EVOLUTIONARY BIOLOGY

Adaptive introgression reveals the genetic basis of a sexually selected syndrome in wall lizards

Nathalie Feiner[1]*, Weizhao Yang[1]†, Ignas Bunikis[2], Geoffrey M. While[3], Tobias Uller[1]

The joint expression of particular colors, morphologies, and behaviors is a common feature of adaptation, but the genetic basis for such "phenotypic syndromes" remains poorly understood. Here, we identified a complex genetic architecture associated with a sexually selected syndrome in common wall lizards, by capitalizing on the adaptive introgression of coloration and morphology into a distantly related lineage. Consistent with the hypothesis that the evolution of phenotypic syndromes in vertebrates is facilitated by developmental linkage through neural crest cells, most of the genes associated with the syndrome are involved in neural crest cell regulation. A major locus was a ~400-kb region, characterized by standing structural genetic variation and previously implied in the evolutionary innovation of coloration and beak size in birds. We conclude that features of the developmental and genetic architecture contribute to maintaining trait integration, facilitating the extensive and rapid introgressive spread of suites of sexually selected characters.

INTRODUCTION

A defining feature of adaptation is that different components of the organism work together. For example, behaviors that serve to defend territories are frequently accompanied by large body size, armaments, and colorful ornaments. Similarly, the ability to remain invisible to predators commonly relies on particular combinations of color, morphology, and behavior. This need for functional phenotypic integration poses a major problem for evolution because coloration, morphology, and behavior are largely underpinned by different molecular mechanisms, cell types, and tissues. In vertebrates, colors are produced by microstructures and pigments in cells of the integument, body size, and shape are a result of growth and development of tissues such as bone and cartilage, and behavior is determined by properties of the brain and the hormonal system. Yet, the ubiquity of complex adaptations suggests that there are genetic and developmental factors that enable characters that need to function together to also evolve together.

One possibility is that different kinds of traits are kept together because of the inheritance of large, nonrecombining regions of the genome that contain several functionally independent genes. For example, in the ruff, a large inversion controls differences in coloration, body size, and reproductive behavior, giving rise to three distinct morphs that are maintained by balancing selection (*1*). In the absence of physical genetic linkage, adaptive phenotypic integration may instead capitalize on developmental mechanisms that allow changes in genes or regulatory regions to cause traits to covary. In vertebrates, it has been suggested that the coevolution of coloration, morphology, physiology, and behavior is facilitated by the developmental biology of neural crest cells (NCCs) [e.g., the mammalian "domestication syndrome" (*2*)]. The neural crest is a transient structure in the early vertebrate embryo whose cells begin to migrate throughout the body and eventually differentiate into many different cell types, including chromatophores, craniofacial mesenchyme, neurons, glia and dorsal root ganglia, and cells of the thymus and adrenal gland. Modifications of genes that regulate the proliferation, migration, or differentiation of NCCs can cause concomitant changes to pigmentation, head morphology, and behavior, thereby facilitating their correlated phenotypic evolution. While comparisons of the genomes of domesticated animals and their wild relatives provide some evidence for this hypothesis (*3*–*5*), it remains controversial [e.g., (*6*)] and comparable data on phenotypic syndromes of neural crest–derived traits in wild animals are lacking.

A promising system for studying how suites of traits are assembled during evolution is found in the common wall lizard, *Podarcis muralis* (Fig. 1A). Common wall lizards are small diurnal lizards with a broad distribution across central and southern Europe. In central Italy, the ancestral phenotype is replaced by animals with exaggerated color ornaments, larger bodies and heads, and more aggressive, dominant behavior (Fig. 1, B and C), all traits that are under sexual selection (*7*–*9*). The most extreme form is a highly distinct phenotype that was described as subspecies *P. muralis nigriventris*. However, the defining characters of this "nigriventris" phenotype or syndrome (e.g., green and black coloration and large head size) vary across the landscape (*10*) and even introgress into a distantly related lineage, across what is otherwise a narrow hybrid zone with limited gene flow (*11*, *12*).

Here, we capitalize on the within-lineage variation and adaptive introgression of coloration and morphology to reveal that the nigriventris phenotype is a syndrome that is underpinned by a genomic architecture with at least one major locus with complex structural variation, and several additional loci distributed across multiple chromosomes. Genes associated with the identified loci have known functions in the proliferation, migration, or differentiation of NCCs, consistent with a role for this cell type in orchestrating the evolutionary origin and adaptive spread of a novel phenotype composed of multiple distinct characters.

RESULTS

The common wall lizard is a widespread species that can be found from Iberia to Asia Minor, and it comprises six major lineages (*13*).

[1]Department of Biology, Lund University, Lund, Sweden. [2]Uppsala Genome Center, Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden. [3]School of Natural Sciences, University of Tasmania, Sandy Bay, Tasmania, Australia.
*Corresponding author. Email: nathalie.feiner@biol.lu.se
†Present address: Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, China.

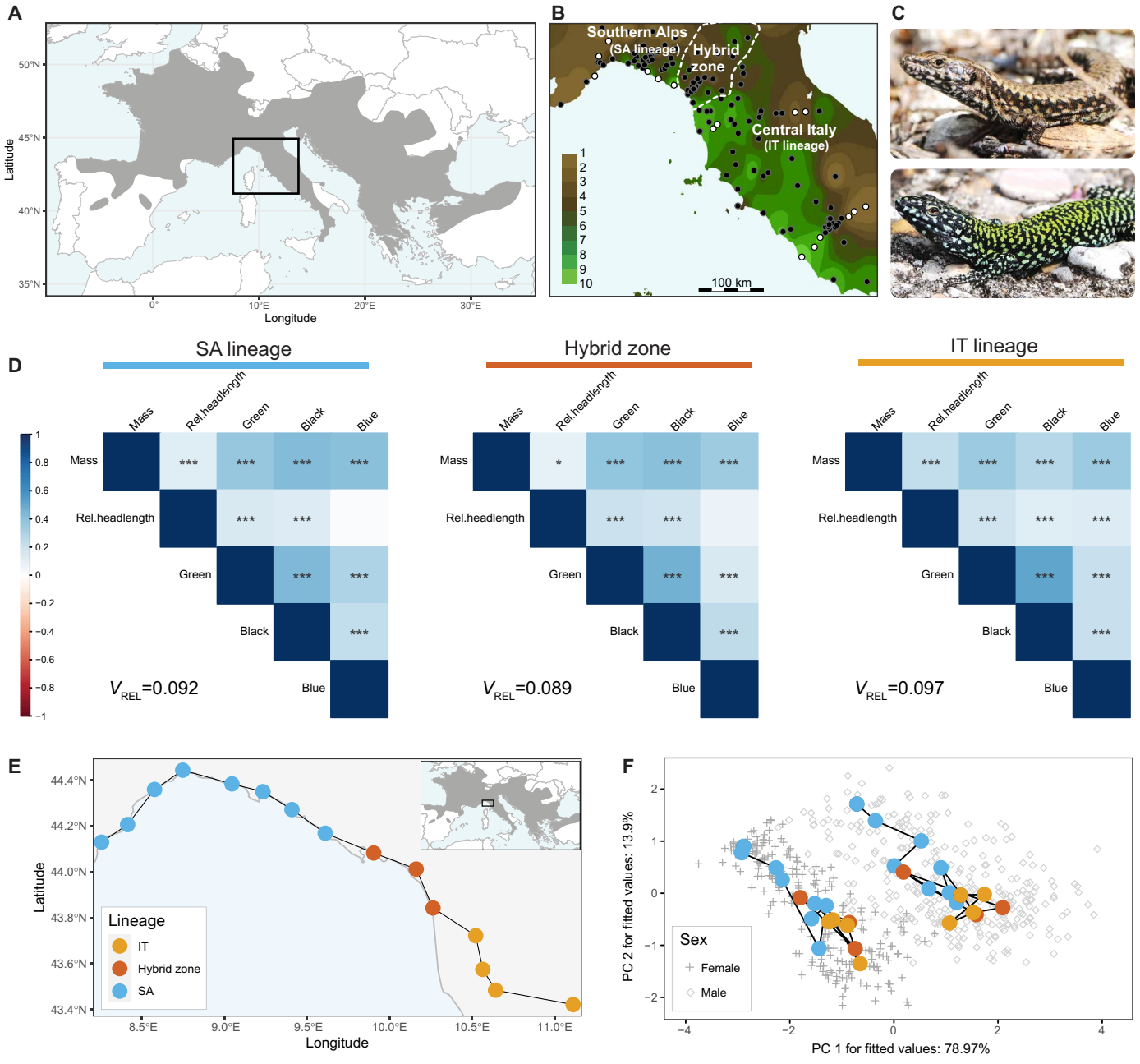

Fig. 1. Geographic distribution and integration of morphology and coloration. (**A**) Distribution map of the common wall lizard (*P. muralis*) across Europe. Black box marks the study area shown in (B). (**B**) Geographic distribution of dorsal coloration (greenness scores from 1 to 10; 1 being pure brown and 10 being pure green), a proxy for the nigriventris phenotype, in male *P. muralis* across 150 populations (N_{males} = 2392). Dots mark the location of sampled populations (black dots for phenotype analyses and white dots for populations used in whole-genome resequencing). (**C**) Typical phenotypes of male common wall lizards with the ancestral (top) and derived nigriventris (bottom) phenotype. (**D**) The five syndrome traits are all significantly correlated at an individual level in males within the central Italian (IT) lineage, the hybrid zone, and the southern Alps (SA) lineage, with the exception of blue spot size and relative head length in the SA lineage and the hybrid zone (SA: $N_{populations}$ = 35, $N_{individuals}$ = 602; hybrid zone: $N_{populations}$ = 31, $N_{individuals}$ = 495; IT: $N_{populations}$ = 71, $N_{individuals}$ = 1128). Note that for the SA lineage, only populations that have been reached by the introgression front (mean greenness scores of males above 3) were included in these analyses. (**E**) Geographic location of the introgression cline along the Ligurian coast with populations color-coded by lineage. (**F**) The five traits show similar patterns of introgression between males and females. Trajectory analysis of the first two PCs of the fitted values of the five traits showed that neither magnitude (path distance $d_{females}$ = 11.47, d_{males} = 14.89, P = 0.21) nor direction (angle = 11.24°, P = 0.24) or shape (distance between females and males = 0.55, P = 0.10) differs significantly between the sexes ($N_{populations}$ = 15; $N_{females}$ = 226; N_{males} = 337).

The data in this paper come from extensive sampling of individuals and populations across the central Italian (IT) and southern Alps (SA) lineages. These lineages appear to have diverged about 5.2 to 6.0 million years ago (*13*) and form a hybrid zone in the northwest of Italy (roughly at the border between Liguria in the northwest and Tuscany in the southeast; Fig. 1B) (*11*, *12*). Using restriction site–associated sequencing (RAD-seq) on 900 individuals across 71 populations, we derived average admixture scores (Q) that were used to assign populations to the IT lineage ($Q > 0.95$), SA lineage ($Q < 0.05$), or hybrid populations [$0.05 \leq Q \geq 0.95$; fig. S1; see also (*10*–*12*)]. Previous work has shown that the nigriventris phenotype originated in the IT lineage and that the characteristic color ornaments and large body and head size have since introgressed into the SA lineage, far beyond the hybrid zone per se (Fig. 1B) (*11*, *14*). Introgression of the nigriventris phenotype is particularly pronounced along the coast, a pattern that is concordant with the geographic distribution of the characters within the IT lineage and that likely reflects climatic effects on the strength of sexual selection (*10*).

Trait integration persists during introgression

While previous work has demonstrated introgression of the nigriventris coloration and morphological traits, it is unknown to what extent these traits remain consistently correlated or if the correlation breaks down during introgression. We used data on five traits that are characteristic of the nigriventris phenotype and that have been previously shown to be under sexual selection in males (*8*, *9*, *14*): three coloration traits (dorsal coloration, ventral coloration, and area of lateral blue spots) and two morphological traits (body mass and relative head length). These traits are all significantly pairwise correlated in males from the IT lineage (all $P < 0.001$; Fig. 1D). An estimate of the overall strength of phenotypic integration (*15*, *16*) among these five traits in males is strongest in the IT lineage ($V_{rel} = 0.097$), but is not significantly weaker in the hybrid zone ($V_{rel} = 0.089$, $P = 0.428$) and in the SA lineage populations reached by the introgression front ($V_{rel} = 0.092$, $P = 0.607$). Despite that the nigriventris phenotype appears to be selected because of male-male competition (*8*, *9*, *14*), the pattern of trait introgression is highly similar in females. Quantifying trait changes along a transect, from IT populations with the nigriventris phenotype across the hybrid zone to SA populations exhibiting the ancestral phenotype, shows that the magnitude, direction, and shape of the phenotypic trajectories are indistinguishable between the sexes (Fig. 1, E and F). The nigriventris phenotype introgresses far beyond the genetic center of the cline and the hybrid zone (as defined above), reaching similar end points for both males and females (*11*, *14*). In sum, different lines of evidence support the existence of a syndrome in both sexes, and that the syndrome remains intact during introgression into a distantly related lineage.

Introgression facilitates the identification of the genomic basis of the syndrome

To identify the genomic basis of the nigriventris phenotype, we exploited the fact that the nigriventris and the ancestral phenotypes both exist in the IT and the SA lineages (in the latter because of introgression; Fig. 1). We could therefore make use of F_{ST} outlier scans in a paired design for both lineages, using whole-genome resequence (WGS) data from a total of 60 males (Fig. 2A). In the SA lineage, we contrasted 10 nigriventris with 10 ancestral phenotype individuals. In the IT lineage, we included two such contrasts (each 10 versus 10) to cover the main axis of genetic differentiation within this lineage (which runs from Rome in the south to Tuscany in the north; fig. S1) but analyzed them together (i.e., 20 versus 20). In each of these three pairs, populations with ancestral and nigriventris phenotypes are genetically very similar, but differ between Rome and Tuscany, which makes this approach powerful (see fig. S1) (*11*). Each group of 10 males was sourced from between two and three different sites to reduce population-specific effects. Sequence data were analyzed using a chromosome-level reference genome stemming from a conspecific from the Western European genetic lineage (henceforth the Spanish reference genome) (*17*), and the main results were validated using two new genome assemblies of Italian individuals (see below).

