

Genome-wide SNVs revealed reticulate evolutionary history of a Eurasian-African songbird radiation (*Saxicola torquatus* complex)

Authors: Chuan Jiang^{1,2}; Yang Liu²; Wenwen Zhu³; Nassoro Mohamed^{1,4}; Jiazheng Jin⁵; Bo Li^{1,6*}

Institutes:

¹*College of Wildlife and Protected Area, Northeast Forestry University, Harbin, Heilongjiang 150040, China*

²*School of Ecology, Sun Yat-Sen University, Shenzhen, Guangdong 518107, China*

³*College of Life Science and Technology, Jinan University, Guangzhou, Guangdong 510632, China*

⁴*College of African Wildlife Management, Mweka, Moshi P.O Box 3031, Kilimanjaro, Tanzania*

⁵*BGI Research, Wuhan, Hubei 430074, China*

⁶*State Forestry and Grassland Administration Detecting Center of Wildlife, Harbin, Heilongjiang 150040, China*

*Corresponding authors, E-mail: libo_770206@nefu.edu.cn



Abstract: Phylogenomics has the power to reveal complex evolutionary hypotheses across taxa, thereby allowing important glimpses into the evolutionary history, such as incomplete lineage sorting (ILS) and introgression among closely related species. The *Saxicola torquatus* complex comprises several widely distributed insectivorous bird species across the Palearctic and Afrotropical realms. Nevertheless, the evolutionary history within this complex remains contentious, and previous findings suggest complications caused by cytonuclear discordance. Here, we conducted a phylogenomic analysis of nearly complete taxa within the complex using genome-wide single nucleotide variants to investigate potential causes of phylogenetic discordance across genomic regions and elucidate its evolutionary history. We detected multiple reticulate events during their diversification, leading to extensive phylogenetic discordance across the genome. The mitochondrial genome and non-recombining region of the W chromosome exhibited a consistent gene tree that conflicted with the species tree, with introgression being the most plausible explanation among several hypotheses. The Z chromosome was less affected by ILS/introgression and exhibited relatively stable local phylogenies compared to autosomes. Furthermore, we found that strongly introgressed regions were predominantly concentrated at the ends of smaller chromosomes, and the strength of introgression among species pairs showed consistently positive correlations, likely attributed to shared recombination hotspots in these regions. These findings stress the heterogeneous genomic landscape influenced by intricate evolutionary processes (e.g., ILS, introgression, and recombination) that endorse a reticulate evolution model for global speciation in the *S. torquatus* complex.

Keywords: Introgression; Phylogenomics; Stonechat; Single nucleotide variants



INTRODUCTION

Early phylogenetic studies, often limited to a few loci, are increasingly challenged by genome-scale data revealing pervasive phylogenetic discordance—discordance among gene trees or between gene and species trees (Edelman et al., 2019; Jiang et al., 2024; Zhang et al., 2024). The two main processes underlying this heterogeneity are incomplete lineage sorting (ILS) and introgression (Steenwyk et al., 2023). ILS results from the random retention of ancestral polymorphisms during population divergence, increasing with larger ancestral effective population sizes (N_e) and shorter internode times (Avice & Robinson 2008; Rivas-González et al., 2023). Introgression occurs through interspecific hybridization and backcrossing, where foreign alleles are incorporated into the recipient genome via recombination, typically following secondary contact between isolated populations and leading to reticulate evolution (Payseur & Rieseberg 2016; Ottenburghs et al., 2017). These two processes often act concurrently but contribute differentially across radiating lineages due to lineage-specific traits (Myers et al., 2024; Baraf et al., 2025). They not only influence the robustness of phylogenetic inference but also shape trait evolution and reflect biogeographic history (Feng et al., 2022; Hu et al., 2022; Rivas-González et al., 2023; Musher et al., 2024). Therefore, evaluating their respective contributions is critical for understanding the evolutionary history of focal clades.

Beyond lineage-specific traits, the effects of ILS and introgression also vary within lineages depending on the inheritance modes of different genomic components (Nikelski et al., 2023; Chase et al., 2024). For example, the Z chromosome, due to its hemizyosity in females, contains more incompatibility loci than autosomes, along with a reduced N_e and recombination rate, making it less prone to ILS and introgression (Edelman et al., 2019; Chase et al., 2024). Two maternally inherited markers, the mitochondrial (mt) genome (mitogenome) and the non-recombining region of the W chromosome (NRW), show stronger effects of hemizyosity and



lower N_e than the Z chromosome and lack recombination, further minimizing ILS and introgression (Moore 1995; Smeds et al., 2015; Mackiewicz et al., 2019). Their strict maternal inheritance, however, creates a distinct dynamic: after hybridization, they may persist in the recipient population and rapidly reach fixation due to their smaller N_e , resulting in phylogenetic discordance from the species tree (e.g., mitogenome capture) (Toews & Brelsford 2012; Nikelski et al., 2023). Moreover, due to the functional coupling between mitochondrial and nuclear genomes (Shi et al., 2025), introgressed mitogenomes may induce mitonuclear incompatibilities (Sloan et al., 2017), driving the co-introgression of mitonuclear genes alongside the mitogenome to maintain mitochondrial function (Hill 2019; Nikelski et al., 2023). Analyzing these phenomena not only facilitates accurate reconstruction of species tree but also highlights the differential roles of genomic components during lineage radiations.

Due to the widespread occurrence of avian hybridization (Ottenburghs et al., 2015; Ottenburghs 2023) and the slow evolution of intrinsic postzygotic isolation—a pattern attributed to their slower rate of molecular evolution and greater chromosomal stability, which delays the accumulation of genetic incompatibilities (Fitzpatrick 2004)—introgression frequently occurs in birds. Introgression plays a dual role in speciation (Musher et al., 2024): on one hand, it can introduce beneficial standing genetic variation from other populations, facilitating adaptation to new environments (Meier et al., 2019; Hu et al., 2022); on the other hand, it may disrupt reproductive isolation barriers between species, thereby hindering speciation (Taylor et al., 2006; Webb et al., 2011). The resulting introgression landscape—that is, the heterogeneous acceptance of foreign genetic material across the genome—may be shaped by multiple factors, including natural selection strength, local recombination rates, and the distribution of reproductive isolation genes (Gante et al., 2016; Ravinet et al., 2017; Martin & Jiggins 2017; Moran et al., 2021; Veller et al., 2023; Musher et al., 2024; Blain et al., 2025). Although introgression has been widely documented across diverse taxa, including birds (Hu



et al., 2022; Leitwein et al., 2024), mammals (Shi et al., 2025), reptiles (Myers et al., 2024), fishes (Gante et al., 2016; Baraf et al., 2025), insects (Edelman et al., 2019; Wang et al., 2025), and plants (Fu et al., 2022), our understanding of its evolutionary consequences and the ecological or genomic factors that drive it remains uneven across lineages, with most insights still coming from a limited number of well-studied systems. Expanding research to include more non-model species will enhance our understanding of the genetic mechanisms underlying speciation and species maintenance.

The *Saxicola torquatus* complex comprises several widely distributed insectivorous bird species across the Palaearctic and Afrotropical realms, with a long-debated taxonomic history (Zink et al., 2009; Opaev et al., 2018). Historically, this complex was considered a superspecies, the common stonechat *S. torquatus*, encompassing multiple subspecies (Zink et al., 2009; Svensson et al., 2012; Opaev et al., 2018). Based on morphology, vocalizations, habitat preference, mitochondrial genes, and genomic data from a limited number of taxa (Wittmann et al., 1995; Urquhart 2002; Illera et al., 2008; Zink et al., 2009; Svensson et al., 2012; Hellström & Norevik 2014; Van Doren et al., 2017; Opaev et al., 2018; Loskot & Bakhtadze 2020; Hellström & Waern 2023; Zhao et al., 2023; Jiang et al., 2025), the *S. torquatus* complex is now tentatively classified into six species: *S. torquatus*, *S. tectes*, *S. rubicola*, *S. dacotiae*, *S. maurus*, and *S. stejnegeri* (Winkler et al., 2020; Gill et al., 2025). However, the species-level status of *S. stejnegeri* remains debated due to its parapatric distribution with *S. maurus* and potential hybridization (Hellström & Norevik 2014; Hellström & Waern 2023), leading to the suggestion that it might be a subspecies of *S. maurus*, namely *S. m. stejnegeri* (Chesser et al., 2022; Gill et al., 2025; AviList Core Team 2025). This ambiguity is partly attributed to previous studies on this species relying on mitochondrial genes, which have not yielded a conclusive phylogeny. Additionally, the phylogenetic position of *S. maurus* shows cytonuclear discordance (Supplementary Figure S1), likely due to ILS, ancient introgression, or other



Zoological
Research

biological processes (Zink et al., 2009; Van Doren et al., 2017), underscoring the complexity of their evolutionary history. Thus, the *Saxicola torquatus* complex represents a promising system for studying how ILS and introgression shape gene-species tree conflicts.

In this study, we performed whole-genome sequencing on the key lineage *S. stejnegeri* within the *S. torquatus* complex and integrated it with published genomic data from several other lineages within this complex. This ultimately covers nearly complete taxa lineages within the *S. torquatus* complex (except *S. tectes*), as well as *S. rubetra* outside the complex. To resolve their evolutionary history, we first reconstructed the phylogeny of this complex using different genomic regions and assessed the levels of ILS and introgression. Based on the discovered reticulate evolution, we further analyzed the introgression landscape and correlations between four species pairs and performed functional enrichment analysis of introgressed genes to explore potential drivers of introgression patterns. We also conducted a mitonuclear co-introgression test for an ancient hybridization event and evaluated the causes of cytonuclear discordance in the phylogenetic position of *S. maurus*. Finally, based on the reconstructed phylogeny and hybridization history, we provided new insights into the taxonomic classification and colonization history of the *S. torquatus* complex.

MATERIALS AND METHODS

Sample Collection and Sequencing

A single *S. stejnegeri* carcass sample was obtained from Bei'an City, Heilongjiang Province, seized by the State Forestry and Grassland Administration Wildlife Detection Center (Harbin, China) in an illegal capture case. Total genomic DNA was extracted from muscle tissue using a tissue extraction kit (UElandy, Suzhou, China). High-quality DNA was



Zoological
Research

fragmented by sonication, and a library with an insert size of ~350 bp was prepared on the DNBSEQ platform following the manufacturer's instructions (MGIEasy Universal DNA Library Preparation Kit, BGI). Sequencing was performed using a 150 bp paired-end strategy on the DNBSEQ-T7 (MGI, China).

Data Processing

In addition to the *S. stejnegeri* sequences we generated, we obtained individual sequencing data for *S. rubetra*, *S. rubicola*, *S. maurus* and *Oenanthe deserti* (used as an outgroup), along with pooled sequencing data for *S. torquatus* (52 individuals) and *S. dacotiae* (56 individuals) from the SRA database (Van Doren et al., 2017; Bergeron et al., 2023) (Supplementary Table S1). All data were filtered using fastp (Chen 2023) with default parameters, and for the single tube Long Fragment Read data of *S. rubetra* and *S. rubicola*, barcode sequences at the 3' end of the reads were trimmed.

Quality-controlled reads were aligned to the chromosomal-level genome of male *O. melanoleuca* (GCA_029582105.1) (Peona et al., 2023) using BWA-MEM v.0.7.18 (Jung & Han 2022) with default parameters. Reads with a mapping quality (MAPQ)>20 were retained and were sorted by genomic coordinates using SAMtools v.1.21 (Danecek et al., 2021). PCR duplicates were removed using Picard v.3.1.1 (McKenna et al., 2010). Alignment rates and coverage for each sample were calculated using the flagstat and coverage modules in SAMtools, respectively. Due to the low mapping rate observed for one species (*S. torquatus*), we extracted 10% of its single-end sequencing reads and aligned them against the core nucleotide database using BLAST v.2.16.0 (Camacho et al., 2009) to check for contamination. Additionally, sex was determined by the normalized Z chromosome depth (Z depth / mean autosomal depth), with expected values of ~1 for males and ~0.5 for females.

