RESEARCH ARTICLE SUMMARY

PIGMENTATION

A mosaic of modular variation at a single gene underpins convergent plumage coloration

Dave Lutgen *et al.*

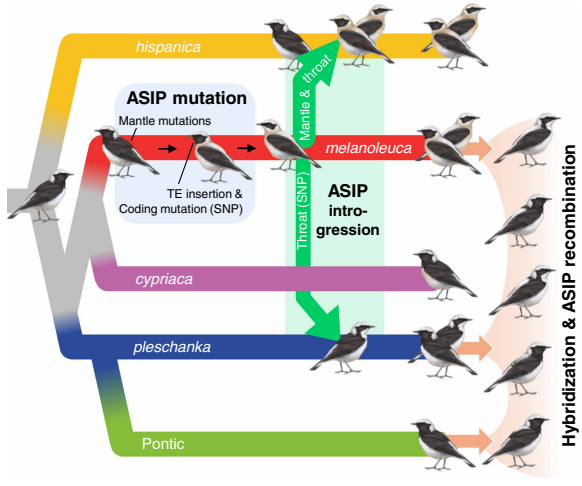
Full article and list of author affiliations: https://doi.org/10.1126/science.ado8005

INTRODUCTION: Phenotypic diversification abundantly reuses existing variation that is available within species or introgressed from other species. Still, insights into how such variation is integrated with new mutations to form new phenotypes and how this process may be limited by genetic architectures remain scarce.

RATIONALE: In this work, we investigated the genetic architecture and evolutionary history of plumage coloration in a group of songbirds. Wheatears (genus *Oenanthe*) show a high incidence of convergent plumage color evolution and frequent hybridization in the *Oenanthe hispanica* complex, within which three species (*O. hispanica*, *O. melanoleuca*, and *O. pleschanka*) exhibit a black or white throat color polymorphism, and the mantle coloration that is divergent black or white between species is abundantly recombined by hybridization. We leveraged the throat color polymorphism and the recombined mantle coloration to infer the genetic architecture underpinning plumage coloration. On the basis of population genetic patterns observed for the color-related genomic variation, we reconstructed plumage colorations' evolutionary histories.

RESULTS: Genome-wide association studies associate throat and mantle coloration with variation at a single gene, the agouti signaling protein (ASIP) locus. Throat coloration is underpinned by the interaction of a new transposable element (TE) insertion 5′ of ASIP with nonsynonymous variation within the ASIP coding sequence. Mantle coloration is governed by additive effects at multiple sites up to 50 kbp 5′ of ASIP. Expression patterns and isoform structures of ASIP in regrowing feathers confirm this role of ASIP in the regulation of the melanin-based plumage coloration. Population genetic analyses based on the ancestral recombination graph place the evolution of white throat and mantle coloration in *O. melanoleuca*, from where both introgressed into *O. hispanica*, and white throat coloration introgressed into *O. pleschanka*. In *O. melanoleuca* and *O. hispanica*, throat phenotypes are associated with different nitrogen stable isotope signatures that suggest an association of throat coloration with alternative foraging niches. At the genus level, the convergent evolution of throat and mantle coloration involved variation different from that recovered in the *O. hispanica* complex; that is, new variation.

CONCLUSION: The convergent plumage coloration in wheatears evolved from a combination of new and introgressed variation at a single gene. Whereas the convergent pattern of coloration predominantly roots in past introgression within the *O. hispanica* complex, it involved new mutations at genus scale. Introgression further contributed to the formation of recombined phenotypes in extant hybrid zones that may constitute a source for future phenotypic evolution. The complex patterns of color evolution in wheatears demonstrate how phenotypes can be decoupled between body parts within the constraints of a monogenic architecture. Despite the strong linkage within a single gene, fine-scale recombination can integrate linked regulatory modules from divergent origins into new, recombined haplotypes. This appears to imperatively involve long-term recombination; be it within species, in their ancestors, or in hybrid zones that ever increasingly come into focus as notable melting pots that enable evolutionary tinkering. □



Plumage color evolution in wheatears. White mantle and throat coloration evolved in *O. melanoleuca* through regulatory mutations (mantle) and coding mutations and TE insertion (throat) at the ASIP gene. Introgression of ASIP variation led to white mantle coloration in *O. hispanica* and a throat polymorphism in *O. hispanica* and *O. pleschanka*. Fine-scale ASIP recombination in hybrid zones underpins the expression of intermediate plumage phenotypes. SNP, single-nucleotide polymorphism.

Corresponding author: Reto Burri (reto.burri@vogelwarte.ch) Cite this article as D. Lutgen *et al.*, *Science* **390**, eado8005 (2025). DOI: 10.1126/science.ado8005

PIGMENTATION

A mosaic of modular variation at a single gene underpins convergent plumage coloration

Dave Lutgen[1,2,3], Valentina Peona[2,4], Madeline A. Chase[2], Niloofar Alaei Kakhki[5], Fritjof Lammers[2], Stacey G. de Souza[6], Anne-Lyse Ducrest[7], Marta Burri[2], Pavlos Andriopoulos[8], Sifiso M. Lukhele[6], Michaella Moysi[6], Elizabeth Yohannes[2], Abdin Abbasov[9], Tamer Albayrak[10], Mansour Aliabadian[11], Nicolas Auchli[2], Vasileios Bontzorlos[12], Ioulios Christoforou[13], José Luis Copete[14], Egidio Fulco[15], Jesus T. Garcia[16], Zura Javakhishvili[17], Anna Kazazou[18], Fumin Lei[19,20], Yang Liu[21], Nika Paposhvili[17], Robert Patchett[22], Áron Péter[23], Raphael Ritter[24], Attila D. Sándor[23,25,26], Fabian Schneider[27], Petar Shurulinkov[28]†, Sergey Sklyarenko[29], Borut Stumberger[30], Abulfaz Tagiyev[31,32], Alessia Uboldi[2], Nikitas Vogiatzis[18], Fanny Taborsak-Lines[33]‡, Joel Gruselius[34], Liqun Yao[33], Catherine L. Peichel[1], Alexander Suh[35,36,37], Pierre-Alexandre Gagnaire[38], Alexander N. G. Kirschel[6], Manuel Schweizer[39,40], Holger Schielzeth[3,41], Reto Burri[1,2,3]*

The reshuffling of genomic variation from multiple origins is an important contributor to phenotypic diversification, yet insights into the evolutionary trajectories of this combinatorial process and their interplay with genetic architecture remain scarce. We show that convergent plumage color evolution in wheatears involves a monogenic architecture with modular variation introgressed at the agouti signaling protein (ASIP) locus. Introgression of a new transposable element insertion and linked protein-coding variation underpin a transspecific throat color polymorphism, which stable isotopes suggest is associated with alternative foraging niches. Cointrogression of linked regulatory ASIP variation resulted in mantle color convergence in one species, whereas convergent color evolution at the genus level required new variation. Our results demonstrate evolutionary trajectories from introgressed variation realized within the constraints of a monogenic architecture.

Phenotypic evolution was long perceived as a slow process driven by a gradual accumulation of new mutations (*1*). However, recent studies across diverse organisms show that the reshuffling of already existing genetic variation, whether standing within species or introgressed from other species ("preexisting variation"), provides an important fast-track to phenotypic evolution (*2*–*6*). Yet, we still have much to learn about the trajectories along which this reshuffling occurs and how it is shaped by the genetic architecture of phenotypes (*6*).

Prominent examples such as Denisovan-derived adaptations in humans, adaptive introgression in butterflies, Darwin's finches and Yellowstone wolves (*3*, *7*–*9*), and adaptation from ancestral variation in stickleback (*10*) all demonstrate preexisting variation as a potent source for phenotypic evolution. These examples highlight how rapid phenotypic evolution can be facilitated by preexisting variation at single genes. By contrast, the evolution of more complex phenotypes, such as recombined phenotypes composed of reshuffled combinations of multiple traits, typically involve genetic variation spanning multiple genes. For example, the adaptive radiation of cichlid fish and convergent color evolution in several bird genera show how individual phenotypic traits can be recombined into rich suites of phenotypes (*2*, *11*–*15*).

Whether the evolution of phenotypic richness typically involves preexisting variation remains a key question. The answer may in part depend on the genetic architecture of phenotypes. Although single genes repeatedly hit by new mutations over time may generate complex phenotypic patterns within single traits (*16*), examples of complex recombined phenotypes underpinned by monogenic architectures remain scarce (*9*, *17*, *18*). This raises the question of how genetic architecture limits the role of preexisting variation in shaping complex phenotypic diversity. Although the cited examples show the power of combinatorial processes in driving phenotypic and species diversity (*6*), substantial gaps remain in our understanding of the evolutionary trajectories along which the reshuffling of preexisting variation into complex phenotypes occurs. The role of genetic architectures in constraining this combinatorial process warrants further investigation (*19*).

To obtain insights into the role of preexisting variation in the evolution of recombined phenotypes and its interplay with genetic architecture, we investigated the genomic basis and evolutionary trajectories of plumage coloration in a group of songbirds, the wheatears (genus *Oenanthe*). The high incidence of convergent evolution [according to (*20*), including all instances of independent evolution of a phenotype; elsewhere referred to as parallel or repeated evolution (*21*)] and hybridization in wheatears offers powerful opportunities to investigate the involvement of combinatorial processes in the evolution of phenotypic diversity (*22*). Males of most wheatear species display plumages combining black and white feather tracts. The coloration of single feather tracts largely switches independently across species, resulting in phenotypes that reappear multiple times across the genus (*22*–*24*). The convergent evolution of these phenotypes may have involved a combination of ancestral, introgressed, and new genetic variation (*22*). However, to fully understand how these processes contribute to phenotypic evolution, it is essential to trace the evolutionary trajectories of the underlying genetic variants.

Three plumage color traits divergent across the wheatear genus are amenable to genetic mapping in the recently radiated *hispanica* complex, where they also evolved convergently (*25*), and are present as polymorphisms in three of four species (throat coloration) or are recombined by hybridization (mantle and neck side coloration) (*25*, *26*) (Fig. 1, A and B). The four species of the *hispanica* complex are distributed from western North Africa and Iberia to eastern China (Fig. 1, A and C): *Oenanthe hispanica*, *O. melanoleuca*, *O. cypriaca*, and *O. pleschanka* (*27*). Hybridization leads to phenotypic and genetic admixture (Fig. 1B) (*25*–*27*) in four regions: the western coast of the Black Sea, the Alborz mountains of northern Iran, the Caucasus, and the Mangyshlak peninsula in southwestern Kazakhstan ("Caspian," Fig. 1C) (*27*). These characteristics distinguish wheatears as an excellent system to investigate the genetic architecture and evolutionary history of plumage coloration.

Cryptic diversity, pervasive hybridization, and idiosyncratic ancestry composition of hybrid zones

To establish the geographic structure of genome-wide variation in the *hispanica* complex, we produced 379 phased genomes (assisted by linked-read information from 235 genomes) sampled over 9000 km across the complex's entire range, with dense sampling of hybrid zones (Fig. 1C and table S1).

Population genomic and phylogenomic analyses establish five ancestry groups that evolved in fast succession (Fig. 1, D to G, and figs. S1 and S2). Principal component analyses (PCAs) of genotype and haplotype frequencies distinguish separate clusters for the four known taxa and uncover a fifth cluster consisting of black-throated *O. pleschanka* phenotypes from the Pontic steppes west and north of the Black Sea (Fig. 1D and figs. S1 and S2), hereafter referred to as the Pontic lineage. The Pontic lineage is sister to *O. pleschanka* (Fig. 1E and fig. S3) but shows a distinct demographic trajectory (Fig. 1F).

Pairwise differentiation (F_{ST}) between lineages ranges from 11 to 41% (table S2). The distinctiveness of the five lineages is supported by (i) high coancestries within lineages in fineSTRUCTURE analyses (*28*) (figs. S4, and S5), (ii) the clear-cut clustering of the lineages based on coancestries (figs. S4 and S5), and (iii) nonlinear ancestry transitions across the three contact zones between *O. melanoleuca*, *O. pleschanka*, and the Pontic lineage (fig. S6). The short interval within which the inverse instantaneous coalescent rate (IICR) trajectories of the five lineages started diverging (Fig. 1F) together with cross-lineage coalescence times estimated from an ancestral recombination graph (ARG) (Fig. 1G) suggests that species split within a time window of about 60 thousand years (kyr) that occurred at the latest around 700 kyr ago (Fig. 1G; range of lowest to highest 95% confidence values, 678 to 727 kyr; table S3).

The *hispanica* complex's evolutionary history is shaped by hybridization. Aside from numerous individuals with intermediate ancestry compositions (Fig. 1D and fig. S4), D and f_4 statistics indicate abundant introgression (table S4). According to the taxa involved, each hybrid zone hosts a distinct ancestry composition, variably combining *O. melanoleuca*, *O. pleschanka*, and/or Pontic ancestries. The abundance of genetic and phenotypic admixture highlights pervasive hybridization in the *hispanica* complex and promises the power to map plumage coloration using individuals with admixed ancestries from different hybrid zones (Fig. 1, D and H, and figs. S2 and S4). All hybrid zones feature individuals with intermediate plumages (*26*, *27*), but individuals of mixed ancestry can be phenotypically indistinguishable from parental taxa (Fig. 1H and fig. S7). This combination of high and partly dissociated variation in plumage coloration and ancestry composition provides the power for the genetic mapping of color phenotypes divergent between parental species (*29*).

Modular architecture of plumage coloration

A transposable element insertion and protein-changing ASIP variation underpin throat coloration

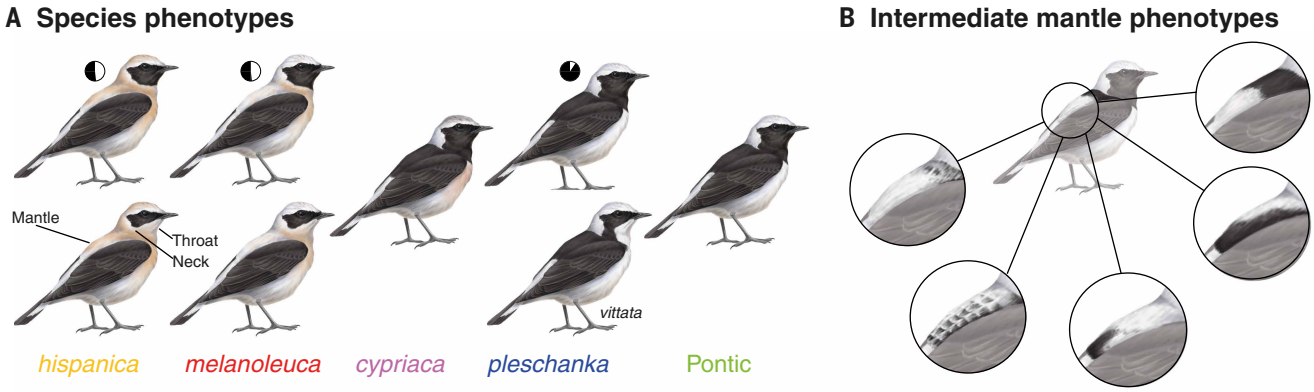
Throat coloration is polymorphic in three species of the *hispanica* complex (Fig. 1A). To investigate its genetic basis, we performed a genome-wide association study (GWAS) on 335 wheatears (57 white- and 278 black-throated) from all five lineages and hybrid zones based on linear mixed models that account for relatedness in GEMMA (*30*). The GWAS identified five single-nucleotide polymorphisms (SNPs) clustered within <15 kilo–base pairs (kbp) on chromosome 20 that are most strongly associated with throat coloration [negative logarithm of the odds ratio for linkage (–LOD) > 27; fig. S8A], all in the genomic region harboring a well-known gene within the melanogenesis pathway, the agouti signaling protein [ASIP; (*31*)] (Fig. 2, A and B, and fig. S9A; see supplementary text). The highest-scoring SNP together with the fully linked neighboring SNP ("throat SNPs") display nonsynonymous variation in consecutive codons in ASIP exon 2 that are otherwise highly conserved across birds (table S5). These throat SNPs are nearly completely associated with throat coloration across all five taxa and within *O. hispanica* and *O. melanoleuca* (Fig. 2C). The sequenced white-throated *pleschanka* ("*vittata*" phenotype, $n = 3$; Fig. 1A) are heterozygous for the same variants. Heterozygotes almost exclusively show white throat coloration, establishing the dominance of the white variant, similar to other species (*32*) and in line with ASIP expression preventing the differentiation or presence of melanocytes (*33*).

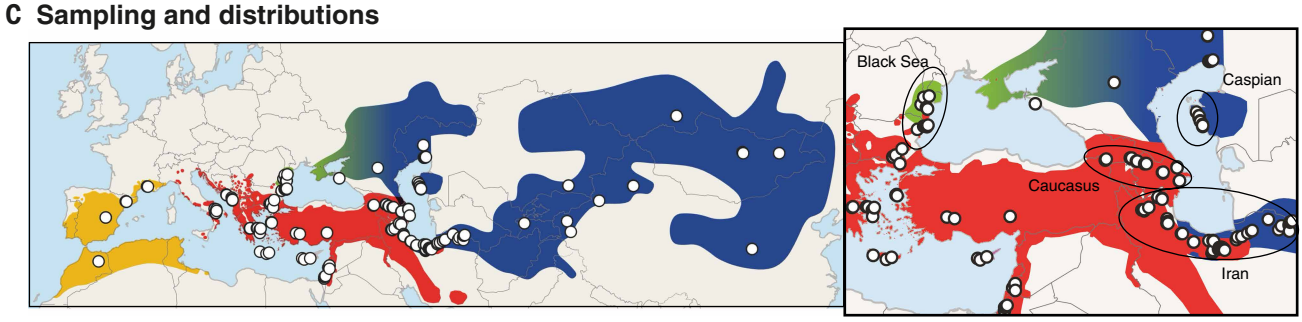
ASIP-related color variation is not usually underpinned by amino acid changes but rather with variation regulating ASIP expression [(*31*), but see (*34*)]. Therefore, we interrogated our data for linked variation hidden from typical GWASs, such as structural variants. A revised repeat annotation of chromosome 20 identified a >9-kbp-long, full-length long terminal repeat (LTR) retrotransposon ("throat LTR") in the reference genome (from a white-throated individual) 617 bp upstream of ASIP exon 2 (Fig. 2B). The presence of the throat LTR was nearly perfectly associated with white throat coloration in *O. hispanica* and *O. melanoleuca* (Fig. 2C).

