



Genome Resources

Genomic resources for comparative analyses of avian obligate brood parasitism

Rachel A. Carroll¹, Edward S. Ricemeyer^{1,2}, LaDeana W. Hillier³, Jeffrey M. DaCosta⁴, Ekaterina Osipova⁵, Sara Smith⁶, Gabriel A. Jamie^{7,8}, Juan G. Martinez⁹, Mercedes Molina-Morales⁹, Tomas Marques-Bonet^{10,11,12,13}, Joseph D. Manthey¹⁴, Diana Haddad¹⁵, Matthew J. Fuxjager¹⁶, Kathleen S. Lynch¹⁷, Jonathan M.D. Wood¹⁸ , Erich Jarvis^{19,20}, Patrick Masterson¹⁴, Francoise Thibaud-Nissen¹⁴, Mark E. Hauber²¹, Claire N. Spottiswoode^{7,8}, Timothy B. Sackton⁵, Christopher N. Balakrishnan^{22,23} , Michael D. Sorenson²⁴, Wesley C. Warren^{1,25,26,*}

¹Division of Animal Sciences, Bond Life Sciences Center, University of Missouri, Columbia, MO, United States, ²Palaeogenomics Group, Institute of Palaeoanatomy, Domestication Research and the History of Veterinary Medicine, Ludwig-Maximilians-Universität, Munich, Germany, ³Department of Genome Sciences, University of Washington, Seattle, WA, United States, ⁴Biology Department, Boston College, Boston, MA, United States, ⁵Informatics Group, Harvard University, Cambridge, MA, United States, ⁶Department of Biology, Mount Royal University, Calgary, AB, Canada, ⁷FitzPatrick Institute of African Ornithology, Department of Biological Sciences, University of Cape Town, Cape Town, South Africa, ⁸Department of Zoology, University of Cambridge, Downing Street, CB2 3EJ Cambridge, United Kingdom, ⁹Departamento de Zoología, Universidad de Granada, Granada, Spain, ¹⁰Institute of Evolutionary Biology (UPF-CSIC), PRBB, Dr. Aiguader 88, Barcelona 08003, Spain, ¹¹Catalan Institution of Research and Advanced Studies (ICREA), Passeig de Lluís Companys, 23, Barcelona 08010, Spain, ¹²CNAG, Centro Nacional de Analisis Genómico, Baldiri i Reixac 4, 08010 Barcelona, Spain, ¹³Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Edifici ICTA-ICP, c/ Columnes s/n, 08193 Cerdanyola del Vallès, Barcelona, Spain, ¹⁴Department of Biological Sciences, Texas Tech University, Lubbock, TX, United States, ¹⁵National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, United States, ¹⁶Department of Ecology, Evolution, and Organismal Biology, Brown University, Providence, RI, United States, ¹⁷Hofstra University, Long Island, NY, United States, ¹⁸Wellcome Sanger Institute, Cambridge, United Kingdom, ¹⁹Laboratory of Neurogenetics of Language, The Rockefeller University, New York, NY, United States, ²⁰The Vertebrate Genome Laboratory, New York, NY, United States, ²¹Advanced Science Research Center and Programs in Biology and in Psychology, Graduate Center of the City University of New York, New York, NY, United States, ²²Department of Biology, East Carolina University, Greenville, NC, United States, ²³Division of Environmental Biology, National Science Foundation, Alexandria, VA, United States, ²⁴Department of Biology, Boston University, Boston, MA, United States, ²⁵Department of Surgery, School of Medicine, University of Missouri, Columbia, MO, United States, ²⁶Institute of Data Science and Informatics, University of Missouri, Columbia, MO, United States

*Corresponding author: Division of Animal Sciences, Bond Life Sciences Center, University of Missouri, Columbia, MO, United States.
Email: warrenwc@missouri.edu

Abstract

Examples of convergent evolution, wherein distantly related organisms evolve similar traits, including behaviors, underscore the adaptive power of natural selection. In birds, obligate brood parasitism, and the associated loss of parental care behaviors, has evolved independently in seven different lineages, though little is known about the genetic basis of the complex suite of traits associated with this rare life history strategy. We generated genome assemblies for ten brood parasitic species plus eight species representatives of their parental/nesting outgroups. This includes nine long-read chromosome-level assemblies, with scaffold N50 sizes ranging from 38.1 to 72.6 MB, and gene representation completeness measures >97%. Leveraging this new catalog of avian genomes, we constructed clade-level alignments that reveal variation in chromosomal synteny, provide first-time or improved annotations of protein-coding and non-coding genes, and define cross-species ortholog reference sets. We also refine estimates for the timing of the seven independent origins of brood parasitism, ranging from recent events such as 1.6–4.5 million years ago in *Molothrus* cowbirds to much earlier origins over 30 million years ago in two of the three cuckoo lineages. These genomic resources lay the foundation for investigating the genetic and genomic underpinnings of brood parasitism, including the loss of parental care, shifts in mating systems, perhaps resulting in heightened sperm competition, elevated annual fecundity, improved spatial cognition related to nest-finding, and the diverse adaptations shaped by intense coevolution with host species.

Key words: assemblies, birds, brood parasitism

Received on 30 July 2025; accepted on 21 October 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of The American Genetic Association. All rights reserved. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Introduction

About 100 avian species, or 1% of all taxa, are obligate brood parasites, reproducing only by laying their eggs in the nests of other bird species and thereby exploiting the parental care behavior of their ‘hosts’ (Davies 2000). Obligate brood parasitism has independent origins in seven different lineages across four divergent avian orders (Payne 1977; Sorenson and Payne 2002), and so provides a compelling opportunity for comparative genomic analysis. Shared among these lineages is the loss of all parental care behaviors, including nest-building, incubation, and the brooding, provisioning, and defending own offspring.

Comparative genomic analyses of adaptation have often considered the convergent evolution of specific traits, such as novel sensory traits (Parker et al. 2013; Sadanandan et al. 2023; Zou et al. 2024), adoption of new diets (Hu et al. 2017; Toda et al. 2021; Eliason et al. 2023), or invasion of extreme environments (Foote et al. 2015; Cole et al. 2022; Lyu et al. 2022). In other cases, convergent adaptation involves the loss of pre-existing traits such as flight (Sackton et al. 2019) or functional vision (Moran et al. 2023). In comparison, the loss of parental care in parasitic birds represents a major behavioral and life history transition with multiple axes of potential convergence. In addition to the loss of parental behaviors, a substantial increase in female fecundity is another trait likely shared across all seven parasitic lineages (Payne 1974; Payne 1976), reflecting a reallocation of time and energy from parental care to egg production (Payne 1977). Other axes of possible convergence are not as well characterized, but may include altered sexual selection associated with loss of parental care (Feeney and Riehl 2019), better spatial ability associated with finding and tracking the status of host nests (Sherry et al. 1993; Reboreda et al. 1996; Nair-Roberts et al. 2006; Guigueno et al. 2014) and accelerated pre- and/or post-hatch development (Birkhead et al. 2011) and novelties in behavioral imprinting and song learning (Payne et al. 1998; Louder et al. 2019). Many brood parasitic birds and their hosts also serve as important models for coevolutionary arms-races and their reciprocal interactions have produced a spectacular diversity of behavioral, morphological and physiological adaptations, and counter-adaptations e.g. (Langmore et al. 2003; Yang et al. 2010; Spottiswoode and Stevens 2011; Langmore et al. 2011; Lund et al. 2023). In some cases, coevolution promotes speciation (Sorenson et al. 2003; Langmore et al. 2024) and in others the evolution and maintenance of host-specific matrilineal (Spottiswoode and Stevens 2011; Fossoy et al. 2016; Spottiswoode et al. 2022), with important genomic consequences.

Robust genomic resources for representative parasitic birds are needed to facilitate analyses of both convergence and the varied lineage-specific adaptations produced in their coevolutionary arms races with hosts (Sorenson and Payne 2002; Ronka et al. 2024). Toward this end, we formed the Brood Parasitic Bird Genomes Consortium (BPBGC) and generated new genome assemblies for ten brood parasites that, combined with a recently published genome assembly for common cuckoo *Cuculus canorus* (Merondun et al. 2024) provides at least one representative for each of the seven evolutionary origins of obligate parasitism in extant birds. In addition, the genomes for eight nesting/parental bird species were assembled to provide outgroup comparisons for different parasitic lineages (Table 1). Three of the 18 new genomes reported here were produced in collaboration with the Vertebrate Genomes

Project (Rhie et al. 2021). Female samples were chosen for all genome assemblies to include the female-specific W chromosome, which along with the mitochondrial DNA (mtDNA), is expected to be an important target of selection in some brood parasitic lineages e.g. (Spottiswoode et al. 2011; Spottiswoode and Stevens 2011; Fossoy et al. 2016; Spottiswoode et al. 2022).

Methods

Biological materials

Samples of female birds were acquired from museum collections, live birds in captivity, or from materials obtained in past or current field research by us and colleagues, with the appropriate permits or Material Transfer Agreements (MTA) (Supplemental Table 1).

