Avian Population Genomics: Latest Findings and Future Prospects

Georg Langebrake[1], Kira Delmore[2,3], Miriam Liedvogel[1,2]

GL – ORCID 0000-0001-8114-5962

KD – ORCID 0000-0003-4108-9729

ML - ORCID 0000-0002-8372-8560

[1] Institute of Avian Research 'Vogelwarte Helgoland', Wilhelmshaven, Germany

[2] MPRG Behavioural Genomics, Max Planck Institute for Evolutionary Biology, Plön, Germany

[3] Columbia University, New York NY, USA

Corresponding author: Georg Langebrake, georg.langebrake@ifv-vogelwarte.de

Abstract

Birds are one of the most recognizable and diverse groups of organisms on Earth. This group has played an important role in many fields, including the development of methods in behavioral ecology and evolutionary theory. The use of population genomics took off following the increased accessibility of high-throughput sequencing across taxa. Several features of bird genomes make them particularly amenable to these approaches, including their nucleated red blood cells, permitting easy DNA extraction, and their small, compact genomes. Here we review the latest findings in the population genomics of birds, emphasizing questions related to behavior, ecology, evolution, and conservation. Additionally, we include insights in trait mapping and the ability to obtain accurate estimates of important summary statistics for conservation (e.g., genetic diversity and inbreeding). We highlight current topics in population

genomics that are starting to be adopted in bird studies, such as the use of ancestral recombination graphs, a focus on more than just single-nucleotide polymorphisms, and interpreting results in a polygenic framework in general. Prospects also include integrating population genomics with other fields.

Introduction

Birds have played a central role in our understanding of many research areas. Notable examples include (1) the development of methods essential for behavior and ecology by Margaret Nice [1] using populations of song sparrows (*Melospiza melodia*), and (2) the definition of species as groups of populations reproductively isolated from one another by Ernst Mayr [2], inspired by the geographic distribution of birds and galvanizing the field of evolutionary biology. Current declines in natural populations of birds worldwide have caused an increase in conservation efforts, further broadening research on birds. Concern about the loss of birds to the millinery trade triggered Harriet Hemenway and Mina Hall to establish one of the first nonprofit environmental organizations, the National Audubon Society in 1886. Together with many international partners through BirdLife International, they produce datasets essential for population monitoring and forecasting to this day (e.g., annual Christmas and Great Backyard Bird Counts to census birds worldwide and eBird, an online database of bird observations). These local efforts remain important and are increasingly coupled with conservation measures on a global scale through policymakers, such as the UN Biodiversity Conference (UNEP-CBD 2022), a landmark event focused on global biodiversity conservation. The conference resulted in the adoption of a new set of goals to guide global action on nature through 2030, aiming to halt and reverse nature loss.

The influential role of birds across varied research fields continued with the development of population genetics approaches. This field emerged in the 1980s following the rapid advancement of sequencing technologies to quantify marker-based genetic variation,

including mitochondrial DNA (mtDNA) sequence variation and microsatellite length polymorphisms. Population genetic approaches can be used to evaluate the role that mutation, genetic drift, selection, and gene flow play in generating variation within and between populations [3, 4], with relevance to behavior, ecology, evolution, and conservation. Avian blood has nucleated red blood cells, making it ideally suited for minimally invasive sampling, efficient extraction of high-integrity genomic DNA, and subsequent population genetic analyses. One early example is the use of mtDNA to identify taxonomic units for conservation (e.g., dusky seaside sparrows [*Ammodramus maritimus nigrescens*], [5, 6]).

The last two decades have seen a change in both the scale and depth of genetic analyses, with the transition from the use of one or a few genetic markers to tens of thousands of markers genome-wide marking the development of population genomics. This transition was stimulated by de novo assembly of reference genomes along with the development of high-throughput sequencing (HTS). HTS is discussed in detail elsewhere (e.g., [7, 8]), but briefly, is a set of platforms that sequence DNA from multiple genomic regions and individuals in parallel. HTS has increased the proportion of the genome that can be sampled and decreased the time and cost of sequencing, allowing its use on most organisms of interest. Again, birds are well-suited for this extension. Not only can good amounts of high-quality genomic DNA be generated easily, but birds across the entire clade also have small (mean ~1.45 billion base pairs; [9]), compact (e.g., fewer transposable elements [TEs] and repetitive regions; [10]) genomes, allowing for relatively easy genome assembly and mapping at high coverages. The first avian reference genomes were the chicken (*Gallus gallus*, [11]) and zebra finch (*Taeniopygia guttata*, [12]), assembled using Sanger sequencing and followed by an explosion of bird genomes assembled using HTS [13], most notably spearheaded by the Bird 10,000 Genomes (B10K) Project, a consortia initiative to generate representative draft genome sequences from all extant bird species [14, 15].

Here, we review the population genomics in birds, emphasizing the role this group has played in both the development of this field and its application to questions in behavior, ecology,

evolution, and conservation. Population genomics can be used to answer questions at both the genome and locus levels. Work at the genome level informs our understanding of population processes (e.g., demography and population structure), while work at the locus level helps identify genomic regions affected by mutation, drift, selection, and/or gene flow [3]. Our chapter follows this division, introducing some of the latest findings from birds, highlighting the benefits of applying population genomics tools to these questions (vs. traditional population genetic techniques with fewer markers), and finishes by outlining future prospects along this trajectory.

Relevance of Genomic Insight for Evolution

Demography is the study of changes in effective population size (Ne) through time, gene flow and divergence, reflecting the evolutionary history of a species or population. Information on these dynamics is essential for understanding the evolution of species, populations, and traits, and is important for setting baselines beyond which evolutionary processes can be examined. This is especially true in the current literature, where genome scans are being used to identify loci associated with phenotypic traits and/or involved in adaptation and speciation (see below; e.g., [16–19]). Early demographic analyses were based on the coalescent (or divergence), assuming that pairwise sequence divergence is proportional to the time of the coalescent, and relying on non-recombining loci (e.g., mtDNA, [20]). These analyses were expanded to include nuclear loci and permit estimates of changes in population size and gene flow, but remained limited to a small number of demographic scenarios and markers and were computationally expensive [21]. The availability of genome-wide data and new tools for their analysis has revolutionized this field and its progression is demonstrated well by a series of studies on collared and pied flycatchers (*Ficedula albicollis* and *F. hypoleuca*).

Early demographic work with collared and pied flycatchers suggested these species diverged during the Pleistocene, expanded their ranges following the last glacial maximum, and came

into secondary contact in Central Europe, where introgression from pied into collared flycatchers is greater [22, 23]. Nadachowska-Brzyska et al. [24] used whole genome resequencing (WGS) data to expand on these findings, comparing 15 demographic models using an Approximate Bayesian Computation (ABC) approach. Two models with recent divergence time (147,000-476,000 years ago [ya]), gene flow (0.16–0.36 pied individuals per generation), and recent reductions in population size provided the best fit to the data. These findings were further validated using the Pairwise Sequentially Markovian Coalescent (PSMC, Li and Durbin [25]; similar to the model described in Chapter 6) on a single collared flycatcher individual. This method can estimate historical changes in effective population size and shows a similar decline as in the ABC analysis in historic population size for collared flycatchers after the estimated split time from the pied flycatcher. This work represents one of the first times an ABC approach with HTS was applied to study demography in a non-model system, providing a level of detail unattainable with earlier methods [24]. Flycatchers exhibit strong reproductive isolation, with evidence for both pre- and postzygotic barriers to gene flow (e.g., nearly complete female sterility, [26]). Accordingly, these results suggest speciation in birds can occur quite rapidly.

These results were later corroborated by another study on the flycatcher complex, this time including the Atlas- and semicollared flycatcher (*F. speculigera* and *F. semitorquata*). Nater et al. [27] used an ABC analysis as well as fastsimcoal2 (a coalescence-based approach that uses site frequency spectra to estimate parameters, [28–30]) and estimated a slightly later divergence time (386,000–888,000 ya), but found further evidence for gene flow between pied and collared flycatcher using a gene-tree-based approach.

Continuously decreasing sequencing costs and improvements of computational methods have facilitated the application of multiple sequentially Markovian coalescent approaches (MSMC/MSMC2, see Chapter 6 and [31]). This method uses information from multiple individuals to more accurately infer population sizes in recent times and also allows for the calculation of cross-coalescence rates, thereby providing an alternative way of estimating

divergence time. MSMC methods have now been successfully applied to birds, giving insight into historic population size developments and population splits in a number of different species (for some examples, see [29–34]).

Many of these studies have confirmed earlier work using data from 38 bird species [32], where many species show population contractions during glacial periods and expansions in between (Fig. 1a). Together these results provide important inferences about the population dynamics of temperate species that have experienced glacial cycles throughout their history and caution against assuming simple demographic history (e.g., constant population sizes, a single expansion, and/or similar trends for populations).

While now possible, the study of demography in non-model organisms does not come without challenges. Nadachowska-Brzyska et al. [33] conducted a thorough analysis of the impact of sequence coverage and missing data on the accuracy of demographic inference with PSMC. They conclude that individuals with less than 18x coverage or more than 25% missing data should be excluded from PSMC analyses. These kinds of considerations are important when implementing analyses with genomic data, as biases arising from poor filtering can significantly impact results. In the case of PSMC, low filtering cutoffs lead to homozygous sites being considered heterozygous, affecting the size of recombination blocks and estimates of Ne (both magnitude and the shape of curves). Additionally, Ishigohoka et al. [34] investigated the impact of high-recombining regions (a common feature in avian genomes, and thus of focal importance here) on demography inference using methods based on the ancestral recombination graph. They found that the inclusion of high-recombining regions affects both the inference of population size history as well as split times, and advise filtering out high-recombining regions when performing demography analyses.

