RESOURCE ARTICLE

A novel exome probe set captures phototransduction genes across birds (Aves) enabling efficient analysis of vision evolution

Noor D. White[1,2,3] | Zachary A. Batz[1] | Edward L. Braun[4] | Michael J. Braun[2,3,5] | Karen L. Carleton[3,5] | Rebecca T. Kimball[4] | Anand Swaroop[1]

[1]Neurobiology Neurodegeneration and Repair Laboratory, National Eye Institute, National Institutes of Health, Bethesda, Maryland, USA

[2]Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, District of Columbia, USA

[3]Behavior, Ecology, Evolution and Systematics Program, University of Maryland, College Park, Maryland, USA

[4]Department of Biology, University of Florida, Gainesville, Florida, USA

[5]Department of Biology, University of Maryland, College Park, Maryland, USA

Correspondence
Noor D. White, Neurobiology Neurodegeneration and Repair Laboratory, National Eye Institute|National Institutes of Health, Bethesda, MD, USA.
Email: noor.white@nih.gov

Abstract

The diversity of avian visual phenotypes provides a framework for studying mechanisms of trait diversification generally, and the evolution of vertebrate vision, specifically. Previous research has focused on opsins, but to fully understand visual adaptation, we must study the complete phototransduction cascade (PTC). Here, we developed a probe set that captures exonic regions of 46 genes representing the PTC and other light responses. For a subset of species, we directly compared gene capture between our probe set and low-coverage whole genome sequencing (WGS), and we discuss considerations for choosing between these methods. Finally, we developed a unique strategy to avoid chimeric assembly by using "decoy" reference sequences. We successfully captured an average of 64% of our targeted exome in 46 species across 14 orders using the probe set and had similar recovery using the WGS data. Compared to WGS or transcriptomes, our probe set: (1) reduces sequencing requirements by efficiently capturing vision genes, (2) employs a simpler bioinformatic pipeline by limiting required assembly and negating annotation, and (3) eliminates the need for fresh tissues, enabling researchers to leverage existing museum collections. We then utilized our vision exome data to identify positively selected genes in two evolutionary scenarios—evolution of night vision in nocturnal birds and evolution of high-speed vision specific to manakins (Pipridae). We found parallel positive selection of *SLC24A1* in both scenarios, implicating the alteration of rod response kinetics, which could improve color discrimination in dim light conditions and/or facilitate higher temporal resolution.

KEYWORDS
birds, exome, phototransduction cascade, sequence capture, vision, whole genome sequencing

1 | INTRODUCTION

A major goal in evolutionary biology is to understand the link between DNA sequences (genotype) and phenotypes across the tree of life. Modern genomics makes it possible to examine patterns of molecular evolution for suites of functionally related genes (e.g., Feng et al., 2020; Zhang et al., 2014), but expanding those efforts to cover the diversity of life remains challenging. Although it might be ideal to generate very high-quality genome assemblies (e.g., Rhie et al., 2021) for large numbers of taxa for such comparisons, the

costs associated with using this strategy remain high. Moreover, the high-quality input DNA necessary for these genome assemblies does not exist for all organisms and obtaining fresh tissue is often not feasible. Thus, methods to obtain sequences for specific genes that can leverage existing museum collections ("museomics", see Colella et al., 2020; Guschanski et al., 2013) are very useful. We examined the potential of two such methods—sequence capture and low-coverage whole genome sequencing (WGS)—to provide sequences for genes involved in avian vision, arguably the most important sensory system in birds.

Birds (Aves) comprise roughly 10,000 species spanning diverse environments (e.g., forest, open grassland, aquatic, polar) and lifestyles (e.g., diurnal, nocturnal, crepuscular, cathemeral). The great diversity of birds presents an excellent system to investigate the evolutionary history of a wide variety of traits and the mechanisms that have allowed species to evolve from one lifestyle to another. The diversity of avian visual phenotypes, in particular, can serve as a framework for studying both the general mechanisms of trait diversification as well as the evolution of vertebrate visual systems, specifically. Indeed, comparing birds that have distinct traits related to vision may reveal the mechanisms that have allowed species to adapt to different lifestyles over evolutionary time. In order to embark on a comparative research program to understand the diversity of avian visual phenotypes, we need efficient, scalable tools to obtain vision gene sequences from a wide variety of species.

Much of the prior work on avian vision has focused on visual pigment opsin genes. Opsins can be "spectrally tuned," whereby the wavelength of light to which a given opsin responds can be altered by as little as a single amino acid change (Imai et al., 1997; Yokoyama, 2000; Yokoyama et al., 2000, 2008). Thus, DNA sequences of opsin genes can shed light on the environment for which a given organism's eyes are molecularly adapted (Hart, 2001a; Hart & Hunt, 2007; Hunt et al., 2009). Previous work has demonstrated that spectral tuning changes can be complex, and conclusions may depend on the phylogenetic scale examined. For example, a 2003 study concluded that the ancestral avian short-wavelength-sensitive opsin 1 (*SWS1*) was violet-sensitive, and that ultraviolet sensitivity independently evolved four times across bird taxa (Ödeen & Håstad, 2003). A decade later, the same authors sampled additional taxa and identified 14 such shifts in *SWS1* by at least four independent molecular mechanisms (Ödeen & Håstad, 2013). Furthermore, it has been demonstrated that changes at spectral tuning sites do not always confer changes in color vision (Bloch, Morrow, et al., 2015; Coyle et al., 2012; Hart, 2001b) and better characterization of both opsin sequence evolution, and analysis of what effect sequence changes have on light response, are needed (e.g., Bloch, Price, et al., 2015). These examples highlight the need for broader taxon sampling as well as more thorough characterization of the phenotypic effect of opsin sequence evolution.

Historically, opsins sequences have been captured via PCR-based methods (Hart et al., 2016; van Hazel et al., 2016; Ödeen & Håstad, 2009; Ödeen et al., 2011; Toomey et al., 2016; Wilkie et al., 2000). While opsin sequences extracted from assembled whole genomes

can also be used, these regions do not appear to assemble robustly, and full opsin sequences cannot always be obtained through this approach (Borges et al., 2015; Feng et al., 2020). This is especially noticeable when sequences are not recovered for opsins for which published microspectrophotometric data exists (Bowmaker et al., 1997; Hart et al., 1999; Hart & Vorobyev, 2005). The challenges with obtaining complete opsin sequences using traditional PCR approaches and from assembled whole genomes are currently a significant obstacle to sampling more species.

Ideally, comparative studies of visual systems should also look beyond the opsins to include other components of the phototransduction cascade (PTC), that is, the gene network that converts light input (beginning with the opsin receptors) into neural signals (Fu & Yau, 2007; Lamb & Hunt, 2017; Lamb et al., 2016; Pugh & Lamb, 2000; Pugh et al., 1999; reviewed in Lamb, 2020). Pugh et al. (1999) predicted eight potential molecular mechanisms involved in vertebrate light adaptation, each associated with a major PTC protein. For example, an increase in the rate of cyclic guanosine monophosphate (cGMP) hydrolysis, such as mediated by phosphodiesterase (*PDE*), is predicted to reduce the photoreceptor operating range, decreasing flash sensitivity, and speeding photoreceptor recovery. This near-instantaneous physiological adjustment within a retina could be genetically assimilated to permanently shift a species' light sensitivity while transitioning to a low-light lifestyle (e.g., from diurnal to nocturnal; from nocturnal to cave-dwelling) and would be an example of visual adaptation. Comparisons of PTC genes in diverse lineages will facilitate testing the mechanisms that Pugh et al. (1999) have hypothesized as underlying light adaptation and help us more fully understand visual adaptation in birds.

In this study, we examine a fundamental question: what is the best way to efficiently sequence the vision exome from diverse taxa for evolutionary comparison? Previous studies have sequenced the PTC by sequencing the retinal transcriptome (e.g., Wu et al., 2016) or by mining data from genome assemblies (Borges et al., 2015). These approaches are effective, yet they can be technically challenging and costly. Retinal transcriptome sequencing requires fresh tissues (i.e., specifically preserved for RNA extraction) acquired by a skilled collector, and this may be difficult or impossible for some taxa. Generating high-quality genome sequences remains expensive and requires relatively high-quality tissue samples (although WGS is sometimes possible for relatively old museum specimens—see Chen et al., 2018). Neither of these techniques is practical for conducting a broad-scale study of vision genes across many species of birds, or within a large intraspecific population of birds. However, targeted exome capture and low-coverage WGS (using read-mapping to obtain sequences from relevant loci, bypassing de novo assembly issues in other WGS data) both represent approaches that may be more feasible.

We address three major goals in this study, the first two of which are technical: (1) compare the effectiveness and efficiency of targeted exome capture with low-coverage WGS, and (2) determine whether there are unexpected challenges associated with the methods we use for assembly of visual system genes. Addressing these first two goals revealed issues that are likely to be applicable to

studies focused on other taxonomic groups and/or genomic targets. Our third major goal is biological: (3) demonstrate the utility of our novel vision exome probe set by analyzing our captured exome data in the context of two evolutionary questions. These evolutionary questions take advantage of the broad diversity of birds to investigate two potentially very different means of visual adaptation.

