mediated through assortative mating and selection against intermediate plumage phenotypes (hybrids), might explain how different plumage traits can be maintained in the face of gene flow in hybrid zones (Poelstra et al., 2014). Lineages showing discordance between phenotypes and genotypes should gradually evolve more concordance between these traits, provided that they remain separated by geographical or reproductive isolating barriers.

Sexually selected plumage traits have potential to diverge rapidly among allopatric, ecologically equivalent populations (Price, 1998), and have been identified as the dominant signals mediating mate-choice and intrasexual aggression in some birds (Seddon et al., 2013). In general, *Motacilla* subspecies vary considerably more in plumage traits closely linked to signalling functions than to ecological (e.g. beak, tarsus, or wing length) traits (Alström and Mild, 2003; del Hoyo et al.,

2004), indicating that sexual selection might have played an important role in phenotypic divergence. Furthermore, the Eurasian clade displays variation in sexual plumage dimorphism, size, seasonal plumage variation, and age-related differences, whereas the African clade does not (del Hoyo et al., 2004; Alström et al., 2015), indicating greater potential for sexual selection in the former group. Indeed, recent work on a hybrid zone between two *M. alba* subspecies provides support for assortative mating based on head plumage maintaining phenotypic boundaries (Semenov et al., 2017).

4.5. Taxonomic implications

Wagtail taxonomy has been much debated over the years, especially the status of the different taxa in the *M. alba* and *M. flava* complexes (reviewed in Alström and Mild, 2003). With the exception of *M. flava* and *M. citreola*, our data support the recent origin and lack of significant differentiation found in previous studies (Pavlova et al., 2005; Voelker, 2002; Pavlova et al., 2003; Li et al., 2016; Ödeen and Björklund, 2003). On account of this, our data does not support treatment as multiple species, except under a "phylogenetic species concept" that uses diagnosability as species criterion.

The non-monophyly of *M. flava* and *M. citreola* in the mtDNA tree has been suggested to be sufficient evidence for splitting the former into three and the latter into two species (Pavlova et al., 2003). Indeed, several authors have recently divided *M. flava* into two species, *M. flava sensu stricto* and *M. tschutschensis*, although with different circumscriptions (del Hoyo and Collar, 2016; Banks et al., 2004; Dickinson and Christidis, 2014). Under the same rationale our mtDNA tree would suggest splitting *M. citreola* into three species, however both nuclear introns and SNP data strongly support the monophyly of *M. citreola*, and therefore invalidate proposals to split this taxon (Pavlova et al., 2003; McKay and Zink, 2010). In contrast, we find inconsistent support for the monophyly of *M. flava*. Our SNP data support a monophyletic *M. flava* complex consisting of three clades with barriers to gene flow, whereas combined analysis of nuclear introns and mtDNA recover *M. flava* as paraphyletic. While it is likely that our SNP data reflects the true species relationships, it is clear that these two species have shared a complex and recent history. However, taxonomic treatment requires more rigorous exploration of demographic histories.

We highlight a few intra-species lineages that our study finds distinct and thus deserve further investigation. In their taxonomic review, Alström and Mild (2003) treated all *M. flava* from northern Scandinavia east to the Kolyma River as *thunbergi*. However, our population genetic structure estimates place all "*thunbergi*" from the Yamal Peninsula eastward in the northeastern clade. This grouping is further supported by mtDNA (Pavlova et al., 2003), analyses of call notes (Bot et al., 2014) and morphology (Red'kin, 2001). Eastern "*thunbergi*" populations have been referred to as *plexa* (Red'kin, 2001) and our results support this distinction. Similarly, the Moroccan endemic *M. a. subpersonata* is distinct amongst *M. alba* and warrants further study. In addition, our SNP data confirm that *M. bocagii* is nested within *Motacilla*, and therefore invalidates recognition of the monotypic genus *Amaurocichla* recognized in the recently published Birdlife Checklist (del Hoyo and Collar, 2016).