Based on the deduplicated mapping results, single nucleotide variants (SNVs) were called using HaplotypeCaller, and joint genotyping was performed on all samples using



Zoological
Research

GenotypeGVCFs in GATK v.4.5 (McKenna et al., 2010). Next, Indels were excluded using the SelectVariants module, and hard-filtering was applied with the VariantFiltration module in GATK using the following criteria: $RPRS < -8$, $SOR > 3.0$, $QD < 2.0$, $FS > 60.0$, $MQ < 40.0$, and $MQRankSum < -12.5$. Finally, SNVs were soft-filtered using Bcftools v.1.20 (Danecek et al., 2021) and Vcftools v.0.1.16 (Danecek et al., 2011), retaining only biallelic loci, excluding SNVs missing in any sample, and removing heterozygous sites on the Z chromosome in any female and pooled samples. Additionally, we calculated both the genome-wide autosomal heterozygosity rate and the heterozygosity rates in 100 kb non-overlapping windows using Vcftools v.0.1.16 and Bedtools v.2.31.1 (Quinlan & Hall 2010).

Phylogenetic analysis of concatenated whole-genome SNV

Based on the high-quality SNVs obtained above, we first constructed species trees using the connected SNVs. Three main approaches were applied: (1) maximum likelihood (ML) methods implemented in RAxML-NG v.1.2.2 (Kozlov et al., 2019) and IQTREE v.2.3.6 (Minh et al., 2020); (2) site-based coalescent methods using quartet topology frequencies, implemented in CASTER v.1.19.1.4 (Zhang et al., 2025a) and PAUP v.4.0 (SVDquartets) (Chifman & Kubatko 2014); and (3) the Bayesian multispecies coalescent method SNAPP, implemented in BEAST v.2.6.3 (Bryant et al., 2012).

In RAxML-NG, the GTR+ Γ nucleotide substitution model was used, with 1 000 bootstrap replicates. In IQTREE, we used the GTR+ASC model to correct for SNP ascertainment bias, also performing 1 000 bootstrap replicates. In CASTER, species trees were constructed based on individual sites patterns (caster-site) and pairwise site patterns (caster-pair), respectively.

For methods requiring unlinked loci (SVDquartets and SNAPP), we used VCFtools v.0.1.16 to thin the dataset, ensuring a minimum inter-SNP distance of 1 000 bp. In SVDquartets, all possible quartets were evaluated, and 1000 bootstrap replicates were performed. SNAPP was run for 10 million generations with sampling every 500 generations.



Zoological
Research

Convergence was assessed using Tracer v.1.7 (Rambaut et al., 2018), in which the first 25% of trees were discarded as burn-in and all parameters were confirmed to have effective sample sizes greater than 200. Posterior SNAPP trees were visualized in DensiTree v.2.2.7 (Bouckaert 2010).

Phylogenetic reconstruction of window-based strategy

We further performed species tree inference using a window-based strategy. First, we constructed genome fragment (GF) trees along the autosomes and Z chromosome with 50 kb non-overlapping windows using `genomics_general` (https://github.com/simonhmartin/genomics_general). For each GF tree, a ML tree was inferred under the GTR+ Γ substitution model using RAXML-NG v.1.2.2, with 100 bootstrap replicates. The species tree was then obtained from the GF trees using Astral-hybrid v.1.19.3.7 (Zhang & Mirarab 2022), based on node support weights under default parameters.

To examine phylogenetic discordance across the genome, TWISST (Martin & Van Belleghem 2017) was used to compute the proportions and genomic distribution of alternative topologies across all GF trees. The GF trees were visualized using DensiTree v.2.2.7, with a randomly selected subset of 1 500 trees.

Phylogenetic analysis of NRW and mitogenome

To examine the phylogeny of NRW, we downloaded the NRW sequences of *Ficedula albicollis* (GCA_900067835.1) (Smeds et al., 2015) and selected all female samples and pooled samples (Supplementary Table S1) to obtain NRW SNVs using the same method as described previously. To exclude potential effects of mapping errors and homologous sequences from the Z chromosome, we retained only reads with a mapping quality of 60 in BWA mapping and kept only SNV sites that were homozygous across all samples during the soft filtering. We then



constructed a ML tree for NRW using RAxML-NG v.1.2.2 under the GTR+ Γ substitution model, with 1 000 bootstrap replicates.

For the mitogenome, we first assembled the 13 protein-coding genes (PCGs) of seven species using MitoFinder v.1.4.1 (Allio et al., 2020) and then aligned them with MAFFT v.7.313 (Kato & Standley 2013). The ML tree of the 13 PCGs was subsequently constructed using RAxML-NG v.1.2.2 under the GTR+ Γ substitution model, with 1 000 bootstrap replicates.

Hybridization detection

We first calculated the frequencies of three possible topologies for each node from genome-wide GF trees using DiscoVista (Zhang et al., 2018). The proportions of the two alternative topologies (discordant with the species tree) were used to infer whether gene tree discordance resulted from ILS or introgression: similar high frequencies suggest ILS, while significant imbalance supports introgression. We then applied QuIBL to further evaluate introgression/ILS levels using branch length information from GF trees (Edelman et al., 2019). This method fits branch length distributions to two models—an “ILS-only” model (single peak at short branches) and an “ILS+introgression” model (additional peak at longer branches)—to distinguish between the two processes. Due to its sensitivity to recombination, we extracted a window every 500 kb to ensure the independence between windows for analysis. Each species triplet was analyzed under both the ‘ILS-only’ and ‘ILS+introgression’ models. Following recommendations, introgression is considered only when the Bayesian information criterion (BIC) of the ‘ILS+introgression’ model is at least 10 units smaller than the BIC of the ‘ILS-only’ model; otherwise, only ILS is present. The results reflecting introgression and/or ILS between species pairs were averaged across all relevant triplets, considering only two alternative topologies.



Zoological
Research

Given their relatively close genetic relationship (Zhao et al., 2023; Jiang et al., 2025), we further assessed introgression using a site-based approach. First, we calculated Patterson's D statistics (ABBA-BABA test) and the f_4 -ratio using Dsuite v.0.5 (The Heliconius Genome Consortium 2012, Patterson et al., 2012; Malinsky et al., 2021). Both D and the f_4 -ratio assume a tree topology typically given as (((P1, P2), P3), O), where the outgroup (O) helps polarize the ancestral allele from the derived allele. The D determines which of P1 or P2 shares significantly more derived alleles with P3, thereby identifying between which two species introgression has occurred. The f_4 -ratio quantifies the proportion of introgression by comparing the observed difference in derived alleles shared between P1 and P2 with P3 to the theoretical maximum difference expected under complete introgression. We fixed *O. deserti* as the outgroup and computed D/f_4 -ratio for all possible combinations based on the species tree topology. We applied a standard block-jackknife procedure to assess the statistical significance of the D statistics and f_4 -ratio by dividing the dataset into 200 Jackknife blocks. Only combinations with a Z-score > 3 were retained, and the heatmap of D statistic and f_4 -ratio was generated using the Ruby script plot_d.rb (https://github.com/mmatschiner/tutorials/tree/master/analysis_of_introgression_with_snp_data). Since different combinations were non-independent due to shared branches on the phylogenetic tree, this may result in highly correlated values when ancient gene flow is present. We therefore utilized the Fbranch module in Dsuite to calculate branch-specific estimates of gene flow (f -branch statistics, f_b), which can indicate introgression between extant species and ancestral lineages. The f_b statistic values for each branch were calculated from f_4 -ratio values that passed the significance threshold ($Z > 3$), and were visualized using the dtools.py script.

We inferred the phylogenetic network using SNaQ implemented in the Julia package PhyloNetworks v.0.16.2 (Solís-Lemus et al., 2017), based on the same set of GF trees used for QuIBL analysis. The search allowed for up to four hybridization events ($h=4$), initialized with



the species tree as the starting topology. For subsequent optimizations, we used the network estimated in the previous step as the starting network. The optimal number of reticulation events was determined by identifying the inflection point showing the steepest improvement in log-pseudolikelihood scores. The optimal number of reticulation events was determined by identifying the inflection point showing the steepest improvement in log-pseudolikelihood scores. Following model selection, we ran 100 bootstrap replicates based on the 100 bootstrap trees generated from each GF tree, employing default settings. The phylogenetic network is also inferred using the ‘InferNetwork_MPL’ command in PhyloNet v.3.8.2 (Yu et al., 2014), allowing for up to four reticulation events and starting from the species tree topology. For each reticulation setting, 100 independent searches are performed to avoid local optima. The optimal network is selected based on the highest log-pseudolikelihood score.

Gene flow was also inferred by utilizing allele frequency information from different species using TreeMix v.1.13 (Pickrell & Pritchard 2012). The same SNV dataset of unlinked loci as in the SANPP analysis was used, with outgroups further removed. First, a set of simulations was performed to define the best migration edges, where migration edges were set from 0 to 7, each with 10 parallel runs. The outputs were then analyzed using the R package OptM (Fitak 2021), and the optimal number of migration edges was selected based on the second-order rate of change in likelihood (Δm).

Introgression pattern and introgressed gene enrichment analysis

We investigated the genomic landscape of several hybridization events using the ABBA-BABA test with 50 kb non-overlapping windows, implemented in *genomics_general*. Four species combinations were used to examine four specific hybridization events, with *O. deserti* fixed as the outgroup: ((*S. torquatus*, *S. rubicola*) *S. maurus*) to examine introgression between *S. maurus* and *S. rubicola*; ((*S. torquatus*, *S. dactylotis*) *S. maurus*) for *S. maurus* and *S. dactylotis*; ((*S. torquatus*, *S. maurus*) *S. stejnegeri*) for *S. maurus* and *S. stejnegeri*; and ((*S. rubicola*, *S.*



Zoological
Research

dacotiae) *S. torquatus*) for *S. dacotiae* and *S. torquatus*. These combinations were chosen to capture strongest signals of introgression among all possible triads while minimizing potential gene flow between P1 and P2. Due to the sensitivity of the D statistic to genomic diversity in window-based analyses, we used f_d , a modified D statistic (Martin et al., 2015), as the introgression measure. Finally, we performed Spearman rank correlation analyses to evaluate the similarity of introgressed regions across species pairs. As negative f_d values are biologically meaningless, we retained only positive f_d values and truncated any values exceeding 1.

We further examined whether introgressed genes between each species pair were significantly enriched for specific functions. Based on the same combinations described above, we calculated the f_d values for 15 782 PCGs located on the autosomes and the Z chromosome for each species pair. Genes with f_d values greater than 0.2 were considered candidate introgressed genes and were subjected to Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses using clusterProfiler v.4.16.0 (Xu et al., 2024). The enrichment significance P values were adjusted for multiple testing using the Benjamini-Hochberg method to control the false discovery rate.

Test for mitonuclear co-introgression

Based on the observed mitonuclear discordance within the *S. torquatus* complex, we propose a hypothesis that this conflict arose from mitogenome capture by *S. maurus* from either the most recent common ancestor (MRCA) of *S. rubicola* and *S. dacotiae* or a sister species to this ancestor—which may represent a cryptic, unsampled lineage or an extinct ghost lineage (hereafter ‘unknown species’). Given the coevolution between mitochondrial and nuclear genomes (Shi et al., 2025), we predict that mitonuclear genes will exhibit co-introgression with the mitogenome due to selection pressures driven by mitonuclear incompatibilities—if mitogenome capture is present. Since the genetic material of the unknown species is unavailable, we use *S. rubicola* and *S. dacotiae* as proxies for the unknown species in our



Zoological
Research

analyses, as they share similar genetic components with it, although this approach may lead to false negatives.

