Molecular Phylogenetics and Evolution 185 (2023) 107822

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Museomics and phylogenomics of lovebirds (Psittaciformes, Psittaculidae, *Agapornis*) using low-coverage whole-genome sequencing

Stella Huynh [a], Alison Cloutier [b], Simon Yung Wa Sin [a,*]

[a] School of Biological Sciences, The University of Hong Kong, Pok Fu Lam Road, Hong Kong Special Administrative Region
[b] Department of Organismic and Evolutionary Biology, Mueum of Comparative Zoology, Harvard University, 26 Oxford Street, Cambridge, MA 02138, USA

ARTICLE INFO

Keywords:
High-throughput sequencing
Historical DNA
Museomics
Museum genomics
African parrots
Whole-genome re-sequencing

ABSTRACT

Natural history collections contain specimens that provide important insights into studies of ecology and evolution. With the advancement of high-throughput sequencing, historical DNA (hDNA) from museum specimens has become a valuable source of genomic data to study the evolutionary history of organisms. Low-coverage whole genome sequencing (WGS) has been increasingly applied to museum specimens for analyzing organelle genomes, but is still uncommon for genotyping the nuclear DNA fraction. In this study, we applied low-coverage WGS to phylogenomic analyses of parrots in the genus *Agapornis* by including both modern samples and historical specimens of ~100-year-old. *Agapornis* are small-sized African and Malagasy parrots with diverse characters. Earlier phylogenetic studies failed to resolve the positions of some key lineages, prohibiting a robust interpretation of the biogeography and evolution of these African parrots. Here, we demonstrated the use of low-coverage WGS for generating both mitochondrial and nuclear genomic data, and evaluated data quality differences between modern and historical samples. Our resolved *Agapornis* phylogeny indicates the ancestor of *Agapornis* likely colonized Madagascar from Australasia by *trans*-oceanic dispersal events before dispersing to the African continent. Genome-wide SNPs also allowed us to identify the parental origins of hybrid *Agapornis* individuals. This study demonstrates the potential of applying low-coverage WGS to phylogenomics and population genomics analyses and illustrates how including historical museum specimens can address outstanding questions regarding the evolutionary history of contemporary lineages.

1. Introduction

Museum specimens represent a valuable and vast source of historical DNA (hDNA) for studying the evolutionary history of organisms (Buerki and Baker, 2016; Habel et al., 2014; Holmes et al., 2016). hDNA is derived from traditional voucher specimens archived in museum collections that are <200 years old (Card et al., 2021; Raxworthy and Smith, 2021), and has been successfully recovered from different parts of dried animal mounts, such as skin, bones, muscle, feathers, and toepads (Card et al., 2021). The availability of hDNA from museum specimens has greatly facilitated genetic research by increasing taxonomic sampling and total sample sizes, and is particularly valuable when studying rare or remote species that can be challenging to sample in the wild (Raxworthy and Smith, 2021). Although hDNA is often degraded and early studies typically yielded short fragments of organellar hDNA such as those originating from the mitochondrial genome (mtDNA), recent advances in DNA sequencing technology have spurred a rapid growth in the field of museomics and have facilitated the use of much larger data sets of nuclear DNA markers (Burrell et al., 2015; Card et al., 2021). High-throughput short-read sequencing (HTS) provides a cost-effective way to sequence the mitochondrial or entire nuclear genome (Grayson et al., 2017; Yao et al., 2017), and the rapidly expanding resource of publicly available high-quality genomes has increased the availability of reference genomes from closely related species for read mapping of historical DNA samples (e.g. Teeling et al., 2018; Zhang, 2015). Museomics is therefore becoming established as an important endeavour to offer insights into the evolution and population biology of organisms, especially for endangered or even extinct species (Murray et al., 2017).

Reduced representation genome sequencing, such as transcriptome-based and hybrid enrichment methods, has been facilitated by HTS and has become the dominant genome-scale data collection approach

* Corresponding author at: School of Biological Sciences, Kadoorie Building, Pok Fu Lam Road, University of Hong Kong, Hong Kong Special Administrative Region.
E-mail address: sinyw@hku.hk (S.Y.W. Sin).

https://doi.org/10.1016/j.ympev.2023.107822
Received 12 January 2023; Received in revised form 7 May 2023; Accepted 19 May 2023
Available online 21 May 2023

(Cheon et al., 2020; Lemmon and Lemmon, 2013; Portik et al., 2016). However, transcriptome sequencing requires fresh tissue and hybrid enrichment is constrained by *a priori* genomic resources (Faircloth, 2017; McCormack et al., 2013). For genomic studies of non-model species, the fast decreasing cost of HTS made whole genome sequencing (WGS) more accessible, and it may soon become less expensive than reduced representation genome sequencing (Irestedt et al., 2022). Although WGS at high coverage remains expensive, WGS at low-coverage can be applied to a large number of samples to increase marker number and decrease costs for phylogenomics and population genomics studies. Low-coverage WGS has been increasingly applied to museum specimens, often for organelle genomes assemblies as they are much smaller and present in high copy-number relative to the nuclear genome (Dodsworth, 2015; Hung et al., 2013; Trevisan et al., 2019), whereas the nuclear genome is usually ignored. Although the often low amount of hDNA extracted from historical museum specimens poses a challenge to use this resource, and a number of biases due to large missing data amount and distantly-related reference genome used can potentially influence the phylogenomics inference (Bertels et al., 2014), careful dataset filtering can make historical and modern specimens comparable (Raxworthy and Smith, 2021; Smith et al., 2020). Recent studies have shown that nuclear genomes at coverage as low as <1–2 × can be used for population genomics analyses and phylogenomics studies across large divergence times (Buerkle and Gompert, 2013; Gilly et al., 2019; Olofsson et al., 2019) which, together with the development of appropriate analytical methods to handle low-coverage sequencing data (Korneliussen et al., 2014), opens new possibilities to incorporate specimens from museum collections into genomic analyses. The possibility of retrieving a large number of genome-wide markers by applying low-coverage WGS on hDNA will greatly facilitate phylogenomics studies covering a large number of taxa, and is an area that requires further investigation.

Museum resources and WGS have greatly facilitated avian genomics research (Irestedt et al., 2022), especially with the increasing availability of high-quality avian genomes (Zhang, 2015). Parrots comprise an avian clade with diverse characters, and at the same time many species are threatened with extinction (Berkunsky et al., 2017; Chan et al., 2021). Earlier studies of several extinct parrot species used hDNA from museum specimens to infer their phylogenetic positions based upon a single or few mitochondrial marker(s) (Jackson et al., 2015; Kirchman et al., 2012), and later studies have sequenced the mitogenome or whole nuclear genome from museum parrot specimens for this purpose (Gelabert et al., 2020; Johansson et al., 2018; Olah et al., 2021). In this study, we evaluate the performance of low-coverage WGS for phylogenetic inference of the parrot genus *Agapornis* including both modern and historical samples. The lovebirds (*Agapornis* spp.) are small-sized parrots native to Africa and Madagascar that are among the most popular species in the bird pet trade (Chan et al., 2021; Mori et al., 2020). *Agapornis* comprise nine species (Dilger, 1960; Forshaw and Cooper, 1989; Moreau, 1948): the black-winged lovebird (*A. taranta*), the grey-headed lovebird (*A. canus*), the red-faced lovebird (*A. pullarius*), the black-collared lovebird (*A. swindernianus*), and the peach-faced lovebird (*A. roseicollis*), as well as the four eye-ringed species including the Nyasa lovebird (*A. lilianae*), the Fischer's lovebird (*A. fischeri*), the masked lovebird (*A. personatus*), and the black-cheeked lovebird (*A. nigrigenis*). Members of *Agapornis* display a breadth of phenotypic and behavioural diversity, and a well-resolved phylogeny of this group is essential to reconstruct ancestral character states and to infer the evolutionary history of these characters (Dilger, 1960; Forshaw and Cooper, 1989; Moreau, 1948). For example, five *Agapornis* species are the only parrots, other than *Myiopsitta monachus* (the Monk parakeet), that build nests (Dilger, 1960; Eberhard, 1998), with *A. roseicollis* building cup-shaped nest and the four eye-ringed species building domed nests within cavities. These five *Agapornis* species are also colonial breeders, and together with *A. swindernianus* they are sexually monomorphic (Dilger, 1960; Forshaw and Cooper, 1989; Moreau, 1948). In addition, all *Agapornis* species are native to the African continent, except for *A. canus* that is native to Madagascar (Dilger, 1960; Forshaw and Cooper, 1989; Moreau, 1948). Since *Agapornis* was found to be the sister group to *Loriculus* of Australasia and Indo-Malaysia (Wright et al., 2008), resolving the phylogeny of *Agapornis* is important to understanding the biogeography and evolution of African parrots. However, earlier phylogenetic studies based on morphological or behavioural data or a single mitochondrial gene could not resolve the phylogeny of *Agapornis*, in particular the positions of *A. canus* and *A. swindernianus*, and the relationships among eye-ringed species were also poorly resolved (Dilger, 1960; Eberhard, 1998; Manegold and Podsiadlowski, 2014; Moreau, 1948).

