Correlated genomic landscapes in birds

ORIGINAL ARTICLE

Correlated patterns of genetic diversity and differentiation across an avian family

Benjamin M. Van Doren¹²* (BVD), Leonardo Campagna¹² (LC), Barbara Helm³ (BH), Juan Carlos Illera⁴ (JCI), Irby J. Lovette¹² (IJL), and Miriam Liedvogel⁵ (ML)

¹ Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA
² Cornell Lab of Ornithology, Cornell University, Ithaca, NY 14850, USA
³ Institute of Biodiversity, Animal Health and Comparative Medicine, University of Glasgow, Glasgow G12 8QQ, UK
⁴ Research Unit of Biodiversity (UO-CSIC-PA), Oviedo University, Campus of Mieres, Research Building, 5th Floor, C/ Gonzalo Gutiérrez Quirós s/n, 33600 Mieres, Asturias, Spain
⁵ Max Planck Institute for Evolutionary Biology, AG Behavioural Genomics, August-Thienemann-Str. 2, 24306 Plön, Germany

* Corresponding author: benjamin.vandoren@zoo.ox.ac.uk. Current address: Edward Grey Institute, Department of Zoology, University of Oxford, Tinbergen Building, South Parks Road, Oxford, OX1 3PS, UK

Keywords: Differentiation peaks, Ficedula, Genetic drift, Genomic scan, Linked selection, Pooled sequencing, Saxicola.
Comparative studies of closely related taxa can provide insights into the evolutionary forces that shape genome evolution and the prevalence of convergent molecular evolution. We investigated patterns of genetic diversity and differentiation in stonechats (genus *Saxicola*), a widely distributed avian species complex with phenotypic variation in plumage, morphology, and migratory behavior, to ask whether similar genomic regions have become differentiated in independent, but closely related, taxa. We used whole-genome pooled sequencing of 262 individuals from 5 taxa and found that levels of genetic diversity and divergence are strongly correlated among different stonechat taxa. We then asked if these patterns remain correlated at deeper evolutionary scales and found that homologous genomic regions have become differentiated in stonechats and the closely related *Ficedula* flycatchers. Such correlation across a range of evolutionary divergence and among phylogenetically independent comparisons suggests that similar processes may be driving the differentiation of these independently evolving lineages, which in turn may be the result of intrinsic properties of particular genomic regions (e.g., areas of low recombination). Consequently, studies employing genome scans to search for areas important for reproductive isolation or adaptation should account for corresponding regions of differentiation, as these regions may not necessarily represent speciation islands or evidence of local adaptation.
INTRODUCTION

Evolutionary biologists seek to understand the genetic basis of speciation and the degree to which the divergence of lineages may involve independent changes on similar loci (Seehausien et al. 2014). Genomic sequencing has made it possible to examine patterns of differentiation across the genomes of organisms at different stages of divergence. Recent comparative studies of genome-wide patterns of variation, or “genomic landscapes,” have identified areas of the genome that are conspicuously differentiated relative to the genomic baseline among closely related taxa (Ellegren et al. 2012; Ruegg et al. 2014; Burri et al. 2015; Wang et al. 2016). It remains uncertain whether these regions are functionally important in speciation, and whether they typically arise during speciation-with-gene-flow or as a consequence of selection in allopatry.

Most empirical studies have used F-statistics (Wright 1965) and other measures that compare allele frequencies between two populations to infer the magnitude of differentiation across the genome. These statistics are influenced by levels of within-population genetic variation and are therefore classified as “relative” measures of divergence (Hedrick 2005). Genomic outlier regions of high differentiation were first described as “islands of speciation” in the face of gene flow (Turner et al. 2005) and hypothesized to harbor loci that were important for reproductive isolation (Nosil et al. 2009; Feder et al. 2012). However, subsequent studies have identified alternate mechanisms by which isolated genomic regions of elevated differentiation can be generated in allopatry, and thus independently of gene flow (e.g., Noor and Bennett 2009; Turner and Hahn 2010; White et al. 2010; Cruickshank and Hahn 2014). For example, post-speciation selective sweeps or background selection, especially in regions of reduced recombination, can
drive the differentiation of these loci relative to the rest of the genome (Nachman and Payseur 2012; Cruickshank and Hahn 2014; Burri et al. 2015).

Selective sweeps bring beneficial alleles to high frequency in a population, greatly reducing genetic diversity at linked sites via “hitchhiking” (Smith and Haigh 1974; Kaplan et al. 1989). The magnitude of the hitchhiking effect is influenced by recombination rate, in addition to the strength of selection, with areas of low recombination experiencing greater linkage and a commensurate reduction in diversity across larger sections of a chromosome ( Begun and Aquadro 1993; Charlesworth et al. 1997; Nielsen 2005). This reduction in local within-population genetic diversity results in high differentiation as measured by FST (Charlesworth 1998; Keinan and Reich 2010; Cruickshank and Hahn 2014). Recurrent sweeps in similar genomic regions across independent populations may be caused by selection for different advantageous alleles at the same locus, or by selection on different but tightly linked loci. Alternatively, they could result from the adaptive introgression of globally advantageous mutations transmitted among populations by gene flow, followed by sweeps due to local adaptation (see Roesti et al. 2014; Delmore et al. 2015). A similar pattern could also arise from the increased establishment probability of beneficial mutations linked to selected sites in areas of low recombination (Yeaman et al. 2016). These related processes can result in corresponding areas of low genetic diversity and high differentiation in independent population comparisons.

Similarly, background (or purifying) selection purges deleterious alleles as they arise and may also independently generate similar genomic landscapes of diversity and differentiation across populations ( Charlesworth 2013; Burri et al. 2015; Wang et al. 2016). Under this scenario, a neutral variant that emerges in a population will subsequently disappear if it is linked to a deleterious mutation, a process that reduces nucleotide diversity (Charlesworth et al. 1993;
Charlesworth et al. 1997; Stephan et al. 1998). As recombination rates decrease, linkage extends over larger genetic distances and the probability of a neutral variant associating with a deleterious mutation (and thus being purged) is higher. Therefore, areas of low recombination will generally exhibit a greater reduction in genetic diversity due to background selection (Charlesworth et al. 1993; Nordborg et al. 1996). If highly conserved genomic regions (e.g., of great functional importance) and/or areas of low recombination are similar across species, the effect of background selection may cause, or contribute to, parallel genomic landscapes of diversity and differentiation in comparisons of independently evolving lineages (Nordborg et al. 1996; Andolfatto 2001; Cruickshank and Hahn 2014).

The effects of background selection and selective sweeps on linked neutral loci have collectively been referred to as “linked selection” (Turner and Hahn 2010; Cutter and Payseur 2013; Cruickshank and Hahn 2014). The frequency with which parallel signatures of linked selection occur in closely related taxa and the contributions of background selection and selective sweeps in shaping genomic landscapes remain actively debated (Keinan and Reich 2010; Burri et al. 2015). In addition, the degree to which this parallelism may extend beyond a few well-studied species complexes is currently unknown.

Genome-wide scans of two independent groups of closely related bird species, *Ficedula* flycatchers and *Phylloscopus* warblers, have identified conspicuous peaks of relative divergence (i.e., genomic regions with very different allele frequencies) in pairwise comparisons of congeners that coincide with “valleys” of absolute divergence (i.e., regions with few sequence differences, not influenced by within-population genetic diversity) (Burri et al. 2015; Irwin et al. 2016). This inverse relationship is inconsistent with the speciation-with-gene-flow paradigm, in which regions of high relative divergence are resistant to gene flow and therefore should show
high absolute divergence (Noor and Bennett 2009; Nachman and Payseur 2012; Cruickshank and Hahn 2014). This suggests that post-speciation selection—not divergence-with-gene-flow—generates differentiation peaks in these systems. Within their respective species complexes, flycatchers and warblers show signatures of selection in similar genomic areas (Burri et al. 2015; Irwin et al. 2016), but neither the specific type of selection, nor their contribution to the speciation process, has been fully characterized. Furthermore, these studies primarily test the correspondence of divergent regions using correlation-based methods, which can be strongly affected by pseudoreplication due to linkage.

Here, we characterize the course of genome-wide molecular evolution in a well-studied group of birds, the *Saxicola* stonechats (Urquhart 2002; Collar 2016a, 2016b). This genus began diversifying during the late Miocene (8.2 million years ago; Illera et al. 2008) and currently comprises 15 recognized species (51 named taxa including subspecies; Gill and Donsker 2016). Some taxa are restricted to small islands, while others span continents, and they range from long distance migrants to year-round residents (Baldwin et al. 2010). The well-documented evolutionary diversity in this clade makes *Saxicola* a powerful system for studying independently evolving lineages across a gradient of differentiation, phenotypic variation, and life histories. We examine five stonechat taxa at disparate stages of divergence, including two that likely still exchange genes and two that diverged ~3.7 million years ago (Illera et al. 2008). These taxa show variation in morphology and behavior (e.g., body size and migratory direction), and we survey both island and continental taxa, which are likely to have varied demographic histories.

Our primary research focus is to investigate the extent to which genome evolution is correlated in independently evolving, but closely related, taxa. We ask: Have the same regions of the genome become differentiated over time in independent stonechat lineages? If so, what role
has natural selection played in driving this correlated differentiation? We further ask if evolution
is correlated at a deeper scale, between stonechats and two Ficedula flycatchers (Sætre and
Sæther 2010). Both genera belong to the family Muscicapidae. We posit that any loci that are
differentiated in both genera are unlikely to arise from parallel ecological selection pressures and
instead stem from intrinsic properties of those genomic regions. Finally, we examine the effects
of life history and demography on the genome by comparing patterns of genetic diversity and
differentiation between continental and island taxa. We hypothesize that an island taxon will
show a weaker overall effect of selection on the genome, reflecting the theoretical prediction of
increased drift with a smaller effective population size. Our goal is to shed light on the processes
that influence the most conspicuous features—the high “peaks” and low “valleys”—of stonechat
genomic landscapes. The results underscore the degree to which broad patterns of genetic
diversity and differentiation are correlated across evolutionary time.