5. Data accessibility

- Raw, demultiplexed ddRAD reads: NBCI SRA under PRJNA356768.
- SNP datasets, input files for analyses, and tree files: Dryad repository doi:https://doi.org//10.5061/dryad.008bq.

Acknowledgements

We thank the museums and those who sent us samples: Ulf Johansson, Swedish Museum of Natural History; Brian Schmidt, USNM; Isao Nishiumi, NMSM; Ben Marks, FMNH; Mark Robbins, KU; Jack Withrow, UAM; Paul Sweet, AMNH; Kristof Zyskowski, YPM; Michael Westberg, Bell Museum; Gary Voelker, TCWC; Sharon Birks, UWBM. We thank Martim Melo for sharing tissues and Dai Shizuka facilitating contacts in Japan.

For helpful discussion and comments, we thank Sievert Rohwer, John Klicka, Nick Sly, Shawn Billerman, Ethan Linck, CJ Battey, Garth Spellman, one anonymous reviewer, and the members of the Leaché Lab. For assistance with analysis, we thank Ryan Gutenkunst, Desislava Petkova, and Tommy Tyrberg.

R.B.H. was supported by the NSF-GRFP. Funding came from the AMNH Chapman Grant, Sigma Xi GIAR, NSF DDIG DEB-1501131, and the Burke Museum of Natural History and Culture. We used the UC Berkeley Vincent J. Coates Genomics Sequencing Laboratory, supported by NIH S10 Instrumentation Grants S10RR029668 and S10RR027303, and the advanced computational, storage, and networking infrastructure provided by the University of Washington Hyak supercomputer system. P.A. was supported by the Swedish Research Council (2015-04402), Jornvall Foundation, and Mark & Mo Constantine.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.ympev.2017.11.020.

References

Alexander, D.H., Novembre, J., Lange, K., 2009. Fast model-based estimation of ancestry in unrelated individuals. Gen. Res. 19, 1655–1664.
Alström, P., Mild, K., 2003. Pipits and Wagtails of Europe, Asia and North America. A&C Black. A&C Black, Princeton University Press.
Alström, P., Jønsson, K.A., Fjeldså, J., Ödeen, A., Ericson, P.G.P., Irestedt, M., 2015. Dramatic niche shifts and morphological change in two insular bird species. Roy. Soc. Open Sci. 2, 140364 (140364–140364).
Alström, P., Ödeen, A., 2002. Incongruence between mitochondrial DNA, nuclear DNA and non-molecular data in the avian genus *Motacilla*: implications for estimates of species phylogenies. In: Alström, P. (Ed.), Species Limits and Systematics in Some Passerine Birds. Uppsala University [Ph.D. thesis].
André, C., Larsson, L.C., Laikre, L., Bekkevold, D., Brigham, J., Carvalho, G.R., et al., 2010. Detecting population structure in a high gene-flow species, Atlantic herring (*Clupea harengus*): direct, simultaneous evaluation of neutral vs putatively selected loci. Heredity 106, 270–280.
Avise, J.C., Robinson, T.J., 2008. Hemiplasy: a new term in the lexicon of phylogenetics. Syst. Biol. 57, 503–507.
Banks, R.C., Cicero, C., Dunn, J.L., Kratter, A.W., Rasmussen, P.C., Remsen Jr, J.V., et al., 2004. Forty-fifth supplement to the American Ornithologists' Union check-list of North American birds. Auk 121, 985–995.
Benestan, L., Gosselin, T., Perrier, C., Sainte-Marie, B., Rochette, R., Bernatchez, L., 2015. RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (*Homarus americanus*). Mol. Ecol. 24, 3299–3315.
BirdLife International and NatureServe, 2015. Bird Species Distribution Maps of the World.
Bivand, R.S., Pebesma, E., Gómez-Rubio, V., 2013. Applied Spatial Data Analysis with R.
Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2130.
Bot, D., Groenendijk, H., van Oosten, S., 2014. Eastern yellow wagtails in Europe: identification and vocalizations. Dutch Birding 35, 295–311.
Britten, R.J., 1986. Rates of DNA sequence evolution differ between taxonomic groups. Science 231, 1393–1398.
Brumfield, R.T., Beerli, P., Nickerson, D.A., Edwards, S.V., 2003. The utility of single nucleotide polymorphisms in inferences of population history. Trends Ecol. Evol. 18, 249–256.
Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N.A., RoyChoudhury, A., 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. Mol. Biol. Evol. 29, 1917–1932.
Bryson Jr, R.W., De Oca, A.N.M., Jaeger, J.R., Riddle, B.R., 2010. Elucidation of cryptic diversity in a widespread Nearctic treefrog reveals episodes of mitochondrial gene capture as frogs diversified across a dynamic landscape. Evolution 64, 2315–2330.
Campagna, L., Benites, P., Lougheed, S.C., Lijtmaer, D.A., Di Giacomo, A.S., Eaton, M.D., et al., 2011. Rapid phenotypic evolution during incipient speciation in a continental avian radiation. Proc. Biol. Sci. 279, 1847–1856.
Carling, M.D., Brumfield, R.T., 2008. Haldane's rule in an avian system: using cline theory and divergence population genetics to test for differential introgression of mitochondrial, autosomal, and sex-linked loci across the passerina bunting hybrid zone. Evolution 62, 2600–2615.
Catchen, J., Amores, A., Hohenlohe, P., Cresko, W., Postlethwait, J., 2011. Stacks: building and genotyping loci *de novo* from short-read sequences. G3 1, 171–182.
Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely

