Molecular Ecology

WILEY

MOLECULAR ECOLOGY

ORIGINAL ARTICLE OPEN ACCESS

Museomics Sheds Light on Evolutionary Diversity in a Critically Endangered Cockatoo Species From Wallacea

Astrid A. Andersson[1] | Arthur F. Sands[1] | Kerry Reid[1] | Taylor Hains[2,3] | Paolo Momigliano[1,4,5] | Jessica G. H. Lee[6] | Geraldine Lee[7] | Frank E. Rheindt[7] | Juha Merilä[1,8] | Caroline Dingle[1,9]

[1]Area of Ecology and Biodiversity, School of Biological Sciences, University of Hong Kong, Hong Kong, People's Republic of China | [2]Committee on Evolutionary Biology, University of Chicago, Chicago, Illinois, USA | [3]Negaunee Integrative Research Center, The Field Museum, Chicago, Illinois, USA | [4]Faculty of Science, Swire Institute of Marine Science, The University of Hong Kong, Hong Kong, People's Republic of China | [5]Department of Ecological and Biological Sciences, University of Tuscia, Viterbo, Italy | [6]Mandai Nature, Mandai, Singapore | [7]Avian Evolution Lab, Department of Biological Sciences, National University of Singapore, Singapore | [8]Ecological Genetics Research Unit, Organismal and Evolutionary Biology Programme, University of Helsinki, Helsinki, Finland | [9]Biology Department, Capilano University, North Vancouver, British Columbia, Canada

Correspondence: Juha Merilä (merila@hku.hk) | Caroline Dingle (cdingle@hku.hk)

Received: 5 February 2024 | **Revised:** 10 October 2024 | **Accepted:** 26 November 2024

Handling Editor: Christine Edwards

Funding: The authors received no specific funding for this work.

Keywords: birds | *Cacatua* | conservation genetics | genetic differentiation | Indonesia | systematics

ABSTRACT

Accurate identification of evolutionarily significant units of rare and threatened organisms provides a foundation for effective management and conservation. Up to seven subspecies of the critically endangered Yellow-crested Cockatoo (*Cacatua sulphurea*) have been described, four of which were commonly recognised pre-2014. In the absence of genotypic data, *C. sulphurea* subspecies delimitation has been based on morphology, behaviour and biogeography. To clarify genetic relationships and shed light on the diversification of this parrot radiation, whole genomes were sequenced for 16 museum specimens, covering the geographic range of the proposed seven subspecies as well as one *C. galerita galerita*. Combined with four museum-derived wild *Cacatua* sequences from NCBI, the results indicate there are three distinct *C. sulphurea* subspecies clusters centred in different biogeographic subregions of Wallacea (Timor; Sumba; as well as the Sulawesi Region and the main Lesser Sunda chain), separated by shallow genetic distances ($d_a < 0.148\%$). The results raise questions about the recent species-level elevation of the phenotypically most distinct subspecies, *C. s. citrinocristata*, and about the origins of *C. s. abbotti*, the only subspecies west of Wallace's Line. Our analyses suggest *C. s. abbotti* is unlikely to be embedded within *C. sulphurea*, suggesting its origin on the remote Masalembu islands may be due to human translocation via historical trade routes. These genomic results inform the prioritisation and streamlining of conservation measures for the critically endangered *C. sulphurea* by identifying and delimiting likely conservation units.

1 | Introduction

Accurately identifying evolutionarily significant units is essential to informing conservation actions (Agnarsson and Kuntner 2007; Thomson et al. 2018; Sadanandan et al. 2020), particularly for species threatened by ongoing rapid anthropogenic and environmental pressures. Approaches integrating multiple characters, including biogeographic, phenotypic and

Juha Merilä and Caroline Dingle share last authorship.

genotypic data, are believed to produce the most reliable delimitation of taxonomic units and the understanding of underlying biological diversity (Wortley and Scotland 2006; Padial et al. 2010; Gwee et al. 2021; Chornelia, Lu, and Hughes 2022) and thereby guide meaningful conservation efforts.

Historic approaches of classifying our planet's biota have been based on morphological, behavioural and biogeographic information, with genetic data being integrated more recently as technologies have evolved. Traditional DNA-based techniques were generally restricted to few mitochondrial or nuclear genes, often producing distorted, incomplete or erroneous pictures of the past because of the well-documented biases associated with phylogenetic analyses based on few markers (e.g., Rheindt and Edwards 2011; Harris et al. 2018; Guo et al. 2019). Such analyses have often remained unable to capture the full complexity of differentiation at variable rates across the genome (Rheindt, Christidis, and Norman 2009; Harrison and Kidner 2011; Jarvis et al. 2014). Recent advances in whole genome sequencing have greatly improved our ability to infer evolutionary relationships among taxonomic units, as well as their delimitation. However, for threatened species it can be challenging to access fresh samples, and in this scenario museomics can help overcome this limitation (Fong et al. 2023).

Classification of evolutionarily significant units in the cockatoos (genus *Cacatua*) has relied heavily on biogeography, since many members of this non-migratory genus reside on separate islands across Australasia. Some studies have characterised cockatoo diversity based on morphology (Smith 1975), behaviour (Courtney 1996) and biochemistry (Adams et al. 1984), while molecular studies have aimed to reconstruct the cockatoo phylogeny based on variation in chromosome number (Christidis, Shaw, and Schodde 1991), mitochondrial DNA, as well as single- (Brown and Toft 1999) and multi-locus nuclear markers (White et al. 2011). However, studies based on entire genomes, which provide much higher resolution, are lacking. As higher-level phylogenies of parrots are becoming better resolved (Olah et al. 2021; Rheindt et al. 2014), there continues to be a dearth of research on intrageneric relationships, particularly in cockatoos (Provost, Joseph, and Smith 2018). Meanwhile, Southeast Asia—and Indonesia in particular—have been identified as regions of high priority for parrot conservation due to their elevated levels of species richness and endemism, and the severity of threats such as deforestation and trapping for trade (Olah et al. 2016).

The Wallacean archipelago is the home of the Yellow-crested Cockatoo (*C. sulphurea*). This species is classified as Critically Endangered by the International Union for the Conservation of Nature (IUCN; BirdLife International 2021a), with an estimated cumulative total of fewer than 2000 individuals (IUCN 2023), mainly due to exploitation for pet trade, which at its peak exceeded 96,700 individuals exported from Indonesia between 1981 and 1992 (Cahill, Walker, and Marsden 2006). Parrot trade has impacted the contemporary demographic structure of cockatoo populations elsewhere in Indonesia (Nandika et al. 2021) and is likely to have also fragmented and reduced the populations of various *C. sulphurea* subspecies at different levels of severity.

Historically, four to six *C. sulphurea* subspecies have been recognised depending on taxonomic authority (e.g., Forshaw 1989; Clements 2000; Dickinson and Remsen 2013): *C. s. sulphurea* (Gmelin, 1788), *C. s. parvula* (Bonaparte, 1850), *C. s. citrinocristata* (Fraser, 1844), *C. s. abbotti* (Oberholser, 1917), and sometimes (e.g., Eaton et al. 2021) also *C. s. occidentalis* Hartert, 1898, and *C. s. djampeana* Hartert, 1897 (Figure 1). A recent taxonomic review based on morphological characters proposed the existence of a seventh subspecies, *C. s. paulandrewi* (Collar and Marsden 2014; Figure 1), which has subsequently been accepted by some modern treatments (e.g., Eaton et al. 2021). One of these seven subspecies, *C. s. citrinocristata* from Sumba, differs substantially from the others phenotypically and has been proposed to warrant elevation to species status (Eaton et al. 2021; BirdLife International 2021b; Reuleaux et al. 2022). The notable differences in biological characters of *C. s. citrinocristata*, such as its distinctly orange rather than yellow crest and its unique juvenile beak colouration, have been mentioned repeatedly in the literature—as has the need for genetic studies to confirm its species status (Eaton et al. 2021; Collar and Marsden 2014).

One of the greatest biogeographic puzzles surrounding Yellow-crested Cockatoos involves the origin and taxonomic validity of *C. s. abbotti* from just beyond the western boundary of Wallacea—a subspecies with a notably larger body size than all other *C. sulphurea* subspecies (Collar and Marsden 2014). It is unknown whether the presence of this most isolated of *C. sulphurea* subspecies, found on the remote Masalembu islands (Figure 1), is due to human-mediated introduction or natural colonisation (Collar and Marsden 2014). The Masalembu Islands fall ~250 km west of Wallace's line, which delimits the westernmost extent of many Australasian-derived faunal components (Mayr 1944; Ali, Aitchison, and Meiri 2020). There are only a few cases of Australasian-centred radiations known to extend slightly beyond Wallacea; other examples include the Philippine cockatoo *Cacatua haematuropygia* on Palawan Island and *Lepidodactylus listeri* geckos on Christmas Island (Ali, Aitchison, and Meiri 2020). Although it is possible that cockatoos self-colonised Masalembu through overwater dispersal, it is more likely that they were brought there via historical trade routes, with trade in cockatoos from this part of the world recorded as far back as the 13th century (Dalton et al. 2018). Though parrot trade has had a considerable impact on cockatoos in Indonesia, the precise impacts vary (Pires et al. 2021)—some populations or subspecies may have been translocated, others over-harvested, fragmented and driven to extinction, and yet others may have become admixed.

The aim of this study was to use genomic data to further our understanding of the evolutionary trajectories of, and relationships among, the seven described subspecies of *C. sulphurea*. To this end, we conducted whole genome resequencing of museum specimens of *C. sulphurea* sampled across its geographical distribution, representing all recognised subspecies. We used this data to test whether 1) there may be fewer than seven genetically distinct groups within *C. sulphurea*, 2) if *C. s. citrinocristata* is sufficiently distinct to warrant species-level elevation and 3) to shed light on the origin of *C. s. abbotti*.