Sequencing

High molecular weight DNA (HMW DNA) was isolated from female muscle samples for each of the following 14 species that were designated for continuous long-read (CLR) or circular consensus sequence (CCS/HiFi) sequencing: CLR (brown-headed cowbird, *Molothrus ater*; village indigobird, *Vidua chalybeata*; greater honeyguide; *Indicator indicator* and downy woodpecker; *Dryobates pubescens*) and HiFi reads (bronzed cowbird, *Molothrus aeneus*; red-winged blackbird, *Agelaius phoeniceus*; cuckoo finch, *Anomalospiza imberbis*, pin-tailed whydah, *Vidua macroura*; great spotted cuckoo, *Clamator glandarius*; striped cuckoo, *Tapera naevia*; pavonine cuckoo, *Dromococcyx pavoninus*; gray-capped cuckoo, *Coccyzus lansbergi*; chestnut-breasted malkoha, *Phaenicophaeus curvirostris*; and rufous-vented ground cuckoo, *Neomorphus geoffroyi*). DNA was isolated in accordance with the 10X Genomics Demonstrated Protocol: DNA Extraction from Single Insects (10X Genomics). The final DNA amount was then measured using the high-sensitivity Invitrogen Qubit Fluorometer protocol, with quality assessed by a 0.7% agarose gel and imaged on the Uvitec Cambridge Uvidoc HD6 UV Fluorescence and Colorimetry instrument. SMRTbell libraries for sequencing were prepared according to the manufacturer’s recommended protocol (Pacific Biosciences-PacBio; Menlo Park, CA). In brief, HMW DNA was sheared to a target size of ~15–20 kb using a Megaruptor 3 system (Diagenode), followed by DNA end-repair and barcode adaptor ligation to generate circular SMRTbell templates for sequencing. Size selection was then performed using BluePippin (Sage Science) to enrich for fragments ≥ 15 kb. All SMRTbell libraries were sequenced on a PacBio instrument. Over the course of this study instrumentation upgrades occurred that included Sequel I, II, RSII, or Revio systems. The relevant sequencing mode, e.g. CLR or HiFi, was run according to the manufacturers instrument-specific protocols. Instrument use per species is specified in Supplemental Table 1. All relevant software used in this study for the purpose of genome assembly and curation, as well as any downstream genome analyses is presented in Table 2.

We also isolated HMW DNA from four duck species: black-headed duck (*Heteronetta atricapilla*), freckled duck (*Stictonetta naevosa*), ruddy duck (*Oxyura jamaicensis*), and African pygmy goose (*Nettapus auritus*) to generate long DNA fragment length linked-read libraries (10x Genomics, Pleasonton, CA) for each. Library preparation was carried

Table 1. Reference genome assembly statistics for all brood parasitic* and parental outgroup assemblies generated by the BPBGC.

Species	Genome Size (Gb)	Contig N50 (Mb)	Scaffold N50 (Mb)	Protein coding genes	BUSCO % Complete
<i>Molothrus aeneus</i> *	1.18	25.8	70.6	16 194	99.1
<i>Molothrus ater</i> *, ^a	1.09	13.3	70.1	16 046	99.3
<i>Agelaius phoeniceus</i>	1.19	22.8	36.4	16 104	98.9
<i>Vidua macroura</i> *	1.12	7.5	61.8	16 296	98.8
<i>Vidua chalybeata</i> *, ^a	1.10	15.8	72.6	16 658	98.9
<i>Anomalospiza imberbis</i> *	1.19	12.3	72.4	16 778	98.6
<i>Indicator indicator</i> *	1.15	8.7	38.1	14 393	97.1
<i>Dryobates pubescens</i> ^a	1.20	19.2	38.7	15 338	98.1
<i>Coccyzus lansbergi</i>	1.30	1.6	1.6	NA	NA
<i>Phaenicophaeus curvirostris</i>	1.31	17.1	70.7	16 797	98.0
<i>Clamator glandarius</i> *	1.40	2.3	2.3	NA	NA
<i>Dromococcyx pavoninus</i> *	1.20	1.8	1.8	NA	NA
<i>Tapera naevia</i> *	1.30	8.5	8.5	NA	NA
<i>Neomorphus geoffroyi</i>	1.20	24	24	NA	NA
<i>Heteronetta atricapilla</i> *	1.20	0.15	23.4	NA	NA
<i>Stictonetta naevosa</i>	1.10	0.14	10.7	NA	NA
<i>Oxyura jamaicensis</i>	1.10	0.13	36.1	16 416	99.0
<i>Nettapus auritus</i>	1.10	0.11	26	NA	NA

Genomes representing the seven independent origins of obligate brood parasitism are indicated in Fig. 3, with the recently published common cuckoo genome (Merondun et al. 2024) representing one of the seven origins. For RefSeq assemblies the number of annotated protein-coding genes and completeness for BUSCO genes are provided. We report for the contig only assemblies the same N50 contig and scaffold length metrics. NA: Not available. ^a Genome assemblies generated in collaboration with the Vertebrate Genomes Project.

Table 2. Study software used.

Software and any non-default options	Version or access date	Reference
De novo assembly		
Canu	2.1	(Koren et al. 2017)
FALCON	1.3.0	(Biosciences 2019)
Flye	2.8.1	(Kolmogorov et al. 2019)
HiFiasm	2021–2023	(Cheng et al. 2021)
Supernova2	2.1.1	(Visendi 2022)
Scaffolding and chromosome formatting		
Chromap	0.2.3	(Zhang et al. 2021)
SALSA	2.2	(Ghurye et al. 2019)
Juicebox	2.3.6	(Robinson et al. 2018)
Juicertools	1.6–3.0.0	(Durand et al. 2016)
YaHS	1.2a.1	(Zhou et al. 2023)
Bionano Solve	3.2.1	(Bionano 2019)
Scaff10x	4.1.0	(Institute 2020)
Longranger Align	2.2.2	(Genomics 2018)
Agptools	1.0	(Ricemeyer 2025; Ricemeyer et al. 2025)
Assembly polishing and purging		
Arrow	6.0	(Pacific Biosciences of California 2019)
freebayes	1.3.1–1.3.5	(Garrison and Marth 2012)
Minimap2	2.21	(Li 2018)
PurgeHaplotigs	1.1.3	(Roach et al. 2018)
Assembly QC		
BUSCO	4.1.4	(Simao et al. 2015)
Gene orthology and chromosome synteny		
OrthoFinder	2.5.4	(Emms and Kelly 2019)
Genespace	2022	(Lovell et al. 2022)
Z chromosome evaluation		
BOWTIE2	2.5.1	(Langmead and Salzberg 2012)
mosdepth	0.3.1	(Pedersen and Quinlan 2017)
ANGSD	0.940	(Korneliusson et al. 2014)
Phylogenetic analysis		
PHYLUCS	1.7.1	(Faircloth 2016)
BEAST	2.7.7	(Bouckaert et al. 2019)

out with the 10× Genomics (Pleasanton, CA) Chromium™ Genome Library and Gel Bead Kit v2 as described in Zhang et al. (2019). We sequenced each linked-read library at 100 bp read length to a targeted 60× depth of coverage on an Illumina NovaSeq 6000 instrument.

All sequences for each species in this study, regardless of type, have been deposited in the NCBI sequence read archive (SRA). All relevant data are retrievable by searching NCBI Bioproject accession numbers provided in Supplemental Table 1).

Genome assembly and curation

Over the course of this project, we performed de novo genome assembly on CLR or HiFi sequence data using a variety of algorithms. Each assembly algorithm was chosen and deployed based on the type of data generated and current best practices at the time of sequence data availability and its type. We specify the assembly algorithm used for each species in [Supplemental Table 1](#). Overall, we used one of four programs to assemble long-reads throughout the course of this study: Canu (Koren et al. 2017), FALCON (Biosciences 2019), Flye (Kolmogorov et al. 2019), or Hifiasm (Cheng et al. 2021). All primary assemblies were subjected to a version of PurgeHaplotigs (Roach et al. 2018) that was run on all contig assemblies to remove potential alternate haploid contigs using a custom nextflow pipeline (Warrenlab 2022).

For the nine species that were designated for chromosome scale assemblies using either CLR or HiFi reads (greater honeyguide, downy woodpecker, brown-headed cowbird, bronzed cowbird, red-winged blackbird, cuckoo finch, pin-tailed whydah, village indigobird, and chestnut-breasted malkoha) chromatin proximity interaction maps were generated for contig scaffolding. High-throughput Chromosome Conformation Capture (Hi-C) libraries were prepared using the Arima Genomics (Carlsbad, CA) or Phase Genomics (Seattle, WA) Hi-C pipelines by following the manufacturer's standard protocols for each. All Hi-C libraries were sequenced on an Illumina Novaseq 6000 (San Diego, CA) instrument using a saturation mapping estimation of 100 M paired end reads per Gb of genome size with bird genomes on average estimated to be 1.2Gb in length.

