Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Accelerated sex chromosome degeneration and mitonuclear coevolution in a brood parasitic bird

Michael Sorenson

msoren@bu.edu

Boston University https://orcid.org/0000-0001-5375-2917

Ekaterina Osipova

Harvard University https://orcid.org/0000-0002-6769-7223

Claire Spottiswoode

University of Cape Town

Gabriel Jamie

University of Cape Town

Tanmay Dixit

University of Cambridge https://orcid.org/0000-0001-5604-7965

Jeffery DaCosta

Boston College

Timothy Sackton

Harvard University https://orcid.org/0000-0003-1673-9216

Wesley Warren

Bond Life Sciences Center, University of Missouri https://orcid.org/0000-0003-2255-2730

Christopher Balakrishnan

National Science Foundation https://orcid.org/0000-0002-0788-0659

Article

Keywords:

Posted Date: September 12th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7330189/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: There is **NO** Competing Interest.

Version of Record: A version of this preprint was published at Nature Ecology & Evolution on October 27th, 2026. See the published version at https://doi.org/10.1038/s41559-026-03216-w.

Sex chromosome degeneration and mitonuclear coevolution in a parasitic bird

Michael D. Sorenson[1], Ekaterina Osipova[2], Claire N. Spottiswoode[3,4], Gabriel A. Jamie[3,4], Tanmay Dixit[3,4], Jeffrey M. DaCosta[5], Timothy B. Sackton[2], Wesley C. Warren[6,7,8], Christopher N. Balakrishnan[9,10]

[1]Department of Biology, Boston University, Boston, MA, USA. [2]Informatics Group, Harvard University, Cambridge, MA, USA. [3]FitzPatrick Institute of African Ornithology, Department of Biological Sciences, University of Cape Town, Cape Town, South Africa. [4]Department of Zoology, University of Cambridge, Cambridge, UK. [5]Biology Department, Boston College, Chestnut Hill, MA, USA. [6]Division of Animal Sciences, Bond Life Sciences Center, University of Missouri, Columbia, MO, USA. [7]Department of Surgery, School of Medicine, University of Missouri, Columbia, MO, USA. [8]Institute of Data Science and Informatics, University of Missouri, Columbia, MO, USA. [9]Department of Biology, East Carolina University, Greenville, NC, USA. [10]Division of Environmental Biology, National Science Foundation, Alexandria, VA, USA.

The coevolution of interacting mitochondrial and nuclear genes involved in the oxidative phosphorylation (OXPHOS) pathway has received growing attention as a critical component of organismal fitness[1] and potential source of genetic incompatibilities[2–4] that promote speciation[5–7]. However, in species with "ZW" sex chromosomes, the mitochondrial genome and female-specific W chromosome are effectively linked by their shared matrilineal history, such that selection on either should subject both to the accumulation of deleterious mutations via genetic hitchhiking[8,9]. Thus, enhanced ecological selection on maternally inherited traits should accelerate both W chromosome degeneration and mitonuclear coevolution, though this prediction has not been tested empirically. Here we show that brood parasitic cuckoo finches *Anomalospiza imberbis*, in which mimetic egg coloration is maternally inherited, have experienced accelerated W chromosome pseudogenization and gene loss in addition to higher rates of amino acid substitution in both their mitochondrial and nuclear-encoded OXPHOS genes. Using a structural model of respiratory complex I, we also show that the degree of acceleration in nuclear-encoded

residues is a function of their proximity to mtDNA-encoded residues, providing strong evidence of compensatory mitonuclear coevolution. Our results link ecological coevolution at the species level to both sex chromosome degeneration and mitonuclear coevolution, highlighting genomic consequences of selection on non-recombining genome components.

Genetic recombination facilitates adaptation both by generating new combinations of genes and by breaking linkage between beneficial and deleterious mutations, allowing positive and purifying selection to proceed simultaneously[10]. When recombination is absent, positive selection can lead to a counterintuitive accumulation of deleterious variants that "hitchhike" along with new, advantageous mutations[11,12]. Indeed, genetic hitchhiking has long been suggested as an explanation for the gradual degeneration of the non-recombining and male-limited Y chromosome[13,14]. The same logic applies to degeneration of the female-limited W chromosome in species with "ZW" sex chromosomes[15], though the effective linkage between the W chromosome and mitochondrial genome (mtDNA) introduces additional complexity[8,9]. Ecological circumstances that increase the intensity of selection on either of these female-limited genome components should accelerate both W chromosome degeneration and the accumulation of slightly deleterious mtDNA mutations, with further consequences extending to nuclear-encoded OXPHOS proteins that interact with mitochondrial gene products[9,16,17], including transfer-RNAs, ribosomal-RNAs and proteins. We tested for these predicted effects in an avian system with clear evidence of diversifying selection on maternally inherited traits.

The cuckoo finch *Anomalospiza imberbis* is an obligate brood parasite that exploits warblers in the family Cisticolidae. The evolution of egg recognition and rejection in the warbler hosts, as a defense against parasitism, sparked an evolutionary arms race, leading in turn to the evolution of mimetic eggs in cuckoo finches, divergence in egg phenotypes among sympatric host species, intraspecific polymorphism in host egg phenotypes (i.e., individualized egg "signatures"), and egg polymorphisms in cuckoo finches that approximate the range of host phenotypes[18-21]. Evidence of ongoing selection over ecological timescales suggests a highly dynamic system[22,23]. Remarkably, it is a single cuckoo finch species — *Anomalospiza* is monotypic — that is simultaneously engaged in arms races with multiple, sympatric host species[21,24]. This is made possible by maternal inheritance of egg phenotype combined with long-term fidelity of matrilines to their respective hosts[21]. Thus, two highly divergent cuckoo finch matrilines,

63 characterized by divergent W chromosomes and divergent mtDNA, are associated with two
64 different host genera (*Prinia* and *Cisticola*), and within each of these major lineages, there has
65 been additional matrilineal diversification of egg phenotypes as cuckoo finches coevolved with
66 different host species[21]. However, this coevolution with multiple hosts has not led to speciation;
67 a complete lack of differentiation in the autosomal genome indicates that cuckoo finch males and
68 females mate randomly with respect to the host species that reared them[21]. Thus, as proposed by
69 Punnett[25], maternal inheritance appears to have solved, for cuckoo finches, the problem of
70 evolving and maintaining host-specific egg mimicry when males and females reared by different
71 hosts interbreed.
72 This solution, however, carries potential costs associated with the lack of recombination in
73 maternally inherited components of the genome. First, future adaptive potential may be reduced
74 as novel phenotypes must arise through de novo mutation rather than new combinations of
75 existing alleles; this may explain the apparent failure of cuckoo finches parasitizing tawny-
76 flanked prinia *Prinia subflava* to evolve olive-colored eggs[21] or fully match the complexity of
77 host egg markings[23]. Second and the focus of this paper, repeated bouts of selection associated
78 with ongoing host-parasite arms races should have subjected cuckoo finches to the consequences
79 of genetic hitchhiking posited above. Specifically, cuckoo finches, in comparison to related
80 species, should have experienced: 1) a higher rate of W chromosome pseudogenization and gene
81 loss; 2) an elevated rate of nonsynonymous substitution in remaining W-linked protein-coding
82 genes; and 3) a faster rate of mtDNA evolution, including relatively more nonsynonymous
83 change in protein-coding genes; all consequences of genetic hitchhiking. Further, as
84 mitochondrial genes function in the OXPHOS pathway: 4) nuclear genes encoding proteins that
85 directly interact with mitochondrial gene products should also show evidence of accelerated
86 evolution in cuckoo finch, reflecting compensatory evolution to maintain OXPHOS function[16,17].
87 We tested these *a priori* predictions about the consequences of genetic hitchhiking using new
88 genome assemblies and genome resequencing data to compare cuckoo finches to their sister
89 lineage, the brood parasitic indigobirds and whydahs (*Vidua* spp.). In *Vidua* species, which lay
90 immaculate white eggs, both male and female offspring mimic the mouth colors and patterns of
91 their respective hosts[26,27]. Expression in both sexes rules out maternal inheritance of this key
92 host-specific adaptation in *Vidua*, providing a clear contrast to the female-limited and maternally
93 inherited phenotype of egg coloration in cuckoo finches.

Accelerated degeneration of the sex-limited chromosome

As predicted, we found that pseudogenization and loss of W chromosome protein-coding genes has been more frequent in the cuckoo finch lineage than in its sister group, the brood parasitic *Vidua* (Fig. 1), in which there is no evidence of maternal inheritance of host-specific adaptations; the difference is striking, with 17 genes experiencing complete loss, pseudogenization or changes incompatible with function in the cuckoo finch lineage compared to just three in the *Vidua* lineage. This results in cuckoo finches having the smallest number of intact W-linked protein-coding genes among the eight species in our analysis (Supplementary Tables 1 and 2). In addition, pseudogenes originating from earlier losses of function are more likely to have been further reduced or lost completely in cuckoo finch than in *Vidua* (Extended Data Fig. 1). Comparisons of sequencing depth among cuckoo finch lineages reveal numerous W chromosome deletions ranging up to ~400 kbp in length, including deletions that coincide with recent losses or reductions of pseudogenes (Extended Data Fig. 2). The *Prinia*-associated cuckoo finch matriline, which is relatively divergent from the reference genome sample (Extended Data Fig. 3), has lost ~1.4 Mbp of W chromosome sequence in over 100 deletions of 1 kbp or more since its divergence from the *Cisticola*-associated matriline. Rare gene conversion events might restore functionality to a W-linked pseudogene[28,29] but we detected just one case of Z to W chromosome gene conversion, and this was in the *Vidua* lineage (Extended Data Fig. 4).