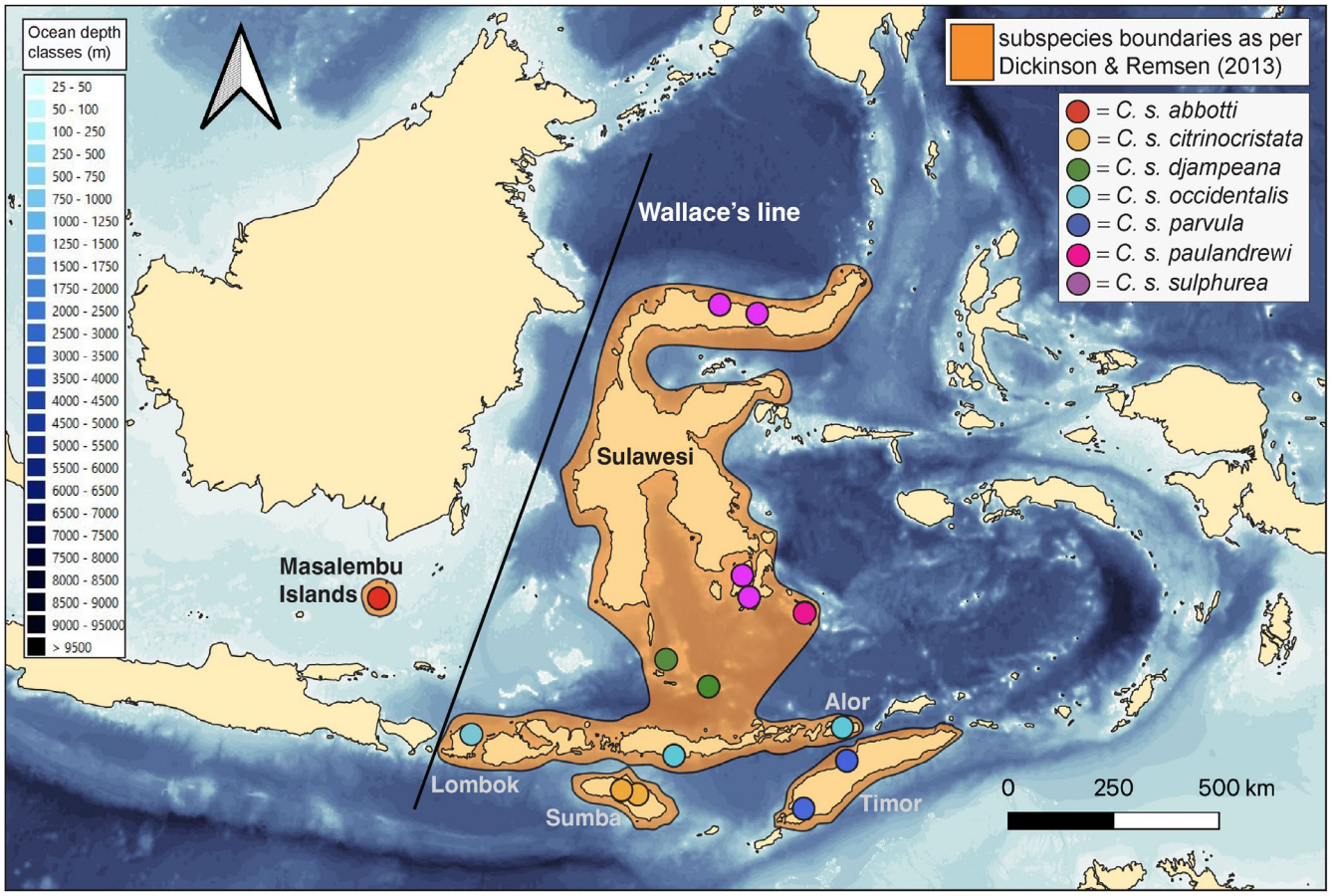

FIGURE 1 | Origins of *Cacatua sulphurea* museum samples used in this study coloured by their subspecies designations as per Collar and Marsden (2014). The most-commonly accepted subspecies treatment pre-2014 (e.g., Dickinson and Remsen 2013) is shown in orange shapes that encompass the four previously recognised subspecies. Bathymetric visualisation and depth legend derived from the General Bathymetric Chart of the Oceans 2021 Grid by OpenDEM. One additional *C. s. djampeana* sample incorporated in the study is not shown due to unconfirmed locality data. The *C. s. occidentalis* specimen from Lombok had insufficient coverage of nuclear DNA, and was only used in analyses for mitochondrial DNA.

2 | Material and Methods

2.1 | Sampling

Cacatua sulphurea is rare and difficult to access in the wild, to the point that collecting modern wild samples for all subspecies has been described as 'virtually or entirely impossible' (Collar and Marsden 2014) and would be ethically questionable from a conservation perspective. Therefore, this study used one to four historical museum specimens to represent each currently named taxon in the *C. sulphurea* complex (Figure 1). There are approximately 125 known individual specimens of *C. sulphurea* subspecies held in museum collections worldwide. Toepad samples from 15 native-range *C. sulphurea* specimens across Indonesia and two Sulphur-crested Cockatoos (*Cacatua galerita galerita*) were incorporated (one from the Australian National Wildlife Collection, Canberra, and the other downloaded from NCBI (Accession number: GCA_035583095.1); Table 1). Raw reads for three additional wild-sourced museum-derived genomes were downloaded from NCBI for the Little Corella (*Cacatua sanguinea*; GCA_030265285.1), Solomons Cockatoo (*Cacatua ducorpsii*; GCA_025448155.1) and Long-billed Corella (*Cacatua tenuirostris*; GCA_030265375.1) and incorporated into certain analyses (details below). For estimating heterozygosity loss associated with decreasing coverage, we used one *C. sulphurea* individual from a feral Hong Kong population sequenced to a higher coverage (Table 1).

2.2 | DNA Extraction and Library Preparation

Historical DNA is usually degraded and damaged, containing low volumes of target DNA in short fragments (Pääbo et al. 2004), but specific methods can be applied to mitigate these issues (e.g., Chattopadhyay et al. 2019; Ewart et al. 2019). Genomic DNA was extracted using a modified DNeasy Blood & Tissue Kit (QIAGEN, Hilden, Germany) protocol following the library preparation procedures of Chattopadhyay et al. (2019), see Appendix. This was done in a dedicated historical DNA facility at the National University of Singapore (NUS), strictly isolated from any post-PCR workspace (Poinar and Cooper 2000). These samples were sequenced by Macrogen (Seoul, South Korea) on an Illumina NovaSeq platform, with a target depth of between 40 and 60× coverage, using PE150 paired-end sequencing.

Three additional samples (RA7636; ZA9265; ZA9267) were extracted at the Institute for Animal Ecology & Systematics, Justus Liebig University Gießen, in a sterile environment

TABLE 1 | Museum samples used in the study, supplied by the American Museum of Natural History, New York (AMNH); Naturalis Biodiversity Center, Leiden (RMNH); Smithsonian National Museum of Natural History, Washington D.C. (USNM); Australian National Wildlife Collection, Canberra (ANWC); Kansas University Natural History Museum, Kansas (KU) and the Field Museum of Natural History, Chicago (FMNH). Samples from KU and FMNH were downloaded from NCBI.

Specimen ID	Museum source	Scientific name	Subspecies	Sex	Material	Latitude (Approx.)	Longitude (Approx.)	Locality (source)	Cov $\bar{x}$	SD	Missing data	Date
181455	USNM	*Cacatua sulphurea*	*C. s. abbotti*	M	toepad	−5.45239	114°42′723	Indonesia	3.5	4.0	0.0169	1907
619675	AMNH	*Cacatua sulphurea*	*C. s. citrinocristata*	M	toepad	−9.63469	119°84′318	Indonesia	10.7	10.8	0.0002	NA
346660	AMNH	*Cacatua sulphurea*	*C. s. citrinocristata*	F	toepad	−9.72716	120°20′172	Indonesia	4.8	6.1	0.0030	1932
266486	AMNH	*Cacatua sulphurea*	*C. s. djampeana*	M	toepad	−6.806588	120°47′42.8	Indonesia	3	2.8	0.319	1927
RA7636	RMNH	*Cacatua sulphurea*	*C. s. djampeana*	F	toepad	NA	NA	Indonesia	1.2	2.0	0.1173	NA
266489	AMNH	*Cacatua sulphurea*	*C. s. djampeana*	F	toepad	−7.36896	121°78′13	Indonesia	0.8	1.7	0.319	1927
154070	FMNH	*Cacatua galerita*	*C. g. galerita*	F	toepad	NA	NA	Australia	11.6	53.21	0.0011	1938
B38028	ANWC	*Cacatua galerita*	*C. g. galerita*	F	toepad	−31.8833	116°00′000	Australia	8.5	9.3	0.0010	1982
619666	AMNH	*Cacatua sulphurea*	*C. s. occidentalis*	M	toepad	−8.87965	121°01′25	Indonesia	8.8	7.7	0.00003	1896
619656	AMNH	*Cacatua sulphurea*	*C. s. occidentalis*	F	toepad	−8.26357	124°76′725	Indonesia	2.1	3.1	0.0440	1897
208266	RMNH	*Cacatua sulphurea*	*C. s. occidentalis*	F	toepad	−8.3583	116.5417	Indonesia	0.6	0.2	NA	1909
345474	AMNH	*Cacatua sulphurea*	*C. s. parvula*	M	toepad	−10.04036	123°90′409	Indonesia	9.3	9.2	0.0001	1932
619677	AMNH	*Cacatua sulphurea*	*C. s. parvula*	F	toepad	−8.9916	124°86′8	Indonesia	0.7	1.7	0.3459	1897
619646	AMNH	*Cacatua sulphurea*	*C. s. paulandrewi*	M	toepad	−5.75435	123°93′189	Indonesia	6.0	6.4	0.0008	1901
153742	AMNH	*Cacatua sulphurea*	*C. s. sulphurea*	M	toepad	0.78742	122°86′955	Indonesia	2.2	2.9	0.0405	1914
ZA9267	RMNH	*Cacatua sulphurea*	*C. s. sulphurea*	M	toepad	−5.175	122°84′17	Indonesia	1.0	1.7	0.1744	1948
ZA9265	RMNH	*Cacatua sulphurea*	*C. s. sulphurea*	F	toepad	−4.9917	122°65′83	Indonesia	0.9	1.7	0.2276	1948
298670	AMNH	*Cacatua sulphurea*	*C. s. sulphurea*	F	toepad	0.98104	122°03′054	Indonesia	7.9	7.2	0.0001	1930
131250	KU	*Cacatua ducorpsii*	NA	F	tissue	−7.03500	155°85′800	Solomon Islands	22.2	116.5	NA	NA
92461	KU	*Cacatua sanguinea*	*C. s. westralensis*	NA	tissue	−24.1200	115°26′667	Australia	14.3	82	0.0007	2002
98482	KU	*Cacatua tenuirostris*	NA	NA	tissue	NA	NA	Australia	1.9	7.1	NA	NA
Fresh sample												
436578	SPCA (HK)	*Cacatua sulphurea*	NA	F	blood	22.2771	114°16′403	Hong Kong	27.1	64.2	NA	2018

Note: Coverage and SD data apply to nuclear DNA. Missing SNP data is retrieved from the genotype likelihood file generated for the phylogenetic tree analysis (including *C. sanguinea* as an outgroup).
Abbreviations: Cov, coverage; F, female; M, male; SD, standard deviation; NA, not available.

where birds have never been processed. Samples were prepared using the same protocols, with library construction carried out by the Beijing Genome Institute (BGI) using the KAPA library prep kit and sequenced using proprietary DNBseq sequencing technologies. Library kits used in each location were tailored for degraded samples and had similar size and fragment specifications. Macrogen library prep was done in-house, so there was more control over the quality of library prep success, while library prep of BGI samples was carried out by BGI. The BGI returns were generally of lower coverage yield, although target coverage was the same for both companies.

2.3 | Whole Genome Sequence Processing

Since the historical samples were 50–150 years old (Table 1) and were likely degraded, we applied protocols specific for processing historical samples and eliminating damaged and low-quality genotype information (see Wu et al. 2020). First adapters were removed with AdapterRemoval v2.2.3 (Schubert, Lindgreen, and Orlando 2016) for each sequencing platform. The raw reads were screened for contamination using FastQ Screen Version 0.11.9 (Wingett and Andrews 2018), retaining only reads that mapped uniquely to cockatoo reference sequences, including the Palm Cockatoo (*Probosciger aterrimus*) GenBank assembly accession: GCA_013397665.1 (Feng et al. 2020) and others (see Appendix). After that mate pairs were repaired with BBTools repair.sh (sourceforge.net/projects/bbmap/). Alignment of raw reads was carried out using Burrows Wheeler Aligner (BWA mem; Li and Durbin 2010), mapping to the reference genome of the Palm Cockatoo (*Probosciger aterrimus*). Merging, sorting, and duplicate marking was conducted with Picard-Tools (broadinstitute.github.io/picard/), and indexing of BAM files were done with SAMtools. The latter tool also provided mapping success statistics (Li et al. 2009) including the average coverage across the genome for each individual sample. We then examined bam files for *post-mortem* damage and rescaled them accordingly using mapDamage2 (Jónsson et al. 2013). A final check was run using Qualimap 2 (Okonechnikov, Conesa, and García-Alcalde 2016), which suggested that trimming 15 base pairs off the ends of each read would be sufficient to remove degradation. We retained 15 *C. sulphurea* samples for downstream analyses, as well as two *C. g. galerita* and three wild-sourced *Cacatua* museum samples available on NCBI. *Cacatua sanguinea* was used as an outgroup to root the phylogenetic tree, as well as for genetic distance estimates along with *C. ducorpsii* and *C. tenuirostris*.

