JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.13305

Extensive phenotypic diversification coexists with little genetic divergence and a lack of population structure in the White Wagtail subspecies complex (*Motacilla alba*)

GEORGY A. SEMENOV*†‡, EVGENIY A. KOBLIK§, YAROSLAV A. RED'KIN§ & ALEXANDER V. BADYAEV†

*Department of Ecology & Evolutionary Biology, University of Colorado, Boulder, CO, USA
†Department of Ecology & Evolutionary Biology, University of Arizona, Tucson, AZ, USA
‡Institute of Systematics and Ecology of Animals, Novosibirsk, Russia
§Department of Ornithology, Zoological Museum of Moscow State University, Moscow, Russia

Keywords:

mito-nuclear discordance;
phenotypic diversification;
plumage evolution;
wagtail.

Abstract

Geographically clustered phenotypes often demonstrate consistent patterns in molecular markers, particularly mitochondrial DNA (mtDNA) traditionally used in phylogeographic studies. However, distinct evolutionary trajectories among traits and markers can lead to their discordance. First, geographic structure in phenotypic traits and nuclear molecular markers can be co-aligned but inconsistent with mtDNA (mito-nuclear discordance). Alternatively, phenotypic variation can have little to do with patterns in neither mtDNA nor nuclear markers. Disentangling between these distinct patterns can provide insight into the role of selection, demography and gene flow in population divergence. Here, we examined a previously reported case of strong inconsistency between geographic structure in mtDNA and plumage traits in a widespread polytypic bird species, the White Wagtail (*Motacilla alba*). We tested whether this pattern is due to mito-nuclear discordance or discrepancy between morphological evolution and both nuclear and mtDNA markers. We analysed population differentiation and structure across six out of nine commonly recognized subspecies using 17 microsatellite loci and a combination of microsatellites and plumage indices in a comprehensively sampled region of a contact between two subspecies. We did not find support for the mito-nuclear discordance hypothesis: nuclear markers indicated a subtle signal of genetic clustering only partially consistent with plumage groups, similar to previous findings that relied on mtDNA. We discuss evolutionary factors that could have shaped the intricate patterns of phenotypic diversification in the White wagtail and the role that repeated selection on plumage 'hotspots' and hybridization may have played.

Introduction

Understanding the processes promoting and maintaining population differentiation is key for understanding speciation (Coyne & Orr, 2004; Butlin *et al.*, 2012). In birds, population differentiation is commonly associated with divergence in plumage patterns and coloration – phenotypic traits targeted by social and sexual selection that play an important role in reproductive barriers (Price, 2008; Seddon *et al.*, 2013). Marked phenotypic differentiation between populations, therefore, generally indicates divergent evolutionary histories, and is often reflected in geographic structuring of genetic markers (Avise, 2000), particularly in mitochondrial DNA (mtDNA) traditionally used in avian phylogeographic studies (Zink & Barrowclough, 2008).

Several mechanisms, however, can cause variation in genetic markers and phenotypes to be incongruent (Zamudio *et al.*, 2016). Differences in contemporary or

Correspondence: Georgy A. Semenov, EBIO, University of Colorado, 334 UCB, Boulder, CO 80309, USA.
Tel.: +1-303-735-1495; fax: +1-303-492-8699;
e-mail: georgy.semenov@colorado.edu

historical gene flow between mitochondrial and nuclear genomes can lead to concordant geographic structure of nuclear DNA markers (nuDNA) and phenotypes, but discordance of these features with mtDNA. This is known as mito-nuclear discordance (Funk & Omland, 2003; Chan & Levin, 2005; Currat *et al.*, 2008; Irwin *et al.*, 2009; Toews & Brelsford, 2012; Toews *et al.*, 2014; Good *et al.*, 2015). Alternatively, phenotypic divergence can be discordant with variation in multiple types of molecular markers. For example, selection can promote rapid divergence in a few loci underlying phenotypes with key functional roles in local adaptation or reproductive isolation (i.e. speciation genes; Nosil & Schluter, 2011). Early in the speciation process, other parts of the genome can undergo unrestricted admixture, or remain relatively undifferentiated as a result of incomplete lineage sorting (Barton & Bengtsson, 1986; Harrison, 1990; Wu, 2001; Feder *et al.*, 2012; Nosil & Feder, 2012; Harrison & Larson, 2014, 2016; Seehausen *et al.*, 2014). In these scenarios, phenotypic differences can be inconsistent with nearly all molecular markers throughout the genome (e.g. Campagna *et al.*, 2012; Mason & Taylor, 2015). In addition, asymmetric introgression of alleles underlying phenotypic signals (e.g. Brumfield *et al.*, 2001; Baldassarre *et al.*, 2014), or hybrid zone movement (e.g. Krosby & Rohwer, 2009) can promote phenotypic and genetic discordance. Unravelling the factors driving conflicting patterns between mtDNA and phenotypic variation can provide insights into the role of demography, selection, gene flow and dispersal in divergent evolution (reviewed in Toews & Brelsford, 2012) and requires the use of nuclear genetic markers to disentangle potential mechanisms (Edwards & Bensch, 2009).

The White Wagtail (*Motacilla alba,* Linnaeus, 1758) represents an intriguing example of discordance between geographic patterns of phenotypes and mtDNA. Nine subspecies of White Wagtail have distinct plumage across the species' range (Fig. 1). Evidence for the predominance of parental plumage phenotypes over intermediates in narrow hybrid zones suggests that there might be reproductive isolation, although incomplete, between some subspecies (*lugens* vs *ocularis* and *leucopsis*; *personata* vs *alba* and *baicalensis*) (Nazarenko, 1968; Stepanyan, 1983, 2003; Koblik *et al.*, 2001; Red'kin, 2003; Semenov & Yurlov, 2010; Semenov *et al.*, 2010; Lobkov, 2011). A certain degree of population structure and subdivision was therefore expected within the White wagtail subspecies complex. Previous analyses of mtDNA markers revealed noticeable population structure, but it was globally incongruent with the geographic distribution of subspecies (Pavlova *et al.*, 2005; Li *et al.*, 2016). Four divergent mtDNA lineages have been found to date, corresponding to (1) northern Eurasia, (2) North Africa, (3) central and western parts of Eurasia (British Isles, Caucasus, Iran and Uzbekistan) and (4) central to southeastern Eurasia. Three of these clades were associated with more than one subspecies, with the fourth North African clade endemic to *subpersonata* (Li *et al.*, 2016). Individual subspecies contained haplotypes belonging to two (*yarrellii*, *alba*, *subpersonata*, *lugens* and *leucopsis*; Li *et al.*, 2016; Pavlova *et al.*, 2005) or even three (*personata*; Semenov *et al.*, 2017) distinct mtDNA lineages. There was little coincidence between the geographic areas where mtDNA lineages and phenotypic clusters came into contact. Two hypotheses were suggested for this striking conflict between traditional taxonomy and molecular markers (Pavlova *et al.*, 2005; Li *et al.*, 2016). First, mtDNA markers may be inconsistent with population history as a result of mito-nuclear discordance. Alternatively, morphological patterns may be poorly associated with those observed in both nuclear and mitochondrial genetic markers, for example due to rapid plumage evolution accompanied by incomplete lineage sorting.

A recent study of genomic and morphological variation in the hybrid zone between *alba* and *personata* subspecies in Siberia provided support for the latter hypothesis (Semenov *et al.*, 2017). There was very shallow genomic divergence between subspecies based on ≈19 500 SNP loci despite divergence in several plumage and morphometric traits. Spatial transition in genomic ancestry was correlated with variation in mtDNA markers (analysed in Pavlova *et al.*, 2005): northern *personata* from the Altai region had genomic ancestry and mtDNA haplotypes grouping them with *alba*, whereas southern *personata* were distinct from northern *personata* in both genomic ancestry and mtDNA. There was therefore no apparent evidence for mito-nuclear discordance. Analysis of geographic clines revealed complex introgression patterns among morphological traits and genomic markers. The cline for coloration of head-and-neck plumage – a trait associated with positive assortative mating – was shifted about 100–300 km from the clines observed in other plumage traits (wing and back coloration), body size and genomic ancestry, suggesting asymmetric introgression of alleles underlying head plumage or hybrid zone movement. Together, available evidence suggests that variation in plumage traits is driven or maintained by processes that have little effect on patterns in both nuclear and mtDNA markers.

However, up until now only one pair of subspecies was studied in detail. Variation in nuclear genetic markers in other White wagtail subspecies has been examined for only a few individuals and relied on a single Z-linked marker (Alström & Ödeen, 2002). It is therefore unknown whether patterns in nuclear genetic markers are globally consistent with phenotype, mtDNA or neither. In addition, *alba* and *personata* subspecies have a broad zone of contact in Siberia, and comprehensive sampling of this region will help us to better understand if the lack of genetic differences and discordant introgression patterns between plumage and genetic markers span the entire range of the subspecies

co-distribution, or if it is restricted to the western Altai region studied in Semenov *et al.* (2017).

In the present study, we assessed population divergence and structure in 17 microsatellite nuclear markers across the distributional range of six subspecies (*alba, personata, baicalensis, ocularis, lugens* and *leucopsis*) in northern Eurasia, where previous studies revealed a lack of mtDNA population structure. In addition, we used microsatellite markers and plumage indices from a densely sampled area of *alba* and *personata* co-occurrence. Presence of population structure in the nuclear markers corresponding to the distribution of subspecies would support the mito-nuclear discordance hypothesis. Alternatively, a lack of population structure and genetic divergence would indicate that the processes affecting plumage evolution in the White Wagtail are substantially distinct from those shaping neutral genetic evolution.