First, do separate nocturnal bird lineages display distinct patterns of positive selection in the vision exome, suggesting independent origins of nocturnality? We know that there are several major lineages of birds with nocturnal members, and that the few that have undergone morphological characterization display unique adaptations to nocturnality (e.g., Martin et al., 2004). However, we do not know the evolutionary history of nocturnality across the whole tree, or whether major nocturnal lineages represent independent evolutionary origins of night vision. We ask this question of four major nocturnal lineages in the bird tree, as well as more specifically within the superorder Strisores, which is comprised of two diurnal lineages (swifts and hummingbirds) nested within a clade of five nocturnal lineages (caprimulgids, frogmouths, oilbird, potoos and owlet-nightjars; White & Braun, 2019). The Strisores provide a useful microcosm of nocturnal evolution in that we know of morphological adaptations to night vision in three lineages (caprimulgids, oilbird and potoos). However, the remaining two lineages have not been morphologically characterized nor do they display adaptations observable in the field (e.g., tapetum lucidum).

Secondly, can we detect a pattern of positive selection in the vision exome unique to manakins, and specifically to their capacity for high-speed vision? Manakins are a diverse family (Pipridae) of colorful, lek-breeding Neotropical passerine birds that have courtship displays characterized by extremely high-speed movements (see Bostwick & Prum, 2003; Day et al., 2021; Fusani et al., 2007; Fuxjager et al., 2016; Lindsay et al., 2016). They appear to respond to extremely rapid displays that can only be captured using high-speed cameras, and thus manakins must have the flicker fusion threshold to detect those displays. Molecular mechanisms underlying this remarkable visual adaptation are unknown, but it is unlikely that they arose merely from opsin spectral tuning. Thus, a broader exploration of the vision exome is necessary, facilitated by the resources (vision exome probe set and low-coverage genomes) that we generate in this study.

2 | MATERIALS AND METHODS

2.1 | Exome probe set design

We designed a probe set that includes complete exon sequences from chicken (*Gallus gallus*) and zebra finch (*Taeniopygia guttata*) for 46 genes including the complete PTC pathway, non-visual opsins, and select genes used in non-image-forming responses to light present in extraocular photoreceptive organs (hereafter called the "vision exome"; Figure 1; Table S1; genes assembled from: Díaz et al., 2015; Hankins & Hughes, 2014; Nakane et al.,

2010; Nakane et al., 2014; Peirson et al., 2009; Sugihara et al., 2016; Tomonari et al., 2008). *GNAT3* was included in this probe set because it serves light-sensitive functions in the parietal eye of some fish, amphibians, and reptiles. Though birds lack a parietal eye, *GNAT3* is present in their genome. Additionally, in humans *GNAT3* is expressed most highly in retinal bipolar cells, and so the gene was included in our probe set. Two vertebrate PTC genes, *PDE6A* and *GNGT1*, have not been identified in birds (Wu et al., 2017) and thus are not present in the probe set. Exons, full genomic DNA (gDNA), and messenger RNA (mRNA) sequences for each targeted gene in chicken and zebra finch were downloaded from Ensembl (Herrero et al., 2016; Yates et al., 2016) and GenBank (NCBI Resource Coordinators, 2018), and then mapped to the chicken genome using the UCSC genome browser to confirm identity (galGal5; Kent, 2002; Kent et al., 2002; Miller et al., 2007). For three genes, *GRK1*, *LWS*, and *SWS2*, genomic sequence was not available for either species, and could not be unambiguously identified in either genome. Thus, probes were designed from the mRNA sequences. Full targeted sequences are available in File S1, and source accession numbers are in Table S1. Chicken and zebra finch span the majority of avian diversity, and thus we hypothesized that these two taxa would be sufficient to generate a probe set useful for all bird species.

Probe design and synthesis were conducted in collaboration with Daciel Arbor Biosciences (myBAITS). Probes are 120 bp long and designed to cover target sequences to a 2× depth. Probes targeting exons ranging from 108–119 bp in length were padded with "T"s on either side to reach 120 bp length. For exons shorter than 108 bp, the exons were mapped to the genome of the respective species in Geneious (v10.0.5; Biomatters Ltd.), and intron sequence was added evenly on either side to reach a total length of 120 bp. For genes with a majority of exons below 108 bp in length, the entire genomic region was included in the probe set.

During multiple design iterations for candidate target sequences, potential probes were aligned to the genomes of 10 species representing avian phyletic diversity using BLAST+ (Camacho et al., 2009) to test in silico capture of the vision exome. These were the downy woodpecker (*Picoides pubescens*), bald eagle (*Haliaeetus leucocephalus*), crested ibis (*Nipponia nippon*), rock dove (*Columba livia*), chuck-will's-widow (*Antrostomus carolinensis*), killdeer (*Charadrius vociferus*), brown kiwi (*Apteryx australis*), common ostrich (*Struthio camelus*), zebra finch and chicken. We hypothesized that our probe set would work across this phylogenetically broad range of bird species, and initial BLAST+ results indicated that our hypothesis was correct. We considered adding sequences from other taxa in the probe set design, however, including probes based on mallard (*Anas platyrhynchos*), collared flycatcher (*Ficedula albicollis*) and turkey (*Meleagris gallopavo*) genomes in the target set did not significantly increase *in silico* capture of the vision exome from the 10 test taxa. Additionally, these three genome assemblies are largely annotated with bioinformatically-predicted proteins, which have low sequence identity with the chicken or zebra finch sequences and display unexpectedly high variation in
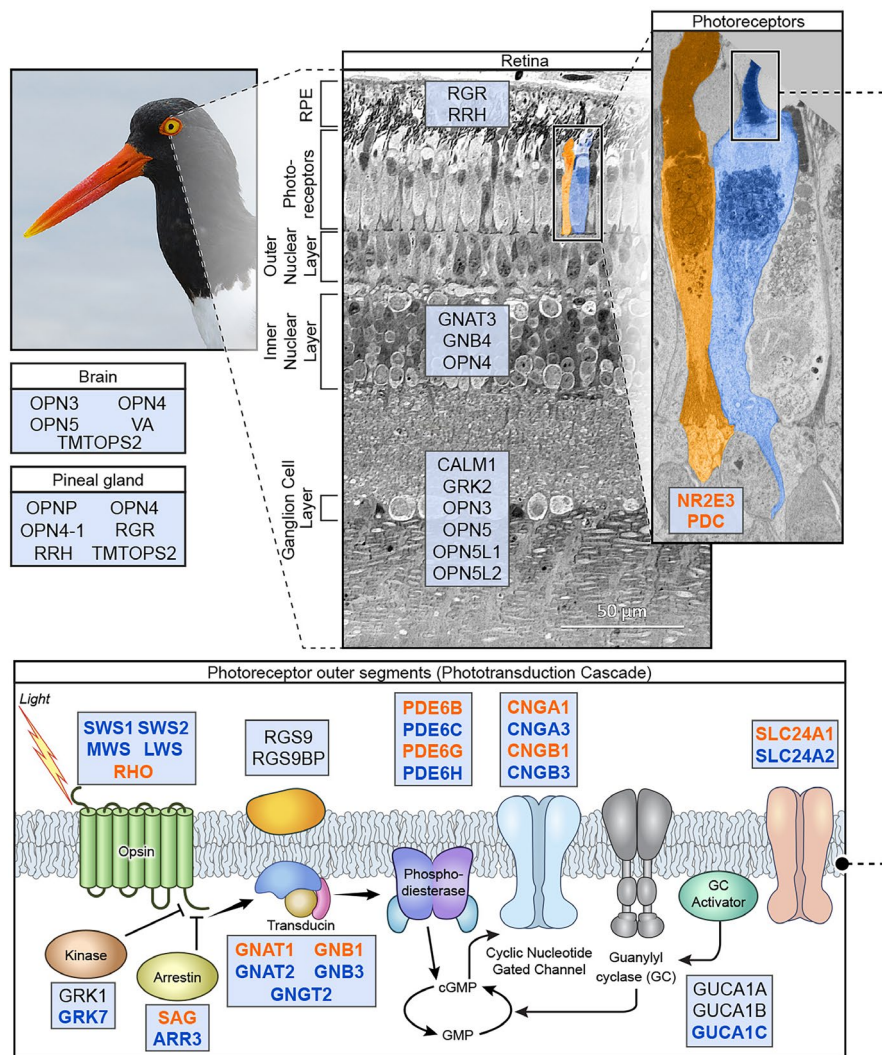
FIGURE 1 Location of expression of genes in the vision exome probe set. Rod-specific genes labelled in orange, cone-specific genes in blue, and genes expressed either in both cell types or outside of photoreceptors in black. RPE, retinal pigment epithelium. American oystercatcher (*Haematopus palliatus*) photo by Daniel J. Field, University of Cambridge

exon number across targeted genes. For these reasons, all taxa other than chicken and zebra finch were eventually excluded from the probe set design.

Potential probe sequences were masked using the RepeatMasker "cross_match" engine, soft-masking all vertebrate simple repeats and low-complexity DNA (Smit et al., 2013–2015). Probe sequences containing more than 10 bp of masked sequences were removed; probes with strings of 10 or fewer masked bases were retained but in those masked regions "N"s were replaced with "T"s. All potential probes were then aligned to the chicken and zebra finch genomes using BLAST+, and a hybridization melting temperature (T_m) was estimated for each hit assuming standard MYbaits reaction conditions and buffers (protocol v3.02). For each probe, one BLAST hit with the highest T_m was assumed to be the genome hit and was retained. Non-specific probes were then further filtered according to the my-BAITS "relaxed filtering" protocol: probe sequences were kept if they had a maximum of 10 BLAST+ matches with a T_m of 62.5–65°C, four matches above 65°C, and fewer than two passing probes on each flank (to avoid biasing recovery of any particular target). The resulting 2,500 probe sequences were synthesized in RNA and constitute our vision exome probe set (probe sequences available in File S2).