NCC regulators are differentiated between nigriventris and ancestral phenotypes

Many genomic regions were differentiated between lizards with nigriventris and ancestral phenotypes within each lineage (Fig. 2B). Within the IT lineage, we identified 1222 islands of differentiation (each at least 5 kb in size) that are highly differentiated between the two phenotypes (i.e., high F_{ST} values). Within the SA lineage, we found 1021 such windows. Intersecting between the two, we found 81 genomic islands shared by both lineages (8.5 times more than expected by chance; expected range, 4 to 15). These 81 genomic regions were associated with 89 genes, and of these, 45 were protein coding and had biologically meaningful annotations. Twenty-four of these genes (53%) could be linked to NCC biology (table S1). For example, the gene *Rbm10* regulates NCC proliferation (*18*–*20*), the gene *Tfcp2l1* is a transcription factor that regulates the switch from a stem cell state to differentiation in NCCs (*21*–*23*), and the gene *Gpc3* has been identified as a "trailblazer" gene crucial for NCC migration (*24*–*26*). Furthermore, several of the genes have described roles in complex human syndromes affecting neural crest–derived features (craniofacial development, behavior, etc.), including the Joubert syndrome (*Tmem237*) (*27*), the TARP syndrome (*Rbm10*) (*28*), the Christianson syndrome (*Slc9a6*) (*29*), or the Charcot-Marie-Tooth syndrome (*Sbf2*) (*30*). This indicates that genomic islands differentiated in lizards expressing the nigriventris phenotype are rich in genes with described functions in the proliferation, migration, or differentiation of NCCs.

The significance of the genomic islands identified by the F_{ST} outlier scans for the nigriventris phenotype was substantiated using lineage-specific genome-wide association studies (GWAS) on an independent dataset consisting of RAD-seq data of 497 individuals from the IT lineage and 188 individuals from the SA lineage. To maximize the power of detecting genomic variants that are associated with the entire suite of traits, we used the first principal component (PC) of the five traits as response variable. Individual traits associated with the nigriventris phenotype load positively and evenly on PC1 (fig. S2), which therefore represents a suitable composite score of the magnitude of expression of the nigriventris phenotype. In the IT lineage, we identified 9191 single-nucleotide polymorphisms (SNPs) (out of a total of ~190,000) associated with the first PC of nigriventris-associated traits (Fig. 2A), and 134 of these are in close proximity (less than 5-kb distance) to islands of differentiation identified in the IT lineage (2.4 times more than expected by chance; expected range, 41 to 70; Fig. 2C). In the corresponding analysis in the SA lineage, we found a total of 4379 SNPs, of which 67 are in close proximity to islands of differentiation (2.9 times more than expected by chance; expected range, 14 to 32; Fig. 2C).

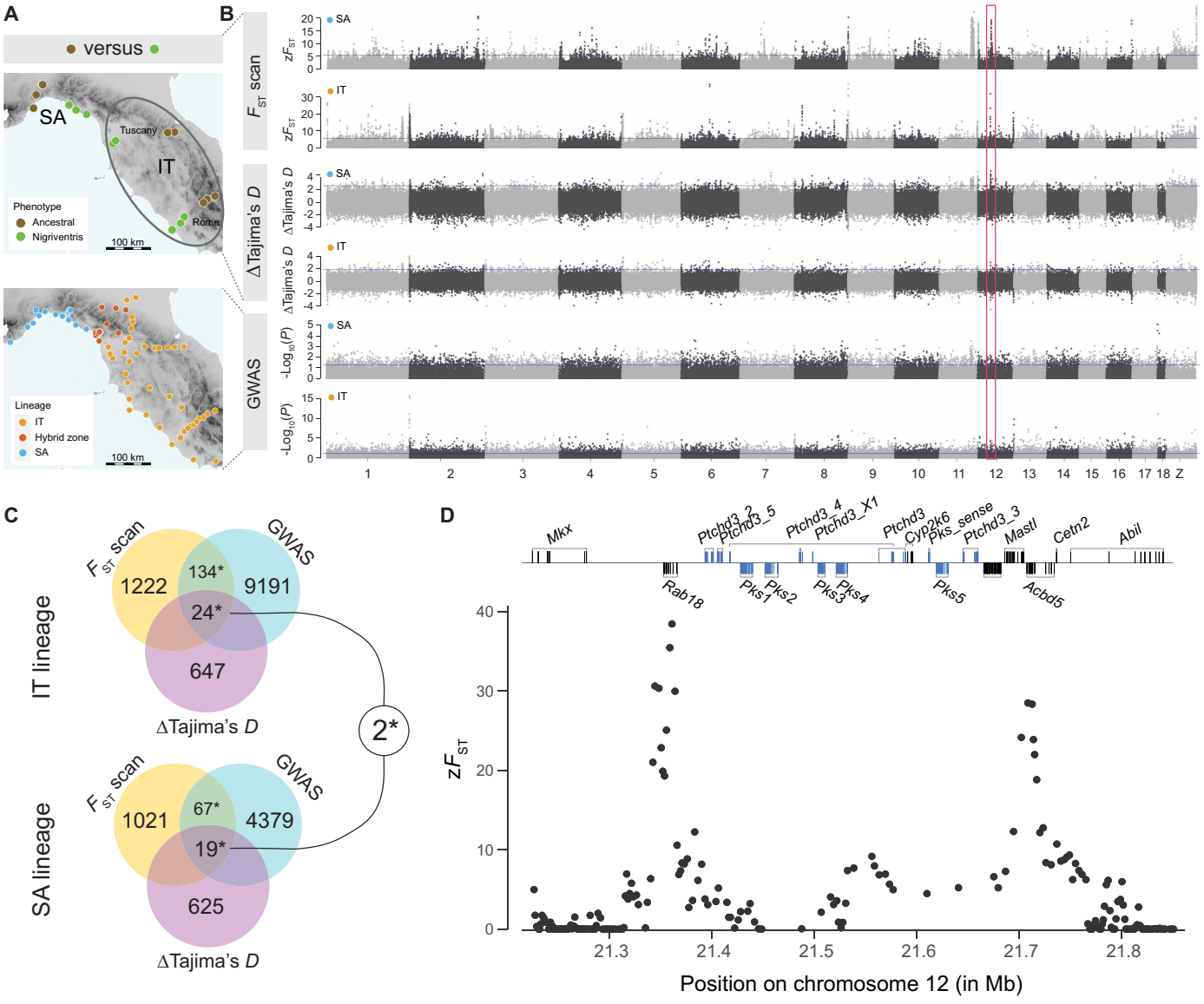

Fig. 2. Genomic regions associated with the syndrome. (**A**) Top panel shows sampling locations for the F_{ST} outlier scan and Tajima's *D* analyses using WGS data. Bottom panel shows sampling locations for the GWAS using RAD-seq data. A paired design was applied that aimed to minimize genetic distance and maximize phenotypic differences within each pair. Three pairs with 10 versus 10 individuals were selected: one pair belonging to the SA, and two pairs to the IT lineage. (**B**) Results of F_{ST} outlier scans, ΔTajima's $D_{(brown\text{-}green)}$, and GWAS analyses. Top, differentiation expressed as zF_{ST} between lizards with nigriventris and ancestral phenotype for each of the lineage-specific comparisons; middle, ΔTajima's $D_{(brown\text{-}green)}$ for the same comparisons; bottom, likelihood of association with the first PC in GWAS on RAD-seq datasets per lineage, all plotted for the 18 autosomes and the Z chromosome. Blue lines mark the 99.5th percentile used as significance threshold. The region on chromosome 12 with the most consistent association is highlighted with a red box. (**C**) Venn diagrams presenting the overlap between different statistical approaches for identifying loci underpinning the nigriventris phenotype in each lineage. Values in circles give total numbers that include the overlapping loci. Asterisks indicate overlaps that are higher than expected by chance. (**D**) Close-up of the genomic region on chromosome 12 with the most consistent association. zF_{ST} values derived across both lineages and summarized in 100-SNP windows. The region is characterized by two peaks with high association, separated by 350 kb that include a number of duplicated *Pks* (polyketide synthase) and *Ptchd3* (patched domain-containing protein 3) genes highlighted in blue.

As a final criterion for the identification of candidate loci underpinning the nigriventris phenotype, we scrutinized differences in genetic diversity. We hypothesized that populations exhibiting the derived (nigriventris) phenotype should have experienced strong selection, leading to removal of genetic variation at loci underpinning the nigriventris phenotype as it spread across the landscape. We therefore calculated Tajima's *D* as a measure of genetic diversity for each phenotype group and each lineage, and screened for genomic windows with particularly low genetic diversity in nigriventris relative to ancestral phenotypes [i.e., high ΔTajima's $D_{(brown\text{-}green)}$]. Of the 134 genomic candidate regions associated with the nigriventris phenotype in the IT lineage identified by F_{ST} outlier scans and GWAS, 24 were also among the 99.5th percentile outliers with particularly high ΔTajima's $D_{(brown\text{-}green)}$ (24 times enriched;

expected range, 0 to 2), and the corresponding number in the SA lineage was 19 genomic regions in total (36 times enriched; expected range, 0 to 1; Fig. 2C).

Finally, when intersecting the lineage-specific outliers that fulfilled all three criteria {genomic differentiation (F_{ST}), association across landscape (GWAS), and high difference in genetic diversity [ΔTajima's $D_{(brown\text{-}green)}$]}, we find two genomic islands (0 to 1 expected by chance) that are strongly and consistently associated with the syndrome (Fig. 2C). These are located on chromosome 12 (21,342,501 to 21,370,000) and chromosome 11 (64,690,001 to 64,697,500). The latter region is small (7.5 kb) and located within 550 kb of the end of the chromosome. Since telomeric regions are suspected to be enriched in false positives in genome scans (*31*), subsequent analyses focused on the region on chromosome 12.

The nigriventris phenotype is associated with a complex genomic region characterized by structural variation

The genomic region that was most strongly and consistently associated with the nigriventris phenotype in both lineages was a region on chromosome 12 that overlaps the gene *Rab18*. Analyses of gene flow revealed that this region likely introgressed from the IT lineage into populations of the SA lineage that strongly express the nigriventris phenotype (fig. S3). The *Rab18* gene encodes a small GTPase (guanosine triphosphatase) that regulates intracellular trafficking of vesicles and is a positive regulator of directed cell migration (*32*). In close vicinity, ~360 kb downstream of *Rab18* lies another peak of strong association with the nigriventris phenotype, coinciding with the gene *Acbd5*, which codes for the *acyl-CoenzymeA binding domain containing 5* protein. This protein functions in the intracellular transport of long-chain acyl–coenzyme A and regulates interactions between peroxisomes and the endoplasmatic reticulum. The region between *Rab18* and *Acbd5* is characterized by multiple gene copies of two genes, *Pks* (polyketide synthase) and *Ptchd3* (patched domain-containing protein 3). This interpeak genomic region containing the gene copies of *Pks* and *Ptchd3* does not show an association with phenotypic variation (Fig. 2D), nor does it harbor obvious neural crest regulators. However, variants within this broader region have been identified as a top candidate for the genomic basis of subspecies differences in color and bill morphology in redpoll finches (*Acanthis* spp.) (*33*) and a color mutant in a parrot (*Melopsittacus undulates*) (*34*). This prompted us to investigate this genomic region in more detail and test if it exhibited features that could account for its association with coloration and morphology in common wall lizards.

Phylogenetic analyses of the protein-coding genes present in the interpeak region, the *Pks* and *Ptchd3* genes (derived from the Spanish reference genome), revealed that these gene copies originated partly in ancient tandem duplications dating back almost to the origin of squamates, and partly in more recent tandem duplications occurring after the split between *P. muralis* and the congeners *P. lilfordi* and *P. raffonei* (fig. S4). However, the *Pks* and *Ptchd3* gene repertoires are likely different in *P. muralis* from the IT and SA lineages since cumulative coverage of all 149 resequenced genomes revealed substantial variation in coverage and deviation from the read depth expected for single-copy loci (fig. S5). Patterns of copy number variation were observed to differ between individuals with the nigriventris and ancestral phenotypes (Fig. 3A). The gene *Cyp2k6*, which is located within the interpeak region in the Spanish reference genome, is likely absent from the genomes of *P. muralis* from the IT and SA lineages as evidenced by a lack of mapped reads at this locus (Fig. 3A and fig. S5). While these results are indicative of copy number variation in the interpeak region, there was no evidence for a large inversion spanning across the two outlier peaks (i.e., peaks are unlikely to correspond to breakpoints of a large inversion), and the region was not characterized by high densities of transposable elements (fig. S6).

To resolve potential structural variations harbored within the ~360 kb interpeak region, we used long-read sequencing (Oxford Nanopore), chromatin conformation capture-based sequencing (Hi-C), and optical mapping (Bionano) on two individuals from the southern range of the IT lineage [where the syndrome originated (*11*)]: one with the ancestral phenotype [from Colle di Sassa (CDS)] and one with the nigriventris phenotype [from Santa Marinella (SMA)]. Despite that the populations are fixed for these different phenotypes, the population genetic differentiation between these locations is very limited and not different from that expected from isolation by distance (*11*). The resulting genome assemblies (using Oxford Nanopore and Hi-C data) were at chromosome level (scaffold $N50_{SMA} = 89.83$ Mb; scaffold $N50_{CDS} = 93.19$ Mb), were highly complete ($BUSCO_{SMA} = 95.8\%$; $BUSCO_{CDS} = 96.6\%$), and reached a continuity and completeness on par with the Spanish reference genome (fig. S7).

Synteny relationships in the candidate region indicated that large portions of the interpeak region are absent from the SMA and CDS genome assemblies (Fig. 3B). Nevertheless, remapping of the WGS data confirmed the strong association of the *Rab18*- and *Acbd5*-containing regions with the nigriventris phenotype (Fig. 3C). Optical maps revealed that both assemblies lack 196-kb (SMA genome) and 215-kb (CDS genome) sequence information within this region (fig. S8). Comparisons of the optical maps to the Spanish reference genome further show that the regions not present in the assemblies are roughly equivalent in length to the interpeak region in the Spanish reference genome, but labels failed to align, indicating substantial sequence divergence (Fig. 3D). Further inspection of the label patterns in the optical maps shows that both SMA and CDS individuals are heterozygote in this genomic region. Repetitive optical marker arrangements indicate complex structural variation segregating in the IT lineage of wall lizards (Fig. 3D). Together, a variety of approaches indicate that the *Rab18- and Acbd5*-containing genomic region is evolutionary labile and characterized by both within-population and between-lineage structural variation.