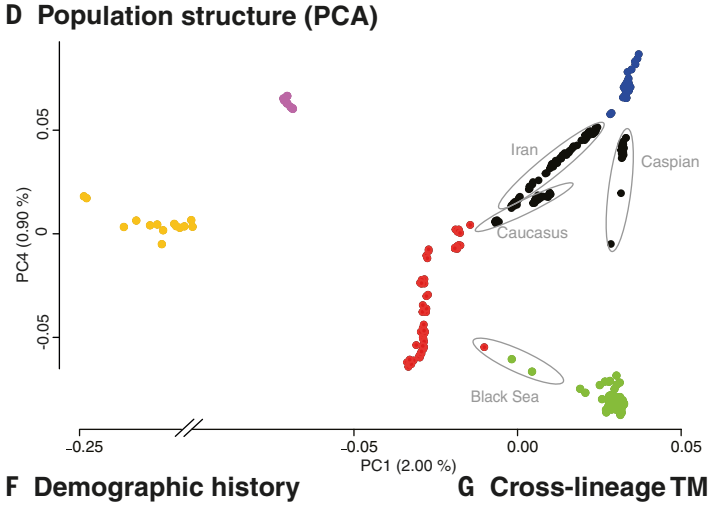
The interaction of throat SNP and LTR genotypes explains all throat color variation in *O. hispanica* and *O. melanoleuca* (associations of coloration with the remaining GWAS top hits were substantially weaker; table S6). From this, we infer that the throat LTR up-regulates ASIP expression, whereas the protein-changing throat SNP variation determines throat coloration (Fig. 2C). The LTR-SNP interaction is supported by rare recombinants for which the throat SNP and LTR genotypes do not match coloration ($n = 4$, genotypes marked by two asterisks in Fig. 2C). These mismatches suggest that the "white" ASIP protein is only expressed if an LTR up-regulates its expression (Fig. 2C, bottom two asterisks). Conversely, ASIP up-regulation by an LTR only produces white plumage if the up-regulated haplotype encodes the "white" ASIP protein (Fig. 2C, right two asterisks). Additional support for this model comes from ASIP haplotype structure. The model predicts that, to up-regulate expression of "white" protein, the LTR must collocate in cis on the same haplotype as the white-coding variant. In line with this prediction, for all heterozygotes with linked reads available from the ASIP region, linked-read barcodes exclusively ($n = 36$) or in majority ($n = 2$) placed the throat LTR and the white throat variant on the same haplotype (table S7).

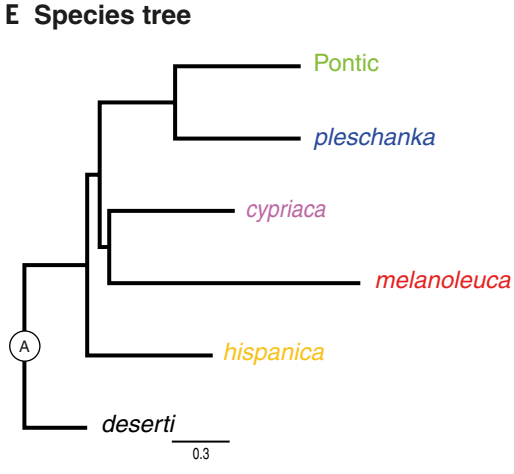
ASIP expression may be regulated differently in the white-throated *vittata* phenotype of *O. pleschanka*. Although the rarity of this phenotype and the accordingly limited data we had at hand prevents us from making firm conclusions, several lines of evidence point in this direction. We found no evidence for the LTR's presence in any black- or white-throated (*vittata*) individuals of *O. pleschanka* and the Pontic lineage. For the three (white-throated) *vittata* individuals, neither a linked-read–based de novo assembly (fig. S10 and data S1)

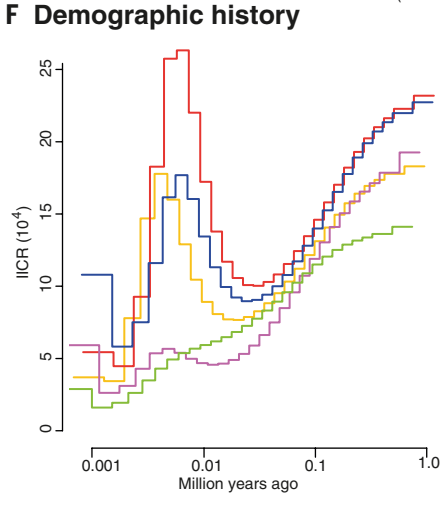
[1]Division of Evolutionary Ecology, Institute of Ecology and Evolution, University of Bern, Bern, Switzerland. [2]Swiss Ornithological Institute, Sempach, Switzerland. [3]Department of Population Ecology, Institute of Ecology and Evolution, Friedrich-Schiller-University Jena, Jena, Germany. [4]Department of Genetics and Bioinformatics, Swedish Natural History Museum, Stockholm, Sweden. [5]Department of Zoology, Stuttgart State Museum of Natural History, Stuttgart, Germany. [6]Department of Biological Sciences, University of Cyprus, Nicosia, Cyprus. [7]Département d'Ecologie et Evolution, Université de Lausanne, Lausanne, Switzerland. [8]Section of Ecology and Systematics, Department of Biology, National and Kapodistrian University of Athens, Athens, Greece. [9]Department of Bioecology, Baku State University, Baku, Azerbaijan. [10]Dokuz Eylül University, Science Faculty, Department of Biology, İzmir, Türkiye. [11]Department of Biology & Zoological Innovations Research Department, Institute of Applied Zoology, Faculty of Sciences, Ferdowsi University of Mashhad, Mashhad, Iran. [12]TYTO – Association for the Management and Conservation of Biodiversity in Agricultural Ecosystems, Larisa, Greece. [13]Pafos Zoo, Paphos, Cyprus. [14]Biblioteca de Ciència i Tecnologia, Universitat Autònoma de Barcelona, Bellaterra, Spain. [15]Studio Naturalistico MILVUS, Altamura, Italy. [16]Instituto de Investigación en Recursos Cinegéticos (IREC), CSIC-UCLM-JCCM, Ciudad Real, Spain. [17]Center for Wildlife Disease Ecology, Institute of Ecology, Ilia State University, Tbilisi, Georgia. [18]Attica Zoological Park, Athens, Greece. [19]Key Laboratory of Zoological Systematics and Evolution, Institute of Zoology, Chinese Academy of Sciences, Beijing, China. [20]University of Chinese Academy of Sciences, Beijing, China. [21]State Key Laboratory of Biocontrol, School of Ecology, Sun Yat-sen University, Shenzhen, China. [22]Centre for Biological Diversity, University of St Andrews, St Andrews, United Kingdom. [23]Department of Parasitology and Zoology, University of Veterinary Medicine, Budapest, Hungary. [24]Department of Forest Sciences, Faculty of Agriculture and Forestry, University of Helsinki, Helsinki, Finland. [25]STAR-UBB Institute, Babes-Bolyai University, Cluj-Napoca, Romania. [26]HUN-REN-UVMB Climate Change: New Blood-Sucking Parasites and Vector-Borne Pathogens Research Group, Budapest, Hungary. [27]Independent researcher, Hauteville, Switzerland. [28]National Museum of Natural History Sofia, Bulgarian Academy of Sciences, Sofia, Bulgaria. [29]Association for the Conservation of Biodiversity of Kazakhstan (ACBK), Almaty, Kazakhstan. [30]Independent researcher, Cirkulane, Slovenia. [31]Department of Zoology and Physiology, Baku State University, Baku, Azerbaijan. [32]Applied Zoology Center, Institute of Zoology, Ministry of Science and Education, Baku, Azerbaijan. [33]Science for Life Laboratory, Division of Gene Technology, KTH Royal Institute of Technology, Stockholm, Sweden. [34]Science for Life Laboratory, Department of Biosciences and Nutrition, Karolinska Institute, Stockholm, Sweden. [35]Centre for Molecular Biodiversity Research, Leibniz Institute for the Analysis of Biodiversity Change, Zoologisches Forschungsmuseum A. Koenig, Bonn, Germany. [36]Bonn Institute for Organismic Biology (BIOB) – Animal Biodiversity, University of Bonn, Bonn, Germany. [37]Department of Organismal Biology – Systematic Biology, Uppsala University, Uppsala, Sweden. [38]ISEM, Université Montpellier, CNRS, EPHE, IRD, Montpellier, France. [39]Natural History Museum Bern, Bern, Switzerland. [40]Division of Population Genetics, Institute of Ecology and Evolution, University of Bern, Bern, Switzerland. [41]German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, Leipzig, Germany. ***Corresponding author. Email: reto.burri@vogelwarte.ch** †Deceased. ‡Present address: Olink Proteomics, part of Thermo Fisher Scientific, Uppsala, Sweden.

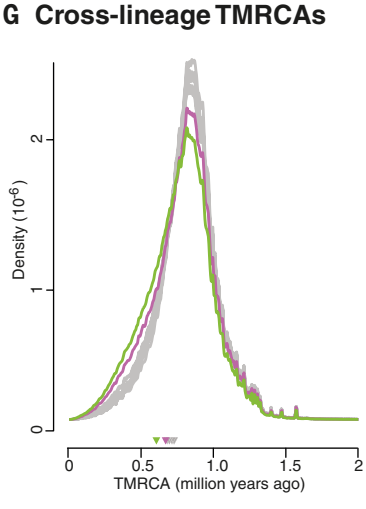












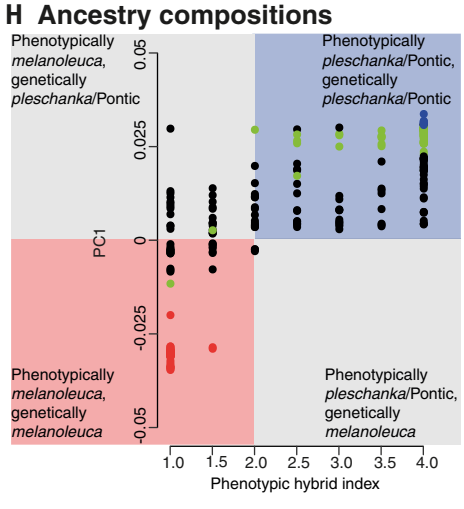


Fig. 1. Plumage color variation and evolutionary history of the *hispanica* complex. (**A**) Color phenotypes and frequencies of throat phenotypes. Mantle and neck sides are white in *O. hispanica* and *O. melanoleuca* and black in the remaining taxa. *O. hispanica*, *O. melanoleuca*, and *O. pleschanka* display a throat color polymorphism [pies indicate species-wide frequencies estimated from (*98*)]. (**B**) Hybrid-type mantle color variation. Black coloration on hybrids' mantles can vary from absent, overmottled, and reduced black to fully black. (**C**) Species ranges and sampling. Colors follow those in (A). Samples are depicted by white dots. Cyprus is shown magnified. The right-hand box zooms onto hybrid zones' locations (encircled). Species distributions are adapted from maps provided by Birdlife International (*99*). (**D**) Population structure inferred by PCA. For further PC axes, see fig. S1. Black, encircled dots depict individuals from hybrid zones. (**E**) Species tree inferred using the multispecies coalescent. The node for which ancestral states of genetic variation underpinning color phenotypes were estimated is highlighted (circled A). All nodes have 100% local posterior probability. (**F**) Demographic history of the lineages. (**G**) Genome-wide distributions of TMRCAs between lineages show that lineages split within a narrow time window. Most distributions are similar (gray); those between *O. melanoleuca* and *O. cypriaca* (purple) and those between Pontic and *O. pleschanka* (green) show more recent TMRCAs. Tick marks indicate mean TMRCAs between lineages (confidence intervals too small to be visible; table S3). (**H**) Highly variable ancestry compositions of color phenotypes provide high power for GWASs. Shown is individuals' ancestry, as estimated from PCA (low values, *O. melanoleuca*; high values, *O. pleschanka*), in relation to phenotypic hybrid index (1, white mantle and neck sides; 4, black mantle and neck sides). Additional PC axes are shown in fig. S7. Colored points follow those in (A).

nor short-read–based LTR genotyping indicated the throat LTR's presence, despite the white throat SNP variant being present. Together with the observation that *vittata* displays a white throat despite featuring the same recombinant haplotype as two black-throated *O. melanoleuca* individuals (Fig. 2C, bottom), this suggests that unknown variation may up-regulate ASIP expression in the *vittata* phenotype of *O. pleschanka*.

Additive ASIP-based architecture of mantle and neck side coloration

Mantle and neck side coloration are divergent between white-mantled *O. melanoleuca* and *O. hispanica* and black-mantled *O. pleschanka*, Pontic, and *O. cypriaca*, with the same color variants shared by non-sister species (Figs. 1, A and E).

To infer the genetic variation underpinning the species-specific mantle and neck side coloration, we leveraged information from 230 individuals from all four hybrid zones (Fig. 1C) in which mantle coloration is recombined into diverse genomic backgrounds (Fig. 1H and fig. S7). GWAS identified 17 variants (14 SNPs and 3 indels, hereafter referred to as "mantle loci;" Fig. 2A and fig. S9, B and C; see supplementary text) strongly associated with mantle coloration (–LOD > 15; fig. S8B), of which 15 are located within <45 kbp upstream of ASIP exon 2, and 2, in intron 2 (Fig. 2B).

Bayesian linear regressions support an additive architecture of mantle coloration. Individuals with a higher number of white variants at mantle loci show increasingly white mantles (Fig. 2D and table S8). Entirely black mantles are expressed with highest probability only in individuals with fewer than 10 white alleles (out of 34 alleles or 17 mantle loci) (Fig. 2D). This suggests partial dominance of white mantle coloration.

The GWAS for neck side coloration identified overall weaker associations (Fig. 2A). The overlap of the identified SNPs with mantle loci (Fig. 2B and fig. S9C; see supplementary text) suggests that neck side coloration is an extension of mantle coloration associated with the same genetic variation. The observation that birds with black neck sides but white mantles are the most frequent hybrid phenotypes leads us to hypothesize that the penetrance of white is lower for neck sides than for the mantle, which may result from embryonic prepatterning by ASIP (*35*). In line with this, more white variation would be required at mantle loci to result in white neck sides than are required to result in a white mantle (see shifted intersection between Fig. 2, D and E). Henceforth, we focus on mantle coloration, which implicitly includes neck side coloration.

ASIP expression and isoform structures corroborate the genetic architecture of wheatear plumage coloration

To corroborate the statistically inferred genetic architectures of throat and mantle coloration, we investigated gene expression patterns in regrowing feathers. To this end, we chose *O. melanoleuca* and *O. cypriaca* as representatives of alternative colorations and performed single-molecule real-time isoform sequencing (Iso-Seq) of RNA for feather pools representing alternatively black or white coloration of throat and mantle, and, in addition, feathers from the shoulders and belly that are homogeneously black (shoulders) or white (belly) across species (Fig. 3A).

ASIP expression profiles of regrowing feathers corroborate ASIP's role in regulating wheatear plumage coloration. ASIP expression is about four times higher in white than in black feathers for the major isoform ($t = 3.6$, $P = 0.007$; adjusted coefficient of determination $R^2 = 0.57$; Fig. 3B; see supplementary text) also when accounting for the feathers' species of origin (color, $t = 3.9$, $P = 0.006$; species, $t = -1.5$, $P = 0.19$; adjusted $R^2 = 0.62$). This result holds when jointly analyzing all detected isoforms ($t = 4.6$, $P = 0.002$; adjusted $R^2 = 0.69$; accounting for species: color, $t = 5.6$, $P < 10^{-3}$; species, $t = -2.0$, $P = 0.08$; adjusted $R^2 = 0.78$; fig. S11). The higher expression in white than in black feathers is in line with ASIP's role in preventing melanin synthesis when acting early during feather growth (*33*).

Moreover, ASIP isoform structures confirm the role of the GWAS hits in ASIP regulation. Eight of the 17 mantle loci, including the top GWAS hits, are situated within or immediately upstream of the three alternative forms of exon 1 (fig. S12). This provides evidence for these genomic regions' involvement in ASIP regulation that is entirely independent of the statistical associations established by GWAS and establishes mantle loci as likely key regulators of ASIP expression in mantle feathers. For white throat color, all transcription start sites are situated further 5′ of ASIP exon 2 than the throat LTR. Together with the reverse complementary orientation of the LTR compared with the ASIP gene (Fig. 2B), this suggests that the LTR is unlikely to provide a transcription start site. Rather, the throat LTR acts as a (throat-specific) enhancer of ASIP expression, as documented for transposable element (TE) insertions upstream of ASIP in other species (*36*).

The patterns of gene expression and isoform variation recovered from regrowing feathers thus provide compelling support for the ASIP gene's key role in regulating melanin-based coloration in wheatears.

Introgression of throat coloration

Throat coloration is polymorphic in three species, with intermediate frequencies of the black and white phenotypes in *O. hispanica* and *O. melanoleuca* and with the black phenotype predominant (>98%) in *O. pleschanka* (*27*) (Fig. 1A). It was unclear whether this polymorphism arose from incomplete lineage sorting of ancestral variation or from introgression of variation that evolved following the species' divergence.

We established that the throat color polymorphism evolved by introgression of white ASIP variation among the throat-polymorphic species. Comparisons of times to the most recent common ancestor (TMRCAs) inferred in tsinfer (*37*) and tsdate (*38*) show that (i) the polymorphism evolved only after species diverged, with (ii) white being the derived state, and that (iii) the ASIP gene body shows distinctive footprints of introgression, with a divergence of the white ASIP haplotypes between *O. hispanica* and *O. melanoleuca* dating this introgression to 403 kyr ago, after *O. melanoleuca* split from *O. cypriaca* (Fig. 4A and tables S3 and S9; see supplementary text).