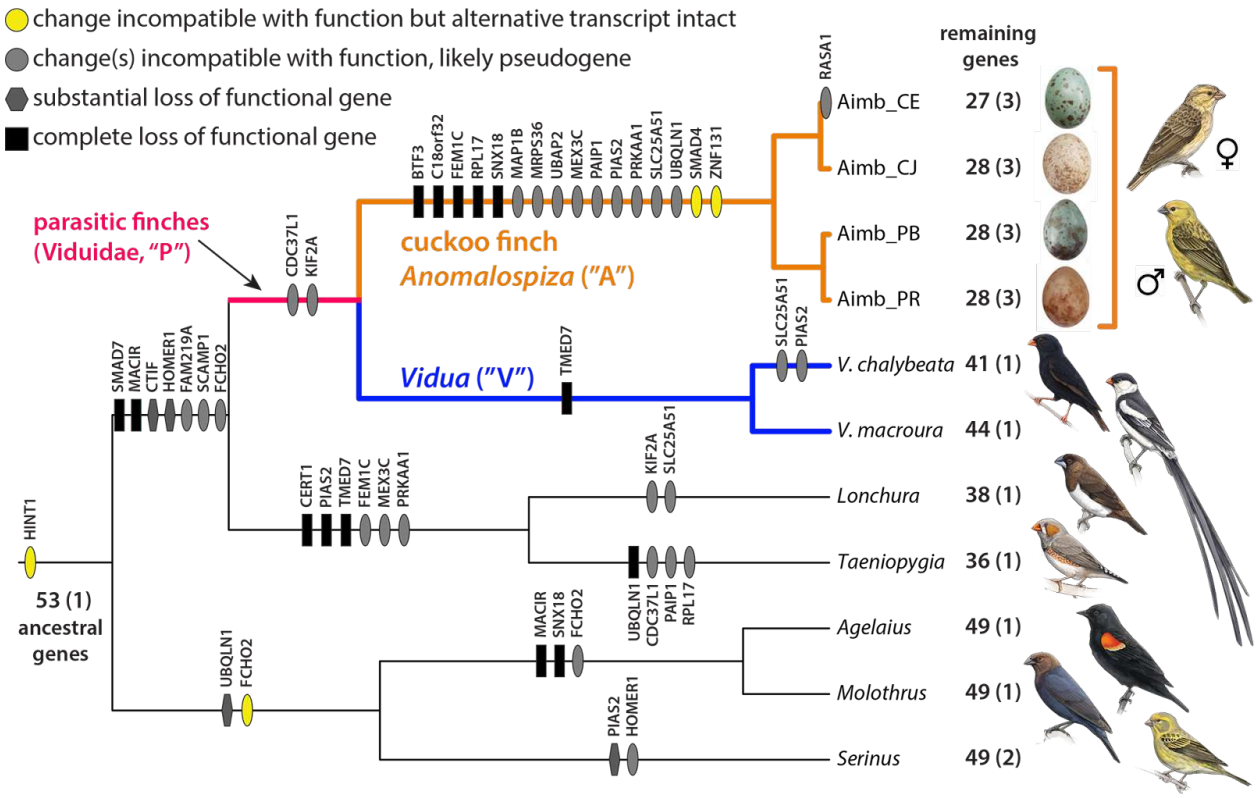


Fig. 1 | Dollo parsimony reconstruction of W-linked gene loss and pseudogenization in parasitic finches and other songbirds. From top, the taxa include: four host-specific cuckoo finch matrilines ("CE" = parasitizing *Cisticola erythrops*; "CJ" = parasitizing *C. juncidis*; "PB" and "PR" = parasitizing *Prinia subflava* and laying eggs with blue and red background color, respectively); two parasitic finches without maternally inherited, host-specific adaptations (village indigobird *V. chalybeata*; pin-tailed whydah *V. macroura*); and as outgroups: two estrildid finches (white-rumped munia *Lonchura striata*; zebra finch *Taeniopygia guttata*), red-winged blackbird *Agelaius phoeniceus*, brown-headed cowbird *Molothrus ater* (a host generalist brood parasite), and canary *Serinus canaria*. Reconstructed gene loss and pseudogenization events are indicated on branches. Pseudogenes are scored as "substantially lost" if < 35% of the original codons remain. Bold numbers indicate the number of intact, putatively functional W-linked protein-coding genes in the common ancestor and each terminal taxon, with the number of additional genes that may remain functional despite putative loss-of-function mutations indicated in parentheses (see Supplementary Table 1 for details). The underlying phylogeny is based on the concatenated coding regions of 29 W-linked genes present in all taxa. Color-coding and labeling of key lineages (with "P" denoting the entire parasitic finch clade = *Anomalospiza* plus *Vidua*) is maintained in subsequent figures. Images of male and female cuckoo finches along with males of the remaining species by Javier Lazaro.

Also as predicted, surviving W-linked protein coding genes have experienced a faster rate of amino acid substitution in the cuckoo finch lineage than in *Vidua* (Fig. 2a). Across 29 W-linked genes that remain putatively functional in all eight species analyzed, *dN* and *dN/dS*, respectively, have been 87% and 46% higher in the cuckoo finch lineage than in *Vidua* (Supplementary Table 3). For these W-linked genes combined, the best model for *dN/dS* rates across the phylogeny is

one in which cuckoo finches differ from all other lineages in the tree (ΔAICc = 18.3, in comparison to a null model with constant *dN/dS* across the phylogeny; see Extended Data Fig. 5 for alternative models tested). Moreover, amino acid substitutions in the cuckoo finch lineage more often involve amino acids with dissimilar chemical properties ($D = 0.174$, $p = 0.0149$; Extended Data Fig. 6a), suggesting a higher rate of potentially deleterious substitutions than in *Vidua*.

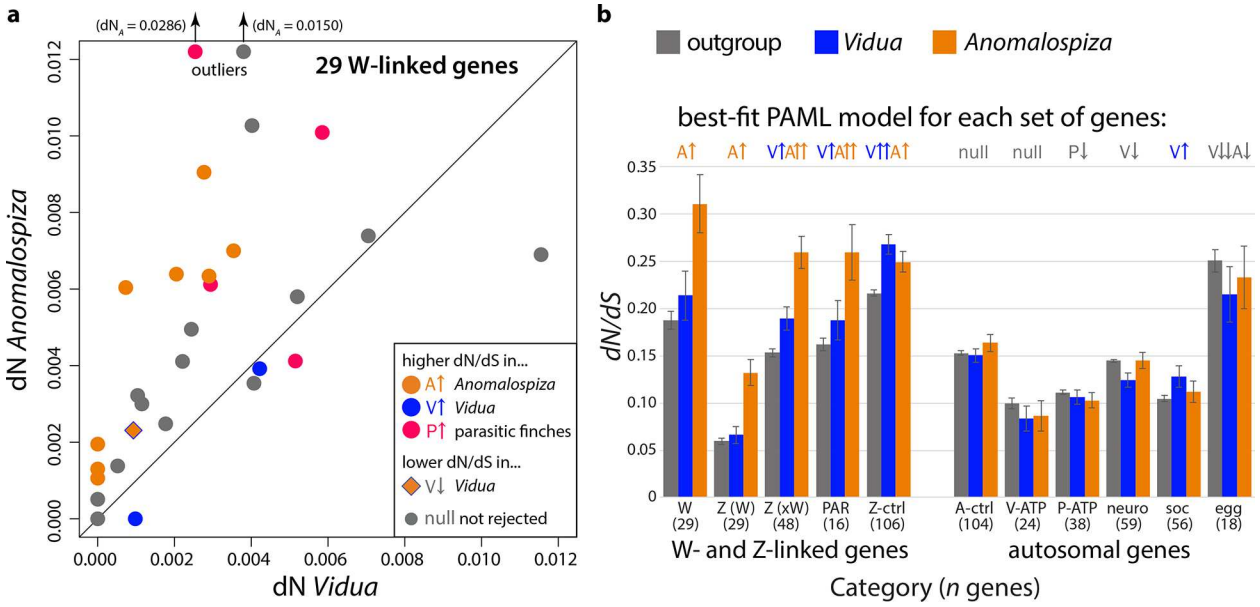


Fig. 2 | Comparisons of evolutionary rate in sex-linked and autosomal genes. a, Rates of nonsynonymous substitution (*dN*) in cuckoo finch (*Anomalospiza*) and its sister lineage (*Vidua* spp.) for 29 W-linked genes. Point color and shape indicate the best model for *dN/dS* for each gene, highlighting significant differences relative to the outgroups (see Extended Data Fig. 5). Note that statistical power increases with gene length and number of variable sites. **b**, *dN/dS* estimates (± SE) for *Vidua*, *Anomalospiza* and outgroup taxa for different sets of sex-linked and autosomal genes, including: 29 W-linked genes ("W"); the Z-linked gametologs of these 29 genes ("Z (W)"); 48 Z-linked genes with lost or non-functional W-linked gametologs in one or more of the eight species analyzed ("Z (xW)"); 16 genes in the pseudoautosomal region ("PAR"); and a control set of 106 Z-linked genes not known to have a W-linked gametolog in any extant songbird ("Z-ctrl"). See Supplementary Table 3 for details on the autosomal gene sets. The best fit model for each gene set is shown above the plots. For example: "A↑" = *Anomalospiza* has a higher rate than both *Vidua* and the outgroup; "V↑A↑↑" = *Vidua* and *Anomalospiza* have higher rates than the outgroup but the rate in *Anomalospiza* is higher than in *Vidua*; "P↓" = parasitic finches combined (*Anomalospiza* plus *Vidua*) have a lower rate than the outgroup.

While we made no a priori predictions for the Z chromosome, we found evidence of rate differences among subsets of Z-linked genes. However, in comparison to the W chromosome results, the effects are smaller (Fig. 2b) and more variable across individual genes (Extended Data Fig. 7). For the set of 29 Z-linked genes with putatively functional W-linked gametologs in

all eight species analyzed, *dN/dS* was significantly higher in cuckoo finch than in *Vidua*, though these genes are generally more conserved than other Z-linked genes (Fig. 2b). Results were similar, but less extreme, for a set of 48 Z-linked genes with W-linked gametologs that have been lost or are no longer functional in one or more species (Fig. 2b). These results may reflect the ongoing evolution of dosage compensation and/or sex-specific expression[30–32] in response to the loss of functional W-linked gametologs. In contrast, *dN/dS* is lower in *Anomalospiza* than in *Vidua* for a representative set of other Z-linked genes ("Z-ctrl" in Fig. 2B); both parasitic finch genera have somewhat higher *dN/dS* than the outgroups for this set of genes. Importantly, across several sets of autosomal genes, *dN/dS* estimates for *Anomalospiza* are generally comparable to both *Vidua* and the outgroup taxa (Fig. 2b and Extended Data Fig. 7 and Supplementary Table 3).

Accelerated evolution of mitochondrial DNA

Using complete mitochondrial genome sequences, we confirm an earlier report of accelerated mtDNA evolution in parasitic finches[33], but further show that the best model for the 13 mtDNA protein-coding genes combined is one in which *dN*/*dS* is higher in cuckoo finch than in *Vidua* (Fig. 3a,b). This model is strongly supported, both in comparison to the null model (ΔAICc = 70.4) and to a model in which the two parasitic finch genera differ from the outgroup but have the same *dN/dS* (ΔAICc = 18.7). As measured by *dN* and *dN/dS*, respectively, evolution of mitochondrial protein-coding genes has been 26% and 33% faster in the cuckoo finch lineage than in *Vidua*. Moreover, amino acid substitutions in the cuckoo finch lineage more often occur at sites conserved in all other sampled species — 45% (163/361) in cuckoo finch as compared to 28% (91/326) in the *Vidua* clade ($\chi^2 = 21.1$, $p < 0.0000043$), though reconstructed amino acid substitutions in the two lineages had similar distributions of Grantham scores (Extended Data Fig. 6b), a measure of amino acid dissimilarity. Finally, sequence evolution in the non-coding mitochondrial genes combined (12S rRNA, 16S rRNA, and 22 tRNAs) has been 65% faster in cuckoo finch than in *Vidua* (Fig. 3b and Supplementary Table 3).

