nature communications

Article

https://doi.org/10.1038/s41467-026-78632-4

Genomic bases of short-term evolution in wild barn owls revealed by long-term monitoring and population-scale sequencing

Received: 4 December 2025

Accepted: 1 October 2026

Published online: 12 October 2026

Check for updates

Tristan Cumer[1,2] ✉, Alexandros Topaloudis[1,2], Eleonore Lavanchy[1,2], Anne-Lyse Ducrest[1], Sonia Sarmiento Cabello[1], Anna Hewett[1,2], Paolo Becciu[1,3], Céline Simon[1], Julien Marquis[4], Aline Charpagne[4], Bettina Almasi[3], Alexandre Roulin[1] & Jérôme Goudet[1,2] ✉

Understanding how wild populations undergo rapid phenotypic evolution requires linking phenotypic change to its underlying genomic architecture over contemporary timescales. Here, we leverage a 30-year study of Swiss barn owls (*Tyto alba*) to examine long-term changes in two melanin-based traits (plumage coloration and spottiness) that have previously been linked to multiple fitness-related traits. We document a consistent temporal shift toward darker and more heavily spotted individuals and show, using whole-genome sequencing of 3,102 individuals, that these traits are controlled by a small number of loci of major effect with partially overlapping architectures. Temporal allele frequency analyses reveal subtle but consistent changes at these loci. By comparing empirical trajectories to simulations, we show that genetic drift alone is insufficient to explain the observed patterns at one major locus, indicating a role for selection. Our results demonstrate how selection acting on a small number of loci can contribute to rapid phenotypic evolution of continuous traits in the wild and highlight the value of long-term genomic and phenotypic monitoring for disentangling evolutionary processes.

Understanding how populations evolve, what mechanisms drive micro-evolutionary changes in phenotypic traits, and how genetic variation translates into variation in fitness are central questions in evolutionary biology. These questions have intrigued evolutionary biology for decades[1–4]. Addressing these questions is not only essential to evolutionary theory but also critical for understanding how populations can rapidly adapt to changing environments. Achieving such comprehensive understanding of short-term evolution requires understanding how genetic variation translates into phenotypic variation and how evolutionary processes, such as drift and selection, shape these traits over time[5–7].

Evolutionary theory predicts that genetically encoded traits evolve toward their optimal values through natural selection, favoring individuals best suited to their environments[1,3]. In practice, however, identifying adaptive phenotypic change across generations and linking it to the underlying genotypes remains challenging. Traditional quantitative genetics approaches can detect heritable changes in life history traits[8] but often fail to pinpoint specific causal loci. Conversely, population genetics approaches can identify genomic regions associated with phenotypic traits[9]. Recent advances in whole-genome sequencing and downstream analyses[10] now enable the study of large cohorts in wild populations[11,12]. By combining quantitative genetic methods with high-quality genomic data, we can explore the genomic basis of phenotypic evolution in the wild with fine-scale resolution.

Coloration offers a powerful system for studying the links between genotype, phenotype, as well as their joint evolution[13]. In

[1]Department of Ecology and Evolution, University of Lausanne, Lausanne, Switzerland. [2]Swiss Institute of Bioinformatics, Lausanne, Switzerland. [3]Swiss Ornithological Institute, Sempach, Switzerland. [4]Genomic Technologies Facility, Faculty of Biology and Medicine, University of Lausanne, Lausanne, Switzerland. ✉e-mail: tristan.cumer@unil.ch; jerome.goudet@unil.ch

particular, avian coloration has emerged as a key model for investigating these processes, owing to the diversity of color phenotypes and the relative ease with which they can be quantified[14]. Recent advances in genomic tools have led to substantial progress in identifying the genetic architecture of coloration in birds and other taxa, revealing that adaptive color traits can often be controlled by a limited number of loci of moderate to large effect[15,16]. Despite this progress, most genomic studies of avian coloration have focused on discrete polymorphisms or categorical traits, often associated with relatively simple genetic architectures[14,17]. In contrast, continuously varying color traits, common within natural bird populations, remain comparatively underexplored[18]. This gap is partly due to the difficulty of disentangling environmental and genetic contributions to quantitative variation, as well as the expectation that such traits are influenced by multiple loci of varying effect sizes and complex regulatory interactions[17]. Nonetheless, continuous variation in avian coloration is widespread and evolutionarily important, reflecting the combined influence of multiple selective pressures (including mate choice, predation, and environmental conditions) that often act simultaneously on the same trait, generating complex evolutionary dynamics across spatial and temporal scales[14,19,20].

Color polymorphism in the barn owl (*Tyto alba*) offers a particularly suitable system to explore the genomic basis of fitness-related traits. In Central Europe, barn owls exhibit continuous variation in two melanin-based traits on the ventral body: plumage coloration (ranging from white to rufous) and spottiness (number and size of black feather spots) (Fig. 1a). Both traits are related to fitness: plumage coloration is linked to fecundity[21,22] and hunting strategy[23,24]. Spottiness, in turn, is related to testosterone levels[25], stress response[26,27] and immunity[28,29]. Both traits are heritable, and a genetic correlation between them suggests shared genomic underpinning[30]. Females tend to be more melanic (redder and spottier) than males[31], and spottiness shows higher heritability in males[30], suggesting sex-linked genetic effects[32].

A recent continental-scale GWAS identified two loci associated with plumage coloration: a major SNP in the *MC1R* gene on chromosome 11 (autosome), and a second locus on the Z sex chromosome[18]. Together, these loci explain a large proportion of the variation in coloration, highlighting a relatively simple genetic architecture for this trait. However, the remaining variation suggests additional genetic factors or interactions that remain to be characterized. In parallel, the genetic basis of spottiness has yet to be investigated. These gaps highlight the importance of investigating whether these two traits share common genetic determinants, if sex-specific effects contribute to their architecture, and how these factors shape their evolutionary trajectories.

Taking advantage of over 30 years of phenotypic and life-history monitoring in a wild barn owl population in western Switzerland, we investigate the microevolutionary dynamics of coloration and spottiness. Over this period, we observe a clear phenotypic trend toward redder and more spotted individuals. To characterize the genetic architecture underlying this trend, and to assess whether the observed temporal allele frequency changes are more consistent with directional selection or stochastic processes, we sequence 3102 individuals, improving resolution of genetic relatedness and refining heritability estimates. We identify three causal loci associated with these traits, including a shared signal on the Z chromosome. Gene expression analyses confirm the link between Z-linked genotypes and phenotypes, suggesting a mechanistic basis for the observed sexual dimorphism. We further model the sex-specific genetic architecture of the traits, revealing differing allelic effects between males and females. Finally, forward simulations show that selection, rather than genetic drift alone, can explain the observed allele frequency shifts.

Results and discussion

Phenotypic changes in the two melanic traits

In the Swiss barn owl population, plumage color and spottiness show substantial variation (Fig. 1a, b). Color scores range from −8 (white) to −1 (red), with a mean of −5.21 (SD = 1.78; Fig. S1), while spottiness, measured as the area covered by black feather spots, average 1.71 cm^2 (SD = 1.54; Fig. S1; $n = 8278$). The two traits are moderately correlated (Pearson's $r = 0.33$; Fig. 1b and Figs. S2–S4), with a significantly stronger correlation in males ($r = 0.37$) than in females ($r = 0.18$; Fig. S2).

Phenotypic data collected over the 30-year study period reveal a clear evolutionary trend toward redder and more heavily spotted individuals (Fig. 1c). Linear mixed-effects models confirm a significant temporal increase in both traits: color scores increased by 0.04 units per year, and spottiness increased by 0.05 cm^2 per year. Over 30 years, these changes represent shifts of 0.66 and 0.91 standard deviations from the respective means (Table S1 and Supplementary Data 1 for color; Table S2 Supplementary Data 2 for spottiness). This trend was consistent across both males and females, showing increased melanism over time (Figs. S5, S6, and Tables S3, S4 for sex-specific models), despite the species' weak sexual dimorphism (Fig. S7) and accounting for the slight coloration difference between fledglings and adults (Fig. S8). Moreover, the trend held when separating individuals born within the study area (residents) from those born outside it (immigrants, Figs. S9 and S10). This pattern further suggests that the observed evolutionary trend is not restricted to the focal population, but likely occurs over a broader spatial scale. Finally, and because sample sizes were smaller in the earliest years of monitoring, we evaluated the robustness of phenotypic trends by re-running the analyses using only the last 25 years of data, during which sample sizes were consistently large. Estimates of temporal change in both coloration and spottiness were nearly identical to those obtained using the full dataset (Tables S7, S8, and Supplementary Data 3, 4), indicating that the inferred phenotypic trends are not driven by early years with limited sampling. Together, these results show that both coloration and spottiness have undergone directional change in this population, with barn owls becoming progressively redder and more spotted over the past three decades. Because years are confounded with any directional environmental change over time, these phenotypic models do not, on their own, allow a full disentanglement of genetic and environmental contributions to the observed trend; we therefore next investigate the genomic architecture underlying these traits.

Population-scale whole-genome sequencing of the Swiss barn owl population

To investigate the genetic architecture underlying variation in coloration and spottiness, we assembled a whole-genome dataset consisting of 3102 barn owls from the Swiss population, sampled over a 31-year period. We leveraged a previously published high-coverage dataset of 502 owls from across the Western Palearctic, including 346 from our focal Swiss population, to construct a reference panel representing continental genomic diversity (10,451,268 phased SNPs; Figs. S11–S14[33]). Using this panel, we imputed genotypes for an additional 2778 Swiss individuals sequenced at low coverage (mean 1.95X; Fig. S15), achieving high imputation accuracy (mean $r^2 = 0.968$). The final dataset comprised 3280 individuals, of which 3102 were phenotyped for coloration and spottiness, providing an comprehensive resource for dissecting the genomic basis and evolutionary dynamics of these traits.

High heritability and large-effect loci shared between the two traits

Using whole-genome data from 3102 owls, we first confirmed the high heritability of both plumage coloration and spottiness. Genetic relatedness was estimated using a genomic relationship matrix (GRM) derived from SNP data, and heritability was estimated using an animal

model. Coloration showed a heritability of $h^2 = 0.82 \pm 0.02$ (Table S9), consistent with previous estimates based on pedigree data ($h^2 = 0.84 \pm 0.02$)[30]. Spottiness exhibited similarly high heritability ($h^2 = 0.80 \pm 0.02$; Table S10), exceeding prior estimates for its two components-number of spots ($h^2 = 0.57 \pm 0.03$) and spot diameter ($h^2 = 0.67 \pm 0.03$)[30], likely due to differences in how the trait was measured. Unlike previous analyses treating the components separately, our approach models total spottiness as an integrated phenotype. Together, these results demonstrate a strong genetic basis for both traits.

We next estimated the genetic correlation between coloration and spottiness to assess the degree of shared genetic architecture. The overall genetic correlation was moderate ($r = 0.24 \pm 0.08$), indicating partial but not complete genetic overlap. However, this correlation was substantially higher in males ($r = 0.55 \pm 0.11$) than in females ($r = 0.11 \pm 0.16$), suggesting sex-specific differences in genetic control. These results, consistent with previous findings in barn owls[30], support the involvement of sex-linked genetic factors, particularly variation on the Z chromosomes, in shaping these traits.

Genome-wide association studies reveal shared and distinct large-effect loci

To identify genomic regions associated with coloration and spottiness, we performed independent genome-wide association studies (GWAS) for each trait (Figs. 2, S16–S21, and Supplementary Data 5–10). For coloration, we identified two significant loci (after Bonferroni threshold: $-\log_{10}(p) > 8.11$; Fig. 2a, c, d). The top signal was on chromosome 11 (Super-scaffold 26) at position 22,522,039, a known nonsynonymous variant in the *MC1R* gene[18,34]. This variant is referred to as $MC1R_{white}$ (G) and $MC1R_{red}$ (A) based on associated phenotypes. A second peak was located on the sex chromosome (chromosome Z—Super-Scaffold 42) in a region (between 29,200,000 and 30,200,000 bp) encompassing 16 genes, with the most associated SNP on position 29,826,792 (Fig. 2a, d, and Supplementary Data 5). This variant is referred to as Z_{white} (G) and Z_{red} (C), based on associated phenotypes. Both loci confirm associations previously reported at the continental scale[18]. To account for spatial heterogeneity in coloration across the ventral surface, we ran additional GWAS separately for breast and belly coloration (Figs. S17, S18, and Table Supplementary Data 6, 7). The breast-specific GWAS revealed an additional locus on chromosome 4 (Super-Scaffold 6) at 27,439,651 bp, located within an intron of the *CORIN* gene ($-\log_{10}(p) = 9.43$), not detected in the full-body analysis.

For spottiness, the GWAS revealed a single significant region on the Z chromosome, overlapping with the region associated with coloration (Fig. 2b, e). The top SNP was located at position 29,811,663 ($-\log10(p) > 8.11$; Fig. 2b, e, and Supplementary Data 8). The associated alleles are referred to as $Z_{immaculate}$ (G) and $Z_{spotted}$ (A), based on phenotype. To further dissect spottiness, we conducted a GWASs on its two components, mean spot diameter and total number of spots. The GWAS on spot diameter showed a pattern similar to the total surface area, with significant associations primarily on the Z chromosome (Fig. S20 and Supplementary Data 9). In contrast, the GWAS for the number of spots identified multiple significant regions, including the MC1R variant on chromosome 11 (Fig. S21 and Supplementary Data 10).

The GWAS results indicate an oligogenic architecture for both traits, with a few large-effect loci driving much of the variation. Notably, both traits share a common genetic architecture, with overlapping signals at the Z chromosome and the *MC1R* gene: (i) there was almost complete linkage disequilibrium ($R^2 = 0.934$) between the top SNPs for Z coloration and Z spottiness, indicating near-identical genotype-phenotype associations; (ii) the *MC1R* locus, classically known for its role in pigmentation across vertebrates, also contributes to variation in both color and spot number. By precisely localizing the shared region on the Z chromosome and confirming its role in both traits, our work refines previous estimates of the genetic correlation between coloration and spottiness[30]. Interestingly, this Z-linked region has also been associated with crown patch color variation in North American rosy finches[35], suggesting a potentially conserved pigmentation pathway across species.