Materials and methods

Field sampling

We used samples from six White wagtail subspecies collected across northern Eurasia during 2003–2011: *alba* (including populations east of the Ural Mountains, sometimes referred to as the *dukhunensis* subspecies; Dement'ev *et al.*, 1954), *personata*, *baicalensis*, *leucopsis*, *lugens* and *ocularis* (Table 1, Fig. 1). There are three other commonly recognized subspecies (*yarrellii*, *subpersonata* and *alboides*) that were not represented in our study. Sample collection was performed as part of several long-term research programmes aimed at describing the biological diversity of northern Eurasia carried out by the Zoological Museum of Moscow State University (ZMMSU Moscow, Russia), the Siberian Zoological Museum (SZM, Novosibirsk, Russia) and the Kirov City Zoological Museum (KCZM, Kirov, Russia). Only locally breeding adults or recently fledged juveniles were used in the analyses. Birds were sexed by gonad examination. Samples of breast muscle were stored in 96% ethanol prior to laboratory analyses. Subspecies were identified in the field based on the authors' expertise or using associated museum skins. The geographic distribution of each subspecies (Fig. 1) was determined using data associated with 561 specimens from collections of ZMMSU, Zoological Institute of the Russian Academy of Sciences (Saint Petersburg, Russia), the SZM, and the literature (Zalesskii, 1927; Sushkin, 1938; Dement'ev *et al.*, 1954; Paludan, 1959; Cheng, 1976; Glutz Von Blotzheim & Bauer, 1985; Ilyashenko, 1986; Cramp, 1988; Roberts, 1992; Badyaev *et al.*, 1996; Romanov, 1996; Alström & Mild, 2003; Red'kin, 2003; Stepanyan, 2003; Tsvetkov *et al.*, 2003; Berezovikov *et al.*, 2007; Porter & Aspinall, 2010; Semenov *et al.*, 2010; Lobkov, 2011; Aye *et al.*, 2012; Berezovikov & Reznichenko, 2014; Ryabitsev *et al.*, 2014).

Laboratory procedures

Total genomic DNA was extracted using the JETQUICK Tissue DNA Spin Kit (Genomed, Loöhne, Germany). A set of 289 microsatellite primer pairs was designed using high-throughput DNA sequencing in GENOSCREEN (http://www.genoscreen.fr/en/). We tested 48 primer pairs using a fluorescently labelled universal primer and a modified locus-specific primer with a 5′ universal primer sequence tail (Blacket *et al.*, 2012). In all, 18 polymorphic loci showing successful amplification were combined into three multiplexes and used for genotyping 299 individuals. See Appendix S1: Table S1 for details on primer sequences, multiplex composition and PCR conditions. All amplifications were carried out using the Multiplex PCR Kit (QIAGEN), PCR products were visualized using an Applied Biosystems 3130 Genetic Analyzer and alleles were scored using GeneMapper v.4.1 (Applied Biosystems, Foster City, CA, USA).

Analysis of microsatellite markers

Correspondence of loci to Hardy–Weinberg equilibrium (HWE, probability test) and linkage equilibrium (LE, G-test) was assessed using GENEPOP On The Web (http://genepop.curtin.edu.au; Raymond & Rousset, 1995). We tested for HWE and LE in three putative populations (groups of geographically clustered samples within the same phenotypic subspecies) with the most representative sample sizes: *alba* from Europe, *alba* from Siberia and *personata* from Siberia (see Table 1 for details). Deviations from HWE and LE were considered significant if present in more than one sample group at $P < 0.01$. Only locus Mot21 deviated from HWE in all three groups and was excluded from further analyses. Three microsatellite loci (Mot5, Mot13, Mot23) were homozygous in all females but heterozygous in many males, suggesting that they may be located on the Z-chromosome. Indeed, sequences of these loci showed high matching scores (110–214) and identity (81.7–95.8%) with regions on the Z-chromosome of the Zebra finch (*Taeniopygia guttata*) genome (http://genome-euro.ucsc.edu). Individuals with > 12% missing data were excluded, resulting in a final dataset of 17 unlinked, presumably selectively neutral microsatellite loci and $n = 250$ individuals (Table 1).

Test of population structure and isolation by distance

We performed Bayesian clustering analysis implemented in STRUCTURE 2.3.4 (Pritchard *et al.*, 2000) and evaluated the number of clusters ranging from 1 to 10 with $n = 10$ replicates per K. We set Markov Chain Monte Carlo (MCMC) to a burn-in period of 1 000 000 followed by 2 000 000 saved iterations. We used the admixture model with correlated allele frequencies and no prior information about sampling location.

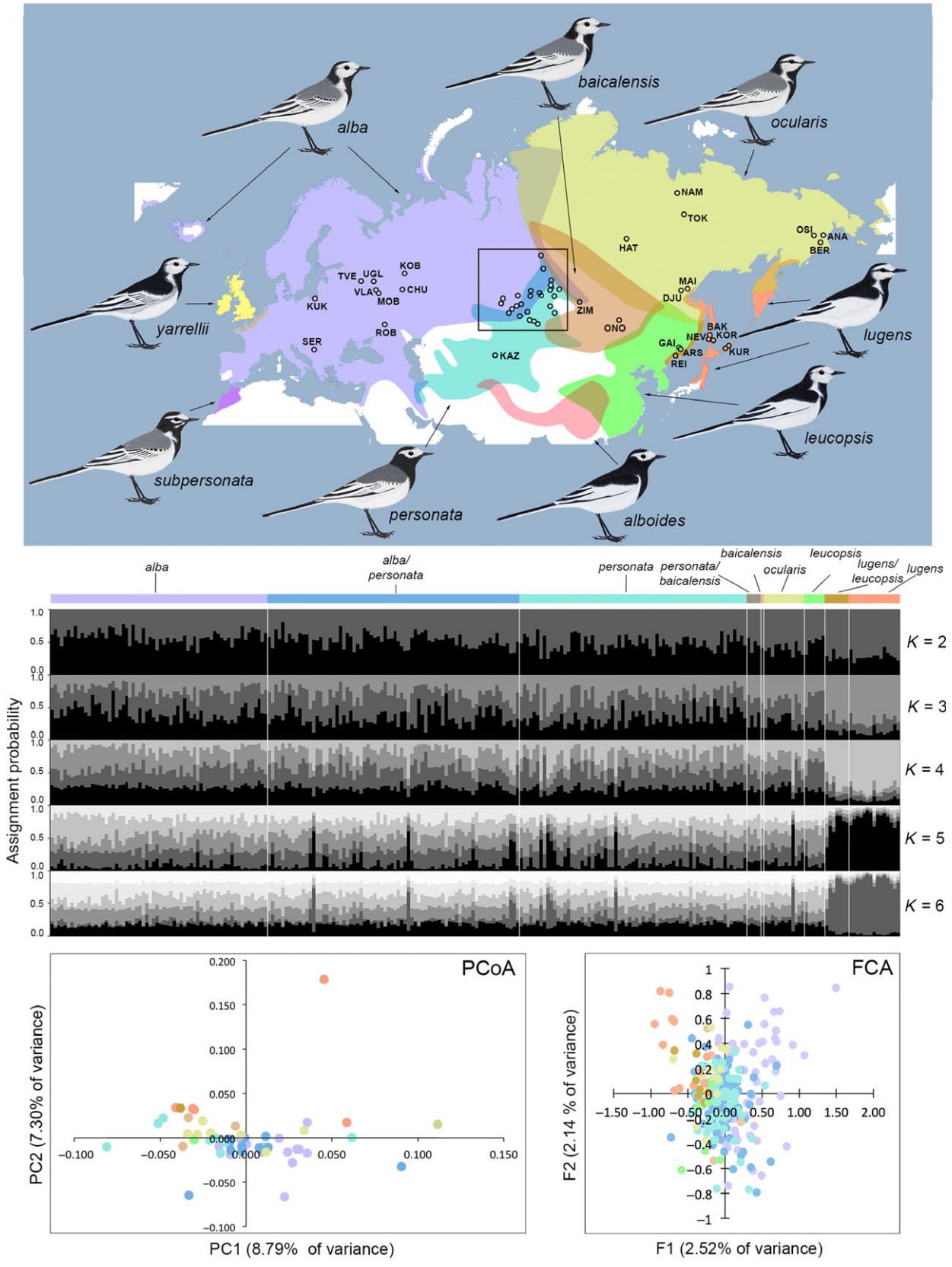
alba
baicalensis
ocularis
yarrellii
lugens
leucopsis
subpersonata
personata
alboides
NAM
TOK
OSI
ANA
BER
HAT
MAI
DJU
ZIM
ONO
BAK
KOR
NEV
GAI
ARS
KUR
REI
KAZ
TVE
UGL
KOB
VLA
CHU
MOB
KUK
ROB
SER
alba
alba/ personata
personata
personata/ baicalensis
baicalensis
ocularis
leucopsis
lugens/ leucopsis
lugens
Assignment probability
K = 2
K = 3
K = 4
K = 5
K = 6
PCoA
PC2 (7.30% of variance)
PC1 (8.79% of variance)
FCA
F2 (2.14 % of variance)
F1 (2.52% of variance)

Fig. 1 Breeding distribution of the White wagtail subspecies and results of population clustering analyses. Top: circles and abbreviations indicate sampling localities and locality codes (see Table 1 for details). Locality codes for the area within rectangle are shown in Figure 3. Middle: STRUCTURE assignment probability plots for the number of clusters, *K*, ranging from 2 to 6 ($n = 250$, 17 microsatellite loci). Each vertical bar represents an individual. Shading shows the probability proportion accounted for by each cluster. The most likely number of clusters identified by the Delta-*K* method was two. Bottom: results of principal coordinates (PCoA, 54 sampling localities, F_{st}) and factorial correspondence analyses (FCA, $n = 248$) based on 17 microsatellite loci. Colour scheme above the STRUCTURE and colour codes of PCoA and FCA correspond to the map. Note that the samples from allopatric *lugens*, *lugens-leucopsis* contact zone, several individuals of *ocularis, personata* and from *alba-personata* contact zone tend to be assigned to the same genetic cluster by STRUCTURE (more noticeable at $K = 5$–6). There is little evidence of population structure across other individuals and sampling localities. Gradual change in STRUCTURE ancestry probabilities is consistent with the pattern of isolation-by-distance supported by Mantel test ($Z = 3153.86$, $R = 0.241$, $P < 0.001$). None of the subspecies appear to constitute distinct genetic clusters based on PCoA and FCA, although differentiation along F1 is consistent with weak differences between western and eastern samples.

Because three loci (Mot5, Mot13, Mot23) were likely located on the Z-chromosome, we accounted for phase information in the STRUCTURE model when simultaneously analysing autosomal and Z-linked loci. To test whether the effects of ploidy affected clustering output, we ran the analysis using two additional datasets. First, since male birds have the same ploidy for sex chromosomes and autosomes, we ran a subset of males with all 17 loci ($n = 168$). Second, we analysed a dataset of both males and females ($n = 250$) with only 14 autosomal loci included.