2.2 | Exome capture and sequencing

Two sets of test samples were prepared; a "diverse set" comprising 32 non-manakin bird species selected to represent a broad sampling of avian diversity at the ordinal level and a "manakin set" comprising 14 species of manakins (Table S2). We also included one crocodilian outgroup, the American alligator (*Alligator mississippiensis*). Tissue samples were assembled through loans from major museum collections (Table S2). For the diverse set, DNA was extracted using a phenol-chloroform protocol (Rosel & Block, 1996). For the manakin set, DNA was extracted from ~25 mg of pectoral muscle tissue using a Qiagen DNeasy Blood and Tissue kit. Extracted DNA from all samples was sheared to 200–650 bp via sonication, and Illumina libraries were prepared using KAPA Biosystems library preparation kits ("LTP" and "HyperPrep"; KAPA Biosystems, Inc.), incorporating a unique single-index adapter for each taxon. Libraries were amplified with 8–10 cycles of PCR and combined into pools of 7–8 species.

Target sequences were captured using the MYcroarray MYbaits protocol (v3.02), with the following modifications. First, chicken *Cot-1* DNA was used instead of human ("BLOCK no. 1"; Applied

Genetics Laboratories, Inc.), and a custom adapter blocker matching the "TruSeq-style" Illumina adapter sequence was used instead of the provided "BLOCK no. 3" (Integrated DNA Technologies, Inc.). Our 2,500 probe sequences were synthesized in a 20,000 array size, and 1 μl of the probe set suspended in 4.5 μl RNase-free H_2O was added per reaction at the 5.5 μl volume specified in the protocol. The last step of the capture protocol includes an amplification step, which was conducted at 14–16 cycles. In every reaction, 500 ng of pooled library DNA went into the capture reaction, and hybridization was carried out at 65°C for 24 hours. Capture success was verified by qPCR: five custom primer pairs, each targeting one chicken exon sequence from five different genes (*CALM1*, *MWS*, *GUCA1A*, *GNB1*, *GRK7*; File S3) present in the probe set were designed and used to quantify (via qPCR) capture of those loci in pooled libraries that underwent probe set capture versus the same pooled libraries that did not. In the diverse set of taxa, mean capture of loci was 534× (224–690×) and in the manakin set the capture of loci was on average 423× (127–1,371×). Our cutoff for successful capture was 100×, and therefore both indicated successful capture of targeted loci. Captured, amplified pools were size-selected via gel extraction to 500–750 bp before sequencing across two, paired-end 300 bp Illumina MiSeq runs. Captured raw reads are available from the NCBI Sequence Read Archive (SRA) under BioProject number PRJNA727529.

2.3 | Low-coverage whole genome sequencing

Shotgun genome libraries of the manakin set of taxa were prepared for low-coverage WGS, starting with the same libraries generated for the vision exome capture reaction. Libraries were size-selected via gel extraction to 400–600 bp, and all 14 species were pooled and sequenced on two, paired-end 150 bp lanes of an Illumina HiSeq X. Four species were resequenced on a paired-end 300 bp Illumina MiSeq run to increase coverage. Raw genome reads are available from the NCBI SRA under BioProject number PRJNA727529.

2.4 | Guided assembly reference generation and the importance of "decoy" sequences

During initial analyses, we identified regions of two vision genes under strong selection in an unlikely phylogenetic pattern. To explore further, we BLASTed the genes against the NCBI non-redundant database and found that the amino acids that appeared to be subject to positive selection differed from the those in published sequence. Instead, our putatively selected amino acids aligned perfectly with closely related paralogs that function in olfaction. These preliminary results indicated our initial assemblies had generated chimeric sequences which have the potential to yield highly misleading results in downstream evolutionary analyses such as the false selection positives seen here.

To avoid the generation of chimeric sequences, we generated reference assembly guides that included "decoy" sequences representing the whole exomes of both chicken (Genome Reference Consortium Chicken Build 6a; RefSeq assembly accession GCF_000002315.6) and zebra finch (bTaeGut2.pat.W.v2; RefSeq assembly accession GCF_008822105.2), downloaded from the NCBI Assembly database (Kitts et al., 2016). Assembly guides were created as follows: first, our targeted vision genes were removed from both decoy sets. Second, in order to remove redundancy from the decoy sequences, only the longest transcript was kept from any given gene. Third, the remaining chicken and zebra finch sequences were combined into one set of off-target decoys (File S4; White, 2021a). Finally, we added back in the species-specific vision genes to the combined decoy set separately for chicken and zebra finch, resulting in two assembly guides (Files S5 and S6). In these genes, we noted an unexpected degree of variation in exon number for the same gene between chicken and zebra finch, and so only exons present in both species were kept in their respective assembly guide.

Following the publications listed in section 2.1 to create a list of target genes, we had added published recoverin (*RCVRN*) sequences for chicken and zebra finch to our probe set (Table S1). Since then, new work has further delineated the neuronal calcium sensor family of proteins, of which *RCVRN* and visinin are members (Lamb & Hunt, 2018). That work determined that *RCVRN* has been lost from Sauropsida (birds and non-avian reptiles) while visinin has been lost from mammals, though mammals have a distinct gene called "visinin-like," further confusing the matter. NCBI's GenBank currently contains sequences identified as all three—*RCVRN*, visinin and visinin-like—for a variety of bird species. Delineating the true identity of these genes is beyond the scope of this work, and so while *RCVRN* appears in our probe set, we dropped it from further analysis. In future, when these proteins have been better characterized, sequence recovered by our *RCVRN* probes may prove useful.

2.5 | Bioinformatic processing and assembly

Read quality control and adapter sequence removal was performed for all data using Trimmomatic (v0.39; Bolger et al., 2014). All guided (read-mapped) assemblies were generated using BWA-MEM (v0.7.17; Li, 2013) with mismatch penalty increased to 8 (from 4; "-B 8") as an additional way to deter the generation of chimeric sequences. Assemblies were generated at the exon level, and subsequently concatenated into full coding sequences (CDS). Detailed steps of our bioinformatic pipeline along with custom scripts written to facilitate the process are available from GitHub (White, 2021b). Briefly, the SAM file output from BWA-MEM is processed using SAMtools (v1.10, Li et al., 2009) and BEDTools (v2.29.2; Quinlan & Hall, 2010) to output consensus files that replace missing bases with "N"s. This involves creating a masking file so that reference sequence is not inserted where our experimental sequences do not cover a particular region of the reference. Our pipeline was modeled after Schott et al. (2017).

De novo assemblies were generated from vision exome capture reads using Trinity (v2.8.4; Grabherr et al., 2011). These assemblies were used to: (1) assess how many probes had captured sequence in each species, and (2) confirm identity of each tissue sample, where possible, by assembling the mitochondrial genome. Mitochondrial genome "bycatch" has been observed in previous in-solution hybridization experiments using DNA from tissues rich in mitochondria (e.g., muscle tissue; Tamashiro et al., 2019). To identify mitochondrial assemblies, we performed a BLAST+ search against the NCBI non-redundant database (accessed January 2021) for each de novo contig ≥10 kb. These mitochondrial genomes are available in File S7.

2.6 | Whole-genome sequencing simulation

To estimate the depth of WGS needed to achieve the same level of coverage of vision exome genes as assembled from our empirical vision exome capture data, WGS reads were randomly subsampled without replacement using SAMtools (v1.10). We used WGS reads from the two of our manakin samples with the highest number of total reads, *Xenopipo atronitens* and *Cryptopipo holochlora*. We generated data sets of 150 bp paired-end reads ranging from 1–10× sequencing depth of the genome, which is about 1.095 GB for birds. Three replicates were generated for every level of coverage and guide-assembled to the finch assembly reference using BWA-MEM as described above. Coverage values were averaged across replicates.

2.7 | Alignment and selection tests

Multiple sequence alignments were prepared for each gene's CDS using the ALigning, Filtering, and eXporting (ALFiX) pipeline implemented in the multiple alignment of coding sequences (MASCE) program (v2.04; Ranwez et al., 2011, 2021). Briefly, unaligned, reference-guided gene assemblies were trimmed to remove non-homologous fragments, then aligned simultaneously at the nucleotide and amino acid levels to prevent the introduction of frameshifts and internal stop codons during alignment. Next, isolated codons and patchy regions of the alignments were filtered out using HmmCleaner and this final, masked alignment was exported with internal stop codons and frameshifts replaced by "NNN". Finally, we aligned these sequences to reference protein sequences from chicken (accession numbers in Table S1) and trimmed 5′- and 3′-untranslated regions (final alignments available in File S8). The alignments for *GNGT2*, *SWS1*, and *RGS9BP* were poor, so those genes were excluded from selection test analyses.