The green and black coloration of the nigriventris phenotype is not associated with differential expression of candidate genes

One of the most conspicuous features of the nigriventris phenotype that sets it apart from the ancestral phenotype of *P. muralis* is the striking green and black dorsal coloration (Figs. 1C and 4A). That *Rab18* and *Acbd5* play important roles in intracellular trafficking is consistent with a role in the accumulation of pigments in chromatophores. Besides regulating directed cell migration, *Rab18* is also involved in lipid droplet dynamics (*35*–*38*), which is consistent with a function in pigment accumulation [intracellular trafficking and cell migration share some key regulators (*39*)]. Another strong candidate gene for coloration differences is *Pks* (or rather the five full-length copies of *Pks*), since this gene has been identified as causal for the synthesis of psittacofulvin, a yellow pigment unique to parrots (*34*). We therefore investigated the molecular basis of differences in

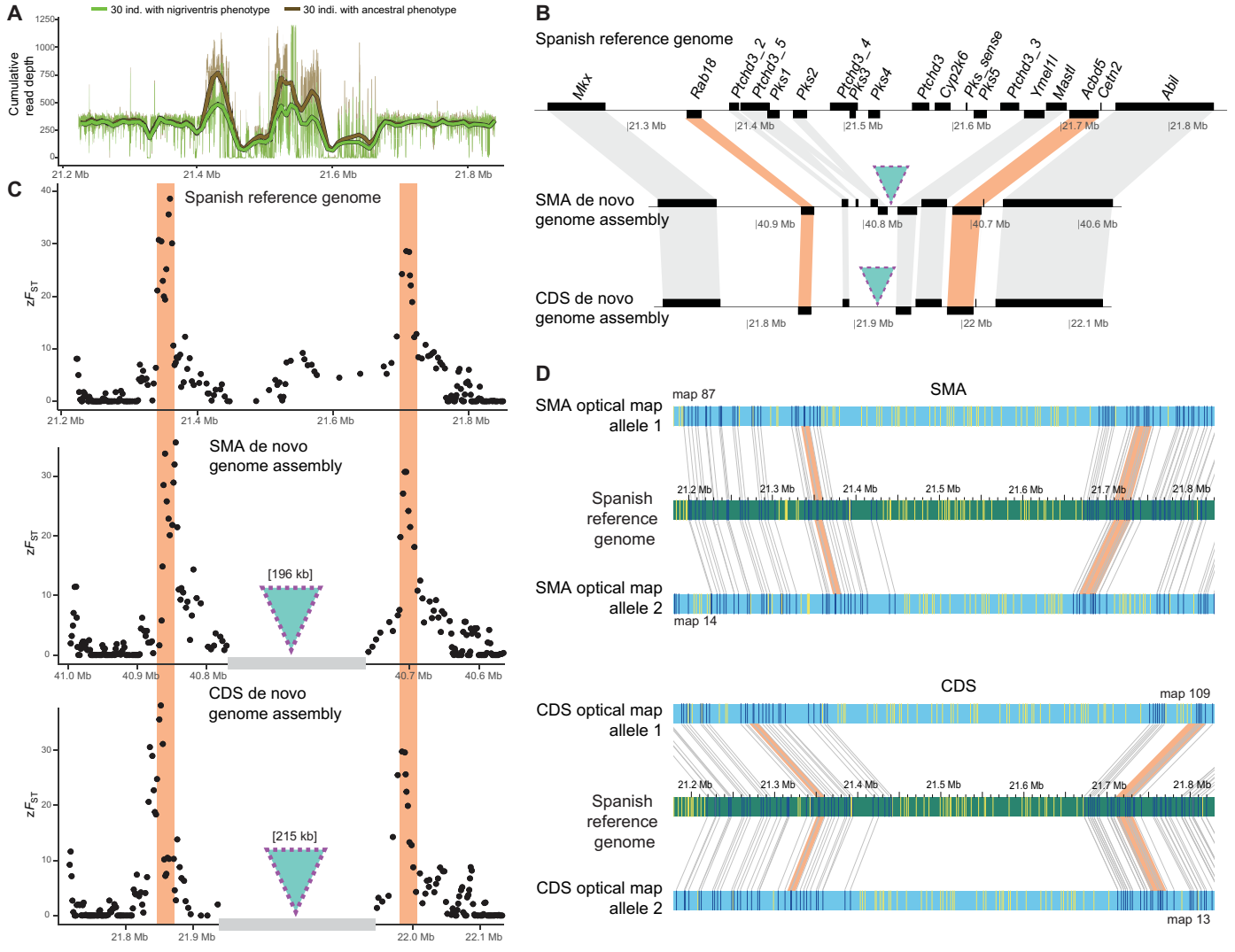

Fig. 3. Structural variation in the genomic region with the highest association with the syndrome. (**A**) Cumulative read depth of 30 individuals with nigriventris phenotypes and 30 with ancestral phenotypes (thick lines show smooth curves obtained using loess method). The interpeak region shows copy number variation that differs between phenotypes (regions where green and brown lines differ) and regions that lack reads from both phenotypes (at 21.6 Mb). (**B**) Gene annotations of the ~420-kb region on chromosome 12 in the Spanish reference genome (top), and the corresponding region in the de novo assembled genome of an individual with the nigriventris phenotype from SMA (middle) and an individual with the ancestral phenotype from CDS (bottom). Turquois triangles mark the genomic regions in SMA and CDS where orthologous genes are missing [also in (C)]. Across all panels, the position of the genes *Rab18* and *Acbd5* is highlighted with orange bars. (**C**) Genomic differentiation (as zF_{ST}) in the focal region between lizards with and without the nigriventris phenotype (30 versus 30) using three different reference genomes. The peaks of highest differentiation consistently lie in the genomic regions overlapping the genes *Rab18* and *Acbd5*, regardless of reference genome. Based on the information of the likely missing regions in the de novo genome assemblies shown in (B), the x axes of the plots were adjusted to include the length and position of the missing sequences. (**D**) Optical maps of the focal region (light blue bars above and below) aligned to the Spanish reference genome (dark green bar) of the SMA individual (top) and the CDS individual (bottom). In each bar, dark blue labels indicate aligned labels (connected by gray lines), whereas the yellow ones indicate unaligned labels.

coloration between the ancestral and nigriventris phenotypes and asked if any gene located in the candidate region could be linked to the characteristic green and black nigriventris coloration.

Using transmission electron microscopy (TEM), we found that black dorsal skin contained markedly more melanin-containing melanocytes than green or brown dorsal skin in the epidermal layer. Both brown and green skin contained a large number of xanthophores filled with pterinosomes (pteridine-filled vesicles), but only xanthophores in green skin contained high concentrations of carotenoid vesicles (Fig. 4B). These differences at the cellular level manifest in chromatic differences as quantified by spectrophotometry on live lizards (Fig. 4C) and extracted pigments (Fig. 4D) (*40*). Pigments extracted from green skin patches showed high concentration of molecules whose absorption curve showed the three-peak pattern characteristic for carotenoids. This absorption curve differed strongly from the absorption curve of the budgerigar feather extract, arguing against a similar role of *Pks* in pigment synthesis in lizards as in parrots (Fig. 4D) (*34*). To further characterize the identity of the yellow carotenoid pigment, we applied spectrophotometry and liquid chromatography–high-resolution mass spectrometry

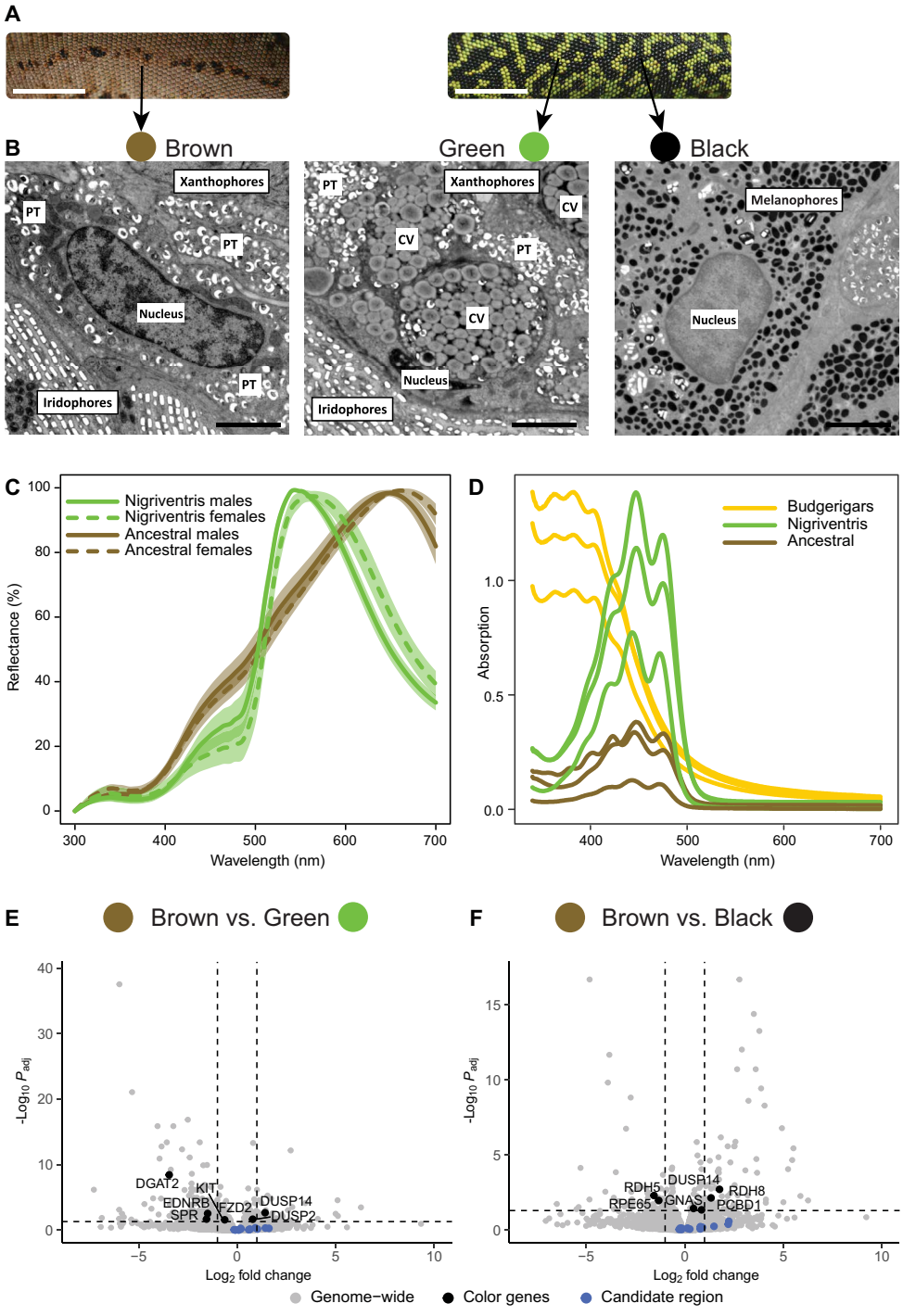

Fig. 4. Cellular and transcriptomic basis of coloration differences. (**A**) Dorsal coloration and color patterns of a lizard with the ancestral phenotype (left) and the nigriventris phenotype (right). Scale bar in both panels is 1 cm. (**B**) TEM of dorsal epidermal cross-sections of pure brown skin of a lizard with the ancestral phenotype and of pure green and pure black skin of a lizard with the nigriventris phenotype show differences in pigment cell composition underlying color differences. Xanthophores of both brown and green skin contain pterinosomes (PT), but only those in green skin are rich in carotenoid vesicles (CV). Black skin shows high densities of melanophores filled with melanosomes. For additional transmission electron micrographs, see figs. S10 and S11. Scale bar in all panels is 2 μm. (**C**) Reflectance of brown and green dorsal patches of lizards with the ancestral and nigriventris phenotypes, respectively, quantified by spectrophotometry. Lizards included here were sampled in the IT lineage close to Rome [ancestral phenotype from CDS: 14 females and 16 males; nigriventris phenotype from Castel Fusano (FU): 14 females and 24 males]. Shaded areas provide 95% confidence intervals for each group. (**D**) Absorption spectra of pigments extracted from three budgerigar feathers (containing psittacofulvin), and biopsies of each three male lizards with the nigriventris and the ancestral phenotypes. (**E** and **F**) Volcano plots of differential gene expression analyses between brown and green (349 differentially expressed genes) (E) and brown and black (234 differentially expressed genes) (F) skin biopsies ($N_{transcripts}$ = 24,947). For more information on the transcriptomic analyses and differentially expressed genes, see fig. S12 and https://doi.org/10.5281/zenodo.10396759.

(LC-HRMS) using electrospray ionization (ESI). These analyses ruled out the presence of psittacofulvin-like pigments in wall lizards and confirmed instead that the pigment enriched in green relative to brown lizard skin is the carotenoid lutein or a lutein stereoisomer based on exact mass, retention time, and tandem mass spectrometry (MSMS) fragmentation (fig. S9).

Finally, we applied an unbiased transcriptomic approach to our data, interrogating gene expression differences between black, green, and brown skin biopsies. This identified differential gene expression of a number of genes with described roles in pigment biosynthesis or transport identified in previous studies of coloration in reptiles (e.g., *Dgat2*, *Ednrb*, or *Rdh5*; Fig. 4, E and F) (*41*). Despite this, only a single gene identified as a candidate for the nigriventris phenotype (see above) was differentially expressed in green and black skin, and this overlap was not different from expectations based on chance alone. The lack of overlap between genes associated with the nigriventris phenotype and differential gene expression in skin of different color substantiates that the genomic analyses identified genes regulating overall syndrome trait expression rather than locally expressed "coloration genes" per se.

DISCUSSION

Adaptation commonly relies on the coexpression of suites of colors, morphological traits, and behaviors. The genetic and developmental mechanisms that enable such different characters to evolve in concert remain poorly understood. Here, we show that the dorsal and ventral coloration, lateral blue spot size, body size, and relative head length that characterize the nigriventris phenotype of common wall lizards remain tightly integrated even during introgression into a distantly related lineage. These results are consistent with correlational sexual selection, or syndrome selection (*42*), on coloration and morphology, presumably because of their joint function in male-male competition (*8*, *9*). However, the consistent trait integration also indicates that there may be genomic or developmental features that are conducive to the introgressive spread of this functionally integrated phenotype.