Uneven sampling of populations can bias results of population assignment in STRUCTURE, with distinct populations represented by fewer individuals tending to cluster with more abundant groups, and individuals from more extensively sampled panmictic populations being artificially split (Puechmaille, 2016). Samples of *alba* and *personata* subspecies and their hybrids were over-represented in our dataset (Table 1) and were also sampled over a broader geographic range. We therefore ran three additional datasets, which subsampled five subspecies relatively evenly (*baicalensis* was excluded due to low sample size). Two datasets used only males and both autosomal and sex-linked loci. The first included 41 males (ten *alba*, five from western and five from eastern parts of the range; ten *personata,* five from northern and five from southern parts of the range; nine *ocularis*; eight *lugens*; four *leucopsis*). The second included 61 males (20 *alba*, ten western and ten eastern; 20 *personata,* ten northern and ten southern; nine *ocularis*; eight *lugens*; four *leucopsis*). Third, we used 55 individuals (males and females) and 14 autosomal loci (12 *alba*; 12 *personata;* 12 *ocularis*; 12 *lugens*; seven *leucopsis*). See Appendix S2: Figs S1–S5 for details on sample selection. We evaluated the most likely number of clusters following the method of (Evanno *et al.*, 2005) using STRUCTURE HARVESTER (Earl & vonHoldt, 2012). Results obtained in repeated runs were summarized by CLUMPP (Jakobsson & Rosenberg, 2007). We examined results for all tested *K* values because hierarchical population structure may bias the results of clustering analysis (Meirmans, 2015).

We further explored patterns of population divergence and structure using two multivariate methods without a pre-defined number of clusters. We ran a principle coordinates analysis (PCoA) based on F_{st} between 54 sampling localities using GENALEX (Peakall & Smouse, 2012) and a factorial correspondence analysis (FCA) for 250 individuals using GENETIX (Belkhir *et al.*, 1998). Our initial FCA revealed that most of the variance along the first two axes was due to several outliers (i.e. individuals carrying rare alleles), obscuring patterns in the majority of the dataset (not shown). We therefore removed two individuals and alleles with global frequencies below 0.01 for consecutive FCA. Finally, we used the neighbour-joining method (Saitou & Nei, 1987) to reconstruct hierarchical genetic relationships between 54 samples using pairwise F_{st} in MEGA7 (Kumar *et al.*, 2016).

We tested for the effects of isolation by distance across the six subspecies using the Mantel test (Mantel, 1967) implemented in Isolation By Distance Web Service (Jensen *et al.*, 2005). We calculated fixation indices in GENALEX (Peakall & Smouse, 2012) and used matrices of transformed genetic distances ($F_{st}/(1 - F_{st})$) and natural logarithms of Euclidean distances (Rousset, 1997). The analysis was carried out with 10 000 permutations.

We determined how genetic differentiation is partitioned among the six subspecies with an AMOVA (Excoffier *et al.*, 1992) implemented in GENALEX (Peakall & Smouse, 2012) using three hierarchical levels: among subspecies, among populations (samples) and within populations (samples). For the AMOVA, we excluded samples with less than four individuals and combined geographically proximate samples with small sample sizes (Table 1). We excluded samples originating from hybrid zones to minimize potential effects of admixture on variance patterns. Statistical significance of pairwise F_{st} was evaluated using 999 permutations.

As initial analyses did not recover population structure consistent with subspecies delineation, we performed further testing using a suite of methods sensitive to a very weak clustering signal. In these analyses, we focused on two rigorously sampled geographic regions: the *alba-personata* contact zone in Central Siberia ($n = 153$) and *ocularis–lugens–leucopsis* distributional ranges in East Siberia and Russian Far East

Table 1 Sampling information. Locality codes correspond to Figs 1 and 3.

Region	Locality	Locality code	Subspecies	Latitude	Longitude	AMOVA grouping	N genetic	N plumage
Europe	Kurshskaya kosa	KUK	*alba*	54.53	20.16	KUK	13	0
Europe	Serbia	SER	*alba*	44.26	19.89	Excluded	2	0
Europe	Tverskaya oblast	TVE	*alba*	57.54	34.41	TVE	6	0
Europe	Uglich	UGL	*alba*	57.50	38.34	Western Russia (WRU)	1	0
Europe	Vladimirskaya oblast	VLA	*alba*	56.06	39.01	Western Russia (WRU)	2	0
Europe	Moscovskaya oblast	MOB	*alba*	55.48	39.86	Western Russia (WRU)	2	0
Europe	Rostovskaya oblast	ROB	*alba*	49.58	41.75	Excluded	1	0
Europe	Kirovskaya oblast	KOB	*alba*	58.84	47.89	West of Ural (WOU)	3	0
Europe	Chuvashia	CHU	*alba*	56.15	47.19	West of Ural (WOU)	1	0
Siberia	Astrodym	AST	*alba*	53.64	77.76	AST	5	5
Siberia	Tchany	TCH	*alba*	54.56	78.33	TCH	15	19
Siberia	Gor'koe	GOR	*alba*	52.60	81.14	Altaiskii krai (AKR)	2	1
Siberia	Klochki	KLO	*alba*	53.18	82.58	Altaiskii krai (AKR)	2	2
Siberia	Novosibirsk	NOV	*alba*	55.05	82.92	NOV	9	10
Central Asia	Kazahstan	KAZ	*personata*	42.00	75.89	KAZ	4	0
Siberia	Yaloman	YAL	*personata*	50.52	86.55	Central Altai (CAL)	4	4
Siberia	Kuraiskaya steppe	KUS	*personata*	50.23	87.94	Central Altai (CAL)	1	1
Siberia	Ulandryk	ULA	*personata*	49.69	89.11	ULA	7	7
Siberia	Tanzybei	TAN	*personata*	53.14	92.92	West and East Sayan, WES	1	1
Siberia	Seserlig	SES	*personata*	51.86	94.33	West and East Sayan, WES	2	3
Siberia	Kuturchin	KUT	*personata*	54.94	94.22	West and East Sayan, WES	2	2
Siberia	Krasnoyarsk	KRA	*personata*	56.05	92.93	KRA	21	31
Siberia	Suhobuz	SUH	*personata*	56.50	93.29	Krasnoyarsk forest-steppe (KFS)	4	3
Siberia	Murta	MUR	*personata*	56.91	93.15	Krasnoyarsk forest-steppe (KFS)	13	7
Siberia	Rybinskoe	RYB	*personata*	55.76	94.80	Excluded	2	2
Siberia	Kansk	KAN	*personata*	56.21	95.73	Excluded	1	1
Siberia	Sayan	SAY	*personata*	52.76	93.35	Excluded	0	2
Siberia	Bor-Forpost	BFP	*alba/personata*	51.85	80.13	Excluded	6	6
Siberia	Zudilovo	ZUD	*alba/personata*	53.53	83.97	Excluded	5	3
Siberia	Maiorka	MIO	*alba/personata*	51.26	83.52	Excluded	3	3
Siberia	Platovo	PLA	*alba/personata*	52.08	85.90	Excluded	6	6
Siberia	Krapivinskii	KPI	*alba/personata*	54.97	86.78	Excluded	12	10
Siberia	Krasnii Yar	KRY	*alba/personata*	55.90	86.94	Excluded	4	4
Siberia	Kornilovo	KNV	*alba/personata*	55.54	89.64	Excluded	3	3
Siberia	Kongarovo	KGR	*alba/personata*	55.10	90.22	Excluded	5	15
Siberia	Kazatchinskoe	KZH	*alba/personata*	57.70	93.26	Excluded	14	23
Siberia	Novonazimovo	NNZ	*alba/personata*	59.56	90.81	Excluded	1	37
Siberia	Bor	BOR	*alba/personata*	61.59	90.10	Excluded	20	19
Siberia	Novyi Gorodok	NGG	*alba/personata*	59.83	90.18	Excluded	0	28
Siberia	Ust-Pit	PIT	*alba/personata*	58.99	91.75	Excluded	0	33
East Siberia	Zima	ZIM	*baicalensis*	53.92	102.05	East Siberia (ESI)	3	0
East Siberia	Onon	ONO	*baicalensis*	50.48	114.27	East Siberia (ESI)	1	0
East Siberia	Haty	HAT	*ocularis*	63.94	116.54	Yakutia and Chukotka (YCH)	1	0
East Siberia	Tokuma	TOK	*ocularis*	67.09	134.35	Yakutia and Chukotka (YCH)	1	0
East Siberia	Namy	NAM	*ocularis*	69.52	132.23	Yakutia and Chukotka (YCH)	1	0
East Siberia	Beringovskiy raion	BER	*ocularis*	63.45	176.57	Yakutia and Chukotka (YCH)	1	0
East Siberia	Osinovaya river	OSI	*ocularis*	64.38	174.55	Yakutia and Chukotka (YCH)	2	0
East Siberia	Anadyr	ANA	*ocularis*	64.43	177.45	Yakutia and Chukotka (YCH)	1	0
Far East	Maimakan	MAI	*ocularis*	56.28	135.49	Djugdjur mountains (DJM)	2	0
Far East	Djugdjur	DJU	*ocularis*	55.98	133.44	Djugdjur mountains (DJM)	3	0
Far East	Baklanovka	BAK	*lugens*	47.50	142.29	Sakhalin and Kuril islands (SKU)	2	0
Far East	Korsakovskiy raion	KOR	*lugens*	46.29	143.45	Sakhalin and Kuril islands (SKU)	8	0
Far East	Nevel'skiy raion	NEV	*lugens*	46.52	141.98	Sakhalin and Kuril islands (SKU)	1	0
Far East	Vodopadnaia bay	KUR	*lugens*	44.56	147.16	Sakhalin and Kuril islands (SKU)	4	0
Far East	Reinike island	REI	*lugens/leucopsis*	42.91	131.74	excluded	7	0
Far East	Arsen'ev	ARS	*leucopsis*	44.36	133.31	Primor'e (PRI)	1	0
Far East	Gaivoron	GAI	*leucopsis*	44.75	132.78	Primor'e (PRI)	6	0

(n = 40). We excluded alleles with global frequency below 0.01 (lumped with the nearest bin) and removed all individuals with missing data to maximize the clustering signal. We used the spatial model in GENELAND 4.0 (Guillot *et al.*, 2005) that accounts for the spatial distribution of genetic variation in population clustering. We ran GENELAND with 500 000 MCMC iterations, every 100th iteration saved, K ranging from one to six, ten replicates per K, DELTA = 2 and null allele model = TRUE. The first 2500 iterations were discarded as burn-in after visual examination of the model fit. As a complementary approach, we performed STRUCTURE analyses using default and LOCPRIOR models. The latter accounts for user-defined sampling information in the clustering process and is more sensitive to a weak clustering signal than the default model (Hubisz *et al.*, 2009). For the LOCPRIOR run, we divided individuals into groups corresponding to their phenotypic and geographic origin. For the *ocularis-lugens-leucopsis* dataset, we grouped individuals into allopatric *ocularis* (n = 12), allopatric *lugens* (n = 14), coastal hybrid zone between *lugens* and *leucopsis* (n = 7) and allopatric *leucopsis* (n = 7). For the *alba* and *personata* dataset, we assigned three groups: *alba* and *personata* hybrid zone (n = 69), allopatric *alba* (n = 30) and allopatric *personata* (n = 54). (Fig. 2, Table 1). We ran STRUCTURE with the settings described above for K = 1–5 and five replicates per K. We also performed a separate FCA for the *ocularis-lugens-leucopsis* and *alba-personata* datasets.