Relevance of Genomic Insight for Conservation

Demographic analyses provide us with reliable information on historical population dynamics. Population genomics can also be used to understand dynamics in present-day populations and are especially relevant for conservation (e.g., helping inform the management of threatened species). The application of these tools is critical to study genetic factors that can compound reductions in population size already experienced by species of concern, potentially aiding in the transfer of findings from common species to those under threat. Here, the availability of HTS allows researchers to obtain far more accurate and precise estimates not only of population structure and dynamics, but also to assess loss of genetic diversity and inbreeding. Using a comparative framework and genome sequences spanning nearly the full phylogenetic spectrum of birds, Li et al. [35] highlighted the potential of these data for conservation. We will discuss this work below, but want to emphasize the importance of the application of population genomics to these questions for birds, where 1,480 out of 11,195 species (13%) are threatened with extinction (numbers retrieved 12/2025 from IUCN red list https://www.iucnredlist.org/). Habitat loss is among the main threats, and the ramifications of extinction in birds will be far-reaching, as they are essential for ecosystem functioning (e.g., as seed dispersers) and serve as important indicators of ecosystem health (e.g., tracking changes in habitat, water, and climate).

Li et al. [35] is one of many papers published based on the 48 reference genomes assembled using HTS during the first phase of the Bird 10k genome project. The authors classified each species as endangered/vulnerable (E/V) or of no conservation concern, and observed that E/V species exhibited lower levels of heterozygosity and more nonsynonymous (and potentially deleterious) mutations compared to species of no conservation concern. The authors showcase the critically endangered crested ibis, a species formerly widespread across North-East Asia, with drastic declines down to only seven individuals from two breeding pairs in 1981. Every individual from the recovering population was sampled for genomic and demographic studies, providing the perfect material for conservation genomics analyses (Fig.

1b). The crested ibis population shows a slow linkage disequilibrium decay, a pattern also observed in highly inbred domestic species. Nonsynonymous mutations were associated with increased linkage disequilibrium (see also Chapter 1) across the genome. Combined, these findings suggest E/V species may be at risk of inbreeding depression. Estimates of inbreeding can also be obtained using pedigrees in the form of inbreeding coefficients (the probability of a locus being identical by descent). Nevertheless, as noted, pedigrees are rare in wild bird populations, so the application of HTS in the framework used by Li et al. [35] will be imperative for understanding risks to natural populations of birds [36].

More recently, researchers have started to use population genomics methods to assess/predict/estimate the effect of future climate change on currently healthy populations. Bay et al. [37] calculate genomic vulnerability, which uses current environmental adaptations and predicted environmental changes from a climate model to estimate the adaptive potential of the existing populations. These findings highlight that many populations exhibit high genomic vulnerability and face drastic threats under the current climatic developments. To validate their findings, the authors hypothesized that populations with high genomic vulnerability were already under pressure from climate change in recent years. As expected, they found a significant correlation between genomic vulnerability and population trend estimates.

While genomics can assess threat levels, it can also help to directly advise conservation measures. The genoscape project is a large-scale effort to connect avian subspecies-specific breeding grounds with non-breeding grounds and use this framework to inform on which areas require the most attention. The genoscape of a bird is a panel of genetic variants that are informative of their breeding range, such that individuals caught on non-breeding grounds can be traced back to their breeding origin. This has been successfully applied to, e.g., the willow flycatcher [38]. New analytical approaches based on machine learning have further refined accuracy [39]. These methods are especially powerful for moving animals, such as birds in general and migratory birds in particular, as they allow us to infer breeding and non-breeding

grounds of individuals caught during migration, which is extremely relevant for conservation efforts.

Locus-Level Work to Examine the Genetic Basis of Phenotypic Traits

The application of population genomics to questions of demography and conservation is an example of genome-wide analyses. Population genomics can also be employed at the locus level and includes bottom-up approaches to examine the genetic basis of phenotypic traits, including morphological and behavioural traits. These analyses involve the identification of populations that differ in a trait of interest with a strong genetic component and the estimation of summary statistics along the genomes of these populations to detect selective sweeps (see also Chapters 1 and 13). Selective sweeps can derive from positive or divergent selection, and evidence for these events includes reductions in nucleotide diversity and increased linkage disequilibrium within populations. Elevated differentiation between populations also provides evidence for selective sweeps [40, 41]. The number of studies examining the genetic basis of phenotypic traits in birds is increasing, and one emerging pattern is the importance of supergenes in maintaining complex phenotypes. A supergene is a genomic region that has a very low recombination rate. Variants located in a supergene are inherited together and, through that, can cause multiple trait changes to coevolve. One common example of supergenes are inversions. Inversions are a special form of rearrangement in which portions of a chromosome are flipped (inverted), disrupting chromosome pairing during meiosis and preventing recombination.

One well-known example of an inversion associated with a phenotypic trait in birds comes from the ruff (*Philomachus pugnax*) where one inversion controls the expression of different reproductive morphs (Fig. 1c, [42, 43]). Additional examples are beginning to accumulate and include the white-crowned sparrow (*Zonotrichia leucophrys gambelii*, [44]), the common quail (*Coturnix coturnix*, [45]), and the redpoll (*Acanthis flammea,* [46]). Tuttle et al. [44]used WGS to identify a series of potential inversions on avian chromosome 2 between two color morphs of white-crowned sparrows that also differ in reproductive behavior (promiscuous vs.

monogamous). This region spans ~100 Mb and includes 1100 genes. FST between morphs is elevated in this region, and linkage disequilibrium is high. One morph is homozygous for alleles at this inversion, and the other is heterozygous. The authors used a phylogenetic analysis to show that the inversion evolved before sparrows diverged from their most closely related relative.

In the quail, the inversion found spans a large proportion of one of the major chromosomes and links many genes that together alter their behaviour, morphology, and coloration [45].

The redpoll finch, previously considered to comprise three separate species, is now classified as a single species by a new study that reveals that the phenotypic differences are solely controlled by a single 55 Mb inversion, with gene flow homogenising the rest of the genome [46]. The inversion here seems to encode both coloration and bill shape, explaining the strong visual differentiation between the different redpolls.

Another type of genomic variant that is increasingly gaining recognition is changes in regulatory elements and not necessarily coding genes themselves. One special case of this is transposable elements (TE). As TEs rely on the transcription machinery of the host organism to self-replicate, they often contain highly active promoters. As a result, their proximity to genes may alter gene transcription [47]. Recently, a regulatory element that may encode migratory direction in the willow warbler was identified. In the willow warbler, Lundberg et al. [48] and Caballero-López et al. [49] identified two inversions and one repeat-rich region (termed “Migration Associated Repeat Block”, MARB) that jointly differentiate two distinct subspecies with different migratory strategies. Sokolovskis et al. [50] then genotyped and tracked individuals in the area where both subspecies co-occur and hybridise, and conclusively demonstrated that migratory direction is largely explained by the number of MARB repeats. While the function of MARB and its location in the genome remains unclear, the repetitive nature and very clear correlation with migratory behaviour suggest that it mainly causes regulatory changes [50].

In a case where the involvement of TEs in phenotype variation is understood in more detail, Lutgen et al. [51] have found that a long terminal repeat (LTR) together with a nonsynonymous mutation is perfectly associated with throat coloration in *Oenanthe hispanica* and *O. melanoleuca*. The authors suggest that the LTR acts as an enhancer of the ASIP gene, which is known to be involved in the melanogenesis pathway. The authors investigate this more in depth in the full *hispanica* complex, involving two additional species, which we will get back to with more detail in a later section.

There are several benefits of using a bottom-up approach of population genomics to study the genetic basis of phenotypic variation. Early work on this topic was limited to a set of candidate genes that were often identified in model organisms only distantly related to the focal species (e.g., [52]). Work with HTS allows researchers to study the entire genome, permitting unbiased assays of genomic variation and allowing for the *de novo* identification of candidate regions and novel biological processes underlying trait variation. This genome-wide perspective also provides a broader understanding of how phenotypic traits are controlled, including the number, size, and distribution of loci that underlie these traits. The bottom-up approach also has considerably greater power than other methods. For example, genome-wide association studies (GWAS) can be used, but often require data from hundreds of individuals as they are conducted in single populations that vary in a trait of interest and have low levels of linkage disequilibrium. Bottom-up approaches can use as few as 10 individuals per population, which can be important for non-standard study systems like birds, where large numbers of individuals may be hard to sample. Nevertheless, some problems and limitations remain associated with the bottom-up approach, including the fact that processes other than positive or divergent selection can generate signals similar to selective sweeps (e.g., background selection, see also Chapter 1). We discuss these problems and potential solutions below.

Locus-Level Work to Understand the Genetics of Adaptation and Speciation

Similar to work focused on identifying the genetic basis of phenotypic traits, the estimation of summary statistics along the genome can be used to study the processes of adaptation and speciation more broadly. Studies typically compare related populations within the same or closely related species and focus on estimates of genomic differentiation, such as FST. One of the major conclusions is that differentiation can be highly variable across the genome, with areas of elevated FST interspersed with areas of reduced FST (e.g., [53–58]). An important inference from these findings is that speciation can proceed through only a few focal changes and does not require divergence across the entire genome [59]. While this conclusion does not seem to be controversial, the processes that generate variable patterns of differentiation have attracted considerable interest and led to two main models: divergence-with-gene-flow and selection-in-allopatry.