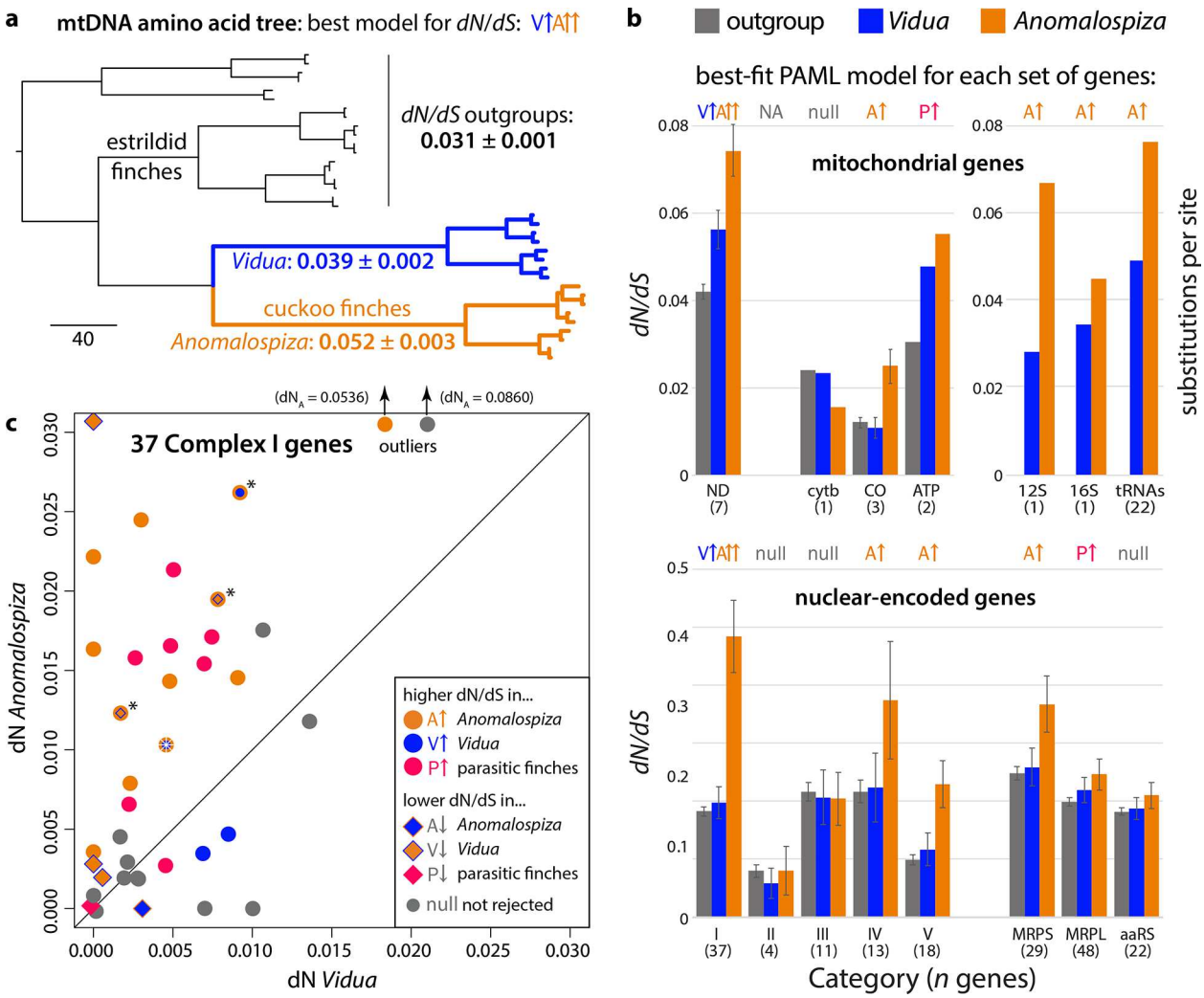


Fig. 3 | Comparisons of evolutionary rate in OXPHOS genes. a, Parsimony tree with branch lengths proportional to number of inferred amino acid substitutions in 13 mitochondrial protein-coding genes. *dN/dS* values (± SE) from PAML are indicated. **b**, *dN/dS* values (± SE, except for "cytb" and "ATP," which have too few codons to yield reliable SE estimates), or substitution rate for non-coding genes, for sets of mitochondrial and nuclear genes. The best-fit model for each gene set is indicated as in Fig. 2. Nuclear gene sets include components of OXPHOS complexes I, II, III, IV, and V, along with small (MRPS) and large (MRPL) ribosomal subunit proteins, and aminoacyl-tRNA synthetase (aaRS) genes. **c**, Rate of nonsynonymous substitution (*dN*) for 37 nuclear-encoded NADH dehydrogenase (Complex I) genes, with PAML results for *dN/dS* coded as in Fig. 2. Asterisks indicate three genes for which the best model included different *dN/dS* rates in cuckoo finch, *Vidua*, and the outgroups, respectively; all three had higher *dN/dS* in cuckoo finch than in either *Vidua* or the outgroups.

Accelerated evolution of nuclear OXPHOS genes

Supporting our final *a priori* prediction, nuclear-encoded OXPHOS genes, and specifically those that directly interact with mitochondrial gene products, have experienced faster evolution in the cuckoo finch lineage (Fig. 3b,c). The largest difference is observed for nuclear-encoded

components of Complex I (NADH dehydrogenase), with the cuckoo finch lineage evolving more than twice as fast as *Vidua*; 125% and 145% faster as measured by *dN* and *dN*/*dS*, respectively (Fig. 2c and Supplementary Table 3). Acceleration in cuckoo finch is also observed in nuclear-encoded components of Complexes IV (cytochrome c oxidase) and V (ATP synthase). Notably, the nuclear components of Complexes II (succinate dehydrogenase) and III (cytochrome c reductase) show no evidence of acceleration; the former is comprised of four nuclear-encoded proteins that do not directly interact with mtDNA-encoded proteins, whereas the latter includes only one mtDNA component, cytochrome *b,* which shows no evidence of acceleration in cuckoo finch (Fig. 3c). Finally, nuclear-encoded MRPS proteins, which complex with the small ribosomal subunit (12S rRNA), are also significantly accelerated in cuckoo finch relative to *Vidua* and the outgroup; among non-coding mtDNA genes, 12S is the most strongly accelerated in cuckoo finch relative to *Vidua* (Fig. 3c). Other categories of nuclear-encoded genes show no evidence of acceleration in cuckoo finch relative to *Vidua* (Fig. 2b and Supplementary Table 3)*, s*upporting the conclusion that the patterns described above are due specifically to mitonuclear coevolution.

To test for compensatory evolution in nuclear OXPHOS genes, we examined the spatial distribution of amino acid substitutions using a structural model of respiratory Complex I in cuckoo finch (Fig. 4a and Extended Data Fig. 8). This model should provide a reasonable approximation of the distances among residues in the overall quaternary structure. We find several examples in which nuclear- and mtDNA-encoded substitutions in cuckoo finch are in close contact (i.e., Cα carbons within 8 Å; Fig. 4B), though distances of 8-20 Å are more common. To explore this further, we divided all nuclear-encoded residues (n = 7,176 in cuckoo finch) into 10 proximity classes of equal size and, for each class, estimated *dN* and compared models for *dN/dS*. In addition to confirming faster evolution in cuckoo finch, we find both a significant effect of proximity and a significant interaction between proximity and lineage; in other words, the magnitude of acceleration in cuckoo finch relative to *Vidua* is greatest for nuclear residues that are closer to mtDNA-encoded residues (Fig. 4c). Moreover, nuclear-encoded amino acid substitutions in cuckoo finch are on average closer to “unique” mtDNA-encoded substitutions (i.e., mtDNA changes at positions conserved in the other seven species we analyzed) than are randomly selected residues within the same nuclear proteins (p = 0.0082 for median distance to the nearest mt-encoded change, and p = 0.0311 for the proportion of nuclear-

encoded changes within 10Å of a mtDNA-encoded change; see Extended Data Fig. 9 for additional details).