Differential gene expression links genotype to phenotype

To explore potential molecular correlates of the GWAS-identified Z-linked region, we conducted exploratory targeted gene expression analyses. Specifically, we quantified expression of 13 candidate genes-*MC1R* and 12 genes located in the Z-linked region associated with both traits (Figs. S22, S23, and Supplementary Data 11), in developing breast feathers from 45 nestlings (28 males and 17 females). Sampling was timed to capture the onset of feather development, when black apical spots and pheomelanin-based coloration first become visible.

The nonsynonymous *MC1R (V126I)* has previously been associated with both coloration and spot number, as well as with variation in *MC1R* gene expression[36]. We included *MC1R* in our expression panel, but due to a sampling strategy designed to maximize Z-linked genotype variation, the resulting *MC1R* genotype distribution was highly unbalanced (38 homozygous $MC1R_{white}$, 6 heterozygous and 1 homozygous $MC1R_{red}$), limited statistical power to assess its regulatory effect. We thus focused on the Z-linked region, where we quantified the relative expression levels of all 12 genes within the top GWAS peak (Fig. 3a). We first designed an assay to genotype the individuals at the Z locus (see Method section and Supplementary Data 12 for details) and confirmed that genotypes at the diagnostic SNP were significantly associated with both coloration and spottiness in these individuals (Fig. 3b, c). We then used a linear mixed model to test whether gene expression levels varied with genotype at the Z locus. As expected, expression levels differed significantly between sexes for all 12 genes, typically higher in males, consistent with incomplete dosage compensation in birds[37–39]. To account for this, we ran sex-specific models to test for genotype-expression associations (Table S11). Of the 12 genes tested, two showed significant genotype-dependent expression: *PDE8B* varied significantly with genotype in males (Fig. 3d and Tables S12, S13), *WDR41* with genotype in females (Fig. 3e and Tables S14, S15). To assess whether these expression differences translated into phenotypic effects, we tested for associations between expression and trait values. Only one significant link emerged with *PDE8B* expression in males associated with spottiness, but not coloration (Figs. 3f, g, S24, and Table S16).

Our findings suggest a potential link between genotype, gene expression, and the spottiness phenotype in barn owls. However, this association remains preliminary and warrants further investigation. Only a small subset of tested genes showed genotype-dependent expression, and only one gene (*PDE8B*) was associated with a phenotypic trait, specifically spottiness in males. Moreover, the strongly unbalanced genotype distribution at the *MC1R* locus in the expression dataset precluded an assessment of *MC1R* expression effects, despite *MC1R* showing the strongest signal of selection at the population level. Despite this constraint, we identified *PDE8B* as a promising candidate gene. *PDE8B* encodes a cAMP-specific phosphodiesterase[40], and although it has not previously been implicated in pigmentation, other phosphodiesterases, such as PDE5, are known to influence melanin synthesis via the cAMP/PKA/CREB/MITF signaling pathway in humans[41,42]. Further research is needed to clarify the molecular role of the Z-linked region. Given its location on the sex chromosome and the slight sexual dimorphism in barn owls, we also considered whether sex-specific genetic architecture might contribute to variation in these traits.

Sexual dimorphism and sex specific genetic architecture of the two traits

The barn owl is a sexually dimorphic species, with males typically exhibiting whiter and fewer ventral spots than females[43]. To investigate

https://doi.org/10.1038/s41467-026-78632-4

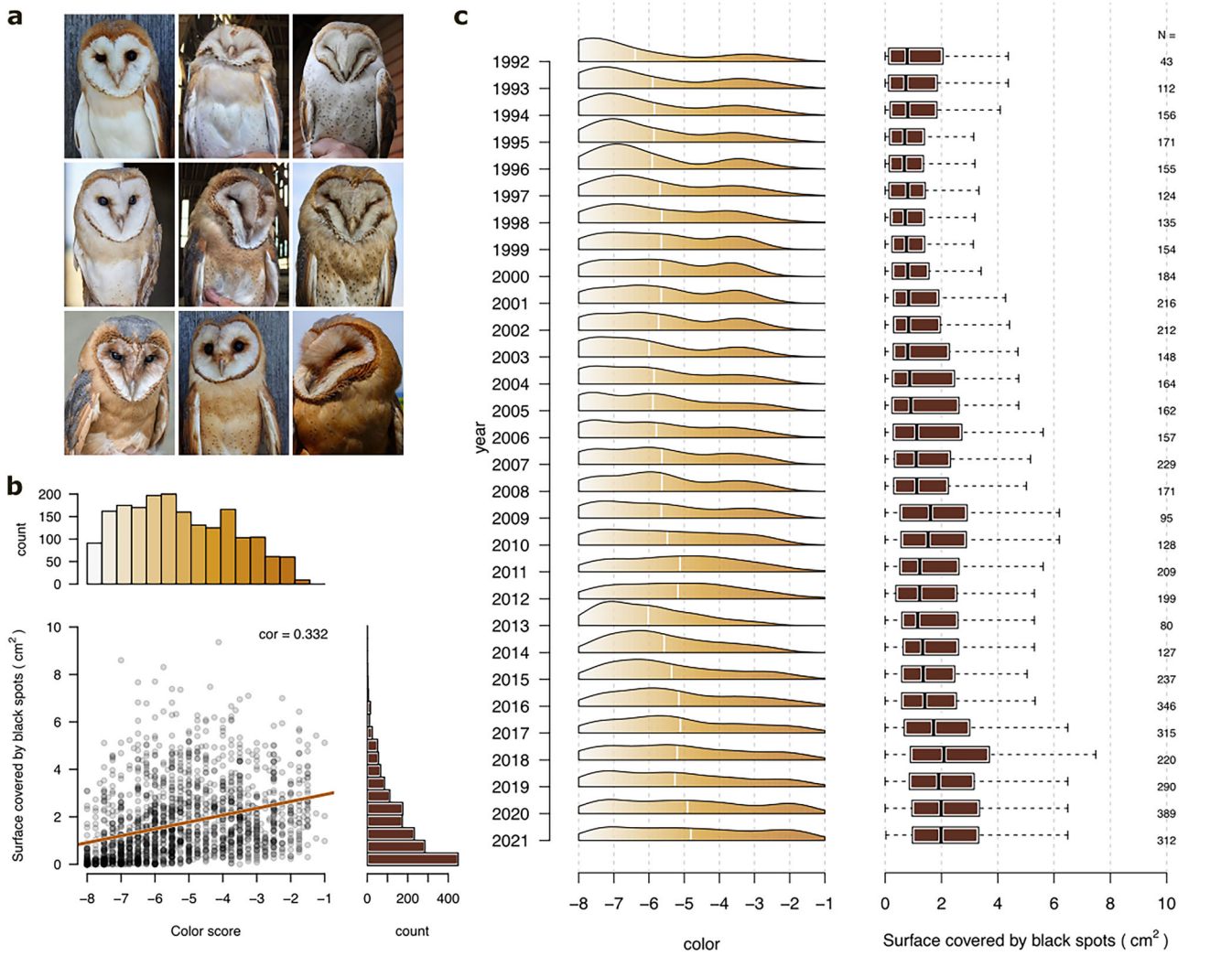


Fig. 1 | Variation in plumage coloration and spottiness of Swiss barn owls over the last 30 years. a Photographs illustrating the range of ventral plumage coloration (rows) and spottiness (columns) observed in Swiss barn owls. Coloration varies from white (−8) to dark reddish (−1), while spottiness reflects variation in the density and size of black spots. **b** Relationship between coloration and spottiness (in cm^2). The upper panel shows histograms of the distribution of coloration, while the right panel shows the distribution of spottiness. The red line represents the fitted linear regression between coloration and spottiness. The correlation coefficient between the two traits is shown in the upper-right corner of the scatterplot ($r = 0.332$). **c** Temporal trends in coloration and spottiness in alive adults between 1992 and 2021 ($n = 1916$). Left: density plots of coloration for each year, with the white line indicating the annual mean. Right: boxplots of spottiness for all individuals alive in each year. Box plots show the median (center line), the 25th and 75th percentiles (box bounds), and whiskers extending to the smallest and largest observations within 1.5 × the interquartile range (IQR) of the lower and upper quartiles, respectively. Sample sizes are reported on the right. Image credit for (**a**), from left to right. Raw 1: Ana Paula Machado (APM)−Guillaume Rapin (GR)−Clément Grandjean. Raw 2: Jeremy Bierer (JB)−(GR)−Lukas Linder. Raw 3: JB−APM−Isabelle Henry.

how the genomic variants identified in our GWAS contribute to phenotypic variation in each sex, and to quantify the relative effect of each locus, we fitted Bayesian generalized linear models separately for each trait. Genotype at each locus was fitted as fixed effects. Given the slight sexual dimorphism and the fact that one of the loci most strongly associated with both traits is located on the Z chromosome, we constructed sex-specific models. In total, we ran four models, one for each trait (coloration and spottiness) and for each sex (Fig. 4). To control for ontogenic changes in phenotype, we included the developmental stage of each individual (fledgling or adult) as a covariate in all models[44].

The two sex-specific models for plumage coloration confirm the contribution of three loci (*MC1R*, the Z-linked region, and *CORIN*) in both males and females. The models explained 80.7% (±1.6) of the total variance in coloration in males and 75.3% (±2.1) in females (Fig. 4a, b). These values closely match the previously estimated heritability of the trait (0.82 ± 0.02; see *High heritability and large-effect loci shared between the two traits* section), supporting the interpretation of plumage color as an oligogenic trait with most of the variance accounted for by few loci. Genotypes at all three loci had significant effect sizes (Tables S17, S18, and Supplementary Data 13, 14). Consistent with earlier findings, *MC1R* emerged as the major contributor to color variation[18,45]), followed by the Z-linked locus[18]. The significant effect of *CORIN*, however, represents a previously unreported result in this population, but is consistent with an emerging role of CORIN in avian coloration more broadly. CORIN has been implicated in pigmentation variation in several bird species, including associations with melanin-based coloration and patterning[46,47], and may act as a modifier locus influencing pigmentation pathways[48]). Its detection here highlights how increasing sample size can reveal loci with smaller effects in oligogenic traits. While all three loci influenced coloration in both sexes, the magnitude and nature of their effects differed. Notably, a significant epistatic interaction between *MC1R* and the Z-linked genotype was detected only in females (Table S18). In females, the

https://doi.org/10.1038/s41467-026-78632-4

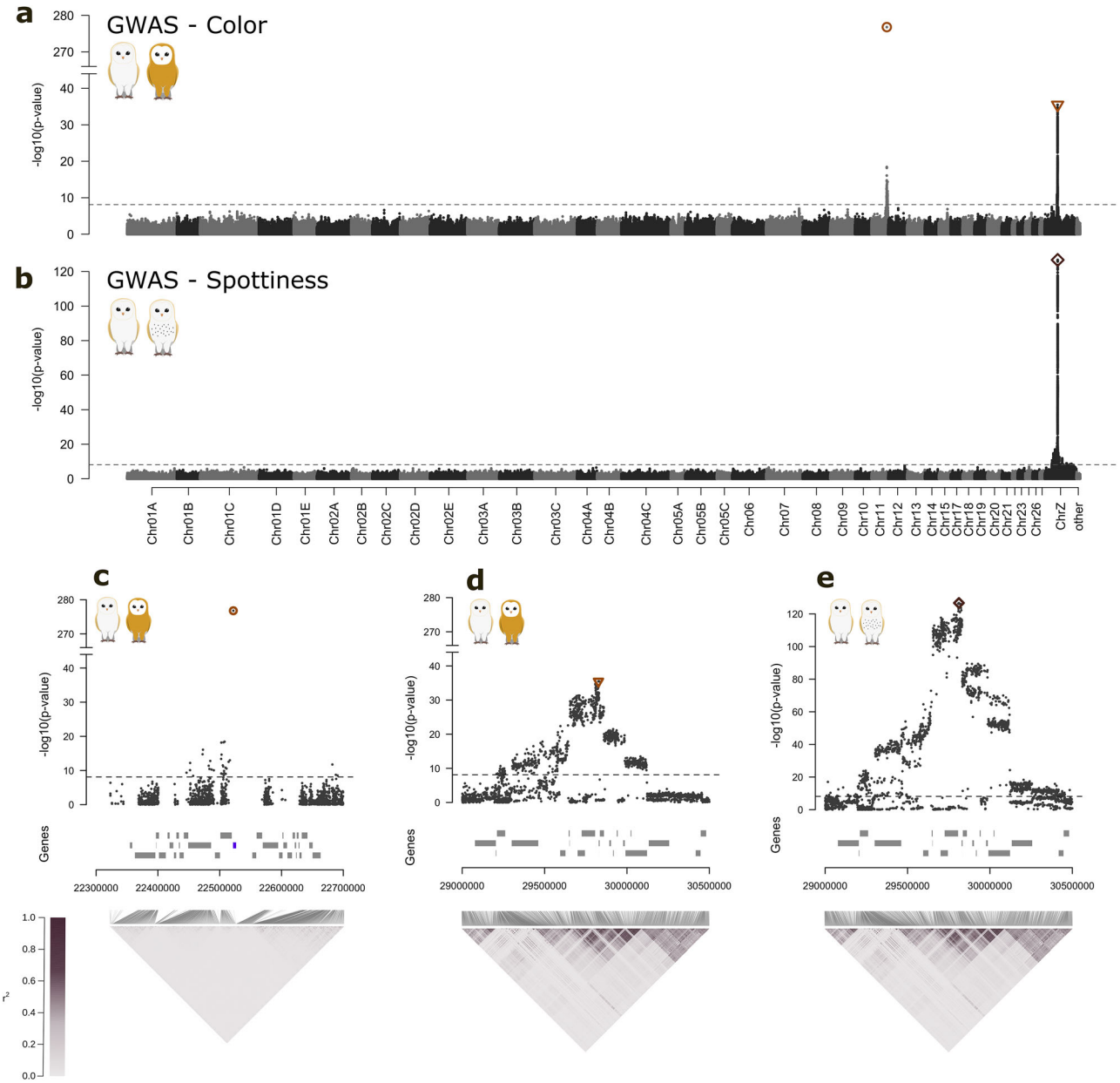


Fig. 2 | Genome-wide association study (GWAS) of plumage traits in Swiss barn owls reveals an oligogenic architecture for the two traits. Association scores ($-\log_{10}$ *p*-values) for each SNP across the genome with **a** plumage coloration and **b** plumage spottiness. Alternating colors distinguish successive chromosomes. The dotted line indicates the Bonferroni-corrected significance threshold ($-\log^{10}$(*p*-value) = 8.11). Circle, triangle, and diamond point to the most significant SNP in each region presented in (**c**–**e**). Zoomed views of regions showing strong associations with coloration (**c**, **d**) or spottiness (**e**). Upper panels display association scores for SNPs within each region, with dotted lines marking the Bonferroni threshold. Rectangles below indicate annotated genes. The blue rectangle in (**c**) indicates the *MC1R* gene. Lower panels show mean linkage disequilibrium (r^2) between SNPs in sliding windows of 10 SNPs. Lines connect each LD block to its genomic position on the x-axis.

accumulation of "red" alleles at *MC1R* and the Z-locus did not result in a linear increase in redness. Instead, coloration reached a plateau (Fig. 4b), consistent with a non-additive (epistatic) relationship between these two major pigmentation genes. This epistatic interaction has been previously hypothesized[45] and is here supported by genomic data.