To assess the distribution of plumage clusters in *alba* and *personata*, we estimated a plumage index for the head and neck (Fig. 3) – these patches likely play a key role in conspecific recognition in these subspecies (Semenov *et al.*, 2017). We used specimens from the Zoological Museum of Moscow State University (Moscow, Russia, n = 35) and the Zoological Institute of the Russian Academy of Sciences (Saint Petersburg, Russia, n = 109) to evaluate variation in allopatric *personata* and *alba* and to develop classification rules for plumage scoring. Using digital photographs of study skins (n = 291, Table 1), we assigned each specimen an index where zero corresponded to *personata*, one to *alba* and four intermediate values to hybrids (Fig. 3). A single observer (GAS) performed phenotypic scoring. Because we did not have museum skin photographs available for many individuals of *ocularis-lugens-leucopsis* needed to perform quantitative estimation of the plumage index, we used literature records on subspecies' distribution and museum specimen data to assess plumage borders.

Results

Population differentiation across six White wagtail subspecies

The six subspecies we examined had little population structure and weak divergence in nuclear genetic markers. The most likely number of genetic clusters identified by STRUCTURE in the complete dataset (n = 250, 17 loci) was two (Delta K = 5.56), and individuals from opposite sides of Eurasia had only slightly different assignment probability scores (Fig. 1). At K = 3–6, samples of *lugens* from Kuril and Sakhalin islands (samples BAK, KOR, NEV, KUR) and the contact zone between *lugens* and *leucopsis* in Primor'e (REI) were differentiated, with the greatest differentiation at K = 6. There were also several individuals of phenotypic *ocularis*, *personata* and from the *alba-personata* contact zone that had a high probability of clustering with *lugens* and the *lugens-leucopsis* contact zone at K = 5–6 (Fig. 1). Higher values of K (7–10) did not reveal a noticeable increase

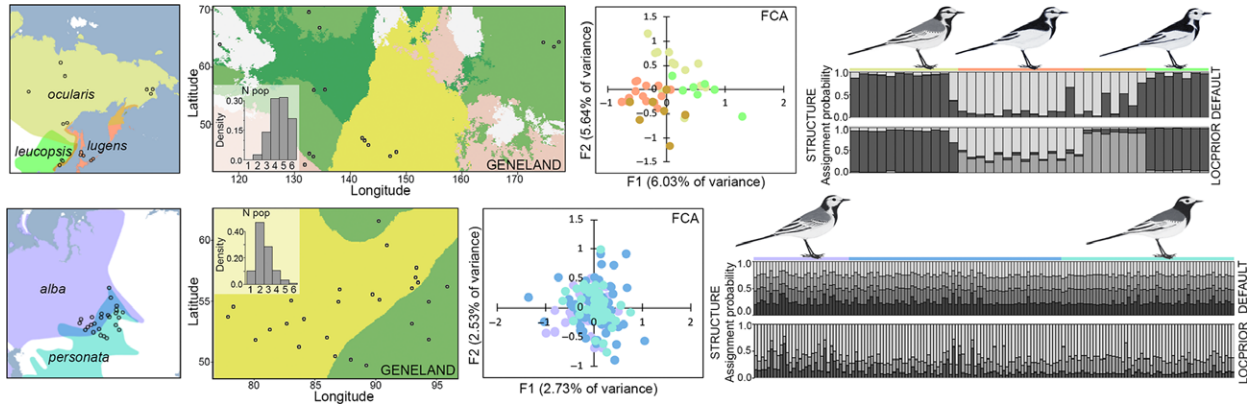

Fig. 2 Population clustering across *ocularis*, *lugens* and *leucopsis* (above) and *alba* and *personata* (below) subspecies. Left to right: subspecies distribution and sampling localities; GENELAND map of estimated population membership for the most likely number of genetic clusters (insertion shows the model density for the number of clusters 1–6), and the dots are sampling localities; results of factorial correspondence analysis (FCA); STRUCTURE assignment probability plots for the default and LOCPRIOR models. Colour scheme of FCA and panels above STRUCTURE correspond to the distributional map. Colour schemes of GENELAND are arbitrary.

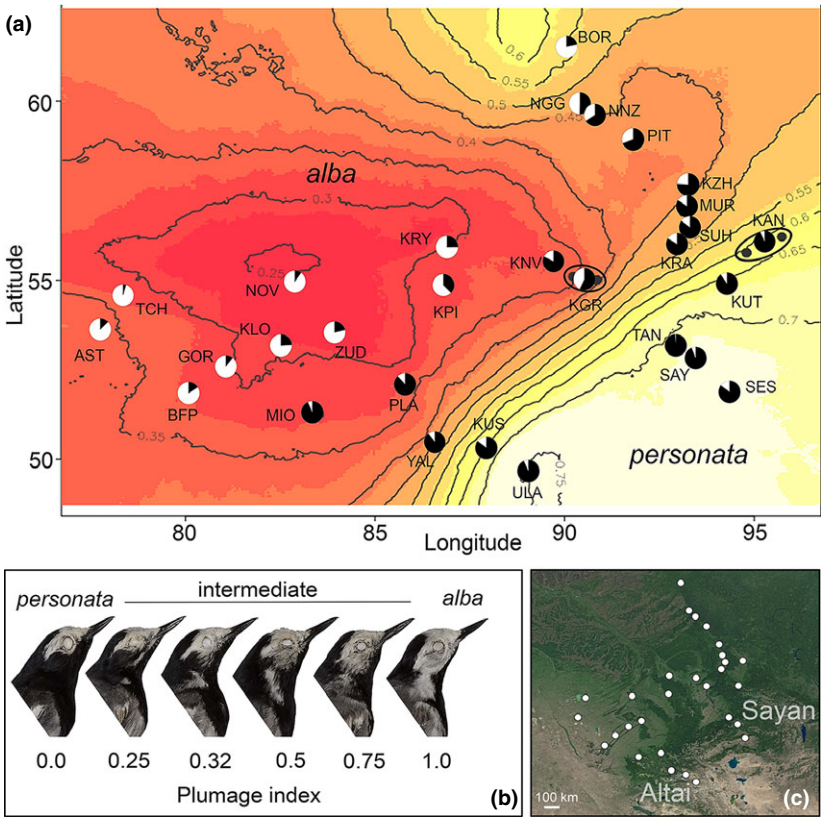

Fig. 3 (a) GENELAND map of population membership probability for the most likely $K = 2$ of *alba* and *personata* contact zone. Pie charts show the average plumage index, and the abbreviations represent the locality codes. (b) Head and neck plumage variation in *personata, alba* and intermediates, and corresponding plumage indices. (c) Landscape features of the sampling area (Map data: Google, DigitalGlobe). Note that the spatial pattern of population clustering is only roughly consistent with the distribution of plumage indices, whereas the majority of clustering signal appears to be driven by *personata* samples from high-altitudinal regions of Altai and Sayan mountains.

in clustering resolution (not shown). The above ancestry pattern was generally consistent across all STRUCTURE subsets although the strength of the clustering signal depended on the number of genetic markers (stronger for 17 loci) and the evenness of subspecies sampling (strongest for sets of 41 and 61 males; Appendix S2: Figs S1–S5). In several subsampling schemes, the most likely number of genetic clusters ranged from 5 to 8, despite most individuals having similar probabilities of belonging to a given cluster at any K (Appendix S2: Figs S1, S3, S4). The Mantel test supported a pattern of isolation-by-distance ($n = 250$, 17 loci, $Z = 3153.86$, $R = 0.241$, $P < 0.001$).

Results of multivariate analyses (PCoA and FCA) did not indicate distinct genetic clusters and showed little consistency between assignment based on morphology and that based on genetic markers (Fig. 1). In both analyses, variation along the first two principal axes explained only a small proportion of the variance (16.09% in PCoA and 4.66% in the FCA, Fig. 1). In the FCA, variation along F1 reflected weak differences between western (*alba*) and eastern (*lugens*) samples, similar to the results of STRUCTURE at $K = 2$ (Fig. 1). Neighbour-joining reconstruction indicated that the eastern samples (*lugens* and *leucopsis* and two easternmost *ocularis*) and western samples (*alba*) were grouped in separate hierarchical clusters; however, none of the subspecies were reciprocally monophyletic (Appendix S2: Fig. S7). In the AMOVA, only a small portion of the total variance was attributable to differentiation among subspecies (5%) and populations (sample groups, 2%), while most of the variance was due to within-population variation (93%). Samples of *lugens* were the most genetically differentiated, with the highest F_{st} in all pairwise comparisons, including those between the geographically proximate *ocularis* (sample DJM) and *leucopsis* (sample PRI) (Table 2). The *alba* samples located at the western edge of the sampling range from Europe (KUK, TVE, WRU) had significant F_{st} values in many pairwise comparisons, particularly with the geographically distant *alba* from Siberia (AST, TCH, AKR, NOV) (Table 2). Most samples in close geographic proximity did not demonstrate significant genetic distances, except for *lugens* (Table 2).