Selection tests were performed to identify evidence of positive selection related to either the evolution of nocturnality or visual adaptation specific to manakins, including high-speed vision. A tree of all 46 of our ingroup taxa was constructed based on published phylogenies of relevant groups (Figure 2; Kimball et al., 2019; Leite et al., 2021; McKay et al., 2010; Ohlson et al., 2013; Reddy et al., 2017; Silva et al., 2018; White & Braun, 2019). For the nocturnal selection tests, all but one manakin species (the one with the best coverage) was removed to avoid overrepresentation of Pipridae (Figure 2a). For the manakin tests, all but one member of
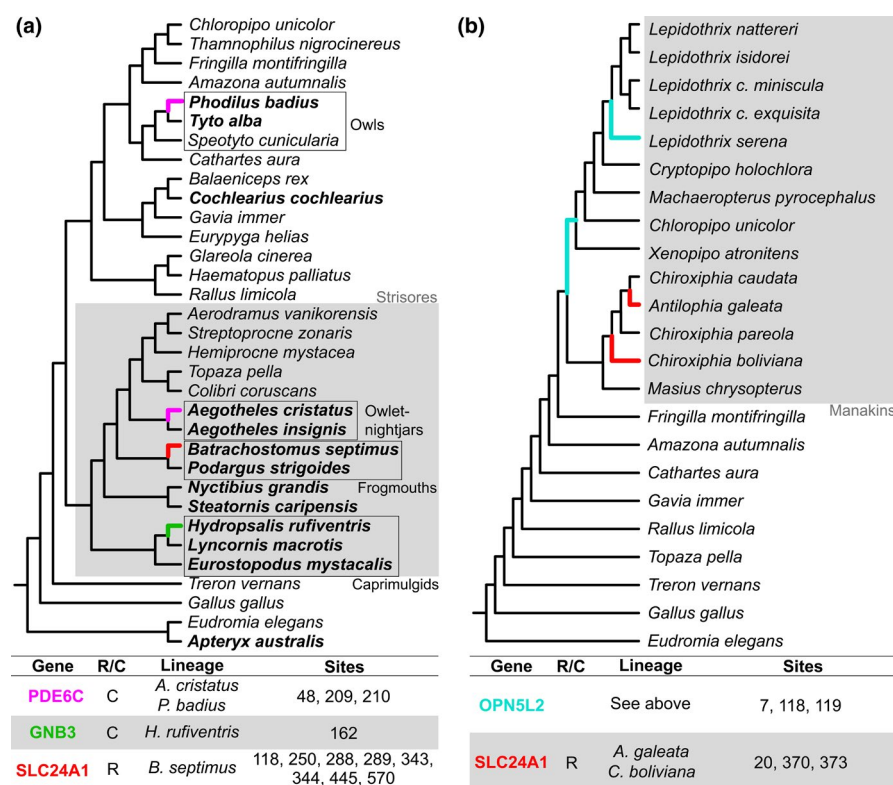

FIGURE 2 Results of selection tests related to the evolution of nocturnality in birds (a) and evolution of visual adaptation specific to manakins (e.g., high-speed vision; b). Colored branches correspond to colored genes in tables below. Numbers are amino acid positions within each alignment. R, rod; C, cone. Results are all significant with *p* ≤ .05. Species in bold are nocturnal. Trees derived from published phylogenies, see main text for references

Table (a):

Gene	R/C	Lineage	Sites
PDE6C	C	*A. cristatus* *P. badius*	48, 209, 210
GNB3	C	*H. rufiventris*	162
SLC24A1	R	*B. septimus*	118, 250, 288, 289, 343, 344, 445, 570

Table (b):

Gene	R/C	Lineage	Sites
OPN5L2		See above	7, 118, 119
SLC24A1	R	*A. galeata* *C. boliviana*	20, 370, 373

each major lineage of the diverse set was removed, to avoid over-representation of any particular lineage (Figure 2b).

All selection tests were performed using the Hypothesis Testing using Phylogenies (HyPhy) software package (v2.5.29; Kosakovsky Pond et al., 2005). To determine if a gene had undergone episodic positive selection within our tree, each gene was tested using the Branch-Site Unrestricted Statistical Test for Episodic Diversification (BUSTED) model (Murrell et al., 2015). Next, to identify individual branches with evidence of positive selection, we used the adaptive Branch-Site Random Effects Likelihood (aBSREL) model (Smith et al., 2015). Finally, to determine if individual sites were subject to episodic positive selection in branches of interest, we applied the Mixed Effects Model of Evolution (MEME) model (Murrell et al., 2012). To clarify, we used BUSTED to identify gene-level selection, and on those genes ran aBSREL and MEME to identify when (in terms of lineages) and where (in terms of sites) that selection occurred, respectively. BUSTED and aBSREL test a priori hypotheses (e.g., do our nocturnal branches differ from our diurnal branches?) while MEME operates without any specific a priori hypotheses. Sites with ≥20% missing or ambiguous amino acids were ignored. All tests were run specifying individual branches or monophyletic lineages of interest relevant to either the nocturnal (see taxa in bold in Figure 2a) or manakin-specific investigations (see grey box in Figure 2b).

3 | RESULTS

3.1 | Sequencing and probe set performance

For the vision exome capture data, an average of 1,190,841 trimmed reads were recovered per ingroup species (range: 435,723–3,189,129; Table S2). For the WGS data, an average of 29,624,754 reads were recovered per species (range: 13,870,766–46,848,995; Table S2).

We mapped our probe sequences to our de novo contigs in order to estimate how many probes were likely to have captured sequences from each species. An average of 715 probes captured sequences from each of our ingroup species (range: 435–1,885). Across all species, 2,332 of the 2,500 unique probes (93%) captured sequences from at least one species in our data set (Table S2). We then mapped our probes to our assembly references (without decoy sequences) and found that our probes covered 85% of the chicken assembly reference and 64% of the finch assembly reference, both to an average 2× depth (Table S3). These numbers are below 100% for two potential reasons: first, in the probe design process, some parts of the reference would not be covered due to its respective probe being repetitive, or not having an appropriate T_m. Second, there was a lapse of time between downloading the published sequences used to design the probes and downloading the published whole exomes used to create the assembly references, and sequences of the former may have been revised or removed in the meantime. Nonetheless, downloading newer exome sequences for our assembly references

allowed us to leverage improved assemblies and utilize potentially more biologically accurate sequences for downstream analysis.

The percent of reads on-target in our exome data—16% for all ingroup taxa (Figure 3)—is comparable to previously published in-solution hybridization studies (e.g., Faircloth et al., 2015; Kieran et al., 2019).

3.2 | Sequencing depth simulation

We recovered an average of 71% of our targeted exome from our manakin samples (Figure 3). We subsampled WGS reads from two manakins at different levels of genome coverage to estimate what depth of sequencing of the whole genome is needed in order to achieve a comparable level of coverage across our vision genes. Our subsampled data indicated that in order to achieve an average of 71% coverage, one would need to sequence a 1.095 GB bird genome at 6–7× depth (Table S4), which is what we observed with our empirical data (Figure 3). Using the probe set, we achieved the same level of coverage (71%) with sequencing data equivalent to a ~0.88× genome coverage depth.

3.3 | Guided assembly reference performance

Initial attempts to align our exome sequences to vision gene references resulted in the generation of chimeric sequences for genes that are part of protein families and have closely related paralogs. Chimeric sequences can yield highly misleading results, such as false positives in analyses of evolutionary selection (see examples illustrated in Figure S1). We hypothesized that including other members of these gene families (off-target "decoys"; see section 2.4) in the assembly references would allow these sequences to be mapped to their correct gene of origin and eliminate chimeric sequences. We decided to include the whole exomes of both chicken and zebra finch since complete exomes would both include the paralogues of interest and provide a general solution to the chimera problem that could be used with any set of genomic targets. After including the whole exomes as decoys, we did not detect any evidence of chimeric sequence in the resulting assemblies.

All but four of our samples displayed a higher degree of on-target gene coverage when assembled to the zebra finch reference compared to assembly to the chicken reference (Table S3). The exceptions were chicken, *Apteryx australis*, *Eudromia elegans*, and *Batrachostomus septimus*. Coverage in *B. septimus* differed by only 0.84% between the two references, a negligible amount. The other three species are unsurprising, as chicken will undoubtedly assemble to itself better than to another species, and the other two species are equally phylogenetically related to chicken and zebra finch (Figure 2a). In that sense they are unique—all of our taxa except for these (chicken, *A. australis* and *E. elegans*) are evolutionarily more closely related to zebra finch than to chicken; therefore, a relatively higher level of coverage is unsurprising. Thus, we proceeded