The large number and genome-wide distribution of loci robustly associated with the nigriventris phenotype rule out the existence of a single supergene governing the coevolution (and cointrogression) of coloration and morphology. Such genomic architecture would have been unexpected given that the individual traits are not discrete but vary continuously across the landscape. The polygenic basis is not only consistent with the quantitative expression of the nigriventris phenotype across the landscape but also suggests that strong ongoing correlational selection is necessary to maintain trait integration. Since literally hundreds of genes can influence the expression of the colors and morphologies associated with the nigriventris phenotype, the spread of these traits would likely cause widespread correlational selection on local standing genetic variation, which can account for the observed mixture of shared and lineage-specific genes associated with the nigriventris phenotype. We also identified at least one genomic region that may be considered a locus of major effect on the syndrome itself. The existence of a small number of loci of major effect would facilitate the adaptive spread of the nigriventris syndrome across the landscape and its introgression into the highly divergent genomic background of the SA lineage of *P. muralis*.

What kind of genetic change could have pleiotropic effects on coloration, morphology, and behavior? One possibility is perturbation of genes that regulate NCC proliferation, migration, and differentiation. For example, an increase or decrease in the number of NCCs will be accommodated by tissues containing neural crest–derived cells (e.g., epidermis or skull), resulting in correlated phenotypic variability (i.e., the propensity to vary) across very different kinds of traits. The expression and form of these traits can be modified further by selection on standing genetic variation in cell- and tissue-specific modifiers [genetic accommodation sensu West-Eberhard 2003 (*43*)]. Consistent with this hypothesis, we found that more than half of the top candidate genes that were both highly differentiated between the ancestral and nigriventris phenotypes and associated with introgression of the nigriventris phenotype into the SA lineage have empirically established links to NCC proliferation, migration, or differentiation. This link to NCC biology remains speculative, but the identification of these candidate genes provides leads for further research on the potential role of NCCs in orchestrating trait coexpression and coevolution.

While the developmental basis of the nigriventris syndrome remains to be established, our results singled out one genomic region with a particular strong association to this phenotype. The strongest candidates are the genes *Rab18* and *Acbd5*. *Rab18* is essential for intracellular trafficking and tethering of lipid vesicles to the endoplasmic reticulum (ER). It is also a regulator of cell migration (*32*, *44*), and thus potentially relevant for NCC functioning. The gene *Acbd5* encodes an acyl–coenzyme A binding domain protein and is involved in lipid metabolism and intracellular trafficking between the ER and other organelles, in particular peroxisomes. As pigment-containing carotenoid vesicles are a kind of lipid vesicle, it is possible that the genetic differences in *Rab18* (and perhaps *Acbd5*) are responsible for the striking color of the nigriventris phenotype, through accumulation of lutein-type carotenoids in the xanthophores of green skin, and perhaps also the accumulation of melanin in black skin. However, while both genes were highly expressed in lizard skin, expression levels did not differ between phenotypes. Only a single gene identified by genome-wide analyses was differentially expressed in skin of different color. Nevertheless, a deeper understanding of the cellular processes of pigment cell development and pigment synthesis is required to evaluate a possible causal role of genes in the Rab18-Acbd5 genomic region.

One feature of the Rab18-Acbd5 genomic region that rendered the analyses challenging is that the region in between those genes was characterized by extensive structural variation. This involved duplication and rearrangement of *Pks* and *Ptchd3* genes. *Ptchd3*, present in five full-length duplicates and one pseudogene, encodes transmembrane receptors that are related and structurally similar to *Patched* genes (*45*), which encode receptors in the sonic hedgehog pathway, and are therefore potentially relevant for embryonic development. While the ortholog of *Pks* in budgerigars appears to be crucial for the synthesis of the yellow psittacofulvin pigments (*34*), our results do not support a role for any of the five full-length *P. muralis Pks* genes in pigment synthesis.

Detailed analyses of two de novo assembled genomes and optical maps of two Italian males with contrasting phenotypes failed to fully resolve the interpeak genomic region between *Rab18* and *Acbd5*, but provided strong support for structural variation segregating within the Italian lineage. Gene duplications, deletions, or less obvious mechanisms such as gene conversions have shaped this

genomic region and possibly affected the epigenetic regulatory landscape pertaining on the larger genomic environment. Together with the fact that the *Rab18*-containing genomic region was also singled out as a candidate explaining differences in beak morphology and coloration, both neural crest–derived traits, in redpolls (*33*), this suggests that this region could well be an evolutionary labile hotspot. Testing this hypothesis will require developmental and comparative genomic analyses across a broader sample of populations, lineages, or species that differ consistently in head morphology, coloration, and behavior.

In conclusion, we capitalized on the well-documented geographic variation and adaptive introgression of color and morphology to identify the genomic basis of a sexually selected syndrome. The genomic architecture of this syndrome appears to be highly polygenic, but a single locus characterized by structural variation was singled out as a major determinant. We speculate that the evolution of the nigriventris syndrome is an example of how NCC biology can facilitate the evolutionary origin and introgressive spread of complex, integrated phenotypes.

MATERIALS AND METHODS

Phenotypic data

Lizards were captured by noosing during breeding season (March to May) between 2012 and 2019. The Ministry of Education, University and Research (MIUR) of Italy provided all authorizations for the study (2012–2013: Aut. Prot. PNM-0009344; 2014–2015: Aut. Prot. PNM-0011379; 2016–2018: Aut. Prot. PNM-0002154). All tissue samples were collected according to the Lund University Local Ethical Review Process. Phenotypic data were collected as previously described (*14*). In brief, body mass was weighed to the nearest 0.01 g, snout-to-vent length (svl) and total length were measured to the nearest mm, and head length was measured to the nearest 0.01 mm. Digital photographs from the dorsal, ventral, and lateral side of each lizard were used to score dorsal coloration ("greenness"), ventral coloration ("blackness"), and the total area of lateral blue spots (on the outer ventral scales). Dorsal coloration was scored based on an intensity scale from 1 to 10 (1 being pure brown, 10 being pure green), which was confirmed to be highly repeatable within and between observers, and highly correlated with scores from digital photographs analyzed in Photoshop CS4 and with values for green chroma extracted using spectrophotometry [see (*14*) for details]. To characterize chromatic differences, reflectance measurements (300 to 700 nm) were taken using a FLAME-S-UV-VIS-ES spectrometer equipped with a QR400-7-VIS-BX reading-illumination probe and a PX-2 Xenon strobe light for full-spectrum illumination (Ocean Optics Inc., Dunedin, FL, USA). A single measurement of the background dorsal coloration was taken at a fixed distance of 3 mm from the skin of lizards from each one population representing the ancestral (CDS) and nigriventris phenotypes (FU). Spectra were smoothened with a span factor of 0.2, normalized, and plotted using the R package pavo 2 [v. 2.8.0; (*46*)]. Ventral coloration was scored as the proportion of black versus nonblack pixels in ImageJ. The area of lateral blue spots was scored as the total area covered by blue pixels (in mm^2). Phenotypic data were collected from a total of 150 populations with a total of 2392 males and 1745 females (average per population, 15.97 males and 11.63 females).

Trait integration was assessed by calculating pairwise correlations between normalized values of mass, relative head length (head length divided by svl), dorsal and ventral coloration, and blue spot size for a given population and a given sex using the function rcorr in the R package Hmisc (v. 4.7-2), and correlations were plotted using the function corrplot in the R package corrplot (v. 0.92). Phenotypic integration of the five traits per group was calculated using the function integration.Vrel, and values were compared between groups using the function compare.ZVrel (*16*), both contained in the R package geomorph (v. 4.0.5). Phenotypic trajectory analyses (*47*) were conducted along a coastal cline (populations from west to east: LO, NL, VA, ME, GN, RA, SL, LE, ST, MG, VI, CA, CR, CN, VE) comparing patterns of trait change between males and females. To this end, we first fitted the model "traits ~ population_order * sex," where "traits" is a scaled matrix of the five traits, using the function lm.rrpp in the RRPP package (v. 1.3.1). We then used the function trajectory.analysis from the same package to estimate sex differences in magnitude and shape, and the angle between female and male trajectories. The first two PCs of the fitted values were plotted to visualize phenotypic trajectories along the coastal cline for males and females. All analyses in R were conducted in R 4.2.2.

Sequencing data

During sampling of lizards, approximately 1 to 2 cm of the tail tip of each lizard were collected for genetic analyses. A subset of all individuals included in the phenotypic dataset was subjected to genetic analyses. DNA was extracted using the DNeasy blood and tissue kit (Qiagen, USA) using the manufacturer's instruction, and quantity and quality were measured using NanoDrop and Qubit fluorometer assays. WGS was performed for 149 male individuals with insert size of 300 to 500 base pairs (bp) on the Illumina HiSeq X platform by NOVOGENE Ltd. (Hong Kong). For an overview of all sequencing data used in this study, see table S2. Of these 149 individuals, 60 were of pure nigriventris or ancestral phenotypes and included in the main analyses, while 89 were selected from populations in the three geographic areas with intermediate phenotypes (30 from the Rome area, 30 from Tuscany, and 29 from the southern Alps area). The RAD-seq libraries were prepared following the protocol of Peterson *et al.* (*48*) with modifications described by Yang *et al.* (*11*).

Genotyping

WGS raw reads were trimmed using trimmomatic, and quality was assessed using FastQC. Preprocessed reads were mapped against the *P. muralis* reference genome [National Center for Biotechnology Information (NCBI) ID GCA_004329235.1; (*17*)] using bwa (*49*). On average, 98.69% of all reads per sample were mapped to the reference genome. The variant calling procedure was adapted from the best practice recommendations for the Genome Analysis Toolkit (GATK) workflow (*50*) provided by the Broad Institute (*51*). Polymerase chain reaction (PCR) duplicates were removed using picard. We used hard-filtering in GATK with the following options: regions close (within 10 bp) of InDels, phred-scaled *P* value using Fisher's exact test to detect strand bias (FS) above 60, quality by depth (QD) below 2, quality (QUAL) below 30, rank sum test for mapping qualities (MQRankSum) below −12.5, read position rank sum (ReadPosRankSum) below −8, mapping quality (MQ) below 40, and minimum depth of 1192 and 4768, which corresponds to an average minimum of 8× and a maximum of 32× coverage per individual [the average coverage per individual was 21.33 (SD 5.29)]. Finally, only biallelic SNPs (-m2–M2 in bcftools), SNPs with a minor allele frequency of minimum 0.03 (-q 0.03:minor in bcftools), SNPs with

a minimum distance of 10 bp to each other (+prune -w 10 bp -n 1 -N rand in bcftools), and SNPs with less than 50% missing individual genotype calls (--e 'F_MISSING > 0.5' in bcftools) were selected, and individual genotypes were filtered for a genotype quality score above 20 (-S. -e 'FMT/GQ<20' in bcftools). All analyses were restricted to the 19 main chromosomes (18 autosomes plus the Z chromosome), and unplaced scaffolds were ignored (unplaced scaffolds make up only 2.4% of the entire genome assembly). This resulted in a final dataset of 20,928,004 high-quality SNPs.

RAD-seq reads were processed using the STACKS pipeline (*52*, *53*). At first, the "process_radtags" module was used to remove reads with low-quality scores (Phred score < 30), ambiguous base calls, or incomplete barcode or restriction site. Clean reads were mapped to the *P. muralis* reference genome [NCBI ID GCA_004329235.1; (*17*)] using bwa (*49*). We used sorted bam files as input for the reference-based STACKS pipeline that contains modules "gstacks" and "populations" to estimate SNPs using a Marukilow model (*54*). For the populations module, we implemented a 95% cutoff for calling individual loci (-R 0.95) and a minimum minor allele frequency of 0.03. The gstacks output was further filtered in plink removing sites with less than 50% genotyping rate across individuals (-geno 0.5), removing individuals with more than 50% missing data (-mind 0.5), and filtering SNPs in linkage disequilibrium (LD) > 0.8 in 1-kb windows (-indep-pairwise 1 kb 1 0.8). This resulted in a final SNP dataset of 189,724 sites for 890 individuals.

To assign populations to genetic lineages (SA, IT, or hybrid populations), population structure was inferred using fastSTRUCTURE [v. 1.0; (*55*)] on the RAD-seq SNP dataset using $K = 2$. The posterior mean of admixture proportions (Q) was averaged across individuals for each population, and based on this, populations were assigned to the IT ($Q_{average} > 0.95$) or SA ($Q_{average} < 0.05$) genetic lineages or hybrid populations ($0.05 \leq Q_{average} \geq 0.95$). In populations for which phenotypic data, but not genetic data, were collected ($N = 79$), assignments were extrapolated from the geographically closest population with genetic information.

F_{ST} outlier scans and GWAS

To identify genomic regions associated with the nigriventris phenotype, we used F_{ST} outlier scans on WGS data and GWAS on RAD-seq data. For the former, we calculated the fixation index F_{ST}, a measure of genetic differentiation, in groups of lizards with high expression of the nigriventris traits versus the ancestral phenotype. These F_{ST} outlier scans were performed using vcftools (v. 0.1.16) on sliding genomic windows of 5 kb (2.5-kb sliding); patterns of LD decay indicate that SNPs further apart than 5 kb are generally not in linkage (fig. S13). The results remain qualitatively similar when other window sizes were tested (2.5 kb, 10 kb, 50 kb, or 100 SNP windows). These analyses were performed for the SA lineage (10 nigriventris phenotype lizards from three populations versus 10 ancestral phenotype lizards from three populations) and the IT lineage spanning Tuscany and Rome (20 nigriventris phenotype lizards from five populations versus 20 ancestral phenotype lizards from five populations). We considered F_{ST} windows as outliers when their score was above the 99.5th percentile.

GWAS on RAD-seq data was performed separately within the IT and SA lineages to minimize the influence of confounding factors. We aimed to identify statistical associations between individual SNPs and the overall expression of the nigriventris phenotype traits. We therefore derived PC scores of the five focal traits separately for the IT and SA lineage. We only used individuals with no missing data in any of the traits, which reduced the dataset from 497 to 484 individuals in the IT lineage, and from 188 to 186 in the SA lineage. All traits were normalized (scale and center in "prcomp" command). Biplots and loadings of individual traits on PC1 were investigated and showed that all traits load positively and evenly on PC1, which makes PC1 a good composite measure of the overall strength of the syndrome. Since males and females differ in both intensity of coloration and morphology, sex was used as covariate in the GWAS analysis. In addition, within the IT lineage, the first PC described by the genetic data (PC analyzed using 497 IT individuals) was also used as covariate to account for the population genetic structure. Within the SA lineage, we did not account for genetic structure since the first PC of genetic variation is largely aligned with the expression of the nigriventris phenotypes (as expected since this represents introgression from the IT lineage). GWAS analyses were performed in plink [v. 1.90b4.9; (*56*)]. We considered SNPs as outliers significantly associated with syndrome traits when their associated P value was below 0.05 after Benjamini-Hochberg correction.