gene trees. PLOS Genet. 2, 762–768.
del Hoyo, J., Collar, N.J., 2016. HBW and BirdLife International Illustrated Checklist of the Birds of the World. Lynx Editions, Barcelona.
del Hoyo, J., Elliott, A., Sargatal, J., 2004. Handbook of the Birds of the World: Cotingas to Pipits and Wagtails.
Dickinson, E.C., Christidis, L. (Eds.), The Howard & Moore Complete Checklist of the Birds of the World, Passerines, vol. 2, forth ed. Aves Press, Eastbourne, UK.
Dong, F., Zou, F.S., Lei, F.M., Liang, W., Li, S.H., Yang, X.J., 2014. Testing hypotheses of mitochondrial gene-tree paraphyly: unravelling mitochondrial capture of the Streak-breasted Scimitar Babbler (*Pomatorhinus ruficollis*) by the Taiwan Scimitar Babbler (*Pomatorhinus musicus*). Mol. Ecol. 23, 5855–5867.
Dray, S., Dufour, A.-B., 2007. The ade4 package: implementing the duality diagram for ecologists. J. Stat. Softw. 22. http://dx.doi.org/10.18637/jss.v022.i04.
Dray, S., Josse, J., 2014. Principal component analysis with missing values: a comparative survey of methods. Plant Ecol. 216, 657–667.
Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.
Edwards, S.V., Liu, L., Pearl, D.K., 2007. High-resolution species trees without concatenation. Proc. Natl. Acad. Sci. USA 104, 5936–5941.
Fogelqvist, J., Niittyvuopio, A., Ågren, J., Savolainen, O., Lascoux, M., 2010. Cryptic population genetic structure: the number of inferred clusters depends on sample size. Mol. Ecol. Resour. 10, 314–323.
Funk, D.J., Omland, K.E., 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34, 397–423.
Gohli, J., Leder, E.H., Garcia-Del-Rey, E., Johannessen, L.E., Johnsen, A., Laskemoen, T., et al., 2015. The evolutionary history of Afrocanarian blue tits inferred from genome wide SNPs. Mol. Ecol. 24, 180–191.
Gompert, Z., Lucas, L.K., Alex Buerkle, C., Forister, M.L., Fordyce, J.A., Nice, C.C., 2014. Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. Mol. Ecol. 23, 4555–4573.
Guillot, G., Rousset, F., 2013. Dismantling the Mantel tests. Meth. Ecol. Evol. 4, 336–344.
Hailer, F., Kutschera, V.E., Hallström, B.M., Klassert, D., Fain, S.R., Leonard, J.A., Arnason, U., Janke, A., 2012. Nuclear genomic sequences reveal that polar bears are an old and distinct bear lineage. Science 336, 344–347.
Heled, J., Drummond, A.J., 2012. Calibrated tree priors for relaxed phylogenetics and divergence time estimation. Syst. Biol. 61, 138–149.
Huang, H., Knowles, L.L., 2014. Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. Syst. Biol. 65, 357–365.
Jánossy, D., 1991. Late Miocene bird remains from Polgárdi (W-Hungary). Aquila 98, 13–35.
Jombart, T., Devillard, S., Dufour, A.-B., Pontier, D., 2008. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. Heredity 101, 92–103.
Jombart, T., Devillard, S., Balloux, F., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 11, 94.
Kearns, A.M., Joseph, L., Toon, A., Cook, L.G., 2014. Australia's arid-adapted butcherbirds experienced range expansions during Pleistocene glacial maxima. Nat. Comm. 5, 3994.
Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56, 17–24.
Lamichhaney, S., Berglund, J., Sällman, A.M., Maqbool, K., Grabherr, M., Martinez-Barrio, A., Promerová, M., Rubin, C.J., Wang, C., Zamani, N., Grant, B.R., Grant, P.R., Webster, M.T., Andersson, L., 2015. Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature 18, 371–375.
Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nat. Meth. 9, 357–359.
Lavretsky, P., McCracken, K.G., Peters, J.L., 2014. Phylogenetics of a recent radiation in the mallards and allies (Aves: Anas): inferences from a genomic transect and the multispecies coalescent. Mol. Phylogenet. Evol. 70, 402–441.
Leaché, A.D., Harris, R.B., Rannala, B., Yang, Z., 2013. The influence of gene flow on species tree estimation: a simulation study. Syst. Biol. 63, 17–30.
Leaché, A.D., Banbury, B.L., Felsenstein, J., de Oca, A.N.-M., Stamatakis, A., 2015. Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. Syst. Biol. 64, 1032–1047.
Legendre, P., Fortin, M.-J., Borcard, D., 2015. Should the Mantel test be used in spatial analysis? Meth. Ecol. Evol. 6, 1239–1247.
Li, X., Dong, F., Lei, F., Alström, P., Zhang, R., Ödeen, A., et al., 2016. Shaped by uneven Pleistocene climate: mitochondrial phylogeographic pattern and population history of white wagtail *Motacilla alba* (Aves: Passeriformes). J. Avian Biol. 47, 263–274.
Lischer, H.E.L., Excoffier, L., Heckel, G., 2014. Ignoring heterozygous sites biases phylogenomic estimates of divergence times: implications for the evolutionary history of *Microtus* voles. Mol. Biol. Evol. 31, 817–831.
Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46, 523.
Mason, N.A., Taylor, S.A., 2015. Differentially expressed genes match bill morphology and plumage despite largely undifferentiated genomes in a Holarctic songbird. Mol. Ecol. 24, 3009–3025.
Masretta-Yanes, A., Arrigo, N., Alvarez, N., Jorgensen, T.H., Pinero, D., Emerson, T.D., 2014. Restriction site-associated DNA sequencing, genotyping error estimation and *de novo* assembly optimization for population genetic inference. Mol. Ecol. Res. 15, 28–41.
McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C., Brumfield, R.T., 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Mol. Phylogenet. Evol. 66, 526–538.
McKay, B.D., Zink, R.M., 2010. The causes of mitochondrial paraphyly. Mol. Phylogenet. Evol. 54, 647–650.
Milá, B., McCormack, J.E., Castañeda, G., Wayne, R.K., Smith, T.B., 2007. Recent post-glacial range expansion drives the rapid diversification of a songbird lineage in the genus *Junco*. Proc. Biol. Sci. 274, 2653–2660.
Moore, W.S., 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. Evolution 49, 718–726.
Nakagawa, S., Freckleton, R.P., 2008. Missing inaction: the dangers of ignoring missing data. Trends Ecol. Evol. 23, 592–596.
O'Connell, J., Schulz-Trieglaff, O., Carlson, E., Hims, M.M., Gormley, N.A., Cox, A.J., 2015. NxTrim: optimized trimming of Illumina mate pair reads. Bioinformatics 31, 2035–2037.