Chromosome scaffolding was performed using combinations of various tools that included Chromap (Zhang et al. 2021), SALSA2 (Ghurye et al. 2019), Juicebox (Robinson et al. 2018), Juicertools (Durand et al. 2016), and YaHS (Zhou et al. 2023). We provide the combination of mapping programs used for each species in [Supplemental Table 1](#). For all draft assemblies, including those based on CLR or HiFi read sources, base error was corrected with Arrow. An additional polishing step to correct assembly base call error for species with CLR data input was performed on greater honeyguide, downy woodpecker, and village indigobird using Illumina sequence data sourced from the same sample. Trimmomatic was used to trim adaptors, low-quality bases, and contaminants from Illumina sequences prior to Burrows-Wheeler Aligner BWA-MEM alignment and base call correction. Iterative polishing was completed with freebayes integrated into the Vertebrate Genomes Project (VGP) assembly pipeline v.1. (VGPassembly 2019) or a custom cextflow pipeline (Ricemeyer 2025). Briefly, this later pipeline aligns short reads to the assembly with minimap2 v2.21 (Li 2018), calls variants using freebayes v1.3.5 Garrison and Marth 2012, and then edits the assembly to match the consensus from the variant calls. We ran this pipeline twice iteratively for each CLR assembly, except for the brown-headed cowbird, for which there was insufficient tissue from the same sample to generate short reads.

After polishing a sequential process was employed to build and format chromosomes by assembling the longest-range scaffolds possible for each species using Hi-C associated procedures previously described in Rhie et al. (2021). Briefly, Hi-C reads were mapped to the draft genome assembly using BWA-MEM, alignments were processed to generate a contact map, and resulting scaffolds iteratively ordered and

oriented primarily with SALSA (Ghurye et al. 2019), followed by manual correction of scaffolding errors the Juicebox Assembly Tools interface (Robinson et al. 2018). The pseudo-chromosomes were aligned to the assembled chromosomes of the most closely related bird species available at that time to manually examine discrepancies and fix gross assembly errors. Agptools was used to finalize chromosome formatting (Ricemeyer 2025). For consistency and future comparative studies, all autosomes were numbered according to total size starting with the largest in accordance with the VGP guidelines (Rhie et al. 2021). As a result, the numerical assignment of chromosomes will not always reflect homology across species.

The genomes of the duck clade including the black-headed duck, freckled duck, ruddy duck, and African pygmy goose were assembled with Supernova v2.1.1 (Visendi 2022) following the methods outlined in Zhang et al. (2019). Initial chromosomal assignments were determined by aligning assembly scaffolds from each species to the reference genomes of domestic chicken (*Gallus domesticus*, GCA_000002315.5) and the mallard (*Anas platyrhynchos*, GCA_003850225.1). To improve chromosomal accuracy, we conducted iterative pairwise alignments among the four duck genomes, allowing each to inform the structural organization of the others. Final chromosome numbering was assigned based on synteny with the chicken genome.

All chromosome assemblies were scanned for repetitive sequences using WindowMasker (Morgulis et al. 2006), enabling the resulting output to be utilized for future genome masking steps in downstream analyses. Assembly completeness was evaluated using BUSCO v4.1.4 with the aves_odb10 lineage dataset to assess conserved gene content coverage (Simao et al. 2015). Additionally, corrections made to coding sequence (CDS) predictions using the Splign alignment algorithm Kapustin et al. 2008 were used as an independent measure of consensus sequence accuracy, with the number of CDS corrections reflecting potential base-calling errors.

Gene annotation

Gene annotations for all chromosome-scale assemblies were generated using the NCBI Eukaryotic Genome Annotation Pipeline, as described in Rhie et al. (2021). RNA-seq datasets from diverse tissue types, publicly available in the NCBI SRA for each species, were integrated into the annotation process to inform and support transcript model predictions. As part of this process, protein alignment to the genome was used to generate preliminary predicted models, and to correct as needed the final models with compensating insertions or deletions to correct for spurious indels in the genome. These errors were tallied as an additional comparative metric of annotation quality.

Gene orthology

To compare gene orthology across the avian genomes in this study, we ran OrthoFinder version 2.5.4 using the RefSeq protein-coding gene predictions for each species as input (Emms and Kelly 2019). To mitigate the effect of multiple transcripts per gene, we used only primary transcripts (the longest version of protein) only, per OrthoFinder recommendations (Emms and Kelly 2019).

Whole genome interspecies synteny analysis

To explore shared gene organization across parasitic species and parental outgroups, we analyzed syntenic orthology

among species using the Genespace R package (Lovell et al. 2022). The parse annotation function in Genespace was used to extract a gene Browser Extensible Data (BED) file from the General Feature Format (GFF) files of each species found in the NCBI Reference Sequence Database. After ensuring proper formatting and matching headers between protein and gene BED files, we initiated Genespace as described in Lovell et al (Lovell et al. 2022). Pairwise comparisons within clades revealed that, in some cases, entire chromosomes appeared to be inverted end-to-end. However, these apparent reversals were attributed to species-specific differences in chromosomal orientation in our assemblies, rather than true biological inversions. To ensure more accurate representation of syntenic relationships across species, such chromosomes were manually reoriented for the purpose of visualization.

Z chromosome evaluation in downy woodpecker

To better understand the unusually long Z-chromosome in downy woodpecker, we mapped short-read sequencing data to the reference genome using BOWTIE2 v. 2.5.1 Langmead and Salzberg 2012 and then used mosdepth v. 0.3.1 Pedersen and Quinlan 2017 to calculate sequencing depth in 10 kbp windows for two male and four female downy woodpeckers. We used ANGSD Korneliussen et al. 2014 to calculate individual heterozygosity in 50 kbp windows for these same six individuals.

Phylogenetic analysis

Sequence data from ultraconserved elements (UCEs) provide a powerful tool for estimating a time-calibrated phylogeny (time tree) of avian species and offer an opportunity to investigate the evolutionary origins of obligate brood parasitism. We used UCEs to estimate a time tree for 52 species comprising representatives of the seven lineages of obligate brood parasites, closely related nesting relatives, and other selected lineages in the avian tree of life (Supplemental Table 2). UCEs were extracted from whole genome assemblies using PHYLUCE v1.7.1 (Faircloth 2016) and a probe set designed to target ~5000 loci (Faircloth et al. 2012). We extracted 1000 bp of the flanking regions for each locus, and sequences for individual loci were aligned and edge trimmed. Loci with < 47 taxa represented were filtered out, resulting in 4491 retained UCE loci that were concatenated and used for phylogenetic inference in BEAST v2.7.7 (Bouckaert et al. 2019). Based on results from a recent phylogenomic study of birds, we constrained the analysis with several monophyletic priors, including monophyly of most families and key higher-level relationships (Stiller et al. 2024). We also adopted six fossil-based node age calibrations from this recent study (Stiller et al. 2024), applying them as log-normal distribution priors to estimate branch lengths in units of millions of years. We initiated 10 independent runs with different starting points in BEAST, each with the GTR+I+G model of sequence evolution, a birth-death coalescent prior, and sampling every 1000 generations. Runs were stopped after 5 M generations, and inspection of posteriors found that six of the runs reached stationarity at similar likelihood scores. After removing the appropriate burn-in for each run, results were combined across these six runs to form a posterior sample of 10 000 trees that were used to estimate a maximum clade credibility time tree.

Results and discussion

Chromosome scale assemblies of long-read sequenced avian genomes

Among the 18 avian genome assemblies we generated over a 4-year span, nine were assembled to the chromosome level using only CLR or HiFi input sequences; these include both brood parasitic and non-parasitic outgroup species (Table 1; Supplemental Table 1). Female samples were chosen for all species to allow analyses of the W-chromosome. Considering all sequence technologies employed, we achieved a total depth of sequence coverage ranging from 30× to 254× irrespective of sequence type (Supplemental Table 1). The total assembly sizes exhibited a narrow range of 1.1–1.3 Gb with N50 contig and scaffold length, respectively, from 7.49 to 25.8 Mb and 38.1 to 72.6 Mb, respectively (Table 1). Improvements in assembly contiguity were achieved as sequencing capacity and accuracy improved along with advancements in de novo assembly algorithms, e.g. PacBio CLR compared to HiFi inputs (Table 1; Supplemental Table 1). This is illustrated in our most recently completed assembly, the bronzed cowbird, which has the highest metrics of contiguity (Table 1). Despite the improvements in overall assembly quality between earlier and later assembled species, we judge all the chromosome scale references to be highly contiguous and of sufficient accuracy for various comparative genomic analyses.

Contig-level assemblies for additional cuckoo species

A total of five contig level assemblies were generated from PacBio/HiFi data for greater representation of cuckoo species, including great spotted cuckoo, gray-capped cuckoo, pavonine cuckoo, striped cuckoo, and rufous-vented ground cuckoo. A HiFi sequence coverage depth of 30× was achieved for all except the great spotted cuckoo at 19× (Supplemental Table 1). Total assembled genome sizes range from 1.2 to 1.39 Gb with contig N50 length of 1.6–8.5 Mb (Supplemental Table 3). For these cuckoo species, no contig scaffolding or chromosome-scale formatting was completed. Nonetheless, these contig-only assemblies offer researchers broader access to the cuckoo clade, including two additional independent origins of obligate brood parasitism, and valuable resources for gene-level analyses and population genetic studies.

Chromosome-scale assemblies of linked-read sequenced duck genomes

Chromosome-scale assemblies were also generated for four duck species: the brood parasitic black-headed duck, and representatives of three closely related, non-parasitic genera, the African pygmy goose, ruddy duck, and freckled duck. However, all these assemblies were generated with long range fragment sequences from the 10× pipeline as described previously (Zhang et al. 2019). Assembly contiguity statistics were as expected for this approach with ranges of N50 contig and scaffold lengths being 105–151 kb and 11–36 Mb, respectively (Supplemental Table 3). Despite the high-level scaffolding contiguity achieved in these assemblies the total contig numbers remain high due to the short-read assembler input (Supplemental Table 3).