We compiled a list of mitonuclear genes by consulting various resources on mitochondrial and mitonuclear dynamics, and prior similar work (Diodato et al., 2014; Greber & Ban 2016; Nikelski et al., 2023). These included only PCGs whose products directly interact with mtDNA or its immediate translation products. After removing genes that were not annotated or lacked specific genomic locations in the reference genome, 202 mitonuclear genes remained (Supplementary Table S2). The remaining 15 580 PCGs were classified as non-mitonuclear.

We used a gene-based approach to calculate the introgression signal strength for each gene and applied a threshold to categorize them as introgressed or non-introgressed. We employed chi-square and Fisher's exact tests to determine whether mitonuclear genes were over-representation among the introgressed genes. This analysis was repeated across different introgression indicators (D , f_d , and f_{dM}), thresholds levels (top 1%, 5%, 10%, 15%, 20%, and 25%), and species combinations (Supplementary Table S3). In all ABBA-BABA test configurations (((P1, P2), P3), Outgroup), *S. torquatus* was fixed as P1 and *S. maurus* as P3, while P2 and the outgroup varied.

Since gene-based methods may introduce computational randomness due to the short length of some genes, and because regulatory elements near genes could also co-introgress adaptively with mitonuclear genes, we repeated the analysis using a window-based approach. Specifically, we calculated introgression signal strength for each 50 kb non-overlapping window and used a threshold to classify the windows as either introgressed or non-introgressed. We assigned each gene to these windows by minimizing the absolute difference between the gene center location and the window center location, thereby categorizing genes as introgressed or non-introgressed. The remaining steps were consistent with the gene-based approach (Supplementary Table S4).



RESULTS

Sex determination and variant discovery

We generated ~33 Gb of data from whole-genome sequencing of *S. stejnegeri* (Supplementary Table S1). Combined with downloaded data from five *Saxicola* species and one outgroup, the total dataset reached ~668G (~87X mean depth). After quality filtering, ~17.27% (~115G) of the bases were trimmed due to low quality or adapter contamination. Mapping to the reference genome showed a high-quality alignment rate (97.4% to 99.08%), except for *S. torquatus* (59.1%). For this species, alignment of sequencing reads to the core nucleotide database indicated that among the top ten target species matched, 31.76% reads were mapped to bacteria, confirming that the low alignment rate was due to bacterial contamination (Supplementary Table S5), as noted by the data releasers (Van Doren et al., 2017). Because these bacterial reads cannot be mapped to the reference genome (indicated by a low mapping rate), they do not affect downstream analyses. Alignment coverage ranged from 95.9% to 99.4%, indicating high completeness. Sex identification based on the normalized Z-depth indicated that only *S. stejnegeri* was male (Supplementary Figure S2). Two pooled samples (*S. torquatus* and *S. dacotiae*) showed intermediate normalized Z-depth values (0.71 and 0.79, respectively), matching their sex ratios (male: female=18:33 and 38:18).

Following SNV filtering, a total of 45 784 604 high-quality SNVs were identified. Apart from *O. deserti* with 8 853 214 SNVs, the six *Saxicola* species each possess between 20 942 377 and 25 633 321 SNVs (Supplementary Table S1). Species-specific SNVs were most abundant in the outgroup (15 636 339) and *S. rubetra* (12 174 382), while the five *S. torquatus complex* species had 1 226 091 to 3 333 538. Genome-wide heterozygosity varied considerably among species, ranging from 0.017 to 0.198, which may partially reflect different demographic



Zoological
Research

histories (Supplementary Figure S3 and Table S1). Notably, the two pooled samples showed moderate-to-low heterozygosity (*S. dacotiae*=0.017, *S. torquatus*=0.058), with distributions similar to other species, indicating their suitability for downstream analysis as individual representatives.

Phylogenetic discordance across genome

All methods based on concatenated autosomal and Z-chromosomal SNVs (45 784 604 SNVs) and the window-based strategy (20 138 windows) yielded consistent species tree topologies with maximal nodal support (Figure 1A; Supplementary Figure S4). The topology revealed a nested structure: *S. dacotiae* and *S. rubicola* were sister taxa, followed sequentially by *S. torquatus*, *S. maurus*, *S. stejnegeri*, and *S. rubetra*. In contrast, the mitogenome phylogeny placed *S. maurus* as sister to *S. dacotiae* and *S. rubicola*, with *S. torquatus* outside, also with full support, indicating cytonuclear discordance. Interestingly, the NRW phylogeny (71 321 SNVs based on all female individuals) mirrored the mitogenome topology, suggesting co-inheritance of the mitogenome and NRW. However, the interbranch lengths in the NRW ML tree were relatively short, possibly due to excessive filtering of SNVs.

Phylogenies reconstructed across genomic windows revealed extensive phylogenetic discordance, failing to resolve the phylogeny as a simple bifurcating tree (Figure 1B, C). Among 20,138 GF trees, seven main topologies accounted for ~86.9% of the total. As expected, the topology concordant with the species tree inferred from autosomal and Z-chromosomal SNVs (blue topology; Figure 1C) was most frequent, comprising ~40.3% of trees. This topology was especially prevalent on the Z chromosome (~56.5%), exceeding its frequency on autosomes (~37.1%) by ~19.4 percentage points. In contrast, the topology consistent with the mitogenome and NRW tree (black topology; Figure 1C) was rare, representing only ~1.7% overall (~1.6% autosomes; ~2.3% Z chromosome).



Reticulate evolution within the *S. torquatus* complex

To elucidate the causes of phylogenetic discordance, we applied DiscoVista and QuIBL to distinguish between introgression and ILS, two major biological sources of discordance (Steenwyk et al., 2023). DiscoVista revealed substantial topological conflict at nodes 2 and 4, with one alternative quartet topology markedly more frequent, implicating introgression—specifically between *S. stejnegeri* and *S. maurus*, and *S. torquatus* and *S. dacotiae*—as the primary source of discordance (Figure 2A). QuIBL analysis of 2 028 GF trees from independent windows corroborated this inference, indicating that introgression accounted for most of discordance (proportions of 0.3 and 0.11 for the two introgression events), while ILS played a minor role, peaking at 0.08 between *S. torquatus* and *S. rubicola* (Figure 2B). D statistics and f_4 -ratio further supported these two introgression events and detected additional, moderate introgression between *S. maurus* and *S. rubicola*, and weak introgression between *S. maurus* and *S. dacotiae* (Figure 2C). Notably, the summarized f_4 -ratio on shared branches (f_b) revealed introgression between *S. maurus* and MRCA of *S. rubicola* and *S. dacotiae* (Figure 2D). This may reflect a hypothesis that the cytonuclear conflict of *S. maurus* may result from mitogenome capture from an unknown species (see Materials and Methods).

The SNaQ species network inference identified one optimal hybridization event (Supplementary Figure S5A), with 21.9% inheritance probability from *S. maurus* to *S. stejnegeri*. However, bootstrap support was weak: the hybrid node appeared in only 7 replicates, and the major and minor hybrid edges were supported in 68 and 7 replicates, respectively, indicating much uncertainty in introgression direction and placement. PhyloNet results showed the best fit with two reticulation events (Supplementary Figure S5B), revealing not only hybridization between *S. maurus* and *S. stejnegeri* but also between *S. maurus* and *S. rubicola*, with inheritance probabilities of 31.72% and 3.71%, respectively.



Zoological
Research

The OptM results indicated that $m=4$ was the best number of hybridization events (migration edges) in Treemix (Supplementary Figure S5C). Although Treemix did not recover the strongest introgression between *S. maurus* and *S. stejnegeri*, it revealed a notable migration edge from the MRCA of *S. rubicola* and *S. dacotiae* to *S. maurus*, which had the highest migration weight and was consistent with the results from f_b . Another migration edge points from *S. maurus* to *S. rubicola*, which was also detected in the D -statistic, f_4 -ratio, f_b , and PhyloNet results. However, the remaining two migration edges were not detected by other methods, including introgression from *S. torquatus* to *S. rubetra* and from *S. rubetra* to *S. rubicola*.

Introgression pattern and mitonuclear co-introgression test

A genome-wide scan based on f_d using 50 kb non-overlapping windows for four detected hybridization events—between *S. maurus* and *S. rubicola*, *S. maurus* and *S. dacotiae*, *S. maurus* and *S. stejnegeri*, and *S. dacotiae* and *S. torquatus*—revealed distinct introgression landscapes; however, all were characterized by clustered introgressed regions (introgression islands), primarily concentrated at the distal ends of smaller chromosomes. The *S. maurus*–*S. stejnegeri* pair exhibited the most introgression islands, followed by *S. dacotiae*–*S. torquatus* and *S. maurus*–*S. rubicola*, while *S. maurus*–*S. dacotiae* showed few islands. Pairwise correlation analyses revealed positive correlations ($P<0.01$) for all pairwise comparisons, except between *S. maurus*–*S. dacotiae* and *S. dacotiae*–*S. torquatus*, which was not statistically significant ($P=0.35$). Notably, the highest correlation coefficient ($\rho=0.44$) was found between *S. maurus*–*S. dacotiae* and *S. maurus*–*S. rubicola*, further suggesting that introgression between *S. maurus* and *S. rubicola*/*S. dacotiae* may have occurred in the MRCA of the latter two species.

Using a threshold of $f_d=0.2$, 894–2 275 (approximately 4.7%–12.0%) introgressed genes were identified across four species pairs. GO enrichment analysis revealed significant



Zoological
Research

functional enrichments in three species pairs (P -adjust<0.05): 1 323 genes between *S. maurus* and *S. rubicola* were enriched in phosphatidylethanolamine metabolic process (GO:0046337); 2 275 genes between *S. maurus* and *S. stejnegeri* in intermediate filament organization (GO:0045109); and 2 093 genes between *S. dacotiae* and *S. torquatus* in ADP-ribosyltransferase (GO:0106274, GO:1990404, GO:0003953) and pentosyltransferase activity (GO:0016763) (Supplementary Figure S6). KEGG analysis showed significant enrichment only for *S. maurus*–*S. stejnegeri* genes in metabolism of xenobiotics by cytochrome P450 (KO00980), while other pairs exhibited no significant enrichment (P -adjust>0.05) (Supplementary Figure S7). Notably, in the GO enrichment results for *S. maurus*–*S. rubicola* introgressed genes, there is a term mitochondrial translation (GO:0032543) ranked third by adjusted P value, though statistically non-significant (P -adjust=0.26). This term encompasses mitochondrial ribosomal (e.g., *MRPS16*, *MRPL23*), tRNA-related (e.g., *SARS2*, *TRMT10C*), and translational regulator (e.g., *MTIF3*, *TSFM*) genes.

Using various methods for introgression gene detection, along with different introgression indicator, thresholds and species combinations, nearly half of the mitonuclear co-introgression tests yielded statistically significant results (P <0.05; Supplementary Tables S3, S4). Considering that *S. rubicola* and *S. dacotiae* may not fully represent their ancestors or its sister lineage, these results may support our hypothesis that the cytonuclear conflict of *S. maurus* originates from mitogenome capture from the MRCA of *S. rubicola* and *S. dacotiae* (or its sister species), accompanied by mitonuclear co-introgression. Notably, majority of significant results were observed when *O. deserti* was used as the outgroup, whereas tests using *S. stejnegeri* or *S. rubetra* as the outgroup were largely non-significant. This may be because these two species are too closely related to the target species (or have hybridized with it) and thus cannot reliably represent ancestral alleles.