To model the contribution of the two loci, *MC1R* and the Z-linked region, to spottiness in each sex, we used a generalized linear model with a hurdle-gamma distribution. This modeling approach separates (i) the effect of each locus on the presence or absence of spots (the hurdle part, estimating the probability that spottiness equals zero) and (ii) the contribution on the total surface area covered by spots in individuals using the gamma part of the model[49]. In males, both loci significantly contributed to the presence/absence of spots. However, the Z-linked locus consistently had a larger effect than *MC1R* on both the likelihood of being spotted and the extent of spottiness (Fig. 4c, Tables S19, S20, and Supplementary Data 15, 16). In females, while the Z-linked locus had a significant impact on both the presence of spots and the total spotted surface, the effect of *MC1R* was not statistically significant (Fig. 4d and Table S19). These results suggest a slightly different architecture for spottiness between the sexes. The models explained 58.3% (±2.5) and 53.9% (±3.1) of the variance in spottiness in males and females, respectively. While this represents a substantial proportion of the trait variance, it falls short of the higher heritability previously estimated for spottiness (0.80 ± 0.02; see *High heritability and large-effect loci shared between the two traits* section). This

https://doi.org/10.1038/s41467-026-78632-4

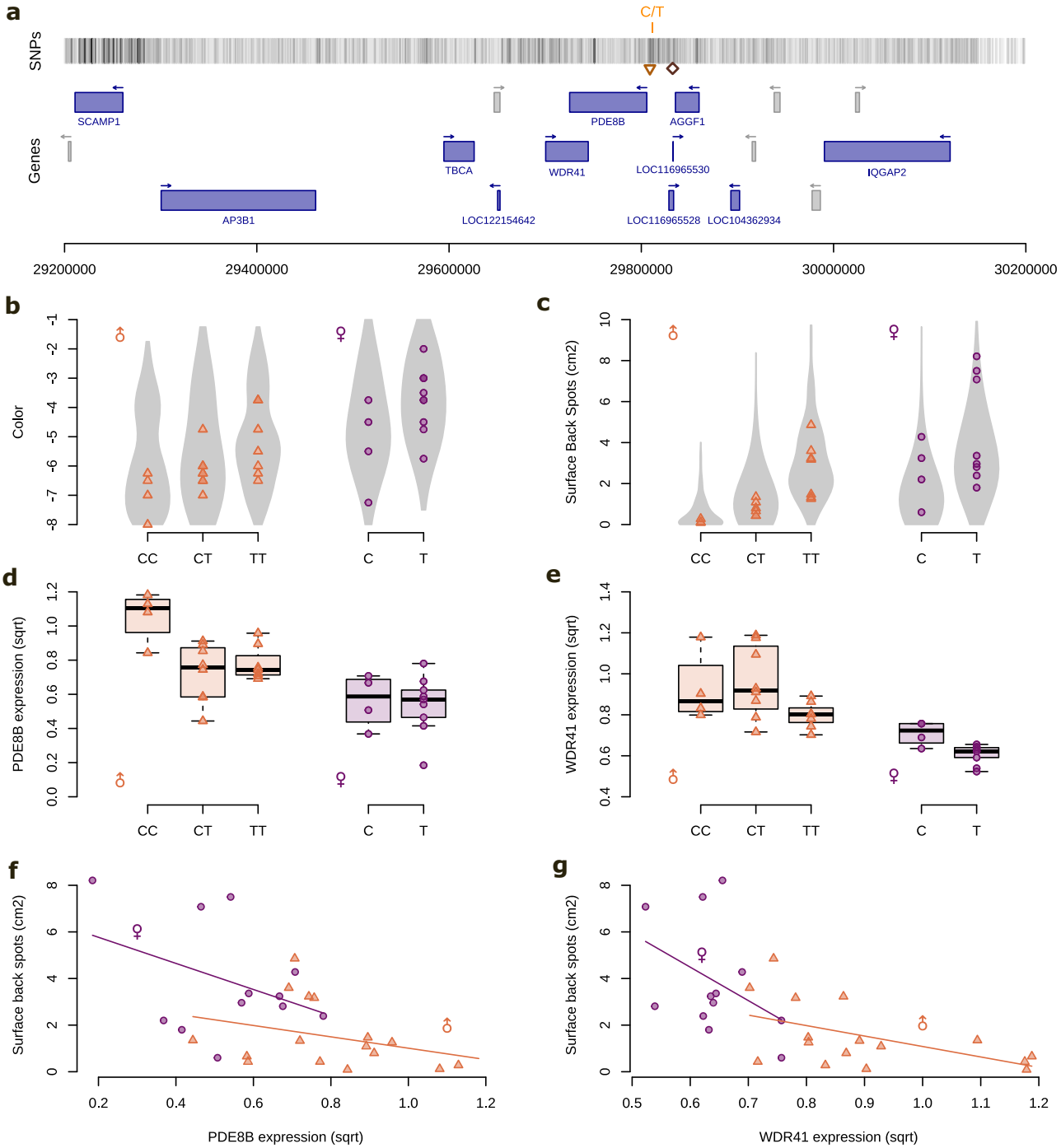


Fig. 3 | Detailed view of the Z-linked genomic region associated with both coloration and spottiness, and the associated genotype-expression-phenotype relationship in developing breast feathers from 45 nestlings (28 males (♂) and 17 females (♀)). a Genomic region on chromosome Z (Super Scaffold 42) containing annotated genes, including *PDE8B* and *WDR41*. The focal SNP (C/T) and the most significant SNP for color and spottiness are indicated. **b**, **c** Violin plots of plumage color and spottiness (black spot area, cm^2) across genotypes (CC, CT, TT in males; C, T in females) for 3102 sequenced individuals. Triangles (for males) and circles (for females) denote the 45 individuals with quantified gene expression. **d**, **e** Boxplots of sqrt-transformed expression of *PDE8B* and *WDR41* across genotypes. Box plots show the median (center line), the 25th and 75th percentiles (box bounds), and whiskers extending to the smallest and largest observations within 1.5 × the interquartile range (IQR) of the lower and upper quartiles, respectively. Points beyond the whiskers are plotted individually as outliers. Triangles (for males) and circles (for females) denote the 45 individuals with quantified gene expression. **f**, **g** Scatter plots showing correlations between sqrt-transformed expression and spottiness for *PDE8B* and *WDR41*; lines depict sex-specific linear models. Triangles (for males) and circles (for females) denote the 45 individuals with quantified gene expression.

discrepancy indicates that additional genetic or environmental factors likely contribute to the trait.

Taken together, our results reveal distinct genetic architectures for both coloration and spottiness between the sexes. The epistatic interaction between *MC1R* and the Z-linked locus in determining coloration in females, alongside the male-specific contribution of *MC1R* to spottiness, suggests that trait expression is influenced not only by individual loci but also by their interactions, interactions that may

https://doi.org/10.1038/s41467-026-78632-4

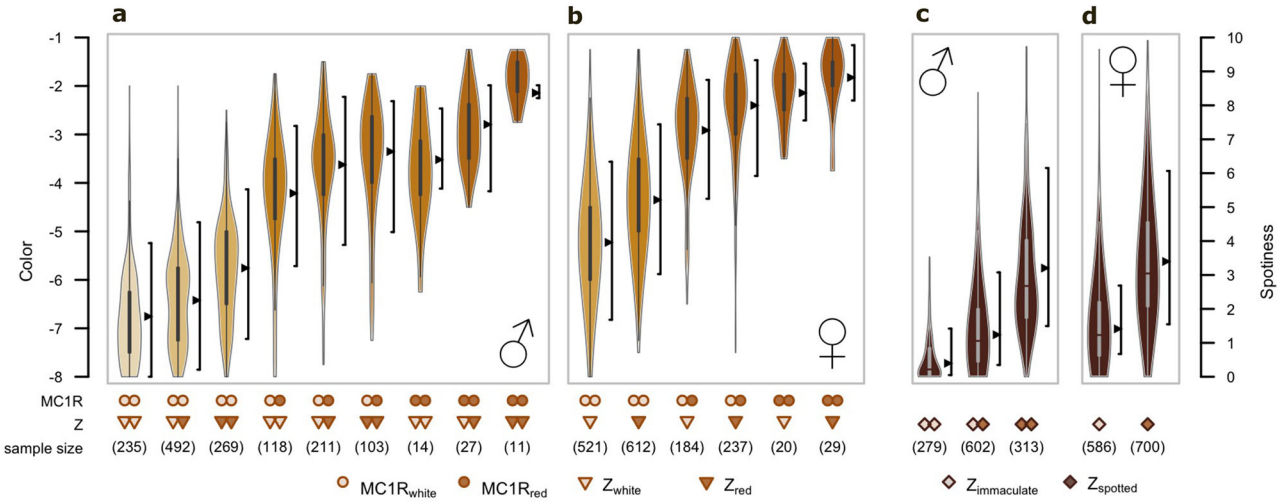


Fig. 4 | Sex-specific genetic architecture of plumage coloration and spottiness in 3087 Swiss barn owls (1484 males (♂) and 1603 females (♀)). Violin plots of observed plumage coloration across genotype combinations for males (**a**) and females (**b**). Violin colors indicate mean phenotype per genotype. Box plots in the violins show the 25th and 75th percentiles (box bounds), and lines extending to the smallest and largest observations within 1.5 × the interquartile range (IQR) of the lower and upper quartiles, respectively. Intervals to the right of each violin show predicted lower and upper whiskers from the genetic architecture model; black triangles indicate median predicted values. The x-axis shows genotype dosage for each locus; sample sizes are in parentheses. **c** Same as (**d**), for plumage spottiness.

differ between sexes with mechanisms, such as incomplete dosage compensation in birds[37–39] and sex-biased gene expression[50,51], both of which can reduce antagonistic sexual conflict[52]. These findings underscore the importance of modeling sex-specific genetic effects and inter-locus interactions when investigating the evolution of complex traits and the maintenance of sexual dimorphism in natural populations.

Given the large effects of these loci on phenotype and the observed changes in trait distribution over time, we next leveraged the temporal dimension of our dataset to assess whether shifts in allele frequencies could account for the observed phenotypic trends.

Characterizing the evolutionary processes driving allelic and phenotypic changes

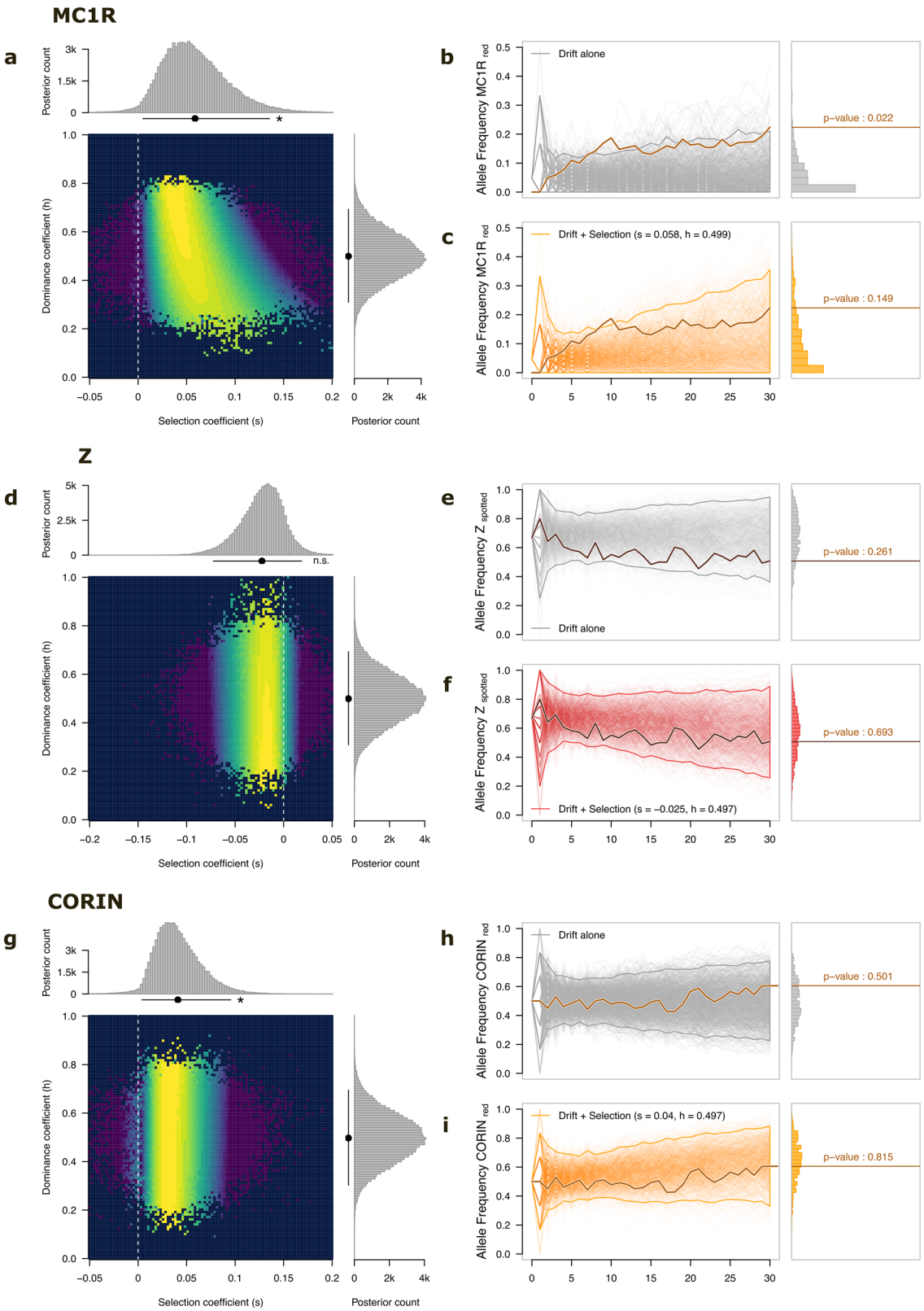
To investigate the evolutionary forces underlying allele frequency changes at the loci associated with color and spottiness, we leveraged the temporal resolution of our 30-year dataset to track allele dynamics over time (see Fig. 5 for *MC1R, CORIN,* and the Z locus associated with spottiness; and Fig. S25 for the Z locus associated with the color). Allele frequencies were estimated annually using genotypes from adult owls nesting in each year (Fig. 5b, e, h). Over the three-decade period, the frequency of the $MC1R_{red}$ allele increased significantly, while the $Z_{spotted}$ allele showed a significant decline (Fig. 5b, e, and Tables S21, S22). Given the strong associations of these loci with plumage coloration and spottiness, respectively, this temporal trend may reflect directional selection acting on these traits. However, change in allele frequencies over time could also be influenced by gene flow, demographic changes, or genetic drift. To disentangle these processes, further analysis is needed to assess the role of selection relative to neutral evolutionary forces.