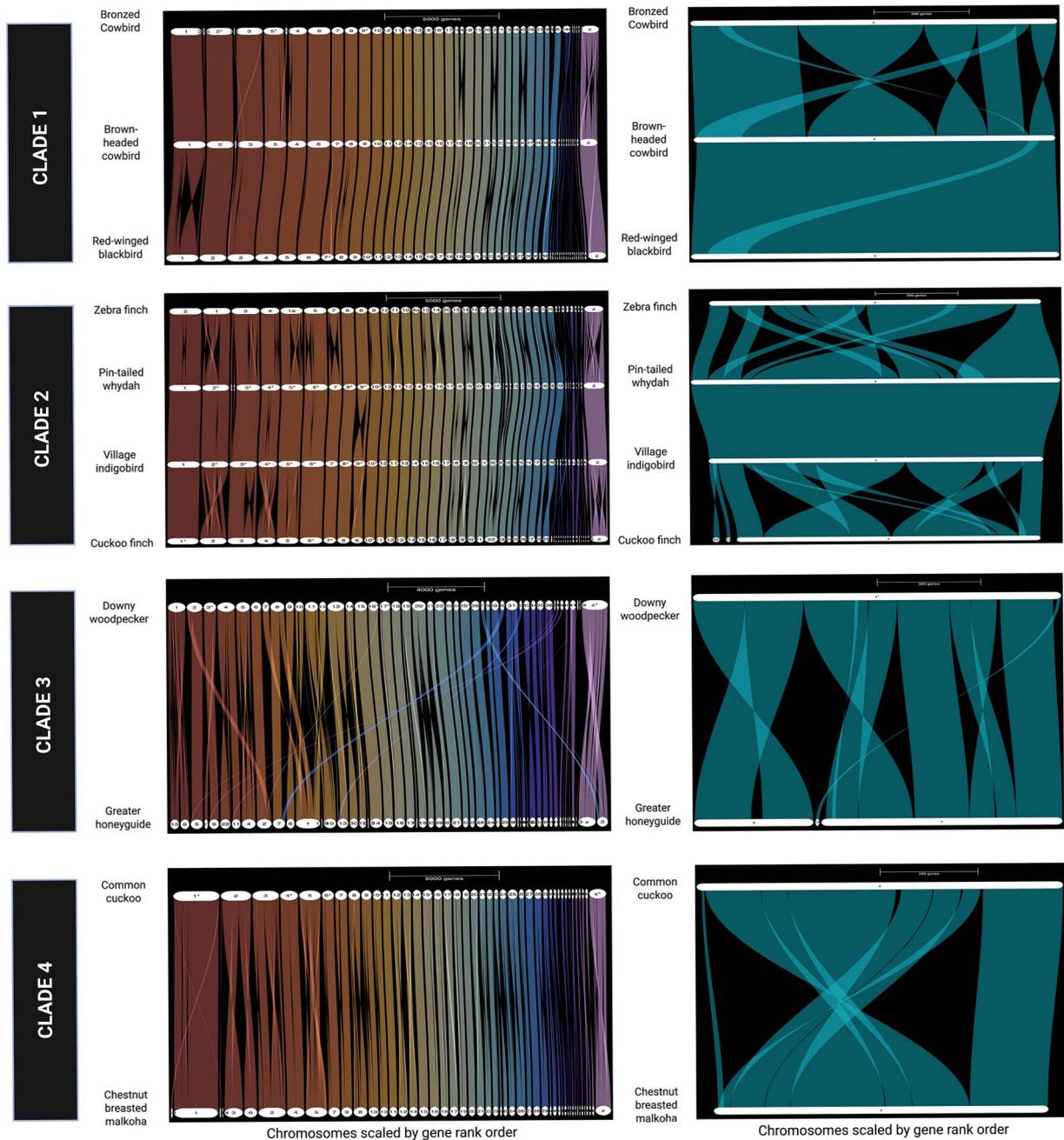


Fig. 1. Chromosomal synteny between avian brood parasite and parental outgroup species. An asterisk (*) denotes chromosomes that were inverted to better visualize conserved synteny across species. Primary syntenic blocks for each designated clade are illustrated. Genomes of the Zebra finch and Common cuckoo were included as outgroup assemblies for clades 2 and 4, respectively. Primary syntenic blocks for the designated clades illustrated. A unified map of orthologous genes across multiple species using RefSeq gene sets for each was used to detect conserved gene order. Plots depict a horizontal arrangement of chromosomes scaled by rank gene order. All chromosomes are presented in the left panel and only Z chromosomes in the right, with comparisons made within the different clades.

Autosomal synteny

Based on an analysis framework of conserved gene order and single-copy orthologs use, we show a substantial level of shared chromosomal synteny within selected phylogenetic clades, but notable species pairwise species differences as well, providing first approximations of how chromosome order and orientation changed over varied spans of avian evolutionary divergence (Fig. 1). The autosomal alignments for piciforms, comprising downy woodpecker and greater honeyguide overall displayed the most extensive number

of rearrangements ('Clade 3' in Fig. 1). For example, the downy woodpecker chromosome 1 corresponds to greater honeyguide chromosomes 8 and 15. Because the greater honeyguide assembly was produced from lower-quality CLR data (Supplemental Table 1), we cannot rule out an assembly artifact as an explanation for this result, although we note that woodpecker genomes are reported to have an unusually high number of rearrangements for birds (de Oliveira et al. 2017; Hruska and Manthey 2021; Forest et al. 2024). In contrast, most other pairwise alignments by clade display relatively

high levels of chromosomal synteny (Fig. 1), interspersed with some inversions between closely related species. We note that additional orthogonal evidence will be needed to confidently establish whether the breaks in synteny that we observed in this study represent assembly error or natural structural variation (Fig. 1).

Sex chromosomes assembly and synteny

Given evidence for maternal inheritance of egg phenotypes (Spottiswoode et al. 2011) phenotypes (Spottiswoode and Stevens 2011; Fossoy et al. 2016; Spottiswoode et al. 2022; Merondun et al. 2024), we evaluated the Z and W chromosomes for possible sources of genetic variation relevant to host-specific adaptation (Fig. 1). The assembled sizes of Z chromosomes ranged from 70 to 84 Mb (Supplemental Table 4), with one notable exception: the downy woodpecker, which was 140 Mb. The downy woodpecker Z chromosome largely aligns with the honeyguide Z chromosome, but a substantial portion of its sequence aligns with chromosome 3, suggesting a possible autosomal translocation (Fig. 1). Autosomal translocations or fusions involving an autosome, and the sex chromosomes, yielding ‘neo-sex chromosomes’, have been reported in songbirds (Sigeman et al. 2021) (Sigeman et al. 2022) and parrots (Huang et al. 2022), though these events remain rare among avian species. Our proximity mapping analysis strongly supports a translocation or fusion of a portion of an autosome to the sex chromosomes in the history of woodpeckers, the precise timing of which is uncertain without comparative data from other species (Supplemental Fig. 1). The assembled size of the downy woodpecker Z chromosome in this study is consistent with woodpecker karyotypes, which include an enlarged Z chromosome Shields et al. 1982. The mapping of whole genome resequencing data for male and female downy woodpeckers also provides support for a Z-autosome fusion (Supplemental Figs 2 and 3). Relative sequencing depth for males is consistent across the full length of the Z-chromosome assembly, whereas relative depth for females increases from 0.5 to 1 at the putative junction between the ancestral Z-chromosome and former autosome (Supplemental Fig. 2). This suggests that the autosomal segment fused to both the Z and W chromosomes and that divergence between the neo-Z and neo-W for the formerly autosomal segment remains sufficiently low that both W-linked and Z-linked sequences in females align to the Z-chromosome. This also assumes that the formerly autosomal portion of the neo-W is not in the genome assembly, such that females appear to be diploid for this region. To the extent that recombination suppression has evolved, divergence between the neo-Z and neo-W chromosomes should generate higher estimates of apparent heterozygosity in females than in males, but we found no evidence to support this (Supplemental Fig. 3), suggesting that the formerly autosomal portions of the neo-Z and neo-W remain pseudoautosomal in downy woodpeckers. Our analysis of other clades detected multiple Z chromosome inversions of potential interest (Fig. 1). As noted earlier for autosomes, apparent Z chromosomes structural rearrangements will also require further experimental validation prior to interpreting their evolutionary significance.

W-chromosome assemblies ranged in size from 3 to 12 Mb, excluding the greater honeyguide W, underscoring the persistent difficulty of fully resolving sex-limited chromosomes (W or Y), even with high-accuracy long-read sequencing.