Zoological
Research

DISCUSSION

Maternal co-inheritance and discordance with species tree

Previous phylogenies based on single mitochondrial genes suggested that *S. maurus* is sister to *S. dacotiae* and *S. rubicola*, relative to *S. torquatus* (Zink et al., 2009). Our mitogenome-based phylogeny further supports this relationship. We also found that the NRW tree aligns with the mitogenome tree but diverges from the species tree, indicating maternal coinheritance. Similar cases have been reported in other avian lineages, such as *Ficedula* (Smeds et al., 2015), *Anas* (Gu et al., 2023), and *Emberiza* (Zhang et al., 2024). Though paternal leakage of mtDNA or recombination in either molecule could potentially weaken the evolutionary association between the NRW and the mitogenome (White et al., 2008). Nevertheless, our findings stress that the NRW and mitogenome exhibit congruent lineage characteristics, and their evolutionary stability in stonechat aligns with mainstream opinions (Ellegren 2007).

One possible perspective is that the mitogenome and NRW represent the true species tree, as these two molecular markers do not recombine and can therefore reflect their true evolutionary history (Moore 1995; Mackiewicz et al., 2019). In contrast, nuclear markers may have undergone genome homogenization due to introgression between *S. torquatus* and *S. rubicola*/*S. dacotiae*, or between *S. maurus* and *S. stejnegeri*, leading to shared genetic material between corresponding species pairs. This would bring the species pairs closer together on phylogenetic trees based on nuclear SNVs. However, this scenario should result in substantial topological conflicts at corresponding nodes (Steenwyk et al., 2023). Our result rejects the mitogenome and NRW tree as the species tree (Figure 2A). Similarly, the very low proportion



Zoological
Research

(~1.7%) of the topology consistent with the mitogenome/NRW tree in the GF trees further rejects this possibility as the species tree (Figure 1C).

We found eight hypotheses that could explain the incongruence between mtDNA/NRW tree and species tree (Figure 4A–H). Considering the introgression and mitonuclear co-introgression signals detected between *S. maurus* and MRCA of *S. rubicola* and *S. dacotiae*, as well as their proximity in the breeding range, the incongruence may result from the mitogenome/NRW of *S. maurus* being captured from the MRCA of *S. dacotiae* and *S. rubicola* (Figure 4A). However, due to incomplete sampling within the breeding ranges of *S. dacotiae* and *S. rubicola*, several alternative capture scenarios cannot be excluded (Figure 4B–E). Except for ILS-generated lineages (Figure 4E), other scenarios (Figure 4B–D) could yield observed introgression and mitonuclear co-introgression signals between *S. maurus* and MRCA of *S. rubicola* and *S. dacotiae*, as the other mitogenome donor lineages share genetic similarity with this MRCA. However, the fourth hypothesis (Figure 4D) is less probable. This is because female-specific selection strong enough to generate additional maternal lineages without accompanying speciation (such as through nest parasitism) (Fossøy et al., 2016) has not been documented in stonechats. Moreover, a high level of male-biased dispersal, which could also lead to the emergence of additional maternal lineages, is unlikely in this species. Stonechats, like many other songbirds, typically display female-biased dispersal (Clarke et al., 1997), further reducing the plausibility of this scenario. The fifth and sixth hypothesis (Figure 5E, F) is also improbable given low detected ILS signals (Figure 2B) and long, well-supported intermediate branches in mitogenome ML trees (Figure 1A), which are inconsistent with ILS expectations (Avisé & Robinson 2008; Rivas-González et al., 2023). The two latest hypotheses (Figure 4G, H) are also less probable, as no introgression signals exist between *S. torquatus* and lineages genetically similar to this MRCA (e.g., *S. stejnegeri*/*S. maurus*). Therefore, overall, compared to the other hypotheses, the first three hypotheses (Figure 4A–C) are the most likely,



Zoological
Research

i.e., the mitogenome/NRW of *S. maurus* may have been captured from lineage related to the MRCA of *S. dacotiae* and *S. rubicola*. Future population genomic analysis with extensive sampling in breeding areas will further validate this and may distinguish the second hypothesis from the other two.

Plausibility of introgression within the *S. torquatus* complex

Species undergoing rapid radiation are prone to hybridization due to frequent geographic overlap with closely related taxa (Edelman et al., 2019). Although we selected only one individual per species except for two pooled samples, which may limit the representation of each species and underestimate introgressive diversity since introgressed fragments may not have become fixed in the populations, we still preliminarily observed extensive hybridization among several species within the *S. torquatus* complex. These findings are plausible when considering the distribution of their breeding ranges. For example, *S. maurus* and *S. stejnegeri* exhibit parapatric distribution along the Yenisei/Angara River (Figure 5), where hybrids between them have been reported (Hellström & Norevik 2014; Hellström & Waern 2023). Similarly, *S. dacotiae* and *S. torquatus* may come into contact in northwestern Africa, as *S. torquatus* has scattered populations in the northern part of the continent. Additionally, *S. maurus* could have hybridized with the MRCA of *S. dacotiae* and *S. rubicola* when considering that *S. dacotiae* and *S. rubicola* may gradually colonize Europe from West Asia (Illera et al., 2008).

Moreover, we identified hybridization between *S. maurus* and the MRCA of *S. rubicola* and *S. dacotiae* via f_b , which was further supported by the high correlation of introgressed regions between *S. maurus*–*S. rubicola* and *S. maurus*–*S. dacotiae*. However, there was a notable discrepancy: a strong introgression signal between *S. maurus* and *S. rubicola*, but a weak signal between *S. maurus* and *S. dacotiae*. This discrepancy may be due to additional introgression between *S. maurus* and *S. rubicola*. Another possibility is a limitation of the



Zoological
Research

ABBA-BABA test (The Heliconius Genome Consortium 2012): since introgression between *S. torquatus* and *S. dacotiae* was detected, using the (((*S. torquatus*, *S. dacotiae*), *S. maurus*), Outgroup) combination to test for introgression between *S. dacotiae* and *S. maurus* might have obscured the signal due to introgression between *S. torquatus* and *S. dacotiae*. The third potential reason is that the island species *S. dacotiae*, with a small population size, has accumulated more derived mutations due to stronger genetic drift, which may lead to the loss of ancestral introgression signals. This is supported by the relatively long terminal branch lengths in genome-wide SNVs and mitogenome ML trees (Figure 1A), and previous studies suggest that differentiation of *S. dacotiae* from other *Saxicola* species has largely been driven by genetic drift (Van Doren et al., 2017). Similarly, in the Phylonet species network, one of the hybridization events, where introgression between *S. maurus* and *S. rubicola*, rather than between *S. maurus* and the MRCA of *S. rubicola* and *S. dacotiae*, could also be influenced by these factors. Furthermore, Treemix identified hybridization events involving *S. rubetra* with *S. torquatus* and *S. rubicola*. As these events were not detected by other methods, they may represent spurious introgression signals due to sampling gaps of intermediate evolutionary lineages or larger fluctuations in allele frequencies due to small sample sizes per species (Pickrell & Pritchard 2012), despite past reports of hybridization between *S. rubetra* and *S. torquatus* complex species (McCarthy 2006).

Drivers of genomic introgression landscapes

The genomic landscape of introgression is typically influenced by the interplay between recombination rates and selection against genomic incompatibilities (Martin et al., 2019; Duranton & Pool 2022; Veller et al., 2023). When selection against genomic incompatibilities is the dominant force, introgression is expected to exhibit a positive correlation with recombination rate (Martin et al., 2019; Edelman et al., 2019); conversely, if foreign ancestry is favored, a negative correlation is anticipated (Duranton & Pool 2022).



Zoological
Research

We found that strong regions tend to be concentrated in island-like forms at the ends of smaller chromosomes, likely due to the higher recombination rates in these regions (Ellegren 2013; Zhang et al., 2014; Guerrero-Cózar et al., 2021; Leitwein et al., 2024), although the lack of high-resolution recombination maps limits further validation. A similar pattern has been observed in other groups. For example, in Princess cichlids, introgressed segments are significantly enriched at the chromosome periphery, consistent with their higher recombination rates compared to the central regions (Gante et al., 2016; Leitwein et al., 2024). Similarly, in Amazonian antbird (*Thamnophilus aethiops*) (Musher et al., 2024) and Swainson's thrushes (*Catharus ustulatus*) (Blain et al., 2025), introgression signals are mainly concentrated on smaller chromosomes, which aligns with their higher recombination rates relative to large chromosomes (Ellegren 2013; Zhang et al., 2014). Furthermore, even when accounting for shared ancestral introgression, the strength of introgression between species pairs remained consistently positively correlated. This further supports the aforementioned hypothesis, as introgressed genes show no consistent functional enrichment across species pairs, suggesting that parallel selection on particular functional candidate genes is unlikely to be the primary driver. Instead, genomic architectures such as recombination hotspots, which are both interspecific conserved (Kawakami et al., 2017) and prone to introgression (Edelman et al., 2019), could underlie this pattern. Although such regions might also experience increased ILS, we found that ILS contributes minimally during divergence of these species compared to introgression. Thus, conserved genomic architectures (e.g., shared recombination hotspots) may directly shape the genomic landscape of these taxa by influencing introgression, independent of adaptive contexts.

This enrichment pattern of introgression in regions with potentially higher recombination rates also suggests that in the *Saxicola* clade, foreign ancestry may have been subject to strong purifying selection due to genetic incompatibilities, consistent with general observations



(Presgraves 2010; Moran et al., 2021). Nonetheless, evidence of functional enrichments indicates a role for adaptive introgression. For example, phosphatidylethanolamine metabolic genes introgressed between *S. maurus* and *S. rubicola* may have optimized cell membrane function, potentially enhancing neural signaling (e.g., migratory navigation or courtship behavior), membrane stability under temperature or altitude stress, or energy metabolism via improved mitochondrial function (Calzada et al., 2016). In the *S. maurus*–*S. stejnegeri* pair, introgressed genes associated with detoxification and intermediate filament organization may relate to metabolism of toxins, feather development, skin toughness, or muscle differentiation, likely improving resilience to chemical and mechanical stressors in shared environments (Nebert et al., 2006; Toivola et al., 2010). Similarly, transferase-related genes introgressed between *S. dacotiae* and *S. torquatus* could facilitate DNA repair, cellular signaling, or metabolic adaptations, supporting responses to pathogens, environmental stress, or cellular homeostasis (Gupte et al., 2017). These findings collectively suggest that introgression provided a source of beneficial alleles for local adaptation to specific ecological challenges. Therefore, while not excluding other factors, the introgression landscape in several *Saxicola* species is likely shaped by the interplay between pervasive purifying selection against genomic incompatibilities and weak spatially heterogeneous adaptive introgression.

Taxonomic and biogeographic implications

Due to the lack of genomic data for *S. stejnegeri*, its taxonomic status has long been controversial, with some considering it a subspecies of *S. maurus* (Chesser et al., 2022; Gill et al., 2024). Our phylogenomic analysis positions *S. stejnegeri* as the most basal lineage within the *S. torquatus* complex, whereas *S. maurus* is sister to all other members. Combined with acoustic evidence (Opaev et al., 2018), we support recognizing *S. stejnegeri* as a distinct species. Additionally, the new phylogenetic relationship, where *S. stejnegeri* and *S. maurus* are sequentially nested at the root, further reinforces the Asian origin of this complex (Figure 5)



Zoological
Research

(Illera et al., 2008; Jiang et al., 2025). Although the precise ancestral ranges of *S. stejnegeri* and *S. maurus* in Siberia remain uncertain, European and African populations likely originated from later colonization by *S. maurus* ancestors. Furthermore, genomic evidence of introgression between the ancestral European lineage (MRCA of *S. rubicola* and *S. dacotiae*) and *S. maurus* favors an Asian (rather than African) origin for European population, followed by Canary Islands colonization. In contrast, African populations may have been colonized either through Europe or via a separate dispersal directly from Asia, followed by colonization to Réunion Island (Illera et al., 2008; Jiang et al., 2025).