As a first step in disentangling the processes driving allele frequency shifts, we tested whether migration could explain the observed changes. Individuals ringed as fledglings were classified as residents, while those first captured as adults, likely born outside the study area, were considered immigrants. Across years, allele frequencies did not differ significantly between these two groups (Fig. S25), suggesting that migration is unlikely to be the primary driver of the observed allele frequency changes. Instead, the trends likely reflect broader-scale evolutionary processes acting across the wider population. This interpretation aligns with previous findings of high connectivity and low genetic differentiation among barn owl populations across continental Europe[53].

To assess whether the observed allele frequency changes could be attributed to selection rather than genetic drift, we considered the relationship between allele effects on fitness and population size. For natural selection to act efficiently, the variant must influence fitness, and the effective population size (N_e) must be large enough for selection to overcome stochastic fluctuations due to drift. In small local populations, drift can dominate, causing allele frequencies to change regardless of selective advantage. Although long-term effective population size reflects historical connectivity at a broader geographic scale, census estimates for the locally breeding Swiss barn owl population remained relatively stable over the study period, ranging from 200 to 1000 breeding pairs[54]. To estimate N_e, we used whole-genome data to calculate nucleotide diversity (π), applying the equation $N_e = \pi/(4 \times \mu)$, where μ is the mutation rate[55]. We adopted the recently estimated mutation of the barn owl ($Mu = 5.6 \times 10^{-9}$)[56]. Across different subsampling, N_e estimates were highly consistent, averaging 77,907 individuals (range: 77,178 in 2016 to 78,825 in 1992; Fig. S26), roughly 200 times larger than the census population. This inflation likely reflects a large historical effective size and high contemporary connectivity among European barn owl populations (low genome-wide differentiation across the Western Palearctic (overall $F_{ST} = 0.047$))[53] and rare occurrence of inbreeding[57].

To test whether genetic drift alone could explain the observed allele frequency changes, we simulated allele frequency trajectories under a neutral model using the census population size as input ($N = 400$, see "Methods"). For the *MC1R* locus, the observed frequency trajectory of the $MC1R_{red}$ allele deviated significantly from the range expected under neutrality. Specifically, the increase in $MC1R_{red}$ frequency exceeded that predicted by drift alone (Fig. 5b), suggesting a likely role for positive selection. Inference of selection at MC1R was robust to population size assumptions, as simulations performed with larger population sizes ($N = 1000$ and $N = 10{,}000$) yielded qualitatively similar conclusions, with the observed trajectory remaining unlikely under drift alone (Fig. S27). Although the observed trajectory exhibited some year-to-year variability, potentially due to limited sample sizes in certain years, family structure within the dataset, or temporal fluctuations in selection intensity, the overall mismatch between simulated and observed patterns indicates that drift alone cannot

https://doi.org/10.1038/s41467-026-78632-4



account for the changes at this locus. In contrast, the frequency dynamics of the $Z_{spotted}$ allele fell within the distribution of trajectories expected under neutrality drift (Fig. 5e), suggesting that the decline in this allele could be explained by stochastic processes only. The same trend was observed for the Z_{red} allele, reflecting the high linkage between the two loci (Fig. S28). Finally, the allele associated with the redness at the CORIN locus also fell within the distribution of trajectories expected under neutrality drift (Fig. 5h).

To further investigate whether the observed allele frequency changes at the different loci reflect directional selection, we applied a probabilistic modeling framework that infers evolutionary parameters from temporal frequency shifts (see “Methods” for details).

Fig. 5 | Likelihood-based inference of selection parameters and simulation-based validation for *MC1R* (top panels), Z_{spot} (middle panels), and *CORIN* (bottom panels). a, **d**, **g** Heatmaps show the log-likelihood surface across a grid of selection coefficients (s) and dominance coefficients (h), with warmer colors indicating higher likelihood of the observed allele frequency trajectory for a given parameter combination. Overlaid bar plots represent the posterior distributions of s and h, summarizing parameter uncertainty; points below the histograms mark posterior means, and bars indicate 95% confidence intervals. Significance was assessed from confidence intervals: estimates with confidence intervals excluding zero were considered significant (*), whereas estimates with confidence intervals including zero were considered non-significant (n.s.). **b**, **c**, **e**, **f**, **h**, **i** Results from forward simulations of allele frequency change under two evolutionary scenarios: genetic drift alone (top) and drift combined with selection (bottom). Each line represents a single simulation replicate (grey for neutral drift, orange or brown for simulations including selection), and the solid line shows the 95% confidence interval across replicates. Red lines indicate the observed allele frequency trajectories over the study period for the *MC1R* red allele and the *CORIN* red allele, and the dark-brown line the allele frequency trajectories for the Z-spotted allele. Histograms on the right show the distribution of final allele frequencies across replicates. *P*-values were obtained from a two-sided empirical test using the null distribution shown in the figure. The empirical *p*-value corresponds to the proportion of simulated values at least as extreme as the observed value relative to the mean of the null distribution.

The *MC1R* locus exhibited a stronger signal of selection. Although only 34.09% of iterations in the combined model (which includes both selection and drift) supported selection, the selection-inclusive models consistently yielded higher log-likelihoods than the drift-only model (Fig. S29). Furthermore, both the combined and selection-only models produced selection coefficient estimates whose 95% credible intervals excluded zero (Figs. 5a and S30), indicating a statistically significant deviation from neutrality. These results provide compelling evidence that positive selection contributed to the observed increase in $MC1R_{red}$ allele frequency. The estimated selection coefficient of $s = 0.054$, with a dominance coefficient of $h = 0.492$, suggesting an additive to slightly dominant effect of the $MC1R_{red}$ allele. To validate our selection inference for *MC1R*, we incorporated the estimated genotype-specific selection coefficients into forward-time simulations. Under a model combining genetic drift and directional selection, the observed allele frequency trajectory for the $MC1R_{red}$ allele fell within the simulated range (Fig. 5c). Importantly, inference of selection at *MC1R* was robust to uncertainty in the initial allele frequency, with sensitivity analyses across a range of plausible starting values and population sizes yielding consistent results (Fig. S27). These results support the role of selection in shaping its increase over time. In contrast, the analysis of allele frequency changes at the Z-linked loci revealed limited evidence for selection. In the model allowing both drift and selection, 96.5% of iterations favored the drift-only scenario for the Z-locus associated with the spottiness (Fig. S31). The log-likelihood difference between the selection and drift models was modest (Fig. S31), and the 95% credible interval for the selection coefficient in the selection-only model included zero (Figs. 5d and S32), indicating no statistically significant deviation from neutrality. Moreover, incorporating the estimated selection coefficient into simulations only marginally improved the fit to the observed frequency trajectory (Fig. 5f). Consistently with the high linkage disequilibrium between the two loci, we obtained similar results with no statistically significant deviation from neutrality for the Z-locus associated with the redness (Figs. S33 and S34).

For the *CORIN* locus, probabilistic modeling inferred a selection coefficient significantly different from zero ($s = 0.04$, 95% credible interval excluding zero, Fig. 5g). However, support for selection at this locus was limited, with only 14.43% of iterations favoring selection in the combined drift-and-selection model (Fig. S35), and incorporating the estimated selection and dominance coefficients into forward simulations did not improve the fit to the observed allele-frequency trajectory (Fig. 5i).

Taken together, these results suggest that the allele frequency dynamics at the *CORIN* locus remain inconclusive and warrant further investigation, that the *MC1R* locus are best explained by a combination of drift and positive selection, and that allele-frequency dynamics at the Z-linked loci are consistent with neutral evolution.

Concluding remarks

Three decades of phenotypic monitoring combined with whole-genome sequencing of over 3000 individuals provide a window into the genomic basis of rapid trait evolution in a wild vertebrate population. We identify an oligogenic architecture underlying two melanin-based traits, coloration and spottiness, both of which have undergone significant shifts over the 30 years monitored. These traits share a partially overlapping genetic basis but exhibit pronounced sex-specific differences in genetic architecture and gene interactions. Allele frequency changes at major loci during this time window are consistent with selection, although the specific fitness consequences and potential trade-offs with other traits remain to be explored. Crucially, this study illustrates the power of integrating genomic, phenotypic, and long-term ecological data to uncover the mechanisms of rapid evolutionary change in natural populations.

Methods

Study system and phenotyping

All animal handling procedures were conducted in accordance with international guidelines for animal welfare and the Swiss Animal Welfare Act. The study protocols were reviewed and approved by the cantonal veterinary authorities of Vaud and Fribourg, Switzerland. Animal experimentation permits were issued by the Veterinary Offices of the Cantons of Vaud and Fribourg under permit numbers 956 (1996), 1146 (1998), 1736 (2004), 1508 (2007), 2177 (2010), 2109 (2012), 2844 (2015), 2934 (2016), 3462 (2020), 3571 (2020 and renewed in 2023), 3213 (2021), and 3884 (2025). Authorizations for work involving protected wildlife, including owl ringing, were issued by the Swiss Federal Office for the Environment (FOEN).

Study system. A wild population of barn owls (*Tyto alba*) in western Switzerland (46°49′ N, 06°56′ E) has been continuously monitored since 1986, spanning over 30 years. These owls typically reach sexual maturity within their first year of life[58], and primarily breed in nest boxes installed on farms. Throughout each breeding season, nest boxes were inspected at 4-week intervals to assess occupancy. When an active nest was detected, females were captured during incubation, while males were typically caught later during the provisioning of nestlings. During these visits, all nestlings and any previously unmarked adults were fitted with uniquely numbered metal rings for individual identification. Blood samples were collected from the brachial vein of each bird and initially stored in liquid nitrogen before being transferred to −80 °C freezers for long-term preservation. Nestling sex was determined using sex-specific molecular markers targeting the SPINDLIN-gene[59]. Breeding females were identified by the presence of a brood patch if its sex was not determined with molecular markers as nestling. The year of birth was known for individuals ringed as nestlings. For adults ringed for the first time, age and year of birth were estimated based on molt patterns of wing flight feathers[60].

Phenotype measurements. At each capture event, two melanin-based plumage traits, namely coloration and spottiness, were assessed on both the breast and the belly of each individual. The degree of reddish pheomelanic coloration was scored using an eight-color chip scale ranging from −8 (white) to −1 (dark reddish). See Fig. S1 for example photographs of barn owls being score using the chip scale. This method has been shown to be highly repeatable, with repeatability estimates averaging 0.90 when individuals are scored twice within the same season[61]. Visual color scores have also been validated against spectrophotometric measurements of pheomelanin-based coloration. In a large subset of individuals, visual scores were strongly correlated with brown chroma (R600-700/R300-700) measured using reflectance spectrometry (Pearson's $r = -0.78$, $P < 0.0001$, $N = 1107$)[44], demonstrating that this scoring system reliably captures biologically meaningful variation in plumage coloration. Spottiness was quantified by counting the number (N) and measuring the mean diameter ($D \pm 0.1$ mm) of eumelanic black spots within a 60×40 mm (2400 mm^2) frame (Fig. S1). For each body part (breast and belly), we calculated the total surface area covered by black spots (referred to as spottiness, S) using the following formula (expressed in cm^2): $S = (N \times (\pi \times (D/2)^2))/100$. The spottiness values of the breast and belly were then summed to obtain a single value representing the total ventral spottiness. To retain a single phenotype value per individual, we averaged trait values across all available adult captures. If no data were available (e.g., due to mortality or dispersal), the fledging measurement was retained. The life stage at which an individual was phenotyped (fledging or adult) was recorded for downstream analyses. Individuals first ringed as adults were classified as "immigrants," whereas those ringed as fledglings were classified as "residents." To examine the relationship between the two traits, we estimated Pearson's correlation coefficient (r) using the *cor()* function R (v4.2.2)[62] across all individuals, and separately for each sex. Differences in trait values between sexes and between life stages (juvenile vs. adult) were tested using the *wilcox.test()* function in R.

Evolution of two melanin-based traits

Modeling the temporal evolution of coloration and spottiness. To assess whether the melanin-based traits of coloration and spottiness have changed over time, we modeled the evolution of each trait across years using a linear mixed-effects model (LMM). Models were fitted via restricted maximum likelihood (REML) using the *lmer()* function from the lme4 package (v1.1-37)[63] in R (v4.2.2). For each trait, we constructed a series of increasingly complex models by sequentially adding explanatory variables. Model comparisons were based on Akaike's Information Criterion (AIC; Supplementary Data 1 and 2). The best supported model for both traits included the following fixed effects: year of birth (continuous) to test for long-term evolutionary trends; sex (categorical) to account for sexual dimorphism in plumage traits[61,64], Fig. S7); and phenotypic stage (categorial: fledging or adult) to control for age-related changes in melanin expression and for age-related variation in coloration known to change slightly during early life but remains largely stable after the first year[44,45], Fig. S8). To account for variation related to observation timing, we included year of first capture as a random effect in all models. We also tested the effect of origin (resident versus immigrant) by including a categorical fixed effect indicating whether an individual was born within the study population. However, adding this variable did not improve model performance and was therefore excluded from the final model.