To maximise accuracy of our low-coverage whole genome sequences (Lou et al. 2021), we utilised genotype likelihood calling in ANGSD (Korneliussen, Albrechtsen, and Nielsen 2014) using the SAMtools genotype model (–GL 1), with parameters set to control for low-quality data, including: (1) only retaining reads that map uniquely to the reference genome (–uniqueOnly 1), (2) removing reads that are marked as duplicates or that have a low base quality score (–remove_bads 1), (3) removing 15 potentially degraded base pairs on either end of the read (–trim 15), (4) filtering out reads that are not properly paired (–only_proper_pairs 1), (5) removing false positives by only retaining SNPs with a p-value of less than 1×10^{-6} (–SNP_pval 1e-6), (6) removing reads with a mapping quality score below 30 (–minMapQ 30) and a base quality score below 30 (–minQ 30), (7) requiring a minor allele frequency of 0.01 to remove sequencing errors (–minMaf 0.01), (8) ensuring all SNPs are present at a read depth of at least 3 (–geno_minDepth 3), (9) requiring that at least three individuals have data at a specific site (–setMinDepthInd 3), and (10) requiring that a SNP is present in at least 90% of the individuals included. This produced a beagle file for two data subsets: (1) *C. sulphurea* plus *C. galerita*, and (2) *C. sulphurea, C. galerita* plus *C. sanguinea*.

After variant calling, filtering, and pruning for linkage disequilibrium (LD), the beagle file from ANGSD for the *C. sulphurea* plus *C. galerita* dataset contained 24,284 high-quality unlinked SNPs. Pruning for linkage disequilibrium (LD) was carried out using a correlation coefficient higher than 0.2 as measured in PLINK 1.90 (Chang et al. 2015), a window size of 50 and a step size of 5 (–indep-pairwise 50 5 0.2).

For phylogenomic analyses, the *C. sanguinea* sample (Table 1) was added as an outgroup, and an additional beagle file was produced for this dataset in ANGSD, using the same parameters as outlined above and pruned for LD by the same means, ultimately retaining 29,256 SNPs.

2.4 | Population-Genomic and Phylogenomic Analyses

2.4.1 | Subspecies Relationships

To assess population structure, a principal component analysis (PCA) was run for the *C. sulphurea* plus *C. galerita* dataset, using PLINK 2 (Chang et al. 2015). We also plotted missing data to determine if this had any influence on the PCA patterns (Yi and Latch 2022), which it did not (Figure S1). PCA plots were generated in ggplot2 (Villanueva and Chen 2019) using R statistical environment 3.6.2 (R Core Team 2021).

To further explore genetic relationships, admixture analyses were conducted to estimate individual admixture proportions among possible groupings (K). NGSadmix (Skotte, Korneliussen, and Albrechtsen 2013) was implemented on the *C. sulphurea* plus *C. galerita* genotype likelihood file for a set number of K ranging from 1 to 11. To compute the optimum K, we followed the method detailed in ngsAdmix_tutorial.md available at alexkrohn/AmargosaVoleTutorials. In brief, we ran NGSadmix 10 times per K value ($K=2$ through $K=11$), created a file of all the log likelihoods and formatted it for CLUMPAK (Cluster Markov Packager Across K; Kopelman et al. 2015). The file was then imported into CLUMPAK to estimate the best K, using methods developed by Evanno, Regnaut, and Goudet (2005). Scenarios $K=3$–7 were plotted using R as the most supported and realistic among assessments.

2.4.2 | Patterns of Divergence

Phylogenetic analyses were run to elucidate the evolutionary and divergence patterns among *C. sulphurea* subspecies using the maximum likelihood approach in IQ-TREE 2 (Minh et al. 2020).

For this we used the genotype likelihood file containing all individuals from *C. sulphurea*, *C. galerita* and *C. sanguinea* to create a consensus fasta with a minimum of four samples per SNP using the python script vcf2phylip.py (Ortiz 2019), designating the *C. sanguinea* sample as the outgroup. We then generated a maximum likelihood consensus tree using IQ-TREE 2 with 10,000 bootstrap replicates (–v –m GTR + ASC –bb 10,000 –bnni –st DNA –nt AUTO), applying the TVM + F + Γ4 best-fit model of sequence evolution as determined during the run. The tree was visualised in SplitsTree Version 3 (Huson and Bryant 2006). Additionally, a bifurcating tree with bootstrap support labels on nodes was generated using FigTree v1.4.4 (Rambaut 2018), see Figure S2.

2.4.3 | Genetic Distances

To explore the level of divergence among putative subspecies compared to divergences typically seen among different species, subspecies and populations, pairwise net nucleotide differences (d_a) were estimated and compared to pairwise comparisons of over 40 other species and subspecies combinations (Table S1), retrieved from Roux et al. (2016). We first generated a genotype likelihood file using ANGSD including only the highest-coverage individual from each of the seven putative *C. sulphurea* subspecies. Adding to the standard flags used to account for museum sample degradation mentioned above, we applied –setMaxDepth 40 –skipTriallelic 1 and adjusted the –minInd to 5 –setMinDepthInd to 2, which produced a beagle file with 233,425,555 common sites that were converted into a position file and added to the subsequent ANGSD run to produce a 1D Site Frequency Spectrum (SFS) file for each putative subspecies, as well as for *C. g. galerita, C. sanguinea, C. ducorpsii and C. tenuirostris*. For this, we used the -doSaf 1 and realSFS ANGSD functions, included the Palm Cockatoo genome mentioned above as the ancestral reference, and provided a contig map and position file. This SFS output was run through an R script to calculate nucleotide diversity (π) and heterozygosity for each individual subspecies, available at Nopaoli/Demographic-Modelling/tree/master/Diversity_fromSFS on Github. The site allele frequency output was then used to compute pairwise SFS between the seven putative subspecies using winsfs (Rasmussen et al. 2022), a tool to infer SFS for low-depth samples. The 2D SFS outputs from these were run through an R script from Momigliano et al. (Momigliano, Florin, and Merilä 2021; also available from the aforementioned GitHub page) to calculate absolute sequence divergence (d_{xy}). Finally, the net nucleotide difference (d_a) between pairs of subspecies was estimated using the formula (Nei and Li 1979):

$$d_a = \frac{d_{xy} - (\pi_x + \pi_y)}{2}$$

The pairwise d_a data points for *C. sulphurea* subspecies were plotted together with pairwise comparisons of different populations, subspecies and species of animals retrieved from Roux et al. (2016). The figure was generated using ggplot2 in R.

2.4.4 | Demographic History

To examine demographic patterns over time in the different subspecies groups, we used the Pairwise Sequentially Markovian Coalescent (PSMC; Li and Durbin 2011) method which examines the density of diverse sites across the genome to infer temporal changes in effective population size (N_e) since it is inversely proportional to coalescent rate. While PSMC typically requires higher-coverage genomes around 18× (Nadachowska-Brzyska et al. 2016) to ensure that all heterozygous sites are appropriately captured, generating genomes at that level of coverage can be a challenge for rare, extinct or non-model species (e.g., Von Seth et al. 2021) and there are means by which to account for heterozygosity loss in low-coverage samples (e.g., Palkopoulou et al. 2015). In our case, we used only *C. sulphurea* samples with > 8× coverage ($n = 3$). We removed pseudo-sex chromosomes and downsampled one 20× *C. sulphurea* individual from the feral Hong Kong population to measure heterozygosity loss at each level of coverage down to 5×. This practice allowed us to input the relevant heterozygosity loss metric for each individual in the PSMC as a false negative rate at the plotting stage, scaling the curve accordingly (Li and Durbin 2011); see Appendix for details.

We called consensus sequences with bcftools mpileup -C 50 -q 30 -Q 20 -Ou and called variants with bcftools call version 1.9 (Li et al. 2009), only including regions with a minimum depth of coverage representing a third of the average depth (-d) and a maximum depth double the average (-D) for each sample. Then we converted the resulting Variant Call Format (VCF) files to Fastq files using the 'vcfutils.pl. vcf2fq' script. We then ran PSMC with settings optimised for birds: -N30 -t5 -r5 -p '4 + 30*2 + 4 + 6 + 10' (Nadachowska-Brzyska et al. 2016). For plotting, we used a mutation rate per generation (μ) of 1.05e-8, which is halfway between the mutation rates of the Blue-throated Macaw (*Ara glaucogularis*; Bergeron et al. 2023) and Kea (*Nestor notabilis*; Martini et al. 2021). The generation time used was 14.3 for *C. sulphurea* (BirdLife International 2021a), and the heterozygosity loss false negative rate metrics were 0.318 for the 8× coverage sample, 0.278 for the 9× coverage sample, and 0.208 for the 10× coverage sample, all of which were factored in during the plotting stage. All PSMC plots were constructed with ggplot2 using a custom R script (Eliason et al. 2022; github.com/thainsCEB/birbscripts).

2.4.5 | Mitogenomic Analyses

To generate bam files for the mitogenomes of each sample, we followed the same pipeline as outlined above for nuclear DNA, except that we mapped to an unpublished reference *C. galerita* mitogenome generated using MitoZ (Meng et al. 2019) from specimen B34682 of the Australian National Wildlife Collection, Canberra. We then created two fasta files per sample using Angsd –doFasta 2 –doCounts 1 and –trim 2 commands. Here, one fasta file was generated with lenient quality requirements (–minQ 10) and the other with stringent quality requirements (–minQ 35 –minMapQ 50) for each sample. We used the Clustal Omega alignment tool in Geneious Prime (geneious.com) to compare the stringent and lenient setting fasta files for each sample, and by referring to these side by side we were able to fill in any gaps that may be present in the stringent fasta file sequences with information from lenient fasta file sequence—since the latter may be more complete, although with less certainty around calls. After validation, we retained the fasta sequence with stringent settings for each individual. Individual gene fastas for COI, ND2 and CytB were then extracted and exported using Geneious Prime, and

haplotype networks were generated in R using a custom script (github.com/AstridAlexAndersson/CacSul).

3 | Results

3.1 | Sequencing

Average depth of sequence coverage for the museum samples ranged from 0.6 to 14.3×. The *C. s. occidentalis* specimen from Lombok with 0.6× coverage was omitted from nuclear DNA analyses, but included in mitochondrial DNA analyses. Genotyping/missingness rates ranged from 0.00003 to 0.3459 (generated with Plink –missing, and calculated based on missing positions in the *C. sulphurea, C. galerita* and *C. sanguinea* beagle file).