The divergence-with-gene-flow model posits that divergent selection at loci involved in speciation must be protecting some regions of the genome from gene flow, elevating an otherwise homogenized (or low) landscape of differentiation [60, 61]. This model received considerable enthusiasm when it was first developed, as it suggests that researchers can identify loci involved with speciation relatively easily by looking for areas of elevated differentiation between closely related populations. Recent work has encouraged caution with this approach, and work with birds has been at the forefront of this wave, promoting a second model to explain variation in FST, selection-in-allopatry. The selection-in-allopatry model assumes that variation in the strength of selection can explain variation in differentiation on its own [62]. This model derives from the observation that FST is a relative measure of differentiation that includes a term for within-population variation. As a result, it can be elevated by reductions in within-population variation alone. These reductions can derive from any form of linked selection, including genetic hitchhiking and background selection that is unrelated to speciation ([62, 63]; also true for elevated neutral variance generated by population structure

313 [64]), and mean that gene flow is not necessarily needed to explain variation in genomic
314 differentiation.

315 Another measure of differentiation is dXY, an absolute measure of genetic differentiation that
316 represents the average number of nucleotide differences between two populations and does
317 not include a term for within-population variation. Contrasts between windowed-estimates of
318 FST, a relative measure, and dXY have been used to support the selection-in-allopatry model.
319 As an absolute measure, dXY should be unaffected by reductions in within-population
320 variation and should show limited associations with estimates of FST across the genome.
321 Work with birds supports this prediction. Burri et al. [65] estimated FST and dXY between
322 collared and pied flycatchers and noted that dXY was not elevated where FST was elevated.
323 In fact, dXY seemed to show the opposite pattern of FST, being reduced where FST was
324 elevated. Burri et al. [65] argued that recurrent linked selection in regions of reduced
325 recombination in ancestral populations was responsible for this pattern (Fig. 1d). This form of
326 selection would reduce dXY to zero prior to population splitting. Similar findings have been
327 documented between Swainson's thrushes [54], greenish warblers (*Phylloscopus trochiloides*,
328 [55]), stonechats (*Saxicola rubicola*, [56]), and Darwin's finches (*Geospiza fortis*, [57]).

329 Burri et al. [65] included an additional dimension to the selection-in-allopatry model. These
330 authors documented an association between FST and recombination rates in flycatchers,
331 suggesting that genomic features - like reduced recombination that extends the effects of
332 linked selection by preventing linked neutral sites from recombining off their shared
333 background - could also play a role in generating variation in FST. Delmore et al. [66] further
334 evaluated this idea using a comparative analysis, estimating genomic differentiation (FST and
335 dXY) between eight bird populations spanning a broad taxonomic scale (sharing a common
336 ancestor ~52 Mya). Features of the local genomic landscape are highly conserved across
337 birds, including chromosome number, recombination rate, and synteny [67–73]. Accordingly,
338 Delmore et al. [66] predicted that if genomic features are generating variation in differentiation
339 across genomes, they should generate correlated patterns of differentiation across population

pairs of birds. In support of this prediction, a significant proportion of variation in windowed estimates of FST and dXY could be explained by correlations across pairs (up to 3% for FST and 26% for dXY). In addition, genomic regions showing high repeatability across pairs were correlated with several genomic features (e.g., reduced recombination rates [approximated using GC content], elevated gene densities, and chromosome size [higher on micro- vs. macrochromosomes]).

As a final note on the genetics of adaptation and speciation, support for the divergence-with-gene-flow model versus the selection-in-allopatry model will likely depend on the geographic context in which speciation occurs. Much of the work on speciation genomics in birds focuses on species in the temperate region that have experienced periods of allopatry and sympatry [74]. Accordingly, a model including allopatric periods (selection-in-allopatry) will likely be more relevant. In addition, there are variants to the selection-in-allopatry model. For example, Delmore et al. [54] and Irwin et al. [55] describe a scenario in which selective sweeps following secondary contact could also reduce dXY in regions of elevated FST, with globally adaptive alleles evolving during allopatric periods and subsequently sweeping across both populations when they come into secondary contact. These alternatives are not mutually exclusive.

Ancestral Recombination Graphs

The ancestral recombination graph (ARG, see Chapters 1 and 11) captures the full evolutionary history of a set of sampled sequences [75]. The availability of ARG reference makes many population genomic analyses, such as demography inference, selection detection, recombination rate estimation, and others, more accurate [76, 77]. Additionally, having an ARG enables novel analyses. Using the time-resolved ancestry of every genomic region, population structure can now be estimated at different time points [78].

One identified limitation that affects ARG reconstruction in birds is the presence of high-recombining regions [34]. ARG estimation methods struggle with these regions because recombination rates often exceed mutation rates, limiting the information available for accurate

tree reconstruction (Fig. 2a, [79–82]). While in humans these hotspots of high recombination rates are comparatively narrow, birds (as some other groups like dogs and percomorph fish) appear to have much wider hotspots [73, 83, 84], making accurate ARG inference in those regions impossible. Genome-wide inferences can still be carried out by masking high-recombining regions (see [34]), but local genealogies at these regions cannot be derived.

Another issue affecting the construction of ARGs is the need for accurately phased data. Many studies in birds rely on statistical phasing, whereas in humans, extensive haplotype reference panels exist, making accurate phasing much simpler than in non-standard organisms. Sequencing techniques such as long-read sequencing and linked-read sequencing can alleviate these issues by providing direct evidence of the haplotype from which a variant originates. These do come with a substantially higher cost than short-read sequencing, but sequencing a few individuals to build a reference panel can already substantially improve phasing accuracy [85].

Recently, a very unique study has been published that fully resolves the ARG and uses tree sequence-based methods in an avian system [51]. In this previously mentioned work on throat coloration in the wheatear *hispanica* complex, the authors use the ARG to resolve the origin of the variants associated with coloration. By comparing the time to the most recent common ancestor of the white coloration haplotype, they find that it is younger than the species split and older in *O. melanoleuca* than in *O. hispanica*, suggesting an introgression of the white haplotype from *O. melanoleuca* into *O. hispanica* [51]. This resolution of haplotype histories is only possible using ARG-based methods, and we expect more studies to make use of this methodology in the future.

Polygenic traits in avian systems

Since the beginning of genetics research, two divergent theories for how information is inherited have been discussed. On the one hand, Mendel's laws, which, with the discovery of DNA, have been found to fit the discrete nature of how information is stored in the genome.

On the other hand, the field of quantitative genetics, founded by Francis Galton around a similar time independently of Gregor Mendel, is well-suited to explain continuous traits and deals with inheritance using statistical methods. Quantitative genetics theory has been used with great success, especially in captive breeding programs, and has been shown to make accurate predictions about the distribution of resulting traits.

A core concept of quantitative genetics is variance partitioning, in which the variance in the phenotype of interest is partitioned between environmental and genetic components. The proportion of variance explained by genetics is also called heritability and indicates how much similarity can be expected in the trait of interest between parent and offspring. Quantitative genetics has also often been applied in avian systems. Early work has primarily focused on easily measurable morphological traits, for which high heritabilities have been reported [86]. Current work also includes non-morphological traits such as behaviour, which has been found to generally exhibit moderate heritability [87, 88]. These studies require either selective or controlled breeding or extensive monitoring to establish a pedigree between the individuals. In wild bird populations this is very resource intensive and requires long-term commitment, which has been done in a few species such as nest box breeding species like the great tits *Parus major* [89], or colony breeding birds like common terns *Sterna hirundo* [90] or larger birds that allow for colour marked readings such as European/common shags *Gulosus aristotelis* [91].

One theoretical paper focusing on the amount of selection pressure needed for a migratory bird species to display a resident phenotype also compares different genomic architectures, assessing how many traits are encoded by many genes vs. few genes. De Zoeten & Pulido [92] find that, while the number of genes involved does not substantially change the speed of adaptation, it does affect the amount of cryptic variation maintained in a system that is governed by a threshold model. In the paper, variants encoding migratory behaviour are much longer maintained in a polygenic setup than when only a few variants are involved, potentially explaining how species can quickly adapt to changing environmental conditions.

With the increased accessibility of genomic sequencing approaches for any study system, the tools are finally available to move beyond variance partitioning and theoretical approaches. With genetic mapping approaches, researchers can investigate how these continuous traits are shaped by the discrete building blocks of DNA and identify specific genetic variants (quantitative trait loci; QTL). Initial applications in humans have identified thousands of loci underlying single traits (e.g., 12,111 variants associated with human height using a sample of 5.4 million individuals, [93]) and have shown that some traits previously thought to be controlled by few loci of large effect are in fact polygenic (e.g., human eye color [94]).

One of the best studied avian traits in this regard is beak shape and size. The first whole-genome population genomics analysis of Darwin's finches, well known for their differences in beak shape, identified 15 regions differentiated between populations with blunt versus pointed beaks. In this study, the authors used methods for detecting selection events to find these regions [95]. Due to the rapid radiation in Darwin’s finches, some of the variation found in different populations and species predates the radiation, and the same haplotype underlies the variation (Fig. 2b, [57, 96]). While some of the regions found might be due to other population differences, some other variants are likely either younger than the radiation or do not form similar-sized blocks and thus remain undetected. Indeed, follow-up studies find additional regions associated with bill size and shape [97].

The main limiting factor in applying genetic mapping methods to avian systems is the required sample size. In model organisms, studies of continuous traits have often found that common variance is composed of a few QTL with large effects and many QTL with small effects. Additionally, rare variants with large effects often explain another large part of the variation. Whenever the effect size is small or the variants are rare, larger sample sizes are required to detect them. Nonetheless, some studies have investigated continuous traits of high heritability and the underlying genetic architecture.