METHODS

Study system and sampling

We included five stonechat taxa in this study: Saxicola rubicola rubicola from Austria (European
stonechat); S. rubicola hibernans from Ireland (European stonechat); S. torquatus axillaris from
Kenya (African stonechat); S. maurus maurus from Kazakhstan (Siberian stonechat); and S.
dacotiae dacotiae from Fuerteventura Island, Spain (Canary Islands stonechat) (Gill and Donsker
from the remaining four taxa about 3.7 mya, Siberian stonechats subsequently split from the
remaining three about 2.5 mya, and Canary Islands stonechats diverged from European
stonechats about 1.6 mya. Illera and colleagues could not distinguish Austrian and Irish
stonechats using mitochondrial DNA. We expect that Canary Islands stonechats have diverged from other taxa without gene flow because this taxon occurs on an oceanic island, and the present-day ranges of Kenyan and Siberian stonechats lead us to expect no ongoing gene flow between these and the other taxa. Conversely, we expect that Austrian and Irish stonechats likely still exchange genes because of their close geographic proximity, lack of mitochondrial divergence, and evidence of breeding dispersal between the British Isles and continental Europe (Helm et al. 2006).

Most of the 262 stonechats included in this study originated from a common-garden experiment that Eberhard Gwinner initiated in 1981 at the Max-Planck Institute in Andechs, Germany, except for Canary Islands stonechats, which were directly sampled in the wild (Table S1). For the other species, most birds were taken into captivity as nestlings, except for Irish stonechats (~50% captured on winter territories). The remaining sampled individuals were offspring of these captive stonechats, hatched between 1988 and 2006. Despite the inclusion of captive birds, relatedness within the pools was low (Table S1). Detailed descriptions of breeding and raising conditions are published elsewhere (Gwinner et al. 1987; Helm 2003; Helm et al. 2009).

The inclusion of second-generation progeny in our study could potentially lower measured levels of genetic diversity relative to a comparable sample of wild individuals. However, we find average genetic diversity (π) to be highest in Siberian stonechats, the species for which we incorporated the most captive-bred birds; this suggests that any putative bias is small and potentially negligible for the purposes of this study.

Draft reference genome
We assembled the genome of a male Siberian stonechat (*S. maurus*) collected in Kazakhstan (44.59° N, 76.609° E) and housed at the Burke Museum (UWBM# 46478). We generated one fragment library with insert sizes of 180 base pairs (bp) and two mate-pair libraries (insert sizes: 3 and 8 kilobases), and we sequenced each of them on one Illumina HiSeq 2500 lane (obtaining 101-bp paired-end reads). We assembled the draft reference genome using the ALLPATHS-LG algorithm (Gnerre et al. 2011) and used HaploMerger (Huang et al. 2012) to improve the assembly by merging homologous scaffolds and removing those resulting from the erroneous split of two haplotypes into separate scaffolds. The final Siberian stonechat assembly comprised 2,819 scaffolds, with a total scaffold length of 1.02 Gb and an N50 scaffold size of 10.0 Mb. Half of the final assembly is represented in 24 scaffolds, and 75% in 65 scaffolds. Ambiguous bases (N’s) make up 4.4% of its total length.

We assembled scaffolds from our stonechat reference genome into draft chromosomes by mapping them to the *Ficedula albicollis* genome assembly, version 1.5 (RefSeq accession GCF_000247815.1) (Kawakami et al. 2014) and used SatsumaSynteny (Grabherr et al. 2010) to align the *Saxicola* draft genome to the *F. albicollis* assembly. Because these species are phylogenetically close and synteny is relatively conserved among birds (Ellegren 2010), this method allowed us to position 97.1% of the stonechat reference genome in the presumed correct order. Inversions and other chromosomal rearrangements occur in birds (Backström et al. 2008), so it is possible that a small percentage of the genome may be ordered or oriented incorrectly.

**Resequencing of five stonechat taxa**

We extracted genomic DNA from stonechat blood or tissue samples using a salt extraction protocol and selected 49-56 individuals (n_{total} = 262, including both males and females) from each
of the five stonechat taxa for sequencing (Table S1). We created five pooled libraries, one per
taxon, from equimolar aliquots of DNA using the Illumina TruSeq DNA kit and sequenced them
on an Illumina NextSeq 500 (Table S2).

We used BWA-MEM (Li 2013) to align sequences to our reference genome and
performed refinement and quality control steps with Picard
(http://broadinstitute.github.io/picard/) and the Genome Analysis Toolkit (GATK) (McKenna et
al. 2010), including filtering by mapping quality and removing duplicate sequences (Supporting
information). Sequences mapped to the draft reference genome at a mean per-pool coverage
between 13.8x and 26.1x, and mean mapping quality was between 45-46 for all taxa (Table S2).
Although this level of coverage is insufficient to sequence every individual at every locus, the
goal of our pooled sequencing strategy was to estimate population allele frequencies by sampling
a subset of chromosomes in a pool. Gautier et al. (2013) found that allele frequencies of
individual SNPs estimated with 10-50x pool coverage (pool size = 30) were strongly correlated
with estimates derived from separate individual-based (n = 20) sequencing at 1x-6x (r = 0.93) and
6-10x (r = 0.94) per individual. Additionally, the effects of pool-derived sampling error are
greatly reduced in window-based analyses where variation and differentiation are summarized
across groups of SNPs (Kofler et al. 2011a). Because we use a windowed approach and therefore
do not rely on the frequencies of individual SNPs, we are confident that we can accurately assess
and compare genome-wide patterns of genetic variation with this level of coverage.

**SNP-based phylogeny and inter-taxa divergence**

Although we used an existing mitochondrial phylogeny (e.g., Illera et al. 2008) as a basis for our
study approach and design, we also constructed a phylogenetic tree for the five focal stonechat
taxa using nuclear markers. This step was designed to confirm the mitochondrial findings and serve as a basis for phylogeny-based inference. We used Pied and Collared Flycatchers (Ficedula hypoleuca and F. albicollis) as outgroups. We constructed a maximum likelihood tree with RAxML (Stamatakis 2014), using 16,876,859 fixed single-nucleotide polymorphisms (SNPs) from across the nuclear genome, of which 330,592 were polymorphic among the stonechat ingroup (Supporting information). We applied the Lewis correction, following the recommendation of Stamatakis (2014), for ascertainment bias resulting from the exclusion of invariant sites.

We generated mean genome-wide $F_{ST}$ and $d_{XY}$ pairwise distance matrices for the five focal stonechat taxa and displayed them graphically using principal coordinate analyses performed with the ape package in R (Paradis et al. 2004).

**Pooled population genomic analyses**

We analyzed sequence data with the software packages npstat (Ferretti et al. 2013) and Popoolation2 (Kofler et al. 2011b), designed specifically for the analysis of pooled sequencing data. With npstat we calculated: (1) Tajima’s $D$, to test for rare variants as a signal of directional or purifying selection or large-scale demographic effects; (2) $\pi$, an estimate of genetic diversity, which is derived from the number of pairwise sequence differences among members of a population; and (3) Fay and Wu’s $H$, a statistic related to Tajima’s $D$ but sensitive only to high frequency derived alleles, thus influenced by positive selection but not by background selection (Fay and Wu 2000). We polarized alleles using the Collared Flycatcher.

We then used Popoolation2 to calculate pairwise $F_{ST}$ among all pairs of taxa. We also estimated $d_{XY}$ (Nei and Li 1979; Cruickshank and Hahn 2014), a measure of absolute divergence,
as $A_XB_Y + A_YB_X$, where $A$ and $B$ are the frequencies of the two alleles at a locus and $X$ and $Y$ denote the two groups being compared.

For all analyses, we excluded bases within 5 bp of indels to reduce the probability of including erroneous genotypes due to misalignments. We calculated all metrics for 50-kb non-overlapping windows (Supporting information), within which we only considered sites with minor allele counts $\geq 2$ and coverage between half and three times that of the pool’s average. We only retained windows in which at least 40% of bases (i.e. 20 kb) satisfied this coverage criterion.

For $d_{XY}$, we calculated the windowed value by summing over the window and dividing by the total number of sites with sufficient coverage (variable or not). We calculated standardized nucleotide diversity for each taxon by dividing $\pi$ by the maximum $d_{XY}$ value from all pairwise comparisons involving that taxon (following Irwin et al. 2016).

Correlation analyses

We first quantified the similarity of genome-wide patterns of genetic diversity and divergence using Spearman rank correlations. Although the p-values of these tests are affected by pseudoreplication due to the inclusion of genetically linked loci, they are nonetheless a valuable summary of genome-wide similarity and provide a means to compare the results of the present study with previous work.

Identification of genomic outlier regions

We identified regions of the stonechat genome showing consistently elevated or lowered values of Tajima’s $D$, $\pi$, Fay and Wu’s $H$, $F_{ST}$, and $d_{XY}$, and therefore may be important in the divergence of stonechat lineages. In particular, we wanted to determine whether any genome-
wide similarities revealed by the correlation analyses were driven by a relatively small number of genomic regions. We applied a kernel-based smoothing algorithm across 50-kb windows (box density with bandwidth of 30; see Supporting information) and compared this smoothed line with 25,000 smoothed lines obtained after permuting the order of the windows (see Ruegg et al. 2014). We called outlier locations where the observed smoothed line was more extreme than the most extreme smoothed value from the null (permutation) distribution. We merged outlier regions separated by four windows or fewer (i.e. by <200 kb). Because the effective population size of the Z chromosome is smaller than that of the autosomes, baseline levels of variation and differentiation are different from those of autosomes (Charlesworth 2001). We therefore permuted windows of the Z chromosome and autosomes separately (see Fig. S1, Supporting information).

Concordance of genomic outlier regions within Saxicola

Once outlier regions were identified, we assessed their overlap among stonechat taxa. For each pairwise comparison, we counted the number of outlier regions that showed any degree of overlap between the two datasets. By considering each region separately, we account for autocorrelation of their constituent windows due to linkage. Although this approach addresses the pseudoreplication that would have resulted from treating the multiple windows within the same outlier region as independent observations, it is important to note that it does not address the larger-scale possibility that multiple outlier regions could be clustered together, e.g. due to a very large region of reduced recombination.