Genetic differentiation in alba-personata and ocularis-lugens-leucopsis subspecies

In the two regions of contact between different subspecies that we studied in detail, there was only weak agreement between the distribution of plumage clusters and signal of population structure, as well as between different assignment methods. In the *ocularis-lugens-leucopsis* dataset, both default (most likely $K = 2$) and

Table 2 Fixation indices (F_{st}) between sampling localities. Only values significant under $*P < 0.05$ and $**P < 0.01$ are shown.

Region	Subspecies	Sample	*alba* Europe KUK	*alba* Europe TVE	*alba* Europe WRU	*alba* Europe WOU	*alba* Siberia AST	*alba* Siberia TCH	*alba* Siberia AKR	*alba* Siberia NOV	*personata* Central Asia KAZ	*personata* Siberia CAL	*personata* Siberia ULA	*personata* Siberia WES	*personata* Siberia KRA	*personata* Siberia KFS	*baicalensis* East Siberia ESI	*ocularis* East Siberia YCH	*ocularis* Far East DJM	*lugens* Far East SKU
Europe	*alba*	KUK	–																	
Europe	*alba*	TVE	0.03*	–																
Europe	*alba*	WRU			–															
Europe	*alba*	WOU				–														
Siberia	*alba*	AST	0.05**				–													
Siberia	*alba*	TCH	0.05**	0.02*	0.04*			–												
Siberia	*alba*	AKR	0.04*		0.03*				–											
Siberia	*alba*	NOV	0.04**		0.04**					–										
Central Asia	*personata*	KAZ	0.04*					0.03*			–									
Siberia	*personata*	CAL	0.05**	0.03*	0.04*							–								
Siberia	*personata*	ULA	0.06**	0.03*	0.04*					0.03*			–							
Siberia	*personata*	WES	0.07**	0.06**	0.06**									–						
Siberia	*personata*	KRA	0.05**	0.02*	0.04**										–					
Siberia	*personata*	KFS	0.05**		0.03*											–				
East Siberia	*baicalensis*	ESI	0.04*														–			
East Siberia	*ocularis*	YCH	0.09**	0.04*	0.06**			0.05**	0.04*	0.05**					0.02*	0.03**		–		
Far East	*ocularis*	DJM	0.07**		0.04*														–	
Far East	*lugens*	SKU	0.13**	0.12**	0.12**	0.06**	0.08**	0.07**	0.06**	0.11**	0.07**	0.06**	0.05**	0.03*	0.07**	0.08**	0.04*	0.06**	0.03*	–
Far East	*leucopsis*	PRI	0.09**	0.06**	0.07**			0.03*		0.05**						0.02*		0.02*		0.07**

LOCPRIOR (most likely $K = 3$) STRUCTURE models revealed that most individuals from allopatric populations of *lugens* and the *lugens-leucopsis* hybrid zone form a cluster (default model) or two clusters (LOCPRIOR model) distinct with *leucopsis* and *ocularis* (Fig. 2). The latter two did not differ from one another according to STRUCTURE (Fig. 2). FCA revealed some clustering among genotypes of all three subspecies although with overlapping F1 and F2 scores (Fig. 2). The spatial clustering model of GENELAND suggested the presence of five genetic clusters although there was only a subtle difference in density scores for $K = 4$ and 5 (Fig. 2). According to GENELAND, allopatric *lugens* formed a distinct genetic cluster, but there was no biologically meaningful clustering among *leucopsis* and *ocularis* (Fig. 2, Appendix S2: Fig. S8).

Both the default (most likely $K = 4$) and LOCPRIOR (most likely $K = 3$) models of STRUCTURE did not recover population structure across *alba, personata* and their hybrid zone (Fig. 2). Consistently, the result of FCA showed broad overlap across the three groups (Fig. 2). In contrast, the spatial model in GENELAND (most likely $K = 2$) identified two weakly differentiated genetic clusters, whose distribution somewhat corresponded to the distribution of *alba* and *personata* plumage indices (Figs 2 and 3). There was a probability above 0.7 that samples of *alba* from west Siberia (KLO, ZUD, NOV) were assigned to one cluster, and a probability above 0.6 that *personata* from Altai and Sayan regions (KUS, ULA, TAN, SES, KUT, KAN) were assigned to a different cluster (Fig. 3).

Discussion

We analysed variation in multiple nuclear markers in the White wagtail, a polytypic bird species, to evaluate potential drivers of previously reported discordance between geographic patterns in morphology and mtDNA markers. On the one hand, observed discrepancies could be due to inconsistent population histories resolved by mtDNA, which could be caused by selective sweeps, asymmetric introgression of mtDNA or demographic factors. On the other hand, extensive phenotypic differentiation between the subspecies could have little to do with the variation observed in either nuclear or mtDNA genetic markers. Most analyses revealed a general lack of population structure or genetic differentiation, as well as few barriers to gene flow between phenotypic subspecies and a pattern of isolation-by-distance. Some approaches detected a signal of population structure across *alba-personata* (GENELAND) and *ocularis-lugens-leucopsis* (STRUCTURE, FCA) somewhat concordant with plumage variation (Figs 2 and 3), but there was subtle divergence between these clusters with the exception of *lugens*. Together, our findings were consistent with those previously reported using mtDNA from the North Eurasian White wagtail subspecies (Pavlova *et al.*, 2005; Li *et al.*, 2016), and they suggest that neutral population structure is weak and only partly congruent with plumage variation and subspecies delineation.

Morphological and genetic variation across the White wagtail subspecies

Several lines of evidence suggested that *lugens* possesses some degree of population subdivision among the White wagtail subspecies. Four samples of *lugens* (BAK, KOR, NEV, KUR; Table 1, Fig. 1) from Kuril and Sakhalin Islands were separated from the mainland by large stretches of seawater, raising the possibility that such barriers could have facilitated genetic isolation and divergence of these populations. However, all of the individuals sampled from Reinike Island (sample REI, Fig. 1), two of which were identified as *lugens* x *leucopsis* intermediates and five as *lugens* based on plumage, had high STRUCTURE probabilities to be assigned to different clusters with a) *leucopsis* sampled approximately 200 km away (samples ARS and GAI) and b) with all but one *ocularis* (Figs 1 and 2). There are instances, therefore, of divergent genotypes maintained in close geographic proximity. The large and significant genetic distances between these populations further suggest restricted gene flow (Table 2). The idea that *lugens* and neighbouring White wagtail races may possess substantial reproductive isolation was previously suggested by observed differences in habitat preferences, life history traits, evidence for assortative mating between *lugens* and *leucopsis* in Primor'e (Nazarenko, 1968), and the apparent lack of intermediate phenotypes in regions where subspecies coexist (Nazarenko, 1968; Stepanyan, 1983, 2003). This evidence led to the assignment of *lugens* as a biological species (AOU, 1983, Badyaev *et al.*, 1996; Stepanyan, 2003; Koblik *et al.*, 2006). However, the species status of *lugens* was later contradicted by findings that it did not form a monophyletic clade with respect to other White wagtail subspecies based on mtDNA markers (Voelker, 2002; Pavlova *et al.*, 2005). Moreover, there was also evidence of shallow divergence between *lugens* and other subspecies in a Z-linked nuclear intron (Alström & Ödeen, 2002) and gene flow between *lugens* and *leucopsis* in Primor'e (Pavlova *et al.*, 2005). Although our sample size is limited, the results suggest that *lugens* may have restricted gene flow with neighbouring *leucopsis* and *ocularis* populations. Whether these genetic signatures are due to substantial reproductive isolation or geographic barriers remains to be determined using more comprehensive sampling from the contact zones between *lugens* and other Asian White wagtail subspecies. Interestingly, our results indicate that the Primor'e region of the Russian Far East is a contact zone between two genetically distinct lineages, similar to the results based on mtDNA (Pavlova *et al.*, 2005). This

finding corroborates the idea that patterns of neutral genetic evolution in the White wagtail may reflect processes related to historical demography rather than processes driving morphological evolution (Pavlova *et al.*, 2005).

Previous findings from Semenov *et al.* (2017) indicated that the geographic transition in plumage is substantially shifted compared to that of genomic ancestry across the *alba* and *personata* contact zone in Siberia. Specifically, transition in head-and-neck plumage from *alba* to *personata* occurred over a highly homogeneous genetic background not distinguishable from allopatric *alba*. Only samples of *personata* located 100–200 km southeast of the transition in plumage were slightly differentiated (western Altai mountains), whereas major genetic transition occurred much further south. Findings of our present study are consistent with these results. First, although the spatial model of GENELAND revealed some degree of population structure between *alba* and *personata*, genetic differences between these clusters were subtle (Fig. 2, Table 2). Second, this clustering pattern appears to be entirely driven by high-altitudinal *personata* samples from Altai and Sayan mountains (KUS, ULA, TAN, SAY, Fig. 3). Thus, despite the fact that spatial distribution of genetic clusters was coarsely co-aligned with distribution of *alba* and *personata* plumage indices, the borders between genetic and plumage variation appear to have a mismatch over several hundred kilometres (Fig. 3). Lack of genetic differentiation and the geographic shift between plumage and genetic markers observed in Semenov *et al.* (2017) may therefore span the broad range of *alba* and *personata* coexistence. Interestingly, our STRUCTURE and FCA analyses did not detect genetic structuring identified by GENELAND. This result is likely due to the GENELAND spatial model being able to account for spatial distribution of genetic variation and hence its superior ability to capture a very weak clustering signal, provided that geographic sampling is comprehensive enough.

Some of our STRUCTURE analyses on the six subspecies subsets revealed that there could be more than two genetic clusters (most likely $K = 5–8$, Appendix S2: Figs S1, S3 and S4). In these results, no given individual or subspecies tended to form a distinct genetic cluster, with the exception of *lugens*, the *lugens-leucopsis* contact zone and several *personata* and *ocularis* individuals that had a high probability of clustering together. This higher number of inferred clusters may indicate additional weak population structure that was undetected by other methods. However, we believe that this pattern may have resulted from rare alleles in combination with very weak overall genetic differentiation. We therefore hesitate to assign it a meaningful biological explanation.