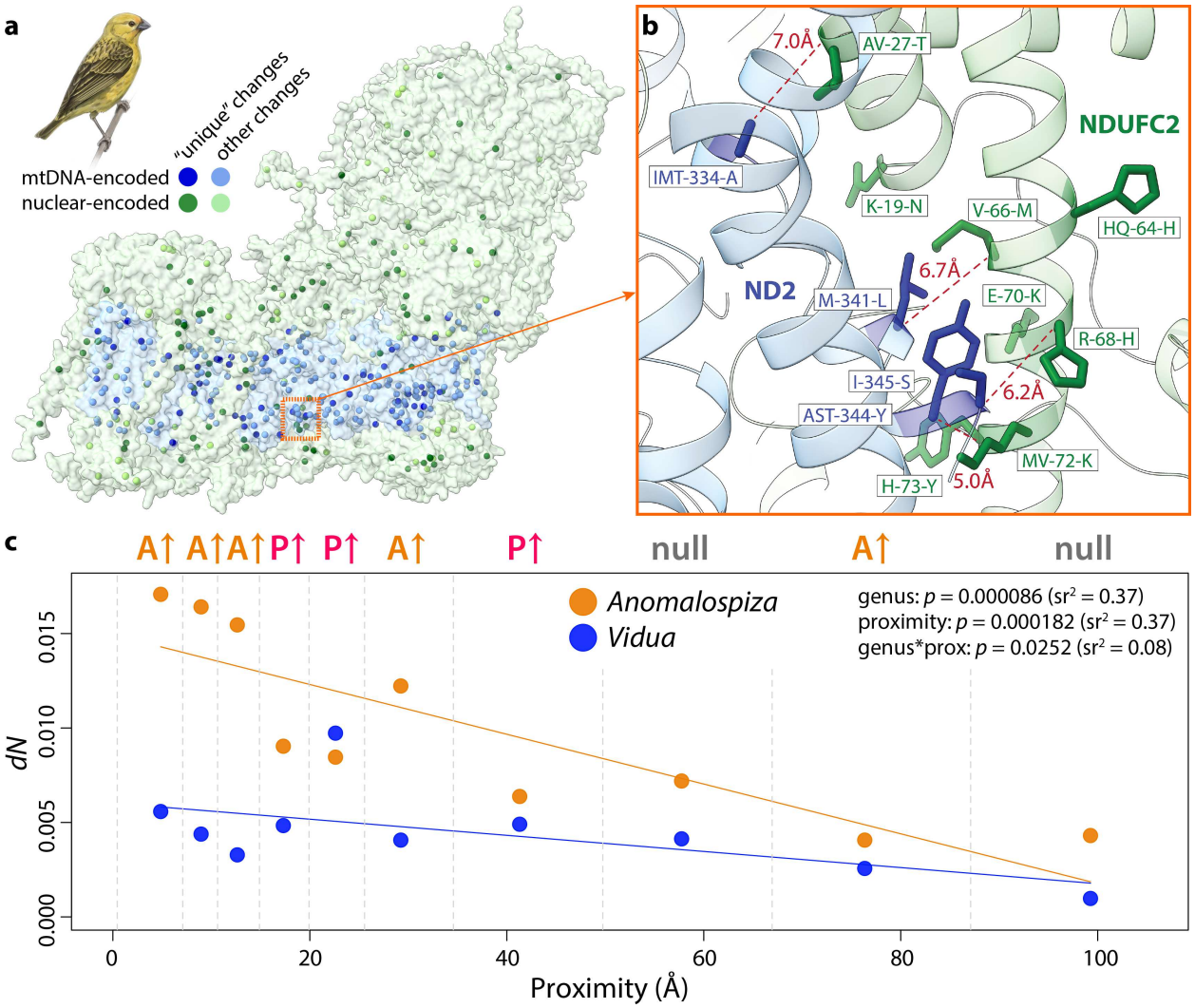


Fig. 4 | Patterns of amino acid substitution in respiratory complex I. a, Structural model of complex I, NADH dehydrogenase, based on sequences from the cuckoo finch reference genome. Nuclear- and mtDNA-encoded proteins are shown in shades of green and blue, respectively. Positions with amino acid changes in the cuckoo finch lineage are highlighted, with darker green and blue indicating unique changes (i.e., occurring at positions conserved in the seven other species in our analysis); lighter green and blue indicate positions that also have a change in at least one other species and/or, in the case of mtDNA-encoded proteins, positions that differ between cuckoo finch matrilines. **b**, Expanded view of the region highlighted in (A) showing examples of nuclear- and mtDNA-encoded substitutions in close proximity. **c**, Rate of amino acid substitution in nuclear-encoded Complex I genes as a function of proximity to the nearest mtDNA-encoded residue. Estimates of *dN* from PAML are plotted for ten proximity classes of equal size (n = 717 or 718 codons per class). The median proximity for each class is plotted on the x-axis with dotted lines showing the boundaries between classes. The best model for *dN*/*dS* is indicated, as in Fig. 2, for each proximity class. P-values and semi-partial correlation coefficients from an ANOVA with median proximity as a covariate are provided.

Demographic effects

It is well known that fixation of slightly deleterious alleles is more likely in small populations[34], prompting an earlier suggestion that faster mtDNA evolution in both *Anomalospiza* and *Vidua* might be due to parasitic birds having relatively small effective population size (N_e)[33]. Comparison of the two parasitic finch genera, however, provides no support for demographic effects as a primary explanation for faster evolution of autosomal OXPHOS genes in cuckoo finches. Analysis of autosomal polymorphism data in PSMC[35] suggests that cuckoo finches have maintained larger N_e than either of the *Vidua* species we sampled for most of the past million years or more (Extended Data Fig. 10a). We also find that Z chromosome nucleotide diversity, autosomal nucleotide diversity, and their ratio are all higher in cuckoo finch than in *Vidua* (Extended Data Fig. 10b). All three species have "Z:A" diversity ratios well below the "all-else-equal" theoretical expectation of $3/4N_e$ [36], but this is consistent with results for other birds and likely reflects a variety of processes that reduce genetic diversity on the Z chromosome[8]. For example, polygynous mating systems in *Vidua*[37] may reduce Z chromosome N_e to a greater extent than in cuckoo finch.

In contrast, patterns of matrilineal diversity in cuckoo finch and *Vidua* are radically different. At the population level, the cuckoo finches we sampled in Zambia have substantially greater mitochondrial and W chromosome diversity, ~20-fold higher than in the *Vidua* populations we sampled (Extended Data Fig. 10), reflecting the long-term maintenance of divergent, host-adapted matrilines. In contrast, mitochondrial nucleotide diversity within cuckoo finch matrilines is lower than in *Vidua* (Extended Data Fig. 10), consistent with the maternally inherited components of the genome experiencing a higher intensity of selection in cuckoo finch.

Discussion

Our results are consistent with a model in which genetic hitchhiking, amplified by coevolutionary selection on egg coloration in cuckoo finches, has resulted in profound consequences for the non-recombining, maternally inherited components of the genome, and in turn the nuclear-encoded proteins that directly interact with mitochondrial gene products. Accelerated rates of amino acid substitution, pseudogenization and gene loss in W-linked genes support the conclusion that the hitchhiking of deleterious mutations during bouts of positive

selection is an important mechanism contributing to degeneration of the sex-limited sex chromosome, at least in the latter stages of degeneration[38]. Contributing to this process, mutations in regulatory or coding regions that negatively impact function may generate selection for silencing that makes subsequent pseudogenization and gene loss effectively neutral[39–41]. Because Z-linked gametologs remain functional and subject to further selection on dosage compensation and sex-specific expression[42,43], the long-term fitness consequences of losing W-linked genes may be minimal.

While selection has long been recognized as a likely factor in sex chromosome degeneration[13,14], we are aware of only one other example connecting lineage specific acceleration to a specific axis of ecological selection; it has been suggested that selection on sperm competition in chimpanzees and bonobos has accelerated Y-chromosome degeneration in comparison to other great apes[44,45]. Other comparisons of gene loss between related lineages have focused on "neo-sex chromosomes" and the timing of recombination cessation as a major determinant of the pace of degeneration[46,47].

The lack of recombination across most of the W chromosome and its effective linkage with the non-recombining mtDNA precludes a standard genome-wide association study (GWAS) and prevents us from identifying the specific genes or regulatory regions underlying egg phenotypes in different cuckoo finch lineages. We did not detect any large additions to the cuckoo finch W chromosome (e.g., via fusion with an autosome or transposition of a portion thereof to the sex chromosomes[47,48]), a potential mechanism for recruiting additional genes to the maternally inherited genome. Though a long-standing hypothesis suggests maternal inheritance via the W chromosome[25], mtDNA variants might also influence egg phenotype[49]. Regardless of the specific targets, however, the predicted consequences of selection on maternally inherited traits remain the same for the W chromosome, mtDNA and nuclear OXPHOS genes.

Correlated acceleration of mtDNA- and nuclear-encoded OXPHOS genes in cuckoo finch is consistent with previous analyses of mitonuclear coevolution[6,50–53], though some have questioned whether correlated rates are due specifically to nuclear compensation given limited evidence of nuclear-encoded changes at positions in direct contact with mtDNA-encoded changes[51–53]. Our finding that *dN/dS* in nuclear-encoded residues of Complex I is strongly related to their proximity to mtDNA-encoded residues is consistent with compensatory coevolution; note that *dN* was ~360% higher in cuckoo finch than *Vidua* for the 30% of nuclear-encoded residues

within ~15Å of a mt-encoded residue (dN = 0.0162 versus 0.0045, Fig. 4C). Many of these changes fall outside a typical 8 Å threshold for "close contact"[54], though the actual distance between interacting side chain atoms may be shorter than the distance between Cα atoms, the proximity measure we used in these analyses. Moreover, research on "protein sectors" suggests that stability and function rely on cooperative amino acid networks that modulate conformation[55,56]. In an intriguing recent example, substitutions surrounding a highly conserved binding site appear to alter its conformation and activity sufficiently to produce alternative male morphs in the ruff *Philomachus pugnax*[57]. Nuclear compensation in OXPHOS genes might be similarly achieved via changes at nearby residues but not necessarily those in direct contact.

Mitonuclear coevolution has received growing attention as a potential source of genetic incompatibilities[3,4,58–60] that promote speciation[5,7,60-62]. In the basic model, faster evolution of mtDNA presents an ever-moving target for coevolving nuclear genes, leading to the evolution of mitonuclear incompatibilities between diverging populations over time[5,7]. Empirical studies, however, often suggest direct selection on OXPHOS performance across an environmental gradient[60–64], which may increase the likelihood of incompatibilities even between closely related lineages. Our results point to a different, non-exclusive mechanism that may increase the rate of mitonuclear coevolution: genetic hitchhiking associated with ecological selection on maternal genome components, but not those directly related to OXPHOS performance. Direct environmental selection on OXPHOS components and hitchhiking may have somewhat different molecular consequences; in the former case, we might expect a small subset of substitutions to have relatively larger effects on protein structure and function and to elicit compensatory changes at sites in close contact, whereas in the latter case, most or all mitochondrial substitutions should be just slightly deleterious perhaps allowing compensation via changes in a larger set of surrounding residues. Moreover, in cuckoo finches, the cumulative effects of hitchhiking are likely more substantial due to ongoing coevolution with hosts.

These considerations may help explain the unusual coexistence in cuckoo finch of highly divergent mitochondrial lineages within a single, interbreeding population, an observation that seems to challenge the expectation that mitochondrial and nuclear genomes must be highly coevolved[65]. Cuckoo finch mitochondrial lineages are more divergent (average of 118 amino acid differences between the "prinia" and "cisticola" lineages) than are other avian lineages with evidence of mitonuclear incompatibility (e.g., as few as 5-7 amino acid differences[60,62] and up to

~78 differences[66]). In cuckoo finch, the diversification of host-specific matrilines combined with a lack of assortative mating implies that selection for compensatory change in nuclear-encoded OXPHOS genes will produce protein structures that tend to work well with a diversity of mitochondrial lineages. This leads to the corollary prediction that OXPHOS performance in cuckoo finches may be slightly compromised in comparison to related species. It is important to note, however, that much of the evolutionary change we analyzed occurs on the ancestral cuckoo finch branch, prior to the diversification of the extant maternal lineages. W chromosome pseudogenization and gene loss maps primarily to this ancestral branch (Fig. 1). Likewise, much of the coevolutionary change in nuclear- and mtDNA-encoded OXPHOS genes maps to this ancestral branch, which represents roughly 70% of cuckoo finch mtDNA evolution (Extended Data Fig. 3B). During much of that earlier history, we speculate that ancestral cuckoo finches were parasitizing and coevolving with multiple host species in a highly dynamic system that included episodic colonization of new hosts along with lineage extinctions, such that the extant matrilines represent just the most recent iteration of an ongoing process. This would imply that maternal inheritance of egg coloration, which we hypothesize is the fundamental cause of the results reported here, also evolved long before the most recent common ancestor of the extant cuckoo finch matrilines.