Next, we established that the white-throated phenotype originated in *O. melanoleuca*. Population genetic expectations are that (i) the species in which the white ASIP haplotype originated displays higher

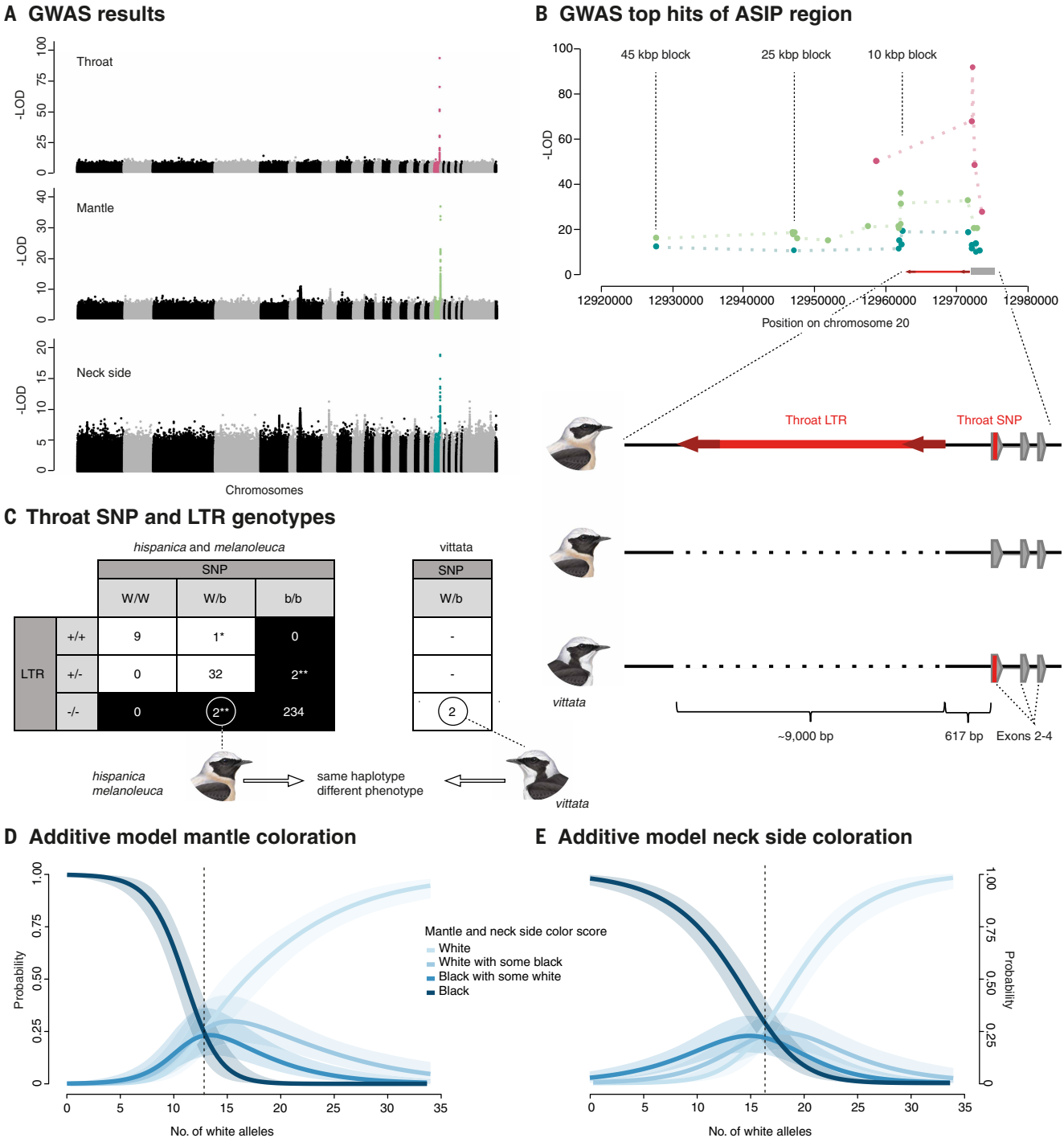


Fig. 2. Genetic architectures of plumage coloration. (**A**) GWAS results across the genome. Chromosomes are marked in black and gray; chromosome 20 is colored. (**B**) Distribution of GWAS top SNPs for the three color phenotypes within the ASIP region [colors follow those in (A)] and schematics of the haplotypes underlying throat color phenotypes in *O. melanoleuca* (top two) and the *vittata* phenotype of *O. pleschanka* (bottom). Arrowed boxes depict annotated ASIP exons. The throat LTR and SNP are highlighted in red. Absence of the throat LTR is depicted by dotted lines. (**C**) Genotypes of the throat SNPs and throat LTR. Boxes are shaded according to throat coloration. Recombinants in *O. hispanica* and *O. melanoleuca* are highlighted by asterisks: Recombinants supporting the interaction of the throat LTR and SNP in determining coloration are indicated with two asterisks, and recombinants not informative in this respect are indicated with one asterisk. All genotypes agree with the phenotypes predicted by a model in which the presence of the throat LTR (+/+ or +/−) up-regulates ASIP expression and, in the presence of a white throat SNP variant (W/W or W/b), confers white throat coloration, whereas an absent throat LTR (−/−) or homozygous black throat SNP (b/b) confers black throat coloration. See also table S6 for matches of observed with predicted phenotypes of the throat SNPs and lower-ranking GWAS hits. (**D**) Bayesian linear regressions of mantle variation. Shading shows 95% confidence intervals. The probability of observing a black mantle decreases, whereas the probability of observing a white mantle increases with an increasing number of white alleles (out of 34 possible at the 17 mantle loci) observed at mantle loci. Intermediate mantle colorations are most likely observed at intermediate numbers of white alleles. (**E**) Bayesian linear regressions of neck side variation. The probability of observing a black neck decreases, whereas the probability of observing a white neck increases with an increasing number of white alleles observed at mantle loci. Intermediate neck colorations are most likely observed at intermediate numbers of white alleles.

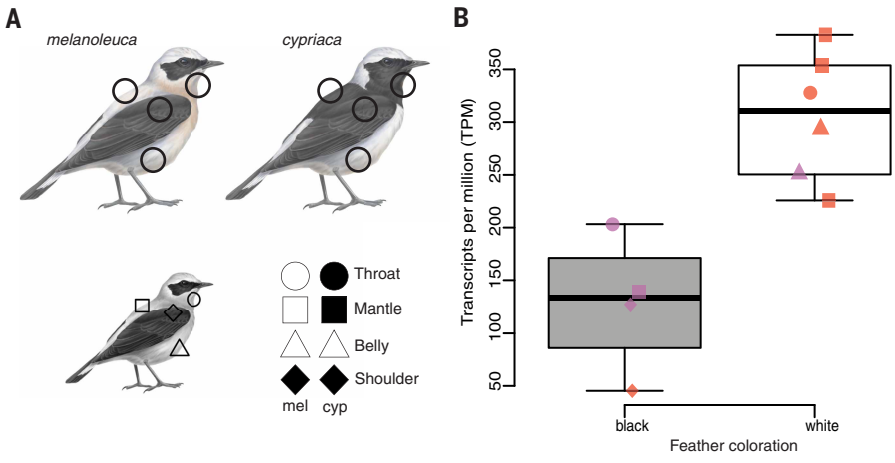


Fig. 3. Association of ASIP expression with feather coloration. (**A**) Design of the Iso-Seq experiment. Iso-Seq was performed once on each combination of feather tract, coloration, and species, except for white mantle, for which three pools from *O. melanoleuca* were sequenced (tables S16 and S17). mel, *melanoleuca*; cyp, *cypriaca*. (**B**) Association of the expression of the major ASIP isoform with feather coloration (TPM, transcripts per million reads). Symbols show individual values and follow those in (A); colors indicate species (red, *O. melanoleuca*; purple, *O. cypriaca*). ASIP expression is higher in white than in black feathers ($t = 3.6$, $P = 0.007$; adjusted R^2=0.57).

diversity (thus, has older TMRCAs) of the derived white ASIP haplotype, whereas, conversely, (ii) introgression increases both the diversity of black ASIP haplotypes and of overall ASIP diversity in the introgressed species. In line with an origin of the white haplotype in *O. melanoleuca*, we found an older TMRCA among white ASIP haplotypes in *O. melanoleuca* (395 kyr) than in *O. hispanica* (320 kyr; Fig. 4A, label 6, and table S9) and older TMRCAs among black ASIP haplotypes and across all ASIP haplotypes in *O. hispanica* (black haplotypes, 564 kyr; all haplotypes, 535 kyr) than in *O. melanoleuca* (black haplotypes, 554 kyr; all haplotypes, 510 kyr; table S10). Despite the confidence intervals of the white haplotypes' TMRCA overlapping between *O. hispanica* and *O. melanoleuca*, this result provides conclusive evidence for an origin of the derived white throat phenotype in *O. melanoleuca* (see supplementary text; fig. S13 and table S10). The white throat phenotype evolved through new protein-coding variation and insertion of the throat LTR in *O. melanoleuca* before being introgressed into *O. hispanica*.

The origin of the white-throated *vittata* phenotype in *O. pleschanka* was likely more intricate. For *O. pleschanka*, owing to the small sample size of white haplotypes ($n = 3$) and their recombined structure, we could not obtain the same level of evidence for introgression as observed in *O. melanoleuca*. Nonetheless, an introgressed origin, similar to that in *O. hispanica*, seems plausible, given well-documented hybridization with *O. melanoleuca* (*26*, *27*) (Fig. 1D and table S4). The absence of the throat LTR in *vittata* (fig. S10 and data S1) suggests that homologous recombination may have dissociated the white coding variant from the throat LTR, as also observed in the five recombinants marked with asterisks in Fig. 2C. In the absence of the throat LTR, it seems reasonable to assume that in *vittata* throat feathers, ASIP expression is up-regulated by undetected new genetic variation.

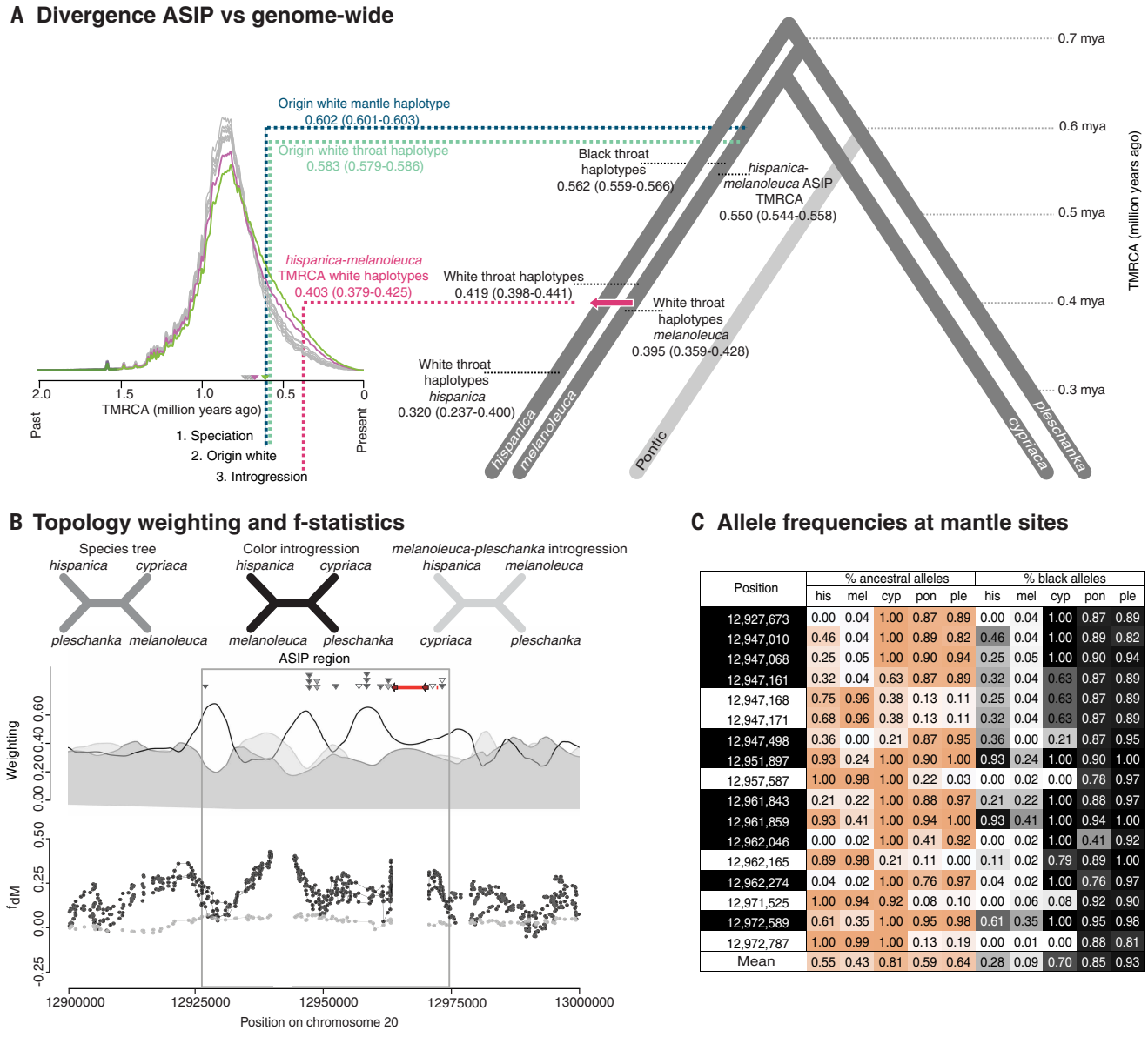
The successful introgression and maintenance of the throat color polymorphism suggest a role of throat coloration in wheatear ecology. Assortative and disassortative mating are unlikely to explain the polymorphism's maintenance: Contrary to expectations under such a scenario, we do not observe an excess of homo- or heterozygotes at the throat SNPs (table S11). Moreover, the two throat phenotypes co-occur on small spatial scales, and no differences in habitat or resource use are known despite rather well-studied species ecology (*27*). We therefore tested the hypothesis that the two throat phenotypes occupy subtly different microniches with different nutrient sources. To this end, we first analyzed triple-element (nitrogen, carbon, and sulfur) stable isotope composition in the whole blood of 161 males of *O. melanoleuca* ($n_{black} = 49$, $n_{white} = 43$) and Iranian hybrid ancestries ($n_{black} = 56$, $n_{white} = 13$). Blood isotope composition provides an integrated isotopic signature over several weeks, reflecting dietary intake and ecological interactions during this period. It can indicate differences in trophic niche [nitrogen, δ^{15}N; (*39*)]; habitat composition [carbon, δ^{13}C; (*40*)]; and foraging niche, dietary source, and habitat use [sulfur, δ^{34}S; (*41*)]. The alternative throat phenotypes showed no significant differences in δ^{13}C and δ^{34}S values (fig. S14). However, irrespective of whether analyzed for all individuals or *O. melanoleuca*, Greek, or Iranian males separately (table S12), white-throated males exhibited a narrower distribution as well as, on average, lower δ^{15}N values than black-throated individuals (Fig. 5 and fig. S15). This result persisted after accounting for individual and geographic correlates (table S12) and for unequal sample sizes of the two throat phenotypes (fig. S16). Second, we were interested in whether the same association persisted in *O. hispanica*, where the white throat phenotype introgressed from *O. melanoleuca*. Despite limited sample sizes ($n_{black} = 15$, $n_{white} = 8$), this analysis shows that white throats are associated with lower δ^{15}N values in *O. hispanica*, too (Fig. 5, fig. S16, and table S12). These results suggest that white-throated individuals prey on insects from a narrower and lower trophic level than black-throated individuals.

On the basis of these results, we speculate that the throat polymorphism is related to the trophic niche of prey and thus the wheatears' foraging strategy, such as observed in other birds: In barn owls, white underbody coloration aids hunting by prolonging prey freezing times in moonlight (*42*), whereas Nearctic warblers and redstarts display white feather tracts to flush prey and enhance foraging success (*43*, *44*). White-throated wheatears may be better at exploiting a specialized niche in a related manner, but further research is required to infer the adaptive role of the throat polymorphism. Alternatively, black-throated birds may be more aggressive (*45*) and exclude white-throated conspecifics from part of the species' niche. However, we deem such a scenario less likely, as it involves a lower fitness of the white throat phenotype and does not explain the polymorphism's maintenance and successful introgression. Therefore, and because the difference in nitrogen isotope composition between throat phenotype seems strongest in *O. hispanica*, the species into which white throat coloration successfully introgressed to high frequencies (Fig. 1A), we favor a scenario of niche specialization of the white throat phenotype. This would suggest that the recruitment of the white throat haplotype from *O. melanoleuca* into *O. hispanica* and *O. pleschanka* constitutes an event of likely adaptive introgression that enabled the exploitation of a specialized foraging niche.

Mantle color evolution by cointrogression of new variation

Mantle coloration is more similar between more distantly related lineages than between sister lineages of the *hispanica* complex (Fig. 1, A and E). We show that this pattern is rooted in the evolution of white mantle coloration in *O. melanoleuca* followed by its cointrogression with white throat coloration into *O. hispanica*.