We reasoned that loci that are causally related to the expression of syndrome traits should show signatures of past or ongoing selection that manifest in changes of overall genetic variation, with lower levels in lizards strongly expressing the nigriventris phenotype. Using vcftools, we calculated Tajima's D, a measure of genetic variation, for 5-kb genomic windows of the WGS data for the same groups as used in the F_{ST} outlier scans, and derived ΔTajima's $D_{(brown\text{-}green)}$ for the ancestral versus nigriventris contrasts within the IT and SA lineages. We considered windows as outliers when their score was above the 99.5th percentile.

To identify genomic regions that are consistently identified as outliers by all three approaches [F_{ST} outlier scan, GWAS, and ΔTajima's $D_{(brown\text{-}green)}$], we intersected the three datasets separately for the IT and SA lineage using bedtools (v. 2.29.2). We first collapsed the 5-kb F_{ST} outlier windows into larger islands in cases where they were maximally 20 kb apart from each other using the bedtools merge command. We then intersected these islands of differentiation with regions spanning ±5 kb of SNPs identified as outlier in the GWAS analyses. Last, we intersected the F_{ST}/GWAS outlier regions with the 5-kb outlier regions with particularly high ΔTajima's D values. These lineage-specific outlier regions satisfying all three criteria were finally intersected to identify a shared set of outliers between the two lineages. To identify genes associated with the outlier regions, we used the R package biomaRt (v. 2.50.3) and extracted genes located within or maximally 10 kb away from the relevant genomic regions. To derive the number of elements shared between datasets for each of the comparisons expected by chance alone, we used 10,000 permutations of drawing genomic windows of the appropriate size and assessed the distribution of these overlapping regions. Similarly, to determine the overlap between the genomic islands of differentiation identified in F_{ST} outlier scans within the IT and SA lineage, we used bedtools intersect and extracted genes located in the vicinity of shared regions using biomaRt as described above.

Characterization of candidate genomic region

Patterns of SNP densities were explored by using the bedtools coverage algorithm on the vcf file before and after filtering. Cumulative read coverages for different sample groups were explored using custom scripts. Abundances of transposable elements (TEs) were

extracted from the Repeat Masker output file (rm.out.gz) made available on the NCBI-FTP server of the reference genome GCF_004329235.1, and the number of TEs per 10-kb windows was summarized. Pattern of linkage disequilibrium was explored by using plink (--r2 --ld-window 100000 --ld-window-kb 1000 --ld-window-r2 0), pairwise R^2 values were summarized for 10-kb windows, and the 95th quantiles were visualized as heatmaps.

We tested if the candidate genomic region has introgressed from the IT lineage into populations of the SA lineage that show the nigriventris phenotype by analyzing pattern of gene flow. We computed D statistics and f_{dM} values using "Dinvestigate" from Dsuite (*57*). Introgression statistics were computed for chromosome 12 in 100-SNP windows (with 50-SNP sliding windows) using the 10 individuals with ancestral phenotype from the SA lineage as P1, 10 individuals with the nigriventris phenotype as P2, 10 individuals with the nigriventris phenotype from Tuscany (IT lineage) as P3, and 10 individuals with the nigriventris phenotype from Rome (IT lineage) as the outgroup. We then tested if the candidate region on chromosome 12 was among the 99.75th percentile of the distribution of values.

De novo genome assemblies and optical maps

To establish if any of the identified patterns were caused by the use of the reference genome stemming from a *P. muralis* individual belonging to the Western European lineage (i.e., the Spanish reference genome) (*17*), we constructed de novo genome assemblies for two individuals from the Italian lineage with contrasting phenotypes: one male with a nigriventris phenotype from SMA (a population with consistent nigriventris phenotypes) and one male with an ancestral phenotype from CDS (a population where all individuals exhibit the ancestral phenotype). The two individuals were captured, processed using our standard procedure (see above), and transferred to Lund University where they were sacrificed and tissues were snap-frozen. Blood samples of both lizards were submitted to the SciLifeLabs Sweden, where high–molecular weight DNA was extracted and subjected to sequencing on an Oxford Nanopore Technology PromethION Platform. For each sample, a 21- and 75-kb library was constructed and sequenced on an individual flow cell (a total of four flow cells). This yielded 120 and 52 Gb of raw reads for the 21- and 75-kb libraries of the SMA sample, and 117 and 90 Gb for the CDS sample, resulting in a raw coverage of 108× for the SMA sample and 129× for the CDS sample.

After completion of sequencing, the raw signal intensity data were basecalled using the Guppy software version 3.5.2 (Oxford Nanopore Technologies). Only reads with a mean qscore (quality) greater than 8 and a read length greater than 15 kb (SMA sample) and 11 kb (CDS sample) were used for assembly. Each dataset amounted to 90 Gb with read N50 of 22 kb (SMA sample) and 18 kb (CDS sample), resulting in approximately 60× coverage of the predicted genome size (1.6 Gb). Wtdbg2 (*58*) was used to assemble the reads with parameters "-p 19 -AS 2 -s 0.05." The initial draft assemblies were polished using the raw nanopore reads with a mean qscore of >8 and a read length of >500 bp with Racon [v. 1.4.11 (*59*)] and one round with Medaka (v. 0.12.1; https://github.com/nanoporetech/medaka) from Oxford Nanopore Technologies.

The scaffolding of the assemblies was further improved by using a conformation capture strategy. Heart tissue of the same two individuals was ground in liquid nitrogen and submitted to the SciLifeLabs Sweden for library preparation using the Omni-C kit (Dovetail Genomics). Libraries were sequenced on one-fourth of an S4 flow cell on a NovaSeq 6000 platform, 2 × 150 bp reads, including the Xp kit. Scaffolding of the initial draft assemblies was performed using Juicer (*60*) (v. 1.6) and 3D-DNA (*61*) (v. 180922). Final assemblies were polished with short read data stemming from the same two individuals. To this end, DNA extracted from blood samples was subjected to library preparation using the TruSeq PCRfree DNA kit (Illumina Inc.). Libraries were sequenced on one SP flowcell (2 × 150 bp) using the NovaSeq 6000 system and v1 sequencing chemistry (Illumina Inc.). This resulted in 145 Gb raw sequence data per sample (approximately 91× coverage) that were used by the software NextPolish. Completeness of the final polished genome assemblies of both samples was evaluated using Benchmarking Universal Single-Copy Orthologs (BUSCO) version 4.0.5 (*62*) with the metazoa odb10 and tetrapoda odb10 lineage gene sets. Statistics on these de novo genome assemblies and comparisons to the Spanish reference genome (*17*) are presented in fig. S7. To substantiate the candidate region, we repeated the variant calling of the WGS data using the de novo assembled genomes as references and performed F_{ST} outlier scans as described above for the Spanish reference genome.

Bionano optical maps were constructed from the same two individuals used for de novo genome assemblies. High–molecular weight DNA extracted from blood samples was transferred to INRAE, France. Labeling and staining of the DNA were performed according to a protocol developed by Bionano Genomics DNA using the Direct Labelling Enzyme (DLE-1), which labels its recognition site without nicking. Loading of the chip and running of the Bionano Genomics Saphyr System were based on the DLS (direct label and stain) labeling and DNA linearization on nanochannel arrays and performed according to the Saphyr System User Guide (https://bionanogenomics.com/support-page/saphyr-system/). Using the software tool Access (v. 1.5.2), optical maps were assessed and in silico aligned to the Spanish reference genome and to the reference genomes derived from the very same individuals. For comparison, the Spanish reference genome and the de novo constructed genome assemblies of an SMA and a CDS individual were subjected to an equivalent in silico digest using the software Access, which enabled alignment of the optical maps.

Phylogenetic analyses of candidate genes

To disentangle the phylogenetic relationships and establish the identity of genes located in the 300-kb region on chromosome 12, we manually curated open reading frames of those genes. We conducted this procedure for the Spanish reference genome whose annotation lacked several genes in this region, and also for the un-annotated de novo genome assemblies of the SMA and CDS samples. The resulting chromosomal coordinates were used for plotting the position of genes that allowed us to compare syntenic relationships between the three genomes, and for deriving the amino acid sequences of the gene products. For the duplicates of the *Pks* and *Ptchd3* genes, amino acid sequences derived from the Spanish reference genome were used in blastp searches against the NCBI database using taxon restrictions to the taxid "squamata" and "crocodilia." The resulting amino acid sequences with high similarity to *Pks* or *Ptchd3* (tables S3 and S4) were aligned with the *P. muralis* sequences derived from all three genomes using the MUSCLE algorithm implemented in MEGA11. Within this software, best protein models were selected using a maximum likelihood (ML) framework, and

ML phylogenies were inferred by using all sites of the alignment and bootstrapping 100 times.

Cellular and transcriptomic basis of coloration

Five lizards exhibiting the nigriventris phenotype and five lizards exhibiting the ancestral phenotype were sacrificed for analyzing the cellular, pigmentary, and transcriptomic basis of coloration. Punch biopsies of 1 or 1.5 mm diameter were taken from dorsal skin and fixed in 0.1 M sodium cacodylate, 2.5% glutaraldehyde, and 2% paraformaldehyde for TEM. For RNA sequencing (RNA-seq), equivalent biopsies were taken and stored in RNAlater. For nigriventris phenotype lizards, these biopsies were taken from exclusively green- and exclusively black-colored patches of skin separately. For pigment extraction, approximately 1 cm^2 of dorsal skin was excised and frozen for downstream chemical analyses. Note that these patches of skin contained both green- and black-colored regions for nigriventris phenotype lizards.

TEM analyses were performed by an in-house imaging facility. In brief, the samples were postfixed in 2% osmium tetroxide in distilled water at 7°C for 1 hour. The specimens were then dehydrated in a graded ethanol series (70% 2 × 10 min, 96% 2 × 10 min, 100% 2 × 15 min) and embedded in Agar 100 via acetone. Ultrathin sections (50 nm) were cut using a Leica UC7 with a diamond knife. Sections were stained with uranyl acetate (2%, 30 min) and Reynolds lead citrate (*63*) (3 min), mounted on copper grids, and viewed with a JEOL 1400 Plus Transmission Electron Microscope at 100 kV. Chemical characterization of pigments was performed in several steps and in parallel on samples of yellow budgerigar (*Melopsittacus undulatus*) feathers to assess the possibility that the yellow pigment in wall lizards has chemical properties similar to psittacofulvin, a parrot-specific yellow pigment. First, the methodology of McGraw *et al.* (*40*) was followed to crudely assess the presence of carotenoids in skin samples. In brief, heated acidified pyridine was used to extract pigments from tissue, followed by transfer to a 1:1 solution of hexane and tert-butyl methyl ether. This procedure resulted in a strongly yellow-colored solution for lizards with green dorsal coloration and an almost transparent solution for the lizards exhibiting brown dorsal coloration, indicating the presence of carotenoids in green, but not brown, skin. We also successfully extracted yellow pigments from lizard skin by incubating biopsies in acetone for 1 day at room temperature [see also (*64*)]. To quantify the different chromatic properties of these pigment extracts (using acetone), absorption spectra were documented using a spectrophotometer for wavelength between 340 and 700 nm. The peaks of the absorption spectra were compared to known absorption spectra of carotenoids. Extracted pigments from budgerigar feathers (using acidified pyridine) and from lizard skin (using acetone) were transferred to the Swedish Metabolomics Center for LC-HRMS using ESI following procedures previously described in (*65*). LC-HRMS analysis was performed on a 1290 Infinity system from Agilent Technologies (Waldbronn, Germany), with an Agilent 6550 QTOF mass spectrometer for MS detection. First, the entire sample set without dilution was analyzed in positive mode. When all samples had been analyzed in this way, the instrument was switched to negative ion mode and a second injection was performed for each sample. For each analysis mode, 2 μl of the dissolved sample was injected onto a 2.1 × 100 mm, 1.7 μm UHPLC Kinetex C18 column (Phenomenex, Torrance, CA, USA). The gradient elution buffers were A (H_2O, 0.1% formic acid) and B [MeOH/IPA (75:25) + 0.1% formic acid], and the flow rate was 400 μl min^{-1}. The initial condition was 20% B, and the proportion of this solvent was linearly increased to reach 100% B at 5 min and it was held at 100% up to 10 min. From 10 to 10.5 min, the column was returned to its initial conditions (20% B) and the flow rate was increased to 0.6 ml min^{-1} for 1 min; these conditions were held for 1 min, after which the flow rate was reduced to 0.4 ml min^{-1} for 0.5 min before the next injection. Compounds were detected with an Agilent 6550 Q-TOF mass spectrometer equipped with a jet stream electrospray ion source operating in positive or negative ion mode. The settings were kept identical between the modes, with exception of the capillary voltage. A reference interface was connected for accurate mass measurements; the reference ions purine (4 μM) and HP-0921 [hexakis(1*H*, 1*H*, 3*H*-tetrafluoropropoxy)phosphazine] (1 μM) were infused directly into the MS at a flow rate of 0.05 ml min^{-1} for internal calibration, and the monitored ions were purine mass/charge ratio (*m*/*z*) 121.05 and *m*/*z* 119.03632; HP-0921 *m*/*z* 922.0098 and *m*/*z* 966.000725 for positive and negative mode, respectively. The gas temperature was set to 150°C, the drying gas flow was set to 16 liters min^{-1}, and the nebulizer pressure was set to 35 psig. The sheath gas temperature was set to 350°C, and the sheath gas flow was set to 11 liters min^{-1}. The capillary voltage was set to 4000 V in negative ion mode. The nozzle voltage was 300 V. The fragmentor voltage was 380 V, the skimmer was 45 V, and the OCT 1 RF Vpp was 750 V. The collision energy was set to 0 V. The *m*/*z* range was 70 to 1700, and data were collected in centroid mode with an acquisition rate of 4 scans s^{-1} (1977 transients/spectrum). Data processing was performed using the Agilent Masshunter Qual version B.07.00 (Agilent Technologies Inc., Santa Clara, CA, USA). The identification of the lutein was confirmed with analysis of a synthetic standard by exact mass, retention time, and MSMS fragmentation.