Avian W chromosomes, like mammal Ys, are enriched in repeat content, structure, and degeneration (Huang et al. 2023; Benham et al. 2024), and both the Y and W represent classic examples of sex chromosome specialization and degenerative evolution, leading to compact, repeat-rich, gene-poor chromosomes shaped by similar evolutionary pressures (Bellott et al. 2017). Even the chicken W chromosome, the most extensively studied in birds remains incomplete despite the application of multiple long-read technologies (Huang et al. 2023). Notably, the greater honeyguide W chromosome was particularly fragmented, with the longest W-linked contig spanning only 0.9 Mb (Supplemental Table 4). This likely reflects both high assembly fragmentation due to the more error-prone CLR data for this genome (Supplemental Table 4), and as a result lower corroboration of mapping evidence. We mostly failed to align greater honeyguide scaffolds at sufficient nucleotide similarity thresholds to the downy woodpecker W assembly, the most closely available genome (Supplemental Table 4). For all avian W chromosomes it will be necessary to implement recent technological developments that leverage complementary long-read platforms, such as PacBio HiFi and Oxford Nanopore ultra-long reads, alongside the use of evolving algorithms like Verkko (Rautiainen et al. 2023). This approach has resulted in generating gap-free assemblies of structurally complex sex chromosomes such as the complete human Y chromosome (Rhie et al. 2023), a result relevant for future research on the avian W.

Gene representation and annotation of chromosome assemblies

Benchmark Universal Single Copy Orthologue (BUSCO) analysis (Simao et al. 2015) revealed all chromosome-level assemblies to have a completeness score >97% (Supplemental Table 5). Importantly, the percentage of missing genes exceeded 2.6% for the greater honeyguide, whereas all others were below 1.6% (Supplemental Table 5). The number of predicted protein-coding genes using the NCBI RefSeq Annotation pipeline (O’Leary et al. 2016) per species ranged from 14 393 to 16 797 (Table 1). Estimates of noncoding gene classifications varied significantly across species (Supplemental Table 6), primarily reflecting the availability of species-specific RNAseq data. The number of long non-coding RNAs (lncRNAs) ranged from 45 to 10 129 (Supplemental Table 6). Notably, the greater honeyguide genome is predicted to have only 45 lncRNAs, likely due to its limited RNA-seq data, which was derived from one blood sample with 99 million reads. In contrast, the cuckoo finch genome was annotated using 10 billion RNAseq reads with greater tissue diversity, resulting in 10 129 predicted lncRNAs. This example highlights the need for more extensive RNA sequencing of diverse tissue sources to further improve genome annotations, both for the species characterized here and more generally.

Gene orthology

Understanding the genetic underpinnings of adaptations linked to the evolution of brood parasitism and the coevolution of brood parasites with host species, depends on a robust gene orthology framework to enable precise investigations of genotype-to-phenotype relationships. Orthologs are genes that diverged at a speciation event, whereas paralogs are genes that arise from gene duplication events. Using

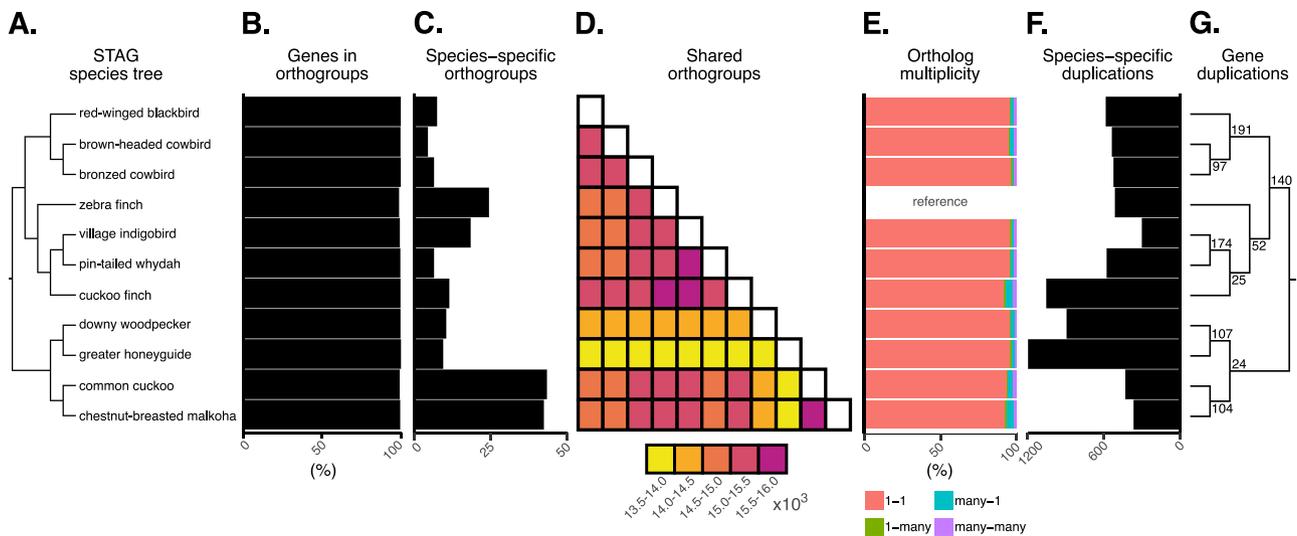


Fig. 2. Summary of the OrthoFinder results for an analysis of genomes from obligate brood parasites and closely related nesting species. (A) The species tree inferred by STAG and rooted by STRIDE. (B) the percentage of genes from each species assigned to orthogroups. (C) The number of species-specific orthogroups. (D) Heat map of the number of orthogroups including each species pair. (E) Ortholog multiplicities for each species using zebra finch as the reference. (F) The number of gene duplication events on each terminal branch of the species tree. (G) The number of duplications on each branch of the species tree that are retained in all descendant species.

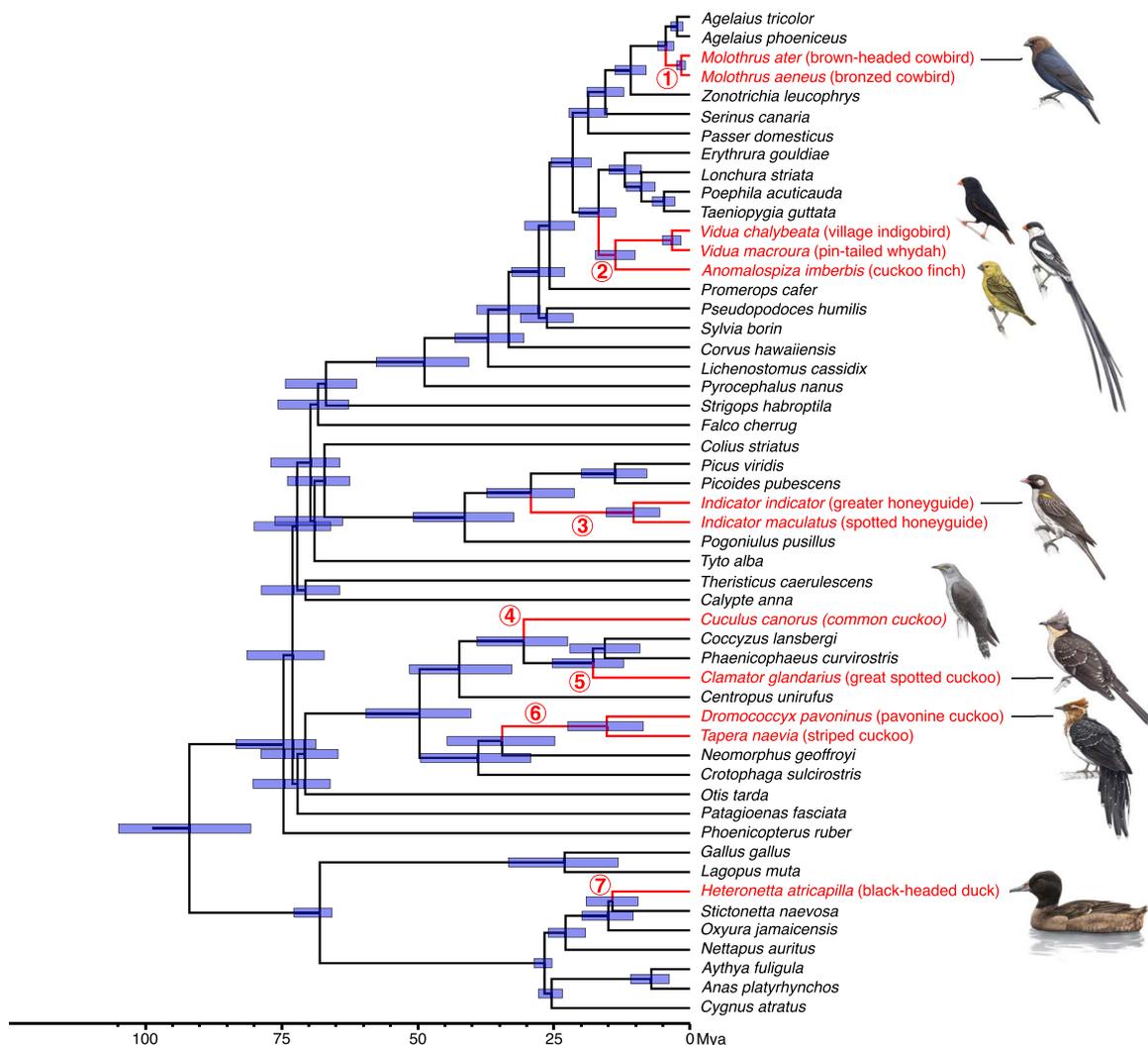


Fig. 3. Evolutionary origins of obligate avian brood parasitism in birds. Obligate brood parasites are labeled in red. the seven independent origins of brood parasitism are numbered. drawings of selected brood parasitic species by Javier Lazaro.