	Species	# Reads (M)	Reads On-target (%)	Coverage (%) On-Target	Coverage (%) Off-Target	Depth On-Target	Depth Off-Target
Exome / Diverse Set	*Aegotheles cristatus*	0.7	14	56	4	140	3
	Aegotheles insignis	1.1	8	57	7	158	3
	Aerodramus vanikorensis	1.1	11	58	7	207	3
	Amazona autumnalis	0.7	11	56	5	138	3
	Apteryx australis	1.3	5	47	7	117	3
	Batrachostomus septimus	0.5	18	54	3	77	2
	Belaeniceps rex	0.4	13	61	3	87	3
	Cathartes aura	0.8	8	65	5	104	3
	Cochlearius cochlearius	0.8	9	65	5	111	3
	Colibri coruscans	1.4	7	57	8	167	3
	Eudromia elegans	0.8	7	46	5	88	3
	Eurostopodus mystacalis	1.1	4	60	8	81	3
	Eurypyga helias	1.1	4	58	7	64	3
	Fringilla montifringilla	1.3	5	86	12	121	3
	Gallus gallus	1.2	6	53	17	132	3
	Gavia immer	1.1	4	63	8	83	3
	Glareola cinerea	0.7	5	59	5	56	3
	Haematopus palliatus	0.7	6	61	5	64	3
	Hemiprocne mystacea	1.0	7	58	6	150	4
	Hydropsalis rufiventris	2.2	7	65	11	268	4
	Lyncornis macrotis	2.2	8	86	16	319	3
	Nyctibius grandis	0.8	9	61	5	125	3
	Phodilus badius	0.9	10	58	7	155	3
	Podargus strigoides	0.9	12	62	6	183	3
	Rallus limicola	1.1	7	59	6	134	3
	Speotyto cunicularia	0.7	9	56	5	120	3
	Steatornis caripensis	0.9	8	60	6	114	3
	Streptoprocne zonaris	0.8	7	60	6	100	3
	Thamnophilus nigrocinereus	1.0	6	68	7	113	3
	Topaza pella	1.0	5	57	7	104	3
	Treron vernans	0.6	7	62	4	72	3
	Tyto alba	1.1	8	58	7	128	3
Exome / Manakins	*Antilophia galeata*	0.7	37	69	1	847	18
	Chiroxiphia boliviana	1.2	38	72	2	1,407	20
	Chiroxiphia caudata	1.4	30	67	1	1,055	25
	Chiroxiphia pareola	1.5	32	70	2	1,476	22
	Chloropipo unicolor	1.2	36	75	2	1,259	21
	Cryptopipo holochlora	3.2	31	74	4	3,030	26
	Lepidothrix c. exquisita	1.7	35	72	2	1,674	21
	Lepidothrix c. miniscula	1.7	33	70	2	1,700	23
	Lepidothrix isidorei	1.6	30	70	2	1,403	20
	Lepidothrix nattereri	1.0	39	69	2	1,217	18
	Lepidothrix serena	1.1	39	70	2	1,174	19
	Machaeropterus pyrocephalus	1.2	40	69	2	1,501	21
	Masius chrysopterus	2.4	33	72	3	2,351	24
	Xenopipo atronitens	2.9	31	73	4	2,685	26
Genome	*Antilophia galeata*	13.9	0	63	37	5	5
	Chiroxiphia boliviana	22.9	0	70	43	6	6
	Chiroxiphia caudata	41.6	0	74	48	12	12
	Chiroxiphia pareola	27.8	0	73	45	10	9
	Chloropipo unicolor	20.7	0	72	41	6	6
	Cryptopipo holochlora	46.0	0	78	50	14	11
	Lepidothrix c. exquisita	30.2	0	75	47	8	8
	Lepidothrix c. miniscula	31.0	0	75	45	9	8
	Lepidothrix isidorei	20.6	0	71	42	7	6
	Lepidothrix nattereri	28.5	0	75	46	8	7
	Lepidothrix serena	22.1	0	70	42	6	6
	Machaeropterus pyrocephalus	23.9	0	71	43	6	6
	Masius chrysopterus	38.8	0	76	49	11	10
	Xenopipo atronitens	46.8	0	78	50	13	12
	Manakin average	**1.6**	**35**	**71**	**2**	**1,627**	**22**
	Exome average	**1.2**	**16**	**64**	**5**	**584**	**9**
	Genome average	**29.6**	**0**	**73**	**45**	**9**	**8**

FIGURE 3 Summary statistics for assemblies of the exome and genome data to the zebra finch reference. *Note:* Reads is total number of trimmed, quality-controlled reads per sample in millions. Shading is done separately for reads on-target, coverage, and depth, with lowest values in white and highest values in dark gray. Shading is done separately for averages at bottom of table. On-target reads are those assembling to the reference PTC gene set; off-target reads are those assembling to the "decoy" sequences (see Methods)

with assemblies made to the zebra finch reference for all downstream analyses. Our manakin set of samples were more successfully captured than the diverse set (percent reads on-target 35% for manakins vs. 8% for the diverse set; Table S3). We are unsure of why this would happen but note that laboratory work was done separately for the manakins and the diverse set samples, though it was done by the same person (NDW) and with the same reagents. Phylogenetic divergence could be a factor, as the manakins are more closely related to the zebra finch probe sequences than other non-passerines (e.g., the other diverse set taxa) would be. Regardless, we consider the manakin set statistics as representative of how well the probe set can perform. We summarize our results for the exome capture data overall as well as for the diverse set and the manakin set, separately.

3.4 | Vision gene capture

With few exceptions, some sequence of every gene was captured from every ingroup species included in this study (exome data average = 45 genes) and 41 out of 46 genes for the outgroup, *Alligator mississippiensis* (Table S2). Guided assembly of vision exome capture data for all ingroup species resulted in 64% capture of targeted sequence on average (60% in the diverse set and 71% in the manakins; Figure 3, Table S3). Guided assembly of WGS data for manakins was comparable to that of the exome data, with 73% coverage. Average depth of coverage for all on-target contigs was 584× for the exome data (127× in diverse set and 1,627× in manakins), whereas depth of coverage averaged 9× for the WGS data. Depth
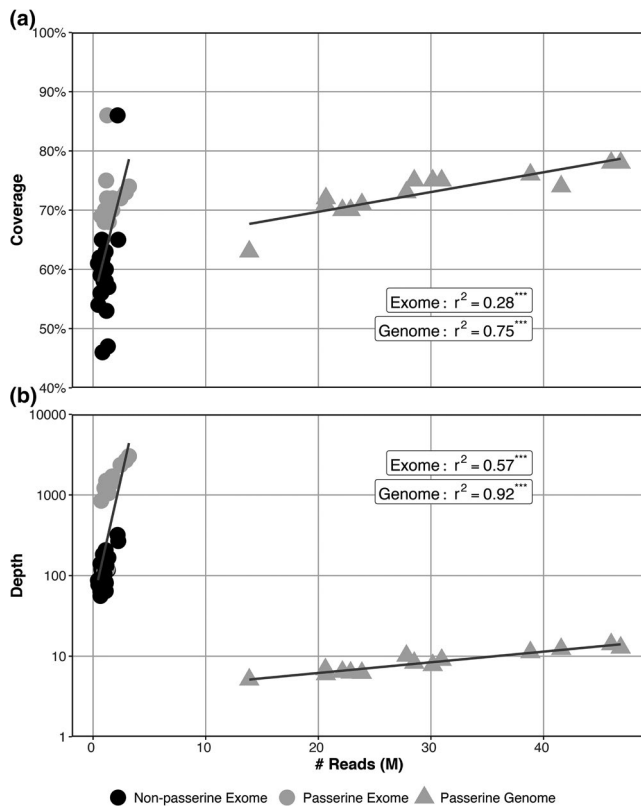
FIGURE 4 Number of reads by percent coverage (a) and depth of sequencing (b) for exome (circle) and genome (triangle) data. Data points from passerine species are colored in grey, non-passerines in black. Depth is in log10 scale. Reads is total number of trimmed, quality-controlled reads per sample in millions. ***$p \leq .001$

of coverage for the WGS data did not vary significantly between on- and off-target contigs ("decoy" sequence contigs). Given that our WGS samples had more than ten times the reads compared to what our exome capture samples had, it is clear that our probe set has successfully captured the targeted sequences. In fact, increasing reads had little effect on either the percent of coverage of on-target genes or the depth of sequence covered (Figure 4).

It has been reported from surveys of de novo genomes that nocturnal birds do not have *SWS1* (Borges et al., 2015; Cho et al., 2019; Feng et al., 2020), but our data directly contradict those reports. We recovered substantial sequence from *SWS1* for the great eared nightjar (*Lyncornis macrotis*) and partial *SWS1* sequences for several other nocturnal birds (*Apteryx australis*, *Cochlearius cochlearius*, *Eurostopodus mystacalis*, *Nyctibius grandis*, *Podargus strigoides* and *Steatornis caripensis*). The great eared nightjar *SWS1* sequence has been submitted to GenBank with accession number OK147115.

3.5 | Analysis of selection

In tests of selection comparing all nocturnal birds to all diurnal birds (see bolded taxa in Figure 2a), we found evidence of positive selection in three vision-associated genes, *PDE6C*, *GNB3* and *SLC24A1*

(Table S5). All three were positively selected within various nightbird (superorder Strisores) lineages, one of which was identified in both owlet-nightjars and owls (*PDE6C*).

Tests of selection on genes related to vision within manakins (see grey box in Figure 2b) also revealed evidence of positive selection on two genes, *OPN5L2* and *SLC24A1* (Table S5).

4 | DISCUSSION

Herein, we developed a probe set that successfully and efficiently captured exonic regions of 46 genes representing the PTC and other light responses from a wide diversity of birds. Our data indicates that this probe set can be used on any extant species of bird. In developing the bioinformatic pipeline for assembling our data, we identified an unexpected problem: simply assembling to our targeted genes yielded some chimeric sequences that could be misleading in downstream analyses. We solved this problem by including "decoy" sequences corresponding to the whole exome in our assembly guides. We also found that low-coverage WGS was able to recover many sequences successfully, albeit with much lower coverage than the exome capture method. Finally, we used our vision gene sequences to elucidate potential genetic mechanisms of visual adaptation in two distinct evolutionary scenarios.