Whole genome sequencing and genotyping

Reference panel: data collection, phasing, and quality control. The reference panel used in this study was previously assembled[65], and includes samples from previous whole-genome sequencing efforts[18,53,65–68]. In total, whole genome sequencing data from 502 samples were processed through a standardized variant discovery pipeline, summarized below (see[65] for further details). Raw sequencing reads were processed with Trimmomatic v0.39[69]: sequencing adapters were removed, and reads shorter than 70 bp were discarded. Cleaned reads were aligned to the *Tyto alba* reference genome (NCBI RefSeq assembly: GCF_018691265.1)[67] using BWA-MEM v0.7.17[70].

Variant discovery followed the GATK v4.2.6 best practices pipeline[71]. Base Quality Score Recalibration (BQSR) was performed using a previously validated truth set[53]. SNPs were initially called per individual using HaplotypeCaller, then merged via joint genotyping with *GenotypeGVCFs*. Variants were filtered to retain only high-quality bi-allelic SNPs that passed the following GATK hard filters: QD < 2.0, QUAL < 30, SOR > 3.0, FS > 60.0, MQ < 40.0, MQRankSum < −12.5, and ReadPosRankSum < −8.0). In addition, regions with low mappability were excluded[72]. Depth-based filtering was performed using BCFTools (v.1.15.1)[73]: any genotype with read depth <5 or > (mean depth + 3 SD) was set to missing. SNPs with a minor allele count (MAC) < 5 were also removed. Following all filtering steps, the final reference panel comprised 28,197,066 high-confidence SNPs. To facilitate downstream analyses, SNPs positions originally mapped to Super-scaffolds were converted to chromosomal coordinates based on the linkage map of the barn owl[65]. Of the retained SNPs, 26,933,469 were located on autosomes, and 1,263,597 were assigned to the Z chromosome (corresponding to Super-scaffold 13 and 42 in the reference assembly).

The full set of 28,197,066 high-confidence variants was phased in two successive steps. In the first step, local phasing was performed using WhatsHap v1.4[74], which uses sequencing reads information to phase variants observed together on the same read or read pair, and incorporates pedigree data when available. Phasing was conducted independently for each individual. For individuals belonging to parent-offspring trios, pedigree-aware phasing was applied using WhatsHap v1.4[74] pedigree mode. In cases where a parent was present in multiple trios (i.e., had more than one offspring in the dataset), the parent was phased together with the offspring that had the highest sequencing depth. Additionally, we incorporated the mean recombination rate of the species[65], by using the *--recombrate 2* option during phasing. After phasing, we applied a second round of variant filtering to minimize bias introduced by relatedness among individuals. Specifically, we retained only biallelic SNPs with minor allele count (MAC) > 5, and missing data <5%, based on a subset of 187 unrelated individuals, defined as having pairwise kinship coefficients <0.03125 (see[65] for kinship estimation details). Following this filtering step, the final dataset comprised 10,451,268 SNPs, of which 10,115,035 were located on autosomes and 336,233 on the Z chromosome.

The set of 10,115,035 autosomal variants was further phased using SHAPEIT v4.1.2[75]. This algorithm extends local phasing by incorporating population-level information using a model based on coalescence and recombination, enabling both statistical phasing of haplotypes and imputation of missing genotypes. To minimize potential bias due to relatedness, phasing was performed in two stages: 187 unrelated individuals were phased together to form a reference haplotype panel. 315 family members were then phased individually against this unrelated set, avoiding the inclusion of family structure in the phasing panel and maintaining statistical independence. SHAPEIT was run according to the authors' recommendations for increased phasing accuracy. Specifically, the number of conditioning haplotypes in the PBWT (Positioning Burrows–Wheeler Transform) was set to 8. The Markov Chain Monte Carlo (MCMC) parameters were set as follows: 10 burn-in iterations, 5 pruning iterations, each separated by 1 additional burn-in iteration, and 10 main iterations for haplotype sampling.

The set of 336,233 SNPs located on the Z chromosome was phased independently to account for the haplo-diploid nature of this chromosome in birds (males: ZZ, females: ZW). For variants located within the pseudo-autosomal region (PAR)—as defined in ref. 66, the phasing procedure followed the same protocol as for autosomal variants, since all individuals are diploid in this region. For non-PAR regions, phasing

was handled differently by sex: diploid males (ZZ) were phased using the same two-step procedure applied to autosomes: first using WhatsHap v1.4[74] for read-based phasing, followed by SHAPEIT v4.1.2[75] for statistical phasing. Haploid females (ZW) possess a single Z chromosome; thus, any heterozygous calls (e.g., due to sequencing or genotyping error) were recoded as haploid genotypes by assigning the most frequent allele observed across individuals at that site. After sex-specific phasing, the male and female datasets were merged, and female genotypes were recoded as fully homozygous diploid for downstream analyses, ensuring compatibility with software that requires diploid input. This yielded a final dataset comprising 10,451,268 phased SNPs, including both autosomal and sex-linked variants.

To evaluate phasing quality on autosomes, we used the switch error-rate (SER) metric[76]. A switch error occurs when a heterozygous site is phased differently between two phase sets for the same individual. For each individual, we compared local phasing inferred either form a read-based or the trio approach (both using WhatsHap v1.4[74] as described previously) with statistical phasing produced by SHAPEIT v4.1.2[75]. In this comparison, SHAPEIT phasing was performed without using read-based or trio information for the target individual, using the same version and parameters detailed in the previous section. This method assumes that short phasing blocks generated by WhatsHap are accurate; thus, any switch in these blocks when compared to SHAPEIT v4.1.2 phasing is considered an error. Switch error rates were calculated using the switchError tool (available at https://github.com/SPG-group/switchError). Our results show that variants present in the reference panel had a low switch error rate, with a mean error SER of 1.83% (Fig. S13).

We also evaluated the ability of the reference panel to capture both SNP diversity and the arrangement of these SNPs into haplotypes within the population. To assess SNP diversity, we downsampled the number of individuals and estimated the number of unique SNPs present in each subset. We observed that the number of SNPs plateaued after approximately 200 individuals, indicating that the reference panel effectively captures the majority of common SNPs found in the Western Palearctic population (Fig. S11). To assess haplotypic diversity, we applied a similar downsampling approach and found that the panel also captures fine-scale haplotype structure (Fig. S14).

Low coverage: DNA extraction, library preparation, and sequencing. We sequenced 2778 owls from a pedigreed population, with samples collected between 1986 and 2020. Individuals were prioritized based on the availability of complete family information and phenotypic data. To assess sequencing performance, we included 32 individuals previously sequenced at high coverage as part of the reference panel (see above). Of these, 10 individuals were sequenced three times at low coverage, once per flow cell, to assess potential batch effects (see below for details). These are referred to as *triplicates*. The remaining 22 individuals, each sequenced once at low coverage, are referred to as *duplicates* and were used to evaluate the accuracy of low-coverage genotyping by comparison with their high-coverage data (see next paragraphs).

Genomic DNA was extracted from blood samples using the DNeasy Tissue Kit (Qiagen, Switzerland) and the Biosprint robot 96 (Qiagen, Switzerland), then stored at −20 °C. We randomized the position of the 2820 DNA samples (2778 owls + 20 triplicates + 22 duplicates) across 30 distinct 96-well plates. Each plate included two empty wells for contamination control. High quality DNA (without degradation; randomly check with Fragment Analyzer HS Large Fragment 50 kb Kit, Agilent, USA) was quantified using the Quant-it PicoGreen dsDNA Assay kit (Thermo Scientific, Switzerland), and diluted in 10 mM Tris-HCl to a final concentration of 1.5 to 2.5 ng/µl. Libraries were prepared using the seqWell−plexWell LP384 and seqWell−plexWell WGS 24 kits (SeqWell, USA) and sequenced across three lanes (10 plates per lane) on the Illumina NovaSeq 6000 platform at the Genomic Technologies Facility (GTF) of the University of Lausanne.

Low coverage: genotyping, phasing, imputation, and validation on autosomes. Raw sequencing reads were trimmed using Trimmomatic v.0.36[69] and aligned to the barn owl reference genome (NCBI RefSeq assembly: GCF_018691265.1)[67] with BWA-MEM v.0.7.15[70]. Read alignment was followed by quality control using Qualimap v2.2.1[77], which was used to estimate mean sequencing coverage across the autosomes and the Z chromosome. Across all individuals, the average genome-wide was 1.95X, with individual coverage ranging from 0.2X to 4.15X (Fig. S14).

We estimated genotype likelihoods for low-coverage individuals at each variant position present in the reference panel using BCFTools v1.19 (*mpileup* and *call* methods), with the *-T* and *-C* options enabled[78]. Phasing and imputation of these low-coverage samples were performed using GLIMPSE v1.1.1[79]. GLIMPSE leverages haplotypes from high-coverage reference panel to impute and phase genotypes in low-coverage samples. We used the 502 high-coverage individuals described earlier as the reference panel. We selected version 1.1.1 of GLIMPSE, as it is better suited for small reference panels according to the official documentation. The GLIMPSE pipeline consists of four steps: chromosomes were divided into overlapping chunks for efficient parallel processing using the GLIMPSE_chunk tool. We applied default parameters (*--window-size 2,000,000* and *--buffer-size 200,000*). Each genomic chunk was phased and imputed using the *GLIMPSE_phase* method. This step iteratively refines genotypes likelihoods and phase information for each individual independently. In each iteration, GLIMPSE phases the low-coverage individual using haplotypes from both the high-coverage reference panel and other imputed low-coverage samples. It identifies the closest haplotypes, those sharing the highest number of identical-by-descent (IBD) segments, and uses them for genotype imputation. Within each diploid sample, haplotypes are imputed separately. A new iteration begins once imputation is complete. We performed this step with an increased number of iterations (*--burnin 100* and *--main 15*) and set the effective population size parameter to (*--ne 10000*). If a recombination map was available for the given chunk[65], it was included using the *--map* option. All samples were considered diploid for autosomal genotyping. Once phasing and imputation were completed for each chunk, results were merged using the *GLIMPSE_ligate* tool to ensure consistent phase across the genome. This step was run with default parameters. The final step involved identifying the most likely haplotype configurations based on posterior genotype likelihoods and phase probabilities, using the GLIMSE_sample method with the *--solve* flag enabled.

Validation of the imputation approach was performed using the 32 individuals that were sequenced at both high and low coverage. As previously mentioned, GLIMPSE selects haplotypes sharing the highest fraction of identical-by-descent (IBD) segments with the target sample for both phasing and imputation. Including the same individual in both the reference panel and the imputed (target) set would therefore lead to artificially elevated accuracy due to circularity. To avoid this bias, we ran the GLIMPSE pipeline five times: we used the full reference panel of 502 individuals and the 2768 unique low-coverage samples (excluding all duplicates and triplicates). This run produced the *main* dataset used throughout the study. The 22 duplicate individuals (those sequenced at both high and low coverage) were removed from the reference panel and added to the low-coverage target set. This resulted in a reference panel of 480 individuals and 2790 low-coverage samples. This setup was used specifically to validate imputation accuracy without self-reference bias. Since GLIMPSE also samples haplotypes from other low-coverage individuals, we further ensured independence by imputing each of the three low-coverage replicates of the 10 *triplicate* individuals separately. In each run (denoted a, b, and c), the

triplicate being imputed was excluded from the rest of the low-coverage data. These runs used as reference panel of 492 individuals (502 minus the 10 triplicates) and 2778 low-coverage samples. To ensure fair evaluation, we excluded 11 individuals from the validation set that had intermediate sequencing depth in the high-coverage dataset (mean ~11x), as their coverage could bias the accuracy assessment (see[33] for details). Using the remaining $n = 21$ individuals sequenced at high coverage (mean ~30x), we estimated the overall imputation accuracy as $r^2 = 0.968$ across all variants. For further methodological details on imputation and validation procedures, refer to ref. 34.

Identification of genomic regions associated with plumage coloration and spottiness

GRM and kinship matrix. We estimated individual-based relatedness (β)[80] and inbreeding coefficients using the R package hierfstat (v0.5-11, R v4.2.2)[81]. These metrics were calculated for all individuals in the dataset. The resulting kinship matrix was then converted into a Genetic Relationship Matrix (GRM) using the *kinship2grm()* function provided by the hierfstat package. This GRM was used in downstream association analyses to account for population structure and relatedness among individuals.

SNPs-based heritability of plumage coloration and spottiness. We estimated SNP-based heritability for the two traits, plumage coloration and spottiness, by fitting two independent animal models. These general linear mixed models partition the total phenotypic variance (V_P) into three components: (1) variance explained by fixed effects, accounting for known environmental sources; (2) additive genetic variance (V_A), estimated via the GRM; and (3) residual variance (V_R). Animal models were implemented in R using the brms package (version 2.34, R version 4.2.2)[49]. Sex and phenotypic stage were included as fixed effects to control for their influence on trait variation. The GRM, previously described, was incorporated as a random effect to estimate V_A. Models were run using 8 independent Markov Chain Monte Carlo (MCMC) chains, each with 10,000 iterations and burn-in of 2000 iterations. Narrow-sense heritability (h^2) was calculated as $h^2 = V_A/V_P$. We also report the 95% credible intervals associated with each h^2 estimate. Results for both traits are presented in Supplementary Tables S9 and S10.

Genome-wide association study (GWAS). To identify genomic variants associated with plumage coloration and spottiness in the Swiss barn owl population, we performed genome-wide association analyses using the *association.test()* function from the *gaston* package (v1.6, R v4.2.2)[82]. A linear mixed model (LMM) was fitted using the Average Information Restricted Maximum Likelihood (AI-REML) algorithm, as implemented in the same package[82]. Both traits were treated as continuous variables. Sex and phenotyping stage (juvenile or adult) were included as fixed covariates, while the Genetic Relationship Matrix (GRM) was included as a random effect to account for population structure, familial relationships, and cryptic relatedness. To avoid proximal contamination and increase statistical power, we applied a Leave-One-Chromosome-Out (LOCO) approach. For each chromosome, we constructed a GRM excluding SNPs from that chromosome and performed the GWAS on the focal chromosome using this modified GRM. This approach ensures that associations are not inflated due to shared signal between the tested SNPs and the GRM[83]. Mixed-model association methods incorporating LOCO have been shown to outperform traditional GWAS approaches in terms of power and control for confounding[84].