Ödeen, A., Alström, P., 2001. Evolution of secondary sexual traits in wagtails (genus *Motacilla*). In: Ödeen, A. (Ed.), Effects of Post-Glacial Range Expansions and Population Bottlenecks on Species Richness. Uppsala University (Ph.D. thesis).
Ödeen, A., Björklund, M., 2003. Dynamics in the evolution of sexual traits: losses and gains, radiation and convergence in yellow wagtails (*Motacilla flava*). Mol. Ecol. 12, 2113–2130.
Olsson, U., Alström, P., Svensson, L., Aliabadian, M., Sundberg, P., 2010. The *Lanius excubitor* (Aves, Passeriformes) conundrum—taxonomic dilemma when molecular and non-molecular data tell different stories. Mol. Phylogenet. Evol. 55, 347–357.
Omland, K.E., Lanyon, S.M., 2000. Reconstructing plumage evolution in orioles (*Icterus*): repeated convergence and reversal in patterns. Evolution 54, 2119–2133.
Pavlova, A., Zink, R.M., Drovetski, S.V., Red'kin, Y., Rohwer, S., 2003. Phylogeographic patterns in *Motacilla flava* and *Motacilla citreola*: species limits and population history. Auk 120, 744–758.
Pavlova, A., Zink, R.M., Rohwer, S., Koblik, E.A., Red'kin, Y.A., Fadeev, I.V., et al., 2005. Mitochondrial DNA and plumage evolution in the white wagtail *Motacilla alba*. J. Avian Biol. 36, 322–336.
Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., Hoekstra, H.E., 2012. Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. PLoS One 7, e37135.
Petkova, D., Novembre, J., Stephens, M., 2016. Visualizing spatial population structure with estimated effective migration surfaces. Nat. Genet. 48, 94–100.
Plummer, M., Best, N., Cowles, K., Vines, K., 2006. CODA: convergence diagnosis and output analysis for MCMC. R News 6, 7–11.
Poelstra, J.W., Vijay, N., Bossu, C.M., Lantz, H., Ryll, B., Müller, I., et al., 2014. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. Science 344, 1410–1414.
Poelstra, J.W., Vijay, N., Hoeppner, M.P., Wolf, J.B.W., 2015. Transcriptomics of colour patterning and coloration shifts in crows. Mol. Ecol. 24, 4617–4628.
Pons, J.M., Sonsthagen, S., Dove, C., Crochet, P.A., 2014. Extensive mitochondrial introgression in North American Great Black-backed Gulls (*Larus marinus*) from the American Herring Gull (*Larus smithsonianus*) with little nuclear DNA impact. Heredity 112, 226–239.
Price, T., 1998. Sexual selection and natural selection in bird speciation. Philos. Trans. R. Soc. Lond. B Biol. Sci. 353, 251–260.
Price, T., 2008. Speciation in Birds. Roberts & Company.
Price, T.D., Bouvier, M.M., 2002. The evolution of F1 postzygotic incompatibilities in birds. Evolution 56, 2083–2089.
Red'kin, Y.A., 2001. A new suggestion on taxonomical structure of "Yellow Wagtails" group. In: Kurochkin, E., Rakhimov, I. (Eds.), Achievements and Problems of Ornithology of Northern Eurasian on a Boundary of Centuries.
Rheindt, F.E., Edwards, S.V., 2011. Genetic introgression: an integral but neglected component of speciation in birds. Auk 128, 620–632.