Here, we used low-coverage WGS data to generate both whole mitochondrial genomes and genome-wide nuclear SNPs to attempt to resolve outstanding questions of *Agapornis* phylogenetic relationships. We compared the data quality of modern specimens, including blood, feather, and tissue samples, and historical museum toepad specimens of >100 years of age to demonstrate the power of applying this approach to samples of different DNA qualities. The four eye-ringed species are known to hybridize with each other and with *A. roseicollis* in captivity (Dilger, 1960; Moreau, 1948; Van der Zwan et al., 2019). We therefore also investigated whether genome-wide SNPs generated from both modern and historical specimens would allow us to determine hybrids and their parental origins, to further demonstrate the possibility of using hDNA for admixture and population genomics analyses. Our findings show that using a low-coverage WGS approach to generate both mitogenomes and nuclear SNPs for phylogenomics and admixture analyses can open up new avenues of investigation for large-scale genomics projects that include both modern and historical museum specimens.

2. Material and methods

2.1. Sample collection, DNA extraction and whole-genome sequencing

Samples were collected from 40 individuals representing all nine *Agapornis* species: *A. canus* (n = 3), *A. fischeri* (n = 12), *A. lilianae* (n = 2), *A. nigrigenis* (n = 3), *A. personatus* (n = 6), *A. pullarius* (n = 3), *A. roseicollis* (n = 6), *A. swindernianus* (n = 2), and *A. taranta* (n = 3). Five *Agapornis* hybrids with known parental origins were also included in the analyses: *A. roseicollis* × *A. fischeri* (n = 3), *A. lilianae* × *A. fischeri* (n = 2). Two closely-related species, the Guaiabero (*Bolbopsittacus lunulatus*) and the Blue-crowned hanging parrot (*Loriculus galgulus*), were included as outgroups, making a total of 47 individuals (Electronic Supplementary Material Table S1).

DNA was extracted from either tissue (n = 28; exact tissue type was not always known) or toepad (n = 12) samples from museum specimens, or blood (n = 3) or feather (n = 4) samples supplied by bird owners. Toepad samples were from specimens collected in year 1915–1933. DNA from tissue and blood (on filter paper) were extracted using DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol. DNA from feather samples were extracted using E.Z.N.A. Tissue DNA kit (Omega Biotek, Norcross, GA, USA), with 40 µl DL-Dithiothreitol (DTT; 1 M) added with buffer TL to each sample followed by an incubation at 55 °C for >3 h for sample lysis. DNA quantity was measured with a Qubit dsDNA HS Assay Kit (Invitrogen, Carlsbad, USA) and quality was visualized using agarose gel electrophoresis. DNA was sheared using Covaris S220 for library preparation. For toepad DNA extractions, samples were washed and rehydrated by adding 100 µl of absolute ethanol to the toepad sample, incubating at room temperature for 5 min, removing ethanol, adding 200 µl of ddH_2O, incubating at room temperature for 5 min, and removing the water. Washed toepad samples were then minced with a sterile razor blade, mixed with 180 µl buffer ATL, 40 µl proteinase K, and 20 µl DTT (1 M), and incubated at 50 °C overnight. DNA from lysed toepad samples was extracted using QIAamp DNA Micro kit (Qiagen) and spin columns from QIAquick PCR purification kit (Qiagen). Three elutions of DNA were obtained using

200 µl prewarmed (56 °C) ddH_2O, 50 µl prewarmed ddH_2O, and 50 µl prewarmed buffer AE. The elutions were pooled and vacuum concentrated to < 100 µl. DNA quantity was measured with a Qubit dsDNA HS Assay Kit (Invitrogen, Carlsbad, USA). The quality of extracted toepad DNA was checked using HS D5000 Screen Tape (Agilent), and the average peak DNA fragment size was 131 bp. Whole-genome library preparation for all extracted DNA followed Sin et al. (2022,2020). The DNA libraries were sequenced on Illumina HiSeq 2500 (High Output 250 kit, PE125bp) or NovaSeq (PE150bp) for ~4 × coverage per sample (Lam and Sin, 2020).

2.2. Phylogenomics analyses

Raw paired-end reads were first filtered and trimmed using *fastp* v0.20.1 (Chen et al., 2018) with the following parameters: –correction –detect_adapter_for_pe –cut_front 3 –cut_tail 3 –qualified_quality-phred 20 –average_qual 20 –length_required 35 –unqualified_percent_limit 10 –n_base_limit 0, and then processed following the Genome Analysis Toolkit (GATK v3.8) workflow (Van der Auwera et al., 2013). Reads were merged with Picard toolkit (https://broadinstitute.github.io/picard/) and then mapped onto the chromosome-level genome assembly of *Melopsittacus undulatus* (bMelUnd1.mat.Z; GenBank accession no.: GCA_012275295.1), which is grouped together with *Agapornis* under Loricoloriinae (Schweizer et al., 2010). Reads were mapped to this *M. undulatus* reference genome using BWA-MEM v.0.7.13 (Li and Durbin, 2009) and marked for duplicates using Picard toolkit. Reads were further realigned around indels using GATK IndelRealigner. The base quality scores of the samples were rescaled using mapDamage2 (Jónsson et al., 2013), with the *M. undulatus* genome as the reference. SNP calling and genotyping was done on the rescaled BAM files using ANGSD v0.933 (Korneliussen et al., 2014) with the following parameters: *-minMapQ 30 -minQ 30 -uniqueOnly 1 -remove bads 1 -trim 0 -only_proper_pairs 0 -baq 1 -C 50 -doMaf 1 -doMajorMinor 1 -skipTriAllelic -doCounts 1 -minMaf 0.05 -SNP_pval 1e-6 -minInd 10 -setMinDepth 24 -maxDepth 480 -dumpCounts 2 -doGeno 23 -doPost 1 -postCutoff 0.34 -GL 1 -doGlf 3*. BAQ computation (*-baq 1*) was shown to greatly reduce false SNP calls around misaligned indels (Li, 2011). To test for the effect of sample type on the number and quality of SNP calls, we compared the number of mapped reads, the average mean per-site depth of coverage, and the proportion of missing genotype data between sample type using two-sided t-tests in R. These metrics were computed using samtools idxstats and coverage (samtools v.1.9; Danecek et al., 2021). The significance of the tests was assessed with the Bonferroni correction for multiple pairwise comparisons.

The species phylogeny was inferred from the nuclear SNPs based on the maximum-likelihood (ML) method in RAxML-NG (Kozlov et al., 2019), using the GTR + G nucleotide substitution model with the ascertainment bias correction following Lewis method (Lewis, 2001) determined by ModelTest-NG (Darriba et al., 2020). After testing different missing data thresholds applied to the whole SNP matrix, e.g. excluding SNPs absent in more than 30% of samples (Smith et al., 2020), we generated a comparable dataset by filtering out SNPs that were missing in the top three samples that have the highest proportion of missing data, i.e. the two *A. swindernianus* toepad samples and the *A. lilianae* toepad sample "Alilianae_S62" (Electronic Supplementary Table S2), resulting in 17,025 SNPs. This threshold allowed us to reduce the uneven distribution of missing data across species while retaining enough phylogenetic information to confidently resolve the species-level phylogeny. Tree searches were done using 10 random and 10 parsimony trees as starting trees. Node supports were computed from 2 × 100 bootstraps on two independent seeds and checked for convergence before annotating the best scoring ML-tree.

To compare with the phylogeny inferred from nuclear markers and to determine the maternal species of the interspecific hybrids, we assembled the mitochondrial genomes of the 47 samples. Raw reads were first trimmed from adapters and quality-filtered using *fastp* v0.20.1 (Chen et al., 2018). The mitochondrial genome assembly of each sample was then performed with MITObim v1.7 (Hahn et al., 2013) following the two-step procedure, consisting of an initial reference-based mapping assembly using MIRA 4 (Chevreux et al., 2004) followed by a baiting and iterative mapping of the raw reads onto the generated *de novo* assembly using MITObim script. All available mitochondrial genomes of *Agapornis* spp. on NCBI were retrieved and the most closely-related species was used as the reference genome for each sample (see Electronic Supplementary Material Table S1). The assembled mitochondrial genomes were then iteratively aligned using PASTA (Mirarab et al., 2015), with MAFFT as aligner (Katoh and Standley, 2013), MUSCLE as merger (Edgar, 2004) and FastTree 2 as tree estimator (Price et al., 2010). The phylogeny was inferred using RAxML-NG (Kozlov et al., 2019) from the aligned mitochondrial genomes after trimming for gaps using *trimal* (Capella-Gutiérrez et al., 2009), using the GTR + I + G4 model selected by ModelTest-NG (Darriba et al., 2020). Tree searches were similarly run and node supports were estimated from 2x200 bootstraps on two independent seeds. The best tree was annotated with the bootstrapped supports before visualized in FigTree v.1.4.3.