3.2 | Subspecies Relationships

The PCA analyses revealed four distinct clusters: (1) *C. s. citrinocristata*, (2) *C. s. parvula*, (3) *C. s. abbotti* with *C. g. galerita* on PCs 1–3, and (4) one tight cluster including all four remaining putative *C. s. sulphurea* subspecies (*C. s. djampeana*, *C. s. paulandrewi*, *C. s. sulphurea* and to a lesser extent *C. s. occidentalis*), particularly along PCs 2–4, which cumulatively accounted for ~37% of variation (Figure 2A,B). The distance among the four putative subspecies grouped together in a tight cluster in group (4) was less than the distance between the two samples of *C. s. parvula*, which separated out very clearly from this cluster and from the equidistant *C. s. citrinocristata*. These same groupings were also apparent in the haplotype network for the mitochondrial DNA gene CytB (Figure 2D) as well as COI and ND2 (Figure S3). Among the *C. sulphurea* subspecies, *C. s. abbotti* emerged as the most distinct unit in all PCA analyses, grouping more closely with the two *C. g. galertia* samples than with its fellow putative *C. sulphurea* subspecies. These patterns were also echoed by the net nucleotide differences (d_a)—with the lowest pairwise differences (< 0.00049) between *C. s. djampeana*, *C. s. paulandrewi*, *C. s. sulphurea* and *C. s. occidentalis*, and between *C. s. abbotti* and *C. g. galerita* (0.00048), see Figure 4 and Table S1. *C. s. citrinocristata* and *C. s. parvula* exhibited a slightly higher pairwise d_a when compared with the other *C. sulphurea* subspecies (0.00095–0.00125).

3.3 | Patterns of Divergence

The maximum likelihood consensus tree revealed that the *C. s. abbotti* sample from Masalembu, west of Wallace's Line, emerged outside the clade comprising all other *C. sulphurea* subspecies in a position which suggests that it may be genetically closer to *C. g. galerita* than to *C. sulphurea* (Figure 2C). The tree shows *C. s. parvula* and *C. s. sulphurea* as the most distant among the populations living east of Wallace's Line, with *C. s. djampeana* being nested within *C. s. sulphurea*, rendering *C. sulphurea* paraphyletic. These findings follow patterns seen in PCA, and d_a estimates (Figure 4); for example, pairwise d_a between *C. s. parvula* and *C. s. sulphurea* is the largest among *C. sulphurea* subspecies (excluding *C. s. abbotti*).

The optimum number of population clusters (K) in the admixture analysis was $K=4$ ($\Delta K=9.061$), followed by $K=5$ ($\Delta K=2.344$; Figure S4). All K scenarios depicted in Figure 3 reveal at least some overlap in inferred ancestry between nominate *C. s. sulphurea* and the putative subspecies in its closest geographic proximity, namely *C. s. djampeana* and *C. s. paulandrewi*, with the latter also showing shared ancestry with its southern neighbour *C. s. occidentalis* in $K=3$, $K=4$ and $K=5$. Admixture analyses also indicated shared ancestry between *C. s. abbotti* and *C. g. galerita*, particularly among $K=3$–7.

3.4 | Genetic Distances

Net nucleotide differences (d_a) in all pairwise combinations involving the seven putative *C. s. sulphurea* subspecies—including *C. s. citrinocristata*—were generally low when contrasted with subspecies comparisons between other animal taxon pairs, and when compared with congeners *C. sanguinea*, *C. ducorpsii* and *C. tenuirostris* (Figure 4; Roux et al. 2016). All *C. sulphurea* subspecies divergence rates fell below 0.5% (0.005 on the plot), meaning they can plausibly be considered to be outside 'the grey zone of speciation' (i.e., the zone where species delimitation based on molecular data is problematic, or where there exist at least partial barriers to gene flow, Roux et al. 2016) and more concordant with divergences among populations within the same species which lack intrinsic barriers to gene flow. Pairwise comparison between *C. g. galerita* and *C. s. sulphurea* was also below this speciation 'grey zone' (Roux et al. 2016). *C. g. galerita* shared the lowest d_a with *C. s. abbotti*. Within the main Wallacean *C. sulphurea* cluster, the lowest pairwise d_a was between *C. s. djampeana* and *C. s. sulphurea* (−0.000085; Table S1). Having a negative d_a is possible and is an artefact of the d_a calculation process—for more information, see Appendix. For nucleotide diversity (π) and heterozygosity data for each individual subspecies, see Table S2.

3.5 | Demographic History

The PSMC plot indicated a parallel increase in N_e in three *C. sulphurea* subspecies throughout most of the last major period of global cooling, up until the last glacial maximum at ~20,000 years ago (Figure 5).

4 | Discussion

As global biodiversity faces unprecedented threats, accurate identification of evolutionarily significant units becomes increasingly important for directing conservation efforts. These issues are particularly acute for rare species where access to live specimens can be prohibitively difficult. Here we obtained genomic data from museum samples to provide a robust dataset to help address questions about the evolutionary history, diversification and conservation status of the *C. sulphurea* radiation. Research delimiting species boundaries requires thoroughly comprehensive sampling—an obstacle for extremely rare non-model species. The availability of museum specimens made this study possible, but also limited the inferences that can be made due to low sample size. Nevertheless, the genomic data

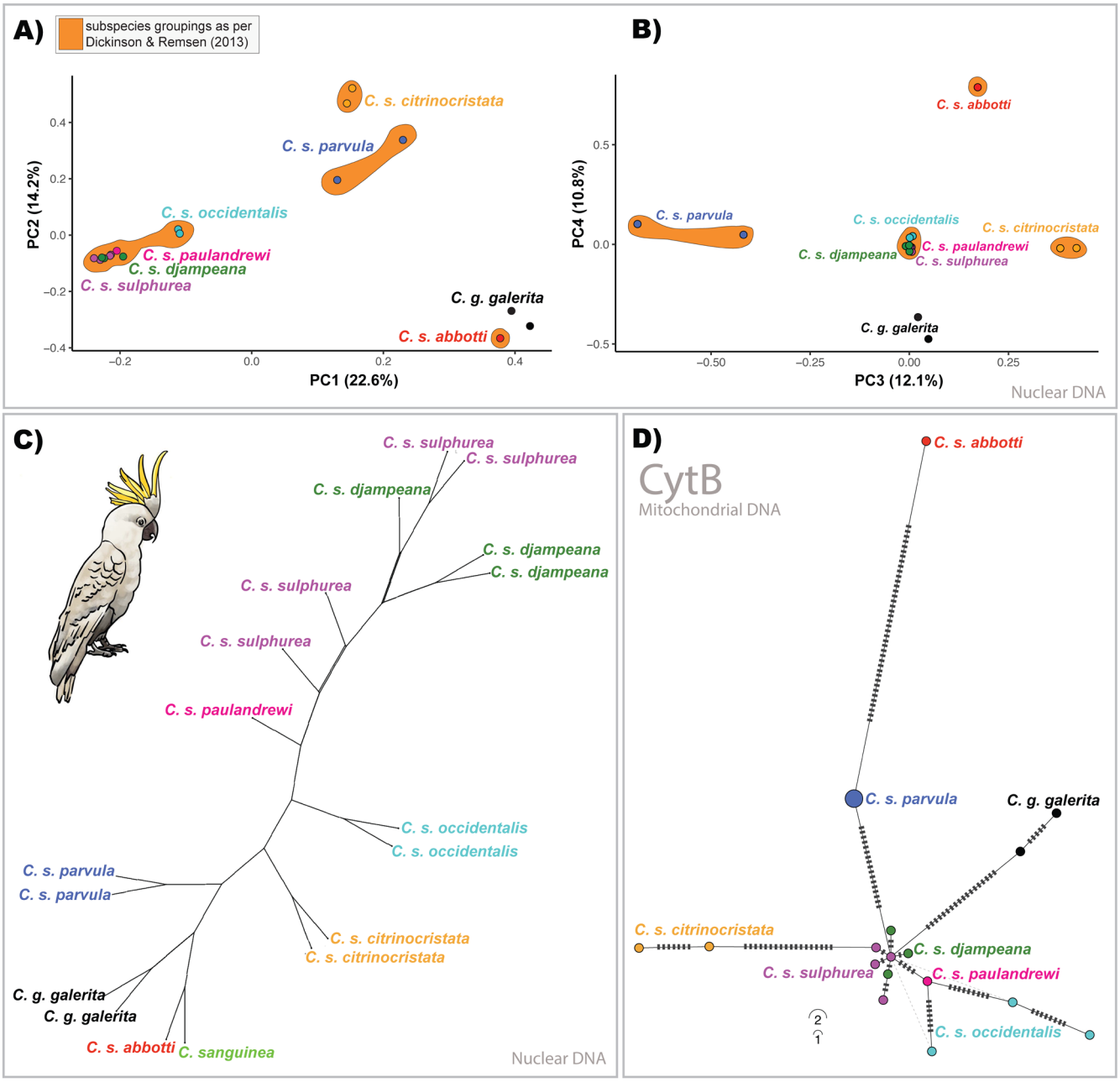

FIGURE 2 | (A) Results of principal component analysis showing the difference among seven putative *Cacatua sulphurea* subspecies and *Cacatua galerita galerita*. The percentage of variance accounted for by each principal component (PC) is stated in brackets along the plot axis. Coloured circles indicate samples that belong to the same putative subspecies as per Collar and Marsden (2014), while the most-commonly accepted subspecies treatment pre-2014 is shown in orange shapes encompassing samples from the four previously recognised subspecies. PC1 and PC2 for all seven putative subspecies of *C. sulphurea* and two *C. g. galerita*. (B) PC3 and PC4 for all seven putative subspecies of *C. sulphurea* and two *C. g. galerita*. (C) Maximum likelihood consensus tree generated with IQ-TREE 2 from 29,256 SNPs showing all individuals and their subspecies designation (indicated by font colour), two *C. g. galerita* samples and one *Cacatua sanguinea* sample as the outgroup. Unrooted tree was generated with SplitsTree. (D) Haplotype network for mitochondrial gene CytB for museum specimens of all *C. sulphurea* subspecies and two *C. g. galerita* samples, generated in R.

presented here does provide a robust interpretation of evolutionarily significant units across this species, particularly in combination with insights provided by prior biogeographical and morphological analyses.