4.1 | Genomic evidence for evolutionary changes in the PTC

We have demonstrated the utility of the vision probe set by applying the exome data towards two evolutionary questions. In the first, we investigated potential signatures of episodic positive selection among PTC genes associated with the independent emergence of nocturnality in four major avian lineages—owls, herons (represented by *Cochlearius cochlearius*), nightbirds (superorder Strisores) and kiwis (represented by *Apteryx australis*; Figure 2a).

The Strisores superorder contains five distinct nocturnal lineages and multiple hypotheses have been offered regarding the number of independent diurnal-to-nocturnal transitions among these lineages (reviewed in White & Braun, 2019). Transitioning from a diurnal to a nocturnal lifestyle presumably requires a suite of morphological and genetic adaptations, likely including specialization of vision genes to support night vision. Morphological evidence supports at least two independent nocturnal transitions in the Strisores—one associated with the development of the tapetum lucidum in caprimulgids and potoos (represented in our dataset by *Nyctibius grandis*), and a second coinciding with the evolution of triplicate photoreceptor layers in the oilbird (*Steatornis caripensis*). Limited research has not identified night vision-associated morphological adaptations in owlet-nightjars or frogmouths. Here, we found three distinct adaptations in PTC genes among nocturnal Strisores lineages (Figure 2a) including adaptation in both calcium-dependent (*SLC24A1*) and calcium-independent (*PDE6C*) gene functions. We also note one selected

site present in caprimulgids for gene *GNB3*, a subunit of the cone version of transducin. Transducin is the gene responsible for light response and altering the rate of cone transducin activation controls cone light saturation (Lobanova et al., 2010). In contrast, we did not identify any genes under selection in the ancestral branches shared by nocturnal Strisores families; these results are consistent with the hypothesis of multiple independent nocturnal transitions within this superorder.

Intriguingly, adaptation in *PDE6C* was observed in both owlet-nightjars and owls (family Strigidae). *PDE6C* encodes a cone-specific homodimer protein that hydrolyses cGMP leading to the closure of cGMP-gated channels and propagation of the visual signal. Cones are not thought to be involved in night vision but instead are involved in day light vision and also during dim light conditions, such as those at dawn and dusk. Pugh et al. (1999) listed *PDE* (of which *PDE6C* is a subunit) as one of eight hypothetical mechanisms of vertebrate light adaptation, citing its potential for speeding photoreceptor recovery and decreasing flash sensitivity, and noting that it is the principal calcium-independent mechanism of photoreceptor adaptation. Thus, changes in *PDE6C* could indicate modification of vision enabling crepuscular activity. Parallel diversification of *PDE6C* in owlet-nightjars and owls suggests that positive selection on primarily diurnal genes may be a convergent feature among diverse nocturnal species. We hypothesize that such a predisposition could underlie the proliferation of species among Strisores and Strigidae, which are the only two speciose lineages of nocturnal birds.

In the second evolutionary scenario, we tested for positive selection associated with the divergence and diversification of manakins—a group of neotropical birds known for their complex lekking displays and high-speed movement (Figure 2b). We observed significant positive selection specific to manakins in a total of two genes. One such gene in the manakins is *OPN5L2*, a light-responsive opsin present in retinal amacrine and ganglion cells, as well as the hypothalamus, where it is involved in non-image-forming responses to light (Ohuchi et al., 2012; Yamashita, 2020).

Surprisingly, we observe positive selection on *SLC24A1* in both of our nocturnal and manakin tests. This gene modulates the rate of rod photoresponse recovery, i.e., the rate at which the cell can return to a dark state and generate a new response. *SLC24A1* forms a channel that removes cytoplasmic calcium and alters rod recovery rate by altering calcium dynamics. What might explain the repeated signatures of positive selection in a rod gene among the diurnal manakin clade? One possibility is that maintaining this gene may incur an energetic or morphological cost which has repeatedly been ameliorated by novel mutations within manakins. Alternatively, manakins may be undergoing adaptation for faster rod recovery and improved rod signal amplification. Rod adaptations could increase sensitivity in dim or mesopic light conditions such as those present under the dense canopy in the neotropical forests where these manakins reside, potentially enhancing temporal resolution (high-speed vision) and even color discrimination (Lamb & Pugh, 2006a,2006b; Umino et al., 2012, 2019; Zele & Cao, 2015). These same adaptations could also enable or enhance crepuscular visual activity in nocturnal birds.

The identification of selection of *SLC24A1* in nocturnal birds is confirmed by Wu et al. (2016), who identified selection of this gene in owls when compared to diurnal raptors.

4.2 | Considerations for selecting a probe set or WGS

While sequencing costs are decreasing as next-generation sequencing technologies continue to develop (Sboner et al., 2011; Schwarze et al., 2020), sequencing high-quality, complete genomes or transcriptomes may be unnecessary and even counterproductive for research focused on a particular set of genes. Massive, unfiltered data sets can introduce bioinformatic challenges including whole genome assembly and annotation of unknown transcripts in non-model species. In contrast, targeted sequence capture primarily allows efficient capture of genomic regions of interest and therefore enables sequencing power to be used for additional species or samples. This is particularly critical for intraspecific studies, which require many individuals of the same species and for which generating whole genomes may be unnecessary, costly, and time-consuming.

The exome probe set presented here is advantageous to anyone wishing to study avian visual adaptation in a broad diversity of species or in many intraspecific samples by efficiently sequencing vision genes. Sequence capture is also feasible with old and/or highly degraded samples (e.g., McCormack et al., 2016), such as might come from a museum specimen (e.g., toe pad of a study skin). This allows the use of existing museum collections as opposed to transcriptome sequencing, which requires the acquisition of fresh material specifically preserved for extraction of high-quality RNA. We used this probe set to sequence genes in birds from 14 orders covering all major avian groups, demonstrating that this probe set should capture vision genes from any extant bird lineage. However, we note an important phylogenetic pattern in the success of sequence assembly and suggest that future users of this probe set create assembly references as closely related to their species of interest as possible.

Looking beyond this specific probe set, we believe that our results provide guidance for investigators that are choosing between sequence capture and low-coverage WGS to recover a particular subset of the genome. Sequence capture probe sets are extremely flexible and can be designed to target any portions of the genome. One could even combine multiple probe sets designed for different purposes (e.g., ultraconserved elements for phylogenomics and the PTC pathway for analysis of the molecular evolution of vision) and capture sequences for multiple projects in one assay. The primary advantage of sequence capture is that this method provides greater depth of coverage for genomic regions of interest and is therefore very cost-effective, particularly when many samples are to be examined. On the other hand, low-coverage WGS avoids the expensive and time-consuming process of probe design (although that has obviously already been done here for the avian PTC) and additional laboratory work for sequence capture. Moreover, by randomly sequencing the genome, WGS can potentially facilitate unplanned

future research efforts. Finally, the costs of synthesizing custom probes may not drop significantly in the future, whereas sequencing costs have been decreasing. As a consequence, the per sample costs of sequence capture, currently lower than WGS unless very few species are examined, could remain relatively flat while WGS becomes more affordable. Of course, the cost of WGS will also vary on genome size; birds have relatively small genomes for vertebrates (see Kapusta et al., 2017), and WGS will be more affordable in birds than in other groups.

With all of these factors in mind, we are left with a question: when should we favor one approach over the other? We argue sequence capture should be favored when: (1) it is necessary to sequence many samples since a large number of samples will defray the costs of the additional labor and probes, (2) the research question is focused on a subset of the genome (e.g., the exome), (3) obtaining fresh, high quality tissues may not be possible, and/or (4) it is necessary to resolve alleles and call heterozygotes accurately such as in intraspecific population studies. WGS should be favored when those factors are less relevant to an investigator and when the potential to randomly sequence other parts of the genome and/or non-coding sequence is desired. Finally, the future trajectory of the costs for sequencing and probes will play into this decision.

The specific probe set that we have constructed provides a way forward to study the thousands of poorly studied bird species with potentially interesting visual phenotypes, opening the door to rapid and inexpensive surveys of the PTC, and subsequent inference into the evolution of visual adaptation. More generally, our results illustrate some lessons for similar studies, especially those that use museum specimens, regardless of whether they are focused on vision or other biological processes. Specifically, we illustrated the advantages for exome capture (the much deeper coverage), provide a solution to cases where assemblies yield chimeras (the use of "decoy" sequences), and offer guidance regarding some of the trade-offs between sequence capture and low-coverage WGS.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

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 in GenBank with project ID PRJNA727529.

DATA AVAILABILITY STATEMENT

Raw read data for exomes and WGS is available in the NCBI SRA archive (BioProject ID PRJNA727529). Probe sequences, primer sequences, guided assembly exon references, decoy sequences and trimmed alignments are available in the supplement (decoy sequences also available as a Zenodo dataset; White, 2021a). Bioinformatic scripts for processing this data are available on Noor D. White's Github page (https://github.com/ndwhite; White, 2021b). The *SWS1* sequence from *Lyncornis macrotis* is available on GenBank with accession number OK147115. Mitochondrial genomes are available in File S7.