We then tested whether the observed number of overlapping regions was significantly greater than expected under the null hypothesis of no association in outlier positions between
datasets, using a custom permutation test. While holding the number and size of outlier regions constant, we randomly permuted their locations across the genome 1,000 times and measured the degree of overlap under these simulated scenarios. The p-value of the test was the proportion of simulations under which the number of overlapping outlier regions was equal to or greater than the observed value; we thus accounted for the varying number and size of outlier regions in each comparison. We applied a false discovery rate correction to each series of tests (Benjamini and Hochberg 1995) and considered tests with corrected p-values less than 0.05 to be statistically significant.

For each comparison, we calculate the proportion of outlier regions in one genomic landscape also present in the other (and vice versa) and report the greater of these two values. Thus, if landscape 1 shows 10 peaks and landscape 2 shows 50 peaks, and 9 out of 10 peaks in landscape 1 are also present in landscape 2, our outlier similarity score will be 9/10 = 0.90.

Previous studies have used scatterplots and correlation analyses as the primary manner of assessing association between outlier regions in independent comparisons (e.g., Burri et al. 2015; Irwin et al. 2016). However, these tests are affected by autocorrelation due to genetic linkage. By considering each outlier region as a single unit, our permutation approach overcomes this issue by treating each contiguous outlier region (instead of each 50-kb window) as an independent observation.

Correspondence of genomic landscapes between Saxicola and Ficedula

To test for conservation of genomic landscapes at a deeper level of divergence, we compared stonechat genomic landscapes to those of the genus Ficedula. We calculated $F_{ST}$, $d_{XY}$, and $\pi$ for Pied Flycatchers ($F. hypoleuca$) and Collared Flycatchers ($F. albicollis$). We downloaded reads
from the Sequence Read Archive (project ERP007074; accession PRJEB7359; http://www.ncbi.nlm.nih.gov/sra) for 10 individuals of each species (Smeds et al. 2015) (Table S3), and processed reads for quality as described for the stonechat analysis (Supporting information).

We filtered, trimmed and aligned *Ficedula* reads to the stonechat draft reference genome so that we could directly compare the locations of outlier regions between genera, using the same tools in GATK and Picard as for stonechat sequences (Supporting information). To calculate \( F_{ST} \), we first generated a VCF file with UnifiedGenotyper from GATK and filtered raw variants with the VariantFiltration tool (settings: QD < 2.0 || FS > 60.0 || MQ < 40.0). We then calculated \( F_{ST} \) with VCFtools (Danecek et al. 2011) from the resulting SNPs across 50-kb non-overlapping windows. We estimated \( d_{XY} \) from minor allele frequencies obtained in ANGSD (Korneiussen et al. 2014), using a custom script to calculate 50 kb windowed averages. We only included sites that had genotype calls for at least 5 out of 10 individuals per species and retained windows for which at least 40% of bases satisfied this criterion.

**RESULTS**

**SNP-based phylogeny and inter-taxa divergence**

The Maximum Likelihood (ML) phylogeny built on fixed nuclear sites showed high support for the placement of the Canary Islands stonechat as the sister taxon to the European stonechat (Austria and Ireland) (Fig. 1 A). The clade comprising European, Canarian, and Kenyan stonechats, to the exclusion of Siberian stonechats, was also strongly supported. This nuclear phylogeny contradicted the existing mtDNA topology. We verified that this result was not an
artifact of sparse taxon sampling or choice of outgroup by constructing an ML tree with
cytochrome-\textit{b} consensus sequences obtained from Austrian, Irish, Kenyan, and Siberian pools;
not enough mitochondrial sequence was recoverable for Canarian stonechats. Here, Kenyan
stonechats were placed as the sister lineage to the remaining stonechats, in agreement with past
mitochondrial studies (not shown).

We calculated mean genome-wide F\textsubscript{ST} and d\textsubscript{XY} to further examine relationships among
stonechat taxa. The first two principal coordinate axes calculated from a distance matrix of mean
pairwise F\textsubscript{ST} (explaining a total of 87\% of variance; 47\% in first axis) revealed Siberian
stonechats to be approximately equidistant from the other taxa in terms of overall allele
frequency differentiation (Fig. 1 B). Stonechats from Austria and Ireland were extremely similar
(with only 7 fixed differences out of 10,164,331 sites with F\textsubscript{ST} > 0, or 7 x 10^{-5} \%), reflecting their
geographic proximity and common evolutionary history. In contrast, Kenyan and Canary Islands
stonechats were most different (1,251,605 fixed differences out of 12,401,462 sites with F\textsubscript{ST} > 0,
or 10.09\%). Overall, Canary Islands stonechats were strikingly dissimilar to even their closest
evolutionary relatives (Austria vs. Canary Islands: 782,967 fixed differences from 12,754,086
variable sites, or 6.14\%). European stonechats were more similar genome-wide to Siberian and
Kenyan stonechats than to those from the Canary Islands, their sister lineage (Austria vs. Siberia:
244,623 fixed out of 15,168,199 variable, or 1.61\%; Austria vs. Kenya: 640,425 fixed out of
12,032,148 variable, or 5.32\%).

The principal coordinate analysis based on d\textsubscript{XY} (68\% of variance explained by first two
axes; 41\% by the first) was similar to the one based on F\textsubscript{ST}, except that Kenyan stonechats were
closer to European stonechats than to Siberian stonechats (Fig. 1 C). This is consistent with the
nuclear tree (Fig. 1 A). Again, Austria and Irish stonechats were nearly identical. Canary Islands stonechats were distant from all stonechat taxa, but most similar to the European taxa.

**Shared regions of high differentiation show low genetic diversity, except in Canary Islands stonechats**

Measures of divergence were strongly correlated among stonechats. $D_{XY}$ showed strong correlations across genomic windows (Fig. 2 A-B), and $d_{XY}$ outlier regions were highly similar (Fig. 2 D; Figs. S2 and S3, Supporting information); mean outlier similarity scores, averaged across all comparisons, were 0.85 for low $d_{XY}$ outliers and 0.79 for high $d_{XY}$ outliers (Fig. S4, Supporting information). $F_{ST}$ was also significantly correlated in all comparisons, but the strength of this correlation varied (Fig. 3 A-D). The association was greatest in comparisons including Siberian stonechats (Fig. 3 A), but $F_{ST}$ was also correlated among independent comparisons (i.e. with no shared taxon) (Fig. 3 B-C). Overall $F_{ST}$ outlier similarity was lower than $d_{XY}$ for both peaks and valleys (means of 0.31 and 0.24, respectively), indicating that approximately one-third of $F_{ST}$ peaks were shared (Fig. 3 E and Fig. S5, Supporting information). Across all comparisons, windows with the lowest $F_{ST}$ showed the most consistent associations. Of note, outlier regions showed significant overlap in several comparisons where the four taxa being compared were all different (Fig. 3 E), implicating common processes in independent stonechat lineages in the generation of differentiation landscapes.

Generally, regions of high $F_{ST}$ showed low genetic diversity, both within ($\pi$) and between ($d_{XY}$) stonechat taxa. $F_{ST}$ and $d_{XY}$ were strongly negatively correlated, especially in comparisons including Siberian stonechats (Fig. 4 A-B), and $F_{ST}$ peaks overlapped strongly with $d_{XY}$ valleys (Fig. 4 A). $F_{ST}$ valleys also overlapped with $d_{XY}$ valleys in some comparisons. Regions of
reduced absolute divergence also showed reduced nucleotide diversity (Fig. S6, Supporting information). Reductions in diversity occurred in the same genomic windows among stonechats, even after standardizing for levels of between-population diversity (Figs. S7, S8 and S9, Supporting information). Note that, due to the high similarity between Austrian and Irish stonechats, we do not present comparisons of Irish stonechats with non-Austrian stonechats.

However, these patterns of genetic variation were often weaker, absent, or even reversed for Canary Islands stonechats. Across the genome, \( F_{ST} \) and \( d_{XY} \) were positively correlated, despite the lowest-\( F_{ST} \) windows showing high \( d_{XY} \) (Fig. 4 C). Canary Islands stonechats showed the weakest associations in diversity correlations (Fig. S6 A,D and Fig. S7 B,D, Supporting information). Standardized nucleotide diversity (\( \pi/d_{XY} \)) was negatively correlated between Canary Islands and Siberian stonechats (Fig. S7 D, Supporting information), indicating that the regions of the Siberian stonechat genome that showed the greatest diversity reductions were, in fact, relatively more diverse in Canary Islands stonechats than the rest of the genome (Fig. S9, Supporting information). No \( \pi/d_{XY} \) valleys regions were shared between Canarian stonechats and other stonechat taxa (Fig. S7 E, Supporting information).

Evidence of selection and effects of demography

Among stonechats, genomic regions of high differentiation (\( F_{ST} \)), low absolute divergence (\( d_{XY} \)), and low genetic diversity (\( \pi \)) coincided with significant decreases in Tajima’s \( D \) and Fay and Wu’s \( H \) (Fig. 5 and Fig. 6). Fay and Wu’s \( H \) showed strong associations with \( F_{ST} \) only in comparisons including Siberian stonechats. Fay and Wu’s \( H \) outlier regions were relatively infrequent but coincided with low Tajima’s \( D \) and \( \pi \) when they occurred, except in Canary Island stonechats (Fig. S10, Supporting information). Tajima’s \( D \) outlier regions were generally shared
across stonechats (Fig. S11, Supporting information). Some distinct low-$H$ outlier regions occurred in only one taxon (e.g., chromosomes 4A and 6) (Fig. S12, Supporting information). Overall, in addition to lacking genetic diversity, outlier regions contained more low frequency alleles than the rest of the genome, which is highly suggestive of a role of positive and/or background selection in shaping differentiation patterns.

The genomic baseline value of Tajima’s $D$ can be biased downward by demographic effects, particularly a population expansion. All stonechat taxa had median Tajima’s $D$ between -0.5 and -1.1, with the exception of Canary Islands stonechats, at -2.8 (Fig. S13, Supporting information). Negative values suggest that all five stonechat taxa have experienced past demographic expansion events, with the signal especially strong in the insular Canary Islands stonechats.