While some success stories exist, various studies have also encountered the limitations of polygenic analyses, finding no significant associations even in cases where the phenotype is

known to be heritable. One study of the collared flycatcher (*Ficedula albicollis*) has combined its apparent negative result with an analysis of the difficulties of detecting QTL [98]. The study aims to find an association of genetic variation with the forehead patch size of the collared flycatcher, a trait known to be under sexual selection [99, 100], with previous studies reporting moderately high heritability based on pedigree [101]. The authors evaluate this relationship using two different datasets. The first set of individuals is sampled for extreme values of forehead patch size and sequenced using whole-genome sequencing. In the second dataset, the genotypes of a much larger number of randomly selected individuals are measured using a 50K SNP chip. In both datasets, no variant was significantly associated with forehead patch size. In their power analysis, the authors show that the WGS extreme sampling scheme dataset outperforms random sampling in terms of detection power; however, the sample size used in their study could only detect variants with an effect of 15% or more on the phenotype. A tripling of the sample size would, however, already increase power significantly to the point where variants of small effect should be detectable. For the SNP chip random sampling, however, the sample size would either have to be increased dramatically or the number of SNPs considered would have to be increased. Given the high effective population size in the sample and therefore low relatedness between samples, randomly selected SNPs were unlikely to be in linkage disequilibrium with QTLs, decreasing the power of the SNP chip assay [98].

Even in systems where pedigrees are known and relatedness is high between samples, QTL mapping is not always successful. In a study of clutch size and egg mass in 969 great tits, the authors found no QTL with a significant effect on the phenotype in question, either using QTL mapping with the known pedigree or in a GWAS [102]. A follow-up study including an additional population, more individuals, and more traits found the same results for clutch size and egg mass, as well as for most other traits studied [103].

All of the previous studies have additionally incorporated a chromosome partitioning approach, where phenotypic variance explained by all variants is calculated separately for each

chromosome. For a highly polygenic trait, the proportion of variance explained is expected to correlate with chromosome size, as larger chromosomes harbour more loci with a potential effect. This method is especially suitable in birds [104], whose karyotypes exhibit substantial variation in chromosome size. But, these methods must also be used with caution regarding sample size, amount of heritability, and skewed effect size distributions, as these factors can lead to a loss of statistical power [104].

The above studies are just a few highlights of cases where no correlation could be found between the genotype and a known heritable trait, leading to the conclusion that many traits are caused by a large number of small-effect loci. However, all of the abovementioned studies use SNPs, and while these may be linked to other variants, there is additional heritable variation that is not detected by the methods employed.

Moving beyond SNPs

Most of the population genomic studies we have highlighted in this chapter focus on genomic variation in SNPs; however, many of the studies on the genomic bases of phenotypic traits find structural variants (SVs) to explain much of the variation. Most of these studies have usually found the SVs of interest by proxy through SNP variation. While this is possible for inversions due to the interruption of recombination, the same is not necessarily true for other SVs, such as insertions/deletions and copy number variations, all of which are often caused by TEs. In the willow warbler study we previously highlighted, which examined a TE affecting a phenotypic trait (migratory direction), the TE in question was by chance included in an AFLP (amplified fragment length polymorphism) panel [105, 106]. Only very few studies [107–109] have specifically scanned for structural variants to explain phenotypic variation, likely because SV detection with short-read sequencing is error-prone and long-read or linked-read sequencing still remains comparatively expensive. Additionally, the studies that identified SVs associated with traits have often found it difficult to narrow down on the exact mechanism by which the phenotype is shaped. In some cases, SVs are close to a gene or even affect the gene structure itself, but in many cases, their effect is likely regulatory and acts through a

complex network of interactions. In these cases, it can be useful to approach the problem from the opposite direction and investigate the effects of the regulatory actions themselves.

Generally speaking, changes in the genome are first detected through changes in gene transcription. For complex traits, it is assumed that distinct phenotypes differ mainly in transcription levels of key genes (some examples where this was found include loss of flight in birds [110] and loss of eyes in cavefishes [111]). To investigate expression levels and characterise transcriptomes, RNA-seq is employed. Here, the sequencing depth of a gene is correlated with its relative abundance in the cell, enabling comparison of expression levels across populations or phenotypes. This methodology has often been applied to avian systems, investigating the transcription differences encoding vocalisation [112], cognition [113], beak morphology [114], and migration [115].

When investigating gene expression patterns, precision in sample collection is extremely important. Expression profiles can vary widely depending on the tissue (and location in the tissue), time of day, and day in the year of the collection, as well as developmental stage, sex, and age of the individual at the time of sampling. Especially in the brain, where differences in behaviour are expected to be regulated and integrated, the exact location and cell type strongly impact the transcription profile, and different types and subtypes of neurons may be located very closely together. Recently, new developments in sequencing technology have enabled a level of detail beyond brain region-specific bulk RNA sequencing. By attaching barcodes to sequencing reads coming from one cell (single-cell sequencing) or one region in space (spatial sequencing), researchers can now pick out signals much more easily, as the noise from cells/regions that are not related to the phenotype in question can be removed.

One area in which this novel method has already been applied extensively and successfully is song learning in songbirds. The brain regions relevant to song learning have been extensively studied, but recent works using single-cell transcriptomics have shed more light on how the exact neurons involved in the song motor pathway have evolved [116] and how hybrids between closely related species can exhibit a broader vocal spectrum [117]. In the

latter case, the transcription profiles of different cell types in zebra finches (*Taeniopygia guttata*), cherry finches (*Aidemosyne modesta*), and their F1 hybrids were examined. The researchers found that most cell types had similar parental transcription profiles, and in cases where differences existed, the hybrids showed an intermediate profile. The only exception was in the GLUTenergic neurons, in which transcription levels in the hybrids clustered independently of those of the parental species. This was driven by an increased abundance of non-additively inherited transcription levels, in which the transcription level of a single gene is not intermediate between the parental species but rather dominant in either direction or over- or under-dominant compared to both parental species (Fig. 2c, [117]).

Bridging the gap between gene expression and genetic variation is challenging. Local interactions between gene expression and genetic variants can be investigated using allele-specific expression. In this method, expression is compared between the two parental alleles, allowing for association with variants on the same strand [118]. For genome wide interactions, techniques such as ChIP-Seq can be employed, where immunoprecipitation is combined with sequencing to find the binding site in the DNA for a specific transcription factor. On the other hand, changes in gene expression do not always arise from genetic variation but also from chemical and molecular modifications, including methylation, phosphorylation, acetylation, chromatin accessibility, and occupancy of regulatory sequences by transcription factors.

Epigenetic studies are beginning to accumulate in birds and include investigations of vocalization (zebra finch [119]); learning and cognition (great tit [Parus major, [120]]), and beak size and shape (Darwin's finches, [121, 122]). In a study focusing on the process of song production, behaviorally regulated gene-expression profiles were found to vary with time and across anatomical structures. It was hypothesized that drivers of this variability are signalling cascades modulated by transcription factors, cis- and trans-regulatory regions, and epigenetic chromatin states [119]. This study focused on identifying transcriptionally active chromatin regions for song nuclei involved in their focal trait (singing). Their results suggest the presence of epigenetic modifications specific to each brain region that prime gene regulation, such that

target regions are in a chromatin state that enables immediate modulation of transcription of behavior-specific genes once the behavior is initiated (upon neuronal firing).

Integrating population genomics with additional fields

We will round up our chapter by highlighting the importance and potential of integrating population genetics with other fields to answer broader questions relevant to ecology, behaviour, conservation, and evolution. Specifically, the fields of landscape genetics and phylogeography were originally developed as bridges to population genetics to study the interaction between landscape features and micro- or macroevolutionary processes, respectively, involving geographic sampling beyond what is done for population genetics. The advent of HTS has blurred the boundaries between these fields, with principles from landscape genetics and phylogeography being used to (1) identify loci under selection, (2) identify correlations between genomic data and the environment, and (3) reconstruct the history of divergence within species [123, 124]. One example comes from Manthey & Moyle [125] who focus on populations of white-breasted nuthatches (*Sitta carolinensis*) that occupy the sky islands of Arizona. RAD sequencing showed that genetic differentiation between these islands was largely mediated by ecological distance rather than geographic distance, identifying eight loci associated with ecological distance (e.g., strong associations with minimum precipitation of the driest month and maximum temperature of the hottest month). This integration provides a more complete picture of adaptation and speciation and could be especially important for conservation, allowing researchers to understand how organisms relate to their environment and how they will respond to future landscape changes.

Conclusion

We look forward to both participating and watching how the field of population genomics evolves for birds in the coming years. As sketched in this chapter, birds have played an important role in the development and application of population genomics, with work on

580 demography and genetic variation at the genome level informing evolution and conservation,
581 and work at the locus level providing information on the genetic basis of phenotypic traits,
582 adaptation, and speciation. We are eagerly expecting the field of trait mapping to move
583 towards a more holistic approach, integrating whole-genome with epigenetic and
584 transcriptional data. On the evolutionary side, the greater data availability to produce accurate
585 phasing will improve our insights into past events, further deepening our understanding of
586 speciation and trait evolution. Integrations with other fields and beyond landscape genetics
587 and phylogenetics (e.g., functional genomics and genome editing) are also on the horizon for
588 some species and will undoubtedly help to unravel many mysteries concerning the behavior,
589 ecology, evolution, and conservation of birds.

