Received: 28 December 2022 | Revised: 15 March 2023 | Accepted: 19 June 2023

DOI: 10.1111/mec.17080

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Mitonuclear discordance results from incomplete lineage sorting, with no detectable evidence for gene flow, in a rapid radiation of *Todiramphus* kingfishers

Devon A. DeRaad[1] | Jenna M. McCullough[2] | Lucas H. DeCicco[1] | Paul M. Hime[1] | Leo Joseph[3] | Michael J. Andersen[2] | Robert G. Moyle[1]

[1]Biodiversity Institute and Natural History Museum, University of Kansas, Lawrence, Kansas, USA

[2]Department of Biology and Museum of Southwestern Biology, University of New Mexico, Albuquerque, New Mexico, USA

[3]Australian National Wildlife Collection, CSIRO National Research Collections Australia, Canberra, Australian Capital Territory, Australia

Correspondence
Devon A. DeRaad, Biodiversity Institute and Natural History Museum, University of Kansas, Lawrence, KS 66045, USA.
Email: devonderaad@gmail.com

Present address
Paul M. Hime, McDonnell Genome Institute, Department of Genetics, Washington University School of Medicine in St. Louis, St. Louis, Missouri, USA

Funding information
Division of Environmental Biology, Grant/Award Number: 1557051, 1557053 and 2112467; National Institutes of Health, Grant/Award Number: 5P20GM103638

Handling Editor: Philine Feulner

Abstract

Many organisms possess multiple discrete genomes (i.e. nuclear and organellar), which are inherited separately and may have unique and even conflicting evolutionary histories. Phylogenetic reconstructions from these discrete genomes can yield different patterns of relatedness, a phenomenon known as cytonuclear discordance. In many animals, mitonuclear discordance (i.e. discordant evolutionary histories between the nuclear and mitochondrial genomes) has been widely documented, but its causes are often considered idiosyncratic and inscrutable. We show that a case of mitonuclear discordance in *Todiramphus* kingfishers can be explained by extensive genome-wide incomplete lineage sorting (ILS), likely a result of the explosive diversification history of this genus. For these kingfishers, quartet frequencies reveal that the nuclear genome is dominated by discordant topologies, with none of the internal branches in our consensus nuclear tree recovered in >50% of genome-wide gene trees. Meanwhile, a lack of inter-species shared ancestry, non-significant pairwise tests for gene flow, and little evidence for meaningful migration edges between species, leads to the conclusion that gene flow cannot explain the mitonuclear discordance we observe. This lack of evidence for gene flow combined with evidence for extensive genome-wide gene tree discordance, a hallmark of ILS, leads us to conclude that the mitonuclear discordance we observe likely results from ILS, specifically deep coalescence of the mitochondrial genome. Based on this case study, we hypothesize that similar demographic histories in other 'great speciator' taxa across the Indo-Pacific likely predispose these groups to high levels of ILS and high likelihoods of mitonuclear discordance.

KEYWORDS
gene flow, genomics, incomplete lineage sorting, mitonuclear discordance, phylogenetics, speciation

1 | INTRODUCTION

Mitonuclear discordance is a phenomenon in which phylogenetic trees inferred from nuclear DNA (nuDNA) and mitochondrial DNA (mtDNA) show conflicting evolutionary histories (Toews & Brelsford, 2012). With the ever-increasing size and scope of molecular datasets, mitonuclear discordance has now been widely documented and accepted as a common phenomenon. Instances of mitonuclear discordance have been documented across the tree of life in taxa including protozoans (Kato, Cáceres, et al., 2021), algae

(Kao et al., 2022), fungi (Bourret et al., 2018), arachnids (Ivanov et al., 2018), insects (Hinojosa et al., 2019; Wham et al., 2021), crabs (Shahdadi et al., 2021), fishes (Berbel-Filho et al., 2022; Kato, Arakaki, et al., 2021), amphibians (Firneno et al., 2020), snakes (Marshall et al., 2021), mammals (Bailey & Stevison, 2021; Ge et al., 2022; Good et al., 2008) and birds (Andersen et al., 2021; Pons et al., 2014). Yet, despite the well-documented ubiquity of mitonuclear discordance, the evolutionary mechanisms driving this phenomenon are often difficult to parse (Bonnet et al., 2017; Kimball et al., 2021).

This difficulty typically stems from the fact that many factors, including confounded taxonomy, sex-biased dispersal, historical and modern gene flow, mitochondrial capture, and incomplete lineage sorting (ILS) are all known to cause patterns of mitonuclear discordance in vertebrates (Toews & Brelsford, 2012). Furthermore, two or more of these processes can act simultaneously under natural conditions (e.g. Phuong et al., 2017), generating conflicting patterns across the genome and making it difficult to identify causative links between pattern and process. Most cases confidently associating mitonuclear discordance with an underlying evolutionary process involve dramatic discordance caused by mitochondrial capture (Andersen et al., 2021; Joseph et al., 2021; Kearns et al., 2014; McElroy et al., 2020; Mikkelsen & Weir, 2022), directional introgression across a hybrid zone (Del-Rio et al., 2022) or recent hybridization resulting in mismatched mitochondrial and nuclear genomes in a single individual (Joseph et al., 2019; Joseph & Moritz, 1993). In contrast, few studies have identified mitonuclear discordance resulting mainly from ILS (except see Firneno et al., 2020; Wang et al., 2018). This is likely because ILS, the process of gene tree sorting which results in stochastic patterns of relatedness in individual gene trees (Scornavacca & Galtier, 2017), is more difficult to identify than short bursts of nuclear gene flow or mitochondrial capture between distantly related taxa, which tend to create more starkly discordant patterns (e.g. Andersen et al., 2021; Denton et al., 2014; Seixas et al., 2018). In other words, ILS is constantly acting, generating the expectation of a diverse set of gene trees throughout the genome under a broad range of conditions (Degnan & Salter, 2005), making it difficult to pin down the exact role of this historical process in cases of mitonuclear discordance.

In this study, we focus on a subset of eight *Todiramphus* kingfisher species, out of 29 extant species in the genus (Gill and Donsker, 2019). Of these eight focal species, seven form a monophyletic clade according to mtDNA (Andersen et al., 2015); the eighth, *Todiramphus leucopygius*, serves as an outgroup. Past studies using multi-locus datasets and thousands of ultraconserved elements (UCEs; Faircloth et al., 2012) have shown that the genus *Todiramphus* harbors high diversification rates compared to all other kingfishers and their allies (Andersen et al., 2015, 2018; McCullough et al., 2019). The rapid and successive speciation of this group has made accurate reconstruction of evolutionary history a challenge. The genus-wide range of *Todiramphus* kingfishers spans thousands of kilometres from the Red Sea eastward to remote islands in French Polynesia, including many species that are endemic to isolated islands. Of the eight focal taxa we focus on, four species, the Mariana Kingfisher (*T. albicilla*), Colonist Kingfisher (*T. colonus*), Ultramarine Kingfisher (*T. leucopygius*) and Melanesian Kingfisher (*T. tristrami*) are endemic to islands of the Indo-Pacific (Figure 1). Meanwhile, two species inhabit wide swaths of coastal habitats: Beach Kingfisher (*T. saurophagus*) is found along the coasts from the Malukus east to the Solomon Islands and Torresian Kingfisher (*T. sordidus*) is endemic to mangrove forests in Australia. Rounding out the group, the Collared Kingfisher (*T. chloris*) and Sacred Kingfisher (*T. sanctus*) are the two most widespread *Todiramphus* kingfishers, and the latter is one of the few migratory taxa within the genus. These eight focal species are highly divergent in their geographical range sizes and respective ecologies, and these taxa display multiple cases of secondary sympatry despite extraordinarily rapid diversification (Andersen et al., 2015, 2018). Among naturally occurring systems, this extreme diversification rate, combined with a wealth of prior genetic and natural history information, make the *Todiramphus* kingfishers an especially amenable target for case studies investigating the speciation process.

In particular, the eight focal *Todiramphus* species described above are an ideal group for studying the evolution of mitonuclear discordance. Despite concordantly recovering the seven ingroup species (*T. leucopygius* is consistently recovered as an outgroup) as monophyletic, previous phylogenies recovered discordant topologies depending on the type of input data (e.g. mitochondrial Andersen et al., 2015; nuclear McCullough et al., 2019). To advance our understanding of mitonuclear discordance, we must move beyond simply documenting individual cases and towards understanding the evolutionary scenarios that lead to mitonuclear discordance across the tree of life (Toews & Brelsford, 2012). The prior evidence for mitonuclear discordance among these eight *Todiramphus* kingfisher species makes them an ideal group for rigorously investigating the evolutionary conditions that precede mitonuclear discordance, with potential to reveal broadly applicable insights about the origin of this commonly observed pattern.

To investigate the evolutionary history of this clade of eight *Todiramphus* kingfishers, we first assembled *de novo* a reference genome for the Mariana Kingfisher (*T. albicilla*). We then sequenced thousands of genome-wide loci from museum-vouchered samples with associated high-quality tissue samples, using a restriction-enzyme-based approach (RADseq; Rochette et al., 2019). Next, we performed descriptive population genetic analyses to validate the reciprocal monophyly of these eight focal species with range-wide genomic sampling and compared the levels of genetic diversity among them. To confirm previously documented mitonuclear discordance (McCullough et al., 2019), we compared patterns of phylogenetic relatedness using our genome-wide nuclear single nucleotide polymorphism (SNP) dataset and a single mitochondrial genome for each species, mined from publicly available UCE data (McCullough et al., 2019). We then investigated whether evolutionary processes commonly invoked in published cases of mitonuclear discordance, such as gene flow, ILS, and incorrect taxonomy (Toews & Brelsford, 2012), could sufficiently explain the

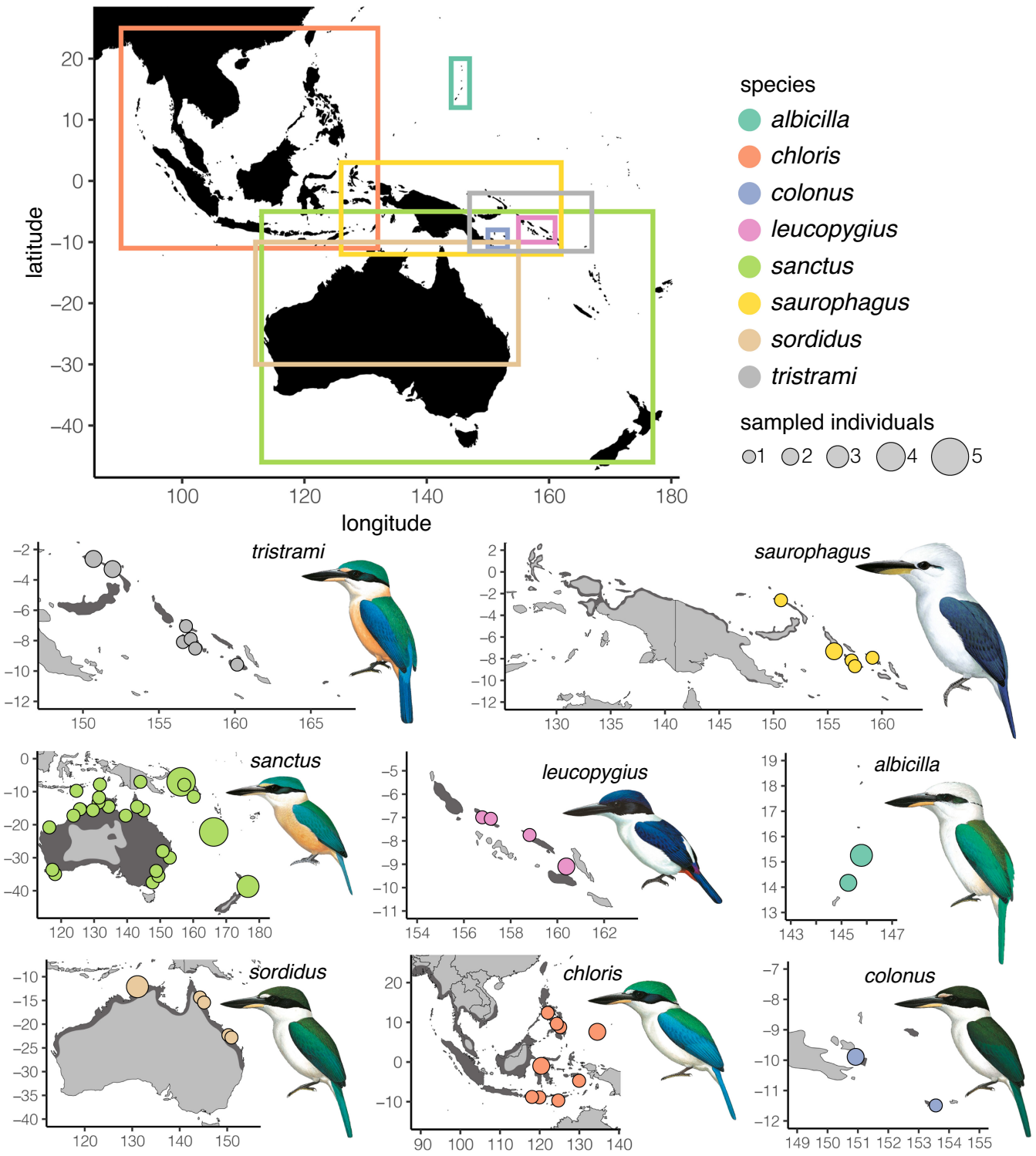

FIGURE 1 Maps showing distributions and sampling for each of the eight focal species. For the migratory taxon *T. sanctus*, only the breeding distribution is shown in darker grey. Range maps adapted from Cornell Lab of Ornithology Birds of the World range maps. We note that the map for *T. chloris* is a subset of the species' whole range, which extends westward to the Red Sea.

topological conflict we observe. Our analyses shed new light on the evolutionary history of this group and suggest that when speciation occurs rapidly and in succession, well-supported topological discordance between the nuclear and mitochondrial genomes can arise via ILS without requiring the invocation of complex, non-tree-like evolutionary processes.