For RNA-seq analyses, total RNA was extracted from three samples per individual and color group (five individuals for brown coloration, five individuals for green and black coloration) each containing up to five punch biopsies and using the RNeasy Fibrous Tissue Mini Kit (Qiagen). The resulting 45 RNA samples were assessed on a Bioanalyzer for RNA integrity before they were transferred to the SciLifeLabs Sweden where they were subjected to Illumina TruSeq library preparation with poly-A selection and subsequent sequencing on a half S4 flow cell on a NovaSeq 6000 platform (2 × 150 bp). We obtained on average 39.19 (SD = 8.96) million raw reads per sample. Raw reads were preprocessed using trimmomatic (v. 0.39) and mapped to the *P. muralis* reference genome using the STAR software (v. 2.7.9a). The gene annotation (gtf-file) available on the NCBI-FTP-server was used to guide the mapping. To account for the fact that the publicly available annotation in the candidate region is incomplete, we complemented with another mapping using a custom-made file that contained the manually curated gene models. Count tables were derived, and the three replicates per sample were collapsed into a mean value of transcript counts. One individual (LO102) was excluded from subsequent analyses since it was found to be an outlier in PC analyses. Differential gene expression between brown versus green, and brown versus black biopsies was inferred using DESeq2 (v. 1.38.3). The differentially expressed genes were cross-checked with an a priori defined list of genes relevant for coloration based on McLean *et al.* (*41*), and with candidate genes identified in the genomic analyses.

Supplementary Materials

This PDF file includes:

Figs. S1 to S13
Tables S1 to S4
References

REFERENCES AND NOTES

1. S. Lamichhaney, G. Fan, F. Widemo, U. Gunnarsson, D. S. Thalmann, M. P. Hoeppner, S. Kerje, U. Gustafson, C. Shi, H. Zhang, W. Chen, X. Liang, L. Huang, J. Wang, E. Liang, Q. Wu, S. M. Y. Lee, X. Xu, J. Höglund, X. Liu, L. Andersson, Structural genomic changes underlie alternative reproductive strategies in the ruff (Philomachus pugnax). *Nat. Genet.* **48**, 84–88 (2016).
2. A. S. Wilkins, R. W. Wrangham, W. T. Fitch, The "domestication syndrome" in mammals: A unified explanation based on neural crest cell behavior and genetics. *Genetics* **197**, 795–808 (2014).
3. A. L. Pendleton, F. Shen, A. M. Taravella, S. Emery, K. R. Veeramah, A. R. Boyko, J. M. Kidd, Comparison of village dog and wolf genomes highlights the role of the neural crest in dog domestication. *BMC Biol.* **16**, 64 (2018).
4. M. J. Montague, G. Li, B. Gandolfi, R. Khan, B. L. Aken, S. M. J. Searle, P. Minx, L. D. W. Hillier, D. C. Koboldt, B. W. Davis, C. A. Driscoll, C. S. Barr, K. Blackistone, J. Quilez, B. Lorente-Galdos, T. Marques-Bonet, C. Alkan, G. W. C. Thomas, M. W. Hahn, M. Menotti-Raymond, S. J. O'Brien, R. K. Wilson, L. A. Lyons, W. J. Murphy, W. C. Warren, Comparative analysis of the domestic cat genome reveals genetic signatures underlying feline biology and domestication. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 17230–17235 (2014).
5. P. Librado, C. Gamba, C. Gaunitz, C. der Sarkissian, M. Pruvost, A. Albrechtsen, A. Fages, N. Khan, M. Schubert, V. Jagannathan, A. Serres-Armero, L. F. K. Kuderna, I. S. Povolotskaya, A. Seguin-Orlando, S. Lepetz, M. Neuditschko, C. Thèves, S. Alquraishi, A. H. Alfarhan, K. al-Rasheid, S. Rieder, Z. Samashev, H. P. Francfort, N. Benecke, M. Hofreiter, A. Ludwig, C. Keyser, T. Marques-Bonet, B. Ludes, E. Crubézy, T. Leeb, E. Willerslev, L. Orlando, Ancient genomic changes associated with domestication of the horse. *Science* **356**, 442–445 (2017).
6. M. Johnsson, R. Henriksen, D. Wright, The neural crest cell hypothesis: No unified explanation for domestication. *Genetics* **219**, iyab097 (2021).
7. H. E. A. MacGregor, R. A. M. Lewandowsky, P. d'Ettorre, C. Leroy, N. W. Davies, G. M. While, T. Uller, Chemical communication, sexual selection, and introgression in wall lizards. *Evolution* **71**, 2327–2343 (2017).
8. H. E. A. MacGregor, G. M. While, J. Barrett, G. Pérez i de Lanuza, P. Carazo, S. Michaelides, T. Uller, Experimental contact zones reveal causes and targets of sexual selection in hybridizing lizards. *Funct. Ecol.* **31**, 742–752 (2017).
9. R. J. P. Heathcote, G. M. While, H. E. A. MacGregor, J. Sciberras, C. Leroy, P. D'Ettorre, T. Uller, Male behaviour drives assortative reproduction during the initial stage of secondary contact. *J. Evol. Biol.* **29**, 1003–1015 (2016).
10. M. R. Miñano, G. M. While, W. Yang, C. P. Burridge, R. Sacchi, M. Zuffi, S. Scali, D. Salvi, T. Uller, Climate shapes the geographic distribution and introgressive spread of color ornamentation in common wall lizards. *Am. Nat.* **198**, 379–393 (2021).
11. W. Yang, G. M. While, H. Laakkonen, R. Sacchi, M. A. L. Zuffi, S. Scali, D. Salvi, T. Uller, Genomic evidence for asymmetric introgression by sexual selection in the common wall lizard. *Mol. Ecol.* **27**, 4213–4224 (2018).
12. W. Yang, N. Feiner, H. Laakkonen, R. Sacchi, M. A. L. Zuffi, S. Scali, G. M. While, T. Uller, Spatial variation in gene flow across a hybrid zone reveals causes of reproductive isolation and asymmetric introgression in wall lizards. *Evolution* **74**, 1289–1300 (2020).
13. W. Yang, N. Feiner, D. Salvi, H. Laakkonen, D. Jablonski, C. Pinho, M. A. Carretero, R. Sacchi, M. A. L. Zuffi, S. Scali, K. Plavos, P. Pafilis, N. Poulakakis, P. Lymberakis, D. Jandzik, U. Schulte, F. Aubret, A. Badiane, G. Perez i de Lanuza, J. Abalos, G. M. While, T. Uller, Population genomics of wall lizards reflects the dynamic history of the Mediterranean Basin. *Mol. Biol. Evol.* **39**, msab311 (2022).
14. G. M. While, S. Michaelides, R. J. P. Heathcote, H. E. A. MacGregor, N. Zajac, J. Beninde, P. Carazo, G. Pérez i de Lanuza, R. Sacchi, M. A. L. Zuffi, T. Horváthová, B. Fresnillo, U. Schulte, M. Veith, A. Hochkirch, T. Uller, Sexual selection drives asymmetric introgression in wall lizards. *Ecol. Lett.* **18**, 1366–1375 (2015).
15. M. Pavlicev, J. M. Cheverud, G. P. Wagner, Measuring morphological integration using eigenvalue variance. *Evol Biol* **36**, 157–170 (2009).
16. M. A. Conaway, D. C. Adams, An effect size for comparing the strength of morphological integration across studies. *Evolution* **76**, 2244–2259 (2022).
17. P. Andrade, C. Pinho, G. Pérez i de Lanuza, S. Afonso, J. Brejcha, C. J. Rubin, O. Wallerman, P. Pereira, S. J. Sabatino, A. Bellati, D. Pellitteri-Rosa, Z. Bosakova, I. Bunikis, M. A. Carretero, N. Feiner, P. Marsik, F. Paupério, D. Salvi, L. Soler, G. M. While, T. Uller, E. Font, L. Andersson, M. Carneiro, Regulatory changes in pterin and carotenoid genes underlie balanced color polymorphisms in the wall lizard. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 5633–5642 (2019).
18. T. E. Forman, B. J. C. Dennison, K. A. Fantauzzo, The role of RNA-binding proteins in vertebrate neural crest and craniofacial development. *J. Dev. Biol.* **9**, 34 (2021).
19. M. M. Al-Qattan, S. A. Almohrij, The pathogenesis of Pierre Robin sequence through a review of SOX9 and its interactions. *Plast. Reconstr. Surg. Glob. Open* **10**, e4241 (2022).
20. J. Rodor, D. R. FitzPatrick, E. Eyras, J. F. Cáceres, The RNA-binding landscape of RBM10 and its role in alternative splicing regulation in models of mouse early development. *RNA Biol.* **14**, 45–57 (2017).
21. M. S. Prasad, R. M. Charney, L. J. Patel, M. I. García-Castro, Distinct molecular profile and restricted stem cell potential defines the prospective human cranial neural crest from embryonic stem cell state. *Stem Cell Res.* **49**, 102086 (2020).
22. K. Liu, Y. Zhang, D. Liu, Q.-L. Ying, S. Ye, TFCP2L1 represses multiple lineage commitment of mouse embryonic stem cells through MTA1 and LEF1. *J. Cell Sci.* **130**, 3809–3817 (2017).
23. S. Ye, P. Li, C. Tong, Q. L. Ying, Embryonic stem cell self-renewal pathways converge on the transcription factor Tfcp2l1. *EMBO J.* **32**, 2548–2560 (2013).
24. R. McLennan, L. J. Schumacher, J. A. Morrison, J. M. Teddy, D. A. Ridenour, A. C. Box, C. L. Semerad, H. Li, W. McDowell, D. Kay, P. K. Maini, R. E. Baker, P. M. Kulesa, Neural crest migration is driven by a few trailblazer cells with a unique molecular signature narrowly confined to the invasive front. *Development* **142**, 2014–2025 (2015).
25. O. Akkermans, C. Delloye-Bourgeois, C. Peregrina, M. Carrasquero-Ordaz, M. Kokolaki, M. Berbeira-Santana, M. Chavent, F. Reynaud, R. Raj, J. Agirre, M. Aksu, E. S. White, E. Lowe, D. Ben Amar, S. Zaballa, J. Huo, I. Pakos, P. T. N. McCubbin, D. Comoletti, R. J. Owens, C. V. Robinson, V. Castellani, D. del Toro, E. Seiradake, GPC3-Unc5 receptor complex structure and role in cell migration. *Cell* **185**, 3931–3949.e26 (2022).
26. J. A. Morrison, R. McLennan, L. A. Wolfe, M. M. Gogol, S. Meier, M. C. McKinney, J. M. Teddy, L. Holmes, C. L. Semerad, A. C. Box, H. Li, K. E. Hall, A. G. Perera, P. M. Kulesa, Single-cell transcriptome analysis of avian neural crest migration reveals signatures of invasion and molecular transitions. *eLife* **6**, e28415 (2017).
27. F. Brancati, B. Dallapiccola, E. M. Valente, Joubert Syndrome and related disorders. *Orphanet J. Rare Dis.* **5**, 20 (2010).
28. J. J. Johnston, J. K. Teer, P. F. Cherukuri, N. F. Hansen, S. K. Loftus, NIH Intramural Sequencing Center (NISC), K. Chong, J. C. Mullikin, L. G. Biesecker, Massively parallel sequencing of exons on the X chromosome identifies RBM10 as the gene that causes a syndromic form of cleft palate. *Am. J. Hum. Genet.* **86**, 743–748 (2010).
29. G. D. Gilfillan, K. K. Selmer, I. Roxrud, R. Smith, M. Kyllerman, K. Eiklid, M. Kroken, M. Mattingsdal, T. Egeland, H. Stenmark, H. Sjøholm, A. Server, L. Samuelsson, A. Christianson, P. Tarpey, A. Whibley, M. R. Stratton, P. A. Futreal, J. Teague, S. Edkins, J. Gecz, G. Turner, F. L. Raymond, C. Schwartz, R. E. Stevenson, D. E. Undlien, P. Strømme, SLC9A6 mutations cause X-linked mental retardation, microcephaly, epilepsy, and ataxia, a phenotype mimicking Angelman syndrome. *Am. J. Hum. Genet.* **82**, 1003–1010 (2008).
30. G. A. Vega-Lopez, S. Cerrizuela, C. Tribulo, M. J. Aybar, Neurocristopathies: New insights 150 years after the neural crest discovery. *Dev. Biol.* **444**, S110–S143 (2018).
31. T. E. Cruickshank, M. W. Hahn, Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol. Ecol.* **23**, 3133–3157 (2014).
32. N. A. Guadagno, A. Margiotta, S. A. Bjørnestad, L. H. Haugen, I. Kjos, X. Xu, X. Hu, O. Bakke, F. Margadant, C. Progida, Rab18 regulates focal adhesion dynamics by interacting with kinectin-1 at the endoplasmic reticulum. *J. Cell Biol.* **219**, e201809020 (2020).
33. E. R. Funk, N. A. Mason, S. Pálsson, T. Albrecht, J. A. Johnson, S. A. Taylor, A supergene underlies linked variation in color and morphology in a Holarctic songbird. *Nat. Commun.* **12**, 6833 (2021).
34. T. F. Cooke, C. R. Fischer, P. Wu, T. X. Jiang, K. T. Xie, J. Kuo, E. Doctorov, A. Zehnder, C. Khosla, C. M. Chuong, C. D. Bustamante, Genetic mapping and biochemical basis of yellow feather pigmentation in budgerigars. *Cell* **171**, 427–439.e21 (2017).
35. S. Y. Dejgaard, J. F. Presley, Rab18: New insights into the function of an essential protein. *Cell. Mol. Life Sci.* **76**, 1935–1945 (2019).
36. S. Ozeki, J. Cheng, K. Tauchi-Sato, N. Hatano, H. Taniguchi, T. Fujimoto, Rab18 localizes to lipid droplets and induces their close apposition to the endoplasmic reticulum-derived membrane. *J. Cell Sci.* **118**, 2601–2611 (2005).
37. C. Li, X. Luo, S. Zhao, G. K. Y. Siu, Y. Liang, H. C. Chan, A. Satoh, S. S. B. Yu, COPI-TRAPPII activates Rab18 and regulates its lipid droplet association. *EMBO J.* **36**, 441–457 (2017).
38. Y. Deng, C. Zhou, A. H. Mirza, A. T. Bamigbade, S. Zhang, S. Xu, P. Liu, Rab18 binds PLIN2 and ACSL3 to mediate lipid droplet dynamics. *Biochim. Biophys. Acta* **1866**, 158923 (2021).
39. T. Maritzen, H. Schachtner, D. F. Legler, On the move: Endocytic trafficking in cell migration. *Cell. Mol. Life Sci.* **72**, 2119–2134 (2015).
40. K. J. McGraw, J. Hudon, G. E. Hill, R. S. Parker, A simple and inexpensive chemical test for behavioral ecologists to determine the presence of carotenoid pigments in animal tissues. *Behav. Ecol. Sociobiol.* **57**, 391–397 (2005).
41. C. A. McLean, A. Lutz, K. J. Rankin, D. Stuart-Fox, A. Moussalli, Revealing the biochemical and genetic basis of color variation in a polymorphic lizard. *Mol. Biol. Evol.* **34**, 1924–1935 (2017).
42. W. Salzburger, Understanding explosive diversification through cichlid fish genomics. *Nat. Rev. Genet.* **19**, 705–717 (2018).