590

591 1. Nice M (1939) The watcher at the nest. Macmillan, New York

592 2. Mayr E (1999) Systematics and the Origin of Species, from the Viewpoint of a
593 Zoologist. Harvard University Press

594 3. Luikart G, England PR, Tallmon D, et al (2003) The power and promise of population
595 genomics: from genotyping to genome typing. Nat Rev Genet 4:981–994.
596 https://doi.org/10.1038/nrg1226

597 4. Begun DJ, Holloway AK, Stevens K, et al (2007) Population Genomics: Whole-
598 Genome Analysis of Polymorphism and Divergence in Drosophila simulans. PLOS Biol
599 5:e310. https://doi.org/10.1371/journal.pbio.0050310

600 5. Avise JC, Nelson WS (1989) Molecular Genetic Relationships of the Extinct Dusky
601 Seaside Sparrow. Science 243:646–648. https://doi.org/10.1126/science.243.4891.646

602 6. Zink RM, Kale HW (1995) Conservation genetics of the extinct dusky seaside sparrow
603 *Ammodramus maritimus nigrescens*. Biol Conserv 74:69–71.
604 https://doi.org/10.1016/0006-3207(95)00010-2

605 7. Stapley J, Reger J, Feulner PGD, et al (2010) Adaptation genomics: the next
606 generation. Trends Ecol Evol 25:705–712. https://doi.org/10.1016/j.tree.2010.09.002

607 8. Davey JW, Hohenlohe PA, Etter PD, et al (2011) Genome-wide genetic marker
608 discovery and genotyping using next-generation sequencing. Nat Rev Genet 12:499–
609 510. https://doi.org/10.1038/nrg3012

610 9. Gregory TR (2005) Synergy between sequence and size in Large-scale genomics.
611 Nat Rev Genet 6:699–708. https://doi.org/10.1038/nrg1674

612 10. Organ CL, Shedlock AM, Meade A, et al (2007) Origin of avian genome size and
613 structure in non-avian dinosaurs. Nature 446:180–184.
614 https://doi.org/10.1038/nature05621

615 11. Hillier LW, Miller W, Birney E, et al (2004) Sequence and comparative analysis of the
616 chicken genome provide unique perspectives on vertebrate evolution. Nature 432:695
617 –716. https://doi.org/10.1038/nature03154

618 12. Warren WC, Clayton DF, Ellegren H, et al (2010) The genome of a songbird. Nature
619 464:757–762. https://doi.org/10.1038/nature08819

620 13. Ellegren H (2014) Genome sequencing and population genomics in non-model
621 organisms. Trends Ecol Evol 29:51–63. https://doi.org/10.1016/j.tree.2013.09.008

622 14. Zhang G (2015) Bird sequencing project takes off. Nature 522:34–34.
623 https://doi.org/10.1038/522034d

624 15. Feng S, Stiller J, Deng Y, et al (2020) Dense sampling of bird diversity increases
625 power of comparative genomics. Nature 587:252–257.
626 https://doi.org/10.1038/s41586-020-2873-9

627 16. Eyre-Walker A, Keightley PD (2009) Estimating the Rate of Adaptive Molecular

628 Evolution in the Presence of Slightly Deleterious Mutations and Population Size
629 Change. Mol Biol Evol 26:2097–2108. https://doi.org/10.1093/molbev/msp119

630 17. Zeng K (2013) A coalescent model of background selection with recombination,
631 demography and variation in selection coefficients. Heredity 110:363–371.
632 https://doi.org/10.1038/hdy.2012.102

633 18. Bank C, Ewing GB, Ferrer-Admettla A, et al (2014) Thinking too positive? Revisiting
634 current methods of population genetic selection inference. Trends Genet 30:540–546.
635 https://doi.org/10.1016/j.tig.2014.09.010

636 19. Beissinger TM, Wang L, Crosby K, et al (2016) Recent demography drives changes in
637 linked selection across the maize genome. Nat Plants 2:16084.
638 https://doi.org/10.1038/nplants.2016.84

639 20. Weir JT, Schluter D (2004) Ice sheets promote speciation in boreal birds. Proc R Soc
640 B Biol Sci 271:1881–1887. https://doi.org/10.1098/rspb.2004.2803

641 21. Pinho C, Hey J (2010) Divergence with Gene Flow: Models and Data. Annu Rev Ecol
642 Evol Syst 41:215–230. https://doi.org/10.1146/annurev-ecolsys-102209-144644

643 22. Borge T, Lindroos K, Nádvorník P, et al (2005) Amount of introgression in flycatcher
644 hybrid zones reflects regional differences in pre and post-zygotic barriers to gene
645 exchange. J Evol Biol 18:1416–1424. https://doi.org/10.1111/j.1420-
646 9101.2005.00964.x

647 23. Sætre G-P, Borge T, Lindell J, et al (2001) Speciation, introgressive hybridization and
648 nonlinear rate of molecular evolution in flycatchers. Mol Ecol 10:737–749.
649 https://doi.org/10.1046/j.1365-294x.2001.01208.x

650 24. Nadachowska-Brzyska K, Burri R, Olason PI, et al (2013) Demographic Divergence
651 History of Pied Flycatcher and Collared Flycatcher Inferred from Whole-Genome Re-

652 sequencing Data. PLOS Genet 9:e1003942.
653 https://doi.org/10.1371/journal.pgen.1003942

654 25. Li H, Durbin R (2011) Inference of human population history from individual whole-
655 genome sequences. Nature 475:493–496. https://doi.org/10.1038/nature10231

656 26. Sætre G-P, Sæther SA (2010) Ecology and genetics of speciation in Ficedula
657 flycatchers. Mol Ecol 19:1091–1106. https://doi.org/10.1111/j.1365-294X.2010.04568.x

658 27. Nater A, Burri R, Kawakami T, et al (2015) Resolving Evolutionary Relationships in
659 Closely Related Species with Whole-Genome Sequencing Data. Syst Biol 64:1000–
660 1017. https://doi.org/10.1093/sysbio/syv045

661 28. Marchi N, Kapopoulou A, Excoffier L (2024) Demogenomic inference from spatially
662 and temporally heterogeneous samples. Mol Ecol Resour 24:e13877.
663 https://doi.org/10.1111/1755-0998.13877

664 29. Excoffier L, Marchi N, Marques DA, et al (2021) fastsimcoal2: demographic inference
665 under complex evolutionary scenarios. Bioinformatics 37:4882–4885.
666 https://doi.org/10.1093/bioinformatics/btab468

667 30. Excoffier L, Dupanloup I, Huerta-Sánchez E, et al (2013) Robust Demographic
668 Inference from Genomic and SNP Data. PLOS Genet 9:e1003905.
669 https://doi.org/10.1371/journal.pgen.1003905

670 31. Schiffels S, Wang K (2020) MSMC and MSMC2: The Multiple Sequentially Markovian
671 Coalescent. Methods Mol Biol Clifton NJ 2090:147–166. https://doi.org/10.1007/978-1-
672 0716-0199-0_7

673 32. Nadachowska-Brzyska K, Li C, Smeds L, et al (2015) Temporal Dynamics of Avian
674 Populations during Pleistocene Revealed by Whole-Genome Sequences. Curr Biol
675 25:1375–1380. https://doi.org/10.1016/j.cub.2015.03.047

33. Nadachowska-Brzyska K, Burri R, Smeds L, Ellegren H (2016) PSMC analysis of effective population sizes in molecular ecology and its application to black-and-white Ficedula flycatchers. Mol Ecol 25:1058–1072. https://doi.org/10.1111/mec.13540

34. Ishigohoka J, Liedvogel M (2025) High-recombining genomic regions affect demography inference based on ancestral recombination graphs. Genetics 229:iyaf004. https://doi.org/10.1093/genetics/iyaf004

35. Li S, Li B, Cheng C, et al (2014) Genomic signatures of near-extinction and rebirth of the crested ibis and other endangered bird species. Genome Biol 15:557. https://doi.org/10.1186/s13059-014-0557-1

36. Kardos M, Taylor HR, Ellegren H, et al (2016) Genomics advances the study of inbreeding depression in the wild. Evol Appl 9:1205–1218. https://doi.org/10.1111/eva.12414

37. Bay RA, Harrigan RJ, Underwood VL, et al (2018) Genomic signals of selection predict climate-driven population declines in a migratory bird. Science 359:83–86. https://doi.org/10.1126/science.aan4380

38. Ruegg K, Anderson EC, Somveille M, et al (2021) Linking climate niches across seasons to assess population vulnerability in a migratory bird. Glob Change Biol 27:3519–3531. https://doi.org/10.1111/gcb.15639

39. Battey C, Ralph PL, Kern AD (2020) Predicting geographic location from genetic variation with deep neural networks. eLife 9:e54507. https://doi.org/10.7554/eLife.54507

40. Siol M, Wright SI, Barrett SCH (2010) The population genomics of plant adaptation. New Phytol 188:313–332. https://doi.org/10.1111/j.1469-8137.2010.03401.x

41. Nielsen R, Williamson S, Kim Y, et al (2005) Genomic scans for selective sweeps

using SNP data. Genome Res 15:1566–1575. https://doi.org/10.1101/gr.4252305

42. Küpper C, Stocks M, Risse JE, et al (2016) A supergene determines highly divergent male reproductive morphs in the ruff. Nat Genet 48:79–83. https://doi.org/10.1038/ng.3443

43. Lamichhaney S, Fan G, Widemo F, et al (2016) Structural genomic changes underlie alternative reproductive strategies in the ruff (Philomachus pugnax). Nat Genet 48:84–88. https://doi.org/10.1038/ng.3430

44. Tuttle EM, Bergland AO, Korody ML, et al (2016) Divergence and Functional Degradation of a Sex Chromosome-like Supergene. Curr Biol 26:344–350. https://doi.org/10.1016/j.cub.2015.11.069

45. Sanchez-Donoso I, Ravagni S, Rodríguez-Teijeiro JD, et al (2022) Massive genome inversion drives coexistence of divergent morphs in common quails. Curr Biol 32:462-469.e6. https://doi.org/10.1016/j.cub.2021.11.019

46. Funk ER, Mason NA, Pálsson S, et al (2021) A supergene underlies linked variation in color and morphology in a Holarctic songbird. Nat Commun 12:6833. https://doi.org/10.1038/s41467-021-27173-z