2 | METHODS

2.1 | Documentation and reproducibility

Filtered datasets used as input for all analyses herein, protocols for molecular laboratory benchwork, plus all code for bioinformatic

analyses performed in this study can be found by following the links on the homepage of this GitHub repository: https://github.com/DevonDeRaad/todiramphus.radseq.

2.2 | *De novo* assembly of a reference genome for *Todiramphus albicilla*

We began by generating a reference genome from a wild-caught specimen of *T. albicilla* (KU116001; NCBI BioProject accession PRJNA992488; https://www.ncbi.nlm.nih.gov/bioproject/PRJNA992488). High molecular weight genomic DNA was extracted from muscle tissue at HudsonAlpha Institute for Biotechnology and prepared into a 10X Chromium sequencing library. This library was sequenced with 150 base pair (bp) paired-end reads across two lanes on an Illumina HiSeq X machine, which generated 365,992,456 linked read pairs passing initial quality filters. The resulting sequence data were assembled *de novo* with *Supernova* v2.1 (Weisenfeld et al., 2017). Assembly completeness and contiguity was assessed *QUAST-LG* v5.0.2 (Mikheenko et al., 2018) and completeness was further assessed using *BUSCO* v5.4.4 (Simão et al., 2015) and the avian ortholog dataset 'aves_odb10'.

Because of the strong conservation of genome architecture among birds (Ellegren, 2010) and the contiguity of our *T. albicilla* genome assembly, we were able to use synteny mapping to assign our scaffolds to putative chromosomal positions. We ran *Satsuma* v1 (Grabherr et al., 2010; https://satsuma.sourceforge.net) using one of the haplotype-resolved assemblies output by supernova and the *Taeniopygia guttata* reference genome assembly (version: TaeGut3.2.4, available from: http://ftp.ensembl.org/pub/release-97/fasta/taeniopygia_guttata/dna/) as input with default parameter settings. This synteny mapped assembly, used here for aligning our single-digest restriction-site-associated DNA (RAD) loci, had a final scaffold N_{50} = 81.77 megabases, indicating that most of the synteny mapped assembly was comprised of large, chromosome-sized scaffolds. We note that the version of the genome assembly publicly released via NCBI was not scaffolded based on synteny, and had an additional ~55 kilobases (~0.005% of the overall assembly) removed/masked by the NCBI Foreign Contamination Screening program (https://github.com/ncbi/fcs). For the purposes of reproducibility, we have also made the exact reference genome assembly used in all analyses performed in this manuscript publicly available (https://doi.org/10.5281/zenodo.8178615).

2.3 | DNA extraction, library preparation and sequencing

We extracted genomic DNA from 100 high-quality, ethanol-preserved frozen tissue samples using a manual magnetic bead-based protocol based on Rohland and Reich (2012). These 100 samples included multiple individuals from each of eight *Todiramphus* species (range = 3–30 retained individuals per species in the final dataset; Table S1; Figure 1) sampled from our focal clade of interest, which forms a monophyletic subset within the dozens of species making up the complete *Todiramphus* radiation (Andersen et al., 2015). We also included two technical replicates to ensure compatibility across sequencing runs. Library preparation was performed at the University of Kansas Genomic Sequencing Core, following the protocol outlined in Manthey and Moyle (2015). Briefly, samples were digested with the enzyme *Nde*I, ligated with custom barcodes, size selected for fragments between 500 and 600 bp in length, PCR amplified and bead purified. Libraries were then pooled across projects and sequenced across multiple flow cells on a NextSeq550 Illumina sequencing platform, using the high-output option to generate single-end 100 bp reads.

2.4 | Read mapping and variant calling from RADseq loci

We used the *Stacks* v2.41 (Rochette et al., 2019) pipeline to process our RAD loci and to call SNPs. We began by demultiplexing our raw sequence files using the 'process.radtags' function, discarding low-quality reads (i.e. phred score <10 at any base) and reads with uncalled bases, and rescuing all possible barcodes. We then used *BWA* v0.7.17 (Li & Durbin, 2009) to map the demultiplexed reads from each sample against our reference genome using the 'mem' command with default settings and used *SAMtools* v1.3.1 (Li et al., 2009) to generate sorted .bam files for downstream use. We used this sorted .bam file as input for the 'gstacks' module from *Stacks* to build a catalogue of individual RAD loci. Finally, we used the 'populations' module from *Stacks* to output an unfiltered variant call format (vcf) file, which contained 100 individual samples and 222,393 SNPs.

2.5 | Quality filtering nuclear SNPs

We then used the R packages *SNPfiltR* v1.0.0 (DeRaad, 2022) and *vcfR* v1.12.0 (Knaus & Grünwald, 2017) to visually identify optimal filtering parameters for this specific dataset. We started with a hard filter coding any genotype falling below a minimum read depth of 5 and a minimum genotype quality of 30 as missing. We then filtered for allele balance, recoding any called heterozygous genotype as missing if it fell outside of an allele balance range of 0.25–0.75. We then converted genotypes with more than 100 reads to missing to mitigate spurious SNPs resulting from the potential mapping of multiple highly similar loci from throughout the genome onto a single paralogous locus. We then set a minimum completeness threshold of 95%, meaning that individual samples may not have a missing genotype at more than 95% of called SNPs to be retained. Additionally, we removed three problematic samples that were missing genotypes at most broadly shared SNPs, and the technical replicate (which visually clustered with its identical replicate despite being split across runs) which passed previous filtering protocols, leaving us with 83 unique samples for downstream analyses. We then explored a series

of potential per-SNP completeness thresholds varying from 30% to 100%, visualizing sample clustering at 75%, 80% and 85% SNP completeness thresholds both with and without removing singletons (i.e. SNPs with a minor allele count [MAC] of one). We determined from these preliminary analyses that patterns of clustering among samples were driven mostly by species identity, and that these patterns were not strongly affected by either per-SNP completeness threshold or the inclusion of singleton SNPs. Based on this investigation, we decided to implement the 85% per-SNP completeness threshold and no MAC cut-off, which seemed to maximize the signal-to-noise trade-off based on our preliminary visual assessment. This resulted in a final complete SNP dataset containing 83 samples (Table S1) and 7430 SNPs across 1892 loci, with 7.57% overall missing data (henceforth referred to as 'complete SNP dataset'). We then filtered these SNPs based on physical distance, retaining only SNPs >100 bp apart (i.e. on different RAD loci), to generate a final, putatively unlinked SNP dataset containing 83 samples and 1892 SNPs, with 8.08% overall missing data (henceforth referred to as 'unlinked SNP dataset'). This dataset of 1892 RAD loci spans roughly 180 Kb in sequence length, covering 0.0156% of our 1.15 Gb *de novo* genome assembly.

2.6 | Population genetics from nuclear SNP data

We constructed sampling maps for each species showing the using the R package *ggplot2* v3.3.5 (Wickham et al., 2020) and color-coded each species map by sampling eight equally spaced values from the color palette 'set2' provided by the R package *RColorBrewer* v1.1-2 (Harrower & Brewer, 2003). For each species, we show the geographical spread of our sample size (range 3–30 samples per species; Table S1) in the context of the species' geographical range. We then constructed an unrooted phylogenetic network to test whether taxonomy accurately reflects the major genetic clades present in this group using our extensive geographical sampling. We used our complete SNP dataset to calculate a pairwise genetic distance (Nei's *D*; Nei, 1972) matrix between all samples using the R package *StAMPP* v1.6.3 (Pembleton et al., 2013). We used this pairwise divergence matrix as input for *SplitsTree4* (v4.15.1; Huson & Bryant, 2006) to build a Neighbor-Net, which visualizes genetic distance between individual samples as a phylogenetic network (Bryant & Moulton, 2004).

To investigate the evidence for detectable shared ancestry among sampled individuals, we used the program *ADMIXTURE* v1.3.0 (Alexander et al., 2009) to assign ancestry proportions to each sample without a priori sample assignments. Starting with our unlinked SNP dataset, we removed outgroup samples (*T. leucopygius*), and all SNPs that became invariant after the removal of these samples, to minimize the potential for hierarchical structure which could mask signals of genetic divergence among our seven ingroup species. We performed 10 *ADMIXTURE* runs with this dataset (78 samples, 1570 SNPs) varying the number of ancestry bins (*K*) from 1 to 10. For all subsequent runs, we considered the optimal *K* value to be the one that minimized cross-validation error (Alexander et al., 2009). We found little support for biologically meaningful ancestry assignments using this dataset, with an optimal $K=5$, which masked species-level divergences apparent in our unrooted phylogenetic network (Figure 2a). To address this issue, we then removed the first branching ingroup taxon (*T. sanctus*) and repeated our *ADMIXTURE* run, varying *K* from 1 to 10 for this downsampled subset of six ingroup species (43 samples, 756 SNPs). Because *T. sanctus* was not included in this iteration, we then performed pairwise assignments of ancestry using only *T. sanctus* and three species with sympatric breeding distributions: *T. tristrami* (44 samples, 1167 SNPs), *T. saurophagus* (41 samples, 1060 SNPs) and *T. sordidus* (43 samples, 1117 SNPs). Ancestry assignments from all tested values of *K* (1–10) and the cross-validation process for each of these unique data subsets are available on the project GitHub repository.

To assess the genetic diversity of each species, we calculated the heterozygosity of each sample and the nucleotide diversity (Pi) of each species using the *Stacks* 'populations' module (Rochette et al., 2019). These diversity estimates were calculated from all sites successfully mapped to the 83 samples included in the final dataset (including invariant sites), before applying any SNP/locus filtering criteria. Shafer et al. (2017) endorse this approach (including all sites, before filtering) for preventing the introduction of biases related to specific filtering criteria, helping to generate estimates of genetic diversity that are comparable across species and sequencing approaches. We visualized these genetic diversity estimates as dot plots using *ggplot2*.

2.7 | Comparing patterns of relatedness between the nuclear and mitochondrial genomes

To compare interspecies divergence between the nuclear and mitochondrial genomes, we calculated pairwise divergence among the eight sequenced species separately for both the nuclear and mitochondrial markers. For mitochondrial inference, we used a nearly whole mitochondrial genome alignment (10 concatenated genes, 11,538 total bp; NCBI accession number for each sample, plus the exact alignment used in nexus format, are available at github.com/DevonDeRaad/todiramphus.radseq/tree/main/data/mitogenomes) assembled from off target UCE reads, which is a subset (eight total mitogenome samples used here, one sample for each of our focal *Todiramphus* taxa) of the broad sampling across the entire order Coraciiformes, performed by McCullough et al. (2019). We chose to use this particular set of mitogenomes because they were all sequenced and assembled using an identical approach, which minimizes the chance of introducing spurious signals of relatedness via different approaches to assembly or alignment between individual samples. Additionally, this dataset (McCullough et al., 2019) represents nearly all of the publicly available mitogenomes for *Todiramphus* species, making it the current best option for comparing patterns of relatedness between the nuclear and mitochondrial genomes. It is also worth noting that although we identified phylogeographical structure within *T. chloris* in our population genetic analyses, we recover

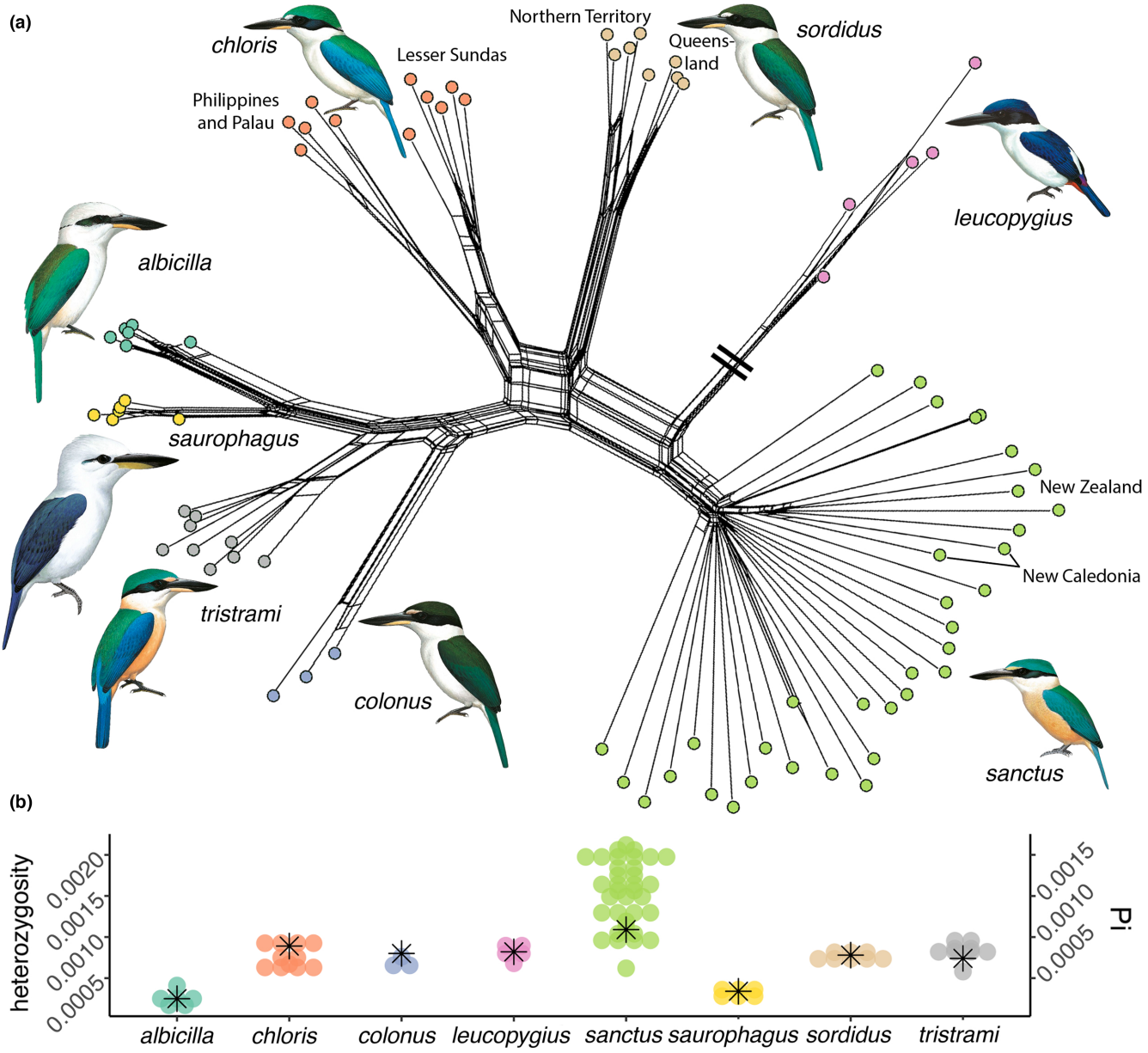

FIGURE 2 Population genetics of this group of *Todiramphus* kingfishers. (a) Unrooted phylogenetic network visualizing pairwise divergence among all samples, color coded according to species. Branch leading to *T. leucopygius* is artificially shortened, as indicated by the double hashmark. (b) Dot plot showing heterozygosity for each individual sample, binned by species. A single measure of overall nucleotide diversity (Pi) within each species is denoted by an asterisk.

all *T. chloris* individuals as reciprocally monophyletic using nuclear SNP data, matching the results of McCullough et al. (2019). These results are also concordant with Andersen et al. (2015), which found that all sampled *T. chloris* samples formed a mitochondrial clade. Based on this evidence, we treat *T. chloris* as a single monophyletic group for all phylogenetic analyses.