2.3. Genetic clustering and admixture of Agapornis hybrids

To identify and determine the admixture of the hybrids, we performed a principal component analysis (PCA) using the called genotypes inferred from ANGSD, from which a covariance matrix was calculated using *ngsCovar* from the *ngsTools* package (Fumagalli et al., 2014) and then used as input for PCA performed in R. The analyses included all the hybrids and their expected parental species only (i.e. *A. roseicollis* and all eye-ringed species, referred to as "dataset Aros + B" hereafter). Considering the much greater phylogenetic distance between *A. roseicollis* and the eye-ringed species than among the eye-ringed species, the analyses were also run on a dataset that only includes the eye-ringed species (referred to as "dataset AgaB" hereafter) in order to better assess the admixture of the hybrids between eye-ringed species. The datasets included 744,407 SNPs for "Aros + B" and 385,238 SNPs for "AgaB" respectively.

Additionally, admixture analyses were performed to assess the level of introgression within each hybrid using *NGSadmix* from the *ngsTools* package, which is similar to STRUCTURE with the admixture model (Pritchard et al., 2000). Using the same genotypes as for the PCAs, we tested the number of genetic clusters (K), ranging from K = 1 to K = 10 for the dataset "Aros + B", and from K = 1 to K = 8 for the dataset "AgaB". *NGSadmix* was run under default parameters with 10 iterations for each K and the best K chosen based on the *deltaK* method from Evanno et al. (2005), i.e. by calculating the second order rate of change of the likelihood Ln P(D) between each K and assessing the best K as the number of K for which the *deltaK* is maximized. The iteration runs were summarized for each K and *deltaK* using CLUMPAK (Kopelman et al., 2015).

3. Results

3.1. Sequencing and read mapping quality of different sample types

High-throughput sequencing yielded a total of 925,089,795 paired-end reads (2 × 125 bp) mapped with on average 2 × per-site read coverage (Electronic Supplementary Material Table S2). The number of reads mapped and the mean read depth varied between sample type (Fig. 1 and Electronic Supplementary Material Table S2). The samples with the highest to lowest read number and depth were blood, feather, tissue, and toepad samples. The proportion of missing data was inversely correlated with read coverage. Toepad samples had the highest proportions of missing data (mean = 64.90%, range = 47.02%–76.84%), whereas the mean was 18.34% among other sample types.

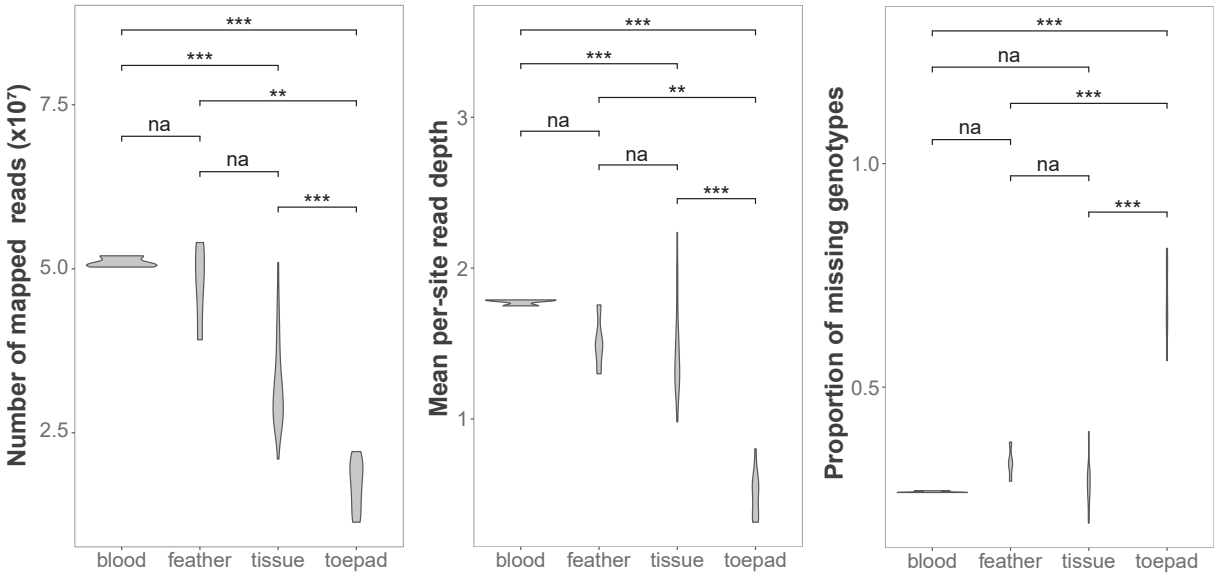

Fig. 1. Comparison of read mapping, read depth, and proportion of missing data among modern and historical samples. The number of mapped reads, mean per-site read depth, and proportion of missing genotypes for blood, feather, tissue, and toepad (~100 years old) samples are shown. X-axis indicates sample types from which the DNA was extracted. Pairwise two-sided t-tests were performed and significance after applying the Bonferroni correction for multiple tests is indicated as: (*) < 0.001, (**) < 0.01, (*) < 0.05, (na) >= 0.05. Detailed data can be found in Electronic Supplementary Material Table S2.

3.2. Phylogenetic relationships

The *de novo* mitochondrial genome assemblies had an average length of 17,264 bp. After sequence alignment and gap trimming, the aligned mitochondrial genomes were 16,455 bp in length. The mitochondrial genome tree was overall well supported, with most bootstrap values >90 for internal nodes (Fig. 2). *A. taranta* and *A. pullarius* formed a monophyletic clade, as well as the four eye-ringed species (*A. lilianae*, *A. nigrigenis*, *A. personatus*, and *A. fischeri*), with *A. roseicollis* being the sister species of the eye-ringed species. *A. swinderianus* was the sister species to the clade comprising *A. taranta*, *A. pullarius*, *A. roseicollis*, and the eye-ringed species. *A. canus* was the sister species to all other *Agapornis* species. The closely-related *A. lilianae* and *A. nigrigenis*, as well as *A. fischeri* and *A. personatus*, remained unresolved based on the mitochondrial genomes. The hybrids were clustered with their expected maternal species, i.e. *A. roseicollis* × *A. fischeri* (abbreviated "ArosxAfis") with *A. roseicollis* and *A. lilianae* × *A. fischeri* (abbreviated "AlilxAfis") with *A. lilianae*.

SNP calling with ANGSD resulted in 2,921,324 SNPs across the 47 samples. The filter to exclude SNPs absent in more than 30% of samples resulted in 285,556 SNPs, but some toepad samples still showed a high proportion of missing data. Subsampling of SNPs to minimize missing data in the top three samples generated a comparable dataset containing 17,025 SNPs. The species phylogeny based on the nuclear genome-wide SNPs was consistent with the mitochondrial phylogeny (Fig. 3), but the nuclear SNPs provided enough power to differentiate the eye-ringed species, i.e. *A. fischeri* and *A. personatus* as sister-species, albeit with low bootstrap support. Although *A. nigrigenis* forms a monophyletic clade, individuals of *A. lilianae* did not group together in a monophyletic clade and one individual forms a low-supported clade with *A. nigrigenis*, likely due to small divergence between the two species and large proportion of missing data of the *A. lilianae* toepad samples.

3.3. Genetic clustering and admixture analyses of hybrids

The PCA (Fig. 4) and admixture (Fig. 5) analyses were largely consistent in differentiating the *Agapornis* species. PCA clearly separated *A. roseicollis* and the four eye-ringed species (Fig. 4A), and also among all the eye-ringed species (Fig. 4B). NGSadmix followed by the *ad hoc* Evanno's method assessed K = 2 and K = 3 as the optimal number of genetic clusters for "Aros + B" and "AgaB" datasets, respectively (Supplementary Material Fig. S1). The admixture analysis also differentiated *A. roseicollis* and the eye-ringed species (Fig. 5A), and among the eye-ringed species except for *A. lilianae* and *A. nigrigenis* (Fig. 5B). The PCA (Fig. 4) and admixture (Fig. 5) analyses consistently identified the admixture of the F1 hybrids, i.e. "ArosxAfis" and "AlilxAfis". The admixture analyses showed equal proportion of admixture in all hybrids from their corresponding parental species (Fig. 5), which was also observed in the PCA with the hybrids located in-between their parental species (Fig. 4).