For every GWAS, we used the score statistics to assess the strength of the association between individual SNP and the phenotype of interest. The significance threshold ($\alpha = 0.05$) was adjusted using a Bonferroni correction to account for multiple testing[85]. Additionally, we visually assessed deviations from the null distribution using quantile-quantile (QQ) plots, focusing on SNPs deviating from the expected 1:1 line as a qualitative check for inflation or strong associations.

To maximize power for detecting loci associated with plumage coloration and spottiness, we conducted six independent GWAS, three for each trait category.

For the coloration, we first ran a GWAS using the mean ventral plumage coloration as the phenotype. However, since the coloration of the belly and breast are not perfectly correlated (Fig. S4), we conducted two additional GWAS, using belly coloration and breast coloration as independent traits. This allowed us to identify loci potentially contributing to regional variation in coloration patterns.

For the spottiness, we began with a GWAS on total spottiness as a composite trait. Given that this phenotype results from the combination of both spot number and spot diameter, we stratified the dataset to explore the genetic basis of each component separately. This involved running two additional GWAS: (i) one using the number of dots, and (ii) one using the mean spot diameter as the phenotype.

Genotyping and gene expression of the Z chromosome region

Feather and blood sampling and genotyping. For gene expression and allelic discrimination, we collected two to three developing breast feathers and blood samples from 28 male and 17 female nestlings born in 2023 ($n = 21$) and 2024 ($n = 24$). We collected growing feathers from nestlings at a similar stage of development (mean age ± SD: 29 days ± 1.8). At this age, the feathers start to develop the typical black spots on the apical part of the feather and the white to reddish pheomelanin coloration. Upon collection, feathers and blood were immediately frozen in liquid nitrogen in the field and stored at −80 °C until molecular analyses. For every individual, genomic DNA was extracted using the Blood and Tissue kit (Qiagen, Switzerland), sexed (as described in ref. 60), and genotyped for *MC1R* (as described in ref. 46). Out of the 45 individuals used for the gene expression analysis, The *MC1R* genotypes are the following: 38 *VV* (GG genotype, homozygous $MC1R_{white}$), 6 *VI* (AG genotype), 1 male *II* (AA genotype, homozygous $MC1R_{red}$). Because of this unequal frequencies, *MC1R* genotypes were not considered for the analysis.

Allelic discrimination assay. To validate associations found on the Z chromosome, we performed an allelic discrimination assay targeting a SNP located on Super-Scaffold 42. The two top SNPs identified in the GWAS for plumage coloration and spottiness could not be directly genotyped due to probe design constraints. We therefore selected a nearby SNP (at position 29,811,381) for which probe design was feasible. This SNP showed strong linkage disequilibrium with the top associated SNPs for both plumage coloration ($r^2 = 0.97$) and spottiness ($r^2 = 0.98$) and was itself significantly associated with both traits ($p < 6.95e\text{-}31$ for coloration; $p < 2.70e\text{-}122$ for spottiness). The following primers and probes were used in the assay: forward primer ChZ29811429Fw: 5′-TATCTTTGGGCTTGACTGGT-3′; reverse primer: ChZ29811308Rev: 5′-AAACACCCAAGAAATAGCAAT-3′; Probe 1 (FAM-labeled): ChZ29811381FamMGBQ530: 5′-ACTCTCACTTTG**T**TGCT CTCTCCT-3′; Label: 5′ FAM fluorophore; 3′ MGB quencher; Probe 2 (HEX-labeled): ChZ29811381HexMGBQ530: 5′-ACTCTCACTTTG**C**TGCT CTCTCCT-3′ Label: 5′ HEX fluorophore; 3′ MGB quencher. All oligonucleotides were synthesized by Microsynth (Balgach, Switzerland).

To enhance signal quality, an initial PCR pre-amplification step was performed using the following thermal cycling conditions: 95 °C for 5 min (initial denaturation) followed by 30 cycles at 95 °C for 30 s, 56 °C for 30 s, 72 °C for 30 s, followed with 72 °C for 10 min (final extension). The reaction mix (10 µL total volume) contained: 200 µM dNTPs, 1 mM $MgCl_2$, 500 nM of each primer, 1× GoTaq buffer, 0.5 U/µl GoTaq polymerase (Promega, Switzerland). Following PCR, the products were diluted 1:1,000,000, and 2 µL if the diluted product was

added to a 20 µL qPCR reaction using Takyon MasterMix (1×) (Eurogentec, Belgium), with 300 nM of each primer and 300 nM of each probe.

qPCR and allelic discrimination were performed on a CFX96 Real-Time PCR Detection System (BIO-RAD, Switzerland) using the standard protocol. The assay was conducted on 184 owls previously sequenced at low coverage. We observed a very high concordance between the genotypes inferred from allelic discrimination and those obtained via imputation from low-coverage whole-genome sequencing, with $r^2 = 0.986$. Only one individual showed a discrepancy; further inspection revealed that this sample had no sequencing coverage at the locus, suggesting the discordant imputed genotype was due to a miscall during imputation rather than an error in the allelic assay.

The allelic discrimination for the region on the Z chromosome (SNP 29,811,381) was then used to genotype the 45 individuals used for RNA extraction and gene expression. The genotypes at the Z locus were for males 5 *CC*, 13 *CT*, 10 *TT*, and for females 5 *C* and 12*T*.

Total RNA extraction and cDNA preparation. Total RNA was extracted from the developing breast feathers of the 45 nestlings used in the gene expression study. For each individual, one to two growing feathers were used. At the time of RNA extraction, a photograph of the basal part of each feather was taken on dry ice to visually document the stage of black spot development. Feathers were ground in liquid nitrogen using a sterile pestle and then resuspended in Qiazol Lysis Reagent. Total RNA was extracted using the miRNeasy Mini Kit (Qiagen, Switzerland) following the manufacturer's protocol, with the following modifications: RNase-free DNase treatment (Qiagen) was included during the extraction to remove genomic DNA; Buffer RWT was prepared using isopropanol instead of ethanol, as recommended for feather samples. Total RNA was eluted in 30 µL of RNase-free H_2O. RNA concentration was quantified using the Qubit fluorometer (Life Technologies, Switzerland) with the RNA Broad Range Assay Kit, and RNA integrity was assessed using the Fragment Analyzer (Agilent, USA). Only samples with high RNA quality (RQN > 7.7) were retained for downstream gene expression analyses. For each individual, 200 ng of total RNA was reverse-transcribed in a total volume of 20 µL using the LunaScript RT SuperMix (New England Biolabs; BioConcept, Switzerland), under the following thermal conditions: 25 °C for 2 min; 60 °C for 10 min; 95 °C for 1 min (inactivation). The resulting cDNA was precipitated using 1 µL of glycogen (20 ng/µL), 5 M ammonium acetate, and ethanol, and resuspended in 20 µL of 10 mM Tris-HCl (pH 8.0) and 0.1 mM EDTA. Because of the low expression levels of some target genes, 5 µL of the precipitated cDNA were preamplified with the complete primer mix used for the qPCRs. Preamplification was performed for 14 cycles using reagents from Life Technologies (ThermoFisher Scientific, Switzerland). The resulting product was diluted 10-fold in 10 mM Tris-HCl (pH 8.0) and 0.1 mM EDTA, and stored at −20 °C until qPCR. Preamplification was validated not to bias relative expression levels for all genes, except one reference gene (HPRT1) that showed altered expression post-preamplification and was therefore excluded from all downstream analyses.

Gene expression analysis. To investigate the molecular basis of plumage coloration and spottiness, we quantified the expression of genes located within and adjacent to the Z chromosome region (Super Scaffold 42) previously identified as being strongly associated with both traits. Specifically, we measured the expression of 12 genes spanning the genomic interval between 29,210,880 and 30,121,259 using reverse transcription quantitative PCR (RT-qPCR). The targeted genes include: *AGGF1* (Angiogenic Factor With G-Patch And FHA Domains 1); *AP3B1* (Adaptor Related Protein Complex 3 Subunit Beta 1; *CRHBP (LOC10436293)* (Corticotropin Releasing Hormone Binding Protein); *F2R* (Coagulation Factor II Thrombin Receptor); *IQGAP2* (IQ Motif Containing GTPase Activating Protein 2); *LOC116965528* (uncharacterized non-coding RNA); *LOC122154642* (uncharacterized non-coding RNA; *PDE8B* (Phosphodiesterase 8B); *SCAMP1* (Secretory Carrier Membrane Protein 1); *SNORA47* (Small Nucleolar RNA, H/ACA Box 47); *TBCA* (Tubulin Folding Cofactor A); *WDR41* (WD Repeat Domain 41). Among these, *PDE8B* and *LOC116965528* are located closest to the SNPs showing the highest associations with spottiness and coloration. Given that multiple splice variants have been annotated for many of these genes in the T. alba_DEE_v4.0, barn owl genome assembly (NCBI RefSeq assembly: GCF_018691265.1)[67], we designed and tested multiple primer and probe sets per gene to evaluate potential transcript-specific expression. All primers and probes used in this study are listed in Supplementary Data 12. All oligonucleotides were synthesized by Microsynth (Balgach, Switzerland). Four additional genes located in the same genomic region could not be amplified despite multiple attempts and were therefore excluded from the expression analyses. To normalize gene expression in the RT-qPCR assays, we evaluated the expression of four commonly used reference genes: *EEF1A* (*Elongation factor 1A); GAPDH* (Glyceraldehyde-3-Phosphate Dehydrogenase); *HPRT1* (Hypoxanthine Phosphoribosyltransferase 1); *TBP* (TATA-Box Binding Protein). However, due to altered expression following preamplification, *HPRT1* was excluded from all analyses. In addition to the genes located on the Z chromosome, we also quantified the expression of four candidate pigmentation genes: *MC1R* (Melanocortin 1 Receptor); *GPR143* (G Protein-coupled receptor 143); *TYR* (Tyrosinase); *TYRP1* (Tyrosinase-related protein 1). Primers and probes for these pigmentation genes were previously described[36,86].

Quantitative PCR (qPCR) reactions were performed on a CFX96 Real-Time PCR Detection System Bio-Rad (BIoRad, Switzerland). Each reaction contained 2 µl of diluted, pre-amplified cDNAs, run in technical duplicates, using the QuantiTect Probe PCR Kit (Qiagen, Switzerland) in a final reaction volume of 20 µL. To ensure technical accuracy, reactions were repeated if the Ct difference between duplicates exceeded 0.25 Ct. To control plate-to-plate variability and potential pre-amplification biases, three pooled pre-amplified cDNA controls (from different individuals) were included on every qPCR plate. QPCR conditions were optimized for each primer and probe set by testing various concentrations using serial dilutions of plasmids or PCR-purified products amplicons of the target gene regions. The goal was to achieve PCR efficiencies between 95% and 105%, in line with MIQE guidelines. Most primer-probe combinations fell within this range (see Supplementary Data 12). One exception was the IQGAP2_T102 primer-probe set, which showed a slightly elevated efficiency of 107%, but was retained for downstream analyses given the acceptable range for relative quantification (see below).

To identify the most stable reference genes for qPCR normalization, we used the RefSeeker package (v1.0.4, R v4.2.2)[87]. RefSeeker evaluates reference gene stability using four established algorithms-Normfinder, geNorm, delta-Ct, and BestKeeper-based on the raw quantification cycle (Cq) values. Among the three tested reference genes (*EEF1A*, *GAPDH*, and *TBP*), *EEF1A* and *TBP* were identified as the most stable, with final stability rankings of 1.2 and 1.4, respectively (Supplementary Data 12). These two genes were therefore used for normalization in subsequent expression analyses. All raw Ct values were log transformed to calculate relative expression assuming a PCR efficiency of 2.0 (corresponding to 100% efficiency, within the acceptable 95–105% range). For IQGAP2_T102, which had a slightly higher empirically determined efficiency, a coefficient of 2.07 was applied. Normalized relative expression values were obtained using the SL1PCR package (v1.72.0 R v4.2.2)[88], with EEF1A and TBP serving as internal controls.

Gene expression analysis and statistical modeling. All qPCR-derived gene expression data were square root-transformed to improve normality of the residuals prior to statistical analyses, using R v4.4.2.

Because multiple transcript variants were detected for several genes, we first assessed the correlation between their expression levels using males and females (see Figs. S21 and S22). Given the high correlation observed, only one transcript per gene was retained for downstream analysis. Where possible, we selected primers and probes that either (i) targeted all transcript variants, or (ii) captured the most highly expressed variant (e.g., for *IQGAP2*, see Supplementary Data 12). To test whether genotype at the Z-linked locus (SNP at position 29,811,381 on Super-Scaffold 42) influenced gene expression, we fitted linear mixed-effects model using the *lme()* function from the *nlme* package (v3.1-168, R v4.2.2)[89].

Because males (ZZ) and females (ZW) differ in their ploidy at the Z chromosome, we ran sex-specific models. Each model included gene expression as the response variable, genotype at the Z locus as a fixed effect (additive coding), brood of origin as a random effect to account for shared environmental and genetic background. For each gene-sex combination, we compared two models: one with and one without genotype as a fixed effect. Model comparison was performed using likelihood ratio tests (ANOVA). The resulting *p*-values were corrected for multiple testing using the Bonferroni method[85]. Only corrected *p*-values below the significant threshold ($\alpha = 0.05$) were considered evidence that genotype significantly affected gene expression.

For genes showing a significant association between genotype and gene expression, we further investigated whether variation in gene expression was associated with individual phenotypic traits. To do so, we fitted linear mixed models using the *lme()* function from the *nlme* package (v3.1-168, R v4.2.2)[89]. As feathers were collected from the breast of the fledglings, we focused this analysis on plumage coloration and spottiness traits measured on the breast only. For each gene, we built two separate models: a baseline model including only sex as fixed effect; a full model including both the sex and gene expression levels as fixed effects. In both models, brood of origin was included as a random effect to account for non-independence among siblings.

Genetic architecture of plumage coloration and spottiness

Variance partitioning among QTLs for color and spottiness. To quantify the contribution of the genomic loci identified in the GWASs and to evaluate the sex-specific genetic architecture of plumage coloration and spottiness, we fitted Bayesian generalized linear models using the R package brms (v2.0.22, R v4.4.2)[49]. Given the well-documented sexual dimorphism in the species, we constructed independent models for males and females for each trait. For each trait/sex combination, we considered multiple model parameterizations to explore the nature of genotype-phenotype relationships. In one set of models, genotypes were encoded as dosage values ranging from 0 to 2, assuming additive effects of alleles. In an alternative set, genotypes were treated as categorical factors, allowing for non-additive and dominance effects at each locus.