We used the R package *phangorn* (Schliep, 2011) to calculate pairwise, uncorrected P distance between all eight focal species across the entire mitochondrial genome alignment. We also calculated the number of pairwise differences between each taxon across the entire mitochondrial alignment using R (R Core Team, 2019). To quantify pairwise nuclear divergence, we calculated pairwise F_{ST} between all named species using our complete SNP dataset as input for *StAMPP*. We also calculated the number of pairwise fixed differences between all named species in the complete SNP dataset using R. We then plotted pairwise divergence in both the nuclear and mitochondrial genomes separate heatmaps using the R package *ggplot2*.

To estimate phylogeny for the mitochondrial genome, we used *IQ-TREE 2* v2.2.0 (Minh, Schmidt, et al., 2020) to construct a maximum likelihood tree from this concatenated alignment, allowing automatic model selection (Chernomor et al., 2016) for each partition (defined as each unique gene) and performing 1000 ultrafast bootstrap replicates (Hoang et al., 2018). We also used *IQ-TREE 2*

to generate a single-locus tree for each partition and calculate gene concordance factors (gCFs) and site concordance factors (sCFs) for each internal branch in the recovered mtDNA species tree (Minh, Hahn, et al., 2020). These concordance factors represent the number of decisive (i.e. parsimony informative) gene trees or sites supporting the given internal branch.

We then used our RAD data to construct a species tree derived from the nuclear genome. Because we needed an alignment with only a single tip per species to generate gCFs and sCFs comparable with our mtDNA alignment, we randomly downsampled our dataset to eight individuals, one per species tip. We generated a 'whitelist' of remaining loci that passed filtering protocols to be retained in our unlinked SNP dataset (1892 loci) and retained variable sites following downsampling (1236 loci). We passed this whitelist to the *Stacks* 'populations' module to generate a concatenated alignment including invariant sites for these eight samples. We then used *IQ-TREE 2* to generate a single-locus gene tree for each individual locus in the concatenated alignment and calculate 1000 ultrafast bootstrap replicates for each gene tree. Because concatenation of thousands of nuclear loci is known to produce misleading species trees in cases of recent and rapid speciation (Kubatko & Degnan, 2007), we used these individual gene trees as input for *ASTRAL-III* v5.7.7 (Zhang et al., 2018) to generate a species tree under the multispecies coalescent model. We then used *IQ-TREE 2* to calculate the gCF and sCF for each internal branch on the species topology recovered with the greatest posterior probability using *ASTRAL-III*.

To evaluate the relative gene tree and site support for the alternative topology, we also calculated mismatched gCFs and sCFs (i.e. mtDNA concordance factors for the nuDNA tree and nuDNA concordance factors for the mtDNA tree). To do so, we used the same approach described above, except for swapping in the alternative tree as the backbone topology on which to calculate concordance factors. We then visualized each tree (mtDNA and nuDNA) with posterior probability/bootstrap support values, gCFs, sCFs, mismatched gCFs and mismatched sCFs labelling each internal branch using *FigTree* v1.4.4 (Rambaut, 2014).

2.8 | Investigating signals of ILS

A hallmark signal of ILS is gene tree discordance. To quantify the exact proportion of concordant topologies recovered from genome-wide gene trees for each internal branch on our species tree, we calculated quartet frequencies using the flag '-t 16' in *ASTRAL-III*. We then visualized the frequency of each alternative quartet topology for each internal branch as a rooted three-tip tree, where the root leads to the clade containing the outgroup *T. leucopygius*.

Both ILS and/or gene flow could result in alternative, discordant species tree topologies with similar posterior probabilities to the tree we recovered with the greatest posterior probability using *ASTRAL-III*. To visualize the entire suite of likely species tree topologies, we used *SNAPP* v.1.5.2 (Bryant et al., 2012) implemented in *BEAST2* v.2.6.4 (Bouckaert et al., 2014) to sample the posterior distribution of species trees and overlaid all sampled tree topologies on a single plot using *DensiTree* v2.2.7 (distributed with *BEAST2*). To generate input for *SNAPP*, we downsampled our unlinked SNP dataset to the same eight samples used to generate input gene trees for *ASTRAL-III* and removed SNPs that became invariant due to downsampling, resulting in 657 unlinked, bi-allelic variable sites. We performed two identical *SNAPP* runs specifying different starting seeds, each with an MCMC chain length of 5,000,000, sampling every 1000 iterations. After discarding the first 1000,000 iterations of each run as burn-in, we used *Tracer* v1.7.1 (Rambaut et al., 2018) to confirm chain stationarity and effective sample size (ESS) >200 for the posterior, likelihood and prior parameters. We then combined these runs using *LogCombiner* v2.5.2 (distributed with *BEAST2*) and visualized all 8000 trees sampled from the post-burn-in posterior distribution of species trees in both runs, emphasizing the 'root canal' tree, or the single bifurcating topology with the greatest posterior probability, using *DensiTree*.

2.9 | Investigating signals of gene flow

To statistically test for excess allele sharing between non-sister species using an ABBA-BABA framework, we used the program *Dsuite* v0.4 (Malinsky et al., 2021). We used our filtered SNP dataset (including 7430 total SNPs shared across 83 samples assigned to eight species tips) as input and calculated a *D* statistic value for every possible (((P1,P2),P3),OG) topology where BBAA was the most common site pattern, meaning a rooted topology of ((P1,P2),P3) was supported by the SNP data under a parsimony framework. We used the *Dtrios* module from *Dsuite* to perform block Jackknifing using the default setting of 20 equally distributed SNP windows across the genome, to estimate standard error and generate a *Z* score for each calculated pairwise *D* statistic, which could then be converted into a *p*-value. Following developer recommendations (Malinsky et al., 2021), we accounted for multiple testing by interpreting these *p* values under a false discovery rate (FDR)-adjusted statistical significance threshold (Benjamini & Hochberg, 1995), which we calculated by dividing a standard .05 significance threshold by the total number of tests implemented (here 35 pairwise tests; FDR-adjusted *p*-value threshold = .00143).

To identify putative migration edges that could explain allele frequency variance among our eight focal species, we used *TreeMix* v4 (Pickrell & Pritchard, 2012) to infer a species tree and then iteratively add putative migration edges. We used our unlinked SNP dataset (83 individuals and 1892 SNPs) as input and assigned all samples to eight species tips matching the same eight tips used for *ASTRAL-III* and *SNAPP* reconstructions. We began by inferring a species tree specifying zero migration edges, and then progressively added the first three putative migration edges that explained the greatest proportion of additional variance in the underlying allele frequency data. Because each additional migration edge adds an additional parameter, each edge will inherently increase the proportion of variance explained by the resulting graph, making it difficult to determine how

many migration edges are plausible and relevant in a given system. For maximal transparency, we present the three most important migration edges, but offer our own interpretation of the biological plausibility and relevance of these migration edge hypotheses.

3 | RESULTS

3.1 | *De novo* reference genome assembly

De novo assembly of all linked read pairs sequenced from our 10X Chromium sequencing library generated an initial phased genome assembly comprising 28,857 contigs contained in 10,881 total scaffolds, with an overall length of 1.15 gigabases (Gb) and 1.6% of the overall assembly comprised of *N*'s (missing base calls). Contiguity metrics revealed a contig N_{50} = 126.7 kilobases (Kb), and a scaffold N_{50} = 4.69 megabases (Mb) for this *de novo* assembly, suggesting that the relatively short contigs we recovered were effectively linked into moderate to large sized scaffolds (largest scaffold = 20.9 Mb) via the long-range linkage information afforded by this linked read approach. A *BUSCO* analysis revealed that 95.9% of all expected avian orthologs were present and complete in our *de novo* assembly (7999 out of 8338 searched; 95.7% present in a single copy and 0.2% present with duplicate copies), while 1.5% were present in a fragmented state, and 2.6% were completely absent. We note that these exact values may differ slightly in the finalized version of this assembly, which had ~0.005% of bases removed or masked via the NCBI Foreign Contamination Screening program. In sum, our *de novo* assembly is highly complete, with a reasonably low proportion of missing base calls, and is comprised of short contigs and moderately sized scaffolds.

3.2 | Population genetics of this kingfisher radiation

An unrooted phylogenetic network based on SNP data for all 83 samples passing our filtering requirements (Table S1) reveals that the current species-level taxonomy accurately describes the diversity of our eight *Todiramphus* focal taxa (Figure 2a). Within species, a deep phylogeographical break divides geographically isolated populations of *T. chloris* into two divergent genomic clades. The species status of these two allopatrically distributed sister clades deserves thorough investigation using an integrative approach to quantifying genomic, morphological and vocal divergence in future studies. Meanwhile, within *T. sanctus*, allopatrically distributed, non-migratory populations on New Caledonia and New Zealand cluster together but are embedded within the overall range-wide genetic diversity of the species (Figure 2a). Plotting heterozygosity and nucleotide diversity for all sampled species reveals that genetic diversity is generally correlated with range size (Figure 2b). This association is mainly evidenced by the reduced diversity in the range-restricted species *T. albicilla*, and the startlingly high levels of nucleotide diversity and heterozygosity in the highly vagile, migratory species *T. sanctus*, whose breeding distribution spans nearly the entire Australian continent and additional islands in the Pacific. Meanwhile, the genetic diversity of *T. colonus* shows noticeable departure from this pattern, with no evidence for reduced diversity despite an incredibly restricted geographical range (Figures 1 and 2b).

Assignments of genomic ancestry show that the seven ingroup species cannot be accurately distinguished by the program *ADMIXTURE* likely due to a combination of differences in sample size between species, large differences in effective population size between species, and hierarchical structure masking recent divergence (Figure 3a). After removing *T. sanctus*, the first branching taxon among these seven ingroup species, optimal ancestry assignments ($K = 8$) for the remaining six ingroup species resulted in no samples with >1% inter-species ancestry assignment (Figure 3b). Further pairwise comparisons between *T. sanctus* and taxa with sympatric breeding distributions (Figure 3c–e) revealed limited evidence for shared ancestry involving *T. sanctus*. Among these pairwise comparisons (Figure 3c–e), only a single sample was assigned >3% interspecies ancestry. This was a single *T. sanctus* sample that was assigned 6.7% *T. sordidus* ancestry in the pairwise comparison between these species (Figure 3e), a result that could indicate shared genomic ancestry or difficulty in accurately distinguishing between closely related species in this pairwise comparison (e.g. Figure 3a).

3.3 | Evidence for mitonuclear discordance

Pairwise comparison of relative divergence between these eight species gives the first glimpse into the differences between the mitochondrial and nuclear genomes in this group of *Todiramphus* kingfishers. Raw mitochondrial divergence percentages (excluding comparisons involving outgroup *T. leucopygius*) range from 0.59% to 1.95% (Figure 4a), implying extraordinarily rapid diversification among these seven ingroup species; within the last one million years based on commonly used avian mtDNA divergence time calibrations (Weir & Schluter, 2008). In contrast, high pairwise F_{ST} values (range 0.3–0.78, again excluding *T. leucopygius*) and dozens to hundreds of fixed differences (range 30–236) throughout the genome indicate a highly differentiated nuclear genome (Figure 4b). Meanwhile, phylogenetic reconstructions imply starkly different evolutionary histories between the nuclear and mitochondrial genomes in this group of *Todiramphus* kingfishers (Figure 4c). For instance, a species tree based on a near-complete mitogenome alignment indicates a sister relationship between *T. sanctus* and *T. sordidus*, while a species tree generated from thousands of nuclear markers indicates that *T. sanctus* is first branching among the ingroup *Todiramphus* species we sequenced. Despite recovering highly discordant topologies between the nuclear and mitochondrial genomes, the posterior probabilities and bootstrap support values are generally high across each respective reconstruction. gCFs and sCFs reveal that a majority of individual sites and genes in the mitochondrial genome concordantly support the recovered mtDNA topology (mean sCF = 55.5, mean gCF = 54;