4. Discussion

4.1. Biogeography and evolution of Agapornis

Based on the mitogenomes and genome-wide SNPs, we have revised the phylogeny of *Agapornis* compared to earlier studies. Our phylogeny differs from an earlier study (Manegold and Podsiadlowski, 2014) that also included a historical toepad sample from *A. swindernianus* and suggested it to be the sister to all other *Agapornis* species, and led to the conclusion that the ancestor of *Agapornis* was an arboreal forest-dweller that originated on the African continent and later colonized Madagascar from Africa. Instead, our placement of *A. canus* as the sister taxon to all other *Agapornis* species revises our understanding of the biogeography and evolution of these African parrots. Only four endemic genera of parrot occur on Africa, including *Agapornis* and *Coracopsis* that are also distributed on Madagascar (Schweizer et al., 2010). The *Agapornis* genus is the sister group to *Loriculus* found in Australasia and Indo-Malaysia, and both group with the Australasian *Melopsittacus*, Loriinae, and Cyclopsittacini to form a monophyletic clade (Schweizer et al., 2010). The common ancestor of *Agapornis* thus likely originated from Australasia. *A. canus* is the only *Agapornis* species endemic to Madagascar, whereas all other species in this genus are distributed on the African continent (Dilger, 1960; Forshaw and Cooper, 1989; Moreau, 1948). Our finding of the early divergence between *A. canus* and other *Agapornis* species is therefore consistent with the hypothesis (Schweizer et al., 2010; Schweizer et al., 2011) proposing that the ancestor of *Agapornis* colonized Madagascar from Australasia through long-distance *trans-*

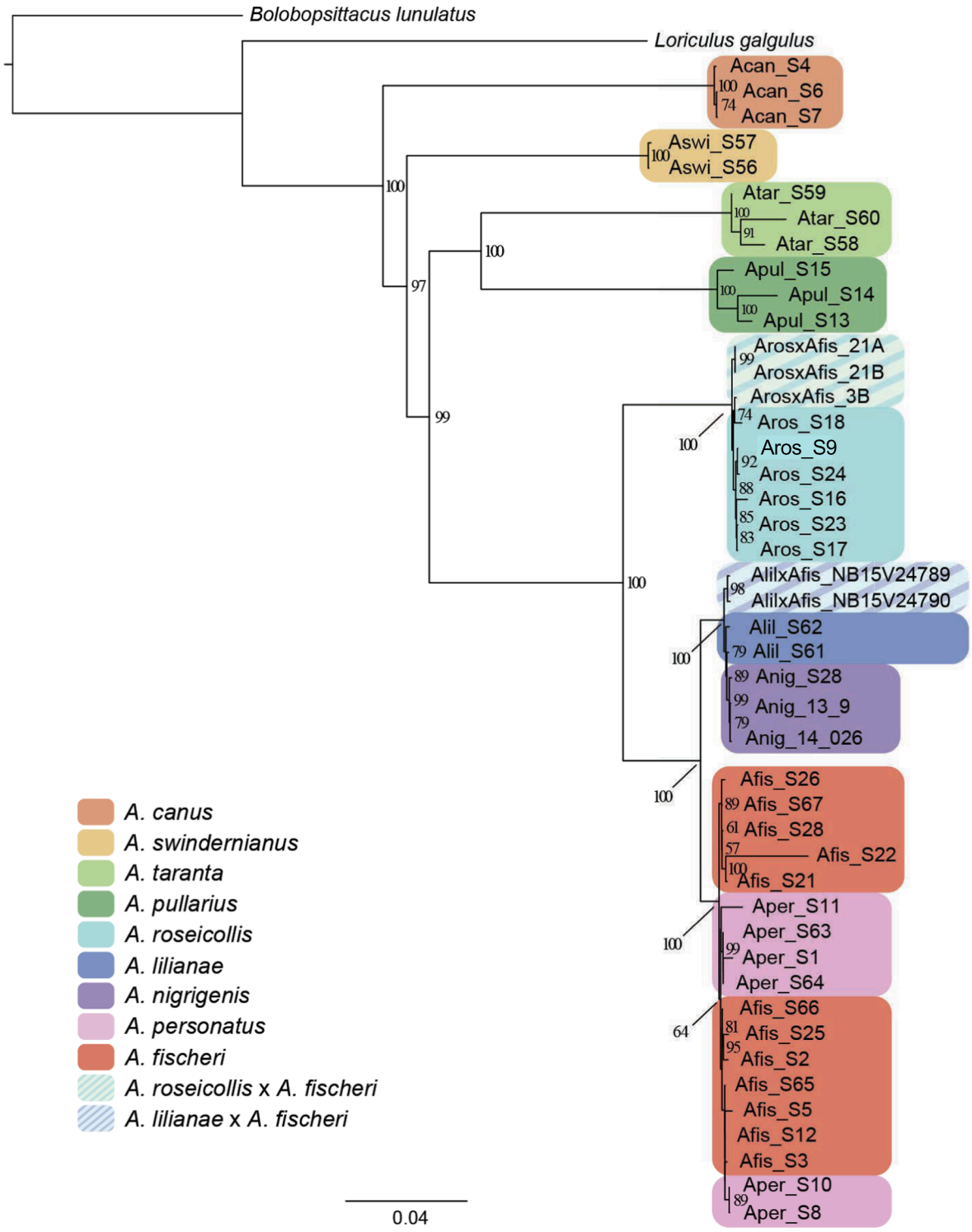

Fig. 2. Phylogenetic tree of *Agapornis* species based on the mitochondrial genome. Alignment of whole mitochondrial genome of each 47 individuals without gaps was used for the inferences with RAxML. Node supports were based on 400 bootstraps. Substitutions per site is indicated by the scale at the bottom. Species are indicated by plain rectangles and hybrids by stripped colored rectangles following the legend.

oceanic dispersal events, despite Schweizer et al. (2010, 2011) not having included *A. swindernianus*, *A. pullarius*, and *A. taranta* in their studies. After colonizing Madagascar, *A. canus* was isolated on Madagascar and diverged from the common ancestor of other lovebirds, which likely dispersed from Madagascar to sub-Saharan Africa and further diversified into the eight extant species on the African mainland.

Our phylogeny is also different from the relationship inferred based on morphological and behavioural characters (Manegold and Podsiadlowski, 2014), which proposes *A. canus*, *A. pullarius*, and *A. taranta* that share characters such as male black underwings to be in the same clade and *A. swindernianus* to be the sister to all other *Agapornis* species. *A. swindernianus* is the only arboreal forest-dweller among the species in this genus (Dilger, 1960; Forshaw and Cooper, 1989; Moreau, 1948). All the other species inhabit wooded grassland and are granivorous. The phylogenetic position of *A. swindernianus*, which is not sister to all other *Agapornis* species as proposed in Manegold and