Our results indicate that, originally, all members of the *O. hispanica* complex displayed black mantle plumage. Ancestral state reconstructions show that black variants were ancestral at 11 mantle loci at the node at which the *O. hispanica* complex split from *O. deserti* (node A, Fig. 1E). Furthermore, except *O. melanoleuca*, all species show moderate



Position	% ancestral alleles					% black alleles				
	his	mel	cyp	pon	ple	his	mel	cyp	pon	ple
12,927,673	0.00	0.04	1.00	0.87	0.89	0.00	0.04	1.00	0.87	0.89
12,947,010	0.46	0.04	1.00	0.89	0.82	0.46	0.04	1.00	0.89	0.82
12,947,068	0.25	0.05	1.00	0.90	0.94	0.25	0.05	1.00	0.90	0.94
12,947,161	0.32	0.04	0.63	0.87	0.89	0.32	0.04	0.63	0.87	0.89
12,947,168	0.75	0.96	0.38	0.13	0.11	0.25	0.04	0.63	0.87	0.89
12,947,171	0.68	0.96	0.38	0.13	0.11	0.32	0.04	0.63	0.87	0.89
12,947,498	0.36	0.00	0.21	0.87	0.95	0.36	0.00	0.21	0.87	0.95
12,951,897	0.93	0.24	1.00	0.90	1.00	0.93	0.24	1.00	0.90	1.00
12,957,587	1.00	0.98	1.00	0.22	0.03	0.00	0.02	0.00	0.78	0.97
12,961,843	0.21	0.22	1.00	0.88	0.97	0.21	0.22	1.00	0.88	0.97
12,961,859	0.93	0.41	1.00	0.94	1.00	0.93	0.41	1.00	0.94	1.00
12,962,046	0.00	0.02	1.00	0.41	0.92	0.00	0.02	1.00	0.41	0.92
12,962,165	0.89	0.98	0.21	0.11	0.00	0.11	0.02	0.79	0.89	1.00
12,962,274	0.04	0.02	1.00	0.76	0.97	0.04	0.02	1.00	0.76	0.97
12,971,525	1.00	0.94	0.92	0.08	0.10	0.00	0.06	0.08	0.92	0.90
12,972,589	0.61	0.35	1.00	0.95	0.98	0.61	0.35	1.00	0.95	0.98
12,972,787	1.00	0.99	1.00	0.13	0.19	0.00	0.01	0.00	0.88	0.81
Mean	0.55	0.43	0.81	0.59	0.64	0.28	0.09	0.70	0.85	0.93

Fig. 4. Introgressed origins of ASIP variation. (**A**) ARG-based reconstruction of of ASIP haplotypes' evolutionary history within the species tree (right) and distributions of interspecific TMRCAs (left). White haplotypes evolved after relevant species splits (blue and green dotted lines). The *O. hispanica–O. melanoleuca* ASIP TMRCA (upper right black dotted line) is more recent than these species' genome-wide TMRCA, indicating ASIP introgression. The TMRCA among white haplotypes is more recent than that among black haplotypes (upper and middle left black dotted lines) and establishes the white haplotype as derived. The TMRCA of white haplotypes in *O. melanoleuca* was older than that of white haplotypes in *O. hispanica* (lowest right and left black dotted lines) and establishes white coloration's origin in *O. melanoleuca*. The TMRCA between white haplotypes in *O. hispanica* and *O. melanoleuca* (pink dotted line) pinpoints the time of ASIP introgression. TMRCAs' confidence intervals are too small to be visualized (fig. S13 and tables S3 and S9). (**B**) Topology weighting with topologies and their interpretation (top) and f_{dM} tests for *O. melanoleuca–O. hispanica* introgression (bottom) variably including one of three black-mantled lineages (black) or exclusively black-mantled lineages (gray). Triangles highlight mantle loci that ancestrally featured black variation (black), polymorphic variation (gray), or white variation (white). Red blocks indicate the throat LTR and throat SNP. Regions harboring color loci show an excess of the introgression topology and footprints of *O. melanoleuca–O. hispanica* introgression (positive f_{dM}). The discrepancy between high f_{dM} in the comparison including *O. melanoleuca* (black) and ones only including black-mantled lineages (gray) consolidates evidence for introgression. (**C**) Frequencies of ancestral and black alleles at mantle loci. Darkening shades indicate increasing allele frequencies. Positions' colors indicate ancestral states. The high frequencies of derived variants and lowest frequencies of black alleles in *O. melanoleuca* and higher frequencies of ancestral black alleles in *O. hispanica* demonstrate the origin of white mantle variation in *O. melanoleuca* and introgression of the white variation into an originally black-mantled background in *O. hispanica*.

(*O. hispanica*) to high frequencies (all black-mantled species) of black variants at most mantle loci (Fig. 4C). The considerable frequencies of black variants at three ancestrally white sites in all species but *O. melanoleuca* (Fig. 4C) suggest that these were polymorphic in the ancestor of the complex. Given the additive genetic architecture (Fig. 2D), the ancestral populations would have likely displayed black mantle coloration.

Then, *O. melanoleuca* evolved derived white variation at the 11 ancestrally black mantle loci, and white variation introgressed into *O. hispanica*. *O. melanoleuca* features white variation at by far the

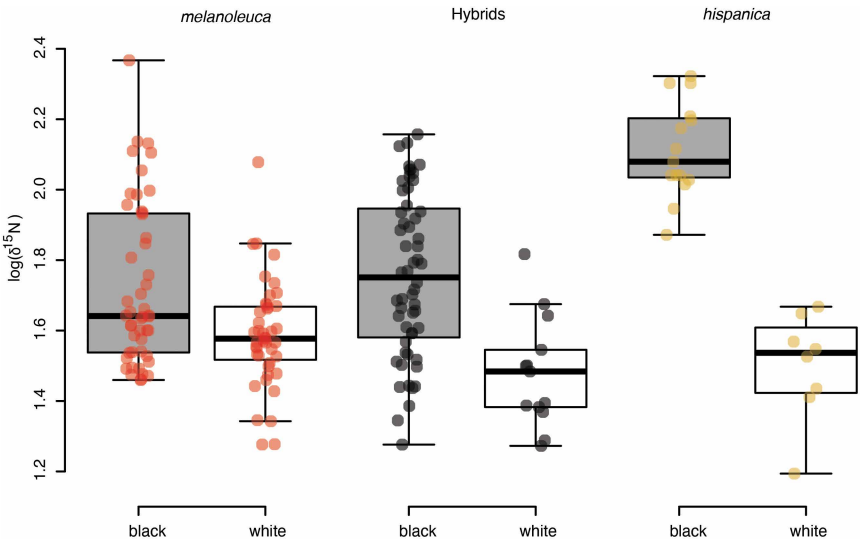


Fig. 5. Association of nitrogen stable isotopes with throat phenotype. Results are shown for analyses including all *O. melanoleuca*, hybrids, and *O. hispanica*. Points show individual values for *O. melanoleuca* (red), hybrid (black), and *O. hispanica* (yellow). All comparisons are significant (*O. melanoleuca*, $t = -3.2$, $P = 0.002$; hybrids, $t = -4.1$, $P < 10^{-3}$; *hispanica*, $t = -8.3$, $P < 10^{-5}$; table S12). Additional comparisons including Greek *O. melanoleuca* separately, all *O. melanoleuca* and hybrids together, and all individuals together are shown in fig. S15.

highest frequencies (91 versus 72% in *O. hispanica*; Fig. 4C). The considerable frequencies of black variation in *O. hispanica* (Fig. 4C) suggests that this species was originally black-mantled but largely swamped by white *O. melanoleuca* variation. This scenario is supported by the same evidence as introgression of throat variation. The genomic regions underpinning mantle coloration display predominantly topologies related to mantle coloration (Fig. 4B, figs. S17 and S18, and table S4); the TMRCAs between black and white haplotypes upstream of ASIP are younger than TMRCAs between species, and the white haplotypes' coalescence between *O. melanoleuca* and *O. hispanica* is younger than these species' split (Fig. 4, figs. S13 and S17, and tables S3, S9, and S10).

Several lines of evidence suggest an involvement of selection in mantle color evolution, especially in *O. melanoleuca*. Compared with the site frequency spectra (SFS) polarized according to ancestry, in all species but *O. cypriaca*, SFS polarized according to coloration show shifts toward the species' coloration (Fig. 6A). The strongest shifts are observed in *O. melanoleuca*, and conspicuous haplotype structures and TMRCAs corroborate an involvement of selection in driving white variation toward fixation, especially upstream of the most distal mantle loci (Fig. 6B, label 1, and fig. S19). Species other than *O. melanoleuca* do not feature comparable footprints of positive selection (Fig. 6B and fig. S19), likely because of a predominant involvement of ancestral variation, and, in *O. hispanica*, introgression.

We conclude that white mantle coloration evolved adaptively in *O. melanoleuca* and introgressed into *O. hispanica* along with the linked white throat variation.

Multistage introgression underpins color convergence and polymorphism

Within the *hispanica* complex, the evolution of species' plumages as we know them today predominantly roots in past introgression. Our results suggest that originally all members of the *hispanica* complex displayed black throat and mantle coloration before white mantle coloration evolved in *O. melanoleuca* through mutations upstream of ASIP that were fixed by positive selection (Fig. 7A). On this white-mantled genomic background, new protein-coding ASIP variation together with an LTR retrotransposon insertion upstream of ASIP led to the emergence of the white-throated phenotype in *O. melanoleuca* (Figs. 4A and 7A). White throat coloration then introgressed into *O. hispanica*, where, as in *O. melanoleuca,* it is maintained as a balanced polymorphism. Possibly owing to its linkage to white throat variation, white mantle variation cointrogressed into *O. hispanica*, where, due to its additive architecture with a partial dominance of white, it confers white mantle coloration despite still segregating black mantle variation (Fig. 4C). Footprints of positive selection (Fig. 6B, label 1), though much subtler than in *O. melanoleuca*, suggest that mantle color introgression might also have been supported by selection in *O. hispanica* (Fig. 6B). In *O. pleschanka*, hybridization with *O. melanoleuca* resulted in the introgression of white coding variation at ASIP, but not cointrogression of variation 5′ of ASIP. We interpret this in terms of selection against introgression of white mantle coloration; nearly fixed derived black variation at three of the mantle loci in *O. pleschanka* and the Pontic lineage (but not *O. cypriaca*) along with slightly longer–than–average haplotypes in the region flanked by two of these sites suggest that black mantle coloration is under positive selection in *O. pleschanka* and the Pontic lineage (Figs. 4C and 6B, label 2). For white throat variation, only the throat variant successfully introgressed into *O. pleschanka*. Introgression thus only planted the seed for the white-throated *vittata* phenotype. It thus appears that introgression shuffled plumage coloration among wheatear species and led to the convergent evolution of throat and mantle color phenotypes, similar to the reshuffling of morphologic or color traits documented from cichlids, *Heliconius* butterflies, Capuchino seedeaters, and Darwin's finches (*3*–*5*, *13*, *14*).

Introgression further contributed to the formation of intermediate phenotypes observed in extant hybrid zones (Figs. 7, B and C), where hybridization and backcrossing recombine ASIP variation into recombined haplotypes and color phenotypes (Figs. 1B and 7, B and C, and data S2). Especially Iranian and Caucasian individuals of mixed ancestries highlight how the fine-scale reshuffling of mantle and throat variation into recombined ASIP haplotypes directs the expression of intermediate phenotypes (Fig. 7C and data S2). Different from these recombined phenotypes of genome-wide admixed individuals, the Pontic and Caspian populations feature white-mantled to *O. melanoleuca*–type (white mantle with white neck sides) phenotypes in individuals with entirely *O. pleschanka* or Pontic ancestries (Fig. 7). This mismatch between phenotype and genome-wide ancestry takes its extreme in an *O. melanoleuca* phenotype individual from the Caspian population whose genome-wide ancestry exactly matches those of *O. pleschanka* phenotypes from the same region (no. 11, Fig. 7). Like similar Pontic and Caspian individuals, this individual displays ASIP haplotypes matching its coloration. This suggests that introgression is restricted to a small proportion of the genome. Accordingly, compared with Pontic and Caspian *O. pleschanka*–type phenotypes, *O. melanoleuca*– and hybrid-type individuals of these populations show higher proportions of topologies congruent with color (*O. melanoleuca*) introgression, mostly in the ASIP region (Pontic, 11.8% more topologies congruent with introgression in *O. melanoleuca* and hybrid phenotypes than in parental phenotypes; Caspian, 6.1%), much less across chromosome 20 (Pontic, 1.7%; Caspian, −0.3%), and not at all genome wide (fig. S20). This suggests that the plumage color polymorphism in these populations is a result of genomically very localized and likely old introgression of white coloration, similar to introgressed wing-pattern variation in *Heliconius* butterflies (*4*, *46*).

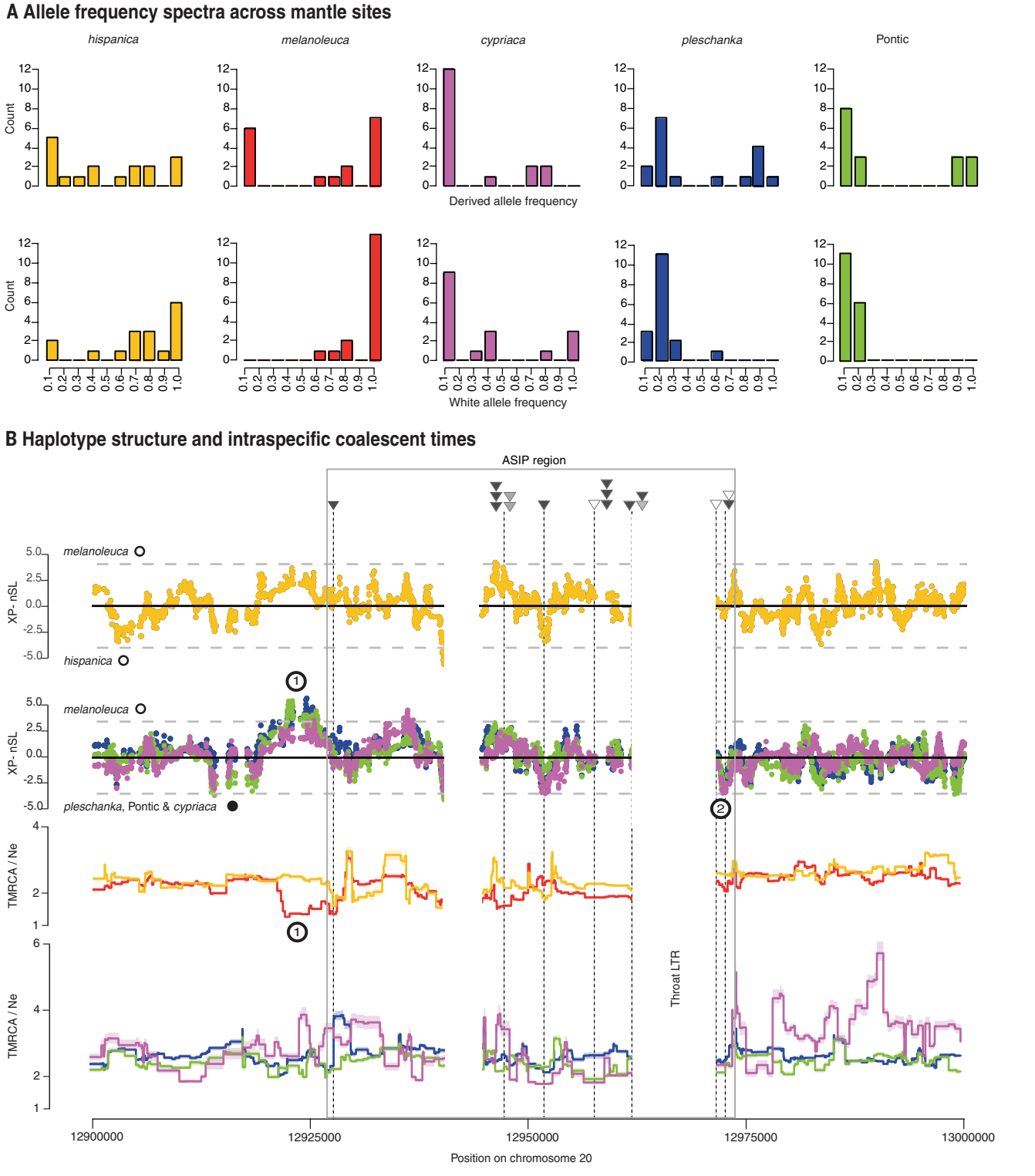


Fig. 6. Evidence for positive selection on ASIP variation. (**A**) Unfolded site frequency spectra for the 17 mantle loci polarized according to ancestry (top) or mantle coloration (bottom). Stronger shifts toward species' colorations than toward ancestral or derived variants (notably strong in *O. melanoleuca*, absent in *O. pleschanka* and the Pontic lineage) are in line with positive selection on coloration. (**B**) (Top) Cross-population haplotype statistics [XP-nSL; (*94*)] between *O. melanoleuca* and *O. hispanica* and between *O. melanoleuca* and all black-mantled species (mantle coloration is indicated by the color of circles besides the species labels). Positive and negative peaks indicate selection in the lineage denoted to the left of the panels. Colors indicate the species to which *O. melanoleuca* was compared (colors follow those in Fig. 1A). (Bottom) ARG-based TMRCAs (*37*,*38*) within species scaled by effective population sizes (estimated based on nucleotide diversity), with low values indicative of positive selection. Footprints indicating positive selection in *O. melanoleuca* (labels 1) and (weaker) in the region flanked by two derived black variants in *O. pleschanka* and the Pontic lineage (label 2) are highlighted. Dashed lines show the positions of mantle loci, with triangles at the top indicating sites that, in the complex's ancestor, were black (black), polymorphic (gray), or white (white). Lines depict multiple mantle loci where the latter are in close proximity.