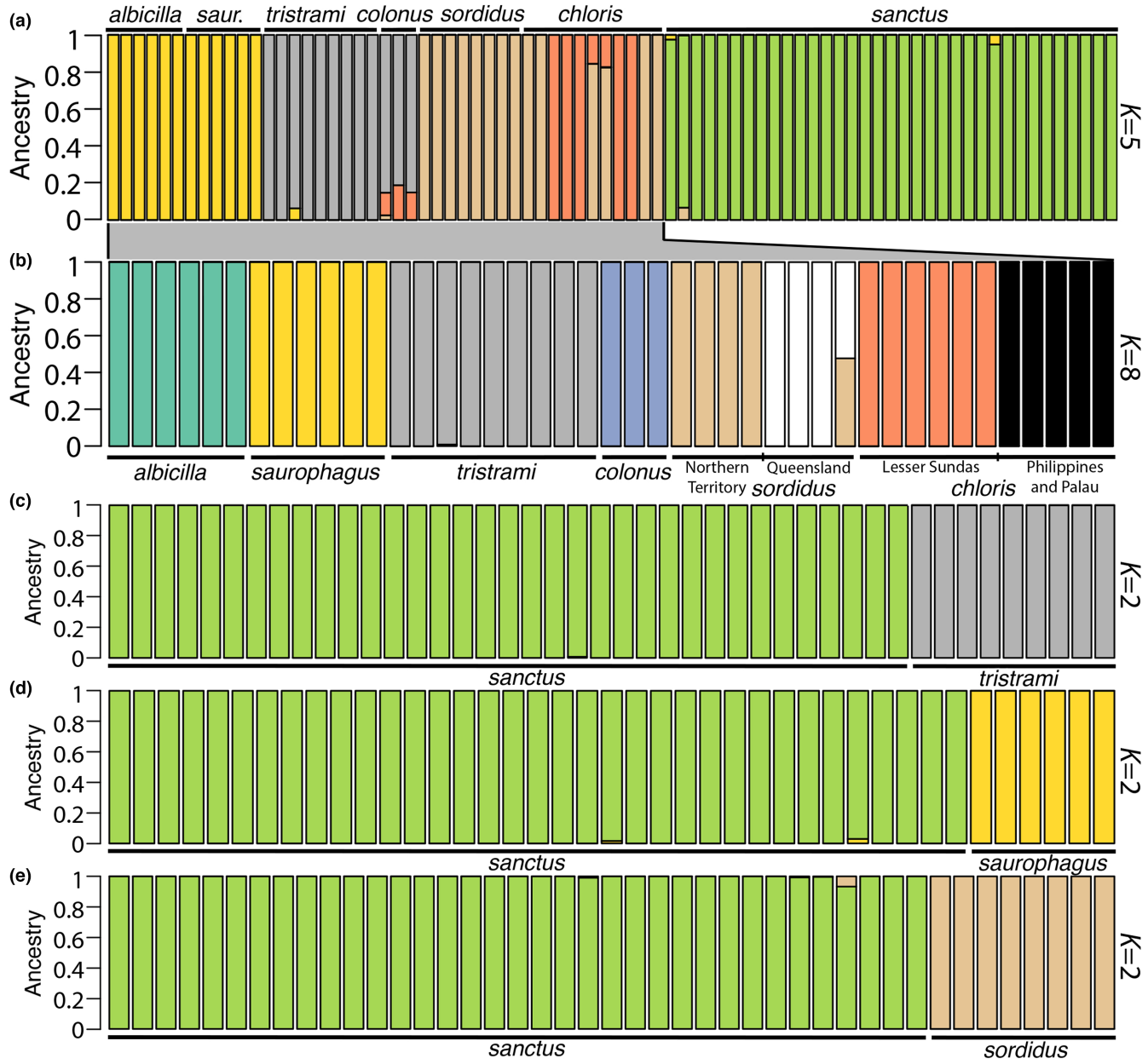

FIGURE 3 Ancestry assignments according to the optimal clustering schemes for various subsets of our overall dataset, which could not be analyzed together because of hierarchical structure masking recent divergence. All color schemes match the colors used to denote species in Figure 2 as closely as possible. (a) Optimal clustering scheme ($K=5$) for all seven ingroup species shows clear signatures of hierarchical structure masking differentiation among closely related species that are distinguishable in the phylogenetic network in Figure 2a. (b) After removing the first-branching species (*Todiramphus sanctus*) to reduce this effect, the optimal clustering scheme ($K=8$) for the six remaining ingroup species shows clear differentiation both between and within species. Geographical substructuring within *T. chloris* and *T. sordidus* is shown in black and white, respectively. (c–e) Ancestry assignments ($K=2$ recovered as optimal for all comparisons) for pairwise comparisons between *T. sanctus* and geographically proximate species with overlapping breeding distributions.

Figure 4c). Meanwhile, the recovered nuclear topology is supported by a majority of decisive variable sites (mean sCF = 54.2), but only a small fraction of genes (mean gCF = 16.1), likely reflecting a combination of real biological uncertainty (i.e. gene tree discordance), and stochastic error in our ability to resolve gene tree topologies from short (~100 bp) RAD loci. gCFs and sCFs calculated on the topology recovered from the alternate genome were generally lower, reflecting reduced support for the alternate topology in both genomes (Figure 4c).

3.4 | Evidence for ILS

To determine the strength of ILS throughout the nuclear genome, we quantified the frequencies of all possible quartet topologies for each internal branch of the recovered species tree built using *ASTRAL-III* (Figure 5a). Quartet frequencies (shown as three tip trees consistently rooted on the clade containing the outgroup, *T. leucopygius*) reveal that the dominant topology is not recovered in >50% of internal

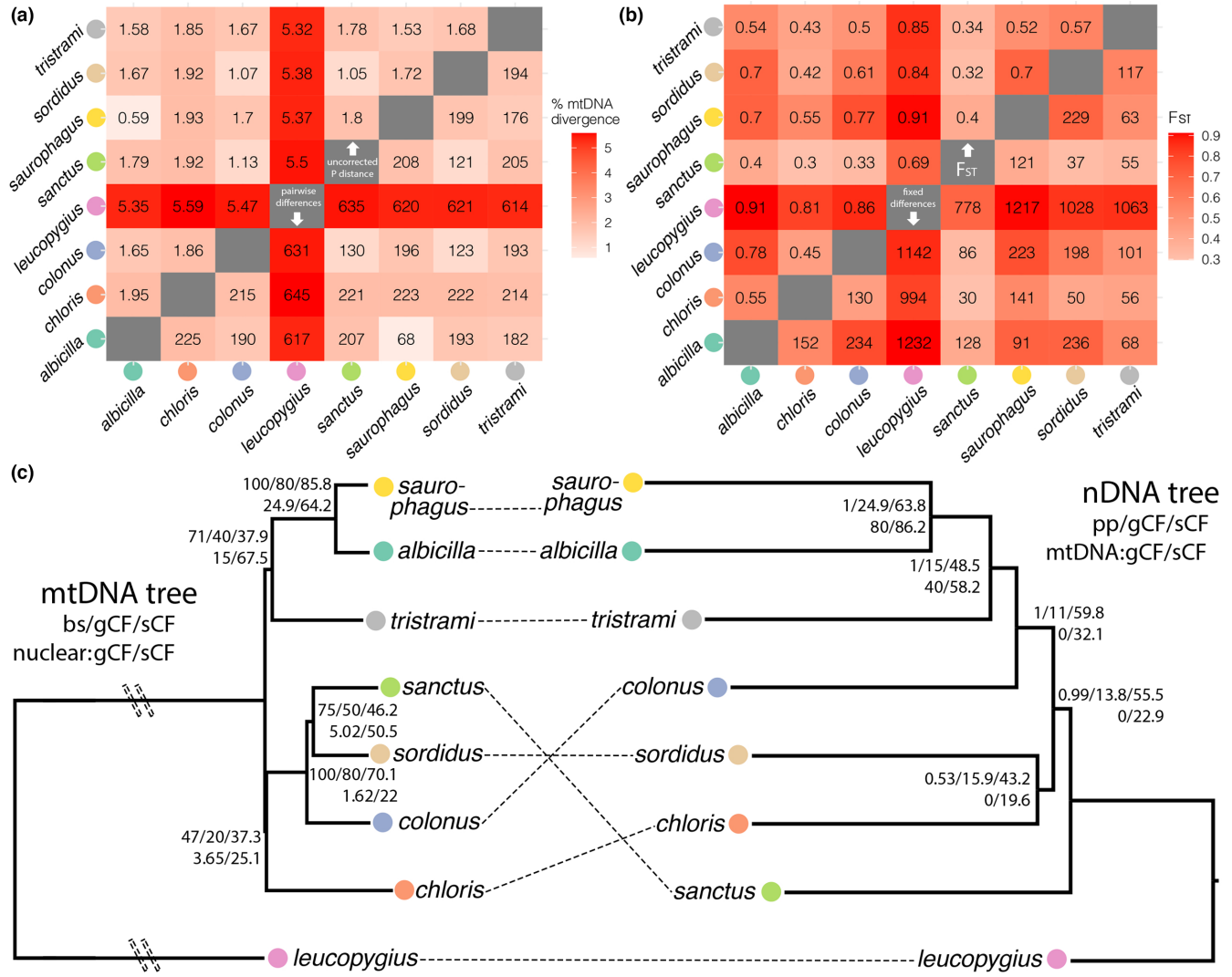

FIGURE 4 Comparing the evolutionary histories of the nuclear and mitochondrial genomes for this group. (a) Heatmap showing pairwise uncorrected P distance in mtDNA between all eight species. Above the diagonal, each cell is labeled with uncorrected P distance, equivalent to raw mtDNA divergence percentage. Below the diagonal, each cell is labeled with the number of pairiwse nucleotide differences. (b) Heatmap showing pairwise F_{ST} among all species. Above the diagonal, each cell is labeled with exact pairwise F_{ST} value, and below the diagonal each cell is labeled with the number of pairwise fixed differences. (c) Left, a mitochondrial species tree inferred from a concatenated alignment of 10 mitochondrial genes. Double hash marks indicate artificially shortened branches. Each internal branch is labeled: bootstrap support (bs)/gene concordance factor (gCF)/site concordance factor (sCF), with corresponding gCF and sCF values calculated from nuDNA gene trees listed below. Right, a summary nuclear species tree generated from 1236 input gene trees, with each internal branch labeled: posterior probability (pp)/gCF/sCF, with corresponding gCF and sCF values calculated from 10 mtDNA gene trees listed below.

gene trees for any internal branch in the consensus species tree, indicating high levels of ILS. For three out of five internal branches, there is an alternate quartet topology found at a genome-wide frequency within 5% of the dominant topology, evidence for extensive genome-wide, gene-tree discordance. In another attempt to assess the prevalence of conflicting signals across the genome, we built a species tree using SNPs as input for *SNAPP* (Figure 5b). The topology recovered by *SNAPP* with the greatest posterior probability was identical to our *ASTRAL-III* reconstruction, but visualization as a cloudogram revealed many discordant species tree topologies sampled from the posterior distribution, which could result from gene flow or ILS. The taxa *T. colonus* and *T. tristrami* are especially fuzzy in the cloudogram visualization, suggesting uncertainty in their phylogenetic placements. This uncertainty is an expected result considering the equivocal quartet frequencies for internal branches determining the placement of these taxa (Figure 5a).

3.5 | Evidence for gene flow

Nuclear ancestry assignments from *ADMIXTURE* indicate that, once hierarchical structure is properly accounted for, there is little sharing of genomic ancestry across species boundaries via recent hybridization (Figure 3b–e). Still, assignments of genomic ancestry

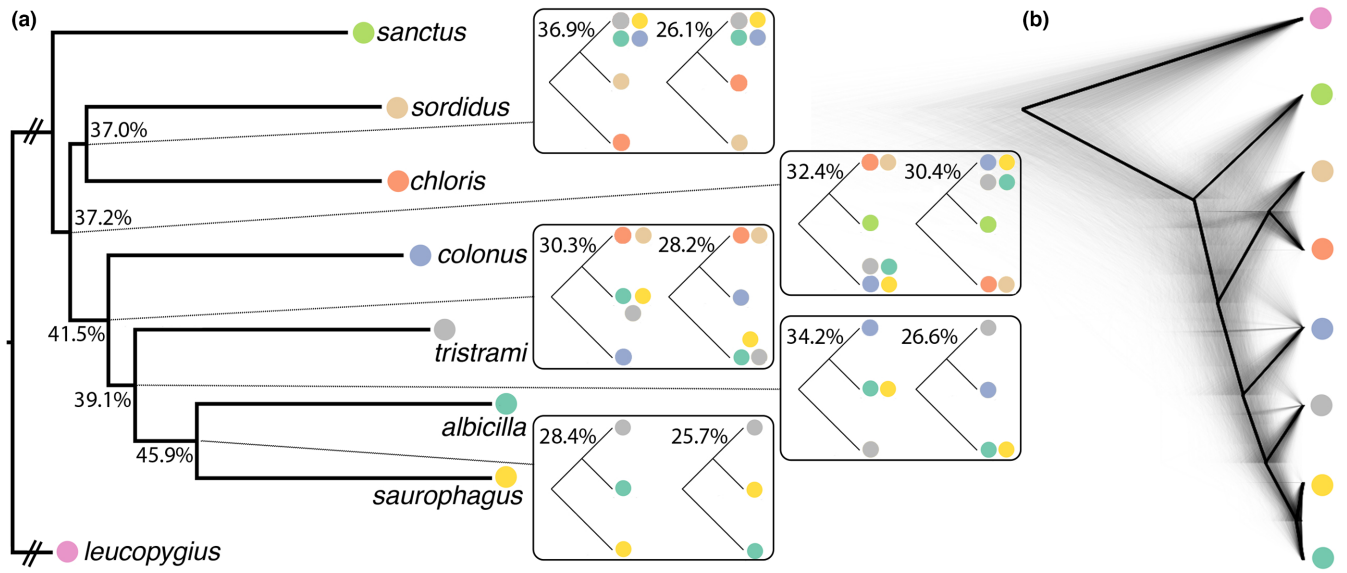

FIGURE 5 Evidence for incomplete lineage sorting among the eight focal *Todiramphus* taxa. (a) Nuclear species tree generated using gene trees as input for *ASTRAL-III*. This tree is identical to Figure 3d, except here internal branches are annotated with quartet frequencies. For each internal branch, alternate quartet topology frequencies are shown. Quartet trees are consistently shown as three tip trees rooted on the clade containing *T. leucopygius* (not shown) to increase interpretability. (b) Cloudogram generated using unlinked SNPs as input for *SNAPP*, showing 8000 samples from the posterior distribution of species trees to visualize the frequency of various alternate topologies generated by gene flow or ILS.

are expected to discriminate between species with high accuracy if recent hybrid individuals are not sampled, even in cases of extensive historical gene flow (e.g. DeRaad et al., 2022). Under these circumstances, histories of gene flow between non-sister taxa can be readily detected using ABBA-BABA tests, which test for a statistical overrepresentation in the number of alleles shared between non-sister lineages, based on a null assumption of ILS. Here, for all possible permutations of a (((P1,P2),P3),OG) topology where BBAA was the most common site pattern (i.e. three-tip, rooted topologies concordant with a parsimony-based species tree), we used an ABBA-BABA framework to calculate a *D* statistic value measuring the proportion of excess allele sharing between P2 and P3 (Figure 6a). The minimum *D* statistic value for each pairwise inter-species comparison hovers near zero (range <0–0.25), which is the expectation under a null model of pure ILS with no gene flow (Figure 6a). None of these minimum pairwise *D* statistic values deviate from the null expectation (i.e. zero), drastically enough to reject the null hypothesis of strictly tree-like evolution across these eight taxa (based on an FDR-adjusted *p*-value threshold = .00143). It is important to note that 17 out of the top 20 pairwise *D*-statistic values we recovered, including the only significant pairwise comparison, specify either *T. albicilla* or *T. saurophagus* as the P2 taxa (Table 1). This is perfectly concordant with recent evidence (Frankel & Ané, 2023) that an increased substitution rate in P2 (we suspect an increased substitution rate in *T. albicilla* and *T. saurophagus*, which display low genetic diversity and assumedly increased rates of fixation due to strong genetic drift) will generate an increased number of ABBA site patterns due to homoplasy, inflating the false-positive rate of ABBA-BABA significance tests (Table 1). The three remaining pairwise comparisons among our top 20 recovered *D*-statistic values indicate gene flow between *T. sordidus* and *T. chloris*, which is concordant with their sister relationship in our recovered species tree (Figure 4c).