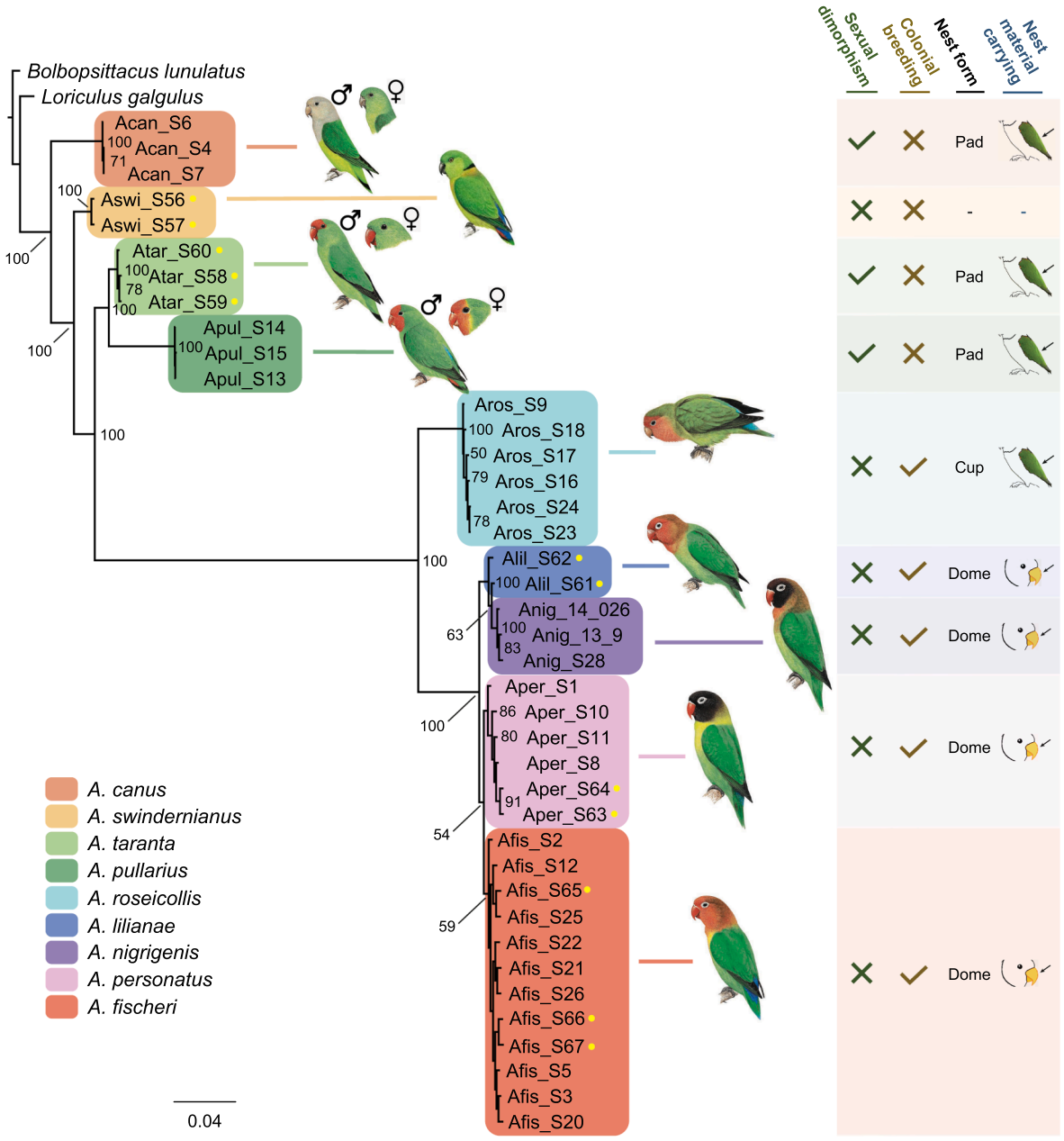

Fig. 3. Phylogenetic tree of Agapornis species based on nuclear genome-wide SNPs. The species tree was inferred by RAxML based on a total of 17,025 SNPs. Node supports, based on 300 bootstraps, are indicated on the figure. Some characteristic traits (Eberhard 1998) are shown on the right panel: sexual dimorphism, colonial breeding, form of the nest, and method of carrying nest material (in between contour feathers or in the beak). Yellow dot indicates toepad sample. Substitutions per site is indicated by the scale at the bottom. Illustrations of lovebirds were reproduced with the permission of Lynx Edicions.

Podsiadlowski (2014), suggests that the specialization to live in tropical rain forest is a derived trait in this genus. This is supported by the morphological character that all *Agapornis* species, including *A. swindernianus*, lack an osseous extremitas sternalis furculae (Mayr, 2010), which is correlated with a shift towards a granivorous diet in other parrots. The common ancestor of *Agapornis* spp. was thus likely a granivore, similar to all other species in this genus except for *A. swindernianus* that became adapted to forests.

A. canus, *A. taranta*, and *A. pullarius*, like the closely related outgroups *Loriculus* and *Bolbopsittacus*, are sexually dimorphic; whereas *A. swindernianus*, *A. roseicollis*, and the four eye-ringed species are sexually monomorphic (Dilger, 1960; Moreau, 1948). Sexual monomorphism therefore likely evolved at least twice from sexual dimorphism independently, in *A. swindernianus* and the common ancestor of *A. roseicollis* and the 4 eye-ringed species. Except for *A. swindernianus*, the other five sexually monomorphic species are colonial breeders (Dilger, 1960). Dilger (1960) proposed that evolution of sexual monomorphism was due to the colonial breeding behavior in these species, which exerts nonsexual social selection (West-Eberhard, 1983) on plumage pattern and color to facilitate group activity coordination. Colonial breeding is also associated with mobbing behavior and nest building (Dilger, 1960; Eberhard, 1998). In the colonial breeders, nest material carrying habits and nest forms also vary among species. *A. roseicollis* carries nesting material in their rump feathers and builds cup-shaped nests, whereas the 4 eye-ringed species carry nesting material in their beaks and build dome-shaped nests (Dilger, 1960; Eberhard, 1998). These derived characters in the colonial breeding *Agapornis* species make them an interesting group to study the evolution of parrot phenotypes and behaviors. Given the colonial breeder species can hybridize, they offer a system allowing interspecific crossing in the

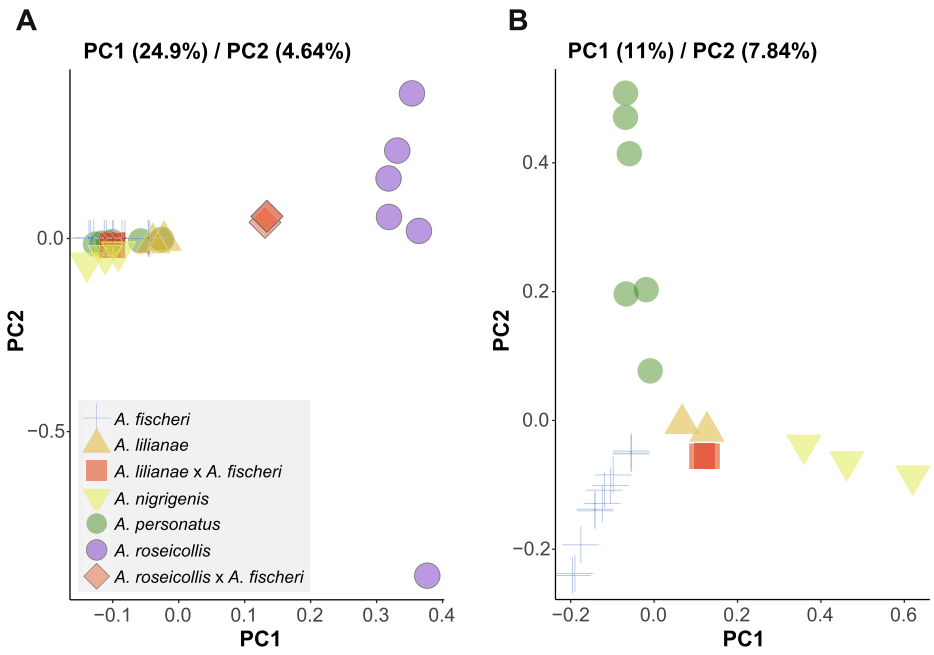

Fig. 4. Genetic clustering of *Agapornis* species and of their hybrids based on PCA. The first two principal components (PCs) are represented here. The percentage of total variance explained by the genomic data are indicated for each PC on top of the plot. Two types of hybrids were present: *A. roseicollis* × *A. fischeri* and *A. lilianae* × *A. fischeri*. PCAs were run (A) on the dataset "Aros + B" of 744,407 SNPs including all the hybrids, *A. roseicollis* and the eye-ringed species (B) and on the dataset "AgaB" of 385,238 SNPs including only *A. lilianae* × *A. fischeri* hybrids and the eye-ringed species.

laboratory for studying factors affecting their characters.

The four eye-ringed species are relatively shallow in their divergence. Our result clearly separates *A. lilianae* + *A. nigrigenis* and *A. personatus* + *A. fischeri*. Among these 4 eye-ringed species, only *A. personatus*, *A. fischeri*, and *A. nigrigenis* form monophyletic clades in the species tree based on genome-wide SNPs. The unresolved relationship between *A. lilianae* and *A. nigrigenis* in the phylogenetic trees can be due to their recent divergence time and/or the low number of markers available for the two *A. lilianae* toepad samples. Nevertheless, PCA is able to separate the 4 species well. Although the species status of these 4 species was questioned (Hampe, 1939) and *A. nigrigenis* was proposed to be a subspecies of *A. lilianae* (Moreau, 1948), they are genetically distinct from each other. Given all 4 eye-ringed species mate readily in captivity and are interfertile (Moreau, 1948), the observation that there is no obvious geographical barrier between the allopatric *A. lilianae* and *A. nigrigenis* and no hybrid was found between them in the wild leave the isolation mechanism an unknown. *A. lilianae* and *A. nigrigenis* only differ in the degree of melanism, and future studies of population genetics and possible introgression between natural populations of the two species will shed light on the speciation process of these parrots.

4.2. *Use of museum specimens and low-coverage WGS for phylogenomics and admixture analyses*

In this study, we combined WGS dataset of low coverage (~2 ×) from both historical (~100 years old) and modern specimens to evaluate its performance to infer phylogenetic relationships among a group of parrots with divergence time ranges from <1 MYA to more than 20 MYA (Kundu et al., 2012; Schweizer et al., 2011). The phylogeny reconstructed using mitogenomes (Fig. 2), which is one of the most widely used markers for phylogenomics analysis due to its high copy number relative to the nuclear genome, is consistent with the phylogeny inferred using nuclear SNPs (Fig. 3). This shows the advantage of using low-coverage WGS to not only obtain the mitogenome but also genome-wide nuclear SNPs for robust phylogeny reconstruction (Bruxaux et al., 2018).