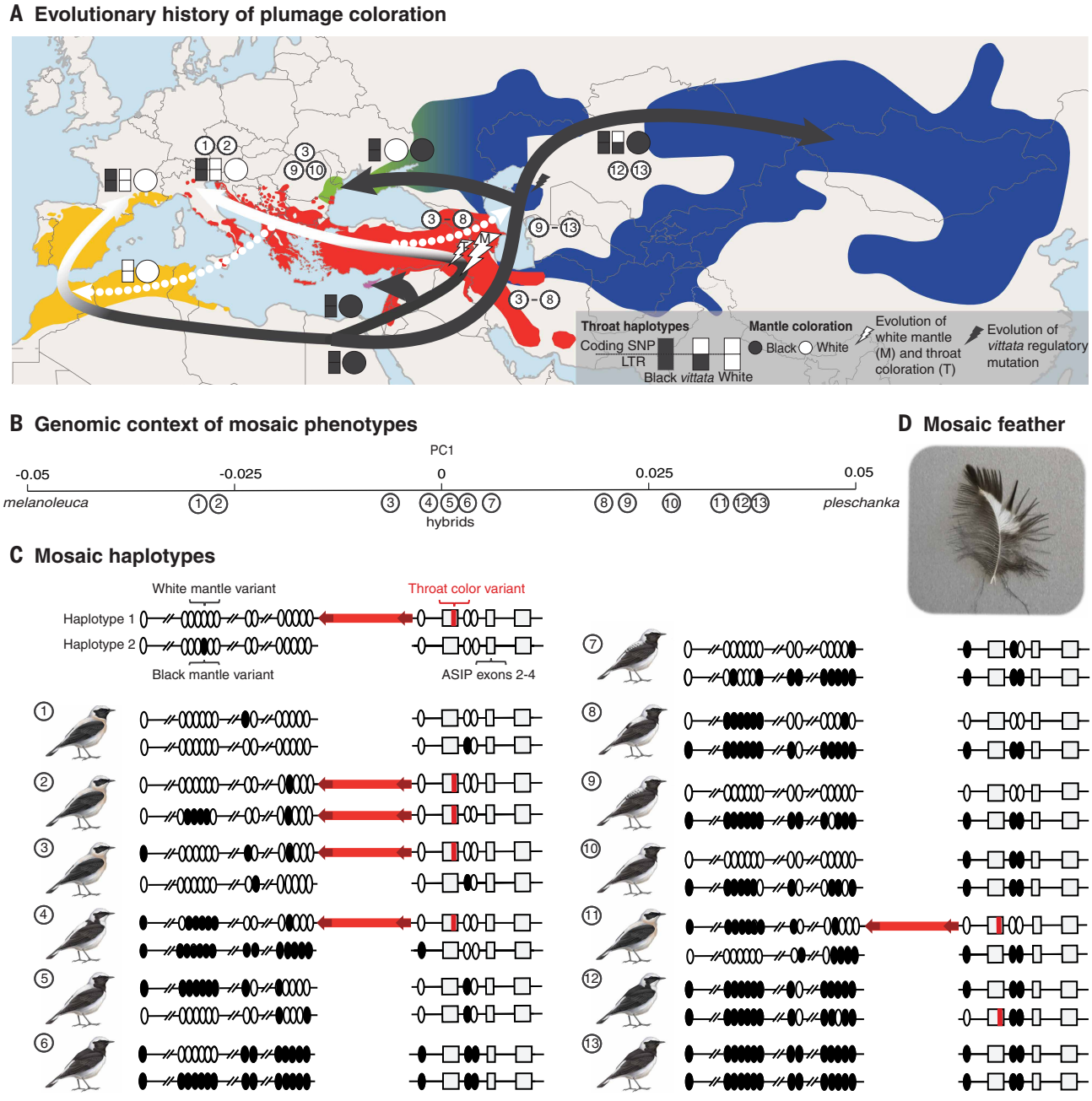


Fig. 7. Evolutionary history of wheatear coloration and recombined phenotypes and haplotypes. (**A**) The evolutionary histories of throat and mantle coloration are shown by arrows in the distribution maps (colors follow those in Fig. 1A). Black arrows depict ancestrally black mantle and throat coloration in *O. pleschanka*, the Pontic lineage, and *O. cypriaca*, with mantle coloration fixed in all three lineages. White mantle and throat coloration evolved in the *O. melanoleuca* ancestor after the split from *O. cypriaca* (white lightning arrows, which transition to white arrow in *O. melanoleuca* range); white mantle coloration is fixed in this species, whereas throat coloration is polymorphic. Then white coloration introgressed from *O. melanoleuca* into *O. hispanica* (dotted arrow), leading to fixed white mantle and polymorphic throat coloration. Further east, the throat SNP introgressed from *O. melanoleuca* into *O. pleschanka* (dotted arrow) and, likely together with a new regulatory mutation (black lightning arrow), led to the evolution of the *vittata* phenotype. Locations where phenotypes depicted in (C) can be observed are indicated by circled numbers. Arrow lengths are not at scale with evolutionary timings. Although the evolutionary processes are projected on current species distributions, geographic locations are not to be interpreted. (**B**) Phenotypes' typical ancestries (several possible) are shown on PC1 for phenotypes and haplotypes from (C). (**C**) Selected range of phenotypes together with examples of the underlying ASIP haplotypes. Ovals depict mantle loci, throat variation is shown in red, and ASIP exons are shown as squares. The depicted haplotypes illustrate the fine-scale recombination of ASIP variation. At mantle loci, variation blends into a continuum of almost exclusively white (number 1), over different degrees of alternating black and white (numbers 2 to 12), to exclusively black variation (number 13). White-throated individuals exhibit at least one haplotype containing the throat LTR (red arrows) and white throat SNP variant (red box), except *vittata* (number 12). (**D**) Mosaic black-and-white back feather, as observed in various hybrid-type phenotypes.

Color mosaics can be observed down to the level of single feathers. In intermediate phenotypes, single mantle feathers commonly display a mosaic of black-and-white coloration (Fig. 7D). This suggests that mosaic ASIP haplotypes may result in unstable ASIP expression.

New mutations underpin genus-level convergence

The plumage coloration in the *hispanica* complex evolved convergently also at the genus level and beyond (*22*, *23*). Therefore, our final interest was to understand whether the same variants as within the *hispanica* complex underpin convergent evolution across all wheatears and related species. To this end, we interrogated alignments of whole-genome resequencing data of 50 individuals from 44 species of wheatears and their close relatives (*22*). At all sites associated with throat and mantle coloration (including the throat LTR) for which data was available, we only observed the ancestral variants in species outside the *hispanica* complex (fig. S21). This highlights that the variation underpinning plumage coloration at the genus level differs from the variation associated with these colorations in the *hispanica* complex. Thus, consistent with results from other bird species (*2*, *12*, *15*), whereas the shorter-term color evolution in the *hispanica* complex is underpinned by interspecific exchange of genetic variants, the longer-term convergent evolution of plumage color diversity beyond the *hispanica* complex is due to new mutations (*19*).

Evolution of recombined phenotypes within the constraints of a monogenic architecture

Wheatear coloration shows how phenotypes can be decoupled between body parts despite being regulated by a single gene. The genetic architecture involving modular regulation of a single gene sets wheatear coloration apart from what is known from other bird systems, where recombined color phenotypes are underpinned by multiple genes (e.g., *2*, *12*, *47*) or a single gene produces uniform coloration across multiple feather tracts (e.g., *11*, *48*, *49*). In the *hispanica* complex, independent colorations of multiple body parts are determined by distinct regulatory modules of ASIP, similar to that in mammals (*17*, *34*, *50*). Also, the grouping of mantle loci in blocks (Fig. 2B) and the alternative transcription start sites observed among isoforms (fig. S12) reflect a pattern known from mammals, where this pattern represents alternative untranslated first ASIP exons (*17*, *50*). The involvement of a TE in ASIP regulation was hitherto only hypothesized for bird coloration (*51*) but has been documented for mammalian ASIP regulatory modules (*36*, *51*). The modular regulation of wheatear coloration is thus similar to what is documented from mammals and insects (*9*, *17*, *50*, *52*) and reflects a likely universal modular regulatory architecture of ASIP that enables decoupled melanin expression between body parts.

However, modular gene regulation alone may not always grant independent evolution of the associated phenotypic elements. The supposed evolutionary history of the *vittata* phenotype suggests that the interplay of modules' linkage with selection against introgression may have prevented opportunities for the recombination required to decouple the evolutionary histories of different body parts. We deem it likely that, in *O. pleschanka*, selection against white mantle coloration prohibited the introgression of the throat LTR that is usually linked to white mantle variation. Therefore, selection limited the opportunity for recombination of the throat LTR onto haplotypes carrying black mantle variation in *O. pleschanka*. Consequently, the evolution of the *vittata* phenotype appears to have required a new element up-regulating ASIP expression in throat plumage. Hence, with selection limiting introgression and opportunities for recombination, the modularity of ASIP expression was not sufficient to completely resolve the pleiotropy between throat and mantle coloration.

For preexisting variation to contribute to phenotypic diversity underpinned by a single gene, sufficient opportunity for recombination to decouple regulatory modules turns out to be key. Together with previous insights from dogs, ladybird beetles, and *Heliconius* butterflies (*9*, *17*, *18*), the evidence from wheatears establishes that, in the long run, even within the constraints of monogenic architectures, fine-scale recombination can integrate linked regulatory modules from divergent origins into new recombined haplotypes. This appears to imperatively involve long-term recombination, be it within species (*9*, *17*, *18*) or their ancestors (*17*) or in hybrid zones (wheatears). In this context, the almost seamless integration of black and white back variation upstream of ASIP in hybrid wheatears (Fig. 7C) demonstrates the power of combinatorial processes in the evolution of phenotypic diversity and underscores the importance of hybrid zones as melting pots that enable evolutionary tinkering.

Convergent plumage color evolution involved diverse types and origins of genetic variation

The patterns, processes, and genetic material involved in wheatear color evolution paint an intricate evolutionary history, exemplifying that "either-or"–type questions that shaped evolutionary genomics for over a decade (*53*–*56*) may rarely find single, entirely black or white answers. Wheatear color evolution involves a single gene, yet different regions associated with this gene; it involves preexisting and new variation, protein coding, and regulatory variation (*53*–*55*), and it is based on both point mutations and structural variation (*51*, *56*). This highlights that the diversity of wheatear coloration is underpinned by a mosaic of genetic variation, evolutionary origins, and evolutionary histories.

We emphasize that a study system spanning variable evolutionary timescales and hybridization contexts was key to obtain insights into evolution from preexisting variation beyond the latter's contributions to phenotypic diversity. Our findings suggest that the role of preexisting variation in wheatear plumage color evolution was limited by its availability and the constraints imposed by a monogenic architecture. First, extant hybrid zones showcase how long-term recombination can break up even tightly linked variation at the fine scale required to reshuffle modular variation into recombined haplotypes. Yet, the resulting phenotypes do not contribute to plumage color diversity outside hybrid zones, such as it was achieved by *O. melanoleuca*–*O. hispanica* hybridization about 400 kyr ago. Second, past hybridization with *O. melanoleuca* had distinct outcomes in *O. hispanica* and *O. pleschanka*. It shaped all color variation of extant *O. hispanica* but contributed only one component for the convergent evolution of the white-throated *vittata* phenotype in *O. pleschanka*. At longer timescales at which preexisting variation was no longer available—that is, across the genus—plumage color evolution was instead driven by new mutations. Similarly, new mutations seem to have been involved where contributions of preexisting variation were likely constrained by linked selection, as seen in *vittata*. Together, these results offer a nuanced perspective on how genetic architectures constrain phenotypic evolution from preexisting variation and, more broadly, how genetic architectures can shape evolutionary trajectories over time.

The detailed insights into the roles and interactions of genetic architecture and diversification processes in phenotypic evolution gained from wheatear coloration may be relevant to our understanding of the reshuffling of preexisting and new genetic variation during the diversification of a wide range of organisms. Our results show that evolution recruits variation from all sources it has at hand, genetic architecture permitting. At deeper timescales, mutations were both frequent enough and required to give rise to convergent plumage phenotypes. Future comparisons of evolutionary dynamics across a wide range of phenotypes and organisms will be required to understand whether this is only possible due to extraordinarily dynamic phenotypes and genetic architectures, or, indeed, mutations may be less limiting than assumed to date (*57*), and/or the regulatory effects of new TE insertions play an even more important role in phenotypic evolution than already commonly acknowledged (*51*, *58*). For example, we suspect that, in line with TEs' effects on color evolution across the

animal kingdom (*51*), extended studies of structural variation across the genus will reveal an important role for TEs in the genetic architecture of coloration in other wheatears and their relatives. Still, preexisting variation and permeable species barriers in wheatears, as in many organisms (*59*), prove to provide abundantly used substrate for adaptation and phenotypic evolution. This highlights humankind's responsibility to conserve a maximum number of species and genetic variation to equip biodiversity with the molecular prerequisites to adapt to the increasing challenges encountered in the Anthropocene.

Materials and methods

Whole-genome resequencing

To obtain a representative sampling of genetic and phenotypic variation across the *hispanica* complex, we collected samples from the major parts of the species' distribution (Fig. 1C). To optimize the power of admixture mapping, hybrid zones were densely sampled. During fieldwork dedicated to the present project, we collected whole blood into >96% ethanol, Queen's lysis buffer, or DNA/RNA Shield (Zymo). Samples obtained from museum collections were either blood or muscle tissue in ethanol, or toe pads collected from skins (table S1).

We extracted DNA from blood and tissues using the DNeasy Blood and Tissue Kit (Qiagen) or the MagAttract HMW DNA kit (Qiagen) following the manufacturer's protocols, with the exception of an adapted digestion of blood samples (*60*). We extracted DNA from toepads using the QIAmp DNA Micro Kit (Qiagen) with a digestion protocol adapted to ensure sufficient DNA quantities (https://www.protocols.io/view/dna-extraction-protocol-for-historical-toe-pad-sam-dm6gpwrdplzp/v1). For quality control, DNA concentrations were determined with a Qubit fluorometer (dsDNA BR assay, Thermo Fisher Scientific), and DNA integrity for linked-read sequencing was evaluated on a TapeStation with the Genomic DNA Screen Tape (Agilent) or on a FEMTO Pulse CE instrument (Agilent). To obtain high quality gametic phase information for a reference haplotype panel, we sequenced 238 individuals using linked-read technology (*60*). Linked-read sequencing libraries were prepared using Chromium Genome library kits (10X Genomics) (table S1). Short-read sequencing libraries for 139 additional individuals for which fresh material was available were prepared with either the Illumina DNA Prep Kit, the Illumina DNA PCR-free kit, or following Baym *et al.* (*61*); sequencing libraries for 12 individuals with degraded DNA were prepared with the xGen ssDNA & Low-Input DNA Library Prep Kit (Integrated DNA Technologies) (table S1). Apart from the seven libraries reported in Lutgen *et al.* (*60*) that were sequenced on an Illumina HiSeq-X instrument, all libraries were sequenced on an Illumina NovaSeq6000 instrument with a target coverage of ca. 15x. After the removal of individuals with a coverage below 8x, this resulted in a dataset comprising a total of 379 individuals (tables S1, S13).

Data preparation

To prepare whole-genome resequencing data for analysis, linked-read data and standard short read data were first treated separately, following the GATK 4.1.4.1 (*62*) best practices pipeline where appropriate. We aligned linked reads to the *O. melanoleuca* reference genome (*63*) using phase-aware alignment in Lariat as implemented in LongRanger 2.2.2 align (*64*). To obtain a SNP exclude set for base quality score recalibration (BQSR), SNPs were then called and genotyped in GATK with HaplotypeCaller as depicted in fig. S22. To include only high-confidence SNPs in the SNP exclude set for BQSR, we retained only SNPs with mapping quality > 40, Fisher strand phred-scaled p-value >60, SNP quality score >20, mapping quality rank sum value > -12.5, read pos rank-sum test value > -8.0 and quality by depth > 5. We retained only biallelic SNPs with at least one homozygous reference and one homozygous alternative genotype or with a minimum of three observations of reference or alternative alleles. BQSR was performed in GATK followed by final SNP calling with HaplotypeCaller. Short reads were adapter-trimmed in fastp 0.20.0 (*65*) before mapping them to the *O. melanoleuca* reference genome (*63*) using bwa-mem 0.7.17 (*66*). We then marked duplications with PicardTools 2.9.1 (http://broadinstitute.github.io/picard) and directly performed BQSR with the SNP exclude set obtained from the geographically and taxonomically representative linked-read data (n=235, table S1). SNPs were called with HaplotypeCaller. To obtain a joint VCF file with unphased SNP data for linked-read and short-read data, we imported the resulting gVCF files into a single genomic database and genotyped SNPs for both data types. The resulting VCF constituted the basis for analyses not requiring gametic phase, such as GWAS.

To infer gametic phase for phase-aware analyses, we leveraged the linked-read information from 235 individuals to establish a high-quality reference haplotype panel (*60*) with which we then supported the statistical phasing of short-read data. To this end, for each individual we phased heterozygous genotypes obtained in the last step either with (i) linked-read-barcode data using HapCUT2 (*67*) for individuals with linked reads available or (ii) paired-end information using whatshap v.1.1 (*68*) for the remaining individuals. To include all originally called SNPs in the individual VCF files, we parsed the phased and unphased data for each individual to reintroduce SNPs homozygous in the individuals. We then parsed individual vcf files into either of two population vcf files, separated by linked-read and short-read data. To establish the final reference haplotype panel, for linked-read data we retained only high-quality SNPs that had a minimum sequencing depth of 5 and minimum genotype quality of 20 in at least 75% of individuals. This left a total of 69,855,099 SNPs (table S13). We then established the gametic phase of phase sets statistically using Shapeit 4.2 (*69*), assuming that phase sets established beforehand based on linked-read info are correct (–use-PS 0). Finally, we introduced the resulting haplotypes as reference haplotype panel into the statistical phasing of short-read data in Shapeit (–use-PS 0.0001). To avoid biases introduced by population stratification we performed this statistical phasing separately for each of the 11 populations (one *cypriaca*, four *melanoleuca*, two *pleschanka*, two Pontic, and two Caucasus populations) using a specific reference panel consisting of individuals from populations closest related as determined from the PCA (Fig. 1D). Short-read data were not filtered prior to statistical phasing, as with the reference haplotype panel and an average sequencing coverage of 18 × per individual, imputation of missing data are expected to yield adequate genotypes.

For analyses that require an outgroup, we added unphased data from three *O. deserti* individuals to the phased dataset. For each of the three *O. deserti*, we generated allsites vcf files in GATK and extracted genotypes for the sites contained in the phased dataset using GATK's SelectVariants. Using bcftools we then merged the outgroup data with the phase dataset and removed *O. deserti*-specific indels that contained no data in the phased dataset. To reconstruct the ancestral states of the *hispanica* complex and the *O. deserti* outgroup, additional allsites vcf files were generated for two *O. monacha* and three *O. oenanthe* serving as distant outgroups.

A characterization of all data sets used for analysis is provided in table S13. Individuals used for specific purposes or analyses are listed in table S14.

Reconstruction of the *hispanica* complex's evolutionary history

We first aimed to obtain insights into the genetic structure of the *hispanica* complex without prior assignment of individuals to populations. To this end, we performed a PCA on genotype frequencies of all 379 individuals using smartpca (*70*) and inferred co-ancestries among individuals and based on this individual clustering in fineSTRUCTURE (*28*).