**Correspondence of genomic landscapes between *Saxicola* and *Ficedula***

Genome-wide patterns of genetic diversity and differentiation were also correlated between stonechats and flycatchers. Absolute divergence was correlated between the two genera ($\rho = 0.37-0.49$, Fig. 2 C-D); $d_{XY}$ outlier similarity was 0.48-0.52 between stonechats and flycatchers. Flycatchers and stonechats shared a significant number of $F_{ST}$ peaks and valleys, but only for a subset of stonechat comparisons (Fig. 3 E). Genome-wide correlations of $F_{ST}$ were significant but weak (Fig. 3 F-G), and the strongest correlations occurred with Siberian stonechats. Finally, within-population genetic diversity ($\pi$) was strongly correlated between stonechat and flycatcher populations; some stonechat-flycatcher correlations were as strong or stronger than stonechat-stonechat correlations (Fig. S7 E-G, Supporting information). Overall, these results suggest that common processes are working independently and in parallel to influence genetic variation in
similar regions of the genome in both genera, although the association between genera is weaker than within *Saxicola*.

**DISCUSSION**

We examined patterns of genetic diversity and differentiation in an avian radiation and identified regions shared among stonechat taxa that were characterized by low within-population diversity, low absolute inter-taxon divergence, and high (or, in some cases, low) differentiation. These patterns are consistent with signatures of natural selection. We found that many stonechat outlier regions also appeared in the closely related genus *Ficedula*. In this genus, genomic regions of low genetic diversity and high differentiation are associated with infrequent recombination (Burri et al. 2015), which suggests that one possible explanation for the parallel patterns of differentiation in these genera is conserved (or convergently evolving) variation in recombination rate (see Singhal et al. 2015). Overall, our results are consistent with linked selection (positive selective sweeps and/or background selection) shaping large-scale patterns of genomic variation in Muscicapid birds. The presence of Fay and Wu’s *H* valleys in differentiation outlier regions supports a role of positive selection in at least some cases. Despite a strong signal of similarity in genomic landscapes, we also found evidence for substantial lineage-specific evolution: Siberian stonechats appear to have experienced the strongest effects of selection, while drift may have shaped Canary Islands stonechats’ genomes.

**Discordance in nuclear and mitochondrial phylogenies**
The phylogenetic tree constructed with SNPs from across the nuclear genome (Fig. 1 A) was highly supported at all nodes, yet it is not fully concordant with previous trees constructed from mitochondrial DNA sequences (Illera et al. 2008; Woog et al. 2008; Zink et al. 2009). These placed Kenyan stonechats (instead of Siberian, as in our reconstruction) as the sister lineage to the remaining stonechats. Branch support for a sister relationship of Siberian stonechats and the European/Canary Islands clade varied by study and tree-building algorithm. Mito-nuclear discordance could be a sign of past admixture, sex biased gene flow, or other biological phenomena (see Toews and Brelsford 2012). This well-resolved nuclear phylogeny serves as a basis for testing broader questions about genome-scale differentiation in this complex: For example, it helps explain why mean $d_{XY}$ between European and Kenyan stonechats is relatively low compared to Siberian stonechats (Fig. 1 C). Finally, although sparse taxon sampling (5 taxa) and choice of outgroup could potentially introduce biases (e.g., Stervander et al. 2015), our cytochrome $b$-only tree (not shown) was consistent with previous mitochondrial studies, which achieved near-complete taxon sampling (e.g., Illera et al. 2008). The topology differences we find between mitochondrial and nuclear-based phylogenies are therefore unlikely to be artifacts of sampling. We note, however, that high bootstrap support does not always indicate a correct species tree (e.g., Suh 2016); further investigation into the larger *Saxicola* clade (e.g., using gene tree-based methods and demographic modeling; Nater et al. 2015) will be required to obtain a better understanding of their phylogenetic affinities.

**Congruent genomic landscapes across a speciation continuum**

Patterns of within- and between-population genetic diversity in stonechats show high levels of parallelism across multiple scales of evolutionary divergence. We found outliers in comparisons
of highly similar taxa in the same regions as comparisons at deeper levels of divergence. The parallel reductions in $d_{XY}$ are highly suggestive of selection before divergence (Cruickshank and Hahn 2014), and analogous patterns in standardized nucleotide diversity ($\pi/d_{XY}$) indicate that common selective forces have continued to reduce diversity on the branches leading to present-day taxa (see Irwin et al. 2016).

Reductions in Fay and Wu’s $H$ in some outlier regions suggest that positive selection has played a role in driving some of these regions of low genetic diversity and high differentiation. Some $H$ outliers are present in multiple taxa, while others occur in only one (as in *Ficedula*, Burri et al. 2015), suggesting that localized selective sweeps may not have occurred in all groups, or that sweeps occurred too far in the past for detection using this method.

Pairwise comparisons that include Siberian stonechats show the most conspicuous $F_{ST}$ peaks, which coincide with regions of low within-population diversity ($\pi$). Together, strongly reduced within-population genetic diversity in specific genomic regions and corresponding peaks of differentiation are consistent with Siberian stonechats experiencing the strongest effects of selection in outlier regions. Most of the larger outlier regions also showed significant decreases in Fay and Wu’s $H$, suggesting that positive selective sweeps have contributed to this pattern. As temperate zone breeders and obligate long-distance migrants, Siberian stonechats are expected to generally show a faster pace of life, larger clutch sizes, and higher metabolic rates, along with a range of specializations associated with a strongly migratory lifestyle (Wikelski et al. 2003; Tieleman et al. 2009; Baldwin et al. 2010; Robinson et al. 2010). It is possible that a combination of these factors has led to a strong footprint of selection on the Siberian stonechat genome.

Kenyan and Canarian stonechats showed the highest genome-wide $F_{ST}$. Notably, however, we also observed conspicuous $F_{ST}$ valleys in the same locations as the $F_{ST}$ peaks of
other pairwise taxon comparisons (Fig. 5 and Fig. S5, Supporting information). In other words, these taxa are differentiated across the vast majority of the genome, but they show low differentiation ($F_{ST}$) in regions of low absolute divergence ($d_{XY}$). This pattern is not unique to this comparison: Austrian and Irish stonechats, and occasionally others, show valleys in similar areas. $F_{ST}$ valleys may occur where $d_{XY}$ (between-group variation) is reduced but $\pi$ (within-group variation) remains high, especially in Canary Islands stonechats, but more work is needed to understand this phenomenon.

Correlation of genomic variation between genera

Genomic landscapes of genetic diversity and differentiation in stonechats are significantly correlated with those in Pied and Collared Flycatchers. These results contrast with recent findings in other passerine birds, for example greenish warblers (Irwin et al. 2016). Nucleotide diversity in greenish warblers is only weakly correlated with that in outgroup comparisons ($\pi$: Pearson’s $r = 0.19$). We found a stronger association in nucleotide diversity between stonechats and flycatchers ($\pi$: Spearman’s $\rho = 0.47-0.60$, excluding Canary Is.). *Saxicola* and *Ficedula* share certain aspects of their life history (e.g., they are insectivores, and the flycatchers and most stonechats are migratory), but the hypothesis that these parallel signatures of selection and differentiation are due to shared ecological selection pressures on the same loci appears unlikely. Burri et al. (2015) demonstrated a clear link between low recombination and areas of high differentiation in *Ficedula*, which suggests that low recombination might also contribute to shared differentiation outliers within *Saxicola*. Although initial evidence suggested that avian recombination landscapes change drastically over time (Backström et al. 2010), recent work has shown that recombination landscapes can be conserved in birds across millions of years of evolution (Singhal et al. 2015). It
is therefore possible that coincident areas of low recombination, in combination with linked
selection, may play a role in shaping the broad patterns of landscapes of genomic variation and
differentiation across both closely related and deeply diverged taxa. However, direct measures of
recombination rates in stonechats are needed to test this hypothesis. While recombination is
reduced in close proximity to avian centromeres (Backström et al. 2010), centromeres do not
explain the recombination deserts in the centers of acrocentric chromosomes (e.g. 4A, 9, 10, 11,
12, 13, and 18; Knief and Forstmeier 2015) (Kawakami et al. 2014; Burri et al. 2015). These
regions frequently show high differentiation among flycatchers (Burri et al. 2015) and stonechats.

Decreases in Fay and Wu’s $H$ in a subset of outlier regions and a subset of taxa suggest
that positive selection has also contributed to this convergent genomic evolution. Indeed, Irwin et
al. (2016) favor positive selection as the likely driver of differentiation landscapes in greenish
warblers, citing exceedingly low nucleotide diversity in differentiation peaks; in one comparison
in that study, regions with $F_{ST} > 0.9$ showed just 6.7% the nucleotide diversity of regions with $F_{ST} < 0.6$. We found diversity reductions in stonechats and flycatchers to be less severe: between
Austrian and Siberian stonechats, which show the greatest reduction in nucleotide diversity in $F_{ST}$
peaks, $\pi$ in regions with $F_{ST}$ above the 95% percentile was reduced to 30-34% of that of regions
with $F_{ST}$ below the 50th percentile. In Ficedula flycatchers, this statistic was 43-50%. Therefore,
we consider background selection, in concert with reduced recombination, to be an additional
plausible driver of correlation in genomic landscapes.

Conserved variation in mutation rate is another possible driver of this correlation. Irwin et
al. (2016) found weak correlations in $d_{XY}$ between greenish warblers and more distant
comparisons (Pearson’s $r = 0.07-0.14$), which does not support this explanation. In contrast, we
found reasonably strong correlations in $d_{XY}$ between stonechat and flycatcher genera (Spearman’s
\[ \rho = 0.37-0.49 \]. Therefore, we cannot rule out a further contribution of conserved variation in mutation rate to these patterns.