Although hDNA was successfully extracted from toepad samples originating from museum specimens prepared in 1915–1933, the sample type had an influence on the data quality, indicated by the number of mapped reads, per-site coverage, and missing data (Fig. 1). Extracted DNA from modern samples, irrespective of whether the source material was blood, feather, or tissue, had higher data quality than that originating from toepads. The fragmented hDNA from historical museum specimens poses challenges for the use of this resource (Billerman and Walsh, 2019; Raxworthy and Smith, 2021), and it was shown that specimen preparation and storage conditions, as well as the species itself, have an impact on the hDNA quality (Billerman and Walsh, 2019; Card et al., 2021; Raxworthy and Smith, 2021). For example, toepads from larger birds tend to have more degraded DNA, likely due to the slower drying of the specimen (Irestedt et al., 2022; Tsai et al., 2020). It is also notable that the modern samples show a wide variation in quality among sample types, such as the significant differences between blood and tissue samples. The variation in quality was in particular greater among the tissue samples relative to the other modern sample types. In addition to the difference in DNA degradation rate among tissue types (Blom, 2021; Card et al., 2021), factors such as varying ages of the material, suboptimal sampling conditions (e.g. not harvesting tissues immediately after death), storage media and conditions, and the number of freeze/thaw cycles gone through could affect the quality of museum tissue collections (Blom, 2021). Since museum specimens are of great value in genomic studies, there is an imperative need to establish best practices to standardize museum sample collection (Card et al., 2021).

Given the wide variation in quality among historical museum specimens, different filtering practices have thus been proposed to handle the WGS data from hDNA (Irestedt et al., 2022), such as subsampling loci that have more comparable missing data in order to include data from modern and historical specimens in the same study (Raxworthy and Smith, 2021; Smith et al., 2020). For example, the high proportion of missing data in the *A. swindernianus* toepad samples had an impact on the *Agapornis* phylogeny inferred. When we only filtered to exclude SNPs absent in more than 30% of samples (Smith et al., 2020), some toepad samples still showed a high proportion of missing data (reaching ~70% for *A. swindernianus* toepads), resulting in a spurious and poorly supported position of *A. swindernianus* (Supplementary Material Fig. S2A) even though the dataset contains 285,556 SNPs. Subsampling of SNPs to minimize missing data in those toepad samples is required to generate a comparable dataset that infers a robust phylogeny (Supplementary Material Fig. S2B), which is also consistent with the mitogenome tree. The accurate grouping of modern and historical samples from the same species together, such as those in the *A. fischeri* and *A. personatus* clades

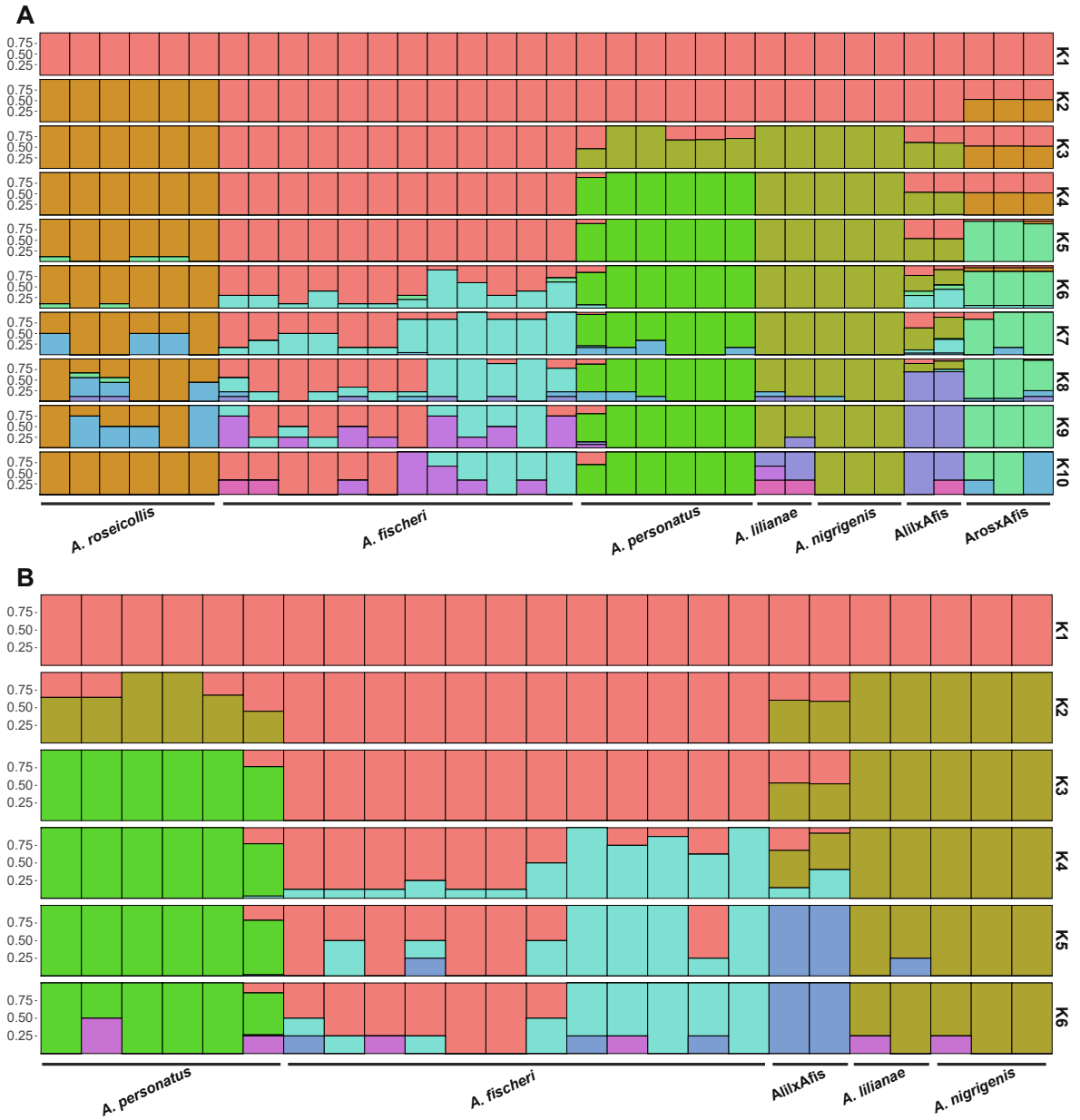

Fig. 5. Genetic structuring and admixture among Agapornis species and their hybrids. Inferences were done by NGSadmix, with the most likely number of genetic cluster (K) chosen following deltaK method from Evanno et al. (2005). Two types of hybrids were present: A. roseicollis × A. fischeri (ArosxAfis) and A. lilianae × A. fischeri (AlilxAfis). The analyses were run for (A) the dataset "Aros + B" of 744,407 SNPs including all hybrids, A. roseicollis and the eye-ringed species (B) and the dataset "AgaB" of 385,238 SNPs including only the AlilxAfis hybrids and the eye-ringed species. The best K estimated was K = 2 and K = 3 for datasets A and B respectively (Supplementary Material Fig. S1).

(Fig. 3), also shows the power and potential of applying low-coverage WGS on historical museum samples for phylogenomics studies.

The genome-wide SNPs generated from hDNA also allows population genomics analyses. Our PCA and the admixture analyses could identify the species and parental origins of hybrids (Figs. 4 and 5). All samples from *A. roseicollis* and the 4 eye-ringed species form their own clusters in both analyses. An exception is *A. lilianae* and *A. nigrigenis* that could not be differentiated well in the admixture analysis, which was likely due to both the recent divergence time between these two species and an insufficient number of markers to provide adequate resolution. However, PCA can separate these two species, suggesting that a larger marker panel could yield more definitive results across analyses. The parental origin of the hybrid individuals could be identified clearly (Fig. 5), with approximately half of the ancestry assigned to the paternal species and half from the maternal species, consistent with the maternal species identified by the mitogenome and the information provided by the bird owners. Since the eye-ringed species inhabit contiguous areas and the possibility that they (e.g. *A. lilianae* and *A. nigrigenis*) may meet at the boundary of their ranges cannot be excluded (Moreau, 1948), our approach can be used to identify hybrids occurring in the wild. Recent population genomics studies also demonstrate the use of hDNA data from museum samples (Ericson et al., 2019; Ernst et al., 2022; Ewart et al., 2019; Garg et al., 2022). In a study that investigates the biogeography of two Southeast Asian songbirds (*Pellorneum* spp.), thousands of SNPs were sequenced from samples including specimens greater than a hundred years old to reveal the gene flow dynamics between populations (Garg et al., 2022). Another study on the extinct passenger pigeon (*Ectopistes migratorius*) has generated high-coverage mitogenomes and nuclear genomes originated from museum specimens to understand the drive of its loss in genetic diversity (Murray et al., 2017). For future

population and species level analyses, such as investigations of introgression between species, low-coverage WGS could be an effective approach to genotype a large number of modern and museum samples (Buerkle and Gompert, 2013; Irestedt et al., 2022). The potential of this approach for population genomics analyses involving historical samples would be a promising and important area to explore.