A high proportion of our samples was collected in the regions close to zones of contact and hybridization between different subspecies (e.g. for *personata*, *leucopsis* and *baicalensis*). In addition, some of the subspecies examined here were represented by a few individuals (*leucopsis* and *baicalensis*) and this unbalanced sample size could potentially have an effect on population assignment patterns (Puechmaille, 2016). At the same time, for the subspecies with a dense sampling of allopatric populations (*alba*, *ocularis* and *lugens*), we observed some degree of clustering signal. It is therefore probable that that there might be a good level of agreement between subspecies established on plumage and neutral structure except for populations close to contact zones, where there seems to be rather strong discordance in cline centres between plumage and neutral markers. Interestingly, another recently published study examining genetic variation in the White wagtail (Harris *et al.*, 2018) reached the same conclusions about genetic patterns across the North Eurasian subspecies, despite using much denser marker sampling across the genome than in our study. Similarly, there was a gradual continent-wide transition in ancestry and no apparent genetic structuring of subspecies (Harris *et al.*, 2018). Together, these results highlight that the use of whole-genome resequencing methods and a comprehensive sampling in allopatric parts of the range and across multiple hybrid zones between different subspecies are needed to better assess the degree of concordance between genetic and phenotypic variation in the White wagtail.

Plumage divergence with little genetic differentiation: a perspective on linking patterns to processes

Pronounced phenotypic diversification coexisting with little genetic divergence has long puzzled scientists, particularly in the cases of divergence in conspicuous plumage signals in birds. Recent advances in sequencing technologies have afforded sufficient resolution to empirically assess what has been implicated for decades – that a comparatively small fraction of the genome can underlie such variation (Poelstra *et al.*, 2014; Mason & Taylor, 2015; Toews *et al.*, 2016; Campagna *et al.*, 2017).

Although this pattern is commonly found, the mechanisms that generate it are less understood. Selection on genomic regions underlying distinct phenotypes on the one hand and gene flow homogenizing selectively neutral genetic variation on the other are thought to be among the major contributing factors (Harrison & Larson, 2014). But why some recently diverged yet actively hybridizing taxa maintain population structure that is generally consistent with phenotypic variation (e.g. Scordato *et al.*, 2017), whereas others – such as wagtails – do not, is unknown. One possibility is that the genetic architecture of traits involved in divergence and reproductive barriers might play a role in shaping these patterns. Selection can repeatedly target a small

number of genetic hotspots of morphological change (Martin & Orgogozo, 2013), such as certain genes, genomic regions or regulatory pathways across multiple divergence events, as in *Heliconius* butterflies (Van Belleghem *et al*., 2017), *Sporophila* Seedeaters (Campagna *et al*., 2017), *Corvus* Crows (Vijay *et al*., 2016) and *Lonchura* Munias (Stryjewski & Sorenson, 2017). Of particular note, a small toolkit of genes can provide the basis for extensive phenotypic diversification due to modification of gene expression by *cis*-regulatory elements (Wray, 2007; Wittkopp & Kalay, 2012; Poelstra *et al*., 2015; Jiggins *et al*., 2017), and re-shuffling of these modular elements due to hybridization and introgression can promote rapid phenotypic evolution (Abbott *et al*., 2013; Seehausen, 2013). Theoretical work supports the idea that traits involved in adaptive radiations might have rather simple genetic architecture (Gavrilets & Vose, 2005). All else being equal, such traits will be more resistant to the destructive effects of recombination in a context of gene flow, compared to polygenic traits. They could therefore more readily evolve and contribute to reproductive barriers, but the effects of linked selection would be limited to a small fraction of the genome, impeding the buildup of genome-wide divergence (Feder *et al*., 2012).

Plumage variation among White wagtail subspecies is confined to a small number of patches – throat, back, and the sides of the head and neck, where coloration takes on alternative states: black *vs* white or black *vs* grey although there is also considerable variation in the intensity of white colour on the wing. There are five types of head patterning (Fig. 4; 1 – *baicalensis*/*leucopsis*; 2 – *ocularis*/*lugens*; 3 – *subpersonata*; 4 – *alba*/*yarrellii*; 5 – *personata*/*alboides*), and this remarkable divergence within the same plumage regions may indicate that these patches were repeated targets of selection. This idea is supported by evidence of strong selection associated with head plumage in the *alba* and *personata* hybrid zone (Semenov *et al*., 2017). Furthermore, there is parallel divergence in black vs. grey back colour in four pairs of subspecies with the same head-and-neck pattern (Fig. 4). Although it is currently unknown whether these morphological hotspots correspond to genetic hotspots (e.g. the same few genes, genomic regions or genetic pathways), several pieces of evidence indicate that this could be the case. Examination of variation in the head and neck sides in the *alba* and *personata* hybrid zone (Semenov *et al*., 2017) and the eye-stripe in the hybrid zones between *lugens* and *leucopsis*, *baicalensis* and *ocularis*, and *alba* and *ocularis* (GAS

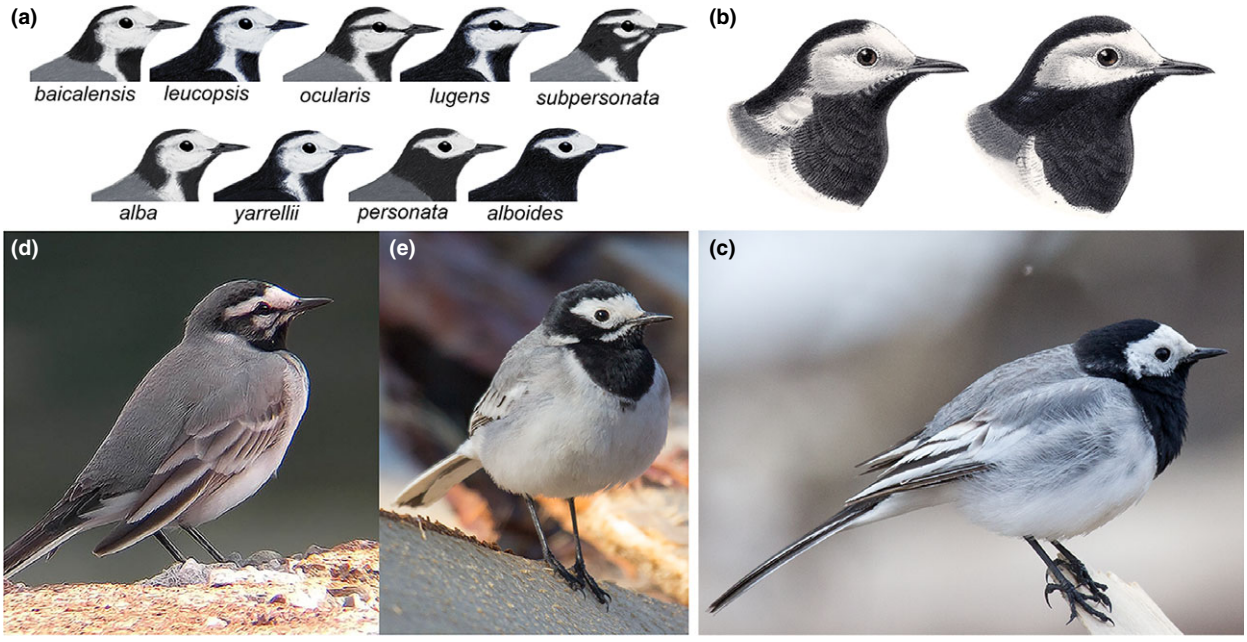

Fig. 4 Plumage variation in the White wagtail mostly occurs within the same few hotspots of evolutionary change – head and neck sides, throat and back (a). Note that grey vs. black back colour repeatedly evolved in four subspecies pairs with the same head, neck and throat pattern (in the *lugens-ocularis* pair, head and neck pattern is the same, but throat colour is different). Although the genetic architecture of these traits is unknown, indirect evidence provides some intriguing insights. Head-and-neck plumage of the *persica* race (b) from Iran (sometimes considered distinct subspecies) is remarkably similar to one type of *alba-personata* hybrids from Siberia (c). Moroccan *subpersonata* (d) appears as another *alba-personata* hybrid variant (e), if this hybrid had an eye-stripe of the *lugens-ocularis*. These striking parallelisms suggest that plumage hotspots may indeed correspond to genetic hotspots. Photo credits: (d) – Angel García; (c, e) – Georgy A. Semenov. (b) – modified from Sharpe, 1885.

unpublished) suggest that both plumage traits demonstrate simple segregation patterns with few intermediate categories and are likely determined by a small number of genes or supergenes. Another interesting insight comes from comparing *alba* and *personata* hybrid phenotypes to some of the other White wagtail subspecies. *Motacilla alba persica* is broadly distributed in Iran and is considered a distinct subspecies in some taxonomic reviews (Tyler, 2004), but not in others (Alström & Mild, 2003). Head-and-neck plumage of this race appears remarkably similar to one *alba* and *personata* hybrid type (Fig. 4b,c), and it has long been suggested that the *persica* population is actually a product of hybridization (Sharpe, 1885; Vaurie, 1959; Stepanyan, 1983), an idea that is bolstered by its distributional range being adjacent to the second *alba* and *personata* hybrid zone (Fig. 1). Another *alba* and *personata* hybrid type looks virtually indistinguishable from the Moroccan *subpersonata,* if this *alba* and *personata* hybrid had an eye-stripe 'module' of a *lugens-ocularis* pair (Fig. 4d, e). In other words, *subpersonata* and *persica* plumages appear as combinations of traits from other subspecies, and their similarity with the actual hybrids indicates that plumage parallelism in the White wagtail may have a shared genetic basis. Uncovering the actual molecular architecture of this variation can provide powerful insights into the mechanisms of phenotypic radiations. Intriguingly, two mtDNA haplotypes with over 2% sequence divergence were found in *subpersonata* (Li *et al*., 2016), suggesting that hybridization could indeed contribute to the evolution of this subspecies.

Phylogenetic reconstructions place the divergence between White wagtail mtDNA lineages in the Pleistocene (Voelker, 2002; Pavlova *et al*., 2005; Li *et al*., 2016), when extensive climate fluctuations could have led to repeated cycles of contact, admixture and geographic isolation (Hewitt, 2000). Such events might provide a context for re-shuffling and fixation to alternative states of plumage alleles and the homogenization of selectively neutral genetic variation between neighbouring subspecies, especially if selection associated with plumage targets only a few genes. The White wagtail shares an interesting feature with its congener, the Yellow wagtail (*Motacilla flava*). In both species, striking plumage diversifications show patterns discordant with those in mtDNA markers (Ödeen & Bjorklund, 2003; Pavlova *et al*., 2003; Harris *et al*., 2018). Moreover, plumage differences between Yellow wagtail subspecies involve the same hotspots as in the White wagtail – head, neck, throat and back. An outstanding question is whether the same evolutionary and molecular mechanisms shaped these parallel patterns, and this question offers a fruitful avenue for future research on the *Motacilla* genus and other avian groups with 'hotspot' patterns of plumage variation (e.g. *Oenanthe* Wheatears, black-and-white *Monarcha* Monarch flycatchers).