ORCID

Noor D. White https://orcid.org/0000-0002-9510-3744
Zachary A. Batz https://orcid.org/0000-0002-4483-2402
Edward L. Braun https://orcid.org/0000-0003-1643-5212
Michael J. Braun https://orcid.org/0000-0001-8844-1756
Karen L. Carleton https://orcid.org/0000-0001-6306-5643
Rebecca T. Kimball https://orcid.org/0000-0001-5449-5481
Anand Swaroop https://orcid.org/0000-0002-1975-1141

REFERENCES

Bloch, N. I., Morrow, J. M., Chang, B. S., & Price, T. D. (2015). SWS2 visual pigment evolution as a test of historically contingent patterns of plumage color evolution in warblers. *Evolution*, *69*(2), 341–356. https://doi.org/10.1111/evo.12572

Bloch, N. I., Price, T. D., & Chang, B. S. (2015). Evolutionary dynamics of Rh2 opsins in birds demonstrate an episode of accelerated evolution in the New World warblers (Setophaga). *Molecular Ecology*, *24*(10), 2449–2462.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Borges, R., Khan, I., Johnson, W. E., Gilbert, M. T. P., Zhang, G., Jarvis, E. D., O'Brien, S. J., & Antunes, A. (2015). Gene loss, adaptive evolution and the co-evolution of plumage coloration genes with opsins in birds. *BMC Genomics*, *16*, 751. https://doi.org/10.1186/s12864-015-1924-3

Bostwick, K. S., & Prum, R. O. (2003). High-speed video analysis of wing-snapping in two manakin clades (Pipridae: Aves). *Journal of Experimental Biology*, *206*, 3693–3706. https://doi.org/10.1242/jeb.00598

Bowmaker, J. K., Heath, L. A., Wilkie, S. E., & Hunt, D. M. (1997). Visual pigments and oil droplets from six classes of photoreceptor in the retinas of birds. *Vision Research*, *37*(16), 2183–2194. https://doi.org/10.1016/S0042-6989(97)00026-6

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Chen, D., Braun, E. L., Forthman, M., Kimball, R. T., & Zhang, Z. (2018). A simple strategy for recovering ultraconserved elements, exons, and introns from low coverage shotgun sequencing of museum specimens: Placement of the partridge genus Tropicoperdix within the galliformes. *Molecular Phylogenetics and Evolution*, *129*, 304–314. https://doi.org/10.1016/j.ympev.2018.09.005

Cho, Y. S., Jun, J. H., Kim, J. A., Kim, H.-M., Chung, O., Kang, S.-G., Park, J.-Y., Kim, H.-J., Kim, S., Kim, H.-J., Jang, J.-H., Na, K.-J., Kim, J., Park, S. G., Lee, H.-Y., Manica, A., Mindell, D. P., Fuchs, J., Edwards, J. S., … Bhak, J. (2019). Raptor genomes reveal evolutionary signatures of predatory and nocturnal lifestyles. *Genome Biology*, *20*(181), 1–11. https://doi.org/10.1186/s13059-019-1793-1

Colella, J. P., Tigano, A., & MacManes, M. D. (2020). A linked-read approach to museomics: Higher quality de novo genome assemblies from degraded tissues. *Molecular Ecology Resources*, *20*(4), 856–870. https://doi.org/10.1111/1755-0998.13155

Coyle, B. J., Hart, N. S., Carleton, K. L., & Borgia, G. (2012). Limited variation in visual sensitivity among bowerbird species suggests that there is no link between spectral tuning and variation in display colouration. *Journal of Experimental Biolbiology*, *215*(7), 1090–1105. https://doi.org/10.1242/jeb.062224

Day, L. B., Helmhout, W., Pano, G., Olsson, U., Hoeksema, J. D., & Lindsay, W. R. (2021). Correlated evolution of acrobatic display and both neural and somatic phenotypic traits in Manakins (Pipridae). *Integrative and Comparative Biology*, icab139.

Díaz, N. M., Morera, L. P., & Guido, M. E. (2015). Melanopsin and the non-visual photochemistry in the Inner Retina of vertebrates. *Photochemistry and Photobiology*, *92*(1), 29–44. https://doi.org/10.1111/php.12545

Faircloth, B. C., Branstetter, M. G., White, N. D., & Brady, S. G. (2015). Target enrichment of ultraconserved elements from arthropods provides a genomic perspective on relationships among Hymenoptera. *Molecular Ecology Resourources*, *15*(3), 489–501.

Feng, S., Stiller, J., Deng, Y., Armstrong, J., Fang, Q., Reeve, A. H., Xie, D., Chen, G., Guo, C., Faircloth, B. C., Petersen, B., Wang, Z., Zhou, Q., Diekhans, M., Chen, W., Andreu-Sanchez, S., Margaryan, A., Howard, J. T., Parent, C., & Zhang, G. (2020). Dense sampling of bird diversity increases power of comparative genomics. *Nature*, *587*(7833), 252–257.

Fu, Y., & Yau, K.-W. (2007). Phototransduction in mouse rods and cones. *Pflügers Archiv - European Journal of Physiology*, *454*(5), 805–819. https://doi.org/10.1007/s00424-006-0194-y

Fusani, L., Giordano, M., Day, L. B., & Schlinger, B. A. (2007). High-speed video analysis reveals individual variability in the courtship displays of male golden-collared manakins. *Ethology*, *113*, 964–972. https://doi.org/10.1111/j.1439-0310.2007.01395.x

Fuxjager, M. J., Goller, F., Dirkse, A., Sanin, G. D., & Garcia, S. (2016). Select forelimb muscles have evolved superfast contractile speed to support acrobatic social displays. *eLife*, *5*, e13544.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Guschanski, K., Krause, J., Sawyer, S., Valente, L. M., Bailey, S., Finstermeier, K., Sabin, R., Gilissen, E., Sonet, G., Nagy, Z. T., Lenglet, G., Mayer, F., & Savolainen, V. (2013). Next-generation museomics disentangles one of the largest primate radiations. *Systematic Biology*, *62*(4), 539–554. https://doi.org/10.1093/sysbio/syt018

Hankins, M. W., & Hughes, S. (2014). Vision: Melanopsin as a novel irradiance detector at the heart of vision. *Current Biology*, *24*(21), R1055–R1057.

Hart, N. S. (2001a). The visual ecology of avian photoreceptors. *Progress in Retinal and Eye Research*, *20*(5), 675–703. https://doi.org/10.1016/S1350-9462(01)00009-X

Hart, N. S. (2001b). Variations in cone photoreceptor abundance and the visual ecology of birds. *Journal of Comparative Physiology A: Sensory, Neural, and Behavioral Physiology*, *187*(9), 685–697. https://doi.org/10.1007/s00359-001-0240-3

Hart, N. S., & Hunt, D. M. (2007). Avian visual pigments: Characteristics, spectral tuning, and evolution. *The American Naturalist*, *169*(1), S7–S26. https://doi.org/10.1086/510141

Hart, N. S., Mountford, J. K., Davies, W. I., Collin, S. P., & Hunt, D. M. (2016). Visual pigments in a palaeognath bird, the emu Dromaius novaehollandiae: Implications for spectral sensitivity and the origin of ultraviolet vision. *Proceedings of the Royal Society B*, *283*(1834), 20161063.

Hart, N. S., Partridge, J. C., & Cuthill, I. C. (1999). Visual pigments, cone oil droplets, ocular media and predicted spectral sensitivity in the Domestic Turkey (Meleagris gallopavo). *Vision Research*, *39*(20), 3321–3328. https://doi.org/10.1016/S0042-6989(99)00071-1

Hart, N. S., & Vorobyev, M. (2005). Modelling oil droplet absorption spectra and spectral sensitivities of bird cone photoreceptors. *Journal of Comparative Physiology A*, *191*(4), 381–392. https://doi.org/10.1007/s00359-004-0595-3

Herrero, J., Muffato, M., Beal, K., Fitzgerald, S., Gordon, L., Pignatelli, M., Vilella, A. J., Searle, S. M. J., Amode, R., Brent, S., Spooner, W., Kulesha, E., Yates, A., & Flicek, P. (2016). Ensembl comparative genomics resources. *Database*, *2016*, bav096.

Hunt, D. M., Carvalho, L. S., Cowing, J. A., & Davies, W. L. (2009). Evolution and spectral tuning of visual pigments in birds and mammals. *Proceedings of the Royal Society B*, *364*(1531), 2941–2955.

Imai, H., Kojima, D., Oura, T., Tachibanaki, S., Terakita, A., & Shichida, Y. (1997). Single amino acid residue as a functional determinant of rod and cone visual pigments. *Proceedings of the National Academy of Sciences*, *94*(6), 2322–2326. https://doi.org/10.1073/pnas.94.6.2322

Kapusta, A., Suh, A., & Feschotte, C. (2017). Dynamics of genome size evolution in birds and mammals. *Proceedings of the National Academy of Sciences*, *114*(8), E1460–E1469. https://doi.org/10.1073/pnas.1616702114

Kent, W. J. (2002). BLAT—The BLAST-like alignment tool. *Genome Research*, *12*(4), 656–664. https://doi.org/10.1101/gr.229202

Kent, W. J., Sugnet, C., Furey, T., Roskin, K., Pringle, T., Zahler, A., & Haussler, A. (2002). The human genome browser at UCSC. *Genome Research*, *12*(6), 996–1006. https://doi.org/10.1101/gr.229102

Kieran, T. J., Gordon, E. R. L., Forthman, M., Hoey-Chamberlain, R., Kimball, R. T., Faircloth, B. C., Weirauch, C., & Glenn, T. C. (2019). Insight from an ultraconserved element bait set designed for

hemipteran phylogenetics integrated with genomic resources. *Molecular Phylogenetics and Evolution*, 130, 297–303. https://doi.org/10.1016/j.ympev.2018.10.026