The data presented here indicates that there are three distinct evolutionary lineages within the *C. sulphurea* complex in Wallacea: *C. s. parvula* from Timor, *C. s. citrinocristata* from Sumba, and the main cluster of subspecies from Sulawesi and most of the Lesser Sunda Islands (Figures 2 and 3, Figure S4). This latter cluster consists of the nominate subspecies *C. s. sulphurea* from Sulawesi in addition to three putative subspecies on relatively nearby islands—*C. s. djampeana, C. s. paulandrewi*, and the somewhat more distant *C. s. occidentalis*. In stark contrast to the Wallacean subspecies, *C. s. abbotti* from the Masalembu Islands emerged as genetically distinct

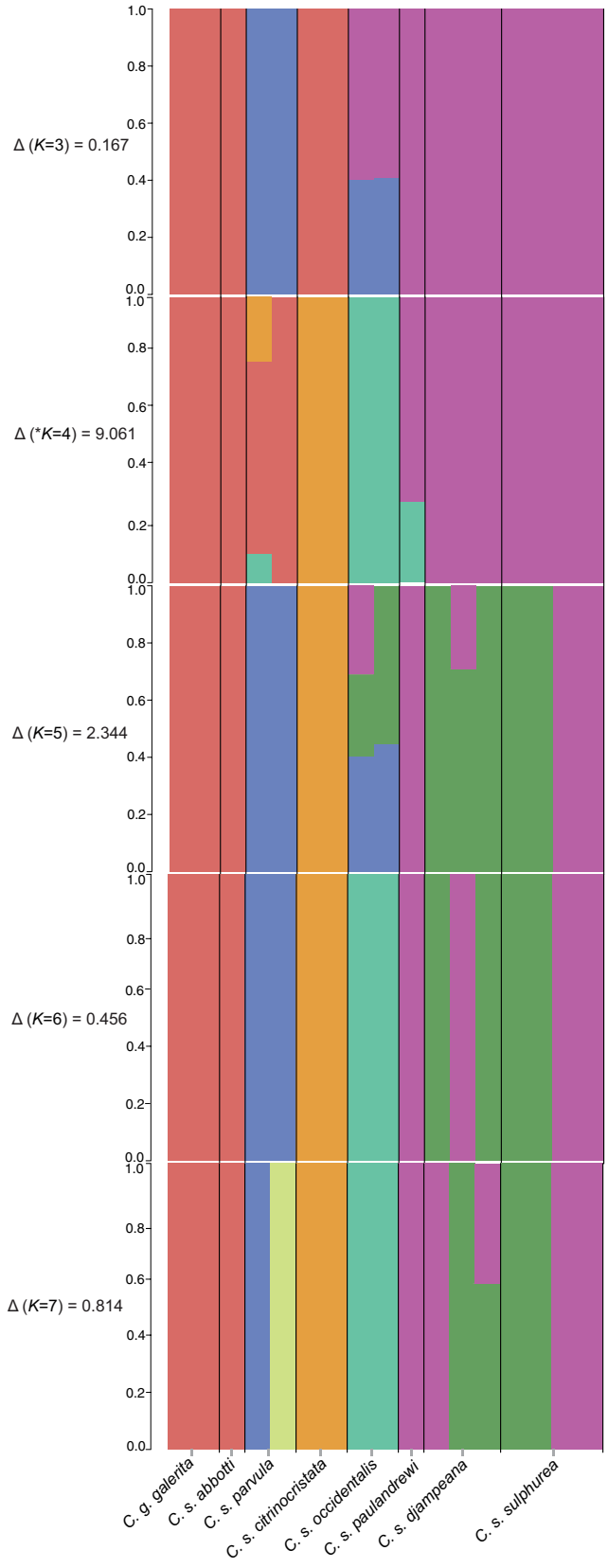

FIGURE 3 | Admixture plot showing groupings under $K = 3$ through $K = 7$ scenarios with individuals representing all seven putative subspecies of *Cacatua sulphurea* plus *Cacatua galerita galerita*. Delta values are shown alongside each K plot and the most-likely scenario ($K = 4$) is indicated with an asterisk.

from all other *C. sulphurea* subspecies, grouping closer to *C. g. galerita* and sharing the lowest genetic distance with *C. galerita* compared to other *C. sulphurea* subspecies. Increased taxon sampling of all *Cacatua* species in the region in the future will be needed to further corroborate that *C. s. abbotti* is not embedded within *C. sulphurea*, determine its closest relatives, and test whether this population originates from human introduction.

The three major *C. sulphurea* subspecies clusters separate neatly into the three principal paleo-island groups within the Lesser Sunda Islands (i.e., Nusa Tenggara) in southeast Indonesia: (1) Timor, the largest Lesser Sunda Island, inhabited by *C. s. parvula*; (2) the isolated outlier Sumba, inhabited by the phenotypically distinct *C. s. citrinocristata*; (3) and the main island chain of the Lesser Sundas ranging from Lombok in the west to Alor in the east, inhabited by *C. s. occidentalis*, which forms part of the nominate *C. s. sulphurea* cluster including *C. s. djampeana, C. s. paulandrewi* and *C. s. occidentalis*. The Lesser Sunda Islands that comprise the range of *C. s. occidentalis*, is today composed of ~5 larger and numerous smaller islands, lower sea levels during the last period of global cooling would have exposed land bridges connecting most of those islands into a group of two major paleo-islands separated by a small strait of ~2–3 km width (Lambeck, Yokoyama, and Purcell 2002; Lambeck et al. 2014). In contrast, expanses of dozens of kilometres of deep sea have separated the ranges of the nominate *C. s. sulphurea* cluster from *C. s. parvula* and *C. s. citrinocristata* since the formation of these islands, and likely acted as a catalyst for their differentiation.

Our findings are in keeping with major biogeographic divisions within Wallacea, and support a model by which the ancestor of *C. sulphurea* is likely to have been of eastern provenance, potentially colonising Wallacea from Australia—a scenario that would be consistent with cockatoos' inferred origin in Australia and Papua (Schweizer et al. 2010; White et al. 2011; Selvatti et al. 2022). It is possible that Timor's *C. s. parvula* may have first been colonised over water from an ancestral population in Australia, whose current overwater distance from Timor (~500 km) was reduced to < 150 km during multiple glacial peaks of the past (Siddall et al. 2003). Other interpretations are also possible, but the data presented here in multiple analyses seems to indicate a degree of shared ancestry between *C. s. parvula* and the Australasian *C. g. galertia*. The lack of deep genomic differentiation between the three subspecies from Sulawesi and satellites (*C. s. sulphurea*, *C. s. paulandrewi*, *C. s. djampeana*) versus *C. s. occidentalis* from the Lesser Sundas suggests a more recent overwater colonisation of the Sulawesi region from the south, likely via stepping stone islands in the Flores Sea nowadays inhabited by *C. s. djampeana*.

For three subspecies (*C. s. parvula*, *C. s. citrinocristata*, *C. s. occidentalis*), we had genomic data of sufficient coverage to trace the trajectory of effective population size across the last couple of periods of global cooling. Our results reveal a parallel trajectory for all three subspecies despite their different geographic ranges. Different factors can influence N_e such as introgression and population subdivision (Nadachowska-Brzyska, Konczal, and Babik 2022). In our case, effective population sizes generally underwent a steady increase from the last interglacial into the last period of global cooling (roughly

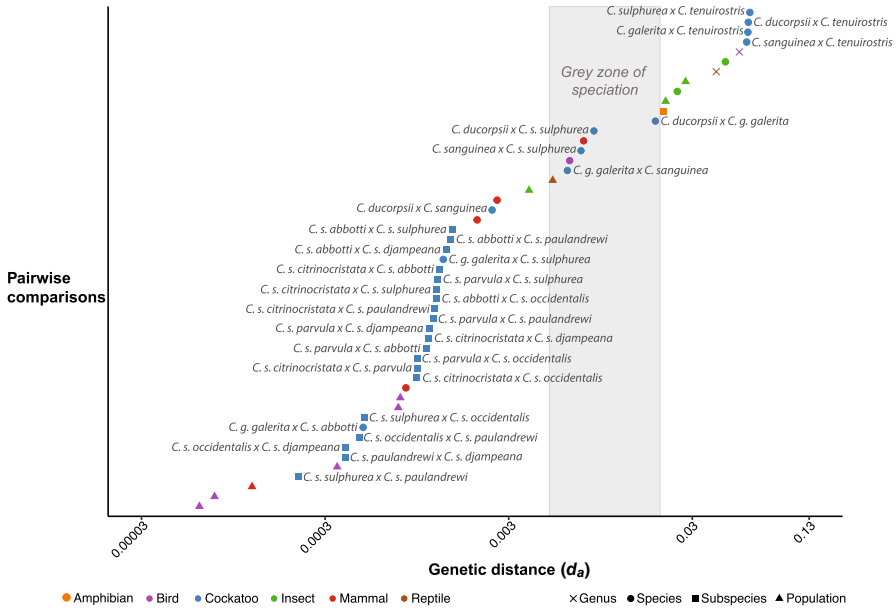

FIGURE 4 | The net genetic distance or net nucleotide difference (d_a) in all possible pairings of the seven putative *Cacatua sulphurea* subspecies plotted (log10 scale) alongside pairwise d_a data from *Cacatua galerita galerita*, *Cacatua sanguinea*, *Cacatua ducorpsii* and *Cacatua tenuirostris*, as well as populations, subspecies and species retrieved from Roux et al. (2016), with the 'grey zone of speciation' at 0.005–0.02 (or 0.5%–2%) demarcated. For additional data see Table S1. Pairwise comparison of *C. s. sulphurea* and *C. s. djampeana* not plotted because it was negative (see Appendix for more information).

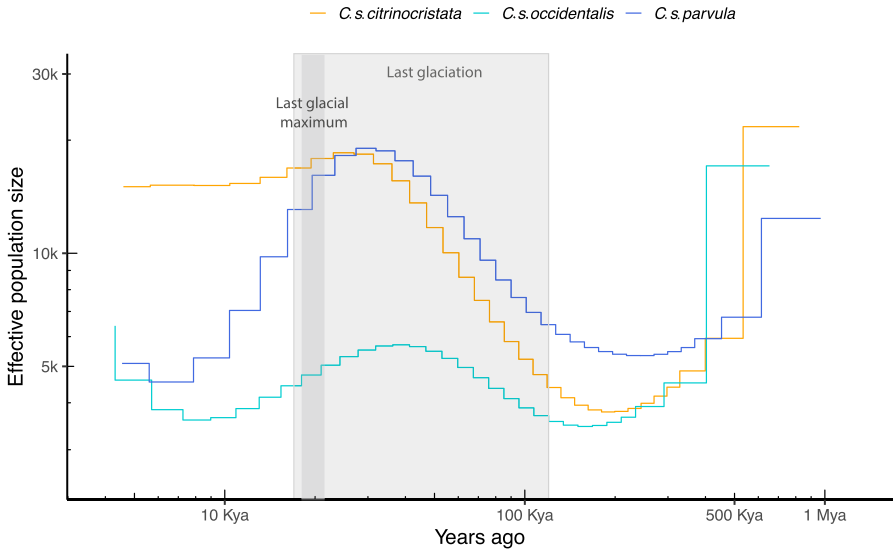

FIGURE 5 | Pairwise Sequentially Markovian Coalescent (PSMC) plot (log10 scale) showing the demographic histories of three *Cacatua sulphurea* subspecies. All plots are based on autosomal positions only and with false negative rate corrections to account for any heterozygosity lost due to low coverage.

from 100 to 30kya), reaching their peaks slightly before the last glacial maximum (~20kya). *Cacatua sulphurea parvula* from Timor and *C. s. citrinocristata* from Sumba exhibited the steepest fluctuations, including the steepest gains during global cooling, consistent with their residence on the most seasonal and most arid islands in the archipelago, on which survival conditions and habitat parameters would dramatically improve for a frugivore such as *C. sulphurea* during colder global periods. In *C. s. occidentalis* from less arid islands, these fluctuations were much less pronounced. Many of the islands which encompass its distribution had been connected during periods of global cooling, supporting higher population densities, while the two subspecies on Timor and Sumba essentially constitute single-island taxa living on more peripheral land masses. Unfortunately, our sample sizes are not applicable to methods allowing estimation of demographic trends since the last glacial maximum.