In summary, the results of our study illustrate that pronounced plumage diversification coexists with little genetic divergence and population structure in multiple subspecies of a widespread bird species. We suggest that this pattern could result from selection repeatedly targeting a few hotspots of morphological change and highlight the potentially pervasive role of hybridization in the evolution and diversification of the White wagtail. The history of incipient diversification, such as the one examined in this paper, should ideally be reconstructed using data from loci linked to reproductive barriers (Williamson *et al*., 2005; Abbott *et al*., 2013). Therefore, future work should use methods with higher genomic resolution to identify genomic regions underlying plumage variation in wagtails to test the above hypotheses.

Acknowledgments

We thank David Khaydarov, Andrey Lissovsky, Alex Katokhin, Evgeniy Khobotov, Alexander Yurlov, Evgeniy Davidyuk, Ilya Murashov, Evgeniy Shemyakin, Vladimir Sotnikov, Marko Rakovic, Georgii Bakhtadze, Nikolay Balatsky and Anatoliy Shapoval for their help with field sampling and for sharing tissue specimens. Authors thank Susana Lopes and Sandra Reis for their help with the lab work and Sergei Drovetski for the financial support in the development of microsatellite primers and the lab work. Four samples of *personata* from Kazakhstan were loaned by the University of Washington Burke Museum of Natural History and Culture (UWBM 46410, 46411, 46419, 46420) and two *baicalensis* by the Yale Peabody Museum of Natural History (YPM 140067, 140317). Authors thank Kristof Zyskowski, Sharon Birks and Sergei Drovetski for their help with sample shipments. We thank Vladmir Loskot for granting access to collections of the Zoological Institute of the Russian Academy of Sciences. Erin Morrison, Ahva Potticary, Kelly Hallinger, David Zonana, Sheela Turbek and three anonymous reviewers provided helpful comments on the manuscript. This work was supported by grants from Russian Foundations for Basic Research 14-04-01492 and 12-04-31042 to GAS, Russian Science Foundation 14-50-00029 to YAR, Moscow State University AAAA-A16-116021660077-3 to YAR and EAK, and by the G.G. Simpson Postdoctoral Fellowship from the University of Arizona to GAS. We are grateful to Yulia Sheina for the wagtail drawings used in Figs 1,2 and 4 and to Angel García for granting permission to publish the photograph of *Motacilla alba subpersonata*. Sampling was carried out with necessary collecting permissions granted to ZMMSU, SZM and KCZM. No authors have any conflict of interest to declare.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J.W., Baird, S.J., Bierne, N. *et al.* 2013. Hybridization and speciation. *J. Evol. Biol.* **26**: 229–246.

Alström, P. & Mild, K. 2003. *Pipits and Wagtails*. Christopher Helm Publishers Ltd, London.

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: *Species Limits and Systematics in Some Passerine Birds* (P. Alström, ed.), pp. 16–21. Acta Universitatis Upsaliensis, Uppsala.

AOU 1983. *Check-List of North American birds*, 4th edn. Am. Ornith. Union, Washington, D.C.

Avise, J.C. 2000. *Phylogeography: The History and Formation of Species*. Harvard University Press, Cambridge, MA.

Aye, R., Schweizer, M. & Roth, T. 2012. *Birds of Central Asia. Kazakhstan, Turkmenistan, Uzbekistan, Kyrgyzstan, Tajikistan and Afganistan*. Christopher Helm, London.

Badyaev, A.V., Gibson, D.D. & Kessel, B. 1996. White Wagtail (*Motacilla alba*). In: *The Birds of North America* (A. Poole & F. Gill, eds). pp. 1–23. Philadelphia: The Academy of Natural Sciences and The American Ornithologists' Union, Washington, D.C.

Baldassarre, D.T., White, T.A., Karubian, J. & Webster, M.S. 2014. Genomic and morphological analysis of a semipermeable avian hybrid zone suggests asymmetrical introgression of a sexual signal. *Evolution* **68**: 2644–2657.

Barton, N.H. & Bengtsson, B.O. 1986. The barrier to genetic exchange between hybridising populations. *Heredity* **57**: 357–376.

Belkhir, K., Borsa, P., Goudet, J., Chikhi, L. & Bonhomme, F. 1998. *Genetix, Logiciel sous Windows pour la Génétique des Populations*. Laboratoire Génome et Populations, CNRS UPR 9060, Université de Montpellier II, Montpellier, France.

Berezovikov, N. & Reznichenko, S. 2014. Nesting of the masked wagtail Motacilla personata and cases of its hybridization with the white wagtail Motacilla alba in Bayanaul (North-East Kazakhstan). *Russ. Ornithol. J.* **23**: 73–78.

Berezovikov, N., Samusev, I., Khrokov, V. & Egorov, V. 2007. Passerine birds of Irtysh floodplains and Altai foothills. *Russ. Ornithol. J.* **16**: 1031–1055.

Blacket, M.J., Robin, C., Good, R.T., Lee, S.F. & Miller, A.D. 2012. Universal primers for fluorescent labelling ofPCR fragments—an efficient and cost-effective approach to genotyping by fluorescence. *Mol. Ecol. Resour.* **12**: 456–463.

Brumfield, R.T., Jernigan, R.W., McDonald, D.B. & Braun, M.J. 2001. Evolutionary implications of divergent clines in an avian (Manacus: Aves) hybrid zone. *Evolution* **55**: 2070–2087.

Butlin, R., Debelle, A., Kerth, C., Snook, R.R., Beukeboom, L.W., Castillo Cajas, R.F. *et al.* 2012. What do we need to know about speciation? *Trends Ecol. Evol.* **27**: 27–39.

Campagna, L., Benites, P., Lougheed, S.C., Lijtmaer, D.A., Di Giacomo, A.S., Eaton, M.D. *et al.* 2012. Rapid phenotypic evolution during incipient speciation in a continental avian radiation. *Proc. Biol. Sci.* **279**: 1847–1856.

Campagna, L., Repenning, M., Silveirad, L.F., Fontana, C.S., Tubaro, P.L. & Lovette, I.J. 2017. Repeated divergent selection on pigmentation genes in a rapid finch radiation driven by sexual selection. https://doi.org/10.1126/sciadv.1602404.

Chan, K.M.A. & Levin, S.A. 2005. Leaky prezygotic isolation and porous genomes: rapid introgression of maternally inherited DNA. *Evolution* **59**: 720–729.

Cheng, C.S. 1976. *Catalog of birds of China and their distribution*. Academia Sinica, Pekin, China.

Coyne, J.A. & Orr, H.A. 2004. *Speciation*. Sinauer Associates, Inc, Sunderland, MA.

Cramp, S. 1988. *Handbook of the Birds of Europe, the Middle East and North Africa*. Oxford Univ. Press, Oxford.

Currat, M., Ruedi, M., Petit, R.J. & Excoffier, L. 2008. The hidden side of invasions: massive introgression by local genes. *Evolution* **62**: 1908–1920.

Dement'ev, G.P., Gladkov, N.A. & Sudilovskaya, A.M. 1954. *Birds of the Soviet Union*. Sovetskaya Nauka, Moscow.

Earl, D.A. & vonHoldt, B.M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* **4**: 359–361.

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

Evanno, G., Regnaut, S. & Goudet, J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**: 2611–2620.

Excoffier, L., Smouse, P.E. & Quattro, J.M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* **131**: 479–491.

Feder, J.L., Egan, S.P. & Nosil, P. 2012. The genomics of speciation-with-gene-flow. *Trends Genet.* **28**: 342–350.

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.

Gavrilets, S. & Vose, A. 2005. Dynamic patterns of adaptive radiation. *Proc. Natl. Acad. Sci. USA* **102**: 18040–18045.

Glutz Von Blotzheim, U.N. & Bauer, K.M. 1985. *Handbuch der Vögel Mitteleuropas*, Vol. **10**. Passeriformes, Akademische Verlagsgesellschaft, Wiesbaden, Germany.

Good, J.M., Vanderpool, D., Keeble, S. & Bi, K. 2015. Negligible nuclear introgression despite complete mitochondrial capture between two species of chipmunks. *Evolution* **69**: 1961–1972.

Guillot, G., Estoup, A., Mortier, F. & Cosson, J.F. 2005. A spatial statistical model for landscape genetics. *Genetics* **170**: 1261–1280.

Harris, R.B., Alström, P., Ödeen, A. & Leaché, A.D. 2018. Discordance between genomic divergence and phenotypic variation in a rapidly evolving avian genus (Motacilla). *Mol. Phylogenet. Evol.* **120**: 183–195.

Harrison, R.G. 1990. Hybrid zones: windows on evolutionary processes. *Oxf. Surv. Evol. Biol.* **t**: 69–128.

Harrison, R.G. & Larson, E.L. 2014. Hybridization, introgression, and the nature of species boundaries. *J. Hered.* **105** (Suppl 1): 795–809.

Harrison, R.G. & Larson, E.L. 2016. Heterogeneous genome divergence, differential introgression, and the origin and structure of hybrid zones. *Mol. Ecol.* **25**: 2454–2466.

Hewitt, G.M. 2000. The genetic legacy of the Quaternary ice ages. *Nature* **405**: 907–913.

Hubisz, M.J., Falush, D., Stephens, M. & Pritchard, J.K. 2009. Inferring weak population structure with the assistance of sample group information. *Mol. Ecol. Resour.* **9**: 1322–1332.

Ilyashenko, V.Y. 1986. Birds of upper Zeya basin. *Proc Zool Inst USSR Acad Sci.* **150**: 77–81.

Irwin, D.E., Rubtsov, A.S. & Panov, E.N. 2009. Mitochondrial introgression and replacement between yellowhammers (*Emberiza citrinella*) and pine buntings (*Emberiza leucocephalos*) (Aves: Passeriformes). *Biol. J. Lin. Soc.* **98**: 422–438.

Jakobsson, M. & Rosenberg, N. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* **23**: 1801–1806.