For PCA, we used the SNPs contained in the genome-wide phased data set. To exclude data affected by poor mapping in repeat regions, selection on genes, and linkage disequilibrium, we removed (i) repeats based on the annotation using the repeat library from Peona *et al.* (*63*),

(ii) gene regions using the annotation bed file published in Peona *et al.* (*63*), and (iii) regions with high linkage disequilibrium (–indep-pairwise 50 5 0.2) using plink v.1.9 (*71*). To ensure that the PCA is not biased by our sampling's high load with individuals from hybrid zones, in addition to the standard PCA including all individuals (fig. S23), we ran smartpca first with 12-15 individuals per each of the five taxa (*hispanica:* 14; *pleschanka:* 15; Pontic: 15; *melanoleuca:* 15; *cypriaca:* 12), with individuals chosen from the core clusters of the former PCA (hereafter referred to as "reference individuals;" fig. S24) and projected the genotypes of the remaining individuals onto the inferred axes using smartpca. Two individuals were strong outliers, with each producing an axis separating it from all other individuals (not shown). We therefore removed these individuals from the final analyses. To estimate the level of differentiation between lineages, we computed pairwise differentiation based on Weir and Cockerham's F_{ST} implemented in vcftools (*72*).

To investigate how ancestries transition across the hybrid zones (Iran, Caucasus, Black Sea, and Caspian), we performed one-dimensional sigmoid geographic cline analyses in the HZAR package in R (*73*). To this end, first, we for each hybrid zone first identified the principal component axis (PC) that best separates parental lineages. The respective PC axes were then used as a proxy for hybrid index by rescaling PC scores to between 0 and 1. Second, we estimated Bayesian hybrid indices in bgchm (*74*) To this end, we for each lineage pair extracted ancestry-informative markers (AIM), with AIM defined as SNP with F_{ST} as estimated in vcftools (*72*) exceeding 0.7. For comparisons exceeding 10,000 AIM, we randomly sampled 10,000. This resulted in each 10,000 AIM for *melanoleuca-pleschanka* across Iran, *melanoleuca-pleschanka* across the Caucasus, and *melanoleuca*-Pontic (Black Sea), and 4,369 AIM for the Pontic-*pleschanka* (Caspian) comparison. Based on these AIM we then estimated Bayesian hybrid indices using the est_hi function in bgchm (*74*), using reference individuals as parentals (table S14, fig. S24). For each hybrid zone, we then projected samples along a one-dimension transect using the distm function implemented in the geosphere R package (https://github.com/rspatial/geosphere). Specifically, we calculated the great circle distance between each individual and the westernmost parental individual averaging across samples for the latitude. We assigned samples to populations along the transect by grouping individuals within a 100 km radius into a single population. Next, we fit three geographic cline models with HZAR using hybrid indices per sampling location as the dependent variable and the geographic distance along the one-dimensional transect (in 100 km) as predictor. We fitted clines using the default MCMC chain length of 100,000 steps with 10,000 burn-in steps and selected the best model by comparing the support for then null model and the three cline models using AIC. We plotted the best model having an AIC threshold score >2.

To investigate admixture and infer population clusters based on individuals' co-ancestries we leveraged our phased data to trace excess haplotype sharing using the powerful approach implemented in fineSTRUCTURE v.4.1.1 (*28*). We visualized the results of this analysis as a clustering tree, haplotype frequency-based PCA, and co-ancestry matrix. For fineSTRUCTURE analyses we corrected excess relatedness among individuals by removing each one individual from nine individual pairs with above second degree relatedness (>0.2; vcftools -relatedness2), because their inclusion would bias population clustering given the extremely high excess of recent haplotype sharing, and we thinned the phased whole-genome dataset using vcftools (*72*) to retain 1 SNP every 500 base pairs (-thin 0.5). Thinning was performed to reach achievable computing times, with parameters based on a comparison of fineSTRUCTURE runs using complete and thinned data sets from (micro)chromosomes 15-28. FineSTRUCTURE was run with a recombination map estimated from *melanoleuca* data using pyrho v.0.1.7 (*75*). Because population structure can confound LD-based recombination rate estimates, we restricted recombination rate estimation to a single population of *melanoleuca* from Greece (N=13). Population size trajectories were estimated with MSMC2 for five individuals from this population to account for demography within the pyrho inference. Optimal hyperparameters were determined through 50 rounds of simulations for all possible combinations of parameters from block penalties of 10, 15, 20, 25, 30, 35, 40, 45, and 50 and window sizes of 30, 40, 50, 60, 70, 80, and 90. A per-generation mutation rate of $4.6 \cdot 10^{-9}$ from collared flycatcher (*76*) was applied to convert population-scaled recombination rate to cM/Mb. The pyrho output was converted to genetic map format with a custom script, and the genome-wide average recombination rate was calculated by taking the average rate between SNP pairs, weighted by the physical distance between SNPs.

To reconstruct the *hispanica* complex's speciation history, we inferred the species tree and estimated demographic trajectories and split times based on the multiple sequentially Markovian coalescent (MSMC) and based on the ancestral recombination graph (ARG). We estimated the species tree based on non-overlapping 10 kbp windows in the reference individuals. Data from three *O. deserti* were included to root the species tree. We only considered windows meeting the following criteria, following Alaei Kakhki *et al.* (*22*): minimum read depth of 5, minimum 50% of the window covered by data, and maximum 20% missing data per site (n=79,127). To include only windows without recombination but free inter-locus recombination, we excluded all windows with recombination signals ($P \leq 0.05$) as inferred from the pairwise homoplasy index Phi (Φw) estimated in PhiPack 1.1 (*77*) and enforced a minimum distance of 10 kbp between consecutive loci, as at this distance no linkage disequilibrium occurs in flycatchers (*78*) (N=29,792). Based on this data, we estimated the species tree with multiple individuals forced together in the species tree (-a option; Fig. 1E), or keeping all individuals as leaves (without the -a option; fig. S3) using the multispecies coalescent in ASTRAL-III (*79*) based on maximum likelihood gene trees estimated in IQ-TREE 2.1.2 (*80*) using a GTR+I+G substitution model and one thousand ultrafast bootstrap approximations. We estimated the MSMC using MSMC2 (*81*) using data of five individuals of each of the five taxa. Where possible, we included individuals with linked-read data available to ensure the best phase information. To make use of the phase information, we first generated sample-specific VCF and mask files using the bamCaller.py from msmc-tools, and into the resulting VCF files injected phase information from our phased VCF file. SNPs without phase information were excluded. MSMC2 was run to evaluate coalescence between haplotypes within individuals for each population (-I) with default time intervals (-p $1 \times 2 + 25 \times 1 + 1 \times 2 + 1 \times 3$) and a ratio of recombination over mutation rate (-r) of 5.65 based on the flycatcher mutation rate of $4.6 \cdot 10^{-9}$ per bp per generation (*76*) and a mean autosomal recombination rate of $2.6 \cdot 10^{-8}$ (as estimated for *melanoleuca*, see above). To convert generations into years, we assumed a generation time of two years.

To infer the relative timing of lineage splits based on coalescence times, we leveraged the power of phased data and ARGs. ARGs were inferred from phased data using tsinfer v. 0.3.1 (*37*) for the reference individuals. We ran tsinfer with a mismatch ratio of 0.1 and provided a uniform recombination rate for each chromosome, using the mean recombination rate estimate for *melanoleuca* (table S15) with pyrho (see above). In this way, we accounted for broadscale interchromosomal variation in recombination rate, where the average recombination rate increases with decreasing chromosome size. We then estimated the age of ancestral nodes in the inferred trees with tsdate v. 0.1.7 (*38*), using the variational gamma method of dating. For each chromosome we provided a mutation rate of $4.6 \cdot 10^{-9}$ per bp per generation (*76*), while a chromosome-specific N_e was estimated from chromosome-level genetic diversity. We then extracted the TMRCA for each pair of haplotypes across the full set of trees based on the ages estimated by tsdate. To obtain genome-wide estimates of divergence times, we calculated the logarithmic mean TMRCA, weighted by tree span, for haplotype pairs of each interspecies comparison. Mean TMRCAs were estimated for all trees genome-wide, restricted to trees

overlapping with genic regions, and for all trees on chromosome 20. We then obtained 95% confidence intervals around the mean values of TMRCA by performing 100 bootstrap replicates, in which we resampled haplotype pairs with replacement and re-estimated the mean TMRCA.

To reconstruct the *hispanica* complex's introgression history, we investigated genome-wide D statistics and f4-ratios using Dsuite (*82*). For these analyses we specified *O. deserti* as outgroup (*22*).

Phenotype scoring

To obtain phenotype scores of throat, mantle (here referring to mantle plus back) and neck coloration for GWAS, we categorized plumage traits based off pictures of field-caught or museum specimens. We scored throat coloration as discrete black or white. For mantle and neck side coloration, intermediate phenotypes are frequent. Hence, we divided these plumage traits into four ordinal categories: white (1), white with some black feathers (2), black with some white feathers (3), and black (4). To visualize the transition of mantle and neck side coloration across geography and admixture proportions, we estimated a phenotypic hybrid index based on the average between individuals' mantle and neck side scores (Fig. 1H, fig. S7).

GWASs

To identify genetic variants associated with polymorphic throat, and divergent mantle and neck side coloration, we conducted GWAS for each of the three plumage color traits. To avoid eliminating relevant SNPs, we preferred a strategy of initially restricting filtering to a minimum but checking for adequate data quality of significant sites recovered by GWAS post hoc. To this end, we determined filtering thresholds based on parameter distributions of Fisher Strand (FS), Strand Odds Ratio (SOR), RMS Mapping Quality (MQ), Mapping Quality Rank Sum Test (MQRankSum), Quality by Depth (QD), Read Position Rank Sum Test (ReadPosRankSum), and combined depth across samples (INFO/DP) and filtered for FS>40.0, SOR>3, MQ<4, MQRankSum>5, QD<2.0, ReadPosRankSum=4, and INFO/DP>13000 using bcftools v1.10 (*72*). We further removed (i) individuals with missing phenotypic information, (ii) females (because they hardly express the phenotype and therefore scoring is difficult or impossible), (iii) individuals with a sequencing coverage below 8x, and (iv) individuals that were outliers in the PCA (ones that formed a principal component by themselves). The final GWAS data set comprised 336 (58 white and 278 black) individuals for throat coloration and 230 individuals for mantle and neck side coloration. Because mantle and neck coloration are fixed in the parental species and only recombine in hybrids, we only included (hybrid) populations that comprise phenotypically admixed individuals in the respective GWAS (Fig. 1H and fig. S7). We conducted GWAS using linear mixed models to account for relatedness as implemented in GEMMA v0.98.5 (*30*). We ran all models with default settings, removing (i) non-polymorphic SNPs, (ii) SNPs with a missingness above 5%, (iii) SNPs with a minor allele frequency below 1%, and (iii) SNPs with an r^2 above 0.9999, which left 40,180,174 SNPs (of originally 162,197,710) SNPs. Throat color variation was almost entirely explained by two SNPs in consecutive codons in ASIP exon 2. Genotypes between the two SNPs were perfectly correlated except for genotypes of the second SNP that showed a third, non-functional indel variant whose pseudogenizing effect fit the associated phenotypes (n=8).

We followed up all GWAS on the global datasets described above by running GWAS split by species and hybrid zone, respectively. For throat coloration, we conducted GWAS using linear mixed models accounting for ancestry in *melanoleuca* (n=46, n_{white}=21, n_{black}=25) and the Iranian hybrids (n=87, n_{white}=17, n_{black}=70). We refrained from modeling GWAS for *hispanica* and *pleschanka* given the limited sampling sizes in *hispanica* (n=11, n_{white}=5, n_{black}=6) and of the white-throated *vittata* phenotype in *pleschanka* (n=2). With these very limited sample sizes, GWAS in GEMMA that model traits as additive did not have the power to identify SNPs associated with the dominant throat color architecture. We therefore reran GWAS under a dominant genetic architecture in angsd 0.940 (*83*). For mantle and neck-side coloration, we performed GWAS on four sets of admixed individuals: Iran (n=87), Black Sea (n=65), Black Sea and Caspian (n=93), and Caucasus (n=50).

Identification, genotyping, and phasing of structural variants

Because ASIP usually affects color phenotypes through regulatory changes that can involve structural variants (SVs) (*17*), we updated the repeat annotation of chromosome 20 with a manually curated version (*84*) of the *O. melanoleuca* library (*63*) (data S3) and investigated the ASIP region for SVs associated with throat coloration. To identify SVs on chromosome 20, we ran Smoove 0.2.8 (https://github.com/brentp/smoove) using all the short-read libraries available. The genotypes of 291 individuals were directly inferred by Smoove (the remaining 44 individuals could not be genotyped due to insufficient sequencing coverage). SVs were then manually curated using Samplot 1.3.0 (*85*) and Plotcritic 1.01 (*86*) to remove false positives. We overlapped the SV annotation with the repeat annotation performed with RepeatMasker/4.1.0 using the new version of the *O. melanoleuca* library. An SV linked to the throat SNP ('throat LTR') was found to be a full-length endogenous retrovirus (LTR retrotransposon) with a target site duplication of 5 bp (ERVL oenMel1-22) (fig. S10). To investigate the presence of the throat LTR in the white-throated *vittata* phenotype of *pleschanka* based on a genome assembly, we investigated the insertion site in a published draft genome assembly of this phenotype [(*60*); ENA biosample SAMEA6823426]. To phase the throat LTR and throat SNP, that is, to infer whether in heterozygous individuals the LTR insertion locates on the same haplotype as the white throat SNP variant, we interrogated linked-read barcodes in 38 heterozygous individuals with linked reads available in the ASIP region. To this end, we first selected all reads supporting the two throat SNP variants and then collected the reads that shared those barcodes separately for the two variants. We then mapped the reads supporting the black and white variants separately to the genomic region comprising the ASIP gene and the throat LTR and in both instances counted the reads mapping to the LTR.

Statistical modeling of genetic architecture

We aimed to understand how variants inferred by GWAS interact to form the associated phenotypes. To this end, we modeled mantle and neck side variation with an additive effect by, for each individual counting the number of white alleles (alleles associated with predominantly white phenotypes) across mantle loci. To model the phenotypic effect of white as dominant, we counted the number of mantle loci with at least one white allele for each individual, assuming that the phenotypic effect of one single white allele is the same as for two white alleles for this site. To model black as dominant, we did the same analysis, but counting the number of sites with at least one black allele for each individual. We then used these genetic effects to predict mantle or neck side phenotypes using Bayesian linear regression models as implemented in the brms package version 2.20.4 (*87*) in R version 4.3.3. We fitted Bayesian linear regression models with the family categorical, using four chains each running 4,000 iterations with a warm-up of 2,000 iterations. We then used the loo_compare() function implemented in loo package to identify the best performing model using leave-one-out (LOO) cross-validation. The cross validation implemented in the loo_compare() function performs pairwise comparisons of each model to the model with the highest estimated log predictive density (ELDP) and generates the difference in ELDP and the associated standard error (*88*).

As evidenced in Fig. 2C the throat LTR and SNP genotypes explained all color variation observed in *hispanica* and *melanoleuca*. We consequently refrained from further statistical modeling (Bayesian models require variation).

Gene expression analysis

To confirm the regulatory nature of GWAS hits and corroborate the genetic architecture of throat and mantle plumage coloration, we investigated ASIP expression and transcript diversity in wheatears of alternative phenotypes. As representatives of alternative plumage coloration, we targeted white-mantled but alternatively white- or black-throated *melanoleuca* and black-mantled and black-throated *cypriaca*. Two white-throated *melanoleuca* individuals were caught in Greece, two black-throated *melanoleuca* individuals were caught each one in Greece and in Cyprus, and three *cypriaca* individuals were caught in Cyprus. Birds were caught between August 22nd and September 18th, 2024, during (*cypriaca*) or after (*melanoleuca*) post juvenile or post-breeding moult when the birds are or have been in a physiological state conducive of feather regrowth and first year birds already moulted feathers to adult-type coloration. Immediately after capture, the birds were transported to bird care facilities at Pafos Zoo (Paphos, Cyprus) or Attica Zoological Park (Athens, Greece).

To characterize ASIP transcript variation and expression during feather growth—that is, when pigments are deposited in the feathers—we induced the regrowth of feathers. To this end, we pulled feathers from the two feather tracts of interest, throat and mantle, as well as from the shoulder and the belly, where both species express black (shoulder) and white colored plumage (belly). Shoulder and belly feathers served as a control to differentiate whether gene expression indeed differed between feathers of different coloration and not between species. Feather regrowth proceeded heterogeneously between species, individuals, and feather tracts in regard to both the onset of regrowth and whether or not feathers regrew. Given that throat and especially mantle feathers do not exhibit uniform color when fresh but particularly in *melanoleuca* show banded color patterns (fig. S25), we pulled feathers at sizes during which we expected them to express the respective white or black color for each feather tract (based on the patterns of the initially pulled feathers; fig. S25). Pulled feathers were immediately stored in liquid nitrogen until further storage and transportation and final storage at -80°C.

Next, to investigate associations of gene expression in regrown feathers with black and white coloration in the different feather tracts, we performed single-molecule real-time isoform sequencing (Iso-Seq) of total RNA on Pacific Biosciences' Revio long-read sequencing platform. To this end, from the available regrown feathers (table S16) we made selections of each 3-9 regrowing feathers to go into ten RNA pools representing alternative colorations and feather tracts (table S17). For *melanoleuca* mantle feathers, we generated three pools with feathers of different sizes and visible coloration (table S17) to ensure that we capture the stage at which white is expressed (fig. S25). From these feather pools we extracted total RNA using the RNeasy Universal kit (Qiagen) following the manufacturer's instructions. The tissues for RNA extraction consisted of the feather tips embedded in the skin that include the follicular sheet, the epidermal collar, part of the dermal papilla, and the pulp of the feather. RNA extracts were cleaned up with the RNeasy mini kit (Qiagen). RNA quantities and qualities were controlled on a Qubit fluorometer (Thermo Fisher Scientific) or an RNA ScreenTape on a TapeStation (Agilent). To prepare the RNAs for sequencing they were transcribed to cDNA and barcoded individually using the Iso-Seq express 2.0 kit (Pacific Biosciences). To maximize sequencing yield, we constructed three sequencing libraries using the KINNEX kit (Pacific Biosciences), with pools grouped into KINNEX libraries as described in table S17. The three resulting libraries were then sequenced each separately on one SMRT cell 25M (according to table S17) at the NGS platform of the University of Berne, Switzerland.