In an orthogonal approach to reconstructing a species tree and identifying potential gene flow events, we used *TreeMix* to infer a phylogenetic backbone and progressively added migration edges that explained the maximum amount of variance in the underlying allele frequency data (Figure 6b). The recovered species tree is perfectly concordant with the species tree we recovered in previous analyses and explains 99.54% of the overall variance in allele frequencies. The first and third most impactful recovered migration edges include the outgroup (*T. leucopygius*), which makes little plausible biological sense considering the apparent rapid evolution of reproductive isolating barriers between sympatric ingroup taxa which share more recent ancestry and geographical proximity. Meanwhile, the second migration edge connects *T. chloris* and *T. colonus*, which is biologically plausible, but explains only 0.1% of overall variance and does not align in any obvious manner with the recovered mitochondrial relationships. Ultimately, there is no way for *TreeMix* to statistically reject the presence of gene flow, as adding migration edges will always increase the proportion of variation explained by the resulting admixture graph. With this context, the most parsimonious interpretation of our *TreeMix* analysis is that the backbone phylogeny explains nearly all of the allele frequency variation in the dataset, while putative migration edge hypotheses can be discounted based on their lack of biological plausibility and low proportion of variance explained. This interpretation is concordant with our *ADMIXTURE* ancestry assignments, and ABBA-BABA test results, and supports a bifurcating

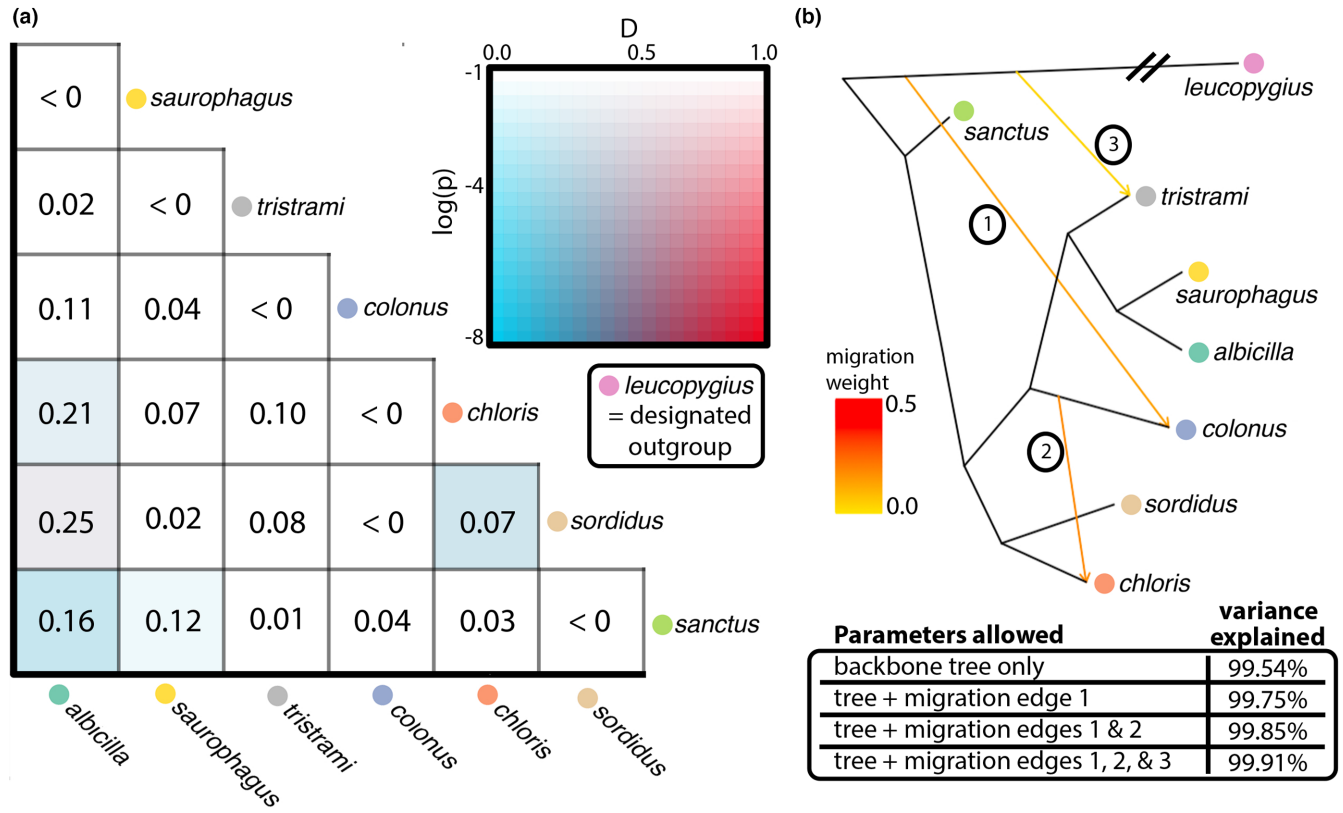

FIGURE 6 Assessing evidence for nuclear gene flow. (a) Heatmap colored according to the *p*-value and *D* statistic value (color-coding shown by inset box) for each pairwise ABBA-BABA comparison, labeled with the minimum pairwise *D* statistic value for each comparison. None of the FDR-adjusted *p*-values associated with these minimum pairwise *D* statistic values reached a false-discovery-rate-adjusted significance threshold ($p = .00143$). (b) Species tree plus potential migration edges inferred from allele frequencies among all 83 samples passing filtering protocols. Migration edges are not statistically supported gene flow events, but rather hypotheses that maximize the percentage of variance explained by the species network. Adding parameters, for example, migration edges, will always increase the percentage variance explained, so these additional parameters must be viewed with skepticism. The double hashmark indicates an artificially shortened branch leading to *T. leucopygius*.

phylogeny (Figure 4c) as the most parsimonious representation of the nuclear species tree for our eight focal *Todiramphus* species, based on this 1892 locus RAD dataset.

4 | DISCUSSION

Many studies have identified topological discordance between phylogenetic reconstructions from nuclear and mitochondrial genomes; however, the causes of this discordance are rarely understood in detail (Toews & Brelsford, 2012). Here, we attempted to understand in depth the evolutionary mechanisms contributing to mitonuclear discordance among a group of recently diverged *Todiramphus* kingfishers (Andersen et al., 2015; McCullough et al., 2019). We first show that this case of mitonuclear discordance is a real biological phenomenon and not simply a result of model violations (Kimball et al., 2021; Weisrock, 2012), incorrect taxonomy, or erroneous species delimitation (Toews & Brelsford, 2012). Concordance factors and quartet frequencies indicate extensive genome-wide nuclear gene tree discordance, and a lack of evidence for gene flow implicates ILS as the main explanation for this discordance. ILS occurs when individual gene trees do not coalesce (looking backwards in time) during the waiting times between speciation events (i.e. deep coalescence). This means that the proportion of gene trees throughout the genome showing patterns of relatedness in conflict with the resolved species tree is a direct estimate of the frequency of ILS (Hibbins & Hahn, 2022). Here, we find that none of the internal branches recovered in our species tree are supported by >50% of genome-wide gene trees, implying extremely short waiting times between speciation events (measured in coalescent units) and high ILS. This positive evidence for ILS combined with a lack of detectable evidence for gene flow events that could plausibly explain the mitonuclear discordance we observe, leads us to conclude that mitonuclear discordance in the *Todiramphus* kingfishers results from deep coalescence of the mitogenome. Although ILS is uncommonly implicated in cases of mitonuclear discordance (McKay & Zink, 2010), we posit that the fragmentary nature of the Indo-Pacific ecoregion, known for driving explosive radiation in birds (Moyle et al., 2016), may present an ideal recipe for generating high levels of genome-wide gene tree discordance, including mitonuclear discordance.

TABLE 1 Detailed results from all pairwise ABBA-BABA tests conducted using *Dsuite*.

P1	P2	P3	*D* statistic	*Z*-score	*p*-value	F_4-ratio	BBAA	ABBA	BABA
saurophagus	*albicilla*	*sordidus*	0.43	2.32	.0101	0.09	177.4	18.4	7.3
colonus	*albicilla*	*sordidus*	0.25	2.31	.0103	0.13	109.3	42.1	25
saurophagus	*albicilla*	*chloris*	0.25	1.78	.0377	0.11	149	18	10.7
colonus	*albicilla*	*chloris*	0.25	2.33	.0099	0.24	89.2	48.4	29.1
tristrami	*albicilla*	*sordidus*	0.25	2.03	.0214	0.1	133.8	30.8	18.6
colonus	*chloris*	*sordidus*	0.23	2.92	.0017	0.18	62.5	61.8	38.4
tristrami	*albicilla*	*sanctus*	0.23	2.24	.0125	0.26	172	27.9	17.4
tristrami	*albicilla*	*chloris*	0.21	1.95	.0253	0.16	107.9	32.9	21.6
saurophagus	*chloris*	*sordidus*	0.18	2.05	.02	0.13	69	54.4	37.8
tristrami	*chloris*	*sordidus*	0.18	2.23	.0129	0.14	69.2	58.6	40.7
colonus	*albicilla*	*sanctus*	0.18	1.5	.0662	0.25	141.2	34.2	24
sordidus	*albicilla*	*sanctus*	0.17	3.03	**.0012**	0.35	96.9	55.8	39.3
chloris	*albicilla*	*sanctus*	0.17	2.63	.0043	0.3	114.7	45.6	32.4
saurophagus	*albicilla*	*sanctus*	0.16	1.1	.1362	0.11	213.3	13.2	9.5
colonus	*saurophagus*	*chloris*	0.16	1.24	.1077	0.15	88.8	46.1	33.7
tristrami	*saurophagus*	*sanctus*	0.14	1.18	.1199	0.17	177	27.3	20.4
sordidus	*saurophagus*	*sanctus*	0.13	1.89	.0291	0.29	92.3	57.1	43.8
chloris	*saurophagus*	*sanctus*	0.13	1.67	.0473	0.24	112	45.5	35
saurophagus	*albicilla*	*colonus*	0.13	1.12	.1314	0.04	110.3	21.6	16.7
colonus	*saurophagus*	*sanctus*	0.12	1.03	.1525	0.18	141.3	34.6	27.2
tristrami	*albicilla*	*colonus*	0.11	0.93	.1766	0.07	75.9	43	34.6
colonus	*tristrami*	*chloris*	0.1	0.87	.1925	0.11	90.5	47.4	38.8
colonus	*saurophagus*	*sordidus*	0.1	0.81	.2089	0.05	110	37.4	30.6
colonus	*tristrami*	*sordidus*	0.08	0.64	.2619	0.04	112.1	42	36.2
albicilla	*chloris*	*sordidus*	0.07	0.76	.225	0.05	70.6	48.7	42.6
tristrami	*saurophagus*	*chloris*	0.07	0.73	.2326	0.06	114.9	31.6	27.6
sordidus	*tristrami*	*sanctus*	0.06	0.96	.1685	0.14	94.6	53.8	47.4
sordidus	*colonus*	*sanctus*	0.05	0.63	.2635	0.12	86.8	51.9	46.7
chloris	*tristrami*	*sanctus*	0.04	0.67	.2512	0.08	112	42.7	39.1
tristrami	*saurophagus*	*colonus*	0.04	0.41	.3394	0.03	79.6	40.9	37.5
chloris	*colonus*	*sanctus*	0.04	0.42	.3379	0.08	103.1	44.1	40.7
sordidus	*chloris*	*sanctus*	0.03	0.5	.3081	0.06	108.4	46	43.2
tristrami	*saurophagus*	*sordidus*	0.02	0.17	.4321	0.01	141.6	27.1	25.9
saurophagus	*albicilla*	*tristrami*	0.02	0.17	.4334	0.02	83.5	27.5	26.6
colonus	*tristrami*	*sanctus*	0.01	0.13	.4492	0.02	142.6	32.8	31.9

Note: Columns P1, P2 and P3 indicate the species used in each position of the (((P1,P2),P3),OG) four-tip topology for each test (*T. leucopygius* was consistently designated as the outgroup taxon for each test). In each test, P1 and P2 are automatically swivelled so that P2 and P3 always share a greater proportion of alleles than P1 and P3. Columns BBAA, ABBA and BABA indicate the count for each of these site patterns under the (((P1,P2),P3),OG) topology (because we included 3–35 samples per species, counts are derived from allele frequencies and can be fractional). Note that a single *p*-value (denoted in bold) is considered significant based on our FDR-adjusted *p*-value threshold ($p = .00143$), but this significant evidence for excess P2-P3 allele sharing is contradicted by multiple non-significant comparisons between the same taxa where P1 is specified as an alternate species.