To identify the most appropriate model for each trait and sex, we compared the prediction accuracy of all fitted models using the widely applicable information criterion (WAIC). WAIC is a Bayesian model selection tool that evaluate the log-likelihood averaged across posterior simulations, providing a robust metric for model comparison that outperforms AIC and DIC in complex hierarchical models[90]. This approach is implemented in the brms function *loo_compare*(). Model comparison and selection were implemented using the loo_compare() functions in the brms package. This approach allowed us to identify the genetic model best explaining trait variance across sexes and traits, offering insights into sex-specific dominance or additive genetic effects at key QTLs.

In the models for plumage coloration, we included as fixed effects the genotype at the three loci identified as significantly associated with the trait in the GWAS analyses: the SNP on Super-Scaffold 26, position 22,522,039 (hereafter referred to as genoMC1R); the SNP on Super-scaffold 42, position 29,808,233 (genoZcol); the SNP on Super-scaffold 6, position 27,439,651 (genoCORIN)). We also included the developmental stage at phenotyping (fledglings versus adults) as a fixed factor, to control for age-related variation in coloration[44]. To account for relatedness among individuals and estimate the additive genetic variance (V_A), the genomic relatedness matrix (GRM) calculated from autosomal SNPs was included in the models as a random effect. To investigate potential epistatic interactions between loci, we additionally fitted two extended models including the interaction term between genoMC1R and genoZcol. All models assumed a Gaussian distribution for the coloration trait. Bayesian inference was performed using 8 independent Markov chains, each run for 10,000 iterations, with a burn-in of 2000 iterations. We set the *adapt_delta* parameter to 0.9 to ensure proper convergence.

For the analysis of spottiness, we included as fixed predictors the genotype of individuals at the two loci significantly associated with the trait in the GWAS: the SNP on Super-scaffold 42, position 29,812,087 (hereafter referred as genoZspot); the SNP on Super-scaffold 26, position 22,522,039 (genoMC1R). We also included the developmental stage at which the phenotype was measured (fledgling or adult) as a fixed effect to control for age-related variation in coloration[44]. To account for genetic relatedness, we included the genomic relatedness matrix (GRM), estimated across the entire genome, as random effect in the models, allowing us to estimate V_A. Because the distribution of the spottiness trait was zero-inflated, with an over-representation of individuals showing no spots, we modeled the phenotype using a hurdle-gamma distribution. In this framework, the presence or absence of spots was modeled as the binary hurdle component, while the degree of spottiness among spotted individuals was modeled with a gamma distribution. All models were run using 8 independent Markov chains, each for 10,000 iterations, with a burn-in of 2000 iterations. The *adapt_delta* parameter was set to 0.9 to ensure robust convergence. For the best-fitting models, we extracted the total variance explained using the *bayes_R2()* function from the *brms* package[49].

Allele frequency changes and inference of evolutionary processes

Temporal evolution of allele frequencies. To assess whether the allele frequencies at the loci identified by the GWAS for coloration and spottiness have changed over time, we estimated allele frequencies for each year based on the adults alive in that year. To account for variation in sample sizes, particularly in earlier years with limited data, we computed 95% confidence intervals around each annual allele frequency estimate using the Clopper–Pearson exact binomial method[91].

To evaluate whether allele frequencies have changed systematically over time, we fitted linear models with allele frequency as the response variable and year as the predictor. To avoid potential bias from years with high uncertainty in frequency due to small sample sizes, we restricted this analysis to the period between 1995 and 2022.

Estimation of effective population size. We estimated the effective population size (N_e) from the whole-genome resequencing data using the standard formula $\pi/(4 \times Mu)$, where π is the nucleotide diversity and μ the mutation rate per site per generation[55]. For each year, π was calculated using only for the adult individuals alive during that year. To ensure consistency with the variant calling pipeline, the total sequence length (L) used for π estimation matched the length of the reference genome after mapping ($L = 1{,}131{,}781{,}302$ bp). Nucleotide diversity (π) was computed using the *pi.dosage()* function from the hierfstat (v0.5-11, R v4.2.2)[81]. We adopted the recently estimated mutation rate for the barn owl (*tyto alba*) ($Mu = 5.6 \times 10^{-9}$)[56]. To obtain 95% confidence intervals around annual Ne estimates, we performed bootstrap resampling of non-overlapping 1 Mb segments along the genome and repeating this process 100 times per year.

Simulation of allele frequency through time. To assess whether the observed allele frequency changes could be explained solely by genetic drift, we simulated allele frequency dynamics under a neutral scenario. This approach allowed us to test whether selection may have also played a role in shaping the temporal trajectory of alleles associated with coloration and spottiness. We modeled drift in a population with overlapping generations, reflecting the demography of barn owls, which have an average lifespan of approximately three years[92]. Accordingly, in our simulations, one-third of the population was renewed each year, while the remaining two-thirds persisted from the previous generation. To account for uncertainty in allele frequency estimates arising from small sample sizes in the early years of the study, simulations began with the initial allele frequencies observed in the population over the first 4 years of the study: $MC1R_{red}$ 0.047; $Z_{spotted}$ 0.685 (based on $n = 32$ individuals). For each locus, we ran 1000 replicates assuming a constant population size of 400 individuals, a conservative estimate consistent with the lowest census sizes of the Swiss barn owl population during the study period[54]. This choice reflects the short-term, contemporary timescale of the analysis, for which stochastic allele frequency change is primarily driven by the number of breeding individuals contributing alleles each year, rather than by long-term effective population size inferred from historical genome-wide diversity and large-scale connectivity. To test if the inference of selection at MC1R was robust to population size assumptions, we also performed simulations with larger population sizes ($N = 1000$, consistent with values inferred from ApproxWF, and $N = 10{,}000$, approaching the order of magnitude of the N_e estimated from nucleotide diversity while remaining computationally tractable). Given that *MC1R* is located on an autosome, simulation assumed diploidy across all individuals.

The Z locus is located on the Z chromosome, which is hemizygous in females (ZW) and homozygous in males (ZZ). We therefore used a sex-aware simulation approach, where the effective number of alleles per individual differed by sex (1 for females, 2 for males), and sexes were sampled at a 1:1 ratio. To mimic the empirical data structure, for each simulated year we subsampled the same number of individuals as were observed alive in the actual dataset for that year. This allowed us to account for sampling variance and generate realistic confidence intervals around the neutral expectation.

Disentangling drift and selection on major QTLs. To evaluate whether the allele frequency changes observed at key QTLs were consistent with neutral expectations or indicative of selection, we used the approxWF framework (comit 85793eb)[93]. This method compares empirical allele frequency trajectories to those expected under a Wright–Fisher diffusion process and is particularly well suited for time-series population genetic data. It allows explicit estimation of selection coefficients while accounting for stochastic fluctuations due to genetic drift. For each focal locus, we fitted two models: (i) a neutral model (drift only), and (ii) a selection incorporating a selection coefficient (s) and, optionally, a dominance coefficient (h). Each Markov Chain Monte Carlo (MCMC) chain was run for 1,010,000 iterations, with the first 10,000 iterations discarded as burn-in, and sampling conducted every 10 steps. For the Z-linked locus, we used the default MCMC parameters. For MC1R, however, the default settings led to low acceptance rates and unstable chains, likely due to lower starting allele frequency. To improve convergence and stability, we reduced the proposal step sizes by half, using logN_step = 0.025 for the log population size, s_step = 0.005 for the selection coefficient (s), and h_step = 0.005 for the dominance coefficient (h).

For each locus, we first fitted a Wright–Fisher model incorporating both genetic drift and selection, and compared the log-likelihoods and the proportion of MCMC iterations supporting selection versus neutrality. A higher posterior support for selection and an improved model likelihood was interpreted as evidence that selection better explains the observed allele frequency trajectory. We then fitted a selection-only model and assessed whether the 95% credible interval of the estimated selection coefficient excluded zero. A locus was considered under directional selection if it met the following three criteria: (i) the selection model had a higher log-likelihood than the drift-only model; (ii) the 95% credible interval did not include zero; and (iii) the MCMC chain spent a greater proportion of iterations in the selection model relative to the baseline proposal rate. In the approxWF reversible-jump framework, transitions between neutral and selection models are proposed at a fixed rate (default 10%), and deviations from this baseline are interpreted as preferential sampling of the model including selection. Loci that failed to meet these criteria, particularly those with overlapping credible intervals including zero and low support for selection, were interpreted as being consistent with neutral drift. To provide further insights into the evolutionary dynamics, effective population size and dominance coefficients were jointly inferred from the MCMC chains.

Finally, we used the simulation and dominance coefficients estimated from the selection-only models to simulate allele frequency trajectories under selection. Simulations followed the scheme outlined in the *Simulation of Allele Frequency Through Time* section. Genotype fitness was parameterized as 1, 1-hs, and 1-s for autosomal loci (MC1R), and as 1 and 1-s for hemizygous females at the Z-linked locus.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The sequencing data underlying this article are available in the GenBank Sequence Read Archive Database at https://www.ncbi.nlm.nih.gov/sra. High coverage sequences used for the reference panel can be accessed within BioProject PRJNA700797, PRJNA727915, PRJNA727977, PRJNA774943, PRJNA925445, and PRJNA1172395. Low coverage sequences can be accessed within BioProject PRJNA1391678. Phenotypic data use throughout the manuscript can be found at https://zenodo.org/records/18037719.

Code availability

Code used to prepare the genomic data is available at https://github.com/cumtr/Cumer_et_al_EvolutionColor_WGS and archived on Zenodo under https://doi.org/10.5281/zenodo.22278101 [https://zenodo.org/records/22278101][94]. Code used for downstream analyses is available at https://github.com/cumtr/Cumer_et_al_EvolutionColor_Analyses and archived on Zenodo under https://doi.org/10.5281/zenodo.22278594 [https://zenodo.org/records/22278594][95].

References

1. Fisher, R. A. *The Genetical Theory of Natural Selection* (The Clarendon Press, 1930).
2. Wright, S. The roles of mutation, inbreeding, crossbreeding, and selection in evolution http://www.esp.org/books/6th-congress/facsimile/contents/6th-cong-p356-wright.pdf (1932).
3. Lewontin, R. C. *The Genetic Basis of Evolutionary Change* (Columbia University Press, 1974).
4. Lande, R. & Arnold, S. J. The measurement of selection on correlated characters. *Evolution* **37**, 1210–1226 (1983).
5. Barrett, R. D. H. & Hoekstra, H. E. Molecular spandrels: tests of adaptation at the genetic level. *Nat. Rev. Genet.* **12**, 767–780 (2011).
6. Rausher, M. D. & Delph, L. F. Commentary: when does understanding phenotypic evolution require identification of the underlying genes? *Evolution* **69**, 1655–1664 (2015).
7. Nelson, T. C., Jones, M. R., Velotta, J. P., Dhawanjewar, A. S. & Schweizer, R. M. Unveiling connections between genotype,