OrthoFinder, we inferred orthogroups, identified species-specific orthogroups, analyzed gene orthogroup multiplicity, and estimated gene duplications at various evolutionary nodes (Fig. 2A–G). The fraction of genes in orthogroups was similar across species (Fig. 2B), although the number of species-specific orthogroups and shared orthogroups were not. The common cuckoo and chestnut-breasted malkoha exhibited the highest number of species-specific orthogroups, defined as sets of genes descended from a single ancestral gene in their last common ancestor (Fig. 2C). Meanwhile, the greater honeyguide appears to have the smallest number of shared orthogroups compared with other species (Fig. 2D). Overall, when using zebra finch as a reference, we observe a high number of shared one-to-one orthologs across all species (Fig. 2E). Among the six parasitic species compared the greater honeyguide followed by cuckoo finch had the greatest number of species-specific gene duplications (Fig. 2F). The smaller number of shared orthogroups and higher number of gene duplications in greater honeyguide is likely due to misannotated genes and genes with split annotations. The number of duplications on each branch of the species tree that are retained in all descendant species was highest for the passeriform clade, comprising cowbirds, village indigobird, and cuckoo finch (Fig. 2G). Overall, our results are in line with other studies showing that greater evolutionary divergence between species will correlates with a reduced number of one-to-one gene orthologs (Emms and Kelly 2019).

Evolutionary origins of brood parasitism in phylogenetic context

Our phylogenomic analysis using concatenated UCE loci resulted in six independent runs that converged in similar parameter space; these runs were combined in a posterior sample of 10 000 trees. Family and higher-level relationships were constrained, guided by the results of Stiller et al. (2024). Our improved sampling of brood parasites provides a single, robust analysis for comparing the timing of the seven independent origins of obligate brood parasitic behavior (Fig. 3). The most recent evolution of brood parasitism was in cowbirds (*Molothrus*), with point estimates of minimum and maximum ages at 1.6 and 4.5 mya, respectively, with the minimum estimate taken from the common ancestor of brown-headed and bronze cowbird and the maximum estimate taken from the common ancestor of cowbirds and the most closely related parental species (i.e. the two ends of the ancestral cowbird lineage). Applying the same logic to estimates for the origin of brood parasitism in parasitic finches (*Anomalospiza-Vidua*; 13.7–16.8 mya), honeyguides (*Indicator*; 10.4–29.2 mya), and the *Dromococcyx-Tapera* clade of cuckoos (15.3–34.5 mya) indicates comparatively older origins of parasitism. With sampling of only a single species, minimum ages could not be estimated for the black-headed duck (*Heteronetta*; max = 14.3 mya), *Calamator* cuckoos (max = 17.8 mya), or the large clade of parasitic cuckoos that includes the common cuckoo (*C. canorus*; max = 30.5 mya). This topology with branch lengths provides an approximate timeframe for future comparative analyses of parasitic birds.

Avian brood parasitism, defined by the deceptively simple act of laying an egg in the nest of an appropriate host species, has figured prominently in evolutionary biology, providing a fascinating diversity of coevolved adaptations in interacting

hosts and parasites. This research has largely been behavioral and ecological with few tests of the genetic basis of the unique adaptations of brood parasites, including their loss of parental care behaviors. This consideration figured prominently in the BPBGC's development of the genomic resources described here. Our broader aim was to fill important taxonomic gaps in the availability of reference genomes to facilitate genomic analyses this rare but widespread reproductive strategy. In doing so, we highlight key genomic features, including chromosomal synteny and gene orthology, and provide a phylogenetic timescale for the evolution of parasitic behavior in each clade. By reporting measures of assembly quality, focusing on contiguity, synteny, and gene representation evaluations, we encourage users to select assemblies described herein based on their specific computational objectives. These genomic resources will enable more detailed testing of diverse hypotheses about the origins of brood parasitism and adaptations of parasitic birds.

Acknowledgments

For their assistance in collecting or providing the samples needed for genome assembly and annotation, we thank Eric Schuppe, Matt Louder, Louisiana State University Museum of Natural Science, Museum of Southwestern Biology, Texas Cooperative Wildlife Collection, Beckman Center for Conservation Research at the San Diego Zoo, Sylvan Heights Bird Park, Museum of Comparative Zoology (Harvard), and the many additional individuals and institutions that supported or assisted in past fieldwork and who contributed samples to museum collections or assisted directly or indirectly in other ways. We thank Maggs X for recommendations on the parameterization of Orthofinder. We thank the members of the Hudson Alpha Institute sequencing core, especially Jane Grimwood, for long read sequencing data generation. Finally, we thank various members of the VGP consortium including Giulio Formenti, Olivier Fedrigo and Bonhwang Koo for assistance during the sequencing and assembly of the brown-headed cowbird, village indigobird, and downy woodpecker. We thank Melissa Rowe, Judith Risse, Simon Grifith and Daniel Hooper for early access to a *Poephila acuticauda* genome assembly. The work of DH, FTN, and PM was supported by the National Center for Biotechnology Information of the National Library of Medicine (NLM), National Institutes of Health.

Author contributions

Rachel A. Carroll (Data curation, Formal analysis, Writing—review & editing), Edward Rice (Data curation, Formal analysis, Resources, Writing—review & editing), Ladeana W. Hillier (Data curation, Formal analysis), Jeffrey M. Dacosta (Conceptualization, Data curation, Formal analysis, Funding acquisition, Visualization, Writing—original draft, Writing—review & editing), Ekaterina Osipova (Data curation, Formal analysis, Writing—review & editing), Sara Smith (Data curation, Formal analysis, Writing—review & editing), Gabriel A. Jamie (Data curation, Resources, Writing—review & editing), Juan Gabriel Martinez Suarez (Resources, Writing—review & editing), Mercedes Molina-Morales (Resources), Tomas Marques-Bonet (Resources, Writing—review & editing), Diana Haddad (Data curation, Formal analysis, Writing—review & editing), Joseph Manthey (Data curation, Resources, Writing—review & editing), Matthew J. Fuxjager (Data curation, Resources, Writing—review & editing), Kathleen S. Lynch (Resources, Writing—review & editing), Jonathan Wood (Data curation, Formal analysis, Writing—review & editing), Erich D. Jarvis (Data curation, Resources, Writing—review & editing), Patrick Masteron (Data curation, Formal analysis), Francoise Thibaud-Nissen (Data curation, Formal analysis, Writing—review & editing), Mark Hauber (Data curation, Resources, Writing—review & editing), Claire N. Spottiswoode (Data curation, Investigation, Resources, Writing—review & editing), Timothy B. Sackton (Conceptualization, Data curation,

Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing—original draft, Writing—review & editing), Christopher Balakrishnan (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing—original draft, Writing—review & editing), Michael D. Sorenson (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing—original draft, Writing—review & editing), and Wesley C. Warren (Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing—original draft, Writing—review & editing)

Supplementary material

Supplementary material is available at *Journal of Heredity* online.

Funding

This project was supported by National Science Foundation collaborative research grants to WCW, MDS, CNB, MEH, TBS, and JMD (NSF DBI 1754311, 1754397, 1754406, 1754546, 1754643, and 1940624 and IOS 1456524 and 1456612).

Data availability

We have deposited the primary data underlying these analyses as follows: chromosome-level or contig-only assemblies were deposited at Genbank under various accession numbers summarized in [Supplemental Table 1](#). All raw sequence data were deposited in the NIH's sequence read archive under the appropriate species Bioproject identification number ([Supplemental Table 1](#)). The WGS data used to evaluate Downy woodpecker Z chromosome is found in GenBank (SRR33614311–20).