5. Conclusions

We demonstrate the application of low-coverage WGS in genomics studies combining historical and modern samples. Toepad samples of avian specimens in museum collections are valuable sources of hDNA for phylogenomics and population genomics studies (Moncrieff et al., 2022), and have allowed us to include all *Agapornis* species in this study. Low-coverage WGS can be applied to hDNA to reduce cost and obtain a large number of markers. In addition, this approach can generate genomic data to isolate markers from both the mitochondrial and nuclear genomes. As mitochondrial DNA is very abundant in cells, it is often the target for studies involving historical samples. At the same time, the same approach can generate useful data from the nuclear genome for genomics analyses that require higher resolution. We show that the same WGS dataset allows phylogenetic inference based on mitogenomes or genome-wide SNPs, allowing more in-depth analyses and cross-validation of the results. By using low-coverage WGS and including historical museum samples, this study has resolved the phylogeny of the lovebirds. Importantly, the revised phylogeny has shed light on the biogeography and evolution of this group of African parrots and reveals that the ancestor of *Agapornis* had likely colonized Madagascar from Australasia by *trans*-oceanic dispersal events before they dispersed to the African continent and underwent diversification with the development of species-specific characters and adaptations.

CRediT authorship contribution statement

Stella Huynh: Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Alison Cloutier:** Formal analysis, Writing – review & editing. **Simon Yung Wa Sin:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Supervision, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The sequencing data has been archived in NCBI under the BioProject accession number PRJNA971241.

Acknowledgement

We would like to acknowledge the American Museum of Natural History, Field Museum of Natural History, Louisiana State University Museum of Zoology, Museum of Comparative Zoology at Harvard University, the University of Kansas Biodiversity Institute & Natural History Museum, San Diego Natural History Museum, and the University of Washington Burke Museum for providing tissue and toepad samples; and Evy Dens and other bird owners for providing feather and blood samples. We also thank Emily Shui Kei Poon for the assistance with laboratory work. This project was supported by the startup fund granted to SYWS. The computations were performed using research computing facilities offered by the Information Technology Services, the University of Hong Kong.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ympev.2023.107822.

References

Berkunsky, I., Quillfeldt, P., Brightsmith, D.J., Abbud, M., Aguilar, J., Alemán-Zelaya, U., Aramburú, R.M., Arias, A.A., McNab, R.B., Balsby, T.J., 2017. Current threats faced by Neotropical parrot populations. Biol. Conserv. 214, 278–287.
Bertels, F., Silander, O.K., Pachkov, M., Rainey, P.B., Van Nimwegen, E., 2014. Automated reconstruction of whole-genome phylogenies from short-sequence reads. Mol. Biol. Evol. 31, 1077–1088.
Billerman, S.M., Walsh, J., 2019. Historical DNA as a tool to address key questions in avian biology and evolution: a review of methods, challenges, applications, and future directions. Mol. Ecol. Resour. 19, 1115–1130.
Blom, M.P., 2021. Opportunities and challenges for high-quality biodiversity tissue archives in the age of long-read sequencing. Mol. Ecol. 30, 5935–5948.
Bruxaux, J., Gabrielli, M., Ashari, H., Pŕŷs-Jones, R., Joseph, L., Milá, B., Besnard, G., Thébaud, C., 2018. Recovering the evolutionary history of crowned pigeons (Columbidae: Goura): Implications for the biogeography and conservation of New Guinean lowland birds. Mol. Phylogenet. Evol. 120, 248–258.
Buerki, S., Baker, W.J., 2016. Collections-based research in the genomic era. Biol. J. Linn. Soc. 117, 5–10.
Buerkle, C., Gompert, Z., 2013. Population genomics based on low coverage sequencing: how low should we go? Mol. Ecol. 22, 3028–3035.
Burrell, A.S., Disotell, T.R., Bergey, C.M., 2015. The use of museum specimens with high-throughput DNA sequencers. J. Hum. Evol. 79, 35–44.
Capella-Gutiérrez, S., Silla-Martínez, J.M., Gabaldón, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972–1973.
Card, D.C., Shapiro, B., Giribet, G., Moritz, C., Edwards, S.V., 2021. Museum genomics. Annu. Rev. Genet. 55, 633–659.
Chan, D.T.C., Poon, E.S.K., Wong, A.T.C., Sin, S.Y.W., 2021. Global trade in parrots–influential factors of trade and implications for conservation. Global Ecol. Conserv. 30, e01784.
Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34, i884–i890.
Cheon, S., Zhang, J., Park, C., 2020. Is phylotranscriptomics as reliable as phylogenomics? Mol. Biol. Evol. 37, 3672–3683.
Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A.J., Müller, W.E., Wetter, T., Suhai, S., 2004. Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs. Genome Res. 14, 1147–1159.
Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H., 2021. Twelve years of SAMtools and BCFtools. Gigascience 10 (2), giab008.
Darriba, D., Posada, D., Kozlov, A.M., Stamatakis, A., Morel, B., Flouri, T., 2020. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. Mol. Biol. Evol. 37, 291–294.
Dilger, W.C., 1960. The comparative ethology of the African parrot genus *Agapornis*. Zeitschrift für Tierpsychologie.
Dodsworth, S., 2015. Genome skimming for next-generation biodiversity analysis. Trends Plant Sci. 20, 525–527.
Eberhard, J.R., 1998. Evolution of nest-building behavior in *Agapornis* parrots. Auk 115, 455–464.
Edgar, R., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucl. Acids Res. 32, 1792–1797.
Ericson, P.G., Qu, Y., Rasmussen, P.C., Blom, M.P., Rheindt, F.E., Irestedt, M., 2019. Genomic differentiation tracks earth-historic isolation in an Indo-Australasian archipelagic pitta (Pittidae; Aves) complex. BMC Evol. Biol. 19, 1–13.
Ernst, M., Jønsson, K.A., Ericson, P.G., Blom, M.P., Irestedt, M., 2022. Utilizing museomics to trace the complex history and species boundaries in an avian-study system of conservation concern. Heredity 128, 159–168.
Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Mol. Ecol. 14, 2611–2620.
Ewart, K.M., Johnson, R.N., Ogden, R., Joseph, L., Frankham, G.J., Lo, N., 2019. Museum specimens provide reliable SNP data for population genomic analysis of a widely distributed but threatened cockatoo species. Mol. Ecol. Resour. 19, 1578–1592.
Faircloth, B.C., 2017. Identifying conserved genomic elements and designing universal bait sets to enrich them. Methods Ecol. Evol. 8, 1103–1112.
Forshaw, J.M., Cooper, W.T., 1989. Parrots of the world. JSTOR.
Fumagalli, M., Vieira, F.G., Linderoth, T., Nielsen, R., 2014. ngsTools: methods for population genetics analyses from next-generation sequencing data. Bioinformatics 30, 1486–1487.
Garg, K.M., Chattopadhyay, B., Cros, E., Tomassi, S., Benedick, S., Edwards, D.P., Rheindt, F.E., 2022. Island biogeography revisited: museomics reveals affinities of shelf island birds determined by bathymetry and paleo-rivers, not by distance to mainland. Mol. Biol. Evol. 39, msab340.
Gelabert, P., Sandoval-Velasco, M., Serres, A., de Manuel, M., Renom, P., Margaryan, A., Stiller, J., De-Dios, T., Fang, Q., Feng, S., 2020. Evolutionary history, genomic adaptation to toxic diet, and extinction of the Carolina parakeet. Curr. Biol. 30 (108–114), e105.
Gilly, A., Southam, L., Suveges, D., Kuchenbaecker, K., Moore, R., Melloni, G.E., Hatzikotoulas, K., Farmaki, A.-E., Ritchie, G., Schwartzentruber, J., 2019. Very low-