CONCLUSION

Our genome-wide analyses encompassed all species-level lineages within the *S. torquatus* complex, except *S. tectes*. We recovered robust phylogeny and revealed widespread phylogenetic discordance across genomic regions, primarily driven by introgression rather than ILS, showing reticulate evolutionary history. Cytonuclear discordance likely resulted from introgression and reflects maternal co-inheritance of mitogenome and the NRW. The Z chromosome exhibited greater phylogenetic stability, being less affected by ILS/introgression, showing its robustness in species tree reconstruction. Introgressed regions were concentrated at the distal ends of smaller chromosomes and correlated across species pairs, likely attributed to shared recombination hotspots in these regions. The revised phylogeny within the *S. torquatus* complex supports the recognition of *S. stejnegeri* as a distinct species and provides a more parsimonious colonization scenario for this complex. Future, more comprehensive sampling of lineages within the *S. torquatus* complex will offer deeper insights into the evolution and colonization history of this group.



DATA AVAILABILITY

The raw sequencing data of *Saxicola stejnegeri* generated in this study have been uploaded to NCBI under BioProject ID: PRJNA1331738.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

C.J. and B.L. designed the research and collected the samples. C.J. performed the experiments and data analysis. W. Z. and Z. J. contributed to data analysis. C. J., Y. L., and W. Z. wrote the manuscript. N.M. and B.L. revised the manuscript and improved the language. All authors have read and agreed to the final version of the manuscript.



ACKNOWLEDGMENTS

We are grateful to Wei Zhang, Yanchun Xu, and Suying Bai for their guidance and logistical support during sample collection. We thank the five anonymous reviewers for their valuable suggestions on this manuscript.

ETHICAL STATEMENT

The bird sample was obtained from carcass related to the case, so animal ethics approval was not necessary.

REFERENCE

Allio R, Schomaker-Bastos A, Romiguier J, et al. 2020. MitoFinder: efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. *Molecular Ecology Resources* **20**(4):892–905.

AviList Core Team. 2025. AviList: The Global Avian Checklist, v2025.

Avise JC, and Robinson TJ. 2008. Hemiplasy: a new term in the lexicon of phylogenetics. *Systematic Biology* **57**(3):503–507.

Baraf LM, Hung JY, and Cowman PF. 2025. Phylogenomics of marine angelfishes: diagnosing sources of systematic discordance for an iconic reef fish family (F: Pomacanthidae). *Systematic Biology*:syaf016.

Bergeron LA, Besenbacher S, Zheng J, et al. 2023. Evolution of the germline mutation rate across vertebrates. *Nature* **615**(7951):285–291.



Zoological
Research

Birds of the World (2025) Birds of the World. In: Billerman SM, Keeney BK, Kirwan GM, Medrano F, Sly ND, Smith MG (Eds) Birds of the World. Cornell Laboratory of Ornithology, Ithaca.

Blain SA, Justen HC, Langdon QK, et al. 2025. Repeatable selection on large ancestry blocks in an avian hybrid zone. *Molecular Biology and Evolution* **42**(3):msaf044.

Bouckaert RR. 2010. DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics* **26**(10):1372–1373.

Bryant D, Bouckaert R, Felsenstein J, et al. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution* **29**(8):1917–1932.

Calzada E, Onguka O and Claypool SM. 2016. Phosphatidylethanolamine metabolism in health and disease. *International Review of Cell and Molecular Biology* **321**:29-88.

Chase MA, Vilcot M, and Mugal CF. 2024. Evidence that genetic drift not adaptation drives fast-Z and large-Z effects in *Ficedula* flycatchers. *Molecular Ecology* **34**(18):e17262.

Chen S. 2023. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta* **2**(2):e107.

Chesser RT, Billerman SM, Burns KJ, et al. 2022. Sixty-third supplement to the American Ornithological Society's *Check-list of North American Birds*. *Ornithology* **139**(3):ukac020.

Chifman J, and Kubatko L. 2014. Quartet inference from SNP data under the coalescent model. *Bioinformatics* **30**(23):3317–3324.

Clarke AL, Sæther B-E, and Røskoft E. 1997. Sex biases in avian dispersal: a reappraisal. *Oikos* **79**(3):429–438.

Danecek P, Auton A, Abecasis G, et al. 2011. The variant call format and VCFtools. *Bioinformatics* **27**(15):2156–2158.



- Danecek P, Bonfield JK, Liddle J, et al. 2021. Twelve years of SAMtools and BCFtools. *GigaScience* **10**(2):giab008.
- Diodato D, Ghezzi D, and Tiranti V. 2014. The mitochondrial aminoacyl tRNA synthetases: genes and syndromes. *International Journal of Cell Biology* **2014**(1):787956.
- Duranton M, and Pool JE. 2022. Interactions between natural selection and recombination shape the genomic landscape of introgression. *Molecular Biology and Evolution* **39**(7):msac122.
- Edelman NB, Frandsen PB, Miyagi M, et al. 2019. Genomic architecture and introgression shape a butterfly radiation. *Science* **366**: 594–599.
- Ellegren H. 2007. Molecular evolutionary genomics of birds. *Cytogenetic and Genome Research* **117**(1–4):120–130.
- Ellegren H. 2013. The evolutionary genomics of birds. *Annual Review of Ecology, Evolution, and Systematics* **44**:239–259.
- Feng S, Bai M, Rivas-González I, et al. 2022. Incomplete lineage sorting and phenotypic evolution in marsupials. *Cell* **185**(10):1646-1660.e18.
- Fitak RR. 2021. OptM: estimating the optimal number of migration edges on population trees using Treemix. *Biology Methods and Protocols* **6**(1):bpab017.
- Fitzpatrick BM. 2004. Rates of evolution of hybrid inviability in birds and mammals. *Evolution* **58**(8):1865–1870.
- Fossøy F, Sorenson MD, Liang W, et al. 2016. Ancient origin and maternal inheritance of blue cuckoo eggs. *Nature Communications* **7**(1):1–6.
- Fu R, Zhu Y, Liu Y, et al. 2022. Genome-wide analyses of introgression between two sympatric Asian oak species. *Nature Ecology & Evolution* **6**(7):924-935.



Zoological
Research

- Gante HF, Matschiner M, Malmstrøm M, et al. 2016. Genomics of speciation and introgression in Princess cichlid fishes from Lake Tanganyika. *Molecular Ecology* **25**(24):6143–6161.
- Gill F, Donsker D, and Rasmussen P. 2025. IOC World Bird List (v15.1).
- Greber BJ, and Ban N. 2016. Structure and function of the mitochondrial ribosome. *Annual Review of Biochemistry* **85**:103–132.
- Gu H, Wen J, Zhao X, et al. 2023. Evolution, inheritance, and strata formation of the W chromosome in Duck (*Anas platyrhynchos*). *Genome Biology and Evolution* **15**(11):evad183.
- Guerrero-Cózar I, Gomez-Garrido J, Berbel C, et al. 2021. Chromosome anchoring in Senegalese sole (*Solea senegalensis*) reveals sex-associated markers and genome rearrangements in flatfish. *Scientific Reports* **11**(1):13460.
- Gupte R, Liu Z and Kraus WL. 2017. PARPs and ADP-ribosylation: recent advances linking molecular functions to biological outcomes. *Genes & Development* **31**(2):101-126.
- Hellström M, and Norevik G. 2014. The uppertail-covert pattern of ‘Stejneger’s Stonechat.’ *British Birds* **107**(11):692–700.
- Hellström M, and Waern M. 2023. Identification of extralimital Siberian and Amur Stonechats. *British Birds* **116**(1):41–53.
- Hill GE. 2019. Reconciling the mitonuclear compatibility species concept with rampant mitochondrial introgression. *Integrative and Comparative Biology* **59**(4):912–924.
- Hu L, Long J, Lin Y, et al. 2022. Arctic introgression and chromatin regulation facilitated rapid Qinghai-Tibet Plateau colonization by an avian predator. *Nature Communications* **13**(1):6413.
- Illera JC, Richardson DS, Helm B, et al. 2008. Phylogenetic relationships, biogeography and speciation in the avian genus *Saxicola*. *Molecular Phylogenetics and Evolution* **48**(3):1145–1154.



Zoological
Research

- Jiang C, Zhu W, Wu S, et al. 2025. Comparative mitogenomics, phylogeny, and biogeography of selected species of *Saxicola* (Aves, Passeriformes). *ZooKeys* **1249**:69–92.
- Jiang Z, Zang W, Ericson PGP, et al. 2024. Gene flow and an anomaly zone complicate phylogenomic inference in a rapidly radiated avian family (Prunellidae). *BMC Biology* **22**(1):49.
- Jung Y, and Han D. 2022. BWA-MEME: BWA-MEM emulated with a machine learning approach. *Bioinformatics* **38**(9):2404–2413.
- Katoh K, and Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**(4):772–780.
- Kawakami T, Mugal CF, Suh A, et al. 2017. Whole-genome patterns of linkage disequilibrium across flycatcher populations clarify the causes and consequences of fine-scale recombination rate variation in birds. *Molecular Ecology* **26**(16):4158–4172.
- Kozlov AM, Darriba D, Flouri T, et al. 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics* **35**(21):4453–4455.
- Leitwein M, Durif G, Delpuech E, et al. 2024. The fate of a polygenic phenotype within the genomic landscapes of introgression in the European seabass hybrid zone. *Molecular Biology and Evolution* **41**(9):msae194.
- Loskot VM, and Bakhtadze GB. 2020. Distribution, systematics and nomenclature of the three taxa of Common Stonechats (Aves, Passeriformes, Muscicapidae, *Saxicola*) that breed in the Caucasian region. *Zoosystematica Rossica* **29**(1):33–57.
- Mackiewicz P, Urantówka AD, Krocak A, et al. 2019. Resolving phylogenetic relationships within Passeriformes based on mitochondrial genes and inferring the evolution of their mitogenomes in terms of duplications. *Genome Biology and Evolution* **11**(10):2824–2849.



Zoological
Research

- Malinsky M, Matschiner M, and Svardal H. 2021. Dsuite - Fast *D*-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources* **21**(2):584–595.
- Martin SH, Davey JW, and Jiggins CD. 2015. Evaluating the use of ABBA–BABA statistics to locate introgressed loci. *Molecular Biology and Evolution* **32**(1):244–257.
- Martin SH, Davey JW, Salazar C, et al. 2019. Recombination rate variation shapes barriers to introgression across butterfly genomes. *PLOS Biology* **17**(2):e2006288.
- Martin SH, and Jiggins CD. 2017. Interpreting the genomic landscape of introgression. *Current Opinion in Genetics & Development* **47**:69–74.
- Martin SH, and Van Belleghem SM. 2017. Exploring evolutionary relationships across the genome using topology weighting. *Genetics* **206**(1):429–438.
- McCarthy EM. 2006. Handbook of avian hybrids of the world. Oxford: Oxford University Press.
- McKenna A, Hanna M, Banks E, et al. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* **20**(9):1297–1303.
- Meier JJ, Stelkens RB, Joyce DA, et al. 2019. The coincidence of ecological opportunity with hybridization explains rapid adaptive radiation in Lake Mweru cichlid fishes. *Nature Communications* **10**(1):5391.
- Minh BQ, Schmidt HA, Chernomor O, et al. 2020. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution* **37**(5):1530–1534.
- Moore WS. 1995. Inferring phylogenies from mtDNA variation: mitochondrial - gene trees versus nuclear - gene trees. *Evolution; international journal of organic evolution* **49**(4):718–726.