We processed the Iso-Seq data using Pacific Biosciences' SMRTLink platform. Consensus reads were generated from subreads and then demultiplexed. Demultiplexing included removal of barcodes and cDNA primer fragments. During the refine step, polyA tails were trimmed and concatemers removed to produce full-length non-concatemer reads (FLNC). FLNC reads were clustered per sequencing library to produce FASTA files containing assembled transcripts for each sample. Tools used: ccs, 8.2.0; pbtrim, 1.1.0; jasmine, 2.3.0; lima, 2.11.0; skera-split, 1.3.0; lima.1, 2.12.0; refine, 4.2.0; pbmerge, 3.4.0; cluster2, 4.2.0. We then generated gene models and transcript quantification using the OenMel1 reference genome and annotation (*63*) in IsoQuant version 3.6.3 (*89*) the currently most efficient tool for isoform detection (*90*). We used the previously assembled transcripts as input data with options–data_type assembly–fl_data –count_exons. IsoQuant reported an updated gene model containing an additional exon upstream of ASIP that matches existing gene models for ASIP (e.g., chicken in the ENSEMBL gene browser). To improve transcript classification, we re-ran IsoQuant with the updated ASIP gene annotation for each sample. We performed downstream data analysis of IsoQuant results in R.

Evolutionary history of color phenotypes

To reconstruct the evolutionary history of throat and mantle coloration, we first reconstructed the underlying variation's ancestral states. To this end we applied the maximum-likelihood approach implemented in est-sfs (*91*) with *O. deserti*, *O. monacha*, and *O. oenanthe* (*22*) as outgroups and the rate 6 model to estimate the ancestral states in the common ancestor of the *hispanica* complex and *O. deserti*. To infer the longer-term conservation of the two throat SNPs, we furthermore queried the NCBI ortholog database for ASIP, downloaded one amino acid sequence per bird species other than wheatear (N=107) and aligned these along with the two wheatear sequences using NCBI's COBALT multiple alignment tool.

To infer whether similar throat and mantle plumage phenotypes in the *hispanica* complex are based on common variation recruited from ancestral variation, by introgression, or evolved independently from novel variation, we ran topology weighting in TWISST (*92*) under the premise that topologies grouping lineages according to coloration reflect a common genetic basis. TWISST was run reducing the phased dataset to the reference individuals and removing either *O. pleschanka* or the Pontic lineage to only explore three (unrooted) topologies at a time: the species tree [(*O. hispanica, O. pleschanka*/Pontic), *O. melanoleuca, O. cypriaca*)], a coloration tree [(*O. hispanica, O. melanoleuca*), *O. pleschanka*/Pontic, *O. cypriaca*)] grouping white and black lineages, and a topology reflecting known introgression (*26*, *27*) between *O. melanoleuca* and *O. pleschanka*/Pontic [(*O. hispanica*, *O. cypriaca*), (*O. melanoleuca*, *O. pleschanka*/Pontic)]. We generated maximum likelihood trees for TWISST using the phyml_sliding_window.py script (*92*) with the GTR model using 50 SNPs windows in PhyML (*93*). We were interested to understand the extent by which *melanoleuca*- and hybrid-type Pontic (n=11) and Caspian (n=3) individuals were more introgressed by *O. melanoleuca* than parent-type phenotypes in these populations. To this end, we counted the frequency of topologies congruent with color introgression, that is, ones in which *O. melanoleuca* is grouped with the according *melanoleuca*-type/hybrid-type, or with parent-type Pontic or Caspian individuals. We obtained these counts genome-wide, for chromosome 20, and for the 50 kbp ASIP region to understand whether introgression occurred genome-wide or concentrated on chromosome 20 or the ASIP region. To obtain further evidence for or against introgression, we investigated patterns of excessive allele sharing across the ASIP region based on f_{dM}, testing for (i) "white" introgression between *O. hispanica* and *O. melanoleuca*, variably including one of the three black-mantled lineages. As a control comparison and to exclude that introgression occurred between the black-mantled lineages, we estimated f_{dM} between *O. cypriaca, O. pleschanka* and Pontic. We computed the f_{dM} statistic in a sliding window containing 50 SNPs with a step of 5 SNPs using Dsuite (*82*) Dinvestigate and *O. deserti* as outgroup.

To infer the age of the color phenotypes and history of ASIP introgression we compared TMRCAs (logarithmic mean weighted by tree span) for different combinations of species and ASIP haplotypes (tables S3

and S9), with TMRCAs estimated from the ARG using tsinfer and tsdate as outlined above. The time scales at which ASIP haplotypes coalesce relative to interspecific variation provides information on whether color phenotypes are ancestral (interspecific variation coalesces before ASIP haplotypes) or introgressed (interspecific variation coalesces after ASIP haplotypes). We used TMRCAs between species to time the history of species splits within which to place the evolution of ASIP haplotypes, (Fig. 4C and table S3). To estimate the age of black and white throat haplotypes, we identified black-throated and white-throated haplotypes based on which allele was present at the first coding SNP within the ASIP gene that is associated with throat coloration (chromosome 20, position 12,972,212). To estimate the time at which the black and white throat haplotypes split, we estimated the TMRCA within and between all black- and white-throated haplotypes, with haplotypes combined across species. To determine the ancestral haplotype and the origin of the derived haplotype, we further compared TMRCAs within and between black- and white-throated haplotypes for both *melanoleuca* and *hispanica*, identifying the ancestral haplotype as the one coalescing further back in time. Finally, to estimate the timing of introgression, we estimated the TMRCA between white *melanoleuca* and white *hispanica* alleles. The analyses targeting the variation underpinning throat coloration were performed both on the ASIP gene body and on the gene body plus the region upstream to the throat LTR. Analyses to determine the evolutionary history of mantle coloration followed the same logic, grouping haplotypes according to species' mantle coloration, and estimating TMRCAs from the ASIP region upstream the LTR and for the entire 50 kbp comprising the ASIP region. All analyses made use of the reference individuals but excluded five individuals from the reference set for the throat phenotype-based analysis due to poor genotype quality at the diagnostic SNP or throat LTR.

To obtain first insights into the role of selection in fixing ASIP variation associated with alternative coloration, we compared each taxon's unfolded (ancestry-polarized) site frequency spectrum (SFS) of mantle loci against the taxon's SFS with alleles polarized according to the coloration they underpin (Fig. 6A). These comparisons visualize whether the frequencies of alleles underpinning the expressed mantle colors are skewed toward higher frequencies compared to skews in the frequencies in ancestry classes. Under the premise that the coloration each allele underpins rather than its ancestry is exposed to selection, this comparison provides insights into the role of selection in the evolution of species' coloration. Finally, we inferred footprints of positive selection in the ASIP region based on haplotype structure and variation in the TMRCA across the ASIP region. To this end, we inferred XP-nSL (*94*) in selscan v2.0.2 (*95*) and estimated within-taxon TMRCAs as stated above.

Comparative phylogenomics

To infer whether the same SNP variants are responsible for throat and mantle color variation across wheatears and relatives, from 50 taxa and phenotypes we extracted genotypes from mapped reads published in Alaei Kakhki *et al.* (*22*). Based on the same data, we genotyped the throat LTR in each species using Paragraph (*96*) to infer whether the same LTR underpins non-melanic throat coloration in species outside the *hispanica* complex.

Stable isotope analysis

To investigate whether the throat color polymorphism could be associated with differences in dietary composition or environmental variables between phenotypes, we investigated stable isotope signatures in the blood of 161 wheatears. To this end, whole blood preserved in >95% ethanol was used to conduct multi-element stable isotope analysis for the quantification of $\delta^{13}C$, $\delta^{15}N$, and $\delta^{34}S$ values. Dried and powdered samples of approximately 0.3-0.4 mg were weighed to the nearest 0.001 mg in small tin cups on a micro-analytical balance. Samples were then combusted in a vario micro cube elemental analyzer (Elementar Analysensysteme GmbH, Germany). Subsequently, the evolved CO_2, N_2, and SO_2 gases underwent gas chromatographic separation and were introduced into the inlet of an Isoprime isotope ratio mass spectrometer (IRMS) for determination of 13C/12C, 15N/14N ratios and 34S/32S. Measurements are reported in δ-notation ($\delta^{13}C$, $\delta^{15}N$, and $\delta^{34}S$), where

$$\delta_{\text{sample}} = \left[\left(\frac{R_{\text{sample}}}{R_{\text{standard}}}\right) - 1\right] \times 1000$$

relative to standards [Pee Dee Belemnite (PDB) for carbon, atmospheric N_2 for nitrogen, and IAEA-S1 and IAEA-S3 for sulfur] in parts per thousand deviations. Thirty-two sulfanilamide (Iso-prime internal) standards were placed between five unknowns in sequence. These Internal laboratory standards indicated measurement errors (SD) of ± 0.08‰ for $\delta^{13}C$, 0.12‰ for $\delta^{15}N$ and 0.23‰ for $\delta^{34}S$. To infer whether throat plumage phenotypes differed in blood isotope composition, we ran linear models with log-transformed $\delta^{13}C$ as response variables, throat coloration as explanatory variable, and coordinates (latitude, longitude), birds' age, day in the year, and year as covariables. Interactions between latitude and longitude, and day in the year and year were included in the models. We performed these analyses on the complete data set (n=161) as well as separately for *melanoleuca* individuals (n=92), Greek *melanoleuca* individuals (n=57), and Iranian individuals (n=69) to account for potential large-scale geographic isotope variation not captured by coordinates. To account for potential biases connected with the smaller sample size of white-throated individuals, we performed a resampling procedure. To this end, we repeated the aforementioned analyses by resampling (10^5 times) as many black-throated individuals as white-throated individuals were available for the respective data partition. To understand whether the mean and variance of $\delta^{15}N$ values observed for the white individuals differ from the ones of the same sample sizes of black-throated individuals, we plotted the mean and variance of the $\delta^{15}N$ values observed for white-throated individuals against the distributions of the according resampled values for black-throated individuals (fig. S16, left and middle). To confirm that these results also hold after accounting for covariates, for each of the 10^5 samples we furthermore recorded the effect size of throat coloration from a linear model including the same covariates as the linear model run on the full data. Based on this, we tested whether the distributions of effect sizes of throat coloration adjusted for sample size and covariates differ from zero (fig. S16, right).

REFERENCES AND NOTES

1. M. Nei, The new mutation theory of phenotypic evolution. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 12235–12242 (2007). doi: 10.1073/pnas.0703349104; pmid: 17640887
2. K. F. Stryjewski, M. D. Sorenson, Mosaic genome evolution in a recent and rapid avian radiation. *Nat. Ecol. Evol.* **1**, 1912–1922 (2017). doi: 10.1038/s41559-017-0364-7; pmid: 29085063
3. S. Lamichhaney *et al.*, A beak size locus in Darwin's finches facilitated character displacement during a drought. *Science* **352**, 470–474 (2016). doi: 10.1126/science.aad8786; pmid: 27102486
4. Heliconius Genome Consortium, Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**, 94–98 (2012). doi: 10.1038/nature11041; pmid: 22722851
5. S. P. Turbek *et al.*, Rapid speciation via the evolution of pre-mating isolation in the Iberá Seedeater. *Science* **371**, eabc0256 (2021). doi: 10.1126/science.abc0256; pmid: 33766854
6. D. A. Marques, J. I. Meier, O. Seehausen, A Combinatorial View on Speciation and Adaptive Radiation. *Trends Ecol. Evol.* **34**, 531–544 (2019). doi: 10.1016/j.tree.2019.02.008; pmid: 30885412
7. T. M. Anderson *et al.*, Molecular and evolutionary history of melanism in North American gray wolves. *Science* **323**, 1339–1343 (2009). doi: 10.1126/science.1165448; pmid: 19197024
8. E. Huerta-Sánchez *et al.*, Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* **512**, 194–197 (2014). doi: 10.1038/nature13408; pmid: 25043035