depth whole-genome sequencing in complex trait association studies. Bioinformatics 35, 2555–2561.
Grayson, P., Sin, S.Y.W., Sackton, T.B., Edwards, S.V., 2017. Comparative Genomics as a Foundation for Evo-Devo Studies in Birds. Humana Press, New York, NY.
Habel, J.C., Husemann, M., Finger, A., Danley, P.D., Zachos, F.E., 2014. The relevance of time series in molecular ecology and conservation biology. Biol. Rev. 89, 484–492.
Hahn, C., Bachmann, L., Chevreux, B., 2013. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. Nucl. Acids Res. 41, e129–e.
Hampe, H., 1939. Zur Systematik von Agapornis. Ornithol. Monatsber. 47, 4–6.
Holmes, M.W., Hammond, T.T., Wogan, G.O., Walsh, R.E., LaBarbera, K., Wommack, E. A., Martins, F.M., Crawford, J.C., Mack, K.L., Bloch, L.M., 2016. Natural history collections as windows on evolutionary processes. Mol. Ecol. 25, 864–881.
Hung, C.-M., Lin, R.-C., Chu, J.-H., Yeh, C.-F., Yao, C.-J., Li, S.-H., 2013. The *de novo* assembly of mitochondrial genomes of the extinct passenger pigeon (*Ectopistes migratorius*) with next generation sequencing. PLoS One 8, e56301.
Irestedt, M., Thörn, F., Müller, I.A., Jønsson, K.A., Ericson, P.G., Blom, M.P., 2022. A guide to avian museomics: Insights gained from resequencing hundreds of avian study skins. Mol. Ecol. Resour. 22, 2672–2684.
Jackson, H., Jones, C.G., Agapow, P.M., Tatayah, V., Groombridge, J.J., 2015. Micro-evolutionary diversification among I ndian O cean parrots: temporal and spatial changes in phylogenetic diversity as a consequence of extinction and invasion. Ibis 157, 496–510.
Johansson, U.S., Ericson, P.G., Blom, M.P., Irestedt, M., 2018. The phylogenetic position of the extinct Cuban Macaw *Ara tricolor* based on complete mitochondrial genome sequences. Ibis 160, 666–672.
Jónsson, H., Ginolhac, A., Schubert, M., Johnson, P.L., Orlando, L., 2013. mapDamage2. 0: fast approximate Bayesian estimates of ancient DNA damage parameters. Bioinformatics 29, 1682–1684.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780.
Kirchman, J.J., Schirtzinger, E.E., Wright, T.F., 2012. Phylogenetic relationships of the extinct Carolina Parakeet (*Conuropsis carolinensis*) inferred from DNA sequence data. Auk 129, 197–204.
Kopelman, N.M., Mayzel, J., Jakobsson, M., Rosenberg, N.A., Mayrose, I., 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. Mol. Ecol. Resour. 15, 1179–1191.
Korneliussen, T., Albrechtsen, A., Nielsen, R., 2014. ANGSD: analysis of next generation sequencing data. BMC Bioinf. 15, 356.
Kozlov, A.M., Darriba, D., Flouri, T., Morel, B., Stamatakis, A., 2019. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics 35, 4453–4455.
Kundu, S., Jones, C.G., Prys-Jones, R.P., Groombridge, J.J., 2012. The evolution of the Indian Ocean parrots (Psittaciformes): extinction, adaptive radiation and eustacy. Mol. Phylogenet. Evol. 62, 296–305.
Lam, D.K., Sin, S.Y.W., 2020. Development and characterization of microsatellite markers in Rosy-faced and other lovebirds (*Agapornis* spp.) using next-generation sequencing. Mol. Biol. Rep. 47 (8), 6417–6427.
Lemmon, E.M., Lemmon, A.R., 2013. High-throughput genomic data in systematics and phylogenetics. Annu. Rev. Ecol. Evol. Syst. 44, 99–121.
Lewis, P.O., 2001. A likelihood approach to estimating phylogeny from discrete morphological character data. Syst. Biol. 50, 913–925.
Li, H., 2011. Improving SNP discovery by base alignment quality. Bioinformatics 27, 1157–1158.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.
Manegold, A., Podsiadlowski, L., 2014. On the systematic position of the Black-collared Lovebird *Agapornis swindernianus* (Agapornithinae, Psittaciformes). J. Ornithol. 155, 581–589.
Mayr, G., 2010. Parrot interrelationships–morphology and the new molecular phylogenies. Emu 110, 348–357.
McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C., Brumfield, R.T., 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Mol. Phylogenet. Evol. 66, 526–538.
Mirarab, S., Nguyen, N., Guo, S., Wang, L.-S., Kim, J., Warnow, T., 2015. PASTA: ultra-large multiple sequence alignment for nucleotide and amino-acid sequences. J. Comput. Biol. 22, 377–386.
Moncrieff, A.E., Faircloth, B.C., Brumfield, R.T., 2022. Systematics of Lepidothrix manakins (Aves: Passeriformes: Pipridae) using RADcap markers. Mol. Phylogenet. Evol. 107525.
Moreau, R., 1948. Aspects of evolution in the parrot genus *Agapornis*. Ibis 90, 206–239.
Mori, E., Cardador, L., Reino, L., White, R.L., Hernández-Brito, D., Le Louarn, M., Mentil, L., Edelaar, P., Pârău, L.G., Nikolov, B.P., 2020. Lovebirds in the air: trade patterns, establishment success and niche shifts of *Agapornis* parrots within their non-native range. Biol. Invas. 22, 421–435.
Murray, G., Soares, A., Novak, B., Schaefer, N., Cahill, J., Baker, A., Demboski, J., Doll, A., Fonseca, R.D., Fulton, T., Gilbert, M., 2017. Natural selection shaped the rise and fall of passenger pigeon genomic diversity. Science 358, 951–954.
Olah, G., Smith, B.T., Joseph, L., Banks, S.C., Heinsohn, R., 2021. Advancing genetic methods in the study of parrot biology and conservation. Diversity 13, 521.
Olofsson, J.K., Cantera, I., Van de Paer, C., Hong-Wa, C., Zedane, L., Dunning, L.T., Alberti, A., Christin, P.A., Besnard, G., 2019. Phylogenomics using low-depth whole genome sequencing: a case study with the olive tribe. Mol. Ecol. Resour. 19, 877–892.
Portik, D.M., Smith, L.L., Bi, K., 2016. An evaluation of transcriptome-based exon capture for frog phylogenomics across multiple scales of divergence (Class: Amphibia, Order: Anura). Mol. Ecol. Resour. 16, 1069–1083.
Price, M.N., Dehal, P.S., Arkin, A.P., 2010. FastTree 2–approximately maximum-likelihood trees for large alignments. PLoS One 5, e9490.
Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. Genetics 155, 945–959.
Raxworthy, C.J., Smith, B.T., 2021. Mining museums for historical DNA: advances and challenges in museomics. Trends Ecol. Evol. 36, 1049–1060.
Schweizer, M., Seehausen, O., Güntert, M., Hertwig, S.T., 2010. The evolutionary diversification of parrots supports a taxon pulse model with multiple trans-oceanic dispersal events and local radiations. Mol. Phylogenet. Evol. 54, 984–994.
Schweizer, M., Seehausen, O., Hertwig, S.T., 2011. Macroevolutionary patterns in the diversification of parrots: effects of climate change, geological events and key innovations. J. Biogeogr. 38, 2176–2194.
Sin, S.Y.W., Lu, L., Edwards, S.V., 2020. De Novo assembly of the Northern Cardinal (*Cardinalis cardinalis*) genome reveals candidate regulatory regions for sexually dichromatic red plumage coloration. G3: Genes Genomes. Genetics 10, 3541–3548.
Sin, S.Y.W., Cloutier, A., Nevitt, G., Edwards, S.V., 2022. Olfactory receptor subgenome and expression in a highly olfactory procellariiform seabird. Genetics 220, iyab210.
Smith, B.T., Mauck III, W.M., Benz, B.W., Andersen, M.J., 2020. Uneven missing data skew phylogenomic relationships within the lories and lorikeets. Genome Biol. Evol. 12, 1131–1147.
Teeling, E.C., Vernes, S.C., Dávalos, L.M., Ray, D.A., Gilbert, M.T.P., Myers, E., Consortium, B.K., 2018. Bat biology, genomes, and the Bat1K project: to generate chromosome-level genomes for all living bat species. Ann. Rev. Anim. Biosci. 6, 23–46.
Trevisan, B., Alcantara, D.M., Machado, D.J., Marques, F.P., Lahr, D.J., 2019. Genome skimming is a low-cost and robust strategy to assemble complete mitochondrial genomes from ethanol preserved specimens in biodiversity studies. PeerJ 7, e7543.
Tsai, W.L., Schedl, M.E., Maley, J.M., McCormack, J.E., 2020. More than skin and bones: comparing extraction methods and alternative sources of DNA from avian museum specimens. Mole. Ecol. Resour. 20, 1220–1227.
Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., 2013. From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. Current protocols in bioinformatics 43, 11.10.1–11.10.33.
Van der Zwan, H., Visser, C., Van der Sluis, R., 2019. Plumage colour variations in the Agapornis genus: a review. Ostrich 90, 1–10.
West-Eberhard, M.J., 1983. Sexual selection, social competition, and speciation. Quart. Rev. Biol. 58, 155–183.
Wright, T.F., Schirtzinger, E.E., Matsumoto, T., Eberhard, J.R., Graves, G.R., Sanchez, J. J., Capelli, S., Müller, H., Scharpegge, J., Chambers, G.K., 2008. A multilocus molecular phylogeny of the parrots (Psittaciformes): support for a Gondwanan origin during the Cretaceous. Mol. Biol. Evol. 25, 2141–2156.
Yao, L., Li, H., Martin, R.D., Moreau, C.S., Malhi, R.S., 2017. Tracing the phylogeographic history of Southeast Asian long-tailed macaques through mitogenomes of museum specimens. Mol. Phylogenet. Evol. 116, 227–238.
Zhang, G., 2015. Bird sequencing project takes off. Nature 522, 34.