Genomics and demography of Canary Islands stonechats

Canary Islands stonechats’ genomic landscapes differed from those of the other stonechats. The valleys of standardized nucleotide diversity (\( \pi/d_{XY} \)) seen in other taxa were completely absent, suggesting that selection has not reduced diversity across the genome in a heterogeneous way. In fact, this ratio was elevated in the same regions in which it was reduced in the other taxa. Tajima’s \( D \) was highly negative genome-wide and showed lower variance than in other stonechats. Combined, these results are most consistent with a demographic history that included a severe population bottleneck (erasing existing patterns of variation), followed by a substantial population expansion. Previous research has found evidence of founder effects and/or bottlenecks in Canary Island birds (Barrientos et al. 2009; Barrientos et al. 2014; Spurgin et al. 2014). The evidence for a bottleneck and expansion and the marked homogeneity of genetic diversity across the genome in Canary Islands stonechats suggest that genetic drift has played a dominant role in its divergence from other stonechats, possibly overpowering selection (see Hansson et al. 2014; Spurgin et al. 2014; Gonzalez-Quevedo et al. 2015; Illera et al. 2016). The unusual pattern seen in standardized nucleotide diversity may be explained by this prevalence of drift over selection. Because selection has not reduced \( \pi \) in the outlier regions shared by other stonechats, this statistic shows little variation across the genome of Canary Islands stonechats. This unusual pattern therefore results from the lack of a reduction in within-population diversity (\( \pi \)) in areas where between-population diversity (\( d_{XY} \)) is still reduced, presumably due to selection in the ancestral stonechat.
Evidence of lineage-specific evolution

Despite striking similarities in the genomic landscapes of stonechats, we also find lineage-specific evolution. At the broadest levels of our analysis, in which we compare genera, we identified conspicuous differentiation peaks that appear in *Ficedula* but not *Saxicola* (e.g., on chromosomes 3, 8, 10, 11, 12, 13, and 18; Fig. S14, Supporting information, shows chromosome 13), and vice versa (e.g., on chromosomes 6, 7, 17, and 20; Fig. S15, Supporting information shows chromosome 20). These outlier regions should be further examined from a functional perspective, as they appear to have resulted from evolutionary processes specific to a particular lineage. The most conspicuous outlier regions shared between these systems should likewise be examined (e.g., on chromosomes 1, 1A, 2, 3, 4, and 4A; Figs. S16 and S17, Supporting information, show chromosomes 1A and 4A).

Conclusion

Few former studies (Burri et al. 2015; Lamichhaney et al. 2015; Irwin et al. 2016; Vijay et al. 2016) have examined genome-wide patterns of differentiation in more than two avian taxa, yet comparative studies of closely related species have great potential to shed light on genome evolution (Cutter and Payseur 2013). We find parallel patterns of selection in the stonechat complex—likely occurring both before and after speciation—and evidence of demography potentially overwhelming signatures of selection in one species. In addition, this study suggests that parallel genomic processes are operating in independent evolutionary systems to drive the differentiation of similar genomic regions across genera. We hypothesize that linked selection coupled with areas of low recombination, which may be conserved across these taxa, have shaped
these broad patterns. Whether concordant outlier regions actually contribute to reproductive
isolation or are otherwise consequential in the speciation process is unknown. Therefore, we
recommend that outlier markers obtained through genome scans and their relevance to speciation
be interpreted with caution. Importantly, our comparative method also identified differentiation
outlier regions that are not widely shared; these may harbor loci important in lineage-specific
evolution and should be examined closely. As genomic comparisons among radiations
accumulate, we will be able to compare the congruence in genomic landscapes and potentially
reveal the phenomena that drive genomic differentiation over evolutionary time.

ACKNOWLEDGEMENTS

We thank the Max Planck Society, the Fuller Evolutionary Biology program at the Cornell Lab of
Ornithology, the Hunter R. Rawlings Presidential Research Scholars Program (Cornell), and the
dextra Undergraduate Research Endowment Fund (Cornell) for funding. We thank the
University of Washington Burke Museum, Heiner Flinks and Bea Apfelbeck for providing
samples, and the Canarian Government and the Cabildo of Fuerteventura for permission to
sample Canary Islands stonechats. We thank Heinke Buhtz, Conny Burghardt, Sven Künzel,
Nicole Thomsen, Mayra Zamora Moreno, and Bronwyn Butcher for guidance in the lab; David
Toews, Scott Taylor, Jacob Berv, and Kira Delmore for valuable discussion; Kelly Zamudio,
Jennifer Walsh, Kristen Ruegg, Sonya Clegg, Eric Anderson, and four anonymous reviewers for
thoughtful feedback on this manuscript; and Susan and Daniel Van Doren for valued support.
REFERENCES


**DATA ACCESSIBILITY**

The Siberian stonechat genome assembly is archived at the European Nucleotide Archive with accession number PRJEB19453. Pooled sequencing data from the five stonechat taxa are archived with accession number PRJEB19452. Both datasets will be made public upon final acceptance.

**AUTHOR CONTRIBUTIONS**

BVD, ML, LC, IJL, and BH designed the research. BH, ML, and JCI supplied samples. BVD and LC performed all analyses and wrote the paper with input from all authors.
Figure 1. (A) Maximum likelihood phylogenetic tree constructed with RAxML from fixed sites across the stonechat nuclear genome, with two Ficedula species used as outgroups. Branch labels denote bootstrap support from 100 rapid bootstrap iterations. This topology places Siberian stonechat (S. maurus) as the sister lineage to the remaining taxa, in contrast to previous trees based on mitochondrial DNA. Canary Islands stonechats are most closely related to European stonechats. Illustrations (males shown) are reproduced with permission from Handbook of Birds of the World Alive (Collar 2016a, 2016b). (B, C) Biplot of two principle coordinate axes derived from analyses of: (B) mean $F_{ST}$ and (C) mean $d_{XY}$. Axes are labeled with percent of variance explained.
Figure 2. Correlation of $d_{XY}$ among stonechats and flycatchers (*Ficedula hypoleuca* and *albicollis*, or “Hyp.” and “Alb.”). (A,B,C) show scatterplots where each point represents one 50-Kb genomic window. Orange lines are best-fit lines, and the Spearman’s rank correlation rho ($\rho$) coefficient is given. (D) shows outlier similarity scores, which quantify the number of low-$d_{XY}$ “valleys” shared among different comparisons. Some comparisons including Irish stonechats are not shown because of their similarity to Austrian stonechats. All tests were significant after applying a false discovery rate correction. Cells with yellow backgrounds indicate that four independent taxa are being compared. Letters in the upper right of cells show which cells correspond to the scatterplots in sections A-C.
Figure 3. Correlation of $F_{ST}$ among stonechats and flycatchers (*Ficedula hypoleuca* and *albicollis*, or “Hyp.” and “Alb.”). (A,B,C,D,F,G) show scatterplots where each point represents one 50-Kb genomic window. Orange lines are best-fit lines, and the Spearman’s rank correlation rho ($\rho$) coefficient is also given. (E) shows outlier similarity scores, which quantify the number of high-$F_{ST}$ “peaks” shared among different comparisons (upper triangle of matrix) and the number of low-$F_{ST}$ “valleys” shared among different comparisons (lower triangle of matrix). Some comparisons including Irish stonechats are not shown because of their similarity to Austrian stonechats. Cells with an ‘X’ indicate tests that were not significant after applying a false discovery rate correction. Cells with yellow backgrounds indicate that four independent taxa are being compared. Letters in the upper right of cells show which cells correspond to the scatterplots in the other sections.
Figure 4. Correlation of $F_{ST}$ and $d_{XY}$ among stonechats and flycatchers (*Ficedula hypoleuca* and *albicollis*, or “Hyp.” and “Alb.”). (A) shows outlier similarity scores, which quantify the number of low-$d_{XY}$ “valleys” that coincide with either high-$F_{ST}$ “peaks” (top row) or low-$F_{ST}$ “valleys” (bottom row). (B,C) show scatterplots where each point represents one 50-Kb genomic window. Refer to Figs. 2-3 for details.
Figure 5. Genomic statistics calculated across stonechat and flycatcher chromosomes 1A and 4A. Yellow and blue boxes indicated shared peaks and valleys, respectively. From top to bottom, the statistics and box details are: $F_{ST}$ among stonechats (peaks shared by 3 or more comparisons), $F_{ST}$ between flycatchers (peaks that also overlap with shared stonechats peaks), $d_{XY}$ among stonechats (valleys shared by 3 or more comparisons), $d_{XY}$ among flycatchers (valleys that also overlap with shared stonechats valleys), Tajima’s $D$ (valleys shared by 2 or more taxa), nucleotide diversity ($\pi$) (valleys shared by 2 or more taxa), and Fay & Wu’s $H$ (valleys shared by 2 or more taxa).
Figure 6. Correlation of $F_{ST}$ and $d_{XY}$ with Tajima’s $D$ and Fay & Wu’s $H$ among stonechats and flycatchers (*Ficedula hypoleuca* and *albicollis*, or “Hyp.” and “Alb.”). (A) shows outlier similarity scores, which quantify the number of high-$F_{ST}$ “peaks” that coincide with low Tajima’s $D$ (top section) and low Fay & Wu’s $H$ (bottom section). Within each section, the top (No. 1) and bottom (No. 2) rows show the results for each of the two taxa being compared. This is necessary because Tajima’s $D$ and Fay & Wu’s $H$ are single-population statistics, while $F_{ST}$ and $d_{XY}$ compare two populations. All comparisons were significant after applying a false discovery rate correction. (B,C) show scatterplots where each point represents one 50-Kb genomic window. Refer to Figs. 2-3 for details.
A. Phylogenetic tree showing the relationships between different subspecies of the Stonechat (S. torquatus). The tree is based on genetic distance (substitutions per site) and is rooted with the species Ficedula albicollis and F. hypoleuca.

B. Genetic diversity plot showing PCoA 1 and PCoA 2. The plot includes different locations such as Canary Islands, Siberia, Austria, and Ireland, with each location represented by a dot.