47. Bourque G, Burns KH, Gehring M, et al (2018) Ten things you should know about transposable elements. Genome Biol 19:199. https://doi.org/10.1186/s13059-018-1577-z

48. Lundberg M, Mackintosh A, Petri A, Bensch S (2023) Inversions maintain differences between migratory phenotypes of a songbird. Nat Commun 14:452. https://doi.org/10.1038/s41467-023-36167-y

49. Caballero-López V, Lundberg M, Sokolovskis K, Bensch S (2022) Transposable elements mark a repeat-rich region associated with migratory phenotypes of willow

724 warblers (Phylloscopus trochilus). Mol Ecol 31:1128–1141.
725 https://doi.org/10.1111/mec.16292

726 50. Sokolovskis K, Lundberg M, Åkesson S, et al (2023) Migration direction in a songbird
727 explained by two loci. Nat Commun 14:165. https://doi.org/10.1038/s41467-023-35788-
728 7

729 51. Lutgen D, Peona V, Chase MA, et al (2025) A mosaic of modular variation at a single
730 gene underpins convergent plumage coloration. Science 390:eado8005.
731 https://doi.org/10.1126/science.ado8005

732 52. Delmore KE, Liedvogel M (2016) Investigating Factors that Generate and Maintain
733 Variation in Migratory Orientation: A Primer for Recent and Future Work. Front Behav
734 Neurosci 10:. https://doi.org/10.3389/fnbeh.2016.00003

735 53. Ellegren H, Smeds L, Burri R, et al (2012) The genomic landscape of species
736 divergence in Ficedula flycatchers. Nature 491:756–760.
737 https://doi.org/10.1038/nature11584

738 54. Delmore KE, Hübner S, Kane NC, et al (2015) Genomic analysis of a migratory divide
739 reveals candidate genes for migration and implicates selective sweeps in generating
740 islands of differentiation. Mol Ecol 24:1873–1888. https://doi.org/10.1111/mec.13150

741 55. Irwin DE, Alcaide M, Delmore KE, et al (2016) Recurrent selection explains parallel
742 evolution of genomic regions of high relative but low absolute differentiation in a ring
743 species. Mol Ecol 25:4488–4507. https://doi.org/10.1111/mec.13792

744 56. Van Doren BM, Campagna L, Helm B, et al (2017) Correlated patterns of genetic
745 diversity and differentiation across an avian family. Mol Ecol 26:3982–3997.
746 https://doi.org/10.1111/mec.14083

747 57. Han F, Lamichhaney S, Grant BR, et al (2017) Gene flow, ancient polymorphism, and

748 ecological adaptation shape the genomic landscape of divergence among Darwin's
749 finches. Genome Res 27:1004–1015. https://doi.org/10.1101/gr.212522.116

750 58. Poelstra JW, Vijay N, Bossu CM, et al (2014) The genomic landscape underlying
751 phenotypic integrity in the face of gene flow in crows. Science 344:1410–1414.
752 https://doi.org/10.1126/science.1253226

753 59. Wu C-I, Ting C-T (2004) Genes and speciation. Nat Rev Genet 5:114–122.
754 https://doi.org/10.1038/nrg1269

755 60. Nosil P (2012) Ecological Speciation. OUP Oxford

756 61. Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous
757 genomic divergence. Mol Ecol 18:375–402. https://doi.org/10.1111/j.1365-
758 294X.2008.03946.x

759 62. Cruickshank TE, Hahn MW (2014) Reanalysis suggests that genomic islands of
760 speciation are due to reduced diversity, not reduced gene flow. Mol Ecol 23:3133–3157.
761 https://doi.org/10.1111/mec.12796

762 63. Noor M a. F, Bennett SM (2009) Islands of speciation or mirages in the desert?
763 Examining the role of restricted recombination in maintaining species. Heredity 103:439
764 –444. https://doi.org/10.1038/hdy.2009.151

765 64. Bierne N, Roze D, Welch JJ (2013) Pervasive selection or is it…? why are FST
766 outliers sometimes so frequent? Mol Ecol 22:2061–2064.
767 https://doi.org/10.1111/mec.12241

768 65. Burri R, Nater A, Kawakami T, et al (2015) Linked selection and recombination rate
769 variation drive the evolution of the genomic landscape of differentiation across the
770 speciation continuum of Ficedula flycatchers. Genome Res 25:1656–1665.
771 https://doi.org/10.1101/gr.196485.115

772 66. Delmore KE, Lugo Ramos JS, van Doren BM, et al (2018) Comparative analysis
773 examining patterns of genomic differentiation across multiple episodes of population
774 divergence in birds. Evol Lett 2:76–87. https://doi.org/10.1002/evl3.46

775 67. Dawson DA, Åkesson M, Burke T, et al (2007) Gene Order and Recombination Rate
776 in Homologous Chromosome Regions of the Chicken and a Passerine Bird. Mol Biol
777 Evol 24:1537–1552. https://doi.org/10.1093/molbev/msm071

778 68. Griffin DK, Robertson LBW, Tempest HG, Skinner BM (2007) The evolution of the
779 avian genome as revealed by comparative molecular cytogenetics. Cytogenet Genome
780 Res 117:64–77. https://doi.org/10.1159/000103166

781 69. Backström N, Qvarnström A, Gustafsson L, Ellegren H (2006) Levels of linkage
782 disequilibrium in a wild bird population. Biol Lett 2:435–438.
783 https://doi.org/10.1098/rsbl.2006.0507

784 70. Stapley J, Birkhead TR, Burke T, Slate J (2008) A Linkage Map of the Zebra Finch
785 Taeniopygia guttata Provides New Insights Into Avian Genome Evolution. Genetics
786 179:651–667. https://doi.org/10.1534/genetics.107.086264

787 71. Kawakami T, Smeds L, Backström N, et al (2014) A high-density linkage map enables
788 a second-generation collared flycatcher genome assembly and reveals the patterns of
789 avian recombination rate variation and chromosomal evolution. Mol Ecol 23:4035–4058.
790 https://doi.org/10.1111/mec.12810

791 72. Zhang G, Li B, Li C, et al (2014) Comparative genomic data of the Avian
792 Phylogenomics Project. GigaScience 3:2047-217X-3–26. https://doi.org/10.1186/2047-
793 217X-3-26

794 73. Kawakami T, Mugal CF, Suh A, et al (2017) Whole-genome patterns of linkage
795 disequilibrium across flycatcher populations clarify the causes and consequences of

fine-scale recombination rate variation in birds. Mol Ecol 26:4158–4172. https://doi.org/10.1111/mec.14197

74. Hewitt G (2000) The genetic legacy of the Quaternary ice ages. Nature 405:907–913. https://doi.org/10.1038/35016000

75. Griffiths R, Marjoram P An ancestral recombination graph. In: Progress in Population Genetics and Human Evolution. pp 257–270

76. Ralph P, Thornton K, Kelleher J (2020) Efficiently Summarizing Relationships in Large Samples: A General Duality Between Statistics of Genealogies and Genomes. Genetics 215:779–797. https://doi.org/10.1534/genetics.120.303253

77. Wong Y, Ignatieva A, Koskela J, et al (2024) A general and efficient representation of ancestral recombination graphs. Genetics 228:iyae100. https://doi.org/10.1093/genetics/iyae100

78. Wohns AW, Wong Y, Jeffery B, et al (2022) A unified genealogy of modern and ancient genomes. Science 375:eabi8264. https://doi.org/10.1126/science.abi8264

79. Terhorst J, Kamm JA, Song YS (2017) Robust and scalable inference of population history from hundreds of unphased whole genomes. Nat Genet 49:303–309. https://doi.org/10.1038/ng.3748

80. Sellinger TPP, Awad DA, Moest M, Tellier A (2020) Inference of past demography, dormancy and self-fertilization rates from whole genome sequence data. PLOS Genet 16:e1008698. https://doi.org/10.1371/journal.pgen.1008698

81. Sellinger TPP, Abu-Awad D, Tellier A (2021) Limits and convergence properties of the sequentially Markovian coalescent. Mol Ecol Resour 21:2231–2248. https://doi.org/10.1111/1755-0998.13416

82. Dutheil JY (2024) On the estimation of genome-average recombination rates. Genetics 227:iyae051. https://doi.org/10.1093/genetics/iyae051

83. Singhal S, Leffler EM, Sannareddy K, et al (2015) Stable recombination hotspots in birds. Science 350:928–932. https://doi.org/10.1126/science.aad0843

84. Bascón-Cardozo K, Bours A, Manthey G, et al (2024) Fine-Scale Map Reveals Highly Variable Recombination Rates Associated with Genomic Features in the Eurasian Blackcap. Genome Biol Evol 16:evad233. https://doi.org/10.1093/gbe/evad233

85. Choi Y, Chan AP, Kirkness E, et al (2018) Comparison of phasing strategies for whole human genomes. PLOS Genet 14:e1007308. https://doi.org/10.1371/journal.pgen.1007308

86. Boag PT, van Noordwijk AJ (1987) Quantitative Genetics. In: Cooke F, Buckley PA (eds) Avian Genetics: A Population and Ecological Approach. Academic Press Inc., London, pp 45–78

87. Dochtermann NA, Schwab T, Anderson Berdal M, et al (2019) The Heritability of Behavior: A Meta-analysis. J Hered 110:403–410. https://doi.org/10.1093/jhered/esz023

88. Dingemanse NJ, Dochtermann NA (2014) Individual behaviour: behavioural ecology meets quantitative genetics. In: Charmantier A, Garant D, Kruuk LEB (eds) Quantitative Genetics in the Wild. Oxford University Press, p 0