Our results connect coevolution at different levels of organization, linking sex-limited ecological selection, a key feature of the ongoing coevolutionary arms races between cuckoo finch matrilines and their respective hosts, to both degeneration of the sex-limited chromosome and accelerated mitonuclear coevolution. While this full chain of genomic consequences may be limited to organisms with ZW sex chromosomes, our study illustrates the potential consequences of genetic hitchhiking when adaptation proceeds without recombination. Finally, our study points to an underexplored connection between selection on the W chromosome and mitonuclear coevolution, along with testable hypotheses for comparative analyses of molecular evolution.

1. Hill, G. E. *et al.* Assessing the fitness consequences of mitonuclear interactions in natural populations. *Biol. Rev.* **94**, 1089–1104 (2019).
2. Burton, R. S. & Barreto, F. S. A disproportionate role for mtDNA in Dobzhansky–Muller incompatibilities? *Mol. Ecol.* **21**, 4942–4957 (2012).

3. Meiklejohn, C. D. *et al.* An incompatibility between a mitochondrial tRNA and its nuclear-encoded tRNA synthetase compromises development and fitness in *Drosophila*. *PLoS Genet.* **9**, e1003238 (2013).
4. Moran, B. M. *et al.* A lethal mitonuclear incompatibility in complex I of natural hybrids. *Nature* 1–9 (2024).
5. Hill, G. E. Mitonuclear coevolution as the genesis of speciation and the mitochondrial DNA barcode gap. *Ecol. Evol.* **6**, 5831–5842 (2016).
6. Barreto, F. S. *et al.* Genomic signatures of mitonuclear coevolution across populations of *Tigriopus californicus*. *Nat. Ecol. Evol.* **2**, 1250–1257 (2018).
7. Burton, R. S. The role of mitonuclear incompatibilities in allopatric speciation. *Cell. Mol. Life Sci.* **79**, 103 (2022).
8. Irwin, D. E. Sex chromosomes and speciation in birds and other ZW systems. *Mol. Ecol.* **27**, 3831–3851 (2018).
9. Wang, S., Mank, J. E., Ortiz-Barrientos, D. & Rieseberg, L. H. Genome architecture and speciation in plants and animals. *Mol. Ecol.* e70004 (2025).
10. Felsenstein, J. The evolutionary advantage of recombination. *Genetics* **78**, 737–756 (1974).
11. Hadany, L. & Feldman, M. W. Evolutionary traction: the cost of adaptation and the evolution of sex. *J. Evol. Biol.* **18**, 309–314 (2005).
12. Hartfield, M. & Otto, S. P. Recombination and hitchhiking of deleterious alleles. *Evolution* **65**, 2421–2434 (2011).
13. Rice, W. R. Genetic hitchhiking and the evolution of reduced genetic activity of the Y sex chromosome. *Genetics* **116**, 161–167 (1987).
14. Bachtrog, D. Evidence that positive selection drives Y-chromosome degeneration in *Drosophila miranda*. *Nat. Genet.* **36**, 518–522 (2004).
15. Smeds, L. *et al.* Evolutionary analysis of the female-specific avian W chromosome. *Nat. Commun.* **6**, 7330 (2015).
16. Hill, G. E. Mitonuclear compensatory coevolution. *Trends Genet.* **36**, 403–414 (2020).
17. Havird, J. C., Whitehill, N. S., Snow, C. D. & Sloan, D. B. Conservative and compensatory evolution in oxidative phosphorylation complexes of angiosperms with highly divergent rates of mitochondrial genome evolution. *Evolution* **69**, 3069–3081 (2015).
18. Spottiswoode, C. N. & Stevens, M. Visual modeling shows that avian host parents use multiple visual cues in rejecting parasitic eggs. *Proc. Natl. Acad. Sci.* **107**, 8672–8676 (2010).
19. Spottiswoode, C. N. & Stevens, M. How to evade a coevolving brood parasite: egg discrimination versus egg variability as host defences. *Proc. R. Soc. B: Biol. Sci.* **278**, 3566–3573 (2011).

20. Caves, E. M., Stevens, M., Iversen, E. S. & Spottiswoode, C. N. Hosts of avian brood parasites have evolved egg signatures with elevated information content. *Proc. R. Soc. B: Biol. Sci.* **282**, 20150598 (2015).
21. Spottiswoode, C. N. *et al.* Genetic architecture facilitates then constrains adaptation in a host–parasite coevolutionary arms race. *Proc. Natl. Acad. Sci.* **119**, e2121752119 (2022).
22. Spottiswoode, C. N. & Stevens, M. Host-parasite arms races and rapid changes in bird egg appearance. *Am. Nat.* **179**, 633–648 (2012).
23. Dixit, T. *et al.* Chase-away evolution maintains imperfect mimicry in a brood parasite–host system despite rapid evolution of mimics. *Nat. Ecol. Evol.* **7**, 1978–1982 (2023).
24. Caves, E. M., Stevens, M. & Spottiswoode, C. N. Does coevolution with a shared parasite drive hosts to partition their defences among species? *Proc. R. Soc. B: Biol. Sci.* **284**, 20170272 (2017).
25. Punnett, R. C. Inheritance of egg-colour in the parasitic cuckoos. *Nature* **132**, 892–893 (1933).
26. Jamie, G. A. et al. Multimodal mimicry of hosts in a radiation of parasitic finches. *Evolution* **74**, 2526–2538 (2020).
27. Payne, R. B. Nestling mouth markings and colors of Old World finches Estrildidae: Mimicry and coevolution of nesting finches and their *Vidua* brood parasites. *Miscellaneous Publications, University of Michigan Museum of Zoology* **194**, 1–45 (2005).
28. Rosser, Z. H., Balaresque, P. & Jobling, M. A. Gene conversion between the X chromosome and the male-specific region of the Y chromosome at a translocation hotspot. *Am. J. Hum. Genet*. **85**, 130–134 (2009).
29. Wright, A. E., Harrison, P. W., Montgomery, S. H., Pointer, M. A. & Mank, J. E. Independent stratum formation on the avian sex chromosomes reveals inter-chromosomal gene conversion and predominance of purifying selection on the W chromosome. *Evolution* **68**, 3281–3295 (2014).
30. Wright, A. E., Moghadam, H. K. & Mank, J. E. Trade-off between selection for dosage compensation and masculinization on the avian Z chromosome. *Genetics* **192**, 1433–1445 (2012).
31. Mank, J. E., Hosken, D. J. & Wedell, N. Conflict on the sex chromosomes: Cause, effect, and complexity. *Cold Spring Harb. Perspect. Biol*. **6**, a017715 (2014).
32. Dean, R., Harrison, P. W., Wright, A. E., Zimmer, F. & Mank, J. E. Positive selection underlies faster-Z evolution of gene expression in birds. *Mol. Biol. Evol*. **32**, 2646–2656 (2015).
33. Sorenson, M. D. & Payne, R. B. A single ancient origin of brood parasitism in African finches: implications for host-parasite coevolution. *Evolution* **55**, 2550–2567 (2001).
34. Ohta, T. Slightly deleterious mutant substitutions in evolution. *Nature* **246**, 96–98 (1973).
35. Li, H. & Durbin, R. Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011).

36. Sayres, M. A. W. Genetic diversity on the sex chromosomes. *Genome Biol. Evol*. **10**, 1064–1078 (2018).
37. Payne, R. B. & Payne, K. Social organization and mating success in local song populations of village indigobirds, *Vidua chalybeata*. *Z. für Tierpsychol*. **45**, 113–173 (1977).
38. Kaiser, V. B. & Charlesworth, B. Muller's Ratchet and the Degeneration of the *Drosophila miranda* Neo-Y Chromosome. *Genetics* **185**, 339–348 (2010).
39. Charlesworth, B. Model for evolution of Y chromosomes and dosage compensation. *Proc. Natl. Acad. Sci*. **75**, 5618–5622 (1978).
40. Bachtrog, D. Y-chromosome evolution: emerging insights into processes of Y-chromosome degeneration. *Nat. Rev. Genet*. **14**, 113–124 (2013).
41. Charlesworth, B. & Charlesworth, D. Evolution: A new idea about the degeneration of Y and W chromosomes. *Curr. Biol*. **30,** R871–R873 (2020).
42. Lenormand, T., Fyon, F., Sun, E. & Roze, D. Sex Chromosome Degeneration by Regulatory Evolution. *Curr. Biol*. **30**, 3001-3006.e5 (2020).
43. D. E. Shaw, A. S. Naftaly, M. A. White, Positive selection drives cis-regulatory evolution across the threespine stickleback Y chromosome. *Mol. Biol. Evol.* **41**, msae020 (2024).
44. Cechova, M. et al. Dynamic evolution of great ape Y chromosomes. *Proc. Natl. Acad. Sci*. **117**, 26273–26280 (2020).
45. Makova, K. D. et al. The complete sequence and comparative analysis of ape sex chromosomes. *Nature* **630**, 401–411 (2024).
46. Sardell, J. M., Josephson, M. P., Dalziel, A. C., Peichel, C. L. & Kirkpatrick, M. Heterogeneous histories of recombination suppression on stickleback sex chromosomes. *Mol. Biol. Evol*. **38**, 4403–4418 (2021).
47. Sigeman, H., Downing, P. A., Zhang, H. & Hansson, B. The rate of W chromosome degeneration across multiple avian neo-sex chromosomes. *Sci. Rep*. **14**, 16548 (2024).
48. Huang, Z. et al. Recurrent chromosome reshuffling and the evolution of neo-sex chromosomes in parrots. *Nat. Commun*. **13**, 944 (2022).
49. Lu, M.-Y., Wang, W.-W., Qi, G.-H., Xu, L. & Wang, J. Mitochondrial transcription factor A induces the declined mitochondrial biogenesis correlative with depigmentation of brown eggshell in aged laying hens. *Poult. Sci*. **100**, 100811 (2021).
50. Yan, Z., Ye, G. & Werren, J. H. Evolutionary rate correlation between mitochondrial-encoded and mitochondria-associated nuclear-encoded proteins in insects. *Mol. Biol. Evol*. **36**, 1022–1036 (2019).