References

- Bellott DW, Skaletsky H, Cho TJ, Brown L, Locke D, Chen N, Galkina S, Pyntikova T, Koutseva N, Graves T, et al. Avian W and mammalian Y chromosomes convergently retained dosage-sensitive regulators. *Nat Genet.* 2017;49:387–394. <https://doi.org/10.1038/ng.3778>
- Benham PM, Cicero C, Escalona M, Beraut E, Fairbairn C, Marimuthu MPA, Nguyen O, Sahasrabudhe R, King BL, Thomas WK, et al. Remarkably high repeat content in the genomes of sparrows: the importance of genome assembly completeness for transposable element discovery. *Genome Biol Evol.* 2024;16. <https://doi.org/10.1093/gbe/evae067>
- Bionano. *Bionano solve theory of operation: variant annotation pipeline.* 2019.
- Biosciences P. *FALCON – genome assembly tool for PacBio long reads.* 2019: GitHub repository.
- Birkhead TR, Hemmings N, Spottiswoode CN, Mikulica O, Moskat C, Ban M, Schulze-Hagen K. Internal incubation and early hatching in brood parasitic birds. *Proc Biol Sci.* 2011;278:1019–1024. <https://doi.org/10.1098/rspb.2010.1504>
- Bouckaert R, Vaughan TG, Barido-Sottani J, Duchene S, Fourment M, Gavryushkina A, Heled J, Jones G, Kuhner D, De Maio N, et al. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Comput Biol.* 2019;15:e1006650. <https://doi.org/10.1371/journal.pcbi.1006650>
- Cheng H, Concepcion GT, Feng X, Zhang H, Li H. Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods.* 2021;18:170–175. <https://doi.org/10.1038/s41592-020-01056-5>
- Cole TL, Zhou C, Fang M, Pan H, Ksepka DT, Fiddaman SR, Emerling CA, Thomas DB, Bi X, Fang Q, et al. Genomic insights into the secondary aquatic transition of penguins. *Nat Commun.* 2022;13:3912. <https://doi.org/10.1038/s41467-022-31508-9>
- Davies N. *Cuckoos, cowbirds and other cheats.* T. & A.D. Poyser; 2000.
- de Oliveira TD, Kretschmer R, Bertocchi NA, Degrandi TM, de Oliveira EH, Cioffi MB, Garner AD, Gunski RJ. Genomic Organization of Repetitive DNA in woodpeckers (Aves, Piciformes): implications for karyotype and ZW sex chromosome differentiation. *PLoS One.* 2017;12:e0169987. <https://doi.org/10.1371/journal.pone.0169987>
- Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, Aiden EL. Juicer provides a one-click system for Analyzing loop-resolution hi-C experiments. *Cell Syst.* 2016;3:95–98. <https://doi.org/10.1016/j.cels.2016.07.002>
- Eliason CM, Mellenthin LE, Hains T, McCullough JM, Pirro S, Andersen MJ, Hackett SJ. Genomic signatures of convergent shifts to plunge-diving behavior in birds. *Commun Biol.* 2023;6:1011. <https://doi.org/10.1038/s42003-023-05359-z>
- Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol.* 2019;20:238. <https://doi.org/10.1186/s13059-019-1832-y>
- Faircloth BC. PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics.* 2016;32:786–788. <https://doi.org/10.1093/bioinformatics/btv646>
- Faircloth BC, McCormack JE, Crawford NG, Harvey MG, Brumfield RT, Glenn TC. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Syst Biol.* 2012;61:717–726. <https://doi.org/10.1093/sysbio/sys004>
- Feeney WE, Riehl C. Monogamy without parental care? Social and genetic mating systems of avian brood parasites. *Philos Trans R Soc Lond Ser B Biol Sci.* 2019;374:20180201. <https://doi.org/10.1098/rstb.2018.0201>
- Foote AD, Liu Y, Thomas GW, Vinar T, Alfoldi J, Deng J, Dugan S, van Elk CE, Hunter ME, Joshi V, et al. Convergent evolution of the genomes of marine mammals. *Nat Genet.* 2015;47:272–275. <https://doi.org/10.1038/ng.3198>
- Forest T, Achaz G, Marbouty M, Bignaud A, Thierry A, Koszul R, Milhes M, Lledo J, Pons JM, Fuchs J. Chromosome-level genome assembly of the European green woodpecker *Picus viridis*. *G3 (Bethesda).* 2024;14:
- Fossoy F, Sorenson MD, Liang W, Ekrem T, Moksnes A, Moller AP, Rutila J, Roskaft E, Takasu F, Yang C, et al. Ancient origin and maternal inheritance of blue cuckoo eggs. *Nat Commun.* 2016;7:10272
- Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. 2012.
- Genomics., x. *Long ranger – software suite for linked-read analysis.* 2018.
- Ghurye J, Rhie A, Walenz BP, Schmitt A, Selvaraj S, Pop M, Phillippy AM, Koren S. Integrating hi-C links with assembly graphs for chromosome-scale assembly. *PLoS Comput Biol.* 2019;15:e1007273. <https://doi.org/10.1371/journal.pcbi.1007273>
- Guigueno MF, Snow DA, MacDougall-Shackleton SA, Sherry DF. Female cowbirds have more accurate spatial memory than males. *Biol Lett.* 2014;10:20140026. <https://doi.org/10.1098/rsbl.2014.0026>
- Hruska JP, Manthey JD. De novo assembly of a chromosome-scale reference genome for the northern flicker *Colaptes auratus*. *G3 (Bethesda).* 2021;11:
- Hu Y, Wu Q, Ma S, Ma T, Shan L, Wang X, Nie Y, Ning Z, Yan L, Xiu Y, et al. Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. *Proc Natl Acad Sci USA.* 2017;114:1081–1086. <https://doi.org/10.1073/pnas.1613870114>
- Huang Z, De OFI, Liu J, Peona V, Gomes AJB, Cen W, Huang H, Zhang Y, Chen D, Xue T, et al. Recurrent chromosome reshuffling and the evolution of neo-sex chromosomes in parrots. *Nat Commun.* 2022;13:944. <https://doi.org/10.1038/s41467-022-28585-1>
- Huang Z, Xu Z, Bai H, Huang Y, Kang N, Ding X, Liu J, Luo H, Yang C, Chen W, et al. Evolutionary analysis of a complete chicken

- genome. *Proc Natl Acad Sci USA*. 2023;120:e2216641120. <https://doi.org/10.1073/pnas.2216641120>
- Institute, W. S. *Scaff10X*. 2020.
- Kapustin Y, Souvorov A, Tatusova T, Lipman D. Splign: algorithms for computing spliced alignments with identification of paralogs. *Biol Direct*. 2008;3:20. <https://doi.org/10.1186/1745-6150-3-20>
- Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol*. 2019;37:540–546. <https://doi.org/10.1038/s41587-019-0072-8>
- Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res*. 2017;27:722–736. <https://doi.org/10.1101/gr.215087.116>
- Korneliussen TS, Albrechtsen A, Nielsen R. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*. 2014;15. <https://doi.org/10.1186/s12859-014-0356-4>
- Langmore NE, Grealy A, Noh HJ, Medina I, Skeels A, Grant J, Murray KD, Kilner RM, Holleley CE. Coevolution with hosts underpins speciation in brood-parasitic cuckoos. *Science*. 2024;384:1030–1036. <https://doi.org/10.1126/science.adj3210>
- Langmore NE, Hunt S, Kilner RM. Escalation of a coevolutionary arms race through host rejection of brood parasitic young. *Nature*. 2003;422:157–160. <https://doi.org/10.1038/nature01460>
- Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. 2012;9:357–359. <https://doi.org/10.1038/nmeth.1923>
- Langmore NE, Stevens M, Maurer G, Heinsohn R, Hall ML, Peters A, Kilner RM. Visual mimicry of host nestlings by cuckoos. *Proc Biol Sci*. 2011;278:2455–2463. <https://doi.org/10.1098/rspb.2010.2391>
- Li H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*. 2018;34:3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>
- Louder MIM, Balakrishnan CN, Louder AAN, Driver RJ, London SE, Hauber ME. An acoustic password enhances auditory learning in juvenile brood parasitic cowbirds. *Curr Biol*. 2019;29:4045–4051.e3. <https://doi.org/10.1016/j.cub.2019.09.046>
- Lovell JT, Sreedasyam A, Schranz ME, Wilson M, Carlson JW, Harkess A, Emms D, Goodstein DM, Schmutz J. GENESPACE tracks regions of interest and gene copy number variation across multiple genomes. *Elife*. 2022;11. <https://doi.org/10.7554/eLife.78526>
- Lund J, Dixit T, Attwood MC, Hamama S, Moya C, Stevens M, Jamie GA, Spottiswoode CN. When perfection isn't enough: host egg signatures are an effective defence against high-fidelity African cuckoo mimicry. *Proc Biol Sci*. 2023;290:20231125. <https://doi.org/10.1098/rspb.2023.1125>
- Lyu T, Zhou S, Fang J, Wang L, Shi L, Dong Y, Zhang H. Convergent genomic signatures of high-altitude adaptation among six independently evolved mammals. *Animals (Basel)*. 2022;12. <https://doi.org/10.3390/ani12243572>
- Merondun J, Marques CI, Andrade P, Meshcheryagina S, Galvan I, Afonso S, Alves JM, Araujo PM, Bachurin G, Balacco J, et al. Evolution and genetic architecture of sex-limited polymorphism in cuckoos. *Sci Adv*. 2024;10:ead15255. <https://doi.org/10.1126/sciadv.adl5255>
- Moran RL, Richards EJ, Ornelas-Garcia CP, Gross JB, Donny A, Wiese J, Keene AC, Kowalko JE, Rohner N, McGaugh SE. Selection-driven trait loss in independently evolved cavefish populations. *Nat Commun*. 2023;14:2557. <https://doi.org/10.1038/s41467-023-37909-8>
- Morgulis A, Gertz EM, Schaffer AA, Agarwala R. WindowMasker: window-based masker for sequenced genomes. *Bioinformatics*. 2006;22:134–141. <https://doi.org/10.1093/bioinformatics/bti774>
- Nair-Roberts RG, Erichsen JT, Reboreda JC, Kacelnik A. Distribution of substance P reveals a novel subdivision in the hippocampus of parasitic south American cowbirds. *J Comp Neurol*. 2006;496:610–626. <https://doi.org/10.1002/cne.20915>
- O'Leary NA, Wright MW, Brister JR, Ciufu S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res*. 2016;44:D733–D745. <https://doi.org/10.1093/nar/gkv1189>
- Pacific Biosciences of California, I. *SMRT link: arrow consensus algorithm*. 2019.
- Parker J, Tsakoggeorga G, Cotton JA, Liu Y, Provero P, Stupka E, Rossiter SJ. Genome-wide signatures of convergent evolution in echolocating mammals. *Nature*. 2013;502:228–231. <https://doi.org/10.1038/nature12511>
- Payne R. The evolution of clutch size and reproductive rates in cuckoos. *Evolution*. 1974;28:169–181. <https://doi.org/10.1111/j.1558-5646.1974.tb00738.x>
- Payne R. Clutch size, egg size, and the consequences of single vs. multiple parasitism in parasitic finches. *Ecology*. 1976;58:500–513
- Payne R. The ecology of brood parasitism in birds. *Ann Rev Ecol Evol Syst*. 1977;8:1–28. <https://doi.org/10.1146/annurev.e.s.08.110177.000245>
- Payne RB, Payne LL, Woods JL. Song learning in brood-parasitic indigobirds *Vidua chalybeata*: song mimicry of the host species. *Anim Behav*. 1998;55:1537–1553. <https://doi.org/10.1006/anbe.1997.0701>
- Pedersen BS, Quinlan AR. Mosdepth: quick coverage calculation for genomes and exomes. *Bioinformatics*. 2017;34:867–868. <https://doi.org/10.1093/bioinformatics/btx699>
- Rautiainen M, Nurk S, Walenz BP, Logsdon GA, Porubsky D, Rhie A, Eichler EE, Phillippy AM, Koren S. Telomere-to-telomere assembly of diploid chromosomes with Verkko. *Nat Biotechnol*. 2023;41:1474–1482. <https://doi.org/10.1038/s41587-023-01662-6>
- Reboreda JC, Clayton NS, Kacelnik A. Species and sex differences in hippocampus size in parasitic and non-parasitic cowbirds. *Neuroreport*. 1996;7:505–508. <https://doi.org/10.1097/00001756-199601310-00031>
- Rhie A, McCarthy SA, Fedrigo O, Damas J, Formenti G, Koren S, Uliano-Silva M, Chow W, Funghamman A, Kim J, et al. Towards complete and error-free genome assemblies of all vertebrate species. *Nature*. 2021;592:737–746. <https://doi.org/10.1038/s41586-021-03451-0>
- Rhie A, Nurk S, Cechova M, Hoyt SJ, Taylor DJ, Altemose N, Hook PW, Koren S, Rautiainen M, Alexandrov IA, et al. The complete sequence of a human Y chromosome. *Nature*. 2023;621:344–354. <https://doi.org/10.1038/s41586-023-06457-y>
- Ricemeyer E. *Assembly error correction workflow*. 2025: Nextflow pipeline.
- Ricemeyer ES, Carroll RA, Warren WC. Agptools: a utility suite for editing genome assemblies. 2025.
- Roach MJ, Schmidt SA, Borneman AR. Purge Haplotigs: allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinform*. 2018;19:460. <https://doi.org/10.1186/s12859-018-2485-7>
- Robinson JT, Turner D, Durand NC, Thorvaldsdottir H, Mesirov JP, Aiden EL. Juicebox.js provides a cloud-based visualization system for hi-C data. *Cell Syst*. 2018;6:256–258.e1. <https://doi.org/10.1016/j.cels.2018.01.001>
- Ronka K, Eroukhanoff F, Kulmuni J, Nouhaud P, Thorogood R. Beyond genes-for-behaviour: the potential for genomics to resolve long-standing questions in avian brood parasitism. *Ecol Evol*. 2024;14:e70335. <https://doi.org/10.1002/ece3.70335>
- Sackton TB, Grayson P, Cloutier A, Hu Z, Liu JS, Wheeler NE, Gardner PP, Clarke JA, Baker AJ, Clamp M, et al. Convergent regulatory evolution and loss of flight in paleognathous birds. *Science*. 2019;364:74–78. <https://doi.org/10.1126/science.aat7244>
- Sadanandan KR, Ko MC, Low GW, Gahr M, Edwards SV, Hiller M, Sackton TB, Rheindt FE, Sin SYW, Baldwin MW. Convergence in hearing-related genes between echolocating birds and mammals. *Proc Natl Acad Sci USA*. 2023;120:e2307340120. <https://doi.org/10.1073/pnas.2307340120>
- Sherry DF, Forbes MR, Khurgel M, Ivy GO. Females have a larger hippocampus than males in the brood-parasitic brown-headed cowbird. *Proc Natl Acad Sci USA*. 1993;90:7839–7843. <https://doi.org/10.1073/pnas.90.16.7839>