C. Genetic diversity plot showing PCoA 1 and d_{XY}. The plot includes different locations such as Canary Islands, Siberia, Austria, and Ireland, with each location represented by a dot.
### Low $d_{XY}$ Outlier Similarity

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyp. Alb.</td>
<td>0.48</td>
<td>0.47</td>
<td>0.53</td>
<td>0.53</td>
<td>0.52</td>
<td>0.55</td>
<td>0.53</td>
<td></td>
</tr>
<tr>
<td>Aus. Can.</td>
<td>0.84</td>
<td>0.94</td>
<td>0.86</td>
<td>0.86</td>
<td>0.84</td>
<td>0.86</td>
<td>0.86</td>
<td></td>
</tr>
<tr>
<td>Aus. Ire.</td>
<td>0.84</td>
<td>0.75</td>
<td>0.74</td>
<td>0.78</td>
<td>0.77</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aus. Ken.</td>
<td>0.84</td>
<td>0.93</td>
<td>0.87</td>
<td>0.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aus. Sib.</td>
<td>0.87</td>
<td>0.9</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ken. Can.</td>
<td>0.87</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ken. Sib.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.95</td>
</tr>
</tbody>
</table>

---

**A**

$\rho = 0.84$

$\rho = 0.88$

$\rho = 0.49$

Four different taxa compared
Hyp. & Alb.

<table>
<thead>
<tr>
<th></th>
<th>Hyp.</th>
<th>Alb.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aus.</td>
<td>0.33</td>
<td>0.22</td>
</tr>
<tr>
<td>Can.</td>
<td>0.27</td>
<td>0.31</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Aus.</th>
<th>Can.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ire.</td>
<td>0.42</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>0.25</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Aus.</th>
<th>Sib.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ken.</td>
<td>0.5</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>0.35</td>
<td>0.35</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Sib.</th>
<th>Can.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.8</td>
<td>0.81</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>High F_{ST}</th>
<th>Low F_{ST}</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Four different taxa compared

Not significant

\[
\rho = 0.51
\]

\[
\rho = 0.13
\]

\[
\rho = -0.04
\]

\[
\rho = 0.26
\]
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Low $d_{XY}$</td>
<td>0.51</td>
<td>0.29</td>
<td>0.56</td>
<td>0.42</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High $F_{ST}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.92</td>
</tr>
<tr>
<td>Low $d_{XY}$</td>
<td>0.17</td>
<td>0.44</td>
<td>0.45</td>
<td>0.56</td>
<td>0.08</td>
<td>0.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low $F_{ST}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.04</td>
</tr>
<tr>
<td>High $d_{XY}$</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.25</td>
</tr>
<tr>
<td>High $F_{ST}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure B**

- Scatter plot of $F_{ST}$ vs. $d_{XY}$ for Aus. & Sib.
- Correlation: $\rho = -0.3$

**Figure C**

- Scatter plot of $F_{ST}$ vs. $d_{XY}$ for Ken. & Can.
- Correlation: $\rho = 0.47$
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High F&lt;sub&gt;ST&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low D (1)</td>
<td>0.54</td>
<td>0.29</td>
<td>0.39</td>
<td>0.5</td>
<td>0.88</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low D (2)</td>
<td>0.33</td>
<td>0.46</td>
<td>0.56</td>
<td>0.38</td>
<td>0.72</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High F&lt;sub&gt;ST&lt;/sub&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low F &amp; W’s H (1)</td>
<td>0.38</td>
<td>0.26</td>
<td>0.22</td>
<td>0.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low F &amp; W’s H (2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Graphs:**

**B:**
- Scatter plot showing the relationship between mean Tajima's D and d<sub>XY</sub> with Aus. & Sib., with a linear trend line and correlation coefficient ρ = 0.59.

**C:**
- Scatter plot showing the relationship between mean Tajima's D and F<sub>ST</sub> with Aus. & Sib., with a linear trend line and correlation coefficient ρ = -0.32.
SUPPORTING INFORMATION

Correlated patterns of genetic diversity and differentiation across an avian family

Benjamin M. Van Doren, Leonardo Campagna, Barbara Helm, Juan Carlos Illera, Irby J. Lovette, and Miriam Liedvogel

ADDITIONAL METHODS

Draft Reference Genome
To extract DNA from Siberian stonechat (S. maurus) muscle tissue, we used the Gentra Puregene Tissue kit, following the protocol for fixed tissue. Gel electrophoresis revealed the DNA to be composed of highly intact molecules (all visible >10 kb). The ALLPATHS-LG algorithm (Gnerre et al. 2011) used 90.1% of the fragment library (total 506,475,396 reads), covering the genome at a mean depth of 46.9x. Combined, the two mate-pair libraries comprised 923,232,904 reads; ALLPATHS-LG used 20.7% of these reads, which covered the genome at 19.6x. This initial assembly required 147.50 hours on a 64-core computer with 512 GB of memory (1735.37 hours of CPU time). ALLPATHS-LG grouped 39,301 contigs into 4,396 scaffolds, with a total scaffold length of 1.027 Gb. The N50 scaffold size was 8.02 Mb, and 4.6% of bases were ambiguous (N’s).

We then used HaploMerger (Huang et al. 2012) to improve the assembly by merging homologous contigs and removing those that had arisen from the erroneous split of two haplotypes. HaploMerger requires “soft-masking” repetitive elements in the genome, which we did with RepeatMasker version open-4.0.2 (Smit et al. 2013-2015). HaploMerger has been used to improve a number of genome assemblies in this manner (e.g., Derks et al. 2015; Davey et al. 2016). After running the original assembly through the HaploMerger pipeline using default settings (and manually breaking two scaffolds that HaploMerger indicated may have been misjoined), the final Siberian stonechat de novo assembly comprised 2,819 scaffolds, with a total scaffold length of 1.020 Gb; the N50 scaffold size increased to 10.0 Mb compared to the original assembly. We verified that the majority of removed scaffolds had fragment library coverage of less than 5x. HaploMerger therefore appears to have been successful in removing a large number of small scaffolds that likely represented duplicates (i.e., heterozygous regions). The 1,577 removed scaffolds spanned only 7.4 Mb (0.7% of the original assembly).

To assess completeness of the reference genome, we used NCBI command-line ‘blastn’ to search for 5561 ultraconserved elements identified by Faircloth et al. (2012) from an analysis of chicken, anole, and zebra finch. The final assembly contained 5486 (98.7%) of these ultraconserved elements. Because they are interspersed throughout the entire genome, this percentage can be considered an approximation for the completeness of the draft assembly; a value of 98.7% is evidence that the assembly covers nearly the entire Siberian stonechat genome.

Sampling
Most birds used in this study originated from the common-garden stonechat study that Eberhard Gwinner initiated in 1981 at the Max-Planck Institute in Andechs, Germany.
Specifically, parental populations originated from the following locations: Austrian stonechats from Lower Austria (48°14'N, 16°22'E); Irish stonechats from Iveragh Peninsula near Killarney, in the County of Kerry, Ireland (c. 52°N, 10°W); African stonechats from Lake Nakuru region, Kenya (0°14'S, 36°0'E), and Mount Meru region, Tanzania (3°50'S, 36°5'E); and Siberian stonechats from the vicinity of Naursum National Park (c. 51.5°N, 63°E), Kazakhstan. All blood samples for the Canary Islands stonechat were collected directly in the field between 2013 and 2016 at various locations on Fuerteventura, Canary Islands, Spain (Barranco de Mal Nombre, Fimbapaire, Norte de Fenimoy, Barranco de Jacomar, Barranco Gran Valle, Barranco de Los Canarios, Barranco de Vinamar).

**Pooled sequencing**

We selected between 49 and 56 individuals (including both males and females) from each stonechat taxon based on a careful assessment of DNA quantity via Trinean DropSense 96 multi-channel spectrophotometer (Trinean, Ghent, Belgium) and quality (check for integrity on a 2% agarose gel), and we created one library of pooled DNA for each taxon following the Illumina TruSeq DNA kit. Each library included an equimolar aliquot of DNA from each individual. We multiplexed 4 of the 5 five groups on one lane of an Illumina NextSeq sequencer (151-bp paired end reads), and ran the fifth with three unrelated samples on a second lane. Thus, each group was sequenced on approximately one-fourth of a lane.

We demultiplexed raw sequence data from the sequencer with the ‘bcl2fastq’ utility by Illumina, using default settings. This utility generates ‘.fastq’ files after removing reads showing a 10% or greater error rate in the adapter sequence or more than 1 error in the barcode. It also masks adapter sequences extending into reads. We then used the program ‘skewer’ to conduct the following additional quality control measures on reads: trim the 3’ end until quality ≥ 20 is reached; and remove reads with normalized error rate > 0.1 (default), indel error rate > 0.03 (default, based on comparison with known adapter sequence), mean base quality < 20, or >15% ambiguous bases (N’s). Approximately 1% of the demultiplexed reads failed these criteria and were removed.

We used BWA-MEM (Li 2013) to align the pooled sequences to the reference genome, marking shorter split hits as secondary. We then converted the alignments to compressed BAM format using ‘samtools view,’ specifying a minimum mapping quality of 20. We sorted BAM files with ‘samtools sort’ and merged them across lanes using Picard’s ‘MergeSamFiles’ (Picard: http://broadinstitute.github.io/picard/). Following this, we marked duplicate reads using Picard’s MarkDuplicates utility; performed local realignment using the Genome Analysis Toolkit (GATK; RealignerTargetCreator and IndelRealigner) (McKenna et al. 2010; DePristo et al. 2011); and fixed mate information in Picard (FixMateInformation). Finally, we took the resulting 5 BAM files (one per taxon) and used ‘samtools mpileup’ to construct an mpileup file comparing the bases in overlapping reads at each position of the genome across populations.

Mapping quality for all stonechat taxa was high (Table S2). Mean mapping quality was lower for Ficedula species (hypoleuca: 35.64; albicollis: 35.82), but a high proportion of reads from these species were successfully mapped to the stonechat reference genome (hypoleuca: 0.93; albicollis: 0.95). This high mapping rate suggests that we are not introducing substantial bias by aligning flycatcher reads to the stonechat
reference genome. We detected bacterial DNA contamination in some of the stonechat pools; these sequences did not map to the reference and were thereafter ignored.