89. McCleery RH, Pettifor RA, Armbruster P, et al (2004) Components of Variance Underlying Fitness in a Natural Population of the Great Tit Parus major. Am Nat 164:E62–E72. https://doi.org/10.1086/422660

90. Moiron M, Charmantier A, Bouwhuis S (2022) The quantitative genetics of fitness in a wild seabird. Evolution 76:1443–1452. https://doi.org/10.1111/evo.14516

91. Acker P, Daunt F, Wanless S, et al (2021) Strong survival selection on seasonal migration versus residence induced by extreme climatic events. J Anim Ecol 90:796–808. https://doi.org/10.1111/1365-2656.13410

92. de Zoeten T, Pulido F (2020) How migratory populations become resident. Proc R Soc B Biol Sci 287:20193011. https://doi.org/10.1098/rspb.2019.3011

93. Yengo L, Vedantam S, Marouli E, et al (2022) A saturated map of common genetic variants associated with human height. Nature 610:704–712. https://doi.org/10.1038/s41586-022-05275-y

94. Simcoe M, Valdes A, Liu F, et al (2021) Genome-wide association study in almost 195,000 individuals identifies 50 previously unidentified genetic loci for eye color. Sci Adv 7:eabd1239. https://doi.org/10.1126/sciadv.abd1239

95. Lamichhaney S, Berglund J, Almén MS, et al (2015) Evolution of Darwin's finches and their beaks revealed by genome sequencing. Nature 518:371–375. https://doi.org/10.1038/nature14181

96. Rubin C-J, Enbody ED, Dobreva MP, et al (2022) Rapid adaptive radiation of Darwin's finches depends on ancestral genetic modules. Sci Adv 8:eabm5982. https://doi.org/10.1126/sciadv.abm5982

97. Enbody ED, Sendell-Price AT, Sprehn CG, et al (2023) Community-wide genome sequencing reveals 30 years of Darwin's finch evolution. Science 381:eadf6218. https://doi.org/10.1126/science.adf6218

98. Kardos M, Husby A, McFarlane SE, et al (2016) Whole-genome resequencing of extreme phenotypes in collared flycatchers highlights the difficulty of detecting quantitative trait loci in natural populations. Mol Ecol Resour 16:727–741. https://doi.org/10.1111/1755-0998.12498

867 99. Qvarnström A (1997) Experimentally increased badge size increases male
868 competition and reduces male parental care in the collared flycatcher. Proc R Soc Lond
869 B Biol Sci. https://doi.org/10.1098/rspb.1997.0169

870 100. Sætre G-P, Moum T, Bureš S, et al (1997) A sexually selected character displacement
871 in flycatchers reinforces premating isolation. Nature 387:589–592.
872 https://doi.org/10.1038/42451

873 101. Qvarnström A (1999) Genotype-by-Environment Interactions in the Determination of
874 the Size of a Secondary Sexual Character in the Collared Flycatcher (ficedula
875 Albicollis). Evolution 53:1564–1572. https://doi.org/10.1111/j.1558-
876 5646.1999.tb05419.x

877 102. Santure AW, De Cauwer I, Robinson MR, et al (2013) Genomic dissection of variation
878 in clutch size and egg mass in a wild great tit (Parus major) population. Mol Ecol 22:3949
879 –3962. https://doi.org/10.1111/mec.12376

880 103. Santure AW, Poissant J, De Cauwer I, et al (2015) Replicated analysis of the genetic
881 architecture of quantitative traits in two wild great tit populations. Mol Ecol 24:6148–
882 6162. https://doi.org/10.1111/mec.13452

883 104. Kemppainen P, Husby A (2018) Inference of genetic architecture from chromosome
884 partitioning analyses is sensitive to genome variation, sample size, heritability and effect
885 size distribution. Mol Ecol Resour 18:767–777. https://doi.org/10.1111/1755-
886 0998.12774

887 105. Bensch S, Grahn M, Müller N, et al (2009) Genetic, morphological, and feather
888 isotope variation of migratory willow warblers show gradual divergence in a ring. Mol
889 Ecol 18:3087–3096. https://doi.org/10.1111/j.1365-294X.2009.04210.x

890 106. Bensch S, Åkesson S, Irwin DE (2002) The use of AFLP to find an informative SNP:

genetic differences across a migratory divide in willow warblers. Mol Ecol 11:2359–2366. https://doi.org/10.1046/j.1365-294X.2002.01629.x

107. Recuerda M, Campagna L (2024) How structural variants shape avian phenotypes: Lessons from model systems. Mol Ecol 33:e17364. https://doi.org/10.1111/mec.17364

108. Delmore KE, Van Doren BM, Ullrich K, et al (2023) Structural genomic variation and migratory behavior in a wild songbird. Evol Lett 7:401–412. https://doi.org/10.1093/evlett/qrad040

109. Weissensteiner MH, Bunikis I, Catalán A, et al (2020) Discovery and population genomics of structural variation in a songbird genus. Nat Commun 11:3403. https://doi.org/10.1038/s41467-020-17195-4

110. Sackton TB, Grayson P, Cloutier A, et al (2019) Convergent regulatory evolution and loss of flight in paleognathous birds. Science 364:74–78. https://doi.org/10.1126/science.aat7244

111. Sifuentes-Romero I, Aviles AM, Carter JL, et al (2023) Trait Loss in Evolution: What Cavefish Have Taught Us about Mechanisms Underlying Eye Regression. Integr Comp Biol 63:393–406. https://doi.org/10.1093/icb/icad032

112. Shi Z, Zhang Z, Schaffer L, et al (2021) Dynamic transcriptome landscape in the song nucleus HVC between juvenile and adult zebra finches. Adv Genet 2:e10035. https://doi.org/10.1002/ggn2.10035

113. Semenov GA, Sonnenberg BR, Branch CL, et al (2024) Genes and gene networks underlying spatial cognition in food-caching chickadees. Curr Biol 34:1930-1939.e4. https://doi.org/10.1016/j.cub.2024.03.058

114. Cheng Y, Gao B, Wang H, et al (2017) Evolution of beak morphology in the Ground Tit revealed by comparative transcriptomics. Front Zool 14:58.

915 https://doi.org/10.1186/s12983-017-0245-6

916 115. Louder MIM, Justen H, Kimmitt AA, et al (2024) Gene regulation and speciation in a
917 migratory divide between songbirds. Nat Commun 15:98.
918 https://doi.org/10.1038/s41467-023-44352-2

919 116. Colquitt BM, Merullo DP, Konopka G, et al (2021) Cellular transcriptomics reveals
920 evolutionary identities of songbird vocal circuits. Science 371:eabd9704.
921 https://doi.org/10.1126/science.abd9704

922 117. Shibata Y, Toji N, Wang H, et al (2024) Expansion of learning capacity elicited by
923 interspecific hybridization. Sci Adv 10:eadn3409.
924 https://doi.org/10.1126/sciadv.adn3409

925 118. Castel SE, Levy-Moonshine A, Mohammadi P, et al (2015) Tools and best practices
926 for data processing in allelic expression analysis. Genome Biol 16:195.
927 https://doi.org/10.1186/s13059-015-0762-6

928 119. Whitney O, Pfenning AR, Howard JT, et al (2014) Core and region-enriched networks
929 of behaviorally regulated genes and the singing genome. Science 346:1256780.
930 https://doi.org/10.1126/science.1256780

931 120. Laine VN, Gossmann TI, Schachtschneider KM, et al (2016) Evolutionary signals of
932 selection on cognition from the great tit genome and methylome. Nat Commun 7:10474.
933 https://doi.org/10.1038/ncomms10474

934 121. Skinner MK, Gurerrero-Bosagna C, Haque MM, et al (2014) Epigenetics and the
935 Evolution of Darwin's Finches. Genome Biol Evol 6:1972–1989.
936 https://doi.org/10.1093/gbe/evu158

937 122. McNew SM, Beck D, Sadler-Riggleman I, et al (2017) Epigenetic variation between
938 urban and rural populations of Darwin's finches. BMC Evol Biol 17:183.