Zoological
Research

- Moran BM, Payne C, Langdon Q, et al. 2021. The genomic consequences of hybridization. *eLife* **10**:e69016.
- Musher LJ, Del-Rio G, Marcondes RS, et al. 2024. Geogenomic predictors of genetree heterogeneity explain phylogeographic and introgression history: a case study in an amazonian bird (*Thamnophilus aethiops*). *Systematic Biology* **73**(1):36–52.
- Myers EA, Rautsaw RM, Borja M, et al. 2024. Phylogenomic discordance is driven by widespread introgression and incomplete lineage sorting during rapid species diversification within rattlesnakes (Viperidae: *Crotalus* and *Sistrurus*). *Systematic Biology* **73**(4):722–741.
- Nebert DW and Dalton TP. 2006. The role of cytochrome P450 enzymes in endogenous signalling pathways and environmental carcinogenesis. *Nature Reviews Cancer* **6**(12):947–960.
- Nikelski E, Rubtsov AS, and Irwin D. 2023. High heterogeneity in genomic differentiation between phenotypically divergent songbirds: a test of mitonuclear co-introgression. *Heredity* **130**(1):1–13.
- Opaev A, Red'kin Y, Kalinin E, et al. 2018. Species limits in Northern Eurasian taxa of the common stonechats, *Saxicola torquatus* complex (Aves: Passeriformes, Muscicapidae). *Vertebrate Zoology* **68**(3):199–211.
- Ottenburghs J. 2023. How common is hybridization in birds? *Journal of Ornithology* **164**(4):913–920.
- Ottenburghs J, Kraus RHS, van Hooft P, et al. 2017. Avian introgression in the genomic era. *Avian Research* **8**(1):30.
- Ottenburghs J, Ydenberg RC, Van Hooft P, et al. 2015. The Avian Hybrids Project: gathering the scientific literature on avian hybridization. *Ibis* **157**(4):892–894.
- Patterson N, Moorjani P, Luo Y, et al. 2012. Ancient Admixture in Human History. *Genetics* **192**(3):1065–1093.



- Payseur BA, and Rieseberg LH. 2016. A genomic perspective on hybridization and speciation. *Molecular Ecology* **25**(11):2337–2360.
- Peona V, Palacios-Gimenez OM, Lutgen D, et al. 2023. An annotated chromosome-scale reference genome for Eastern black-eared wheatear (*Oenanthe melanoleuca*). *G3: Genes, Genomes, Genetics* **13**(6):jkad088.
- Pickrell J, and Pritchard J. 2012. Inference of population splits and mixtures from genome-wide allele frequency data. *Nature Precedings* **8**(11):e1002967.
- Presgraves DC. 2010. Darwin and the origin of interspecific genetic incompatibilities. *The American Naturalist* **176**(S1):S45–S60.
- Quinlan AR, and Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**(6):841–842.
- Rambaut A, Drummond AJ, Xie D, et al. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology* **67**(5):901–904.
- Ravinet M, Faria R, Butlin RK, et al. 2017. Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *Journal of Evolutionary Biology* **30**(8):1450–1477.
- Rivas-González I, Rousselle M, Li F, et al. 2023. Pervasive incomplete lineage sorting illuminates speciation and selection in primates. *Science* **380**(6648):eabn4409.
- Shi X, Ma C, Chen N, et al. 2025. Selection increases mitonuclear DNA discordance but reconciles incompatibility in African cattle. *Molecular Biology and Evolution* **42**(2):msaf039.
- Sloan DB, Havird JC, and Sharbrough J. 2017. The on-again, off-again relationship between mitochondrial genomes and species boundaries. *Molecular Ecology* **26**(8):2212–2236.
- Smeds L, Warmuth V, Bolivar P, et al. 2015. Evolutionary analysis of the female-specific avian W chromosome. *Nature Communications* **6**(1):7330.



Zoological
Research

- Solís-Lemus C, Bastide P, and Ané C. 2017. PhyloNetworks: A package for phylogenetic networks. *Molecular Biology and Evolution* **34**(12):3292–3298.
- Steenwyk JL, Li Y, Zhou X, et al. 2023. Incongruence in the phylogenomics era. *Nature Reviews Genetics* **24**(12):834–850.
- Svensson L, Shirihai H, Frahnert S, et al. 2012. Taxonomy and nomenclature of the Stonechat complex *Saxicola torquatus sensu lato* in the Caspian region. *Bulletin of the British Ornithologists' Club* **132**:260–269.
- Taylor EB, Boughman JW, Groenenboom M, et al. 2006. Speciation in reverse: morphological and genetic evidence of the collapse of a three-spined stickleback (*Gasterosteus aculeatus*) species pair. *Molecular Ecology* **15**(2):343–355.
- The Heliconius Genome Consortium. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**(7405):94–98.
- Toews DPL, and Brelsford A. 2012. The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology* **21**(16):3907–3930.
- Toivola DM, Strnad P, Habtezion A, et al. 2010. Intermediate filaments take the heat as stress proteins. *Trends in Cell Biology* **20**(2):79–91.
- Urquhart E. 2002. Stonechats: a guide to the genus *Saxicola*. London: Christopher Helm.
- Van Doren BM, Campagna L, Helm B, et al. 2017. Correlated patterns of genetic diversity and differentiation across an avian family. *Molecular Ecology* **26**(15):3982–3997.
- Veller C, Edelman NB, Muralidhar P, et al. 2023. Recombination and selection against introgressed DNA. *Evolution* **77**(4):1131–1144.
- Wang J, Zhou X, Dietrich CH, et al. 2025. Extensive hybridisation and complex evolutionary history in the leafhopper genus *Agnesiella* (Hemiptera: Cicadellidae: Typhlocybinae). *Systematic Entomology* **50**(4): 903–919.



Zoological
Research

- Webb WC, Marzluff JM, and Omland KE. 2011. Random interbreeding between cryptic lineages of the Common Raven: evidence for speciation in reverse. *Molecular Ecology* **20**(11):2390–2402.
- White DJ, Wolff JN, Pierson M, et al. 2008. Revealing the hidden complexities of mtDNA inheritance. *Molecular Ecology* **17**(23):4925–4942.
- Winkler DW, Billerman SM, and Lovette IJ. 2020. Old World Flycatchers (Muscicapidae), version 1.0. *Birds of the World*.
- Wittmann U, Heidrich P, Wink M, et al. 1995. Speciation in the stonechat (*Saxicola torquata*) inferred from nucleotide sequences of the cytochrome-b gene. *Journal of Zoological Systematics and Evolutionary Research* **33**(3–4):116–122.
- Xu S, Hu E, Cai Y, et al. 2024. Using clusterProfiler to characterize multiomics data. *Nature Protocols* **19**(11):3292–3320.
- Camacho C, Coulouris G, Avagyan V, et al. 2009. BLAST+: architecture and applications. *BMC Bioinformatics*, **10**(1): 421.
- Yu Y, Dong J, Liu KJ, et al. 2014. Maximum likelihood inference of reticulate evolutionary histories. *Proceedings of the National Academy of Sciences* **111**(46):16448–16453.
- Zhang C, and Mirarab S. 2022. Weighting by gene tree uncertainty improves accuracy of quartet-based species trees. *Molecular Biology and Evolution* **39**(12):msac215.
- Zhang C, Nielsen R, and Mirarab S. 2025. CASTER: Direct species tree inference from whole-genome alignments. *Science* **387**(6737):eadk9688.
- Zhang C, Rabiee M, Sayyari E, et al. 2018. ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics* **19**(6):153.
- Zhang D, She H, Wang S, et al. 2024. Phylogenetic conflict between species tree and maternally inherited gene trees in a clade of *Emberiza* buntings (Aves: Emberizidae). *Systematic Biology* **73**(2):279–289.



Zoological
Research

Zhao M, Gordon Burleigh J, Olsson U, et al. 2023. A near-complete and time-calibrated phylogeny of the Old World flycatchers, robins and chats (Aves, Muscicapidae). *Molecular Phylogenetics and Evolution* **178**:107646.

Zink RM, Pavlova A, Drovetski S, et al. 2009. Taxonomic status and evolutionary history of the *Saxicola torquata* complex. *Molecular Phylogenetics and Evolution* **52**(3):769–773.

Accepted



Zoological
Research

Supporting Information

Additional supporting information can be found online in the Supporting Information section.

Figure S1. The phylogeny of several species within the *Saxicola torquatus* complex revealed by past studies. A: A phylogeny based on the mitochondrial *ND2* identified *S. stejnegeri* as the most basal lineage of the *S. torquatus* complex, while *S. maurus* was placed as sister to the clade comprising *S. rubicola* and *S. dacotiae*. B: In contrast, a phylogeny based on genomic data indicated that *S. maurus* is sister to the clade formed by *S. rubicola*, *S. dacotiae*, and *S. torquatus*, indicating a cytonuclear discordance. However, due to the absence of genomic data for *S. stejnegeri*, the phylogenetic relationship between *S. stejnegeri* and *S. maurus* remains unresolved. As a result, the status of *S. stejnegeri* as an independent species has been questioned.

Figure S2. Normalized Z chromosome depth for different samples. The red horizontal dashed line at a normalized Z-depth of 0.5 indicates the expected normalized Z-depth for females, and the green horizontal dashed line at a normalized Z-depth of 1 indicates the expected normalized Z-depth for males. The type/sex of the samples is annotated beside each sample points.

Figure S3. Violin plots showing autosomal heterozygosity rates in 100-kb non-overlapping windows for seven samples.

Figure S4. The species tree constructed through multiple methods, with numerical values on the nodes representing node support.

Figure S5. Hybridization count testing for species network inference. A: Log-pseudolikelihood results from the SNaQ analysis. The steepest change in log-likelihood suggesting the optimal number of reticulations as 1. B: Log-pseudolikelihoods of the best network inferred under different maximum numbers of reticulations in PhyloNet, suggesting the optimal number of reticulations as 2. C: The OptM result of multiple migration edges in TreeMix, suggesting the best number of migration edges as 4.

Figure S6. The bubble plot and Cnetplot of GO enrichment results for introgressed genes between species: (A) *S. rubicola* and *S. maurus*, (B) *S. dacotiae* and *S. maurus*, (C) *S. maurus* and *S. stejnegeri*, (D) *S. dacotiae* and *S. torquatus*. Terms marked in blue indicate statistical significance ($P\text{-adjust}<0.05$), and terms marked in red are related to mitochondrial function.

Figure S7. The bubble plot and Cnetplot of KEGG enrichment results for introgressed genes between species: (A) *S. rubicola* and *S. maurus*, (B) *S. dacotiae* and *S. maurus*, (C) *S. maurus* and *S. stejnegeri*, (D) *S. dacotiae* and *S. torquatus*. Terms marked in blue indicate statistical significance ($P\text{-adjust}<0.05$).

Table S1. Information on genome data, mapping, heterozygosity, and species-special SNV statistics.

Table S2. Detailed information on the genomic coordinates and functional annotations of the 202 mitonuclear genes investigated in this study.

Table S3. Results of mitonuclear co-introgression tests based on the gene-based approach. Values highlighted with an orange background indicate statistical significance ($P<0.05$).

Table S4. Results of mitonuclear co-introgression tests based on the window-based approach. Values highlighted with an orange background indicate statistical significance ($P<0.05$).



Zoological
Research

Table S5. Alignment of partial *Saxicola torquatus* sequencing reads against the core nucleotide database; only the top 10 species by hit count are displayed. Rows with orange background indicate bacterial matches.