4.1 | Evidence for mitonuclear discordance

Understanding the relative divergence between lineages in the mitochondrial and nuclear genomes can help identify the evolutionary processes contributing to mitonuclear discordance (Morales et al., 2015). Here, we find shallow mitochondrial divergence among all seven ingroup taxa, indicating a common ancestor roughly within the last million years, assuming strict clock-like evolution (Weir & Schluter, 2008). This speciation frequency would support one of the highest diversification rates known among all

extant birds (Andersen et al., 2015). Meanwhile, high pairwise nuclear F_{ST} values could be explained by repeated purging of shared variation via repeated population bottlenecks. Although an in-depth investigation of historical demography is outside of the scope of this study, the repeated island colonization events necessary to persist across a fragmented Indo-Pacific distribution (Filardi & Moyle, 2005) suggest complex demographic histories in this group. Furthermore, the uneven distribution of genetic diversity among the eight focal taxa we sequenced indicates divergent current and historical effective population sizes between these species. Patterns of relatedness throughout the genome could be affected in myriad ways by the massive variance in effective population size between isolated island taxa, and closely related, widely distributed migratory continental taxa in our focal clade of interest (e.g. Schrago & Seuánez, 2019). These confounding forces make historical demography a promising avenue for future investigations into the evolutionary history of this species complex.

Beyond differences in degree of divergence, we also identified strongly supported topological differences between species trees reconstructed using nuclear and mitochondrial data. Using concordance factors (Minh, Hahn, et al., 2020), we identified multiple internal branches supported by a large proportion of the nuclear and very little of the mitochondrial genome, or vice versa. This represents a new use-case for gCFs and sCFs as an objective method for quantifying the proportion of discordant loci between the nuclear and mitochondrial genomes for specific internal branches. By establishing this discordance at both the individual site and gene levels, we distinguish this case of mitonuclear discordance from cases where resolving the nuclear species tree is simply intractable (e.g. Scherz et al., 2022). Our approach highlights the value of understanding genome-wide support for specific topologies, and we suggest future investigations incorporate concordance factors and quartet frequencies into the arsenal of modern genomic tools available for understanding mitonuclear discordance.

4.2 | Could confounded taxonomy create patterns of mitonuclear discordance?

One potential explanation for mitonuclear discordance is confounded taxonomy and species limits (Després, 2019; Marshall et al., 2021). This can occur in cases of speciation reversal where the nuclear genomes of previously divergent lineages become homogenized via gene flow upon secondary contact, if divergent mitochondrial haplotypes remain segregating in the population (e.g. Hinojosa et al., 2019; Kearns et al., 2018; Slager et al., 2020). In these cases, taxonomy and species limits determined by either mitochondrial or nuclear loci alone will conflict with signals of relatedness in markers sampled from the alternate genome. Here, population genetic approaches using nuclear SNP data as input find support for the eight currently recognized focal species, which were delimited based on plumage, geography and mitochondrial DNA sequence data (Andersen et al., 2015). These results indicate that mitonuclear discordance in *Todiramphus* is not a result of confused species limits or taxonomy, despite the complex biogeographical history of the group and multiple cases of sympatry among the eight focal taxa we investigated. Based on the lack of statistical evidence for gene flow, it seems that isolating barriers have prevented the extensive sharing of genetic material across species boundaries within *Todiramphus*, allowing us to rule out mitochondrial capture or speciation reversal as potential explanations for the mitonuclear discordance we observe.

4.3 | The potential role of gene flow

Another potential explanation for the mitonuclear discordance we observe is partial reproductive isolation allowing introgression of nuclear alleles across species boundaries despite the maintenance of reciprocal monophyly in the mitogenome (McKay & Zink, 2010). This is a commonly invoked explanation for mitonuclear discordance, especially in ZW systems (e.g. birds) where female (i.e. the heterogametic sex) hybrid offspring are predicted to have lower fitness than their male counterparts (Haldane's Rule; Orr, 1997). In these systems, we expect reduced allele sharing in the (maternally inherited) mitochondrial genome across species boundaries (Carling & Brumfield, 2008) compared with nuclear alleles. This differential gene flow could result in different patterns of relatedness between the mitochondrial and nuclear genomes (McKay & Zink, 2010; Toews & Brelsford, 2012). Here we found little indication of recently shared inter-species ancestry in the nuclear genome (Figure 3), indicating strong reproductive isolation, even between recently diverged sympatric *Todiramphus* species (Andersen et al., 2015).

Despite this limited evidence for recent introgression, individual ancestry assignments could miss signatures of rare or historical gene flow events (e.g. DeRaad et al., 2022). In these cases, genomic signatures of ILS and gene flow can still be differentiated via statistical approaches (e.g. ABBA-BABA tests; Durand et al., 2011; Green et al., 2010; Malinsky et al., 2021). Although we find evidence for some pairwise comparisons deviating from the null hypothesis of the ABBA-BABA test (i.e. $D=0$), nearly all of these deviations can be explained by substitution rate variation between our ingroup taxa. Specifically, elevated frequencies of ABBA patterns were consistently observed in cases where P2 is likely to have an elevated substitution rate compared to P1 (i.e. P2 was a taxon with reduced genetic diversity like *T. albicilla* or *T. saurophagus*), a phenomenon known to generate spurious false-positive *D* statistic values in pairwise ABBA-BABA tests (Frankel & Ané, 2023). Furthermore, a phylogenetic approach allowing reticulation events fails to identify biologically plausible horizontal gene flow events in the nuclear genome that could explain the mitonuclear discordance we observe (Figure 6b). These multipronged lines of reasoning culminate in scant evidence for nuclear gene flow among this group of *Todiramphus* kingfishers.

4.4 | The biogeography of mitonuclear discordance

Previous research on mitonuclear discordance has often focused on contact zones in continental regions where lineages come into secondary contact (McKay & Zink, 2010; Toews & Brelsford, 2012). These studies have proved fruitful, often identifying cases where differential introgression between the nuclear and mitochondrial genomes has led to discordance (e.g. Denton et al., 2014; Gompert et al., 2008; Mao & Rossiter, 2020; Shu et al., 2022; among many others). Here we highlight a fundamentally different evolutionary scenario, in which rapid and repeated speciation across the highly fragmented Indo-Pacific ecoregion has led to extensive genome-wide ILS (Lopes et al., 2021; Suh et al., 2015) and well-supported mitonuclear discordance. Much empirical and theoretical research has documented that rapid, frequent speciation events create the ideal conditions for rampant ILS, because alleles do not have time to coalesce between speciation events (Maddison & Knowles, 2006; Mendes & Hahn, 2016). So far, genomic research on ILS has typically focused on its ability to mislead nuclear species tree reconstruction, specifically concatenation-based phylogenomic approaches (Degnan & Rosenberg, 2006; Kubatko & Degnan, 2007; Kutschera et al., 2014; McCormack et al., 2009; Mirarab et al., 2016; Suh et al., 2015).

Meanwhile, the role of ILS in cases of mitonuclear discordance has only rarely been investigated in depth (except see Firneno et al., 2020; Wang et al., 2018). Here we posit that ILS, resulting largely from the explosive diversification history of the *Todiramphus* kingfishers (Andersen et al., 2015), did not allow sufficient time for the accurate sorting of most loci between speciation events (Maddison & Knowles, 2006) including the mitochondrial genome which can be thought of as a single locus because it is non-recombining. These high ILS conditions were likely further exacerbated by the extreme variance in ecology, range size and genetic diversity among our eight focal species (e.g. Schrago & Seuánez, 2019). Our results suggest that other systems with short waiting times between speciation events (e.g. anomaly zone cases; Degnan & Rosenberg, 2006) and complex demographic histories are also likely candidates for mitonuclear discordance. We suggest that the Indo-Pacific ecoregion may be a hotspot for ILS and mitonuclear discordance, because the fragmented and dynamic biogeographical history of the region regularly facilitates rapid speciation (Andersen, Naikatini, et al., 2014; Andersen, Nyári, et al., 2014; Andersen et al., 2013, 2015; Moyle et al., 2009, 2016). Other radiations of Indo-Pacific birds with similar magnitude and breadth (i.e. 'great speciators' such as *Ceyx*, *Zosterops*, *Pachycephala*, *Myzomela*, *etc.*; Diamond et al., 1976) are excellent candidate clades for future studies assessing the frequency of mitonuclear discordance across this ecoregion.

5 | CONCLUSIONS

The eight focal *Todiramphus* kingfisher species that we studied here provide an ideal opportunity to investigate the evolutionary forces that lead to mitonuclear discordance. We confirmed well-supported mitonuclear discordance using range-wide sampling and genomic data, and found little support for commonly invoked phenomena such as confounded species limits, modern or historical nuclear gene flow, or mitochondrial capture events, that could explain the mitonuclear discordance we observe (Toews & Brelsford, 2012). Instead, we suggest that extensive ILS, resulting from the explosive diversification and complex demographic history of the *Todiramphus* kingfishers (Andersen et al., 2015, 2018; McCullough et al., 2019), led to confidently supported topological differences between phylogenetic reconstructions from the nuclear and mitochondrial genomes.

A remaining question is whether our mitochondrial or nuclear species tree reconstruction is more likely to reflect the true bifurcating history of this group. Features of the mitochondrial genome including its haploid, non-recombining (Barr et al., 2005) nature and exclusively uni-parental mode of inheritance, lead to an effective population size ¼ that of a comparable nuclear locus (Birky et al., 1989). The smaller effective population size of the mitochondria means that it has a greater probability of coalescing during the speciation waiting period (i.e. lineage sorting) than any single nuclear locus (Hedrick, 2013; Peris et al., 2016). Yet, the recombining nature of the nuclear genome allows us to sample thousands of unlinked loci, which can effectively serve as independent replicates tracking the bifurcating species tree with greater combined accuracy than any individual locus, including the mitochondrial genome. All evidence considered, the most parsimonious explanation for the mitonuclear discordance we observe is deep coalescence of the mitochondrial genome, which is contradicted by a well-supported nuclear species tree reconstructed from thousands of independent genomic loci which are likely to accurately converge on the true bifurcating history of the group in the absence of gene flow and other non-tree-like processes. Determining exactly how confident we can be in supporting this nuclear reconstruction over the mitochondrial reconstruction depends on how many independent nuclear loci are sampled, and we believe this to be a fruitful avenue for future investigation in the realm of population genetic theory. Additional genomic investigations will be necessary to assess the frequency of mitonuclear discordance in other species radiations, and to better understand how the coalescent sorting process shapes patterns of relatedness between co-inherited cytonuclear genomes at macroevolutionary scales.

AUTHOR CONTRIBUTIONS

All authors collaboratively designed the research. DAD and LHD generated the RADseq data. PMH assembled and evaluated the *de novo* reference genome assembly for *T. albicilla*. DAD performed bioinformatic analyses on the RADseq data. DAD and JMM collaboratively wrote the first draft of the manuscript. All authors reviewed literature and contributed substantial feedback on each manuscript draft.

ACKNOWLEDGEMENTS

This work was supported by National Science Foundation grants to RGM (DEB-1557053) and MJA (DEB-1557051, DEB-2112467). All images reproduced with permission from Lynx Ediciones. We thank Mark Holder for guidance on species tree reconstruction approaches. We thank the KU Natural History Museum, the Australian National

Wildlife Collection, the Bernice Pauahi Bishop Museum, the American Museum of Natural History, the Western Australian Museum, the Smithsonian National Zoological Park, and the Burke Museum of Natural History and Culture for providing tissue samples necessary to complete this research. P.M.H. thanks the University of Kansas Biodiversity Institute and Natural History Museum for postdoctoral fellowship support. We thank the KU Genome Sequencing Core (supported by National Institutes of Health grant 5P20GM103638 to E.A. Lundquist) for access to lab equipment and services.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest related to this work.

OPEN RESEARCH BADGES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at Data Availability Statement for details.

DATA AVAILABILITY STATEMENT

All input data files, molecular laboratory protocols, and code necessary to reproduce the results presented herein are available at: https://github.com/DevonDeRaad/todiramphus.radseq. This repository is thoroughly documented, guiding the reader to subdirectories that show code and results in .html format for each of the analyses presented in this paper, allowing for maximal transparency and reproducibility. The entire repository is also permanently versioned and accessioned via Zenodo at: https://doi.org/10.5281/zenodo.8122439. Raw sequence data (RADseq) for each of the 83 individual samples used in analyses in this manuscript are available via the National Center for Biotechnology Information (NCBI BioProject accession PRJNA992577; https://www.ncbi.nlm.nih.gov/sra/PRJNA992577). Additionally, the finalized version of our *de novo* genome assembly for sample KU116001 is publicly available for download via NCBI (NCBI BioProject accession PRJNA992488; https://www.ncbi.nlm.nih.gov/bioproject/PRJNA992488), and the exact iteration of the reference genome used for all analyses presented here is permanently archived via Zenodo (https://doi.org/10.5281/zenodo.8178615).

ORCID

Devon A. DeRaad https://orcid.org/0000-0003-3105-985X
Paul M. Hime https://orcid.org/0000-0001-5322-4161
Leo Joseph https://orcid.org/0000-0001-7564-1978
Michael J. Andersen https://orcid.org/0000-0002-7220-5588

REFERENCES

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664.