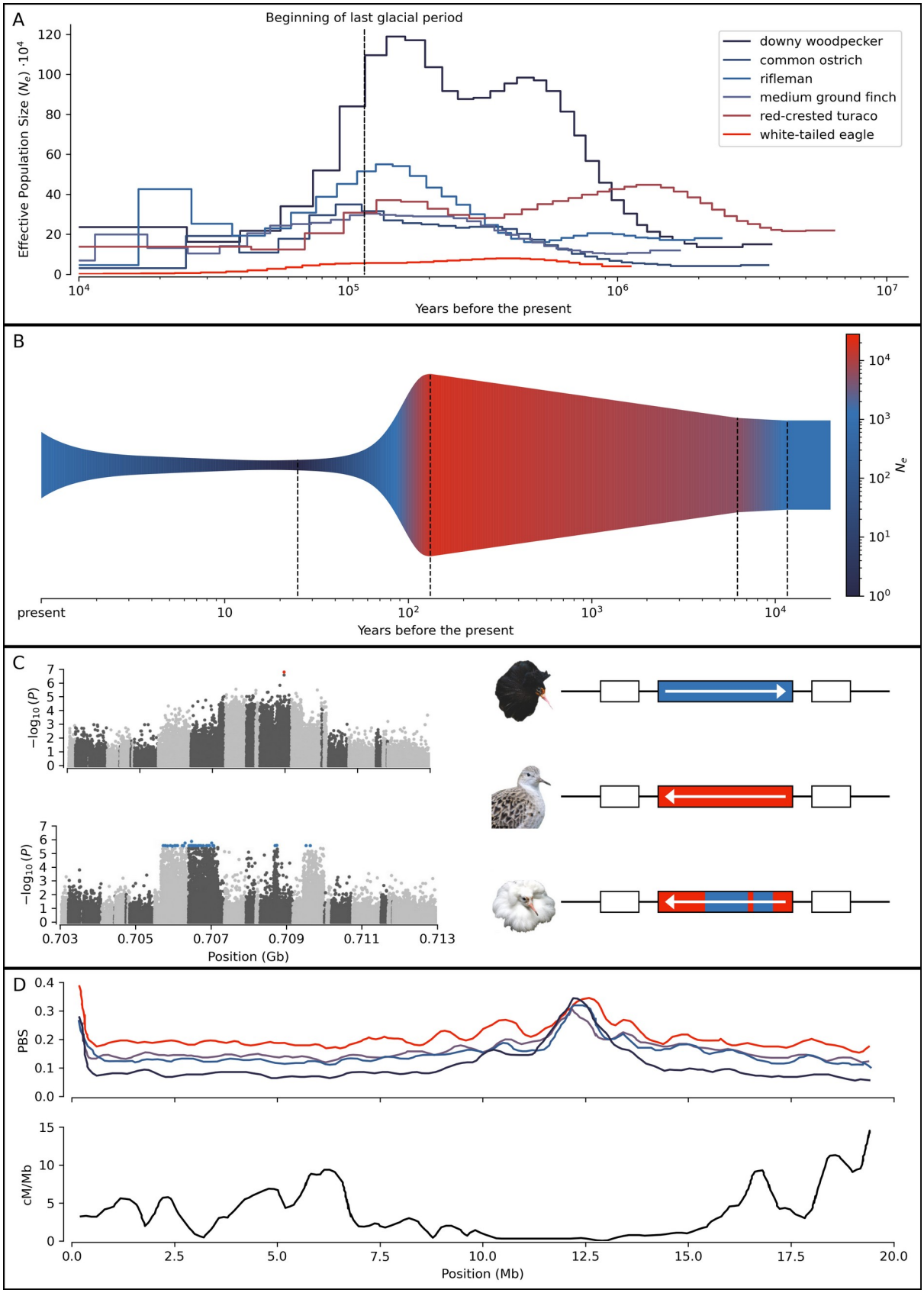
https://doi.org/10.1186/s12862-017-1025-9

123. Edwards SV, Shultz AJ, Campbell-Staton SC (2015) Next-generation sequencing and the expanding domain of phylogeography. Folia Zool 64:187–206. https://doi.org/10.25225/fozo.v64.i3.a2.2015

124. Edwards SV, Potter S, Schmitt CJ, et al (2016) Reticulation, divergence, and the phylogeography–phylogenetics continuum. Proc Natl Acad Sci 113:8025–8032. https://doi.org/10.1073/pnas.1601066113

125. Manthey JD, Moyle RG (2015) Isolation by environment in White-breasted Nuthatches (Sitta carolinensis) of the Madrean Archipelago sky islands: a landscape genomics approach. Mol Ecol 24:3628–3638. https://doi.org/10.1111/mec.13258

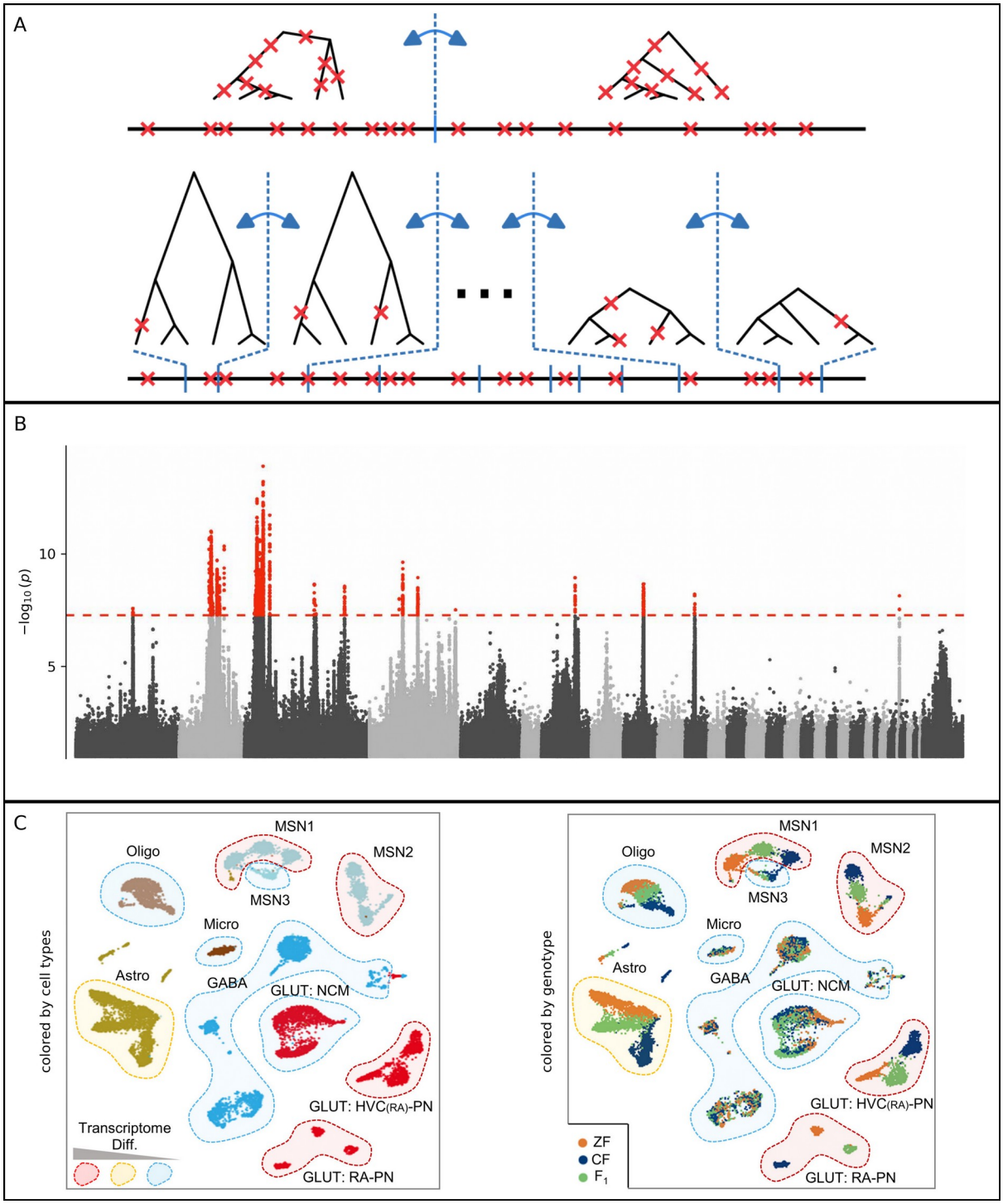
126. Loveland JL, Lank DB, Küpper C (2021) Gene Expression Modification by an Autosomal Inversion Associated With Three Male Mating Morphs. Front Genet 12:. https://doi.org/10.3389/fgene.2021.641620



958

959 Figure 1 a) Demographic history of 8 holarctic bird species, all of which show a recent

population decrease following the beginning of the last glacial period 100 ka ago. Adapted from [32]. b) Recent demographic history of the crested ibis, a once critically endangered bird. The inferred demography matches historical records, showing a rapid decline about 100 years ago following a steady expansion after the last glaciation around 10 ka ago. The genomic data also shows the recent increase in population size following human-assisted recovery programs. Adapted from [35]. c) A genomic inversion in the Ruff encodes divergent morphs and reproductive strategies. The left figure shows an association analysis of variants in the region of the inversion with the faeder phenotype (top) and satellite phenotype (bottom), both compared to the independent. The right figure shows the evolution of this inversion schematically, with the inversion itself causing the faeder phenotype to develop, and subsequent recombination mixing variants between the separated genotypes causing the satellite phenotype. Adapted from [42, 126]. d) Regions of genomic divergence correlate with low recombination rate. The top figure shows lineage-specific FST (PBS) along a chromosome of semi-collared (red), atlas (purple), pied (blue), and collared (dark grey) flycatcher. At around 12.5 Mb along the chromosome, there is a peak of differentiation in all 4 species, which coincides with a low recombination rate as shown in the bottom panel. Adapted from [65].



976 Figure 2 a) Regions of high recombination cause non-resolvability of the ancestral
977 recombination graph. Recombination events are indicated by blue arrows, mutations by red
978 crosses. The top figure shows a region of low recombination rate, with many mutations within
979 the different haplotypes, allowing for a high resolution of the local genealogy. The bottom
980 figure shows a region of high recombination, where haplotypes are short and thus capture few
981 mutations within them. This means that local genealogies cannot be observed, and the

ancestral recombination graph is not well resolved. Adapted from [34]. b) Results of a genome-wide admixture mapping study, where variants are found that contribute to the differentiation of 3 species of closely related Darwin's finches, which differ primarily in beak and body size. Many of the peaks found here overlap with a previous study investigating genomic association with beak and body size within a single species. Together, these results point towards a shared polygenic basis of beak and body size in this recent radiation. Adapted from [96]. c) Single-cell RNA-seq data allows finding distinct patterns in different cell types. In the left figure, color distinguishes different cell types, which form distinct clusters in a UMAP dimension reduction. In the right figure, the same clustering is shown with cells colored by their genotype (pure zebra finch in orange, cherry finch in blue, and hybrids between zebra finch and cherry finch in green). Some cell types show no clear distinction between genotypes (background shaded blue), others show intermediate transcription profiles in hybrids compared to the parental species (background shaded yellow), and in some cases, hybrids cluster into distinct groups, clearly separating them from the parental species (background shaded red). Adapted from [117].