Jensen, J.L., Bohonak, A.J. & Kelley, S.T. 2005. Isolation by distance, web service. *BMC Genet.* **6**: 13.

Jiggins, C.D., Wallbank, R.W. & Hanly, J.J. 2017. Waiting in the wings: what can we learn about gene co-option from the diversification of butterfly wing patterns? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **372**: 1–10.

Koblik, E.A., Rohwer, S., Drovetski, S.V., Wood, C.S., Andreev, A.V., Banin, D.A. *et al.* 2001. Faunistic records from the eastern regions of Russia. *Ornithologia* **29**: 47–57.

Koblik, E.A., Red'kin, Y.A. & Arkhipov, V.Y. 2006. *List of Birds of Russian Federation*. KMK, Moscow.

Krosby, M. & Rohwer, S. 2009. A 2000 km genetic wake yields evidence for northern glacial refugia and hybrid zone movement in a pair of songbirds. *Proc. Biol. Sci.* **276**: 615–621.

Kumar, S., Stecher, G. & Tamura, K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**: 1870–1874.

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.

Lobkov, E.G. 2011. Kamchatka wagtail *Motacilla* (*alba*) *lugens* (Gloger, 1829): variability, Relationships with the Spectacled white wagtail *Motacilla alba ocularis* (Swinhoe, 1860), and the Taxonomic Status. *Far East. J.Orn.* **2**: 27–55.

Mantel, N.A. 1967. The detection of disease clustering and a generalized regression approach. *Can. Res.* **27**: 209–220.

Martin, A. & Orgogozo, V. 2013. The Loci of repeated evolution: a catalog of genetic hotspots of phenotypic variation. *Evolution* **67**: 1235–1250.

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.

Meirmans, P.G. 2015. Seven common mistakes in population genetics and how to avoid them. *Mol. Ecol.* **24**: 3223–3231.

Nazarenko, A.A. 1968. On the character of interrelations of two forms of pied wagtails in south Ussurilans. *Problemi Evolutsii*. **1**: 195–201.

Nosil, P. & Feder, J.L. 2012. Genomic divergence during speciation: causes and consequences. *Phil. Trans. R. Soc. B* **367**: 332–342.

Nosil, P. & Schluter, D. 2011. The genes underlying the process of speciation. *Trends Ecol. Evol.* **26**: 160–167.

Ödeen, A. & Bjorklund, 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.

Paludan, K. 1959. *On the birds of Afghanistan*. Zoologisk Museum, Copenhagen.

Pavlova, A., Zink, R.M., Drovetski, S.V., Red'kin, Y.A. & 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., V., F.I. *et al.* 2005. Mitochondrial DNA and plumage evolution in the white wagtail *Motacilla alba*. *J. Avian Biol.* **36**: 322–336.

Peakall, R. & Smouse, P.E. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* **28**: 2537–2539.

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. 2015. Transcriptomics of colour patterning and coloration shifts in crows. *Mol. Ecol.* **24**: 4617–4628.

Porter, R. & Aspinall, S. 2010. *Birds of the Middle East*. Christopher Helm, London.

Price, T. 2008. *Speciation in Birds*. Roberts & Company Publishers, Greenwood Village, CO, 470 pp.

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

Puechmaille, S.J. 2016. The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Mol. Ecol. Resour.* **16**: 608–627.

Raymond, M. & Rousset, F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Heredity* **86**: 248–249.

Red'kin, Y.A. 2003. New data on spatial and reproductive relationships of some close forms of Passeriformes in Tyva. In: *Material of the 1st Russian scientific conference in the memory of A. A. Buturlin*. pp. 201–206. Ulyanovsk.

Roberts, T.J. 1992. *The Birds of Pakistan*. Vol. **2** Oxford University Press, Oxford.

Romanov, A.A. 1996. *Birds of Putorana Plateau*. Putorana Nature Preserve, Moscow, Russia.

Rousset, F. 1997. Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**: 1219–1228.

Ryabitsev, V.K., Kovshar, A.F., Kovshar, V.A. & Berezovikov, N.N. 2014. *Field Guide to the Birds of Kazakhstan*. Menzbier Ornithological Society and Kazakhstan Birds Conservation Union, Almaty.

Saitou, N. & Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406–425.

Scordato, E.S.C., Wilkins, M.R., Semenov, G., Rubtsov, A.S., Kane, N.C. & Safran, R.J. 2017. Genomic variation across two barn swallow hybrid zones reveals traits associated with divergence in sympatry and allopatry. *Mol. Ecol.* **26**: 5676–5691.

Seddon, N., Botero, C.A., Tobias, J.A., Dunn, P.O., Macgregor, H.E., Rubenstein, D.R. *et al.* 2013. Sexual selection accelerates signal evolution during speciation in birds. *Proc. Biol. Sci.* **280**: 20131065.

Seehausen, O. 2013. Conditions when hybridization might predispose populations for adaptive radiation. *J. Evol. Biol.* **26**: 279–281.

Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E., Boughman, J.W., Hohenlohe, P.A. *et al.* 2014. Genomics and the origin of species. *Nat. Rev. Genet.* **15**: 176–192.

Semenov, G.A. & Yurlov, A.K. 2010. Relationships between masked and white wagtails on the south of Siberia. *Ornithologia* **36**: 7–21.

Semenov, G., Yurlov, A. & Khaydarov, D. 2010. Hybridization of *Motacilla alba* Linnaeus, 1758, and *M.* (a.) *personata* Gould,

1861, in the south of Siberia. *Contemp. Probl. Ecol.* **3**: 579–586.

Semenov, G.A., Scordato, E.S.C., Khaydarov, D.R., Smith, C.C.R., Kane, N.C. & Safran, R.J. 2017. Effects of assortative mate choice on the genomic and morphological structure of a hybrid zone between two bird subspecies. *Mol. Ecol.* **26** (22): 6430–6444.

Sharpe, R.B. 1885. *Catalogue of the Birds in the British Museum*. The Trustees of the British Museum, London.

Stepanyan, L.S. 1983. *Superspecies and twin-species in the avifauna of USSR*. Nauka, Moscow.

Stepanyan, L.S. 2003. *Conspectus of the ornithological fauna of Russia and adjacent territories (within the borders of the USSR as a historic region)*. Academkniga, Moscow.

Stryjewski, K.F. & Sorenson, M.D. 2017. Mosaic genome evolution in a recent and rapid avian radiation. *Nat. Ecol. Evol.* **1**: 1912–1922.

Sushkin, P. 1938. *Birds of Soviet Altai and Adjacent Parts of North West Mongolia*. Academy of Sciences of the USSR, Moscow-Leningrad.

Toews, D.P. & Brelsford, A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Mol. Ecol.* **21**: 3907–3930.

Toews, D.P., Mandic, M., Richards, J.G. & Irwin, D.E. 2014. Migration, mitochondria, and the yellow-rumped warbler. *Evolution* **68**: 241–255.

Toews, D.P., 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.

Tsvetkov, A.V., Red'kin, Y.A. & Koblik, E.A. 2003. To distribution and biology of Wagtails in Tyva. *Russ. Ornithol. J.* **229**: 768–787.

Tyler, S. 2004. Family Motacillidae (Pipits and Wagtails). In: *Handook of The Birds of the World*, (J. del Hoyo, ed.). pp. 777–778. Lynx Edicions, Barselona.

Van Belleghem, S.M., Rastas, P., Papanicolaou, A., Martin, S.H., Arias, C.F., Supple, M.A. *et al.* 2017. Complex modular architecture around a simple toolkit of wing pattern genes. *Nat. Ecol. Evol.* **1**: 0052.

Vaurie, C. 1959. *The Birds of Paleactic Fauna*. Passeriformes, London.

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**: 725–739.

Williamson, S.H., Hernandez, R., Fledel-Alon, A., Zhu, L., Nielsen, R. & Bustamante, C.D. 2005. Simultaneous inference of selection and population growth from patterns of variation in the human genome. *Proc. Natl Acad. Sci. USA* **102**: 7882–7887.

Wittkopp, P.J. & Kalay, G. 2012. Cis-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Genet.* **31**: 59–69.

Wray, G.A. 2007. The evolutionary significance of cis-regulatory mutations. *Nat. Rev. Genet.* **8**: 206–216.

Wu, C.-I. 2001. The genic view of the process of speciation. *J. Evol. Biol.* **14**: 851–865.

Zalesskii, I. 1927. Distribution borders of the Masked wagtail (*Motacilla personata* Gould.) in West Siberia. *Uragus* **2**: 7–8.

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.

Zink, R.M. & Barrowclough, G.F. 2008. Mitochondrial DNA under siege in avian phylogeography. *Mol. Ecol.* **17**: 2107–2121.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Appendix S1:

Table S1 Primer sequences and PCR conditions for microsatellite amplification.

Table S2 Sample Size (*N*), Number of alleles (Na), Number of effective alleles (Ne), Information Index (I), Observed Heterozygosity (Ho), Expected (He) and Unbiased Expected Heterozygosity (uHe), and Fixation Index (F) for 17 microsatellite loci.

Table S3 Allele frequency and sample size by subspecies.

Appendix S2:

Figure S1 Results of STRUCTURE analysis for 168 males using 14 autosomal and 3 Z-linked microsatellite loci for $K = 2$–6.

Figure S2 Results of STRUCTURE analysis for 250 individuals (both sexes) using 14 autosomal microsatellite loci for $K = 2$–6.

Figure S3 Results of STRUCTURE analysis for a subset of 41 males using 14 autosomal and 3 Z-linked microsatellite loci for $K = 2$–6.

Figure S4 Results of STRUCTURE analysis for a subset of 61 males using 14 autosomal and 3 Z-linked microsatellite loci for $K = 2$–6.

Figure S5 Results of STRUCTURE analysis for 55 individuals (both sexes) using 14 autosomal microsatellite loci for $K = 2$–6.

Figure S6 Results of Principal Coordinate Analysis using genetic distances (F_{st}) between 54 sampling localities on 17 microsatellite loci.

Figure S7 Neighbor-Joining tree for pairwise F_{st} between 54 sampling localities.

Figure S8 GENELAND maps of population membership probability for the most likely $K = 5$ for the *ocularis-lugens-leucopsis* subset.

Data deposited at Dryad: https://doi.org/10.5061/dryad.nn34452.

Received 2 May 2017; revised 27 March 2018; accepted 28 May 2018