Accepted



Zoological
Research

Figures:

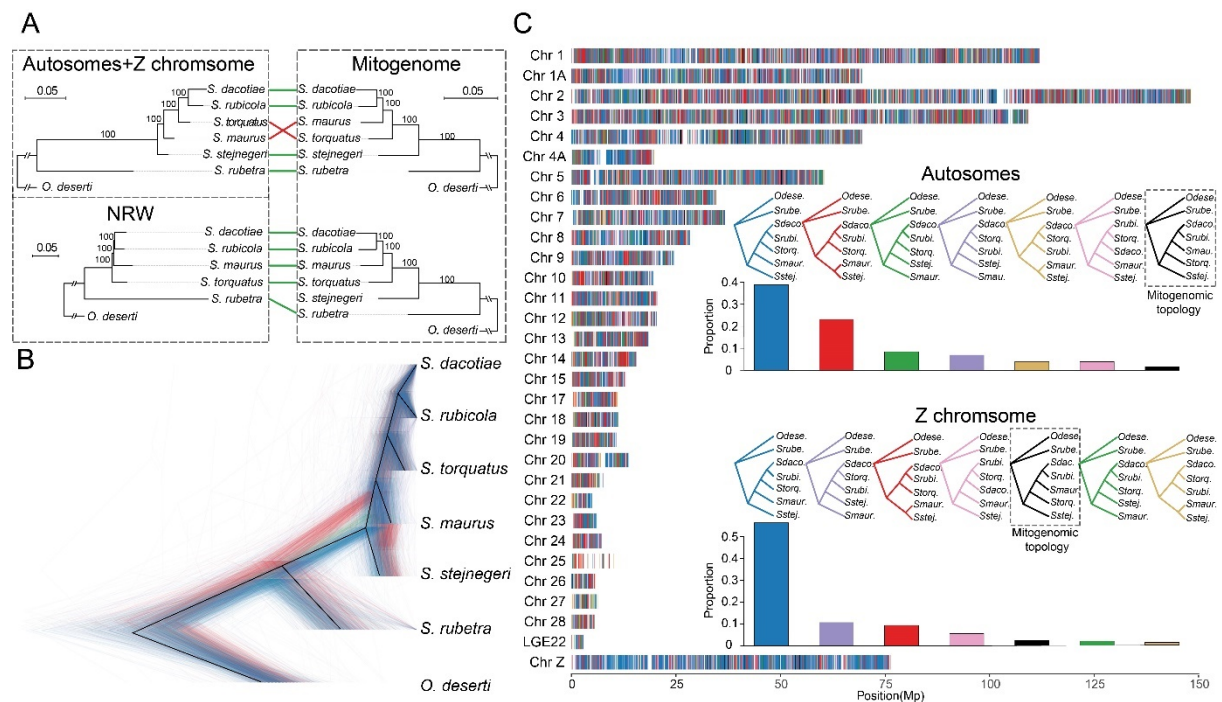


Figure 1 Phylogenies across different genomic regions. A: Phylogenetic trees of the Autosomes, Z chromosome, NRW, and mitogenome constructed using the ML method in RAxML-NG. Solid lines connect the same taxa across trees, with green indicating phylogenetic congruence and red indicating incongruence across different trees. B: DensiTree visualization of 1 500 sampled GF trees across the genome. Different colors represent distinct tree topologies. C: The distribution and proportions of the seven main tree topologies across the autosomes and Z chromosome based on a 50 kb sliding window. Alternating colors along each chromosome represent different tree topologies, with colors matching those shown in the panel. White regions represent tree topologies outside the seven main topologies. The black topology, outlined by a dashed black box, corresponds to the mitogenome tree. The histograms show the proportions of each tree topology on autosomes and the Z chromosome. The abbreviations *Odese.*, *Srube.*, *Sdaco.*, *Srubi.*, *Storq.*, *Smaur.*, *Sstej.* represent *O. deserti*, *S. rubetra*, *S. dacotiae*, *S. rubicola*, *S. torquatus*, *S. maurus*, and *S. stejnegeri*, respectively. The vertical axis shows the

chromosome number, while the horizontal axis represents the chromosome length scaled in million base pairs (Mb).

Accepted



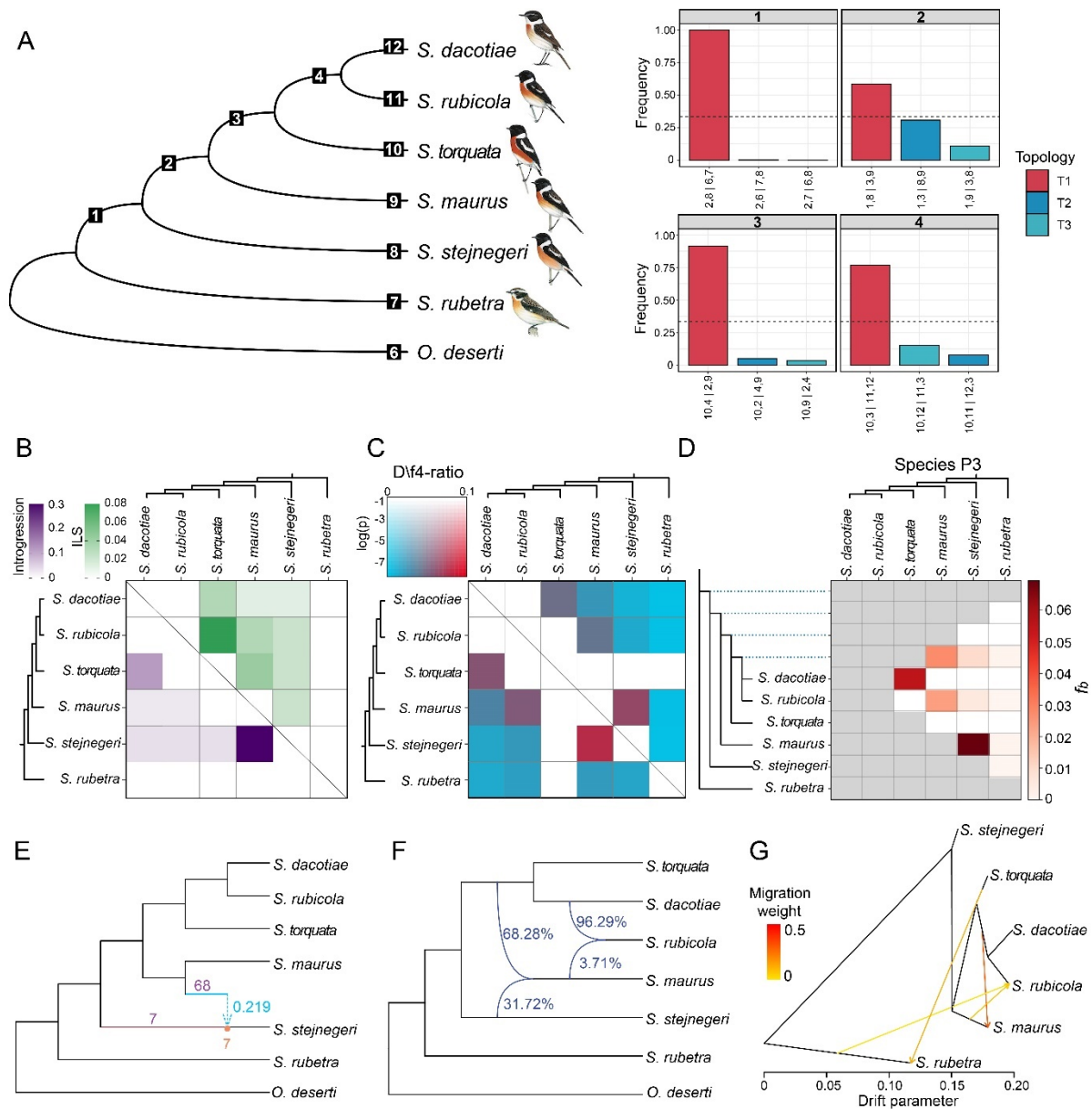


Figure 2 Topologies distribution per node and interspecific introgression. A: Species tree of six *Saxicola* species (left) and the frequency of three topologies per node, calculated by DiscoVista (right). Main quartet topologies are in red, alternatives in blue. Dotted lines represent the 1/3 threshold. The bird images were sourced from Birds of the World (2025). B: Species pairwise metrics for the proportions of introgression (below diagonal) and ILS (above diagonal) estimated by QuIBL. Darker purple and green indicate higher proportions of introgression and ILS, respectively. C: Species pairwise metrics show Patterson's D (below diagonal) and f_4 -ratio (above diagonal) inferred from every trio of the six species. The color

legend corresponds to the P value (block-jackknife) and the magnitude of introgression (D and f_4 -tario). D: The f_b statistic identifies possible introgression events from the branch of the tree on the y axis to the species on the x axis. E: Inferred species networks using SNaQ with one hybridization event. Bootstrap values for hybrid nodes and hybrid edges are shown in orange and purple, respectively. Light blue dashed lines with numbers denote minor edges and their inheritance probabilities. F: Inferred species networks using Phylonet with a maximum of 2 reticulations. Blue percentages show inheritance probabilities. G: TreeMix result when 4 migration edges were allowed. Migration edge color indicates migration weight.

Accepted Article



Zoological
Research

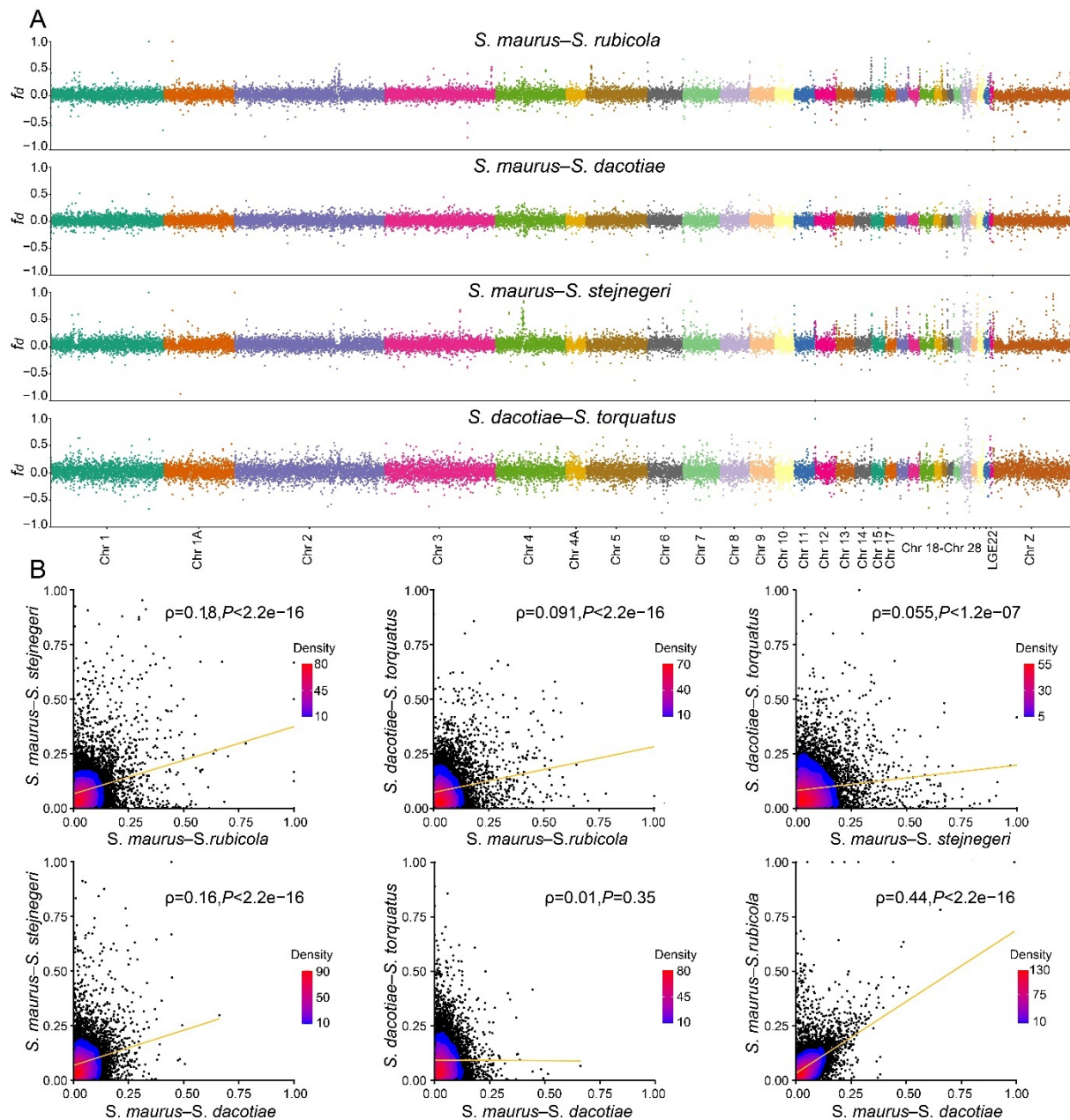


Figure 3 The introgression landscape of four species pairs and their pairwise correlations. A: The introgression landscape of the four species pairs revealed by f_d calculated based on the ABBA test. Alternating colors indicate different chromosomes corresponding to the x-axis. The y-axis represents the f_d value, with values greater than 0 indicating introgression between the target two species, while values less than 0 have no biological significance. B: Pairwise correlation of f_d among the four species pairs. Each point in the scatterplots represents one 50 kb genomic window. Orange lines are best-fit lines, and the Spearman's rank correlation rho

(ρ) coefficient and P value are provided. The gradient from blue to red indicates increasing point density.