51. Zhang, F. & Broughton, R. E. Mitochondrial–nuclear interactions: Compensatory evolution or variable functional constraint among vertebrate oxidative phosphorylation genes? *Genome Biol. Evol*. **5**, 1781–1791 (2013).
52. Piccinini, G. et al. Mitonuclear coevolution, but not nuclear compensation, drives evolution of OXPHOS complexes in bivalves. *Mol. Biol. Evol*. **38**, msab054- (2021).
53. Weaver, R. J., Rabinowitz, S., Thueson, K. & Havird, J. C. Genomic signatures of mitonuclear coevolution in mammals. *Mol. Biol. Evol*. **39**, msac233 (2022).
54. Adhikari, B. & Cheng, J. Protein residue contacts and prediction methods. *Methods Mol. Biol*. **1415**, 463–476 (2016).
55. McLaughlin, R. N., Poelwijk, F. J., Raman, A., Gosal, W. S. & Ranganathan, R. The spatial architecture of protein function and adaptation. *Nature* **491**, 138–142 (2012).
56. Salinas, V. H. & Ranganathan, R. Coevolution-based inference of amino acid interactions underlying protein function. *eLife* **7**, e34300 (2018).
57. Loveland, J. L. et al. A single gene orchestrates androgen variation underlying male mating morphs in ruffs. *Science* **387**, 406–412 (2025).
58. Chang, C.-C., Rodriguez, J. & Ross, J. Mitochondrial–nuclear epistasis impacts fitness and mitochondrial physiology of interpopulation *Caenorhabditis briggsae* hybrids. *G3: Genes, Genomes, Genet*. **6**, 209–219 (2016).
59. Healy, T. M. & Burton, R. S. Strong selective effects of mitochondrial DNA on the nuclear genome. *Proc. Natl. Acad. Sci*. **117**, 6616–6621 (2020).
60. McDiarmid, C. S., Hooper, D. M., Stier, A. & Griffith, S. C. Mitonuclear interactions impact aerobic metabolism in hybrids and may explain mitonuclear discordance in young, naturally hybridizing bird lineages. *Mol. Ecol*. **33**, e17374 (2024).
61. Morales, H. E. et al. Concordant divergence of mitogenomes and a mitonuclear gene cluster in bird lineages inhabiting different climates. *Nat. Ecol. Evol*. **2**, 1258–1267 (2018).
62. Wang, S. et al. Signatures of mitonuclear coevolution in a warbler species complex. *Nat. Commun*. **12**, 4279 (2021).
63. Lamb, A. M. et al. Climate-driven mitochondrial selection: A test in Australian songbirds. *Mol. Ecol*. **27**, 898–918 (2018).
64. Gonçalves, L. T., Pezzi, P. H., Deprá, M. & Françoso, E. Mitonuclear coevolution in bumblebees (*Bombus*): Genomic signatures and its role in climatic niche adaptation. *Genome Biol. Evol*. **17**, evaf123 (2025).
65. Hill, G. E. The mitonuclear compatibility species concept. *Auk* **134**, 393–409 (2017).

66. Morales, H. E., Pavlova, A., Joseph, L. & Sunnucks, P. Positive and purifying selection in mitochondrial genomes of a bird with mitonuclear discordance. *Mol. Ecol*. **24**, 2820–2837 (2015).

Methods

Genome assemblies

Our analyses are based in part on new, high quality genome assemblies for five avian species[67]. These include three parasitic finches (family Viduidae): cuckoo finch *Anomalospiza imberbis* (NCBI genome accession: GCA_031753505.1), pin-tailed whydah *Vidua macroura* (GCF_024509145.1) and village indigobird *V. chalybeata* (GCF_026979565.1); along with two icterid species, the brown-headed cowbird *Molothrus ater* (GCF_012460135.2) and red-winged blackbird *Agelaius phoeniceus* (GCF_020745825.1). We also used previously published genomes for common canary *Serinus canaria* (GCF_022539315.1) and two estrildid finches: zebra finch *Taeniopygia guttata* (GCF_003957565.2) and white-rumped munia *Lonchura striata* (GCF_005870125.1). Estrildid finches (family Estrildidae) are the sister group to the parasitic finches (family Viduidae)[33]; the remaining species serve as more distant outgroup taxa. One of the outgroups, brown-headed cowbird, is well known as a generalist brood parasite of many different passerine species[68].

Population re-sequencing data

To allow for correction of errors and assembly of genes missing from the reference genomes, to characterize divergence among cuckoo finch matrilines, and to estimate population genetic parameters, we generated whole genome re-sequencing (WGS) data for population samples of cuckoo finch (n = 20), pin-tailed whydah (n = 16), Cameroon indigobird *V. camerunensis* (n = 19), a close relative of village indigobird, brown-headed cowbird (n = 10), and red-winged blackbird (n = 10) at 9-72x genome coverage per sample (median = 21x). These samples were collected during prior fieldwork[21,69,70] or obtained from the United States National Museum of Natural History. Comparable re-sequencing data for zebra finch (n = 19)[71], white-rumped munia (n = 4)[72] and common canary (n = 6, including one pooled sample of 39 individuals)[73,74] were downloaded from GenBank. See Supplementary Table 4 for details on all samples. To process the WGS data, we used NGmerge[75] to trim adapter sequences, then aligned per sample data to the relevant reference genome using BOWTIE2 v. 2.5.1[76], followed by sorting with samtools v.

1.12[77] and removal of duplicate read pairs with Picard v. 2.25.2 (https://broadinstitute.github.io/picard/). We used mosdepth v. 0.3.1[78] to calculate per sample sequencing depth for genomic regions of interest.

Retrieving, assembling, and correcting gene sequences

To examine the evolution of genes of interest, we first retrieved full-length gene sequences and annotations from as many of the reference genomes as possible. For any missing genes, we next used BLAST v. 2.12.0[79], with the coding sequence from a related taxon as the query, to locate genes present in genome assemblies but either not annotated or annotated as a low quality or uncharacterized locus ("LOC…"). We used this approach for all cuckoo finch genes as most of the analyses presented here were completed before the cuckoo finch genome was annotated. For any genes not located with BLAST, we used BOWTIE2 to align population resequencing data to the gene sequence of the most closely related taxon with available data. For almost all autosomal and Z-linked genes, this approach allowed us to assemble at least the coding regions of any missing genes.

We used Geneious Prime v. 2023.2.1 (https://www.geneious.com) to align and carefully examine the sequences for each gene, checking for any changes incompatible with function, including unexpected stop codons or frameshift mutations, changes in conserved exon-intron boundaries, or missing exons. Such changes were corrected, or less often confirmed, by aligning WGS data to the gene in question; unexpected frameshift mutations and stop codons were most often revealed to represent sequencing errors in the reference genome, whereas missing exons could be assembled in almost all cases. For genes or portions of genes corrected or assembled from WGS data, we used the consensus sequence for each species or lineage for further analysis, whereas the sequence from the reference genome was used when no issues were encountered.

For genes with alternative transcripts and/or inconsistent annotations among taxa, we selected a single annotation/isoform for analysis; generally, the transcript with the largest number of included exons and codons, provided this annotation implied a functional protein in all eight of our focal species, including conserved exon-intron boundaries. When this test failed, we used protein BLAST in GenBank to identify a "consensus annotation" across a broader diversity of passeriform species and used that to inform the selection of an alternative transcript and/or to modify annotations in our focal species (e.g., moving the position of the start codon to

the right was a typical change). A few genes with one or more confirmed frameshift mutations near the end of the final exon were assumed to remain functional with an alternate, nearby stop codon (see ref. 80); we trimmed these alignments to the last conserved codon when alignment at the end of the gene was ambiguous.

Z-W gametologs

We used a similar approach to build alignments for genes with Z- and W-linked gametologs, but with several additional considerations. We started with a set of 83 genes observed to have Z-W gametologs in previous studies of songbirds[15,81], including two (IFNW1, SMC2) present as pseudogenes on the collared flycatcher *Ficedula albicollis* W chromosome[15] and another annotated pseudogene (KCNN2) in the canary genome. Of these 83 genes, five (ZNF532, NEDD4L, ST8SIA3, TXNL1, TCF4) were determined to be in the pseudoautosomal region (PAR) or immediately adjacent to it (see Supplementary Fig. 1). These and twelve additional protein-coding genes in the PAR were in most cases included only in the Z chromosome assembly for each species, and we found no evidence of divergent Z- and W-linked versions of these PAR genes in any species. Thus, we examined the remaining set of 78 candidate genes that might have Z-W gametologs (see Supplementary Table 2).

Each of the seven reference genomes with annotated W chromosome assemblies and/or W-linked contigs included additional annotated protein-coding genes (n = 11 to 33 per genome) that could not be assigned to one of the 78 candidate genes; most of these genes were listed as "uncharacterized" and/or "low quality." Reciprocal BLAST searches indicate that almost all these additional genes represent repetitive sequences, typically generating multiple matches on the W chromosome along with additional matches on both the Z chromosome and various autosomes. Exceptions to this typical result included: 1) a processed pseudogene limited to the *V. chalybeata* W chromosome (LOC128782888 matches the coding regions of the autosomal ASB8 gene); and 2) an incomplete, W-linked copy of the autosomal TACR3 gene, a pseudogene shared by *Agelaius* (LOC129132391) and *Molothrus* (LOC118701484). These may represent examples of relatively recent transpositions to the W chromosome. Overall, however, we found that none of the additional W-linked annotated genes were consistent with expectations for single-copy Z-W gametologs.