Ritchie, A.M., Lo, N., Ho, S.Y.W., 2016. The impact of the tree prior on molecular dating of data sets containing a mixture of inter- and intraspecies sampling. Syst. Biol. 66, 413–425.
Rokas, A., Carroll, S.B., 2006. Bushes in the tree of life. PLoS Biol. 4, e352.
Royal, C.D., Novembre, J., Fullerton, S.M., Goldstein, D.B., Long, J.C., Bamshad, M.J., et al., 2010. Inferring genetic ancestry: opportunities, challenges, and implications. Am. J. Hum. Genet. 86, 661–673.
Seddon, N., Botero, C.A., Tobias, J.A., Dunn, P.O., Macgregor, H.E.A., Rubenstein, D.R., et al., 2013. Sexual selection accelerates signal evolution during speciation in birds. Proc. Biol. Sci. 280, 20131065.
Semenov, G.A., Scordato, E.S.C., Khaydarov, D.R., Smith, C.C.R., Kane, N.C., Safran, R., 2017. Effects of assortative mate choice on the genomic and morphological structure of a hybrid zone between two bird subspecies. Mol. Ecol. 26, 6430–6444.
Shipham, A., Schmidt, D.J., Joseph, L., Hughes, J.M., 2016. A genomic approach reinforces a hypothesis of mitochondrial capture in eastern Australian rosellas. Auk 134, 181–192.
Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313.
Tang, Q.Y., Liu, S.Q., Yu, D., Liu, H.Z., Danley, P.D., 2012. Mitochondrial capture and incomplete lineage sorting in the diversification of balitorine loaches (Cypriniformes, Balitoridae) revealed by mitochondrial and nuclear genes. Zool. Scripta. 41, 233–247.
Toews, D.P.L., Brelsford, A., 2012. The biogeography of mitochondrial and nuclear discordance in animals. Mol. Ecol. 21, 3907–3930.
Toews, D.P.L., Taylor, S.A., Vallender, R., Brelsford, A., Butcher, B.G., Messer, P.W., et al., 2016. Plumage genes and little else distinguish the genomes of hybridizing warblers. Curr. Biol. 26, 2313–2318.
Vijay, N., Bossu, C.M., Poelstra, J.W., Weissensteiner, M.H., Suh, A., Kryukov, A.P., et al., 2016. Evolution of heterogeneous genome differentiation across multiple contact zones in a crow species complex. Nat. Commun. 7, 13195.
Voelker, G., 2002. Systematics and historical biogeography of wagtails: dispersal versus vicariance revisited. Condor 104, 725.

Wagner, C.E., Keller, I., Wittwer, S., Selz, O.M., Mwaiko, S., Greuter, L., et al., 2013. Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. Mol. Ecol. 22, 787–798.

Wang, W., Dai, C., Alström, P., Zhang, C., Qu, Y., Li, S.H., Yang, X., Zhao, N., Song, G., Lei, F., 2014. Past hybridization between two East Asian long-tailed tits (*Aegithalos bonvaloti* and *A. fuliginosus*). Front. Zool. 11, 40.

Waples, R.S., Gaggiotti, O., 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. Mol. Ecol. 15, 1419–1439.

Wright, S., 1943. Isolation by distance. Genetics 28, 114–138.

Zamudio, K.R., Bell, R.C., Mason, N.A., 2016. Phenotypes in phylogeography: Species' traits, environmental variation, and vertebrate diversification. Proc. Natl. Acad. Sci. USA. 113, 8041–8048.