**Mapping to *Ficedula* chromosomes**

We assembled scaffolds from the stonechat assembly into draft chromosomes by mapping them to the *Ficedula albicollis* genome assembly, version 1.5 (RefSeq accession GCF_000247815.1; http://www.ncbi.nlm.nih.gov/assembly/GCF_000247815.1; http://www.ncbi.nlm.nih.gov/genome/?term=txid59894[orgn]) (Kawakami et al. 2014). We used SatsumaSynteny (Grabherr et al. 2010) to align the *Saxicola* draft genome to the *F. albicollis* assembly. This method unambiguously placed nearly all scaffolds of sufficient size (> 10 Kb) on a *F. albicollis* chromosome, with 85% of scaffolds mapping to single chromosomes across >70% of their extents. In the rare cases (~1%) where scaffolds mapped to more than one *Ficedula* chromosome across greater than 20% the scaffold length, we assigned the scaffold to the chromosome with the greatest amount of sequence aligned (always a majority of the scaffold). SatsumaSynteny thus allowed us to position scaffolds from the stonechat genome in the correct order and orientation along the chromosomes, assuming that synteny is conserved in these taxa (Ellegren 2013). This assumption appears robust given the high conservation of sequence within scaffolds. SatsumaSynteny successfully mapped 97.1% of the stonechat reference genome to a *Ficedula* chromosome. Most unmapped scaffolds had not passed the 10 Kb threshold.

**Coverage heterogeneity**

To rule out the possibility that variation in coverage could be driving differentiation patterns, we compared read depth in FST outlier regions to read depth outside of those regions. We selected the comparison of Irish and Siberian stonechats because this comparison showed arguably the most conspicuous FST peaks, and therefore any effect of coverage should be most pronounced. Because allele frequencies in adjacent 50 Kb windows are autocorrelated due to linkage and therefore contribute to pseudoreplication, we subsampled the genome at a ratio of 1:10. We used t-tests to test for differences inside and outside of outlier regions. Read depth was not significantly different within and outside of FST peaks for both taxa (Irish: \( t = 0.78, \) df = 118.37, \( P = 0.44 \); Siberian: \( t = 0.89, \) df = 119.56, \( P = 0.39 \)). Specifically, for Irish stonechats, mean within-peak coverage was 26.01 and mean outside-of-peak coverage was 26.38. For Siberian stonechats these values were 15.08 and 15.23, respectively.

**Phylogeny**

We aligned raw reads from Pied and Collared Flycatcher re-sequencing data to the stonechat genome in order to call genotypes. We then selected 16,876,859 sites across the genome which satisfied the following criteria: minimum coverage of 5 in all populations; fixation of a single allele at the locus (allowing a maximum count of 1 of another allele because of the possibility for sequencing error); and variation in the fixed allele among the 7 taxa. Using these SNPs, we generated a phylogenetic tree with RAxML v. 8.2.6 (Stamatakis 2014) on CIPRES (http://www.phylo.org). We applied the Lewis correction, following the recommendation of Stamatakis (2014), for ascertainment bias resulting from the exclusion of constant sites and using 100 bootstrapped replicates to assess branch support.
Choice of window size and bandwidth size for genome-wide scans

We used a window size of 50 Kb for our genomic analyses because it provided us sufficiently fine resolution across the genome while still averaging over hundreds of SNPs per window. We felt it was important not to rely heavily on allele frequencies of individual SNPs because of the random variation in allele frequencies introduced by our pooled sequencing approach. We conducted a sensitivity analysis (not shown) and found that we identified fewer and larger outlier regions as we increased window size, but that the level of overlap detected between genomic landscapes did not systematically vary. We feel that this justifies a window size of 50 Kb because it allows us to capture relatively small outlier regions while keeping the number of regions to a manageable size for this whole-genome analysis of multiple taxa.

We selected a bandwidth of 30 because it allowed us to identify relatively small regions of differentiation while still providing a benefit by smoothing out apparent noise in the data. We conducted a sensitivity analysis (not shown) and found that the median size of outlier regions identified by our analysis stayed relatively constant until a bandwidth of about 50, after which we observed an increase. Therefore, we do not believe that we are biased towards detecting large outlier regions by using a bandwidth of 30. We also did not observe any systematic effect of bandwidth size on the level of overlap detected between genomic landscapes.
SUPPLEMENTARY FIGURES

Figure S1. Boxplots comparing $F_{ST}$ and $\pi$ between the Z chromosome and autosomes. Outlier values are not shown. Under neutral expectation, the equilibrium level of neutral variability is proportional to the effective population size, and the effective population size of the Z chromosome is expected to be three-fourths that of the autosomes because females only have one copy (Charlesworth 2001). We used a $t$-test for ratios ($t.test.ratio$ function in the mratios package) to test whether the ratio of $\pi$ on the Z chromosome to $\pi$ on the autosomes was significantly different from 0.75 (Djira et al. 2012). Stonechats and flycatchers showed $\pi$ ratios between 0.74-0.83; Kenyan, Siberian and Canary Islands stonechats and Pied Flycatchers ($Ficedula hypoleuca$) had $\pi$ ratios that did not significantly differ from 0.75, while Austrian and Irish stonechats and Collared Flycatchers ($Ficedula albicollis$) showed slightly more diverse Z chromosomes than expected by theory. In all cases, $F_{ST}$ on the Z chromosome was elevated over that of the autosomes, with ratios in stonechats between 1.04-1.29, and a much higher ratio in flycatchers of 1.80.
Figure S2. Genome-wide landscape of d_{XY} for pairwise comparisons of stonechats and Pied and Collared Flycatchers (*Ficedula albicollis* and *F. hypoleuca*). All stonechat comparisons showed very similar genomic landscapes of d_{XY}. Many outlier regions were also shared with *Ficedula*, especially on the larger chromosomes. For clarity, comparisons including Irish stonechats are not included (with the exception of Austria-Ireland), because of the high degree of similarity between Austrian and Irish taxa. The colored lines are kernel-based density smoothers. Individual points represent 50-Kb windows; scaffolds alternate dark gray and light gray coloring. Chromosomes (based on alignment to *Ficedula albicollis*) are delineated by thick dark gray or light gray lines on the upper border of each plot and are labeled above this line. Z* indicates a flycatcher Z chromosome linkage group that could not be exactly placed in the flycatcher genome assembly. Shaded orange rectangles show d_{XY} peaks and blue rectangles show d_{XY} valleys.
Figure S3. $D_{XY}$ across stonechat chromosome 1A. All stonechat comparisons show very similar fluctuations, including two pronounced valleys. The largest valley is also apparent in the comparisons of Pied and Collared Flycatchers (*Ficedula albicollis* and *F. hypoleuca*). Blue rectangles indicate significant $d_{XY}$ valleys. See Figure S2 for other details. For clarity, comparisons including Irish stonechats are not included (with the exception of Austria-Ireland) because of the high degree of similarity between Austrian and Irish populations.
Figure S4. Correlation of high $d_{XY}$ regions among stonechats and flycatchers (*Ficedula hypoleuca* and *albicollis*, or “Hyp.” and “Alb.”). Matrix shows outlier similarity scores, which quantify the number of high-$d_{XY}$ “peaks” shared among different comparisons. Some comparisons including Irish stonechats are not shown because of their similarity to Austrian stonechats. All tests were significant after applying a false discovery rate correction. Cells with yellow backgrounds indicate that four independent taxa are being compared.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyp.</td>
<td>0.46</td>
<td>0.43</td>
<td>0.3</td>
<td>0.43</td>
<td>0.3</td>
<td>0.39</td>
<td>0.39</td>
<td>0.39</td>
<td>0.39</td>
<td>0.39</td>
<td>0.39</td>
<td>0.39</td>
<td>0.39</td>
<td>0.39</td>
</tr>
<tr>
<td>Alb.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aus.</td>
<td>0.71</td>
<td>0.54</td>
<td>0.85</td>
<td>0.77</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>Can.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aus.</td>
<td>0.57</td>
<td>0.71</td>
<td>0.85</td>
<td>0.77</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>Ire.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aus.</td>
<td>0.96</td>
<td>0.79</td>
<td>0.96</td>
<td>0.96</td>
<td>0.83</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ken.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aus.</td>
<td>0.97</td>
<td>0.98</td>
<td>0.91</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sib.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Can.</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Can.</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Four different taxa compared.
Figure S5. Genome-wide landscape of $F_{ST}$ for pairwise comparisons of stonechats and Pied and Collared Flycatchers (Ficedula albicollis and F. hypoleuca). Pairs including Siberian stonechats showed the most conspicuous peaks; other comparisons showed less distinct outlier regions. Some comparisons (e.g., Kenya-Canary Is.) showed $F_{ST}$ valleys in the same regions as the $F_{ST}$ peaks of other comparisons. Shaded orange rectangles show $F_{ST}$ peaks and blue rectangles show $F_{ST}$ valleys. See Figure S2 for other details.
Figure S6. Correlation of nucleotide diversity ($\pi$) and $d_{XY}$ among stonechats and flycatchers (*Ficedula hypoleuca* and *albicollis*, or “Hyp.” and “Alb.”). (A) shows outlier similarity scores, which quantify the number of low-$d_{XY}$ “valleys” that coincide with low $\pi$ “valleys” in the two taxa being compared (top row and bottom row). The top (No. 1) and bottom (No. 2) rows show the results for each of the two taxa being compared. This is necessary because $\pi$ is a single-population statistic, while $F_{ST}$ and $d_{XY}$ compare two populations. All comparisons were significant after applying a false discovery rate correction. (B,C,D) show scatterplots where each point represents one 50-Kb genomic window. Refer to Figs. 2-3 for details.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Low $d_{XY}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low $\pi$ (1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.75</td>
<td>0.88</td>
<td>0.94</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>0.73</td>
<td>0.55</td>
<td>0.96</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
</tbody>
</table>