606 Using the process described above, we retrieved or assembled as many of the 78 candidate
607 genes as possible. Because we made a specific, a priori prediction of accelerated
608 pseudogenization and loss of W-linked genes in cuckoo finch, it was critical that we tested
609 whether missing genes or portions of genes were lost evolutionarily or were simply missing from
610 a given genome assembly, particularly given the difficulty of assembling the W chromosome[15,82].
611 Thus, for any missing gene or portion of a gene in any of the focal taxa, we aligned WGS data to
612 the relevant gene sequence of the most closely related taxon with available data, aligning
613 simultaneously to reference sequences for both the Z and W gametologs and checking sequence
614 reads aligned to the Z-linked reference for divergent reads in female samples that might represent
615 a W-linked gametolog. If a given gene or individual coding region was not recovered using more
616 stringent alignment settings (--very-sensitive, --no-discordant, --no-mixed), we tried again using
617 less stringent options (--very-sensitive-local with no paired read requirements). In many cases,
618 we successfully assembled the entire coding sequence of a missing gene and confirmed that
619 sequences were either Z- or W-linked based on comparison of male and female samples. Our
620 success in assembling missing genes (including all W-linked genes in *L. striata*, the reference
621 genome for which is from a male) suggests that our inferences of the full or partial evolutionary
622 loss of a given gene, based on a complete lack of short-read sequences aligning to one or more of
623 a gene's coding regions, are robust. Finally, given the divergent maternal lineages within cuckoo
624 finch, we aligned WGS data and generated consensus sequences for each of four W chromosome
625 lineages and assessed whether gene losses or changes incompatible with function have occurred
626 after diversification of the extant, host-specific cuckoo finch lineages.
627 Alignments of Z- and W-linked genes were carefully checked and corrected as outlined
628 above, using WGS data to either correct errors or confirm changes incompatible with function.
629 For genes with alternative transcripts/isoforms, we added to the criteria described above a
630 preference for consensus annotations with the same number of homologous coding regions in the
631 Z- and W-linked copies of each gene, along with homologous intron-exon boundaries; this was
632 possible for almost all gametolog pairs with functional copies on both chromosomes.
633 The above steps allowed us to assign each W-linked gene in each species to one of five
634 categories: 1) full-length, functional copy with exon-intron structure homologous to the Z-linked
635 gametolog; 2) one or more changes incompatible with function, but with a putatively functional
636 alternative transcript still intact; 3) a likely pseudogene, most often based on multiple changes

637 incompatible with function; 4) a “remnant” pseudogene with less than 35% of the original coding
638 sequence remaining in the genome; and 5) completely lost from the genome, no detectable W-
639 linked coding regions remaining. Changes among these five states were then reconstructed on a
640 W chromosome phylogeny using Dollo parsimony[83]. A phylogeny based on the concatenated
641 coding regions of 29 W-linked genes with complete, putatively functional copies in all eight
642 species was generated using BEAST v. 2.7.7 with partitioning by codon position, model
643 averaging and an optimized relaxed clock[84,85]. We included divergence time priors for three basal
644 nodes, reflecting estimates and 95% HPD intervals from ref. 86, to calibrate divergence time
645 estimates.
646
647 **Analysis of OXPHOS genes and other sets of nuclear genes**
648 In addition to Z-W gametologs, we analyzed various genes involved in oxidative
649 phosphorylation (OXPHOS) and mitochondrial translation. These included: 1) mitochondrial
650 protein-coding genes; 2) nuclear-encoded genes comprising portions of respiratory complexes I-
651 V; 3) mitochondrial ribosomal proteins that interact with the mitochondrial small (12S) and large
652 (16S) ribosomal RNAs[87]; and 4) aminoacyl-tRNA synthetases that interact with the
653 mitochondrial tRNAs[3]. We also analyzed genes that function as assembly factors for coenzyme
654 Q10 and for OXPHOS complexes I and IV. Finally, we analyzed several sets of autosomal genes
655 to characterize genome-wide evolutionary rates, including V-type and P-type ATPases as
656 controls for the mitochondrial F-type ATP synthase (complex V) and other sets of genes that
657 might be under selection in parasitic birds. The latter included: 1) 59 genes with neurological
658 functions previously found to be accelerated in zebra finch relative to chicken[88]; 2) 56 genes
659 putatively involved in parental care and/or social behavior generally (e.g., refs. 89-91); and 3) 18
660 genes associated with egg production in galliforms or anseriforms[92,93].
661
662 **Selection of additional control sets of Z-linked and autosomal genes**
663 For comparison to the above gene sets, we also generated alignments for “control” sets of Z-
664 linked and autosomal genes. For the Z-linked control set, we used annotations for the zebra finch
665 Z chromosome and started by excluding the following genes: 1) 78 genes included in the list of
666 Z-W candidates noted above; 2) genes within 150 kbp of the above 78 genes; 3) genes annotated
667 as uncharacterized or low quality (“LOC…”); and 4) genes annotated in fewer than six of our

eight focal species. We then selected the closest remaining genes, both upstream and downstream, of each of the 29 W-linked genes with a putatively functional copy in all eight of our focal species (n = 2 x 29 = 58 genes) plus the closest gene (either upstream or downstream) to each of 48 additional Z-W candidate genes (n = 48 genes; note that MEX3C is not included in the zebra finch Z chromosome assembly and we were not able to find or assemble it), yielding a total of 106 genes in the Z-linked control set. Thus, the "control" genes had a similar distribution across the Z chromosome as the 78 Z-W candidate genes but were not tightly linked to them. We used a similar approach to assemble a control set of 102 autosomal genes with a genomic distribution similar to a set of 102 autosomal OXPHOS genes; the latter included genes encoding subunits of Complexes I-V as well as assembly proteins for Complexes I and IV but excluded several Z-linked OXPHOS genes (ATP5F1A, ATP5ME, COX7C and NDUFS4). For this autosomal control set, we randomly chose "upstream" or "downstream" and then selected the first annotated gene at least 500 kbp from the OXPHOS gene. A small number of the initially selected control genes had variable and/or ambiguous annotations across species and were replaced with the nearest available gene meeting the above criteria. Supplementary Table 5 lists each of the ~750 genes we assembled, aligned and analyzed.

Assembly and analysis of mitochondrial genomes

Mitochondrial genomes for all samples of the three parasitic finches and for representative individuals of the remaining species were assembled from WGS data (see Supplementary Table 6 for details). We used BOWTIE2 to align WGS data to available reference genomes and examined the results in Geneious, taking care to discriminate between mtDNA sequences and nuclear copies or "numts"[94]. Numts were particularly evident in data derived from blood samples of zebra finch, long-tailed finch, cuckoo finch, and Cameroon indigobird, and prevented the use of automated methods for mitochondrial genome assembly (e.g., ref. 95). For these species, we used an iterative process that involved constructing draft mtDNA and numt sequences, recognizing the latter based on changes incompatible with function, the presence of heterozygous sites within individuals, and linkage among variants observable across overlapping read pairs. Variable ratios of mtDNA and numt sequences across samples, patterns of phylogenetic divergence, and comparison of data derived from blood versus muscle tissue (when available) provided additional confirmation. As needed, we iteratively refined our inference of the mtDNA

sequence by realigning WGS data to both the mtDNA and numt draft sequences. We augmented the mtDNA data set with two additional estrildid finch species (long-tailed finch *Poephila acuticauda* and grand munia *Lonchura grandis*), each represented by two subspecific lineages, so that analyses of mitochondrial evolution included a comparable diversity of lineages in the estrildid and parasitic finch clades, respectively. We downloaded WGS data for long-tailed finch[71] and previously assembled mtDNA sequences for grand munia[96]. A phylogeny based on a concatenated alignment of the 12 light-strand mtDNA protein-coding genes was generated using the same approach as described above for W-linked genes.

We used Geneious to construct amino acid sequence alignments for both mitochondrial genes and nuclear-encoded OXPHOS genes and reconstructed amino acid substitutions in each lineage using parsimony in PAUP* v. 4.0a[97]. At positions where *Vidua* and *Anomalospiza* differ from each other and from the outgroup taxa, which yields three equally parsimonious reconstructions, we used the reconstruction with substitutions in both *Vidua* and *Anomalospiza*; thereby avoiding bias in estimating the number of amino acid substitutions for the key comparison between *Vidua* and *Anomalospiza*.

Analyses of nucleotide and protein sequence evolution

We analyzed nucleotide and protein sequence evolution using, respectively, the baseml and codeml modules within PAML v. 4.8a[98]. For each gene and various concatenated sets of genes, we estimated parameters for six models of rate variation among branches in the phylogeny. These models included: 1) a null model with the same nucleotide substitution rate or *dN/dS* ratio on all branches; and five additional models with: 2) a different rate/ratio in the parasitic finch lineage (comprising both *Vidua* and *Anomalospiza*); 3) a different rate/ratio in *Anomalospiza* only; 4) a different rate/ratio in *Vidua* only; or 5) and 6) two models with three distinct rates/ratios for *Vidua*, *Anomalospiza* and the outgroup taxa, respectively (Extended Data Fig. 5). Because models 2, 3 and 4 have the same number of parameters, we used the corrected Akaike Information Criterion (AICc)[99], with the number of aligned codons as the sample size, to select the best fit model for each gene or set of genes. For visualization of results and to focus on our *a priori* predictions of rate differences between *Vidua* and *Anomalospiza*, we used parameter estimates from model 5 (three different rates/ratios estimated for *Vidua*, *Anomalospiza* and the outgroups combined) to obtain point estimates of *dN* and *dN*/*dS* for each genus, whether there

was a significant difference in *dN/dS* or not. Single gene estimates of *dN/dS* were sometimes undefined (e.g., $dN > 0$, $dS = 0$), so for visualization, we plot *dN* values for individual genes while also reporting the best model for *dN/dS*. Where relevant (i.e., for *Vidua* generally and for maternally inherited genes in *Anomalospiza*), *dN* was calculated as the average total branch length on the "*dN* tree" from the base of a given clade to each of the terminals in that clade.

Structural analysis of respiratory complex I

Based on PAML results showing that nuclear-encoded components of respiratory complex I, NADH dehydrogenase, showed the greatest differences in evolutionary rate between *Anomalospiza* and *Vidua*, we modeled the structure of cuckoo finch complex I using the relevant amino acid sequences from the reference genome and associated mitochondrial genome. Each of the seven mtDNA-encoded and 37 nuclear-encoded proteins were submitted for homology modeling using the SwissModel server (https://swissmodel.expasy.org/)[100], allowing the server to identify each gene based on the submitted amino acid sequence. We used PyMOL v2.5.5 (https://www.pymol.org/) to assemble the modeled subunits, using the crystal structure of complex I from sheep as a guide (PDB ID: 5LNK, RCSB Protein Data Bank). Character-state reconstructions from PAUP* were used to identify positions with "unique" amino acid substitutions in *Anomalospiza* and *Vidua*, respectively — i.e., substitutions at positions conserved in all other species we analyzed — along with "other" substitutions occurring at positions with more than one change across the phylogeny, or at positions that differ between cuckoo finch matrilines (mtDNA only) or between the two *Vidua* species. We used ChimeraX v. 1.9[101] to visualize the distribution of these sites on the structural model. We used the Biopython[102] PDB module to calculate the distance between every nuclear-encoded residue and every mtDNA-encoded residue, using coordinates of the Cα atom for each residue.