All putative subspecies of *C. s. sulphurea* exhibited pairwise genetic distances (d_a) across genome-wide SNPs much below

0.5%—placing all *C. s. sulphurea* subspecies below the 'grey zone' of the speciation continuum delineated by Roux et al. (2016). The d_a among *C. s. sulphurea*. *C. s. djampeana, C. s. paulandrewi* and *C. s. occidentalis* (−0.0008–0.00049) are comparable to the within-subspecies, population-level d_a for other species including Common Teal (*Anas crecca nimia*) on the Aleutian Islands (between 0.00006–0.00035) and the d_a of 0.0007 between populations of Eurasian Blue tit (*Cyanistes caeruleus*) across France (Roux et al. 2016). Changes in sea levels would have offered more opportunity for these taxa to exchange genetic information, which would also contribute to their lower levels of genetic divergence compared to other subspecies, and the sharing of (or similarity between) their mitochondrial haplotypes which parallels the nuclear DNA patterning in identifying three main genetic groups within the core *C. sulphurea* radiation (Figure S3). Notably, the d_a between *C. g. galerita* and *C. s. abbotti* is among the lowest in our analyses, lower than the pairwise d_a between *C. s. abbotti* and the other putative *C. s. sulphurea* subspecies (Table S1), and the distinctiveness of *C. s. abbotti* is also supported by the mitochondrial haplotypes (Figure S3). The genetic distance between *C. g. galerita* and *C. s. sulphurea* is also below the 'grey zone of speciation' (Roux et al. 2016)—suggesting more research into the relationships among taxa comprising the *C. galerita* and *C. sulphurea* species complex is warranted. In contrast, genetic distances between *C. sulphurea/C. galerita* and congeners *C. sanguinea, C. ducorpsii* and *C. tenuirostris* fall within or above this 'grey zone', which is to be expected given that these are comparisons between different subgenera of white cockatoos and corellas, recognised as Cacatua and Licmetis, respectively.

Subspecies within the *C. sulphurea* complex have been described often on the basis of subtle morphological characters. Considering the fitful nature of evolution, and the difficulties in accurately identifying evolutionarily significant units, it is possible that some bird species may have experienced taxonomic inflation (Holt and Jønsson 2014)—being over-assigned units based on phenotypic traits that, though visually present, account for shallow genomic divergence (e.g., Safran et al. 2016; Semenov et al. 2018). Subspecies is the most pliant category in the Linnaean hierarchical system, and valid subspecies may range anywhere from populations defined on the basis of minute character differences to highly distinct units that are close to the species boundary (Remsen 2010). Our genomic data, particularly the patterns revealed on the PCA, admixture analyses, phylogenetic tree and genetic distance plots, suggest that some of the subspecies in *C. sulphurea*, especially those on Sulawesi satellites, lack substantial genomic differentiation and may not warrant subspecies rank recognition, but in some cases their synonymisation should await confirmatory evidence from morphological datasets. At a minimum, we conclude that *C. s. djampeana* and *C. s. paulandrewi* from small islands geographically peripheral to Sulawesi consistently emerge as embedded within *C. s. sulphurea*—the subspecies from Sulawesi. Indeed these smaller islands in the Flores Sea and Banda Sea region are likely to have been stepping stones from which the species recently invaded, or spread out from, the much larger landmass of Sulawesi. To preclude unnecessary taxonomic and conservation focus on undifferentiated lineages, it appears safe to propose the synonymisation of *C. s. djampeana* and *C. s. paulandrewi* under *C. s. sulphurea* since they are weakly differentiated, and retaining them as separate subspecies likely constitutes taxonomic inflation. Moreover, synonymisation would enable these populations to function as a potential source of individuals for reintroduction and rescue for the near-extinct nominate Sulawesi population (IUCN 2023; Reuleaux et al. 2022).

The distinct morphological characters of *C. s. citrinocristata* (i.e., orange versus yellow crest, wing and tail length, darker juvenile beak coloration) have prompted its classification as a separate species by some scholars and the IUCN (BirdLife International 2021b; Reuleaux et al. 2022). However, genetically it is embedded within the subspecies radiation of *C. sulphurea*. In the PCA, *C. s. citrinocristata* is as separate from the other *C. sulphurea* subspecies as *C. s. parvula* is, while in terms of pairwise net nucleotide differences (d_a), *C. s. citrinocristata* shows a lower or equal divergence than that between the other putative subspecies. Moreover, if *C. s. citrinocristata* is to be maintained at species level (as per IOC v14.1), based on the presented phylogenetic analysis, this would necessitate that *C. s. parvula* be elevated to species level in order to avoid a paraphyletic *C. sulphurea*. However, this is ill-advised given the low levels of divergence between *C. s. citrinocristata*, *C. s. parvula* and the remaining *C. s. sulphurea* subspecies, which fall well below the threshold typically seen in species-level divergences (Figure 4). All this supports the view that *C. s. citrinocristata* deserves recognition as a highly distinct subspecies embedded within *C. sulphurea*.

While we deployed complete taxon sampling within the *C. sulphurea* radiation, our sampling of closely related cockatoo species was sparse. This sampling design limits our ability to assess the evolutionary position of the mysterious taxon *C. s. abbotti*, which appears to be genetically distinct from all the other *sulphurea* subspecies, and seems likely to be more closely related to *C. galerita* based on the PCA, phylogenetic tree, admixture analyses and genetic distance plots. If *C. s. abbotti* is an odd natural population of the remote Masalembu islands, on the 'wrong' side of Wallace's Line, it would have been there for a long period of time, likely going back much further than the history of human colonisation of southeast Indonesia. Isolation and genetic drift may explain its genetic differentiation from all the other *C. sulphurea* subspecies, but its unusual divergence—grouping closer to *C. g. galerita*—would necessitate a time of divergence on the order of hundreds of thousands of years. Future studies should include a denser sampling of the *Cacatua* genus to investigate whether *C. s. abbotti* may in fact be a population introduced by humans travelling through the Wallacean region for trade (Dalton et al. 2018).

The genotypic data presented here for *C. sulphurea* can help direct the prioritisation and practical application of conservation for this species. While the overall IUCN Red List status 'Critically Endangered' remains intact, our results suggest that initiatives to safeguard this species can target three distinct *C. sulphurea* subspecies clusters rather than upholding the barely divergent island populations like *C. s. paulandrewi* and *C. s. djampeana* as distinct subspecies and therefore dividing limited resources and coordinating logistically demanding efforts to manage seven distinct subspecies as separate units. In addition, our results suggest that *C. s. abbotti* may need to be removed

from any discussions and efforts directed to manage *C. sulphurea* populations. This Masalembu Islands population could instead be the recipient of dedicated and separate conservation action, although its possible status as an introduced, non-native population should also be considered.

Identifying evolutionarily distinct entities with whole genome analyses is particularly pertinent to heavily traded and threatened species such as *C. sulphurea*, since it can inform conservation efforts including translocation, genetic rescue and conservation breeding (Collar et al. 2012; Collar and Butchart 2014) to a much more comprehensive degree than single-gene data. For example, genetic reference points from this study can be applied to develop toolkits for subspecies identification by frontline conservationists and authorities (e.g., Cardeñosa et al. 2018). Hence it is particularly important from a conservation perspective to examine whole genome data for taxa that are distributed across multiple islands and are targets for illegal trade, and which may therefore be frequently seized, rehabilitated and released (e.g., Zein et al. 2017). In this scenario, single-gene data is not adequate for accurate repatriation to subspecies level—and releasing individuals that originated from another island on an entirely different island could lead to admixture and unnatural dilution of the subspecies.

Accurate native-range repatriation is crucial for a long-lived highly traded species with low fecundity, such as cockatoos, that are intercepted by customs and also potentially for *C. sulphurea* individuals present in *ex-situ* populations in Hong Kong, Singapore and Taiwan (Lin and Lee 2006; Neo 2012; Andersson et al. 2021). The whole genome data presented here makes it possible to assess to what extent these introduced populations are functioning as genetic reservoirs or admixture hot spots for this otherwise rare species, in a level of detail and precision that would not be possible with multiple-gene data. Such evidence-based conservation interventions have been leveraged to support other threatened species in the genus *Cacatua*, including *C. moluccensis* on Seram, where rehabilitation and release of individuals confiscated by customs helped repopulate areas (Nandika et al. 2021), and for *C. s. abbotti*, where similar action has been instrumental in avoiding the complete extinction of this population, of which just 20–30 individuals remain (IUCN 2023).

One of the benefits of museomics is the ability to examine genetic data from species that are extinct, rare or inaccessible (Garg et al. 2022; Fong et al. 2023; Tan et al. 2023). It also provides an opportunity for consistency and continuity in the advancement of scientific knowledge—for example, in our case we were able to examine the genomes of some of the exact same specimens used in the morphological study on *C. sulphurea* by Collar and Marsden (2014). However, use of historical specimens from museum collections also introduces challenges such as mislabelling or genetic degradation—the latter of which can necessitate limiting the amount of genomic markers used, for example, to only include sites in ultra-conserved elements of the genome (e.g., Smith et al. 2020). When examining degraded samples, it is important to employ a genome-scale study, since from a technical perspective testing single-locus primers on degraded museum DNA would be time-consuming and result in less genetic information.

Ultimately, the genomic data retrieved in this study was of reasonable quality (Table 1), and missing data had little-to-no influence on overall patterns shown in the analyses (e.g., Figure S1). In addition, the SNPs were filtered to account for damage, and we included multiple individuals from the same subspecies for five of the seven putative subspecies for cross-validation—with different samples from the same subspecies consistently grouping together in all analyses despite varying levels of DNA degradation, coverage and missing data. We were therefore able to make use of 15 *C. sulphurea* samples and genotypes from across the entire genome in downstream analyses, adding valuable data from both mitochondrial genes and whole-genome derived SNPs to our collective knowledge of this critically endangered species.

4.1 | Conclusions

Our comprehensive genomic analyses of all putative *C. sulphurea* subspecies support the recognition of three genetically distinct subspecies clusters. One of these contains four closely related subspecies from the Sulawesi and Lesser Sunda region (*C. s. sulphurea*, *C. s. djampeana*, *C. s. occidentalis* and *C. s. paulandrewi*), of which *C. s. djampeana* and *C. s. paulandrewi* consistently emerged as poorly supported and embedded within *C. s. sulphurea* so as to suggest its synonymisation; other subspecies may require synonymisation pending reanalyses of morphological datasets. Our results also call into question the designation of *C. s. citrinocristata* as an independent species, as it emerged embedded within *C. sulphurea*. Our analyses further highlight the potential misclassification of *C. s. abbotti*, suggesting that it may not be a member of the *C. sulphurea* radiation. Its taxonomic status remains to be addressed. Apart from resolving genetic and evolutionary affiliations among the different populations and subspecies of this iconic and critically endangered species, our results should be useful in guiding conservation and management of the remnant wild *C. sulphurea* populations.

Author Contributions

All authors contributed to writing and approved this manuscript. A.A.A., C.D., F.E.R., J.L. and J.M. conceptualised the study. A.A.A. and C.D. coordinated sample collection. A.A.A. led the analyses and writing. A.F.S. helped with sampling specimens from museums, lab work at JLU and with analyses. K.R., T.H. and P.M. contributed analytical support and technical advice. G.L. conducted the lab work at NUS. F.E.R. contributed laboratory equipment and resources for DNA extraction/library prep at NUS, as well as guidance on analyses. C.D. and J.M. provided funding for the study and supervision.