**Graphs:***

- **B**: $d_{XY}$ Aus. & Ire. vs mean $\pi$, $\rho = 0.97$
- **C**: $d_{XY}$ Aus. & Sib. vs mean $\pi$, $\rho = 0.95$
- **D**: $d_{XY}$ Ken. & Can. vs mean $\pi$, $\rho = 0.7$
Figure S7. Correlation of nucleotide diversity ($\pi$) and standardized nucleotide diversity ($\pi/d_{XY}$) among stonechats and flycatchers (*Ficedula hypoleuca* and *albicollis*, or “Hyp.” and “Alb.”). (A,B,C,D,F,G) show scatterplots where each point represents one 50-Kb genomic window. (E) shows outlier similarity scores, which quantify the number of $\pi/d_{XY}$ valleys shared among different comparisons (upper triangle of matrix) and the number of $\pi$ valleys shared among different comparisons (lower triangle of matrix). Refer to Figs. 2-3 for details.
<table>
<thead>
<tr>
<th></th>
<th>$\pi / d_{XY}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alb.</td>
<td>0.85 0.39 X X 0.39 0.34 0.63</td>
</tr>
<tr>
<td>Hyp.</td>
<td>0.79 0.33 X X 0.31 0.27 0.46</td>
</tr>
<tr>
<td>Aus.</td>
<td>X X X 0.93 0.7 0.75</td>
</tr>
<tr>
<td>Can.</td>
<td>X X X X X X X</td>
</tr>
<tr>
<td>Ire.</td>
<td>0.48 0.54 0.94 0.68 0.67 0.77</td>
</tr>
<tr>
<td>Ken.</td>
<td>0.48 0.52 0.75 0.56 0.74 0.67</td>
</tr>
<tr>
<td>Sib.</td>
<td>0.56 0.57 0.81 0.61 0.79 0.73</td>
</tr>
</tbody>
</table>

$\rho = 0.77$

$\rho = 0.34$

$\rho = 0.55$

$\rho = -0.36$

$\rho = 0.51$

$\rho = 0.6$
Figure S8. Genome-wide landscape of $\pi$ for five stonechat taxa. Canary Islands stonechats generally did not share the valleys present in the genomes of the other taxa. See Fig. S2 for other details.
Figure S9. Genome-wide landscape of standardized nucleotide diversity ($\pi/d_{XY}$) for five stonechat and two flycatcher taxa. Canary Islands stonechats did not share the valleys present in the genomes of the other taxa. See Fig. S2 for other details.
Figure S10. Correlation of nucleotide diversity ($\pi$) with Tajima’s $D$ and Fay & Wu’s $H$ stonechats. Shown are outlier similarity scores, which quantify the number of low Fay & Wu’s $H$ “valleys” that coincide with low $\pi$ (top section) and low Tajima’s $D$ (bottom section). Refer to Figs. 2-3 for details.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Low $\pi$</td>
<td>0.79</td>
<td>×</td>
<td>0.71</td>
<td>0.6</td>
<td>0.74</td>
</tr>
<tr>
<td>Low $F &amp; W's H$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Tajima’s $D$</td>
<td>0.6</td>
<td>×</td>
<td>0.44</td>
<td>0.4</td>
<td>0.52</td>
</tr>
<tr>
<td>Low $F &amp; W's H$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure S11. Genome-wide landscape of Tajima’s $D$ for five stonechat and two flycatcher taxa. Austrian, Irish, Siberian, and Kenyan stonechats shared similar genomic landscapes of Tajima’s $D$. Canary Islands stonechats showed a different pattern, with very low Tajima’s $D$ across the entire genome. See Fig. S2 for other details.
Austria

Ireland

Kenya

Siberia

Canary

F. hyp.

F. alb.
Figure S12. Genomic landscape of Fay and Wu’s $H$ for each of five stonechat taxa. See Fig. S2 for other details.
Figure S13. Boxplot of Tajima’s $D$ for five stonechat taxa. Each data point represents one 50-Kb window. Canary Islands stonechats showed the lowest median Tajima’s $D$. 
Figure S14. $F_{ST}$ across stonechat chromosome 13, including Pied and Collared Flycatchers (*Ficedula albicollis* and *F. hypoleuca*). The *Ficedula* comparison shows a distinct peak, which is not present in any stonechat comparison. This suggests that the evolutionary processes driving differentiation in this chromosome are potentially unique to *Ficedula*. See Fig. S2 for other details.
Austria & Ireland

Austria & Canary

Austria & Kenya

Kenya & Canary

Siberia & Canary

Kenya & Siberia

Austria & Siberia

F. hypoleuca & F. albicollis
Figure S15. $F_{ST}$ across stonechat chromosome 20, including Pied and Collared Flycatchers (*Ficedula albicollis* and *F. hypoleuca*). Comparisons including the Siberian population show a distinct peak at the right end of the chromosome. Notably, this peak is absent in *Ficedula*, suggesting that the evolutionary processes driving divergence in this chromosome are potentially unique to the stonechat radiation. Also note that the Kenya-Canary comparison shows a valley in the center of the chromosome, where there is a peak in other comparisons. See Fig. S2 for other details.
Figure S16. $F_{ST}$ across stonechat chromosome 4A, including Pied and Collared Flycatchers ($Ficedula albicollis$ and $F. hypoleuca$). Comparisons including Siberian stonechats, $Ficedula$, and Austria-Ireland show a distinct peak. The other comparisons show a valley in the same region. See Fig. S2 for other details.
Figure S17. $F_{ST}$ across stonechat chromosome 1A, including Pied and Collared Flycatchers (*Ficedula albicollis* and *F. hypoleuca*). Comparisons including Siberian stonechats and *Ficedula* show a distinct peak. Other comparisons show a valley in the same region. See Fig. S2 for other details.
**SUPPLEMENTARY TABLES**

Table S1. Origin, sex, and relatedness information of stonechats included in this study. Kinship matrices were calculated with the `kinship` function in the R package `kinship2` (Therneau and Sinnwell 2015) using a pedigree of captive stonechats, and values presented are the mean (+SD) values from each kinship matrix. Inbreeding coefficients were calculated with the `calcInbreeding` function in the R package `pedigree` (Coster 2013). IQR stands for interquartile range.

<table>
<thead>
<tr>
<th>Origins</th>
<th>Direct from wild</th>
<th>Hatched in captivity</th>
<th>Sex</th>
<th>Male</th>
<th>Female</th>
<th>Mean kinship</th>
<th>SD kinship</th>
<th>Median kinship</th>
<th>IQR kinship</th>
<th>Mean inbreeding</th>
<th>SD inbreeding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Austria</td>
<td>1</td>
<td>48</td>
<td>27</td>
<td>22</td>
<td>0.014</td>
<td>0.045</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.006</td>
<td>0.018</td>
</tr>
<tr>
<td>Ireland</td>
<td>27</td>
<td>27</td>
<td>26</td>
<td>28</td>
<td>0.009</td>
<td>0.043</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.002</td>
<td>0.012</td>
</tr>
<tr>
<td>Kenya</td>
<td>1</td>
<td>50</td>
<td>18</td>
<td>33</td>
<td>0.009</td>
<td>0.039</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.012</td>
<td>0.05</td>
</tr>
<tr>
<td>Siberia</td>
<td>0</td>
<td>52</td>
<td>30</td>
<td>22</td>
<td>0.033</td>
<td>0.064</td>
<td>0</td>
<td>0.0625</td>
<td>0.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canary</td>
<td>56</td>
<td>0</td>
<td>38</td>
<td>18</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
Table S2. Summary of alignment of Illumina 150-bp reads from five stonechat taxa to the draft reference genome. Mapping quality is given after filtering out alignments with a mapping quality of 20 or lower.

<table>
<thead>
<tr>
<th>Taxon</th>
<th>Reads Mapped</th>
<th>Mean (Median) Coverage</th>
<th>Mean Mapping Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kenya</td>
<td>98,758,285</td>
<td>13.8 (12.7)</td>
<td>45.61</td>
</tr>
<tr>
<td>Ireland</td>
<td>185,976,416</td>
<td>26.1 (24.8)</td>
<td>45.26</td>
</tr>
<tr>
<td>Austria</td>
<td>135,110,173</td>
<td>18.8 (17.7)</td>
<td>45.14</td>
</tr>
<tr>
<td>Siberia</td>
<td>107,623,583</td>
<td>14.9 (13.9)</td>
<td>45.80</td>
</tr>
<tr>
<td>Canary Islands</td>
<td>176,167,216</td>
<td>24.7 (23.9)</td>
<td>45.64</td>
</tr>
</tbody>
</table>
Table S3. *Ficedula* individuals included in the study (data from the Sequence Read Archive, or SRA, project ERP007074, published in Smeds et al. (2015)).

<table>
<thead>
<tr>
<th>Species</th>
<th>SRA Run</th>
<th>Mean coverage of <em>Ficedula albicollis</em> genome</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>F. hypoleuca</em></td>
<td>ERR637490</td>
<td>16x</td>
</tr>
<tr>
<td></td>
<td>ERR637491</td>
<td>12x</td>
</tr>
<tr>
<td></td>
<td>ERR637492</td>
<td>11x</td>
</tr>
<tr>
<td></td>
<td>ERR637493</td>
<td>11x</td>
</tr>
<tr>
<td></td>
<td>ERR637494</td>
<td>13x</td>
</tr>
<tr>
<td></td>
<td>ERR637495</td>
<td>16x</td>
</tr>
<tr>
<td></td>
<td>ERR637496</td>
<td>18x</td>
</tr>
<tr>
<td></td>
<td>ERR637501</td>
<td>14x</td>
</tr>
<tr>
<td></td>
<td>ERR637503</td>
<td>14x</td>
</tr>
<tr>
<td></td>
<td>ERR637504</td>
<td>12x</td>
</tr>
<tr>
<td><em>F. albicollis</em></td>
<td>ERR637505</td>
<td>14x</td>
</tr>
<tr>
<td></td>
<td>ERR637506</td>
<td>14x</td>
</tr>
<tr>
<td></td>
<td>ERR637508</td>
<td>10x</td>
</tr>
<tr>
<td></td>
<td>ERR637511</td>
<td>16x</td>
</tr>
<tr>
<td></td>
<td>ERR637512</td>
<td>15x</td>
</tr>
<tr>
<td></td>
<td>ERR637513</td>
<td>14x</td>
</tr>
<tr>
<td></td>
<td>ERR637515</td>
<td>11x</td>
</tr>
<tr>
<td></td>
<td>ERR637519</td>
<td>12x</td>
</tr>
<tr>
<td></td>
<td>ERR637522</td>
<td>11x</td>
</tr>
<tr>
<td></td>
<td>ERR637523</td>
<td>13x</td>
</tr>
</tbody>
</table>