To test for evidence of compensatory change in the nuclear-encoded Complex I genes, we used the 2,121 (mt) x 7,174 (nuclear) matrix of physical distances between Cα atom along with the locations of phylogenetically inferred amino acid substitutions. Sorting individual nuclear residues by the distance to the closest mtDNA-encoded residue, we created 10 proximity bins or categories, each with 717 or 718 nuclear-encoded positions. The corresponding codons for each of these 10 bins were then concatenated and analyzed as 10 different "genes" in PAML to determine the best model for variation in *dN*/*dS* among lineages and to estimate *dN* for

761 *Anomalospiza* and *Vidua*, respectively (Fig. 3c). We also conducted a randomization test to ask if
762 nuclear-encoded amino acid substitutions are closer, on average, to a unique amino acid
763 substitution in a mtDNA-encoded residue than are randomly selected residues within the same
764 nuclear genes (Extended Data Fig. 8); in other words, the randomization test maintained the
765 observed number of cuckoo finch amino acid substitutions in each of 30 nuclear-encoded genes;
766 seven nuclear genes with all residues > 20Å from any mt-encoded residue were excluded from
767 this analysis. P-values for this test are one-tailed given a clear *a priori* hypothesis.
768
769 **Demography**
770 To compare demographic history in the three parasitic finch species, we analyzed individual
771 samples in PSMC[35] following the authors' instructions (https://github.com/lh3/psmc). Initially,
772 we estimated population size for 100 historical intervals (-p "100*1"). Then, based on the initial
773 results, we combined recent time intervals with highly variable results among individuals of the
774 same population. The extent of variation in estimates for recent time intervals varied among
775 species, which led us to use a slightly different formulation for each of the three species for the
776 final analyses: cuckoo finch (-p "1*4+96*1"); pin-tailed whydah (-p "1*2+98*1"); and
777 Cameroon indigobird: (-p "1*6+1*3+91*1").
778
779 **Nucleotide diversity**
780 We used ANGSD[103] to estimate nucleotide diversity for the three parasitic finch species. We
781 used data for female samples only to allow for comparable estimates for the Z and W
782 chromosomes (both analyzed with the "-isHap 1" option). For the autosomal genome, we
783 sampled 20 Mbp for each species (= 400 randomly sampled 50kbp regions). For the W
784 chromosome, given differences in gene content, repetitive elements, and assembly quality across
785 species, we sampled 20 homologous, putatively single-copy regions, each comprising the entire
786 length of one of 20 W-linked genes that were present as complete copies in the genome
787 assemblies of all three parasitic finch species; this resulted in a sample of ~2.1 Mbp for each
788 species. As ANGSD estimates might be biased due to differences in sequencing depth, we down-
789 sampled the data to ~53M paired-end reads per sample for this analysis. Further, given sample
790 sizes of three to five females for the host-specific cuckoo finch lineages in our data set, we
791 analyzed sets of four *V. camerunensis* and *V. macroura* females, respectively, to ensure that

estimates were not biased by sample size. The averages across these sets were nearly identical to estimates for all females combined, so we report the latter. The same was true for three "mixed" sets of cuckoo finch females, each comprising one female from each of four different maternal lineages, as compared to results for all cuckoo finch females combined.

Data availability

All genome sequencing data generated during this project have been uploaded to NCBI. Details, including all relevant BioProject numbers and individual accession numbers for both new data and previously available data used in this study, are provided in Supplementary Tables 4 and 6.

Code availability

Derived data sets (e.g., coding region alignments for all genes analyzed), custom scripts, and computational pipelines supporting each of the results reported in the main text, figures, extended data figures and supplementary tables are available at https://figshare.com {link to be updated}.

67. Carroll, R. A. et al. Genomic resources for comparative analyses of obligate avian brood parasitism. *J. Heredity* (submitted).
68. Lowther, P. E. Brown-headed cowbird (*Molothrus ater*), version 1.0. *Birds of the World* (2020).
69. Balakrishnan, C. N., Sefc, K. M. & Sorenson, M. D. Incomplete reproductive isolation following host shift in brood parasitic indigobirds. *Proc. R. Soc. B: Biol. Sci*. **276**, 219–228 (2009).
70. Schuetz, J. G. Low survival of parasite chicks may result from their imperfect adaptation to hosts rather than expression of defenses against parasitism. *Evolution* **59**, 2017–2024 (2005).
71. Singhal, S. et al. Stable recombination hotspots in birds. *Science* **350**, 928–932 (2015).
72. Lu, C., Yao, C. & Hung, C. Domestication obscures genomic estimates of population history. *Mol. Ecol*. **31**, 752–766 (2022).
73. Madison, F. N., Conte, M. A., Brown, J. A., Carleton, K. L. & Dooling, R. J. Whole genome sequencing identifies genetic candidates for high-frequency hearing loss in canaries (*Serinus canaria*). *J. Acoust. Soc. Am*. **157**, 2330–2335 (2025).
74. Lopes, R. J. et al. Genetic basis for red coloration in birds. *Curr. Biol*. 26, 1427–1434 (2016).
75. Gaspar, J. M. NGmerge: merging paired-end reads via novel empirically-derived models of sequencing errors. *BMC Bioinform*. **19**, 536 (2018).

76. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
77. Danecek, P. et al. Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
78. Pedersen, B. S. & Quinlan, A. R. Mosdepth: quick coverage calculation for genomes and exomes. *Bioinformatics* **34**, 867–868 (2017).
79. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinform*. **10**, 421 (2009).
80. Ng, P. C. et al. Genetic variation in an individual human exome. *PLoS Genet*. **4**, e1000160 (2008).
81. Xu, L. et al. Dynamic evolutionary history and gene content of sex chromosomes across diverse songbirds. *Nat. Ecol. Evol*. **3**, 834–844 (2019).
82. Benham, P. M. et al. Remarkably high repeat content in the genomes of sparrows: The importance of genome assembly completeness for transposable element discovery. *Genome Biol. Evol*. **16**, evae067 (2024).
83. Farris, J. S. Phylogenetic analysis under Dollo's law. *Syst. Biol*. **26**, 77–88 (1977).
84. Bouckaert, R. R. & Drummond, A. J. bModelTest: Bayesian phylogenetic site model averaging and model comparison. *BMC Evol. Biol*. **17**, 42 (2017).
85. Bouckaert, R. et al. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol*. **15**, e1006650 (2019).
86. Oliveros, C. H. et al. Earth history and the passerine superradiation. *Proc. Natl. Acad*. Sci. **116**, 7916–7925 (2019).
87. Barreto, F. S. & Burton, R. S. Evidence for compensatory evolution of ribosomal proteins in response to rapid divergence of mitochondrial rRNA. *Mol. Biol. Evol*. **30**, 310–314 (2012).
88. Nam, K. et al. Molecular evolution of genes in avian genomes. *Genome Biol*. **11**, R68 (2010).
89. Lopes, P. C. & Bruijn, R. de. Neurotranscriptomic changes associated with chick-directed parental care in adult non-reproductive Japanese quail. *Sci. Rep*. **11**, 15481 (2021).
90. Lynch, K. S., O'Connell, L. A., Louder, M. I. M., Balakrishnan, C. N. & Fischer, E. K. Understanding the loss of maternal care in avian brood parasites using preoptic area transcriptome comparisons in brood parasitic and non-parasitic blackbirds. *G3: GenesGenomesGenet*. **9**, 1075–1084 (2019).
91. Lynch, K. S. et al. Examining the disconnect between prolactin and parental care in avian brood parasites. *Genes, Brain Behav*. **19**, e12653 (2020).
92. Mishra, S. K. et al. Transcriptome analysis reveals differentially expressed genes associated with high rates of egg production in chicken hypothalamic-pituitary-ovarian axis. *Sci. Rep*. **10**, 5976 (2020).
93. Bhavana, K. et al. Comparative transcriptome analysis of Indian domestic duck reveals candidate genes associated with egg production. *Sci. Rep*. **12**, 10943 (2022).

94. Sorenson, M. D. & Quinn, T. W. Numts: A challenge for avian systematics and population biology. *Auk* **115**, 214–221 (1998).
95. Hahn, C., Bachmann, L. & Chevreux, B. Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res*. **41**, e129–e129 (2013).
96. Stryjewski, K. F. & Sorenson, M. D. Mosaic genome evolution in a recent and rapid avian radiation. *Nat. Ecol. Evol*. **1**, 1912–1922 (2017).
97. Swofford, D. W. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods)*. (Sinauer Associates, Sunderland, Massachusetts, 2003).
98. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol*. **24**, 1586–1591 (2007).
99. Hurvich, C. M. & Tsai, C.-L. Regression and time series model selection in small samples. *Biometrika* **76**, 297–307 (1989).
100. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. **46**, W296–W303 (2018).
101. Meng, E. C. et al. UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci*. 32, e4792 (2023).
102. Cock, P. J. A. et al. Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics* **25**, 1422–1423 (2009).
103. Korneliussen, T. S., Albrechtsen, A. & Nielsen, R. ANGSD: Analysis of next generation sequencing data. *BMC Bioinform*. **15**, 356 (2014).
104. Ho, S. Y. W., Phillips, M. J., Cooper, A. & Drummond, A. J. Time dependency of molecular rate estimates and systematic overestimation of recent divergence times. *Mol. Biol. Evol*. **22**, 1561–1568 (2005).

Acknowledgements We thank Justin Schuetz for providing the population sample of pin-tailed whydahs and the United States National Museum of Natural History for the population samples of brown-headed cowbird and red-winged blackbird. Dacey Mercer, Sara Smith and Rachel Carroll assisted with DNA sequencing. Many others made this study possible by contributing to or supporting past field research[21,69], or applying their expertise to the assembly of new reference genomes[67]. Justin Schuetz kindly provided comments on an earlier draft, and Justin Havird provided helpful guidance on modeling protein structures. Drawings of birds were provided by Javier Lazaro. This work was supported by a National Science Foundation collaborative research

grant to M.D.S. (DBI 1754311), T.B.S. (DBI 1754397), C.N.B. (DBI 1754406), J.M.D. (DBI 1754643) and W.C.W. (DBI 1940624).

Author contributions M.D.S., C.N.B., C.N.S., J.M.D., T.B.S. and W.C.W. contributed to the original research proposal and development of the *a priori* hypotheses tested in this research. G.A.J., T.D. and C.N.S. provided critical samples allowing assembly and annotation of the cuckoo finch genome. J.M.D., M.D.S., C.N.B. and W.C.W. coordinated collection of population resequencing data. E.O. and M.D.S. completed the structural analysis of respiratory Complex I. M.D.S. completed other data analyses and drafted the manuscript. All authors edited and approved the final manuscript.

Competing interests The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Michael D. Sorenson.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTables47.xlsx
- SupplementaryInfo.pdf
- ExtendedDataFigs.pdf