Andersen, M. J., McCullough, J. M., Gyllenhaal, E. F., Mapel, X. M., Haryoko, T., Jønsson, K. A., & Joseph, L. (2021). Complex histories of gene flow and a mitochondrial capture event in a nonsister pair of birds. *Molecular Ecology*, *30*(9), 2087–2103. https://doi.org/10.1111/mec.15856

Andersen, M. J., McCullough, J. M., Mauck, W. M., III, Smith, B. T., & Moyle, R. G. (2018). A phylogeny of kingfishers reveals an Indomalayan origin and elevated rates of diversification on oceanic islands. *Journal of Biogeography*, *45*(2), 269–281. https://doi.org/10.1111/jbi.13139

Andersen, M. J., Naikatini, A., & Moyle, R. G. (2014). A molecular phylogeny of Pacific honeyeaters (Aves: Meliphagidae) reveals extensive paraphyly and an isolated Polynesian radiation. *Molecular Phylogenetics and Evolution*, *71*, 308–315. https://doi.org/10.1016/j.ympev.2013.11.014

Andersen, M. J., Nyári, Á. S., Mason, I., Joseph, L., Dumbacher, J. P., Filardi, C. E., & Moyle, R. G. (2014). Molecular systematics of the world's most polytypic bird: The *Pachycephala pectoralis/melanura* (Aves: Pachycephalidae) species complex. *Zoological Journal of the Linnean Society*, *170*(3), 566–588. https://doi.org/10.1111/zoj.12088

Andersen, M. J., Oliveros, C. H., Filardi, C. E., & Moyle, R. G. (2013). Phylogeography of the Variable Dwarf-Kingfisher *Ceyx lepidus* (Aves: Alcedinidae) inferred from mitochondrial and nuclear DNA sequences. *The Auk*, *130*(1), 118–131. https://doi.org/10.1525/auk.2012.12102

Andersen, M. J., Shult, H. T., Cibois, A., Thibault, J.-C., Filardi, C. E., & Moyle, R. G. (2015). Rapid diversification and secondary sympatry in Australo-Pacific kingfishers (Aves: Alcedinidae: *Todiramphus*). *Royal Society Open Science*, *2*(2), 140375. https://doi.org/10.1098/rsos.140375

Bailey, N. P., & Stevison, L. S. (2021). Mitonuclear conflict in a macaque species exhibiting phylogenomic discordance. *Journal of Evolutionary Biology*, *34*(10), 1568–1579. https://doi.org/10.1111/jeb.13914

Barr, C. M., Neiman, M., & Taylor, D. R. (2005). Inheritance and recombination of mitochondrial genomes in plants, fungi and animals. *New Phytologist*, *168*(1), 39–50. https://doi.org/10.1111/j.1469-8137.2005.01492.x

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Berbel-Filho, W. M., Pacheco, G., Tatarenkov, A., Lira, M. G., Garcia de Leaniz, C., Rodríguez López, C. M., Lima, S. M. Q., & Consuegra, S. (2022). Phylogenomics reveals extensive introgression and a case of mito-nuclear discordance in the killifish genus *Kryptolebias*. *Molecular Phylogenetics and Evolution*, *177*, 107617. https://doi.org/10.1016/j.ympev.2022.107617

Birky, C. W., Fuerst, P., & Maruyama, T. (1989). Organelle gene diversity under migration, mutation and drift: Equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics*, *121*, 613–627.

Bonnet, T., Leblois, R., Rousset, F., & Crochet, P.-A. (2017). A reassessment of explanations for discordant introgressions of mitochondrial and nuclear genomes. *Evolution*, *71*(9), 2140–2158. https://doi.org/10.1111/evo.13296

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M. A., Rambaut, A., & Drummond, A. J. (2014). Beast 2: A software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, *10*(4), e1003537. https://doi.org/10.1371/journal.pcbi.1003537

Bourret, T. B., Choudhury, R. A., Mehl, H. K., Blomquist, C. L., McRoberts, N., & Rizzo, D. M. (2018). Multiple origins of downy mildews and mito-nuclear discordance within the paraphyletic genus *Phytophthora*. *PLoS One*, *13*(3), e0192502. https://doi.org/10.1371/journal.pone.0192502

Bryant, D., Bouckaert, R., Felsenstein, J., Rosenberg, N. A., & RoyChoudhury, A. (2012). Inferring species trees directly from biallelic genetic markers: Bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution*, *29*(8), 1917–1932. https://doi.org/10.1093/molbev/mss086

Bryant, D., & Moulton, V. (2004). Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution*, *21*(2), 255–265. https://doi.org/10.1093/molbev/msh018

Carling, M. D., & Brumfield, R. T. (2008). Haldane's rule in an avian system: Using cline theory and divergence population genetics to test for differential introgression of mitochondrial, autosomal, and sex-linked loci across the Passerina bunting hybrid zone. *Evolution*, *62*(10), 2600–2615. https://doi.org/10.1111/j.1558-5646.2008.00477.x

Chernomor, O., von Haeseler, A., & Minh, B. Q. (2016). Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology*, *65*(6), 997–1008. https://doi.org/10.1093/sysbio/syw037

Degnan, J. H., & Rosenberg, N. A. (2006). Discordance of species trees with their most likely gene trees. *PLoS Genetics*, *2*(5), e68. https://doi.org/10.1371/journal.pgen.0020068

Degnan, J. H., & Salter, L. A. (2005). Gene tree distributions under the coalescent process. *Evolution*, *59*(1), 24–37. https://doi.org/10.1111/j.0014-3820.2005.tb00891.x

Del-Rio, G., Rego, M. A., Whitney, B. M., Schunck, F., Silveira, L. F., Faircloth, B. C., & Brumfield, R. T. (2022). Displaced clines in an avian hybrid zone (Thamnophilidae: *Rhegmatorhina*) within an Amazonian interfluve. *Evolution*, *76*(3), 455–475. https://doi.org/10.1111/evo.14377

Denton, R. D., Kenyon, L. J., Greenwald, K. R., & Gibbs, H. L. (2014). Evolutionary basis of mitonuclear discordance between sister species of mole salamanders (*Ambystoma* sp.). *Molecular Ecology*, *23*(11), 2811–2824. https://doi.org/10.1111/mec.12775

DeRaad, D. A. (2022). snpfiltR: An R package for interactive and reproducible SNP filtering. *Molecular Ecology Resources*, *22*(6), 2443–2453. https://doi.org/10.1111/1755-0998.13618

DeRaad, D. A., McCormack, J. E., Chen, N., Peterson, A. T., & Moyle, R. G. (2022). Combining species delimitation, species trees, and tests for gene flow clarifies complex speciation in scrub-jays. *Systematic Biology*, *71*(6), 1453–1470. https://doi.org/10.1093/sysbio/syac034

Després, L. (2019). One, two or more species? Mitonuclear discordance and species delimitation. *Molecular Ecology*, *28*(17), 3845–3847. https://doi.org/10.1111/mec.15211

Diamond, J. M., Gilpin, M. E., & Mayr, E. (1976). Species-distance relation for birds of the Solomon Archipelago, and the paradox of the great speciators. *Proceedings of the National Academy of Sciences of the United States of America*, *73*(6), 2160–2164. https://doi.org/10.1073/pnas.73.6.2160

Durand, E. Y., Patterson, N., Reich, D., & Slatkin, M. (2011). Testing for ancient admixture between closely related populations. *Molecular Biology and Evolution*, *28*(8), 2239–2252. https://doi.org/10.1093/molbev/msr048

Ellegren, H. (2010). Evolutionary stasis: The stable chromosomes of birds. *Trends in Ecology & Evolution*, *25*(5), 283–291. https://doi.org/10.1016/j.tree.2009.12.004

Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology*, *61*(5), 717–726. https://doi.org/10.1093/sysbio/sys004

Filardi, C. E., & Moyle, R. G. (2005). Single origin of a pan-Pacific bird group and upstream colonization of Australasia. *Nature*, *438*(7065), 216–219. https://doi.org/10.1038/nature04057

Firneno, T. J., O'Neill, J. R., Portik, D. M., Emery, A. H., Townsend, J. H., & Fujita, M. K. (2020). Finding complexity in complexes: Assessing the causes of mitonuclear discordance in a problematic species complex of Mesoamerican toads. *Molecular Ecology*, *29*(18), 3543–3559. https://doi.org/10.1111/mec.15496

Frankel, L. E., & Ané, C. (2023). Summary tests of introgression are highly sensitive to rate variation across lineages. *bioRxiv*. https://doi.org/10.1101/2023.01.26.525396

Ge, D., Feijó, A., Wen, Z., Lissovsky, A., Zhang, D., Cheng, J., Yan, C., Mu, D., Wu, X., Xia, L., & Yang, Q. (2022). Ancient introgression underlying the unusual mito-nuclear discordance and coat phenotypic variation in the Moupin pika. *Diversity and Distributions*, *28*, 2593–2609. https://doi.org/10.1111/ddi.13479

Gill, F. B., & Donsker, D. (2019). *IOC World Bird List. Version 9.1*. International Ornithologists' Union.

Gompert, Z., Forister, M. L., Fordyce, J. A., & Nice, C. C. (2008). Widespread mito-nuclear discordance with evidence for introgressive hybridization and selective sweeps in *Lycaeides*. *Molecular Ecology*, *17*(24), 5231–5244. https://doi.org/10.1111/j.1365-294X.2008.03988.x

Good, J. M., Hird, S., Reid, N., Demboski, J. R., Steppan, S. J., Martin-Nims, T. R., & Sullivan, J. (2008). Ancient hybridization and mitochondrial capture between two species of chipmunks. *Molecular Ecology*, *17*(5), 1313–1327. https://doi.org/10.1111/j.1365-294X.2007.03640.x

Grabherr, M. G., Russell, P., Meyer, M., Mauceli, E., Alfoldi, J., Di Palma, F., & Lindblad-Toh, K. (2010). Genome-wide synteny through highly sensitive sequence alignment: *Satsuma*. *Bioinformatics*, *26*(9), 1145–1151. https://doi.org/10.1093/bioinformatics/btq102

Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M. H., Hansen, N. F., Durand, E. Y., Malaspinas, A. S., Jensen, J. D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H. A., … Pääbo, S. (2010). A draft sequence of the Neandertal genome. *Science*, *328*(5979), 710–722. https://doi.org/10.1126/science.1188021

Harrower, M., & Brewer, C. A. (2003). ColorBrewer.org: An online tool for selecting color schemes for maps. *The Cartographic Journal*, *40*(1), 27–37. https://doi.org/10.1179/000870403235002042

Hedrick, P. W. (2013). Adaptive introgression in animals: Examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology*, *22*(18), 4606–4618. https://doi.org/10.1111/mec.12415

Hibbins, M. S., & Hahn, M. W. (2022). Phylogenomic approaches to detecting and characterizing introgression. *Genetics*, *220*(2), iyab173. https://doi.org/10.1093/genetics/iyab173

Hinojosa, J. C., Koubínová, D., Szenteczki, M. A., Pitteloud, C., Dincă, V., Alvarez, N., & Vila, R. (2019). A mirage of cryptic species: Genomics uncover striking mitonuclear discordance in the butterfly *Thymelicus sylvestris*. *Molecular Ecology*, *28*(17), 3857–3868. https://doi.org/10.1111/mec.15153

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the ultrafast bootstrap approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. https://doi.org/10.1093/molbev/msx281

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution*, *23*(2), 254–267. https://doi.org/10.1093/molbev/msj030

Ivanov, V., Lee, K. M., & Mutanen, M. (2018). Mitonuclear discordance in wolf spiders: Genomic evidence for species integrity and introgression. *Molecular Ecology*, *27*(7), 1681–1695. https://doi.org/10.1111/mec.14564

Joseph, L., & Moritz, C. (1993). Hybridization between the white-browed and atherton scrubwrens: Detection with mitochondrial DNA. *Emu*, *93*(2), 93–99. https://doi.org/10.1071/mu9930093

Joseph, L., Campbell, C. D., Drew, A., Brady, S. S., Nyári, Á., & Andersen, M. J. (2021). How far east can a Western Whistler go? Genomic data reveal large eastward range extension, taxonomic and nomenclatural change, and reassessment of conservation needs. *Emu*, *121*(1–2), 90–101. https://doi.org/10.1080/01584197.2020.1854047

Joseph, L., Dolman, G., Iova, B., Jønsson, K., Campbell, C. D., Mason, I., & Drew, A. (2019). Aberrantly plumaged orioles from the Trans-Fly savannas of New Guinea and their ecological and evolutionary significance. *Emu*, *119*(3), 264–273. https://doi.org/10.1080/01584197.2019.1605831

Kao, T.-T., Wang, T.-H., & Ku, C. (2022). Rampant nuclear–mitochondrial–plastid phylogenomic discordance in globally distributed

calcifying microalgae. *New Phytologist*, *235*(4), 1394–1408. https://doi.org/10.1111/nph.18219

Kato, H., Cáceres, A. G., Gomez, E. A., Tabbabi, A., Mizushima, D., Yamamoto, D. S., & Hashiguchi, Y. (2021). Prevalence of genetically complex *Leishmania* strains with hybrid and mito-nuclear discordance. *Frontiers in Cellular and Infection Microbiology*, *11*, 625001. https://doi.org/10.3389/fcimb.2021.625001

Kato, S., Arakaki, S., Kikuchi, K., & Hirase, S. (2021). Complex phylogeographic patterns in the intertidal goby *Chaenogobius annularis* around Kyushu Island as a boundary zone of three different seas. *Ichthyological Research*, *68*(1), 86–100. https://doi.org/10.1007/s10228-020-00772-4

Kearns, A. M., Joseph, L., Toon, A., & Cook, L. G. (2014). Australia's arid-adapted butcherbirds experienced range expansions during Pleistocene glacial maxima. *Nature Communications*, *5*(1), 3994. https://doi.org/10.1038/ncomms4994

Kearns, A. M., Restani, M., Szabo, I., Schrøder-Nielsen, A., Kim, J. A., Richardson, H. M., Marzluff, J. M., Fleischer, R. C., Johnsen, A., & Omland, K. E. (2018). Genomic evidence of speciation reversal in ravens. *Nature Communications*, *9*(1), 906. https://doi.org/10.1038/s41467-018-03294-w

Kimball, R. T., Guido, M., Hosner, P. A., & Braun, E. L. (2021). When good mitochondria go bad: Cyto-nuclear discordance in landfowl (Aves: Galliformes). *Gene*, *801*, 145841. https://doi.org/10.1016/j.gene.2021.145841