43. M. J. West-Eberhard, *Developmental Plasticity and Evolution* (Oxford Univ. Press, 2003), p. 794.
44. Q. Wu, X. Sun, W. Yue, T. Lu, Y. Ruan, T. Chen, D. Zhang, RAB18, a protein associated with Warburg Micro syndrome, controls neuronal migration in the developing cerebral cortex. *Mol. Brain* **9**, 19 (2016).
45. Y. Zhong, L. J. Gu, X. G. Sun, S. H. Yang, X. H. Zhang, Comprehensive analysis of patched domain-containing genes reveals a unique evolutionary pattern. *Genet. Mol. Res.* **13**, 7318–7331 (2014).
46. R. Maia, H. Gruson, J. A. Endler, T. E. White, pavo2: New tools for the spectral and spatial analysis of colour in r. *Methods Ecol. Evol.* **10**, 1097–1107 (2019).
47. M. L. Collyer, D. C. Adams, Phenotypic trajectory analysis: Comparison of shape change patterns in evolution and ecology. *Hystrix-Italian J. Mammal.* **24**, 75–83 (2013).
48. B. K. Peterson, J. N. Weber, E. H. Kay, H. S. Fisher, H. E. Hoekstra, Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLOS ONE* **7**, e37135 (2012).
49. H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
50. A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M. A. DePristo, The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
51. M. A. DePristo, E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernytsky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, M. J. Daly, A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43**, 491–498 (2011).
52. J. M. Catchen, A. Amores, P. Hohenlohe, W. Cresko, J. H. Postlethwait, Stacks: Building and genotyping loci de novo from short-read sequences. *G3 (Bethesda)* **1**, 171–182 (2011).
53. N. C. Rochette, A. G. Rivera-Colon, J. M. Catchen, Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Mol. Ecol.* **28**, 4737–4754 (2019).
54. T. Maruki, M. Lynch, Genotype calling from population-genomic sequencing data. *G3 (Bethesda)* **7**, 1393–1404 (2017).
55. A. Raj, M. Stephens, J. K. Pritchard, fastSTRUCTURE: Variational inference of population structure in large SNP data sets. *Genetics* **197**, 573–589 (2014).
56. C. C. Chang, C. C. Chow, L. C. A. M. Tellier, S. Vattikuti, S. M. Purcell, J. J. Lee, Second-generation PLINK: Rising to the challenge of larger and richer datasets. *Gigascience* **4**, 7 (2015).
57. M. Malinsky, M. Matschiner, H. Svardal, Dsuite—Fast D-statistics and related admixture evidence from VCF files. *Mol. Ecol. Resour.* **21**, 584–595 (2021).
58. J. Ruan, H. Li, Fast and accurate long-read assembly with wtdbg2. *Nat. Methods* **17**, 155–158 (2020).
59. R. Vaser, I. Sovic, N. Nagarajan, M. Sikic, Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* **27**, 737–746 (2017).
60. N. C. Durand, M. S. Shamim, I. Machol, S. S. P. Rao, M. H. Huntley, E. S. Lander, E. L. Aiden, Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).
61. O. Dudchenko, S. S. Batra, A. D. Omer, S. K. Nyquist, M. Hoeger, N. C. Durand, M. S. Shamim, I. Machol, E. S. Lander, A. P. Aiden, E. L. Aiden, De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. *Science* **356**, 92–95 (2017).
62. F. A. Simao, R. M. Waterhouse, P. Ioannidis, E. V. Kriventseva, E. M. Zdobnov, BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* **31**, 3210–3212 (2015).
63. E. S. Reynolds, The use of lead citrate at high pH as an electron-opaque stain in electron microscopy. *J. Cell Biol.* **17**, 208–212 (1963).
64. S. V. Saenko, J. Teyssier, D. van der Marel, M. C. Milinkovitch, Precise colocalization of interacting structural and pigmentary elements generates extensive color pattern variation in Phelsuma lizards. *BMC Biol.* **11**, 105 (2013).
65. R. Diamanti, V. Srinivas, A. I. Johansson, A. Nordström, J. J. Griese, H. Lebrette, M. Högbom, Comparative structural analysis provides new insights into the function of R2-like ligand-binding oxidase. *FEBS Lett.* **596**, 1600–1610 (2022).
66. J. Garcia-Porta, I. Irisarri, M. Kirchner, A. Rodríguez, S. Kirchhof, J. L. Brown, A. MacLeod, A. P. Turner, F. Ahmadzadeh, G. Albaladejo, J. Crnobrnja-Isailovic, I. de la Riva, A. Fawzi, P. Galán, B. Göçmen, D. J. Harris, O. Jiménez-Robles, U. Joger, O. Jovanović Glavaš, M. Karış, G. Koziel, S. Künzel, M. Lyra, D. Miles, M. Nogales, M. A. Oğuz, P. Pafilis, L. Rancilhac, N. Rodríguez, B. Rodríguez Concepción, E. Sanchez, D. Salvi, T. Slimani, A. S'khifa, A. T. Qashqaei, A. Žagar, A. Lemmon, E. Moriarty Lemmon, M. A. Carretero, S. Carranza, H. Philippe, B. Sinervo, J. Müller, M. Vences, K. C. Wollenberg Valero, Environmental temperatures shape thermal physiology as well as diversification and genome-wide substitution rates in lizards. *Nat. Commun.* **10**, 4077 (2019).
67. S. Kumar, G. Stecher, M. Suleski, S. B. Hedges, TimeTree: A resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* **34**, 1812–1819 (2017).
68. L. C. Barritt, J. M. Miller, L. R. Scheetz, K. Gardner, M. L. Pierce, G. A. Soukup, S. M. Rocha-Sanchez, Conditional deletion of the human ortholog gene Dicer1 in Pax2-Cre expression domain impairs orofacial development. *Indian J. Hum. Genet.* **18**, 310–319 (2012).
69. A. Zehir, L. L. Hua, E. L. Maska, Y. Morikawa, P. Cserjesi, Dicer is required for survival of differentiating neural crest cells. *Dev. Biol.* **340**, 459–467 (2010).
70. X. Nie, Q. Wang, K. Jiao, Dicer activity in neural crest cells is essential for craniofacial organogenesis and pharyngeal arch artery morphogenesis. *Mech. Dev.* **128**, 200–207 (2011).
71. T. Huang, Y. Liu, M. Huang, X. Zhao, L. Cheng, Wnt1-cre-mediated conditional loss of Dicer results in malformation of the midbrain and cerebellum and failure of neural crest and dopaminergic differentiation in mice. *J. Mol. Cell. Biol.* **2**, 152–163 (2010).
72. Z.-P. Huang, J. F. Chen, J. N. Regan, C. T. Maguire, R. H. Tang, X. R. Dong, M. W. Majesky, D. Z. Wang, Loss of microRNAs in neural crest leads to cardiovascular syndromes resembling human congenital heart defects. *Atertio. Thromb. Vasc. Biol.* **30**, 2575–2586 (2010).
73. A. M. J. Weiner, N. L. Scampoli, T. J. Steeman, C. M. Dooley, E. M. Busch-Nentwich, R. N. Kelsh, N. B. Calcaterra, Dicer1 is required for pigment cell and craniofacial development in zebrafish. *Biochim. Biophys. Acta Gene. Regul. Mech.* **1862**, 472–485 (2019).
74. H. Azzedine, A. Bolino, T. Taïeb, N. Birouk, M. di Duca, A. Bouhouche, S. Benamou, A. Mrabet, T. Hammadouche, T. Chkili, R. Gouider, R. Ravazzolo, A. Brice, J. Laporte, E. LeGuern, Mutations in MTMR13, a new pseudophosphatase homologue of MTMR2 and Sbf1, in two families with an autosomal recessive demyelinating form of Charcot-Marie-Tooth disease associated with early-onset glaucoma. *Am. J. Hum. Genet.* **72**, 1141–1153 (2003).
75. A. Niemann, P. Berger, U. Suter, Pathomechanisms of mutant proteins in Charcot-Marie-Tooth disease. *Neuromolecular Med.* **8**, 217–241 (2006).
76. A. A. Ng, A. M. Logan, E. J. Schmidt, F. L. Robinson, The CMT4B disease-causing phosphatases Mtmr2 and Mtmr13 localize to the Schwann cell cytoplasm and endomembrane compartments, where they depend upon each other to achieve wild-type levels of protein expression. *Hum. Mol. Genet.* **22**, 1493–1506 (2013).
77. H. Ozyurek, O. E. Kayacik, O. Gungor, F. Karagoz, Rare association of Hirschsprung's disease and Joubert syndrome. *Eur. J. Pediatr.* **167**, 475–477 (2008).
78. B. L. Maria, E. Boltshauser, S. C. Palmer, T. X. Tran, Clinical features and revised diagnostic criteria in Joubert syndrome. *J. Child Neurol.* **14**, 583–590 (1999).
79. S. R. Braddock, K. M. Henley, B. L. Maria, The face of Joubert syndrome: A study of dysmorphology and anthropometry. *Am. J. Med. Genet. A* **143a**, 3235–3242 (2007).
80. M. Elmali, Z. Ozmen, M. Ceyhun, O. Tokatlioğlu, L. Incesu, B. Diren, Joubert syndrome with atrial septal defect and persistent left superior vena cava. *Diagn. Interv. Radiol.* **13**, 94–96 (2007).
81. G. Jadhav, D. Teguh, J. Kenny, J. Tickner, J. Xu, Morc3 mutant mice exhibit reduced cortical area and thickness, accompanied by altered haematopoietic stem cells niche and bone cell differentiation. *Sci. Rep.* **6**, 25964 (2016).
82. Y. Redhead, D. Gibbins, E. Lana-Elola, S. Watson-Scales, L. Dobson, M. Krause, K. J. Liu, E. M. C. Fisher, J. B. A. Green, V. L. J. Tybulewicz, Craniofacial dysmorphology in Down syndrome is caused by increased dosage of Dyrk1a and at least three other genes. *Development* **150**, dev201077 (2023).
83. S. Edvardson, Y. Murakami, T. T. M. Nguyen, M. Shahrour, A. St-Denis, A. Shaag, N. Damseh, F. le Deist, Y. Bryceson, B. Abu-Libdeh, P. M. Campeau, T. Kinoshita, O. Elpeleg, Mutations in the phosphatidylinositol glycan C (PIGC) gene are associated with epilepsy and intellectual disability. *J. Med. Genet.* **54**, 196–201 (2017).
84. M. Lukacs, T. Roberts, P. Chatuverdi, R. W. Stottmann, Glycosylphosphatidylinositol biosynthesis and remodeling are required for neural tube closure, heart development, and cranial neural crest cell survival. *eLife* **8**, e45248 (2019).
85. P. Louphrasitthiphol, I. Ledaki, J. Chauhan, P. Falletta, R. Siddaway, F. M. Buffa, D. R. Mole, T. Soga, C. R. Goding, MITF controls the TCA cycle to modulate the melanoma hypoxia response. *Pigment Cell Melanoma Res.* **32**, 792–808 (2019).
86. R. R. Anholt, Olfactomedin proteins: Central players in development and disease. *Front. Cell Dev. Biol.* **2**, 6 (2014).
87. A. Þ. Pétursdóttir, "MITF in the mouse central nervous system: Mitf expression and target genes in the olfactory bulb," thesis, University of Iceland (2010).
88. B. Zhang, K. Peng, J. Che, N. Zhao, L. Jia, D. Zhao, Y. J. Huang, Y. G. Liao, X. He, X. Gong, B. Bao, Single-nucleotide polymorphisms responsible for pseudo-albinism and hypermelanosis in Japanese flounder (Paralichthys olivaceus) and reveal two genes related to malpigmentation. *Fish Physiol. Biochem.* **47**, 339–350 (2021).
89. R. Mei, L. Huang, M. Wu, C. Jiang, A. Yang, H. Tao, K. Zheng, J. Yang, W. Shen, X. Chen, X. Zhao, M. Qiu, Evidence that ITPR2-mediated intracellular calcium release in oligodendrocytes regulates the development of carbonic anhydrase II + type I/II oligodendrocytes and the sizes of myelin fibers. *Front. Cell. Neurosci.* **15**, 1 (2021).