Acknowledgements

For the provision of samples used in the study, we thank Paul Sweet (American Museum of Natural History; AMNH; CITES no. US 055), Pepijn Kamminga (Naturalis Biodiversity Center; RMNH; CITES no. NL 001), Chris Milensky (Smithsonian National Museum of Natural History; USNM; CITES no. US 005) and Christopher A. Wilson (Australian National Wildlife Collection; ANWC; CITES no. AU 022). We thank USNM for sample USNM181456, from which we were not able to successfully extract DNA despite repeated attempts. For provision of the Hong Kong cockatoo sample, we thank the Society for the Prevention of Cruelty to Animals, Hong Kong. For providing facilities,

we thank the Avian Evolution Lab and Lee Kong Chian Natural History Museum at the National University of Singapore (NUS), CITES registration number SG 002, and Thomas Wilke and Christian Albrecht of the Institute for Animal Ecology & Systematics, Justus Liebig University Gießen (JLU) CITES registration number DE 206-41. With gratitude we acknowledge the support from the Finnish IT Centre Scientific Computing (CSC) and funding from the University of Hong Kong (JM, CD) and General Research Fund of Research Grants Council of Hong Kong, grant number 17104824 (JM, AFS), and Wildlife Reserves Singapore (CD, AAA).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

Scripts relevant to this paper can be found at github.com/AstridAlexAndersson/CacSul and data can be accessed on Figshare project: 10.6084/m9.figshare.27199266.

References

Adams, M., P. R. Baverstock, D. A. Saunders, R. Schodde, and G. T. Smith. 1984. "Biochemical Systematics of the Australian Cockatoos (Psittaciformes: Cacatuinae)." *Australian Journal of Zoology* 32, no. 3: 363–377.

Agnarsson, I., and M. Kuntner. 2007. "Taxonomy in a Changing World: Seeking Solutions for a Science in Crisis." *Systematic Biology* 56, no. 3: 531–539.

Ali, J. R., J. C. Aitchison, and S. Meiri. 2020. "Redrawing Wallace's Line Based on the Fauna of Christmas Island, Eastern Indian Ocean." *Biological Journal of the Linnean Society* 130, no. 1: 225–237.

Andersson, A. A., L. Gibson, D. M. Baker, et al. 2021. "Stable Isotope Analysis as a Tool to Detect Illegal Trade in Critically Endangered Cockatoos." *Animal Conservation* 24, no. 6: 1021–1031.

Bergeron, L. A., S. Besenbacher, J. Zheng, et al. 2023. "Evolution of the Germline Mutation Rate Across Vertebrates." *Nature* 615, no. 7951: 285–291.

BirdLife International. 2021a. *Cacatua sulphurea*. The IUCN Red List of Threatened Species 2021: e.T200296187A178119524. https://doi.org/10.2305/IUCN.UK.2021-3.RLTS.T200296187A178119524.en.

BirdLife International. 2021b. *Cacatua citrinocristata*. The IUCN Red List of Threatened Species 2021: e.T200296433A201545199 https://doi.org/10.2305/IUCN.UK.2021-3.RLTS.T200296433A201545199.en.

Brown, D. M., and C. A. Toft. 1999. "Molecular Systematics and Biogeography of the Cockatoos (Psittaciformes: Cacatuidae)." *Auk* 116, no. 1: 141–157.

Cahill, A. J., J. S. Walker, and S. J. Marsden. 2006. "Recovery Within a Population of the Critically Endangered Citron-Crested Cockatoo *Cacatua sulphurea citrinocristata* in Indonesia After 10 Years of International Trade Control." *Oryx* 40, no. 2: 161–167.

Cardeñosa, D., J. Quinlan, K. H. Shea, and D. D. Chapman. 2018. "Multiplex Real-Time PCR Assay to Detect Illegal Trade of CITES-Listed Shark Species." *Scientific Reports* 8, no. 1: 16313.

Chang, C. C., C. C. Chow, L. C. Tellier, S. Vattikuti, S. M. Purcell, and J. J. Lee. 2015. "Second-Generation PLINK: Rising to the Challenge of Larger and Richer Datasets." *GigaScience* 4, no. 1: s13742.

Chattopadhyay, B., K. M. Garg, I. H. Mendenhall, and F. E. Rheindt. 2019. "Historic DNA Reveals Anthropocene Threat to a Tropical Urban Fruit Bat." *Current Biology* 29, no. 24: R1299–R1300.

Chornelia, A., J. Lu, and A. C. Hughes. 2022. "How to Accurately Delineate Morphologically Conserved Taxa and Diagnose Its Phenotypic Disparities: A Species Delimitation in Cryptic Rhinolophidae (Chiroptera) of Asia Lineages." *Frontiers in Ecology and Evolution* 10: 854509.

Christidis, L., D. D. Shaw, and R. Schodde. 1991. "Chromosomal Evolution in Parrots, Lorikeets and Cockatoos (Aves: Psittaciformes)." *Hereditas* 114, no. 1: 47–56.

Clements, J. F. 2000. *Birds of the World: A Checklist*. Fifth ed. Robertsbridge, UK: Pica Press.

Collar, N. J., and S. H. M. Butchart. 2014. "Conservation Breeding and Avian Diversity: Chances and Challenges." *International Zoo Yearbook* 48, no. 1: 7–28.

Collar, N. J., L. Gardner, D. F. Jeggo, et al. 2012. "Captive Breeding and the Most Threatened Birds in Asia." *BirdingASIA* 18: 50–57.

Collar, N. J., and S. J. Marsden. 2014. "The Subspecies of Yellow-Crested Cockatoo *Cacatua sulphurea*." *Forktail* 30: 23–27.

Courtney, J. 1996. "The Juvenile Food-Begging Calls, Food-Swallowing Vocalisation and Begging Postures in Australian Cockatoos." *Australian Bird Watcher* 16, no. 6: 236–249.

Dalton, H., J. Salo, P. Niemelä, and S. Orma. 2018. "Frederick II of Hohenstaufen's Australasian Cockatoo: Symbol of Detente Between East and West and Evidence of the Ayyubids' Global Reach." *Parergon* 35, no. 1: 35–60.

Dickinson, E. C., and J. V. Remsen Jr., eds. 2013. *The Howard & Moore Complete Checklist of the Birds of the World, 1*. Fourth ed. Eastbourne: Aves Press.

Eaton, J. A., B. van Balen, N. W. Brickle, and F. E. Rheindt. 2021. *Birds of the Indonesian Archipelago: Greater Sundas and Wallacea*. Second ed. Spain: Lynx Edicions.

Eliason, C. M., T. Hains, J. McCullough, M. J. Andersen, and S. J. Hackett. 2022. "Genomic Novelty Within a "Great Speciator" Revealed by a High-Quality Reference Genome of the Collared Kingfisher (*Todiramphus chloris collaris*)." *G3* 12, no. 11: jkac260.

Evanno, G., S. Regnaut, and J. Goudet. 2005. "Detecting the Number of Clusters of Individuals Using the Software STRUCTURE: A Simulation Study." *Molecular Ecology* 14, no. 8: 2611–2620.

Ewart, K. M., R. N. Johnson, R. Ogden, L. Joseph, G. J. Frankham, and N. Lo. 2019. "Museum Specimens Provide Reliable SNP Data for Population Genomic Analysis of a Widely Distributed but Threatened Cockatoo Species." *Molecular Ecology Resources* 19, no. 6: 1578–1592. https://doi.org/10.1111/1755-0998.13082.

Feng, S., J. Stiller, Y. Deng, et al. 2020. "Dense Sampling of Bird Diversity Increases the Power of Comparative Genomics." *Nature* 587, no. 7833: 252–257.

Fong, J. J., M. P. Blom, A. Aowphol, J. A. McGuire, C. Sutcharit, and P. S. Soltis. 2023. "Recent Advances in Museomics: Revolutionizing Biodiversity Research." *Frontiers in Ecology and Evolution* 11: 349.

Forshaw, J. M. 1989. *Parrots of the World*. Third (revised) ed. London: Blandford Press.

Garg, K. M., B. Chattopadhyay, E. Cros, et al. 2022. "Island Biogeography Revisited: Museomics Reveals Affinities of Shelf Island Birds Determined by Bathymetry and Paleo-Rivers, Not by Distance to Mainland." *Molecular Biology and Evolution* 39, no. 1: msab340.

Guo, B., B. Fang, T. Shikano, et al. 2019. "A Phylogenomic Perspective on Diversity, Hybridization and Evolutionary Affinities in the Stickleback Genus Pungitius." *Molecular Ecology* 28, no. 17: 4046–4064.

Gwee, C. Y., Q. L. Lee, S. P. Mahood, et al. 2021. "The Interplay of Colour and Bioacoustic Traits in the Differentiation of a Southeast Asian Songbird Complex." *Molecular Ecology* 30, no. 1: 297–309.

Harris, R. B., P. Alström, A. Ödeen, and A. D. Leaché. 2018. "Discordance Between Genomic Divergence and Phenotypic Variation in a Rapidly

Evolving Avian Genus (*Motacilla*)." *Molecular Phylogenetics and Evolution* 120: 183–195.

Harrison, N., and C. A. Kidner. 2011. "Next–Generation Sequencing and Systematics: What Can a Billion Base Pairs of DNA Sequence Data Do for You?" *Taxon* 60, no. 6: 1552–1566.

Holt, B. G., and K. A. Jønsson. 2014. "Reconciling Hierarchical Taxonomy With Molecular Phylogenies." *Systematic Biology* 63, no. 6: 1010–1017.

Huson, D. H., and D. Bryant. 2006. "Application of Phylogenetic Networks in Evolutionary Studies." *Molecular Biology and Evolution* 23, no. 2: 254–267.

IUCN. 2023. The IUCN Red List of Threatened Species *Cacatua Sulphurea* Yellow-crested Cockatoo. Version 2020-1 www.iucnredlist.org.

Jarvis, E. D., S. Mirarab, A. J. Aberer, et al. 2014. "Whole-Genome Analyses Resolve Early Branches in the Tree of Life of Modern Birds." *Science* 346, no. 6215: 1320–1331.

Jónsson, H., A. Ginolhac, M. Schubert, P. L. Johnson, and L. Orlando. 2013. "mapDamage2. 0: Fast Approximate Bayesian Estimates of Ancient DNA Damage Parameters." *Bioinformatics* 29, no. 13: 1682–1684.

Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg, and I. Mayrose. 2015. "Clumpak: A Program for Identifying Clustering Modes and Packaging Population Structure Inferences Across K." *Molecular Ecology Resources* 15, no. 5: 1179–1191.

Korneliussen, T. S., A. Albrechtsen, and R. Nielsen. 2014. "ANGSD: Analysis of Next Generation Sequencing Data." *BMC Bioinformatics* 15, no. 1: 1–13.

Lambeck, K., H. Rouby, A. Purcell, Y. Sun, and M. Sambridge. 2014. "Sea Level and Global Ice Volumes From the Last Glacial Maximum to the Holocene." *Proceedings of the National Academy of Sciences USA* 111, no. 43: 15296–15303.

Lambeck, K., Y. Yokoyama, and T. Purcell. 2002. "Into and out of the Last Glacial Maximum: Sea-Level Change During Oxygen Isotope Stages 3 and 2." *Quaternary Science Reviews* 21: 343–360. https://doi.org/10.1016/S0277-3791(01)00071-3.

Li, H., and R. Durbin. 2010. "Fast and Accurate Long-Read Alignment With Burrows–Wheeler Transform." *Bioinformatics* 26, no. 5: 589–595.

Li, H., and R. Durbin. 2011. "Inference of Human Population History From Individual Whole-Genome Sequences." *Nature* 475, no. 7357: 493–496.

Li, H., B. Handsaker, A. Wysoker, et al. 2009. "The Sequence Alignment/Map (SAM) Format and SAMtools." *Bioinformatics* 25, no. 16: 2078–2079.