Knaus, B. J., & Grünwald, N. J. (2017). vcfr: A package to manipulate and visualize variant call format data in R. *Molecular Ecology Resources*, *17*(1), 44–53. https://doi.org/10.1111/1755-0998.12549

Kubatko, L. S., & Degnan, J. H. (2007). Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Systematic Biology*, *56*(1), 17–24. https://doi.org/10.1080/10635150601146041

Kutschera, V. E., Bidon, T., Hailer, F., Rodi, J. L., Fain, S. R., & Janke, A. (2014). Bears in a forest of gene trees: Phylogenetic inference is complicated by incomplete lineage sorting and gene flow. *Molecular Biology and Evolution*, *31*(8), 2004–2017. https://doi.org/10.1093/molbev/msu186

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lopes, F., Oliveira, L. R., Kessler, A., Beux, Y., Crespo, E., Cárdenas-Alayza, S., Majluf, P., Sepúlveda, M., Brownell, R. L., Franco-Trecu, V., Páez-Rosas, D., Chaves, J., Loch, C., Robertson, B. C., Acevedo-Whitehouse, K., Elorriaga-Verplancken, F. R., Kirkman, S. P., Peart, C. R., Wolf, J. B. W., & Bonatto, S. L. (2021). Phylogenomic discordance in the eared seals is best explained by incomplete lineage sorting following explosive radiation in the southern hemisphere. *Systematic Biology*, *70*(4), 786–802. https://doi.org/10.1093/sysbio/syaa099

Maddison, W. P., & Knowles, L. L. (2006). Inferring phylogeny despite incomplete lineage sorting. *Systematic Biology*, *55*(1), 21–30. https://doi.org/10.1080/10635150500354928

Malinsky, M., Matschiner, M., & Svardal, H. (2021). Dsuite—Fast D-statistics and related admixture evidence from VCF files. *Molecular Ecology Resources*, *21*(2), 584–595. https://doi.org/10.1111/1755-0998.13265

Manthey, J. D., & Moyle, R. G. (2015). Isolation by environment in White-breasted Nuthatches (*Sitta carolinensis*) of the Madrean Archipelago sky islands: A landscape genomics approach. *Molecular Ecology*, *24*(14), 3628–3638. https://doi.org/10.1111/mec.13258

Mao, X., & Rossiter, S. J. (2020). Genome-wide data reveal discordant mitonuclear introgression in the intermediate horseshoe bat (*Rhinolophus affinis*). *Molecular Phylogenetics and Evolution*, *150*, 106886. https://doi.org/10.1016/j.ympev.2020.106886

Marshall, T. L., Chambers, E. A., Matz, M. V., & Hillis, D. M. (2021). How mitonuclear discordance and geographic variation have confounded species boundaries in a widely studied snake. *Molecular Phylogenetics and Evolution*, *162*, 107194. https://doi.org/10.1016/j.ympev.2021.107194

McCormack, J. E., Huang, H., & Knowles, L. L. (2009). Maximum likelihood estimates of species trees: How accuracy of phylogenetic inference depends upon the divergence history and sampling design. *Systematic Biology*, *58*(5), 501–508. https://doi.org/10.1093/sysbio/syp045

McCullough, J. M., Moyle, R. G., Smith, B. T., & Andersen, M. J. (2019). A Laurasian origin for a pantropical bird radiation is supported by genomic and fossil data (Aves: Coraciiformes). *Proceedings of the Royal Society B: Biological Sciences*, *286*(1910), 20190122. https://doi.org/10.1098/rspb.2019.0122

McElroy, K., Black, A., Dolman, G., Horton, P., Pedler, L., Campbell, C. D., Drew, A., & Joseph, L. (2020). Robbery in progress: Historical museum collections bring to light a mitochondrial capture within a bird species widespread across southern Australia, the Copperback Quail-thrush *Cinclosoma clarum*. *Ecology and Evolution*, *10*(13), 6785–6793. https://doi.org/10.1002/ece3.6403

McKay, B. D., & Zink, R. M. (2010). The causes of mitochondrial DNA gene tree paraphyly in birds. *Molecular Phylogenetics and Evolution*, *54*(2), 647–650. https://doi.org/10.1016/j.ympev.2009.08.024

Mendes, F. K., & Hahn, M. W. (2016). Gene tree discordance causes apparent substitution rate variation. *Systematic Biology*, *65*(4), 711–721. https://doi.org/10.1093/sysbio/syw018

Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D., & Gurevich, A. (2018). Versatile genome assembly evaluation with QUAST-LG. *Bioinformatics*, *34*(13), i142–i150. https://doi.org/10.1093/bioinformatics/bty266

Mikkelsen, E. K., & Weir, J. T. (2022). Phylogenomics reveals that mitochondrial capture and nuclear introgression characterizes skua species proposed to be of hybrid origin. *Systematic Biology*, *72*, 78–91. https://doi.org/10.1093/sysbio/syac078

Minh, B. Q., Hahn, M. W., & Lanfear, R. (2020). New methods to calculate concordance factors for phylogenomic datasets. *Molecular Biology and Evolution*, *37*(9), 2727–2733. https://doi.org/10.1093/molbev/msaa106

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., von Haeseler, A., & Lanfear, R. (2020). IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Molecular Biology and Evolution*, *37*(5), 1530–1534. https://doi.org/10.1093/molbev/msaa015

Mirarab, S., Bayzid, M. S., & Warnow, T. (2016). Evaluating summary methods for multilocus species tree estimation in the presence of incomplete lineage sorting. *Systematic Biology*, *65*(3), 366–380. https://doi.org/10.1093/sysbio/syu063

Morales, H. E., Pavlova, A., Joseph, L., & Sunnucks, P. (2015). Positive and purifying selection in mitochondrial genomes of a bird with mitonuclear discordance. *Molecular Ecology*, *24*(11), 2820–2837. https://doi.org/10.1111/mec.13203

Moyle, R. G., Filardi, C. E., Smith, C. E., & Diamond, J. (2009). Explosive Pleistocene diversification and hemispheric expansion of a "great speciator". *Proceedings of the National Academy of Sciences of the United States of America*, *106*(6), 1863–1868.

Moyle, R. G., Oliveros, C. H., Andersen, M. J., Hosner, P. A., Benz, B. W., Manthey, J. D., Travers, S. L., Brown, R. M., & Faircloth, B. C. (2016). Tectonic collision and uplift of Wallacea triggered the global songbird radiation. *Nature Communications*, *7*, 12709. https://doi.org/10.1038/ncomms12709

Nei, M. (1972). Genetic distance between populations. *The American Naturalist*, *106*, 283–292. https://doi.org/10.1086/282771

Orr, H. A. (1997). Haldane's rule. *Annual Review of Ecology and Systematics*, *28*(1), 195–218. https://doi.org/10.1146/annurev.ecolsys.28.1.195

Pembleton, L. W., Cogan, N. O. I., & Forster, J. W. (2013). StAMPP: An R package for calculation of genetic differentiation and structure of

mixed-ploidy level populations. *Molecular Ecology Resources*, *13*(5), 946–952. https://doi.org/10.1111/1755-0998.12129

Peris, D., Langdon, Q. K., Moriarty, R. V., Sylvester, K., Bontrager, M., Charron, G., Leducq, J. B., Landry, C. R., Libkind, D., & Hittinger, C. T. (2016). Complex ancestries of lager-brewing hybrids were shaped by standing variation in the wild yeast *Saccharomyces eubayanus*. *PLoS Genetics*, *12*(7), e1006155. https://doi.org/10.1371/journal.pgen.1006155

Phuong, M. A., Bi, K., & Moritz, C. (2017). Range instability leads to cytonuclear discordance in a morphologically cryptic ground squirrel species complex. *Molecular Ecology*, *26*(18), 4743–4755. https://doi.org/10.1111/mec.14238

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genetics*, *8*(11), e1002967. https://doi.org/10.1371/journal.pgen.1002967

Pons, J.-M., Sonsthagen, S., Dove, C., & Crochet, P.-A. (2014). Extensive mitochondrial introgression in North American Great Black-backed Gulls (*Larus marinus*) from the American Herring Gull (*Larus smithsonianus*) with little nuclear DNA impact. *Heredity*, *112*(3), 226–239. https://doi.org/10.1038/hdy.2013.98

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/

Rambaut, A. (2014). *FigTree, a graphical viewer of phylogenetic trees*. Retrieved August 4, 2020, from http://tree.bio.ed.ac.uk/software/figtree/

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, *67*(5), 901–904. https://doi.org/10.1093/sysbio/syy032

Rochette, N. C., Rivera-Colón, A. G., & Catchen, J. M. (2019). Stacks 2: Analytical methods for paired-end sequencing improve RADseq-based population genomics. *Molecular Ecology*, *28*(21), 4737–4754. https://doi.org/10.1111/mec.15253

Rohland, N., & Reich, D. (2012). Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Research*, *22*(5), 939–946. https://doi.org/10.1101/gr.128124.111

Scherz, M. D., Masonick, P., Meyer, A., & Hulsey, C. D. (2022). Between a rock and a hard polytomy: Phylogenomics of the rock-dwelling mbuna cichlids of Lake Malaŵi. *Systematic Biology*, *71*(3), 741–757. https://doi.org/10.1093/sysbio/syac006

Schliep, K. P. (2011). phangorn: Phylogenetic analysis in R. *Bioinformatics*, *27*(4), 592–593. https://doi.org/10.1093/bioinformatics/btq706

Schrago, C. G., & Seuánez, H. N. (2019). Large ancestral effective population size explains the difficult phylogenetic placement of owl monkeys. *American Journal of Primatology*, *81*(3), e22955. https://doi.org/10.1002/ajp.22955

Scornavacca, C., & Galtier, N. (2017). Incomplete lineage sorting in mammalian phylogenomics. *Systematic Biology*, *66*(1), 112–120. https://doi.org/10.1093/sysbio/syw082

Seixas, F. A., Boursot, P., & Melo-Ferreira, J. (2018). The genomic impact of historical hybridization with massive mitochondrial DNA introgression. *Genome Biology*, *19*(1), 91. https://doi.org/10.1186/s13059-018-1471-8

Shafer, A. B. A., Peart, C. R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C. W., & Wolf, J. B. W. (2017). Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution*, *8*(8), 907–917. https://doi.org/10.1111/2041-210X.12700

Shahdadi, A., Mvogo Ndongo, P. A., & Schubart, C. D. (2021). Mito-nuclear discordance in West African mangrove crab species (Decapoda: Brachyura: Sesarmidae) suggests uni-directional mitochondrial introgression, despite prolonged evolutionary independence. *Marine Biology Research*, *17*(5–6), 503–512. https://doi.org/10.1080/17451000.2021.1990959

Shu, X.-X., Hou, Y.-M., Cheng, M.-Y., Shu, G.-C., Lin, X.-Q., Wang, B., Li, C., Song, Z. B., Jiang, J. P., & Xie, F. (2022). Rapid genetic divergence and mitonuclear discordance in the Taliang knobby newt (*Liangshantriton taliangensis*, Salamandridae, Caudata) and their driving forces. *Zoological Research*, *43*(1), 129–146. https://doi.org/10.24272/j.issn.2095-8137.2021.299

Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

Slager, D. L., Epperly, K. L., Ha, R. R., Rohwer, S., Wood, C., Van Hemert, C., & Klicka, J. (2020). Cryptic and extensive hybridization between ancient lineages of American crows. *Molecular Ecology*, *29*(5), 956–969. https://doi.org/10.1111/mec.15377

Suh, A., Smeds, L., & Ellegren, H. (2015). The dynamics of incomplete lineage sorting across the ancient adaptive radiation of neoavian birds. *PLoS Biology*, *13*(8), e1002224. https://doi.org/10.1371/journal.pbio.1002224

Toews, D. P. L., & Brelsford, A. (2012). The biogeography of mitochondrial and nuclear discordance in animals. *Molecular Ecology*, *21*(16), 3907–3930. https://doi.org/10.1111/j.1365-294X.2012.05664.x

Wang, K., Lenstra, J. A., Liu, L., Hu, Q., Ma, T., Qiu, Q., & Liu, J. (2018). Incomplete lineage sorting rather than hybridization explains the inconsistent phylogeny of the wisent. *Communications Biology*, *1*(1), 1–9. https://doi.org/10.1038/s42003-018-0176-6

Weir, J. T., & Schluter, D. (2008). Calibrating the avian molecular clock. *Molecular Ecology*, *17*(10), 2321–2328. https://doi.org/10.1111/j.1365-294X.2008.03742.x

Weisenfeld, N. I., Kumar, V., Shah, P., Church, D. M., & Jaffe, D. B. (2017). Direct determination of diploid genome sequences. *Genome Research*, *27*(5), 757–767. https://doi.org/10.1101/gr.214874.116

Weisrock, D. W. (2012). Concordance analysis in mitogenomic phylogenetics. *Molecular Phylogenetics and Evolution*, *65*(1), 194–202. https://doi.org/10.1016/j.ympev.2012.06.003

Wham, B. E., Rahman, S. R., Martinez-Correa, M., & Hines, H. M. (2021). Mito-nuclear discordance at a mimicry color transition zone in bumble bee *Bombus melanopygus*. *Ecology and Evolution*, *11*(24), 18151–18168. https://doi.org/10.1002/ece3.8412

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., Yutani, H., & Dunnington, D. (2020). *ggplot2: Create elegant data visualisations using the grammar of graphics*. R package, version 3.3.2. Retrieved from https://CRAN.R-project.org/package=ggplot2

Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, *19*(6), 153. https://doi.org/10.1186/s12859-018-2129-y

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: DeRaad, D. A., McCullough, J. M., DeCicco, L. H., Hime, P. M., Joseph, L., Andersen, M. J., & Moyle, R. G. (2023). Mitonuclear discordance results from incomplete lineage sorting, with no detectable evidence for gene flow, in a rapid radiation of *Todiramphus* kingfishers. *Molecular Ecology*, *00*, 1–19. https://doi.org/10.1111/mec.17080