90. D. Anastasiadi, F. Piferrer, M. Wellenreuther, A. Benítez Burraco, Fish as model systems to study epigenetic drivers in human self-domestication and neurodevelopmental cognitive disorders. *Genes (Basel)* **13**, 1 (2022).
91. P. Davoodi, A. Ehsani, R. Vaez Torshizi, A. A. Masoudi, New insights into genetics underlying of plumage color. *Anim. Genet.* **53**, 80–93 (2022).
92. H. Sugawara, Y. Murata, T. Ikegame, R. Sawamura, S. Shimanaga, Y. Takeoka, T. Saito, M. Ikeda, A. Yoshikawa, F. Nishimura, Y. Kawamura, C. Kakiuchi, T. Sasaki, N. Iwata, M. Hashimoto, K. Kasai, T. Kato, M. Bundo, K. Iwamoto, DNA methylation analyses of the candidate genes identified by a methylome-wide association study revealed common epigenetic alterations in schizophrenia and bipolar disorder. *Psychiatry Clin. Neurosci.* **72**, 245–254 (2018).
93. T. Karayannis, E. Au, J. C. Patel, I. Kruglikov, S. Markx, R. Delorme, D. Héron, D. Salomon, J. Glessner, S. Restituito, A. Gordon, L. Rodriguez-Murillo, N. C. Roy, J. A. Gogos, B. Rudy, M. E. Rice, M. Karayiorgou, H. Hakonarson, B. Keren, G. Huguet, T. Bourgeron, C. Hoeffer, R. W. Tsien, E. Peles, G. Fishell, Cntnap4 differentially contributes to GABAergic and dopaminergic synaptic transmission. *Nature* **511**, 236–240 (2014).
94. F.-T. Yin, Caspr4 interaction with LNX2 modulates the proliferation and neuronal differentiation of mouse neural progenitor cells. *Stem Cells Dev.* **24**, 640–652 (2014).
95. P. Ha, X. Luo, C. Li, J. W. Mok, S. S. Lee, N. K. Golnazarian, E. Chen, Z. Zheng, K. Ting, X. Zhang, C. Soo, Inactivation of Cntnap4 in cranial neural crest cells results in craniofacial bone deformities and hydrocephalus. *Plast. Reconstr. Surg. Glob. Open* **7**, 52–53 (2019).
96. C. Li, Z. Zheng, P. Ha, X. Chen, W. Jiang, S. Sun, F. Chen, G. Asatrian, E. A. Berthiaume, J. K. Kim, E. C. Chen, S. Pang, X. Zhang, K. Ting, C. Soo, Neurexin superfamily cell membrane receptor contactin-associated protein like-4 (Cntnap4) is involved in neural EGFL-like 1 (Nell-1)-responsive osteogenesis. *J. Bone Miner. Res.* **33**, 1813–1825 (2018).
97. L. Zeng, H. He, M. Sun, X. Gong, M. Zhou, Y. Hong, Y. Wu, X. Chen, Q. Chen, Runx2 and Nell-1 in dental follicle progenitor cells regulate bone remodeling and tooth eruption. *Stem Cell Res. Ther.* **13**, 486 (2022).
98. L. J. Grange, J. J. Reynolds, F. Ullah, B. Isidor, R. F. Shearer, X. Latypova, R. M. Baxley, A. W. Oliver, A. Ganesh, S. L. Cooke, S. S. Jhujh, G. S. McNee, R. Hollingworth, M. R. Higgs, T. Natsume, T. Khan, G. Á. Martos-Moreno, S. Chupp, C. G. Mathew, D. Parry, M. A. Simpson, N. Nahavandi, Z. Yüksel, M. Drasdo, A. Kron, P. Vogt, A. Jonasson, S. A. Seth, C. Gonzaga-Jauregui, K. W. Brigatti, A. P. A. Stegmann, M. Kanemaki, D. Josifova, Y. Uchiyama, Y. Oh, A. Morimoto, H. Osaka, Z. Ammous, J. Argente, N. Matsumoto, C. T. R. M. Stumpel, A. M. R. Taylor, A. P. Jackson, A. K. Bielinsky, N. Mailand, C. le Caignec, E. E. Davis, G. S. Stewart, Pathogenic variants in SLF2 and SMC5 cause segmented chromosomes and mosaic variegated hyperploidy. *Nat. Commun.* **13**, 6664 (2022).
99. A. Benítez-Burraco, M. Fernández-Urquiza, M. S. Jiménez-Romero, Language impairment with a partial duplication of DOCK8. *Mol. Syndromol.* **11**, 243–263 (2021).
100. Y. Miyamoto, T. Torii, K. Kawahara, A. Tanoue, J. Yamauchi, Dock8 interacts with Nck1 in mediating Schwann cell precursor migration. *Biochem. Biophys. Rep.* **6**, 113–123 (2016).
101. D. Martinez, N. Zuhdi, M. Reyes, B. Ortega, D. Giovannone, V. M. Lee, M. E. de Bellard, Screen for Slit/Robo signaling in trunk neural cells reveals new players. *Gene Expr. Patterns* **28**, 22–33 (2018).
102. D. Bem, S. I. Yoshimura, R. Nunes-Bastos, F. F. Bond, M. A. Kurian, F. Rahman, M. T. W. Handley, Y. Hadzhiev, I. Masood, A. A. Straatman-Iwanowska, A. R. Cullinane, A. McNeill, S. S. Pasha, G. A. Kirby, K. Foster, Z. Ahmed, J. E. Morton, D. Williams, J. M. Graham, W. B. Dobyns, L. Burglen, J. R. Ainsworth, P. Gissen, F. Müller, E. R. Maher, F. A. Barr, I. A. Aligianis, Loss-of-function mutations in RAB18 cause Warburg micro syndrome. *Am. J. Hum. Genet.* **88**, 499–507 (2011).
103. J. Lu, E. Swearingen, M. Hardy, P. Collins, B. Wu, E. Yuan, D. Lu, C. M. Li, S. Wang, M. Ollmann, RAB18 is a key regulator of GalNAc-conjugated siRNA-induced silencing in Hep3B cells. *Mol. Ther. Nucleic Acids* **28**, 423–434 (2022).
104. S. Takáts, L. Lévay, A. Boda, S. Tóth, Z. Simon-Vecsei, A. Rubics, Á. Varga, M. Lippai, P. Lőrincz, G. Glatz, G. Juhász, The Warburg Micro Syndrome-associated Rab3GAP-Rab18 module promotes autolysosome maturation through the Vps34 Complex I. *FEBS J.* **288**, 190–211 (2021).
105. A. Gerondopoulos, R. N. Bastos, S. I. Yoshimura, R. Anderson, S. Carpanini, I. Aligianis, M. T. Handley, F. A. Barr, Rab18 and a Rab18 GEF complex are required for normal ER structure. *J. Cell Biol.* **205**, 707–720 (2014).
106. M. E. Taskinen, E. Närvä, J. R. W. Conway, L. S. Hinojosa, S. Lilla, A. Mai, N. de Franceschi, L. L. Elo, R. Grosse, S. Zanivan, J. C. Norman, J. Ivaska, MASTL promotes cell contractility and motility through kinase-independent signaling. *J. Cell Biol.* **219**, 1 (2020).
107. J. Yang, M. A. Price, G. Y. Li, M. Bar-Eli, R. Salgia, R. Jagedeeswaran, J. H. Carlson, S. Ferrone, E. A. Turley, J. B. McCarthy, Melanoma proteoglycan modifies gene expression to stimulate tumor cell motility, growth, and epithelial-to-mesenchymal transition. *Cancer Res.* **69**, 7538–7547 (2009).
108. A. Nishiyama, A. Lee, C. B. Brunquell, NG2 (Cspg4): Cell surface proteoglycan on oligodendrocyte progenitor cells in the developing and mature nervous system, in *Neural Surface Antigens: From Basic Biology towards Biomedical Applications*, J. Pruszak, Ed. (Academic Press, 2015), pp. 103–112.
109. E. Ampofo, B. M. Schmitt, M. D. Menger, M. W. Laschke, The regulatory mechanisms of NG2/CSPG4 expression. *Cell. Mol. Biol. Lett.* **22**, 4 (2017).
110. M. R. Sun, H. M. Chung, V. Matsuk, D. M. Fink, M. J. Stebbins, S. P. Palecek, E. V. Shusta, R. J. Lipinski, Sonic Hedgehog signaling in cranial neural crest cells regulates microvascular morphogenesis in facial development. *Front. Cell Dev. Biol.* **8**, 590539 (2020).
111. F. Girolamo, I. de Trizio, M. Errede, G. Longo, A. d'Amati, D. Virgintino, Neural crest cell-derived pericytes act as pro-angiogenic cells in human neocortex development and gliomas. *Fluids Barriers CNS* **18**, 14 (2021).
112. K. Lichti-Kaiser, G. ZeRuth, A. M. Jetten, Transcription factor gli-similar 3 (glis3): Implications for the development of congenital hypothyroidism. *J. Endocrinol. Diabetes Obes.* **2**, 1024 (2014).
113. D. W. Scoville, H. S. Kang, A. M. Jetten, Transcription factor GLIS3: Critical roles in thyroid hormone biosynthesis, hypothyroidism, pancreatic beta cells and diabetes. *Pharmacol. Ther.* **215**, 107632 (2020).
114. Y. S. Kim, G. Nakanishi, M. Lewandoski, A. M. Jetten, GLIS3, a novel member of the GLIS subfamily of Krüppel-like zinc finger proteins with repressor and activation functions. *Nucleic Acids Res.* **31**, 5513–5525 (2003).
115. J. Y. Beak, H. S. Kang, Y.-S. Kim, A. M. Jetten, Functional analysis of the zinc finger and activation domains of Glis3 and mutant Glis3(NDH1). *Nucleic Acids Res.* **36**, 1690–1702 (2008).
116. I. C. Welsh, J. Hart, J. M. Brown, K. Hansen, M. Rocha Marques, R. J. Aho, I. Grishina, R. Hurtado, D. Herzlinger, E. Ferretti, M. J. Garcia-Garcia, L. Selleri, Pbx loss in cranial neural crest, unlike in epithelium, results in cleft palate only and a broader midface. *J. Anat.* **233**, 222–242 (2018).
117. E. Ferretti, B. Li, R. Zewdu, V. Wells, J. M. Hebert, C. Karner, M. J. Anderson, T. Williams, J. Dixon, M. J. Dixon, M. J. Depew, L. Selleri, A conserved Pbx-Wnt-p63-Irf6 regulatory module controls face morphogenesis by promoting epithelial apoptosis. *Dev. Cell* **21**, 627–641 (2011).
118. H. J. Parker, B. de Kumar, S. A. Green, K. D. Prummel, C. Hess, C. K. Kaufman, C. Mosimann, L. M. Wiedemann, M. E. Bronner, R. Krumlauf, A Hox-TALE regulatory circuit for neural crest patterning is conserved across vertebrates. *Nat. Commun.* **10**, 1189 (2019).
119. A. G. A. I. V. Howard, An atlas of neural crest lineages along the posterior developing zebrafish at single-cell resolution. *eLife* **10**, e60005 (2021).
120. O. Machon, J. Masek, O. Machonova, S. Krauss, Z. Kozmik, Meis2 is essential for cranial and cardiac neural crest development. *BMC Dev. Biol.* **15**, 40 (2015).
121. L. Selleri, V. Zappavigna, E. Ferretti, 'Building a perfect body': Control of vertebrate organogenesis by PBX-dependent regulatory networks. *Genes Dev.* **33**, 258–275 (2019).
122. A. P. Thiery, A. L. Buzzi, E. Hamrud, C. Cheshire, N. M. Luscombe, J. Briscoe, A. Streit, scRNA-sequencing in chick suggests a probabilistic model for cell fate allocation at the neural plate border. *eLife* **12**, e82717 (2023).
123. E. Vincent, S. Chatterjee, G. H. Cannon, D. Auer, H. Ross, A. Chakravarti, L. A. Goff, Ret loss-of-function decreases neural crest progenitor proliferation and restricts developmental fate potential during enteric nervous system development. bioRxiv 2021.2012.2028.474390 [Preprint] (2022). https://doi.org/10.1101/2021.12.28.474390.
124. K. K. Hozyasz, A. Mostowska, P. Wojcicki, M. Lianeri, P. P. Jagodzinski, Polymorphic variants of genes related to arginine metabolism and the risk of orofacial clefts. *Arch. Oral Biol.* **55**, 861–866 (2010).
125. J. Topczewski, R. M. Dale, B. E. Sisson, Planar cell polarity signaling in craniofacial development. *Organogenesis* **7**, 255–259 (2011).
126. P. P. Dwivedi, R. H. Grose, J. Filmus, C. S. T. Hii, C. J. Xian, P. J. Anderson, B. C. Powell, Regulation of bone morphogenetic protein signalling and cranial osteogenesis by Gpc1 and Gpc3. *Bone* **55**, 367–376 (2013).
127. A. Cagan, T. Blass, Identification of genomic variants putatively targeted by selection during dog domestication. *BMC Evol. Biol.* **16**, 10 (2016).
128. C. Theofanopoulou, S. Gastaldon, T. O'Rourke, B. D. Samuels, A. Messner, P. T. Martins, F. Delogu, S. Alamri, C. Boeckx, Self-domestication in Homo sapiens: Insights from comparative genomics. *PLOS ONE* **12**, e0185306 (2017).
129. D. T. Ahern, P. Bansal, I. V. Faustino, H. R. Glatt-Deeley, Y. Kondaveeti, E. C. Banda, S. F. Pinter, Isogenic hiPSC models of Turner syndrome reveal shared roles of inactive X and Y in the human cranial neural crest network. bioRxiv 2023.03.08.531747 [Preprint] (2023). https://doi.org/10.1101/2023.03.08.531747.
130. M. R. Miñano, Population genetic differentiation and genomic signatures of adaptation to climate in an abundant lizard. *Heredity* **128**, 271–278 (2022).

Acknowledgments: We thank A. Kelber for providing us with feather samples of budgerigars; O. Gustafsson for TEM service; A. Johansson (Swedish Metabolomic Centre, Umeå) for performing LS-MS analyses; J. Abalos for assistance with spectrophotometry; M. Zuffi, R. Sacci, S. Scali, and D. Salvi for logistic and administrative support; field workers who assisted in the collection of phenotypic data; and H. Laakkonen for assistance in the molecular laboratory. We would like to acknowledge support of the National Genomics Infrastructure (NGI)/Uppsala Genome Center and UPPMAX for providing assistance in massive parallel sequencing and

computational infrastructure. Work performed at NGI/Uppsala Genome Center has been funded by RFI/VR and Science for Life Laboratory, Sweden. **Funding:** This work was funded by the Swedish Research Council (2020-03650 to N.F. and E0446501 to T.U.), the European Research Council (948126 to N.F.), the British Ecological Society, the National Geographic Society, the Royal Society of London, and a Wallenberg Academy Fellowship (all to T.U.).
Author contributions: N.F., G.M.W., and T.U. conceived the study. N.F., W.Y., G.M.W., and T.U. conducted field work. N.F. and T.U. conducted phenotypic analyses. N.F. and W.Y. conducted laboratory work. N.F. conducted genetic analyses. W.Y. and T.U. provided input on genetic analyses. I.B. assembled new reference genomes. N.F. and T.U. wrote the first draft of the manuscript with input from G.M.W. All authors commented on the draft and approved the final manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. RNA-seq data are available at NCBI under BioProject ID PRJNA1005299. All data related to the construction of genome assemblies (Oxford Nanopore long reads, Hi-C sequence data, short reads used for polishing, and final genome assemblies) are available under BioProject PRJNA662234. WGS data are available under BioProject PRJNA1008404. RAD-seq data are available under BioProject PRJNA486080. Code to replicate the results presented in this publication is available at Zenodo (https://doi.org/10.5281/zenodo.10396759).

Submitted 20 September 2023
Accepted 28 February 2024
Published 3 April 2024
10.1126/sciadv.adk9315