phenotype, and fitness in natural populations. *Mol. Ecol.* **28**, 1866–1876 (2019).
8. Bonnet, T. et al. Genetic variance in fitness indicates rapid contemporary adaptive evolution in wild animals. *Science* **376**, 1012–1016 (2022).
9. Stinchcombe, J. R. & Hoekstra, H. E. Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity* **100**, 158–170 (2008).
10. Homburger, J. R. et al. Low coverage whole genome sequencing enables accurate assessment of common variants and calculation of genome-wide polygenic scores. *Genome Med.* **11**, 74 (2019).
11. Strickland, K. et al. Microevolutionary change in wild stickleback: using integrative time-series data to infer responses to selection. *Proc. Natl. Acad. Sci. USA.* **121**, e2410324121 (2024).
12. Enbody, E. D. et al. Community-wide genome sequencing reveals 30 years of Darwin's finch evolution. *Science* **381**, eadf6218 (2023).
13. Cuthill, I. C. et al. The biology of color. *Science* **357**, eaan0221 (2017).
14. Price-Waldman, R. & Stoddard, M. C. Avian coloration genetics: recent advances and emerging questions. *J. Hered.* **112**, 395–416 (2021).
15. San-Jose, L. M. & Roulin, A. Genomics of coloration in natural animal populations. *Phil. Trans. R. Soc. B.* **372**, 20160337 (2017).
16. Orteu, A. & Jiggins, C. D. The genomics of coloration provides insights into adaptive evolution. *Nat. Rev. Genet.* **21**, 461–475 (2020).
17. Campagna, L. & Toews, D. P. L. The genomics of adaptation in birds. *Curr. Biol.* **32**, R1173–R1186 (2022).
18. Cumer, T. et al. The genomic architecture of continuous plumage colour variation in the European barn owl (*Tyto alba*). *Proc. Biol. Sci.* **291**, 20231995 (2024).
19. Dunning, J., Sheard, C. & Endler, J. A. Viewing conditions predict evolutionary diversity in avian plumage colour. *Proc. Biol. Sci.* **292**, 20241728 (2025).
20. Mason, N. A. & Bowie, R. C. K. Plumage patterns: ecological functions, evolutionary origins, and advances in quantification. *Auk* **137**, ukaa060 (2020).
21. Roulin, A., Dijkstra, C., Riols, C. & Ducrest, A.-L. Female- and male-specific signals of quality in the barn owl. *J. Evol. Biol.* **14**, 255–266 (2001).
22. Becciu, P., Séchaud, R., Schalcher, K., Plancherel, C. & Roulin, A. Prospecting movements link phenotypic traits to female annual potential fitness in a nocturnal predator. *Sci. Rep.* **13**, 5071 (2023).
23. Charter, M., Leshem, Y., Izhaki, I. & Roulin, A. Pheomelanin-based colouration is correlated with indices of flying strategies in the barn owl. *J. Ornithol.* **156**, 309–312 (2015).
24. San-Jose, L. M. et al. Differential fitness effects of moonlight on plumage colour morphs in barn owls. *Nat. Ecol. Evol.* **3**, 1331–1340 (2019).
25. Béziers, P., Ducrest, A.-L., Simon, C. & Roulin, A. Circulating testosterone and feather-gene expression of receptors and metabolic enzymes in relation to melanin-based colouration in the barn owl. *Gen. Comp. Endocrinol.* **250**, 36–45 (2017).
26. Almasi, B., Roulin, A., Korner-Nievergelt, F., Jenni-Eiermann, S. & Jenni, L. Coloration signals the ability to cope with elevated stress hormones: effects of corticosterone on growth of barn owls are associated with melanism. *J. Evol. Biol.* **25**, 1189–1199 (2012).
27. Béziers, P., Ducrest, A.-L., San-Jose, L. M., Simon, C. & Roulin, A. Expression of glucocorticoid and mineralocorticoid receptor genes co-varies with a stress-related colour signal in barn owls. *Gen. Comp. Endocrinol.* **283**, 113224 (2019).
28. Roulin, A., Jungi, T. W., Pfister, H. & Dijkstra, C. Female barn owls (*Tyto alba*) advertise good genes. *Proc. Biol. Sci.* **267**, 937–941 (2000).
29. Roulin, A., Riols, C., Dijkstra, C. & Ducrest, A.-L. Female plumage spottiness signals parasite resistance in the barn owl (*Tyto alba*). *Behav. Ecol.* **12**, 103–110 (2001).
30. Roulin, A. & Jensen, H. Sex-linked inheritance, genetic correlations and sexual dimorphism in three melanin-based colour traits in the barn owl. *J. Evol. Biol.* **28**, 655–666 (2015).
31. Roulin, A., Richner, H. & Ducrest, A.-L. Genetic, environmental, and condition-dependent effects on female and male ornamentation in the barn owl *Tyto alba*. *Evolution* **52**, 1451 (1998).
32. Roulin, A. & Dijkstra, C. Genetic and environmental components of variation in eumelanin and phaeomelanin sex-traits in the barn owl. *Heredity* **90**, 359–364 (2003).
33. Topaloudis, A. et al. Benchmarking imputation accuracy in the presence or absence of a reference panel. *Mol. Biol. Evol.* **43**, nsag094 (2026).
34. San-Jose, L. M., Ducret, V., Ducrest, A.-L., Simon, C. & Roulin, A. Beyond mean allelic effects: a locus at the major color gene MC1R associates also with differing levels of phenotypic and genetic (co) variance for coloration in barn owls. *Evolution* **71**, 2469–2483 (2017).
35. Funk, E. R. et al. The genetic basis of plumage coloration and elevation adaptation in a clade of recently diverged alpine and arctic songbirds. *Evolution* **77**, 705–717 (2023).
36. San-Jose, L. M. et al. MC1R variants affect the expression of melanocortin and melanogenic genes and the association between melanocortin genes and coloration. *Mol. Ecol.* **26**, 259–276 (2017).
37. Ellegren, H. et al. Faced with inequality: chicken do not have a general dosage compensation of sex-linked genes. *BMC Biol.* **5**, 40 (2007).
38. Wolf, J. B. & Bryk, J. General lack of global dosage compensation in ZZ/ZW systems? Broadening the perspective with RNA-seq. *BMC Genom.* **12**, 91 (2011).
39. Sigeman, H. et al. Insights into avian incomplete dosage compensation: sex-biased gene expression coevolves with sex chromosome degeneration in the common whitethroat. *Genes* **9**, 373 (2018).
40. Tsai, L.-C. L., Shimizu-Albergine, M. & Beavo, J. A. The high-affinity cAMP-specific phosphodiesterase 8B controls steroidogenesis in the mouse adrenal gland. *Mol. Pharm.* **79**, 639–648 (2011).
41. Zhang, X. et al. PDE5 inhibitor promotes melanin synthesis through the PKG pathway in B16 melanoma cells. *J. Cell. Biochem.* **113**, 2738–2743 (2012).
42. Wei, B., Zhang, Y.-P., Yan, H.-Z., Xu, Y. & Du, T.-M. Cilostazol promotes production of melanin by activating the microphthalmia-associated transcription factor (MITF). *Biochem. Biophys. Res. Commun.* **443**, 617–621 (2014).
43. Roulin, A. Geographic variation in sexual dimorphism in the barn owl *Tyto alba*: a role for direct selection or genetic correlation? *J. Avian Biol.* **34**, 251–258 (2003).
44. Dreiss, A. N. & Roulin, A. Age-related change in melanin-based coloration of barn owls (*Tyto alba*): females that become more female-like and males that become more male-like perform better. *Biol. J. Linn. Soc.* **101**, 689–704 (2010).
45. San-Jose, L. M. et al. Effect of the MC1R gene on sexual dimorphism in melanin-based colorations. *Mol. Ecol.* **24**, 2794–2808 (2015).
46. Bourgeois, Y. X. C. et al. Candidate gene analysis suggests untapped genetic complexity in melanin-based pigmentation in birds. *J. Hered.* **107**, 327–335 (2016).
47. Wood, A. W., Szpiech, Z. A., Lovette, I. J., Smith, B. T. & Toews, D. P. L. Genomes of the extinct Bachman's warbler show high divergence and no evidence of admixture with other extant *Vermivora* warblers. *Curr. Biol.* **33**, 2823–2829 (2023).
48. Enshell-Seijffers, D., Lindon, C. & Morgan, B. A. The serine protease Corin is a novel modifier of the agouti pathway. *Development* **135**, 217–225 (2008).

49. Bürkner, P.-C. brms: an R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* **80**, 1–28 (2017).
50. van der Bijl, W. & Mank, J. E. Widespread cryptic variation in genetic architecture between the sexes. *Evol. Lett.* **5**, 359–369 (2021).
51. Tosto, N. M., Beasley, E. R., Wong, B. B. M., Mank, J. E. & Flanagan, S. P. The roles of sexual selection and sexual conflict in shaping patterns of genome and transcriptome variation. *Nat. Ecol. Evol.* **7**, 981–993 (2023).
52. Roulin, A., Altwegg, R., Jensen, H., Steinsland, I. & Schaub, M. Sex-dependent selection on an autosomal melanic female ornament promotes the evolution of sex ratio bias. *Ecol. Lett.* **13**, 616–626 (2010).
53. Cumer, T. et al. Landscape and climatic variations shaped secondary contacts amid barn owls of the Western Palearctic. *Mol. Biol. Evol.* **39**, msab343 (2022).
54. Swiss Breeding Bird Atlas | Swiss Ornithological Institute. *vogelwarte.ch* https://www.vogelwarte.ch/modx/en/projects/monitoring/swiss-breeding-bird-atlas.
55. Wright, S. Size of population and breeding structure in relation to evolution. *Science* **87**, 430–431 (1938).
56. Topaloudis, A. et al. The effect of age and sex on the rate of de novo mutations in barn owls. *Genome Biol. Evol.* **18**, evag175 (2026).
57. Hewett, A. et al. Inbreeding depression in morphological and growth traits of wild barn owls (Tyto alba). *Proc. R. Soc. B.* **293**, 20252708 (2026).
58. Béziers, P. & Roulin, A. Sexual maturity varies with melanic plumage traits in the barn owl. *J. Avian Biol.* **52**, jav.02715 (2021).
59. Py, I., Ducrest, A.-L., Duvoisin, N., Fumagalli, L. & Roulin, A. Ultraviolet reflectance in a melanin-based plumage trait is heritable. *Evol. Ecol. Research* **8**, 483–491 (2006).
60. Taylor, I. R. Age and sex determination of barn owls *Tyto alba alba*. *Ringing & Migration* https://doi.org/10.1080/03078698.1993.9674050 (1993).
61. Roulin, A. Nonrandom pairing by male barn owls (*Tyto alba*) with respect to a female plumage trait. *Behav. Ecol.* **10**, 688–695 (1999).
62. R Core Team. *R: A Language and Environment for Statistical Computing* (3.5. 1)[Computer Software] (R Foundation for Statistical Computing, 2020).
63. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
64. Roulin, A. Delayed maturation of plumage coloration and plumage spottedness in the barn owl (*Tyto alba*). *J. Ornithol.* **140**, 193–197 (1999).
65. Topaloudis, A. et al. The recombination landscape of the barn owl, from families to populations. *Genetics* **229**, iyae190 (2025).
66. Cumer, T. et al. Genomic basis of insularity and ecological divergence in barn owls (*Tyto alba*) of the Canary Islands. *Heredity* **129**, 281–294 (2022).
67. Machado, A. P. et al. Unexpected post-glacial colonisation route explains the white colour of barn owls (*Tyto alba*) from the British Isles. *Mol. Ecol.* **31**, 482–497 (2022).
68. Machado, A. P. et al. Genomic consequences of colonisation, migration and genetic drift in barn owl insular populations of the eastern Mediterranean. *Mol. Ecol.* **31**, 1375–1388 (2022).
69. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
70. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
71. Van der Auwera, G. A. et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr. Protoc. Bioinform.* **43**, 11.10.1–11.10.33 (2013).
72. Corval, H. et al. Genetically diverse populations hold the keys to climatic adaptation in the Western barn owl (Tyto alba). *Commun. Biol.* **9**, 1172 (2026).
73. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).
74. Martin, M. et al. Read-Based Phasing and Analysis of Phased Variants with WhatsHap. *Methods Mol. Biol.* **2590**, 127–138 (2023).
75. Delaneau, O., Zagury, J.-F., Robinson, M. R., Marchini, J. L. & Dermitzakis, E. T. Accurate, scalable and integrative haplotype estimation. *Nat. Commun.* **10**, 5436 (2019).
76. Choi, Y., Chan, A. P., Kirkness, E., Telenti, A. & Schork, N. J. Comparison of phasing strategies for whole human genomes. *PLoS Genet.* **14**, e1007308 (2018).
77. Okonechnikov, K., Conesa, A. & García-Alcalde, F. Qualimap 2: advanced multi-sample quality control for high-throughput sequencing data. *Bioinformatics* **32**, 292–294 (2016).
78. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**, 2987–2993 (2011).
79. Rubinacci, S., Ribeiro, D. M., Hofmeister, R. J. & Delaneau, O. Efficient phasing and imputation of low-coverage sequencing data using large reference panels. *Nat. Genet.* **53**, 120–126 (2021).
80. Goudet, J., Kay, T. & Weir, B. S. How to estimate kinship. *Mol. Ecol.* **27**, 4121–4135 (2018).
81. Goudet, J. hierfstat, a package for r to compute and test hierarchical F-statistics. *Mol. Ecol. Notes* **5**, 184–186 (2005).
82. Perdry, H. & Dandine-Roulland, C. gaston: Genetic Data Handling (QC, GRM, LD, PCA) & Linear Mixed Models. 1.6 https://doi.org/10.32614/CRAN.package.gaston (2015).
83. Yang, J., Zaitlen, N. A., Goddard, M. E., Visscher, P. M. & Price, A. L. Advantages and pitfalls in the application of mixed-model association methods. *Nat. Genet.* **46**, 100–106 (2014).
84. Sul, J. H., Martin, L. S. & Eskin, E. Population structure in genetic studies: confounding factors and mixed models. *PLoS Genet.* **14**, e1007309 (2018).
85. Bonferroni, C. E. *Teoria Statistica Delle Classi e Calcolo Delle Probabilità* (Seeber, 1936).
86. Ducrest, A.-L. et al. Melanin and neurotransmitter signalling genes are differentially co-expressed in growing feathers of white and rufous barn owls. *Pigment Cell Melanoma Res.* **38**, e70001 (2025).
87. Petersen, P. H. D., Lopacinska-Jørgensen, J., Høgdall, C. K. & Høgdall, E. V. Expression stability analysis of candidate references for normalization of RT-qPCR data using RefSeeker R package. *Bio Protoc.* **13**, e4801 (2023).
88. Kohl M. SLqPCR: Functions for analysis of real-time quantitative PCR data at SIRS-Lab GmbH. SIRS-Lab GmbH, Jena. www.sirs-lab.com (2007).
89. Pinheiro, J., Bates, D. & R Core Team. nlme: Linear and Nonlinear Mixed Effects Models. 3.1-168 https://doi.org/10.32614/CRAN.package.nlme (1999).
90. Vehtari, A., Gelman, A. & Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* **27**, 1413–1432 (2017).
91. Clopper, C. J. & Pearson, E. S. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* **26**, 404–413 (1934).
92. Altwegg, R., Schaub, M. & Roulin, A. Age-specific fitness components and their temporal variation in the barn owl. *Am. Nat.* **169**, 47–61 (2007).
93. Ferrer-Admetlla, A., Leuenberger, C., Jensen, J. D. & Wegmann, D. An approximate Markov model for the Wright–Fisher diffusion and its application to time series data. *Genetics* **203**, 831–846 (2016).
94. Cumer, T. cumtr/Cumer_et_al_EvolutionColor_WGS: Release v1.0.0. Preprint at https://doi.org/10.5281/zenodo.22278101 (2026).
95. Cumer, T. cumtr/Cumer_et_al_EvolutionColor_Analyses: Release v1.0.0. Preprint at https://doi.org/10.5281/zenodo.22278594 (2026).

Acknowledgements
We thank all individuals who contributed to this long-term project, including field assistants, former PhD students, postdoctoral researchers, and collaborators (full list in Supplementary Note 1: Acknowledgments for Sample Collection). We are also grateful to the Genomic Technologies Facility (GTF), Lausanne, for sequencing support and to Olivier Delaneau for insightful discussions.

Author contributions
T.C. and J.G. were responsible for the conceptualization of the study, with contributions from A.T., E.L., A.L.D. and A.R. B.A. and A.R. supervised sample acquisition and individual phenotyping over the long-term study period. A.L.D. supervised sample preparation for sequencing and carried out the gene expression experiment, with contributions from C.S. J.M. and A.C. carried out the sequencing and generated the raw sequencing data. A.T. assembled the reference panel, with contributions from T.C. and E.L. E.L. performed genotyping and imputation of the low-coverage samples, with contributions from T.C., A.T. and S.S.C. T.C. conducted the formal analysis, with assistance from P.B. for phenotypic analyses, A.H. for the quantitative genetic approach, and A.L.D. for the gene expression analyses. A.T., E.L. and J.G. contributed to the analyses throughout the process. A.R. and J.G. were responsible for funding acquisition. T.C. wrote the original draft, with contributions from A.T., E.L., A.L.D. and J.G. All authors read, provided feedback on, and approved the final version of the paper.

Funding
This work was supported by SNSF grants number 31003A_173178, 310030_200321/1 and ICC00I0-23168 to A.R. and 31003A_179358 and 310030_215709 to J.G. Open access funding provided by University of Lausanne.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information The online version contains Supplementary Material available at https://doi.org/10.1038/s41467-026-78632-4.

Correspondence and requests for materials should be addressed to Tristan Cumer or Jérôme Goudet.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2026