9. R. W. R. Wallbank *et al.*, Evolutionary Novelty in a Butterfly Wing Pattern through Enhancer Shuffling. *PLOS Biol.* **14**, e1002353 (2016). doi: 10.1371/journal.pbio.1002353; pmid: 26771987
10. P. F. Colosimo *et al.*, Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science* **307**, 1928–1933 (2005). doi: 10.1126/science.1107239; pmid: 15790847
11. S. Wang *et al.*, Selection on a small genomic region underpins differentiation in multiple color traits between two warbler species. *Evol. Lett.* **4**, 502–515 (2020). doi: 10.1002/evl3.198; pmid: 33312686
12. M. D. Baiz, A. W. Wood, A. Brelsford, I. J. Lovette, D. P. L. Toews, Pigmentation Genes Show Evidence of Repeated Divergence and Multiple Bouts of Introgression in *Setophaga* Warblers. *Curr. Biol.* **31**, 643–649.e3 (2021). doi: 10.1016/j.cub.2020.10.094; pmid: 33259789
13. J. I. Meier *et al.*, Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nat. Commun.* **8**, 14363 (2017). doi: 10.1038/ncomms14363; pmid: 28186104
14. J. I. Meier *et al.*, Cycles of fusion and fission enabled rapid parallel adaptive radiations in African cichlids. *Science* **381**, eade2833 (2023). doi: 10.1126/science.ade2833; pmid: 37769075
15. L. Campagna *et al.*, Repeated divergent selection on pigmentation genes in a rapid finch radiation. *Sci. Adv.* **3**, e1602404 (2017). doi: 10.1126/sciadv.1602404; pmid: 28560331
16. B. Prud'homme *et al.*, Repeated morphological evolution through cis-regulatory changes in a pleiotropic gene. *Nature* **440**, 1050–1053 (2006). doi: 10.1038/nature04597; pmid: 16625197
17. D. L. Bannasch *et al.*, Dog colour patterns explained by modular promoters of ancient canid origin. *Nat. Ecol. Evol.* **5**, 1415–1423 (2021). doi: 10.1038/s41559-021-01524-x; pmid: 34385618
18. M. Gautier *et al.*, The Genomic Basis of Color Pattern Polymorphism in the Harlequin Ladybird. *Curr. Biol.* **28**, 3296–3302.e7 (2018). doi: 10.1016/j.cub.2018.08.023; pmid: 30146156
19. M. Bohutínská, C. L. Peichel, Divergence time shapes gene reuse during repeated adaptation. *Trends Ecol. Evol.* (2023). doi: 10.1016/j.tree.2023.11.007; pmid: 38155043
20. J. Arendt, D. Reznick, Convergence and parallelism reconsidered: What have we learned about the genetics of adaptation? *Trends Ecol. Evol.* **23**, 26–32 (2008). doi: 10.1016/j.tree.2007.09.011; pmid: 18022278
21. J. Cerca, Understanding natural selection and similarity: Convergent, parallel and repeated evolution. *Mol. Ecol.* **32**, 5451–5462 (2023). doi: 10.1111/mec.17132; pmid: 37724599
22. N. Alaei Kakhki *et al.*, A Phylogenomic Assessment of Processes Underpinning Convergent Evolution in Open-Habitat Chats. *Mol. Biol. Evol.* **40**, msac278 (2023). doi: 10.1093/molbev/msac278; pmid: 36578177
23. M. Aliabadian *et al.*, Convergent evolution of morphological and ecological traits in the open-habitat chat complex (Aves, Muscicapidae: Saxicolinae). *Mol. Phylogenet. Evol.* **65**, 35–45 (2012). doi: 10.1016/j.ympev.2012.05.011; pmid: 22634240
24. M. Schweizer *et al.*, Genome-wide evidence supports mitochondrial relationships and pervasive parallel phenotypic evolution in open-habitat chats. *Mol. Phylogenet. Evol.* **139**, 106568 (2019). doi: 10.1016/j.ympev.2019.106568; pmid: 31349102
25. M. Schweizer *et al.*, Parallel plumage colour evolution and introgressive hybridization in wheatears. *J. Evol. Biol.* **32**, 100–110 (2019). doi: 10.1111/jeb.13401; pmid: 30421480
26. J. Haffer, Secondary contact zones of birds in Northern Iran. *Bonn. Zool. Monogr.* **10**, 1–64 (1977).
27. E. N. Panov, *Wheaters of the Palearctic. Ecology, Behaviour and Evolution of the genus Oenanthe* (Pensoft, 2005).
28. D. J. Lawson, G. Hellenthal, S. Myers, D. Falush, Inference of population structure using dense haplotype data. *PLOS Genet.* **8**, e1002453 (2012). doi: 10.1371/journal.pgen.1002453; pmid: 22291602
29. C. A. Buerkle, C. Lexer, Admixture as the basis for genetic mapping. *Trends Ecol. Evol.* **23**, 686–694 (2008). doi: 10.1016/j.tree.2008.07.008; pmid: 18845358
30. X. Zhou, M. Stephens, Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44**, 821–824 (2012). doi: 10.1038/ng.2310; pmid: 22706312
31. H. E. Hoekstra, Genetics, development and evolution of adaptive pigmentation in vertebrates. *Heredity* **97**, 222–234 (2006). doi: 10.1038/sj.hdy.6800861; pmid: 16823403
32. G. A. Semenov *et al.*, Asymmetric introgression reveals the genetic architecture of a plumage trait. *Nat. Commun.* **12**, 1019 (2021). doi: 10.1038/s41467-021-21340-y; pmid: 33589637
33. S. J. Lin *et al.*, Topology of feather melanocyte progenitor niche allows complex pigment patterns to emerge. *Science* **340**, 1442–1445 (2013). doi: 10.1126/science.1230374; pmid: 23618762
34. C. R. Linnen *et al.*, Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science* **339**, 1312–1316 (2013). doi: 10.1126/science.1233213; pmid: 23493712
35. N. Haupaix *et al.*, The periodic coloration in birds forms through a prepattern of somite origin. *Science* **361**, eaar4777 (2018). doi: 10.1126/science.aar4777; pmid: 30237324
36. D. Liang *et al.*, Genomic Analysis Revealed a Convergent Evolution of LINE-1 in Coat Color: A Case Study in Water Buffaloes (*Bubalus bubalis*). *Mol. Biol. Evol.* **38**, 1122–1136 (2021). doi: 10.1093/molbev/msaa279; pmid: 33212507
37. J. Kelleher *et al.*, Inferring whole-genome histories in large population datasets. *Nat. Genet.* **51**, 1330–1338 (2019). doi: 10.1038/s41588-019-0483-y; pmid: 31477934
38. A. W. Wohns *et al.*, A unified genealogy of modern and ancient genomes. *Science* **375**, eabi8264 (2022). doi: 10.1126/science.abi8264; pmid: 35201891
39. R. Inger, S. Bearhop, Applications of stable isotope analyses to avian ecology. *Ibis* **150**, 447–461 (2008). doi: 10.1111/j.1474-919X.2008.00839.x
40. K. A. Hobson, Tracing origins and migration of wildlife using stable isotopes: A review. *Oecologia* **120**, 314–326 (1999). doi: 10.1007/s004420050865; pmid: 28308009
41. E. Yohannes, J.-L. Berthoud, F. Woog, Trait based niche differentiation in tetrakas (Bernieridae) endemic to Madagascar: A multi-isotope approach. *Front. Ecol. Evol.* **11**, 1082226 (2023). doi: 10.3389/fevo.2023.1082226
42. L. M. San-Jose *et al.*, Differential fitness effects of moonlight on plumage colour morphs in barn owls. *Nat. Ecol. Evol.* **3**, 1331–1340 (2019). doi: 10.1038/s41559-019-0967-2; pmid: 31477846
43. R. L. Mumme, White tail spots and tail-flicking behavior enhance foraging performance in the Hooded Warbler. *Auk* **131**, 141–149 (2014). doi: 10.1642/AUK-13-199.1
44. P. G. Jablonski, A rare predator exploits prey escape behavior: The role of tail-fanning and plumage contrast in foraging of the painted redstart (*Myioborus pictus*). *Behav. Ecol.* **10**, 7–14 (1999). doi: 10.1093/beheco/10.1.7
45. A.-L. Ducrest, L. Keller, A. Roulin, Pleiotropy in the melanocortin system, coloration and behavioural syndromes. *Trends Ecol. Evol.* **23**, 502–510 (2008). doi: 10.1016/j.tree.2008.06.001; pmid: 18644658
46. N. Rosser *et al.*, Hybrid speciation driven by multilocus introgression of ecological traits. *Nature* **628**, 811–817 (2024). doi: 10.1038/s41586-024-07263-w; pmid: 38632397
47. C. Estalles *et al.*, Concerted variation in melanogenesis genes underlies emergent patterning of plumage in capuchino seedeaters. *Proc. Biol. Sci.* **289**, 20212277 (2022). doi: 10.1098/rspb.2021.2277; pmid: 35016545
48. D. P. L. Toews *et al.*, Plumage Genes and Little Else Distinguish the Genomes of Hybridizing Warblers. *Curr. Biol.* **26**, 2313–2318 (2016). doi: 10.1016/j.cub.2016.06.034; pmid: 27546575
49. S. Y. W. Sin *et al.*, Genetic basis and evolution of structural color polymorphism in an Australian songbird. *Mol. Biol. Evol.* **41**, msae046 (2024). doi: 10.1093/molbev/msae046; pmid: 38415852
50. H. Vrieling, D. M. Duhl, S. E. Millar, K. A. Miller, G. S. Barsh, Differences in dorsal and ventral pigmentation result from regional expression of the mouse agouti gene. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 5667–5671 (1994). doi: 10.1073/pnas.91.12.5667; pmid: 8202545
51. J. D. Galbraith, A. Hayward, The influence of transposable elements on animal colouration. *Trends Genet.* **39**, 624–638 (2023). doi: 10.1016/j.tig.2023.04.005; pmid: 37183153
52. M. Rebeiz, J. E. Pool, V. A. Kassner, C. F. Aquadro, S. B. Carroll, Stepwise modification of a modular enhancer underlies adaptation in a *Drosophila* population. *Science* **326**, 1663–1667 (2009). doi: 10.1126/science.1178357; pmid: 20019281
53. R. D. H. Barrett, D. Schluter, Adaptation from standing genetic variation. *Trends Ecol. Evol.* **23**, 38–44 (2008). doi: 10.1016/j.tree.2007.09.008; pmid: 18006185
54. H. E. Hoekstra, J. A. Coyne, The locus of evolution: Evo devo and the genetics of adaptation. *Evolution* **61**, 995–1016 (2007). doi: 10.1111/j.1558-5646.2007.00105.x; pmid: 17492956
55. S. B. Carroll, Endless forms: The evolution of gene regulation and morphological diversity. *Cell* **101**, 577–580 (2000). doi: 10.1016/S0092-8674(00)80868-5; pmid: 10892643
56. C. Mérot, R. A. Oomen, A. Tigano, M. Wellenreuther, A Roadmap for Understanding the Evolutionary Significance of Structural Genomic Variation. *Trends Ecol. Evol.* **35**, 561–572 (2020). doi: 10.1016/j.tree.2020.03.002; pmid: 32521241
57. A. Eyre-Walker, P. D. Keightley, The distribution of fitness effects of new mutations. *Nat. Rev. Genet.* **8**, 610–618 (2007). doi: 10.1038/nrg2146; pmid: 17637733
58. E. B. Chuong, N. C. Elde, C. Feschotte, Regulatory activities of transposable elements: From conflicts to benefits. *Nat. Rev. Genet.* **18**, 71–86 (2017). doi: 10.1038/nrg.2016.139; pmid: 27867194
59. J. V. Peñalba *et al.*, The Role of Hybridization in Species Formation and Persistence. *Cold Spring Harb. Perspect. Biol.* **16**, a041445 (2024). doi: 10.1101/cshperspect.a041445; pmid: 38438186
60. D. Lutgen *et al.*, Linked-read sequencing enables haplotype-resolved resequencing at population scale. *Mol. Ecol. Resour.* **20**, 1311–1322 (2020). doi: 10.1111/1755-0998.13192; pmid: 32419391
61. M. Baym *et al.*, Inexpensive multiplexed library preparation for megabase-sized genomes. *PLOS ONE* **10**, e0128036 (2015). doi: 10.1371/journal.pone.0128036; pmid: 26000737
62. A. McKenna *et al.*, The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010). doi: 10.1101/gr.107524.110; pmid: 20644199
63. V. Peona *et al.*, An annotated chromosome-scale reference genome for Eastern Black-eared Wheatear (Oenanthe melanoleuca). *G3* **17**, jkad088 (2023).
64. A. Bishara *et al.*, Read clouds uncover variation in complex regions of the human genome. *Genome Res.* **25**, 1570–1580 (2015). doi: 10.1101/gr.191189.115; pmid: 26286554
65. S. Chen, Y. Zhou, Y. Chen, J. Gu, fastp: An ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018). doi: 10.1093/bioinformatics/bty560; pmid: 30423086

66. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009). doi: 10.1093/bioinformatics/btp324; pmid: 19451168
67. P. Edge, V. Bafna, V. Bansal, HapCUT2: Robust and accurate haplotype assembly for diverse sequencing technologies. *Genome Res.* **27**, 801–812 (2017). doi: 10.1101/gr.213462.116; pmid: 27940952
68. M. Martin, M. Patterson, S. Garg, S. O. Fischer, N. Pisanti, G. W. Klau, A. Schöenhuth, T. Marschall, WhatsHap: fast and accurate read-based phasing. bioRxiv 085050 [Preprint] (2016); https://doi.org/10.1101/085050.
69. O. Delaneau, J.-F. Zagury, M. R. Robinson, J. L. Marchini, E. T. Dermitzakis, Accurate, scalable and integrative haplotype estimation. *Nat. Commun.* **10**, 5436 (2019). doi: 10.1038/s41467-019-13225-y; pmid: 31780650
70. N. Patterson, A. L. Price, D. Reich, Population structure and eigenanalysis. *PLOS Genet.* **2**, e190 (2006). doi: 10.1371/journal.pgen.0020190; pmid: 17194218
71. S. Purcell *et al.*, PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007). doi: 10.1086/519795; pmid: 17701901
72. P. Danecek *et al.*, Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021). doi: 10.1093/gigascience/giab008; pmid: 33590861
73. E. P. Derryberry, G. E. Derryberry, J. M. Maley, R. T. Brumfield, HZAR: Hybrid zone analysis using an R software package. *Mol. Ecol. Resour.* **14**, 652–663 (2014). doi: 10.1111/1755-0998.12209; pmid: 24373504
74. Z. Gompert, D. A. DeRaad, C. A. Buerkle, A Next Generation of Hierarchical Bayesian Analyses of Hybrid Zones Enables Model-Based Quantification of Variation in Introgression in R. *Ecol. Evol.* **14**, e70548 (2024). doi: 10.1002/ece3.70548; pmid: 39583044
75. J. P. Spence, Y. S. Song, Inference and analysis of population-specific fine-scale recombination maps across 26 diverse human populations. *Sci. Adv.* **5**, eaaw9206 (2019). doi: 10.1126/sciadv.aaw9206; pmid: 31681842
76. L. Smeds, A. Qvarnström, H. Ellegren, Direct estimate of the rate of germline mutation in a bird. *Genome Res.* **26**, 1211–1218 (2016). doi: 10.1101/gr.204669.116; pmid: 27412854
77. T. C. Bruen, H. Philippe, D. Bryant, A simple and robust statistical test for detecting the presence of recombination. *Genetics* **172**, 2665–2681 (2006). doi: 10.1534/genetics.105.048975; pmid: 16489234
78. H. Ellegren *et al.*, The genomic landscape of species divergence in *Ficedula* flycatchers. *Nature* **491**, 756–760 (2012). doi: 10.1038/nature11584; pmid: 23103876
79. C. Zhang, M. Rabiee, E. Sayyari, S. Mirarab, ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19** (Suppl 6), 153 (2018). doi: 10.1186/s12859-018-2129-y; pmid: 29745866
80. B. Q. Minh *et al.*, IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020). doi: 10.1093/molbev/msaa015; pmid: 32011700
81. S. Schiffels, R. Durbin, Inferring human population size and separation history from multiple genome sequences. *Nat. Genet.* **46**, 919–925 (2014). doi: 10.1038/ng.3015; pmid: 24952747
82. M. Malinsky, M. Matschiner, H. Svardal, Dsuite - Fast D-statistics and related admixture evidence from VCF files. *Mol. Ecol. Resour.* **21**, 584–595 (2021). doi: 10.1111/1755-0998.13265; pmid: 33012121
83. T. S. Korneliussen, A. Albrechtsen, R. Nielsen, ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* **15**, 356 (2014). doi: 10.1186/s12859-014-0356-4; pmid: 25420514
84. C. Goubert *et al.*, A beginner's guide to manual curation of transposable elements. *Mob. DNA* **13**, 7 (2022). doi: 10.1186/s13100-021-00259-7; pmid: 35354491
85. J. R. Belyeu *et al.*, Samplot: A platform for structural variant visual validation and automated filtering. *Genome Biol.* **22**, 161 (2021). doi: 10.1186/s13059-021-02380-5; pmid: 34034781
86. J. R. Belyeu *et al.*, SV-plaudit: A cloud-based framework for manually curating thousands of structural variants. *Gigascience* **7**, giy064 (2018). doi: 10.1093/gigascience/giy064; pmid: 29860504
87. P.-C. Bürkner, brms: An R Package for Bayesian Multilevel Models Using Stan. *J. Stat. Softw.* **80**, (2017). doi: 10.18637/jss.v080.i01
88. A. Vehtari, A. Gelman, J. Gabry, Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* **27**, 1413–1432 (2017). doi: 10.1007/s11222-016-9696-4
89. A. D. Prjibelski *et al.*, Accurate isoform discovery with IsoQuant using long reads. *Nat. Biotechnol.* **41**, 915–918 (2023). doi: 10.1038/s41587-022-01565-y; pmid: 36593406
90. Y. Su *et al.*, Comprehensive assessment of mRNA isoform detection methods for long-read sequencing data. *Nat. Commun.* **15**, 3972 (2024). doi: 10.1038/s41467-024-48117-3; pmid: 38730241
91. P. D. Keightley, B. C. Jackson, Inferring the Probability of the Derived *vs.* the Ancestral Allelic State at a Polymorphic Site. *Genetics* **209**, 897–906 (2018). doi: 10.1534/genetics.118.301120; pmid: 29769282
92. S. H. Martin, S. M. Van Belleghem, Exploring Evolutionary Relationships Across the Genome Using Topology Weighting. *Genetics* **206**, 429–438 (2017). doi: 10.1534/genetics.116.194720; pmid: 28341652
93. S. Guindon *et al.*, New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010). doi: 10.1093/sysbio/syq010; pmid: 20525638
94. Z. A. Szpiech, T. E. Novak, N. P. Bailey, L. S. Stevison, Application of a novel haplotype-based scan for local adaptation to study high-altitude adaptation in rhesus macaques. *Evol. Lett.* **5**, 408–421 (2021). doi: 10.1002/evl3.232; pmid: 34367665
95. Z. A. Szpiech, selscan 2.0: Scanning for sweeps in unphased data. *Bioinformatics* **40**, btae006 (2024). doi: 10.1093/bioinformatics/btae006; pmid: 38180866
96. S. Chen *et al.*, Paragraph: A graph-based structural variant genotyper for short-read sequence data. *Genome Biol.* **20**, 291 (2019). doi: 10.1186/s13059-019-1909-7; pmid: 31856913
97. D. Lutgen *et al.*, Supporting material for: A mosaic of modular variation at a single gene underpins convergent plumage coloration, Zenodo (2025); https://doi.org/10.5281/zenodo.15487817.
98. F. Suárez, El dimorfismo de la collalba rubia (Oenanthe hispanica L.): Variaciones geograficas y entre habitats. *Ardeola* **37**, 291–298 (1990).
99. BirdLife International, *Handbook of the Birds of the World, Bird species distribution maps of the world, version 6.0* (Lynx, 2016).

ACKNOWLEDGMENTS

We thank A. Khani, R. Ramezani, M. Monfared, S. Huang, V. Khimonen, C. Nikiforou, M. Dretakis, S. Iezekiel, C. Grande, M. Illa, M. Pérez, M. Dimaki, C. Ninou, and J. Rodriguez Fuentes for support with fieldwork and M. Alaei Kakhki; the regional water company of Khorasan Razavi, Northern Khorasan, Semnan, Ardabil, and Tabriz; Ferdowsi University of Mashhad; I. Grigoryeva; V. Voronova; and F. Jiguet for logistic support. We thank P. Korner for statistics support, A. Gomez for support with DNA extractions, S. Schiffels for support with data analysis, and the Peichel group, the Suh group, and Jochen Wolf for discussion. M. Carneiro and S. Yin provided guidance on feather sampling for transcriptome analysis. We thank the University of Erlangen, Applied and Environmental Geology, and the University of Graz, Department of Earth Sciences, for the support in the triple-element stable isotope measurements. We thank all permit-issuing authorities. The Burke Museum of the University of Washington, the Museum of Comparative Zoology of Harvard University, the Steinhardt Museum of Natural History in Tel Aviv, the Royal Ontario Museum in Toronto, and Hadoram Shirihai kindly provided samples. Computations were performed at the Genetic Diversity Centre (GDC), ETH Zurich, the High-Performance Computing Cluster EVE, a joint effort of the Helmholtz Centre for Environmental Research (UFZ) and the German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig; at UPPMAX with resources provided by the project NAISS 2024/22-351 from the National Academic Infrastructure for Supercomputing in Sweden (NAISS), partially funded by the Swedish Research Council through grant agreement no. 2022-06725; and at the University of Cyprus High Performance Cluster. We thank the administration and support staff of GDC, EVE, and NAISS. We thank the sequencing facilities of NGI Sweden in Solna and the NGS platform of the University of Berne and their staff for their support. **Funding:** This research was supported by German Research Foundation (DFG) research grant BU3456/3-1 and Swiss National Science Foundation (SNSF) grant 310030_207417 to R.B., National Research Fund (FNR) Luxembourg grant 14575729 to D.L., an A.G. Leventis Foundation Grant to A.N.G.K., a Georg Foster Research Stipend of the Alexander von Humboldt Foundation and a scholarship for female researchers from Friedrich-Schiller-University Jena to N.A.K., and a Swedish Research Council (Vetenskapsrådet) grant 2022-06195 to V.P. **Author contributions:** R.B. conceptualized the study with input from D.L., A.N.G.K., M.S., and H.S. D.L., V.P., M.A.C., N.A.K., F.T.-L., E.Y., and S.G.d.S. performed analyses. P.-A.G. supported data analysis. D.L., A.-L.D., M.B., and A.U. performed lab work. R.B., M.S., H.S., P.-A.G., A.N.G.K., A.S., and C.L.P. supervised analyses. P.A., S.M.L., M.M., E.Y., A.A., T.A., M.A., N.A., V.B., I.C., J.L.C., E.F., J.T.G., Z.J., A.K., F.T.-L., Y.L., N.P., R.P., Á.P., R.R., A.D.S., F.S., P.S., S.S., B.S., A.T., N.V., F.T.-L., J.G., and L.Y. provided resources. R.B. and D.L. wrote the paper with input from M.A.C., P.-A.G., A.N.G.K., F.T.-L., C.L.P., V.P., H.S., M.S., and A.S. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Sequencing data and the linked read-based *vittata* genome assembly are available on the short-read archive, bioprojects PRJNA1128009 (whole-genome resequencing data) and PRJNA1260766 (Iso-Seq data). Stable isotope data and scripts used for analyses are available on Zenodo (*97*).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ado8005
Supplementary Text; Figs. S1 to S25; Tables S1 to S17; References (*100*); Data S1 to S3; MDAR Reproducibility Checklist

Submitted 24 May 2024; resubmitted 24 March 2025; accepted 17 July 2025
10.1126/science.ado8005