- Shields Gerald F, Jarrell Gordon, H, Redrupp Elizabeth. Enlarged Sex Chromosomes of Woodpeckers (piciformes). *The Auk*. 1982;99:28.
- Sigeman H, Strandh M, Proux-Wera E, Kutschera VE, Ponnikas S, Zhang H, Lundberg M, Soler L, Bunikis I, Tarka M, et al. Avian neo-sex chromosomes reveal dynamics of recombination suppression and W degeneration. *Mol Biol Evol*. 2021;38:5275–5291. <https://doi.org/10.1093/molbev/msab277>
- Sigeman H, Zhang H, Ali Abed S, Hansson B. A novel neo-sex chromosome in *Sylvietta brachyura* (Macrosphenidae) adds to the extraordinary avian sex chromosome diversity among Sylvioidea songbirds. *J Evol Biol*. 2022;35:1797–1805. <https://doi.org/10.1111/jeb.14096>
- Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*. 2015;31:3210–3212. <https://doi.org/10.1093/bioinformatics/btv351>
- Sorenson MD, Payne RB. Molecular genetic perspectives on avian brood parasitism. *Integr Comp Biol*. 2002;42:388–400. <https://doi.org/10.1093/icb/42.2.388>
- Sorenson MD, Sefc KM, Payne RB. Speciation by host switch in brood parasitic indigobirds. *Nature*. 2003;424:928–931. <https://doi.org/10.1038/nature01863>
- Spottiswoode CN, Stevens M. How to evade a coevolving brood parasite: egg discrimination versus egg variability as host defences. *Proc Biol Sci*. 2011;278:3566–3573. <https://doi.org/10.1098/rspb.2011.0401>
- Spottiswoode CN, Stryjewski KF, Quader S, Colebrook-Robjent JF, Sorenson MD. Ancient host specificity within a single species of brood parasitic bird. *Proc Natl Acad Sci USA*. 2011;108:17738–17742. <https://doi.org/10.1073/pnas.1109630108>
- Spottiswoode CN, Tong W, Jamie GA, Stryjewski KF, DaCosta JM, Kuras ER, Green A, Hamama S, Taylor IG, Moya C, et al. Genetic architecture facilitates then constrains adaptation in a host-parasite coevolutionary arms race. *Proc Natl Acad Sci USA*. 2022;119:e2121752119. <https://doi.org/10.1073/pnas.2121752119>
- Stiller J, Feng S, Chowdhury AA, Rivas-Gonzalez I, Duchene DA, Fang Q, Deng Y, Kozlov A, Stamatakis A, Claramunt S, et al. Complexity of avian evolution revealed by family-level genomes. *Nature*. 2024;629:851–860. <https://doi.org/10.1038/s41586-024-07323-1>
- Toda Y, Ko MC, Liang Q, Miller ET, Rico-Guevara A, Nakagita T, Sakakibara A, Uemura K, Sackton T, Hayakawa T, et al. Early origin of sweet perception in the songbird radiation. *Science*. 2021;373:226–231. <https://doi.org/10.1126/science.abf6505>
- VGPassembly. *Veterebrate genome project assembly pipeline*. 2019: version 1.6.
- Visendi P. De novo assembly of linked reads using supernova 2.0. *Methods Mol Biol*. 2022;2443:233–243. https://doi.org/10.1007/978-1-0716-2067-0_12
- Warrenlab. *purge-haplotigs-nf*. 2022. From <https://github.com/WarrenLab/purge-haplotigs-nf>
- Yang C, Liang W, Cai Y, Shi S, Takasu F, Moller AP, Antonov A, Fossoy F, Moksnes A, Roskaft E, et al. Coevolution in action: disruptive selection on egg colour in an avian brood parasite and its host. *PLoS One*. 2010;5:e10816. <https://doi.org/10.1371/journal.pone.0010816>
- Zhang H, Song L, Wang X, Cheng H, Wang C, Meyer CA, Liu T, Tang M, Aluru S, Yue F, et al. Fast alignment and preprocessing of chromatin profiles with Chromap. *Nat Commun*. 2021;12:6566. <https://doi.org/10.1038/s41467-021-26865-w>
- Zhang L, Zhou X, Weng Z, Sidow A. Assessment of human diploid genome assembly with 10x linked-reads data. *Gigascience*. 2019;8. <https://doi.org/10.1093/gigascience/giz141>
- Zhou C, McCarthy SA, Durbin R. YaHS: yet another hi-C scaffolding tool. *Bioinformatics*. 2023;39. <https://doi.org/10.1093/bioinformatics/btac808>
- Zou D, Huang S, Tian S, Kilunda FK, Murphy RW, Dahn HA, Zhou Y, Lee PS, Chen JM. Comparative genomics sheds new light on the convergent evolution of infrared vision in snakes. *Proc Biol Sci*. 2024;291:20240818. <https://doi.org/10.1098/rspb.2024.0818>