Accepted



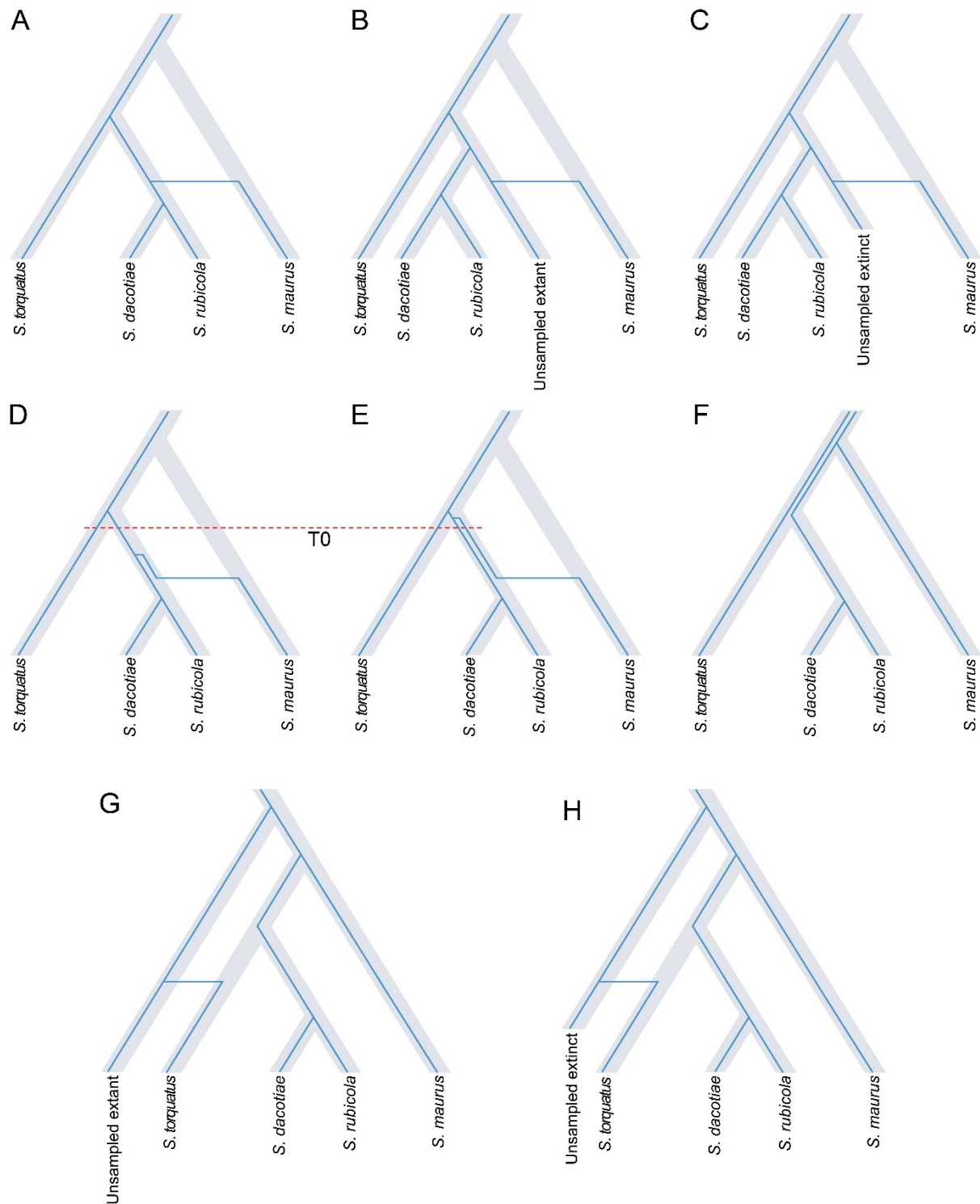


Figure 4 Eight hypotheses potentially explaining the observed incongruence between the species tree and mitogenome/NRW trees. The incongruence may result from the mitogenome/NRW of *S. maurus* being captured from: (A) the MRCA of *S. dacotiae* and *S. rubicola*; (B) an unsampled extant lineage sister to the MRCA of *S. dacotiae* and *S. rubicola*; (C) an extinct lineage sister to the MRCA of *S. dacotiae* and *S. rubicola*; (D) an unsampled

lineage within the MRCA of *S. dacotiae* and *S. rubicola* generated by female-specific selection or male-biased dispersal; (E) an unsampled lineage within the MRCA of *S. dacotiae* and *S. rubicola* generated by ILS from ancestral polymorphisms in the MRCA of *S. torquatus*, *S. rubicola*, and *S. dacotiae*. The incongruence may have arisen directly from ILS resulting from ancestral polymorphisms in the MRCA of *S. torquatus*, *S. dacotiae*, *S. rubicola*, and *S. maurus* (F). The mitogenome/NRW of *S. torquatus* may have been captured from an unsampled extant (G) or extinct (H) lineage that is sister to the MRCA of *S. dacotiae*, *S. rubicola*, *S. maurus* and *S. torquatus*. The species tree is shown in grey with the blue lines representing mitogenome/NRW lineages. The divergence time (T0) between different lineages of mitogenome/NRW within the MRCA of *S. dacotiae* and *S. rubicola* would be shallower than the root (D) or around the root divergence time (E).

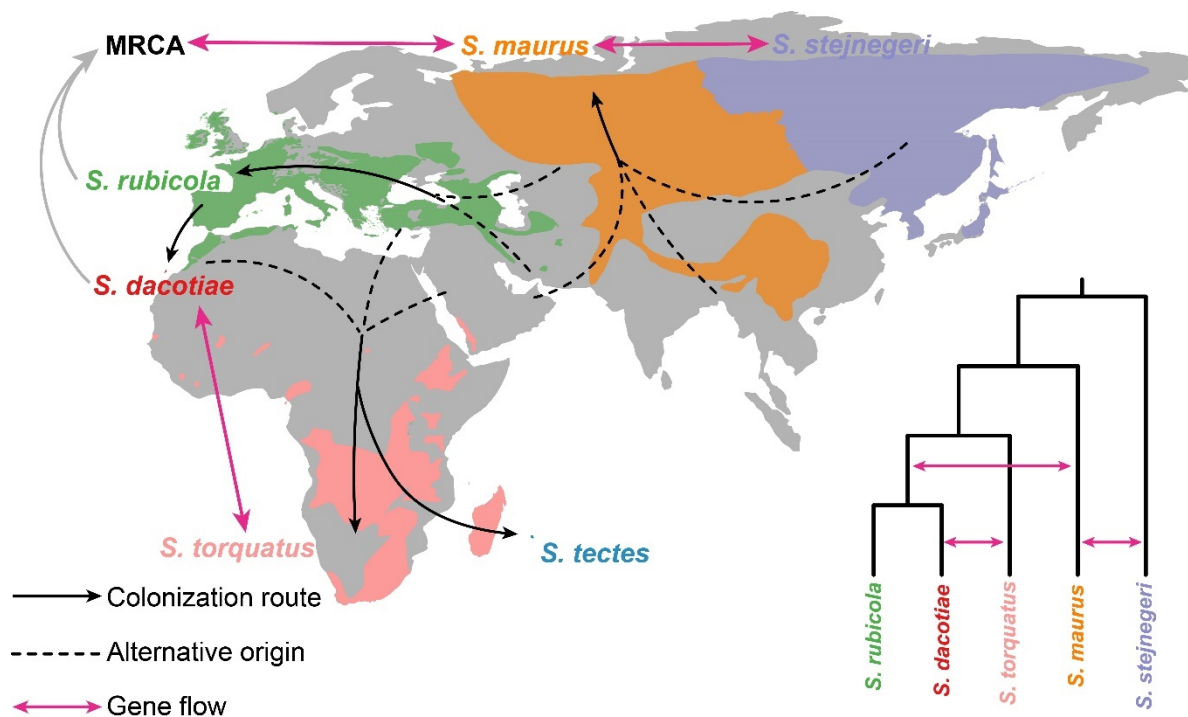


Figure 5 Schematic diagram of colonization and hybridization in five *S. torquatus* complex species. Colored regions denote the breeding areas of each species, based on data from Birds of the World (2025). The breeding ranges of *S. maurus* and *S. stejnegeri* are separated by the Yenisey/Angara rivers. Black arrows indicate colonization routes, with dashed lines at the starting points representing alternative origins. A gray arrow points from *S. rubicola* and *S. dacotiae* to their MRCA. Red bidirectional arrows on both the map and the phylogenetic tree denote gene flow.

题目：系统发育基因组学揭示了一个欧亚-非洲鸣禽辐射的网状演化历史

作者：姜川^{1,2}; 刘阳²; 朱文文³; Nassoro Mohamed^{1,4}; 晋家正⁵; 李波^{1,6*}

¹野生动物与自然保护地学院, 东北林业大学, 哈尔滨, 黑龙江 150040, 中国

²生态学院, 中山大学, 深圳, 广东 518107, 中国

³生命科学技术学院, 暨南大学, 广州, 广东 510632, 中国

⁴College of African Wildlife Management, Mweka, Moshi P.O Box 3031, Kilimanjaro, Tanzania

⁵华大基因, 武汉, 湖北 430074, 中国

⁶国家林业和草原局野生动植物检测中心, 哈尔滨, 黑龙江 150040, 中国

*通讯作者, E-mail: libo_770206@nefu.edu.cn

中文摘要:

系统发育基因组学能够揭示分类单元间复杂的演化假说, 为重构演化历史提供关键证据, 如不完全谱系分选 (ILS) 和基因渐渗等现象。黑喉石鹇复合种由 *Saxicola* 属内几个广泛分布于古北界与非洲热带界的近缘物种组成, 其演化历史仍不清楚。先前研究提示该类群存在核质不一致现象, 反映出复杂的演化历史。本研究基于全基因组单核苷酸变异数据, 对该复合种中近乎所有类群进行了系统发育基因组学分析, 旨在解析不同基因组区域系统发育信号不一致的成因, 并进一步阐明其演化历史。研究结果发现, 该复合种分化过程中存在多次网状演化事件, 导致全基因组范围内出现广泛的系统发育不一致。线粒体与 W 染色体显示了一致的基因树结构, 但与物种树不一致, 其中基因渗入是多个假设中最合理的解释。Z 染色体受 ILS/基因渗入的影响较小, 表现出相较于常染色体更稳定的局部系统发育。此外, 强渗入区域主要集中在较小染色体的末端, 且物种对间的基因渗入强度均呈显著正相关, 可能与这些区域共享的重组热点有关。这些发现证实了由复杂演化过程塑造的异质性基因组景观, 支持黑喉石鹇复合体物种形成遵循网状演化模型。

关键词: 基因渐渗; 系统发育基因组学; 石鹇; 单核苷酸变异



Zoological
Research