Lin, R. S., and P. F. Lee. 2006. "Status of Feral Populations of Exotic Cockatoos (Genus Cacatua) in Taiwan." *Taiwania* 51, no. 3: 188–194.

Lou, R. N., A. Jacobs, A. P. Wilder, and N. O. Therkildsen. 2021. "A beginner's Guide to Low-Coverage Whole Genome Sequencing for Population Genomics." *Molecular Ecology* 30, no. 23: 5966–5993.

Martini, D., N. Dussex, B. C. Robertson, N. J. Gemmell, and M. Knapp. 2021. "Evolution of the "world's Only Alpine Parrot": Genomic Adaptation or Phenotypic Plasticity, Behaviour and Ecology?" *Molecular Ecology* 30, no. 23: 6370–6386.

Mayr, E. 1944. "Wallace's Line in the Light of Recent Zoogeographic Studies." *Quarterly Review of Biology* 19, no. 1: 1–14.

Meng, G., Y. Li, C. Yang, and S. Liu. 2019. "MitoZ: A Toolkit for Animal Mitochondrial Genome Assembly, Annotation and Visualization." *Nucleic Acids Research* 47, no. 11: e63.

Minh, B. Q., H. A. Schmidt, O. Chernomor, et al. 2020. "IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era." *Molecular Biology and Evolution* 37, no. 5: 1530–1534.

Momigliano, P., A. B. Florin, and J. Merilä. 2021. "Biases in Demographic Modeling Affect Our Understanding of Recent Divergence." *Molecular Biology and Evolution* 38, no. 7: 2967–2985.

Nadachowska-Brzyska, K., R. Burri, L. Smeds, and H. Ellegren. 2016. "PSMC Analysis of Effective Population Sizes in Molecular Ecology and Its Application to Black-And-White *Ficedula* Flycatchers." *Molecular Ecology* 25, no. 5: 1058–1072.

Nadachowska-Brzyska, K., M. Konczal, and W. Babik. 2022. "Navigating the Temporal Continuum of Effective Population Size." *Methods in Ecology and Evolution* 13, no. 1: 22–41.

Nandika, D., D. Agustina, R. Heinsohn, and G. Olah. 2021. "Wildlife Trade Influencing Natural Parrot Populations on a Biodiverse Indonesian Island." *Diversity* 13, no. 10: 483.

Nei, M., and W.-H. Li. 1979. "Mathematical Model for Studying Genetic Variation in Terms of Restriction Endonucleases." *Proceedings of the National Academy of Sciences USA* 76: 5269–5273.

Neo, M. L. 2012. "A Review of Three Alien Parrots in Singapore." *Nature in Singapore* 5: 241–248.

Okonechnikov, K., A. Conesa, and F. García-Alcalde. 2016. "Qualimap 2: Advanced Multi-Sample Quality Control for High-Throughput Sequencing Data." *Bioinformatics* 32, no. 2: 292–294.

Olah, G., S. H. M. Butchart, A. Symes, et al. 2016. "Ecological and Socio-Economic Factors Affecting Extinction Risk in Parrots." *Biodiversity and Conservation* 25, no. 2: 205–223.

Olah, G., B. T. Smith, L. Joseph, S. C. Banks, and R. Heinsohn. 2021. "Advancing Genetic Methods in the Study of Parrot Biology and Conservation." *Diversity* 13, no. 11: 521.

Ortiz, E. M. 2019. vcf2phylip v2.0: Convert a VCF Matrix Into Several Matrix Formats for Phylogenetic Analysis https://doi.org/10.5281/zenodo.2540861.

Pääbo, S., H. Poinar, D. Serre, et al. 2004. "Genetic Analyses From Ancient DNA." *Annual Review of Genetics* 38: 645–679.

Padial, J. M., A. Miralles, I. De la Riva, and M. Vences. 2010. "The Integrative Future of Taxonomy." *Frontiers in Zoology* 7, no. 1: 1–14.

Palkopoulou, E., S. Mallick, P. Skoglund, et al. 2015. "Complete Genomes Reveal Signatures of Demographic and Genetic Declines in the Woolly Mammoth." *Current Biology* 25, no. 10: 1395–1400.

Pires, S. F., G. Olah, D. Nandika, D. Agustina, and R. Heinsohn. 2021. "What Drives the Illegal Parrot Trade? Applying a Criminological Model to Market and Seizure Data in Indonesia." *Biological Conservation* 257: 109098.

Poinar, H. N., and A. Cooper. 2000. "Ancient DNA: Do It Right or Not at all." *Science* 5482, no. 1139: 416.

Provost, K. L., L. Joseph, and B. T. Smith. 2018. "Resolving a Phylogenetic Hypothesis for Parrots: Implications From Systematics to Conservation." *Emu-Austral Ornithology* 118, no. 1: 7–21.

R Core Team. 2021. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rambaut, A. 2018. FigTree v1. 4.4, a Graphical Viewer of Phylogenetic Trees. 2014. Java https://github.com/rambaut/figtree.

Rasmussen, M. S., G. Garcia-Erill, T. S. Korneliussen, C. Wiuf, and A. Albrechtsen. 2022. "Estimation of Site Frequency Spectra From Low-Coverage Sequencing Data Using Stochastic EM Reduces Overfitting, Runtime, and Memory Usage." *Genetics* 222, no. 4: iyac148.

Remsen, J. V., Jr. 2010. "Chapter 6: Subspecies as a Meaningful Taxonomic Rank in Avian Classification." *Ornithological Monographs* 67, no. 1: 62–78.

Reuleaux, A., N. J. Collar, M. J. Jones, et al. 2022. "Correlates of Persistence in Remnant Populations of Two Critically Endangered Cockatoos." *Animal Conservation* 25, no. 6: 811–824.

Rheindt, F. E., L. Christidis, S. Kuhn, S. de Kloet, J. A. Norman, and A. Fidler. 2014. "The Timing of Diversification Within the Most Divergent Parrot Clade." *Journal of Avian Biology* 45: 140–148.

Rheindt, F. E., L. Christidis, and J. A. Norman. 2009. "Genetic Introgression, Incomplete Lineage Sorting and Faulty Taxonomy Create Multiple Cases of Polyphyly in a Montane Clade of Tyrant-Flycatchers (Elaenia, Tyrannidae)." *Zoologica Scripta* 38, no. 2: 143–153.

Rheindt, F. E., and S. V. Edwards. 2011. "Genetic Introgression: An Integral but Neglected Component of Speciation in Birds." *Auk* 128, no. 4: 620–632.

Roux, C., C. Fraisse, J. Romiguier, Y. Anciaux, N. Galtier, and N. Bierne. 2016. "Shedding Light on the Grey Zone of Speciation Along a Continuum of Genomic Divergence." *PLoS Biology* 14, no. 12: e2000234.

Sadanandan, K. R., G. W. Low, S. Sridharan, et al. 2020. "The Conservation Value of Admixed Phenotypes in a Critically Endangered Species Complex." *Scientific Reports* 10, no. 1: 15549.

Safran, R. J., E. S. C. Scordato, M. R. Wilkins, et al. 2016. "Genome-Wide Differentiation in Closely Related Populations: The Roles of Selection and Geographic Isolation." *Molecular Ecology* 25, no. 16: 3865–3883.

Schubert, M., S. Lindgreen, and L. Orlando. 2016. "AdapterRemoval v2: Rapid Adapter Trimming, Identification, and Read Merging." *BMC Research Notes* 9, no. 1: 1–7.

Schweizer, M., O. Seehausen, M. Güntert, and S. T. Hertwig. 2010. "The Evolutionary Diversification of Parrots Supports a Taxon Pulse Model With Multiple Trans-Oceanic Dispersal Events and Local Radiations." *Molecular Phylogenetics and Evolution* 54, no. 3: 984–994.

Selvatti, A. P., A. Galvão, G. Mayr, C. Y. Miyaki, and C. A. D. M. Russo. 2022. "Southern Hemisphere Tectonics in the Cenozoic Shaped the Pantropical Distribution of Parrots and Passerines." *Journal of Biogeography* 49, no. 10: 1753–1766.

Semenov, G. A., E. A. Koblik, Y. A. Red'kin, and A. V. Badyaev. 2018. "Extensive Phenotypic Diversification Coexists With Little Genetic Divergence and a Lack of Population Structure in the White Wagtail Subspecies Complex (*Motacilla alba*)." *Journal of Evolutionary Biology* 31, no. 8: 1093–1108.

Siddall, M., E. J. Rohling, A. Almogi-Labin, et al. 2003. "Sea-Level Fluctuations During the Last Glacial Cycle." *Nature* 423: 853–858. https://doi.org/10.1038/nature01690.

Skotte, L., T. S. Korneliussen, and A. Albrechtsen. 2013. "Estimating Individual Admixture Proportions From Next Generation Sequencing Data." *Genetics* 195, no. 3: 693–702.

Smith, B. T., W. M. Mauck III, B. W. Benz, and M. J. Andersen. 2020. "Uneven Missing Data Skew Phylogenomic Relationships Within the Lories and Lorikeets." *Genome Biology and Evolution* 12, no. 7: 1131–1147.

Smith, G. A. 1975. "Systematics of Parrots." *Ibis* 117, no. 1: 18–68.

Tan, H. Z., J. J. F. J. Jansen, G. A. Allport, et al. 2023. "Megafaunal Extinctions, Not Climate Change, May Explain Holocene Genetic Diversity Declines in Numenius Shorebirds." *eLife* 12: e85422.

Thomson, S. A., R. L. Pyle, S. T. Ahyong, et al. 2018. "Taxonomy Based on Science Is Necessary for Global Conservation." *PLoS Biology* 16, no. 3: e2005075.

Villanueva, R. A. M., and Z. J. Chen. 2019. "ggplot2: Elegant Graphics for Data Analysis." *Measurement: Interdisciplinary Research and Perspectives* 17, no. 3: 160–167.

Von Seth, J., N. Dussex, D. Díez-Del-Molino, et al. 2021. "Genomic Insights Into the Conservation Status of the world's Last Remaining Sumatran Rhinoceros Populations." *Nature Communications* 12, no. 1: 2393.

White, N. E., M. J. Phillips, M. T. P. Gilbert, et al. 2011. "The Evolutionary History of Cockatoos (Aves: Psittaciformes: Cacatuidae)." *Molecular Phylogenetics and Evolution* 59, no. 3: 615–622.

Wingett, S. W., and S. Andrews. 2018. "FastQ Screen: A Tool for Multi-Genome Mapping and Quality Control." *F1000 Research* 7.

Wortley, A. H., and R. W. Scotland. 2006. "The Effect of Combining Molecular and Morphological Data in Published Phylogenetic Analyses." *Systematic Biology* 55, no. 4: 677–685.

Wu, M. Y., G. W. Low, G. Forcina, et al. 2020. "Historic and Modern Genomes Unveil a Domestic Introgression Gradient in a Wild Red Junglefowl Population." *Evolutionary Applications* 13, no. 9: 2300–2315.

Yi, X., and E. K. Latch. 2022. "Nonrandom Missing Data Can Bias Principal Component Analysis Inference of Population Genetic Structure." *Molecular Ecology Resources* 22, no. 2: 602–611.

Zein, M. S. A., T. Haryoko, Y. S. Fitriana, E. Sulistyadi, and D. M. Prawiradilaga. 2017. "Aplikasi kajian DNA molekuler dan fenotipik pada program pelepasliaran burung kakatua." *Jurnal Biologi Indonesia* 13, no. 1.

Supporting Information

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