Kimball, R. T., Oliveros, C. H., Wang, N., White, N. D., Barker, F. K., Field, D. J., Ksepka, D. T., Chesser, R. T., Moyle, R. G., Braun, M. J., Brumfield, R. T., Faircloth, B. C., Smith, B. T., & Braun, E. L. (2019). A phylogenomic supertree of birds. *Diversity*, 11(7), 109. https://doi.org/10.3390/d11070109

Kitts, P. A., Church, D. M., Thibaud-Nissen, F., Choi, J., Hem, V., Sapojnikov, V., Smith, R. G., Tatusova, T., Xiang, C., Zherikov, A., DiCuccio, M., Murphy, T. D., Pruitt, K. D., & Kimchi, A. (2016). Assembly: A resource for assembled genomes at NCBI. *Nucleic Acids Research*, 44(D1), D73–D80. https://doi.org/10.1093/nar/gkv1226

Kosakovsky Pond, S. L., Frost, S. D. W., & Muse, S. V. (2005). HyPhy: Hypothesis testing using phylogenies. *Bioinformatics*, 21(5), 676–679. https://doi.org/10.1093/bioinformatics/bti079

Lamb, T. D. (2020). Evolution of the genes mediating phototransduction in rod and cone photoreceptors. *Progress in Retinal and Eye Research*, 76, 100823. https://doi.org/10.1016/j.preteyeres.2019.100823

Lamb, T. D., & Hunt, D. M. (2017). Evolution of the vertebrate phototransduction cascade activation steps. *Developmental Biology*, 431(1), 77–92. https://doi.org/10.1016/j.ydbio.2017.03.018

Lamb, T. D., & Hunt, D. M. (2018). Evolution of the calcium feedback steps of vertebrate phototransduction. *Open Biology*, 8, 180119. https://doi.org/10.1098/rsob.180119

Lamb, T. D., Patel, H., Chuah, A., Natoli, R. C., Davies, W. I. L., Hart, N. S., Collin, S. P., & Hunt, D. M. (2016). Evolution of vertebrate phototransduction: Cascade activation. *Molecular Biology and Evolution*, 33(8), 2064–2087. https://doi.org/10.1093/molbev/msw095

Lamb, T. D., & Pugh, E. N. Jr (2006a). Phototransduction, dark adaptation, and rhodopsin regeneration the proctor lecture. *Investigative Ophthalmology & Visual Science*, 47(12), 5138–5152. https://doi.org/10.1167/iovs.06-0849

Lamb, T. D., & Pugh, E. N. Jr (2006b). Avoidance of saturation in human cones is explained by very rapid inactivation reactions and pigment bleaching. *Investigative Ophthalmology & Visual Science*, 47(13), 3714.

Leite, R. N., Kimball, R. T., Braun, E. L., Derryberry, E. P., Hosner, P. A., Derryberry, G. E., Anciães, M., McKay, J. S., Aleixo, A., Ribas, C. C., Brumfield, R. T., & Cracraft, J. (2021). Phylogenomics of manakins (Aves: Pipridae) using alternative locus filtering strategies based on informativeness. *Molecular Phylogenetics and Evolution*, 155, 107013. https://doi.org/10.1016/j.ympev.2020.107013

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

Li, H. (2013). Aligning sequence reads, clone sequences and assembly Contigs with BWA-MEM. *arXiv*, 1303.3997v1.

Lindsay, W. R., Houck, J. T., Giuliano, C. E., & Day, L. B. (2016). Acrobatic courtship display coevolves with brain size in Manakins (Pipridae). *Brain, Behavior and Evolution*, 85, 29–36. https://doi.org/10.1159/000369244

Lobanova, E. S., Herrmann, R., Finkelstein, S., Reidel, B., Skiba, N. P., Deng, W. T., Jo, R., Weiss, E. R., Hauswirth, W. W., & Arshavsky, V. Y. (2010). Mechanistic basis for the failure of cone transducin to translocate: Why cones are never blinded by light. *Journal of Neuroscience*, 30(20), 6815–6824. https://doi.org/10.1523/JNEUROSCI.0613-10.2010

Martin, G., Rojas, L. M., Ramírez, Y., & McNeil, R. (2004). The eyes of oilbirds (Steatornis caripensis): pushing at the limits of sensitivity. *Naturwissenschaften*, 91, 26–29. https://doi.org/10.1007/s00114-003-0495-3

McCormack, J. E., Tsai, W. L., & Faircloth, B. C. (2016). Sequence capture of ultraconserved elements from bird museum specimens. *Molecular Ecology Resources*, 16, 1189–1203. https://doi.org/10.1111/1755-0998.12466

McKay, B. D., Barker, F. K., Mays, H. L., Doucet, S. M., & Hill, G. E. (2010). A molecular phylogenetic hypothesis for the manakins (Aves: Pipridae). *Molecular Phylogenetics and Evolution*, 55(2), 733–737. https://doi.org/10.1016/j.ympev.2010.02.024

Miller, W., Rosenbloom, K., Hardison, R. C., Hou, M., Taylor, J., Raney, B., Burhans, R., King, D. C., Baertsch, R., Blankenberg, D., Kosakovsky Pond, S. L., Nekrutenko, A., Giardine, B., Harris, R. S., Tyekucheva, S., Diekhans, M., Pringle, T. H., Murphy, W. J., Lesk, A., ... Kent, W. J. (2007). 28-way vertebrate alignment and conservation track in the UCSC Genome Browser. *Genome Research*, 17(12), 1797–1808. https://doi.org/10.1101/gr.6761107

Murrell, B., Weaver, S., Smith, M. D., Wertheim, J. O., Murrell, S., Aylward, A., Eren, K., Pollner, T., Martin, D. P., Smith, D. M., Scheffler, K., & Kosakovsky Pond, S. L. (2015). Gene-wide identification of episodic selection. *Molecular Biology and Evolution*, 32(5), 1365–1371. https://doi.org/10.1093/molbev/msv035

Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., & Kosakovsky Pond, S. L. (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics*, 8(7), e1002764. https://doi.org/10.1371/journal.pgen.1002764

Nakane, Y., Ikegami, K., Ono, H., Yamamoto, N., Yoshida, S., Hirunagi, K., Ebihara, S., Kubo, Y., & Yoshimura, T. (2010). A mammalian neural tissue opsin (Opsin 5) is a deep brain photoreceptor in birds. *Proceedings of the National Academy of Sciences*, 107(34), 15264–15268. https://doi.org/10.1073/pnas.1006393107

Nakane, Y., Shimmura, T., Abe, H., & Yoshimura, T. (2014). Intrinsic photosensitivity of a deep brain photoreceptor. *Current Biology*, 24(13), R596–R597. https://doi.org/10.1016/j.cub.2014.05.038

NCBI Resource Coordinators (2018). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Research*, 46(D1), D8–D13. https://doi.org/10.1093/nar/gkx1095

Ödeen, A., & Håstad, O. (2003). Complex distribution of avian color vision systems revealed by sequencing the SWS1 Opsin from total DNA. *Molecular Biology and Evolution*, 20(6), 855–861. https://doi.org/10.1093/molbev/msg108

Ödeen, A., & Håstad, O. (2009). New primers for the avian SWS1 pigment opsin gene reveal new amino acid configurations in spectral sensitivity tuning sites. *Journal of Heredity*, 100(6), 784–789. https://doi.org/10.1093/jhered/esp060

Ödeen, A., & Håstad, O. (2013). The phylogenetic distribution of ultraviolet sensitivity in birds. *BMC Evolutionary Biology*, 13(1), 36. https://doi.org/10.1186/1471-2148-13-36

Ödeen, A., Pruett-Jones, S., Driskell, A. C., Armenta, J. K., & Håstad, O. (2011). Multiple shifts between violet and ultraviolet vision in a family of passerine birds with associated changes in plumage coloration. *Proceedings of the Royal Society B*, 279(1732), 1269–1276.

Ohlson, J. I., Fjeldså, J., & Ericson, P. G. P. (2013). Molecular phylogeny of the manakins (Aves: Passeriformes: Pipridae), with a new classification and the description of a new genus. *Molecular Phylogenetics and Evolution*, 69(3), 796–804. https://doi.org/10.1016/j.ympev.2013.06.024

Ohuchi, H., Yamashita, T., Tomonari, S., Fujita-Yanagibayashi, S., Sakai, K., Noji, S., & Shichida, Y. (2012). A non-mammalian type opsin 5 functions dually in the photoreceptive and non-photoreceptive organs of birds. *PLoS ONE*, 7(2), e31534. https://doi.org/10.1371/journal.pone.0031534

Peirson, S. N., Halford, S., & Foster, R. G. (2009). The evolution of irradiance detection: Melanopsin and the non-visual opsins. *Proceedings of the Royal Society B*, 364(1531), 2849–2865.

Pugh, E. N. Jr, & Lamb, T. D. (2000). Phototransduction in vertebrate rods and cones: molecular mechanisms of amplification, recovery and

light adaptation. In: D. G. Stavenga, W. J. de Grip, & E. N. Pugh Jr, eds. *Handbook of Biological Physics* (Vol. 3, pp. 183–255). Elsevier Science B.V.

Pugh, E. N. Jr, Nikonov, S., & Lamb, T. D. (1999). Molecular mechanisms of vertebrate photoreceptor light adaptation. *Current Opinions in Neurobiology*, 9(4), 410–418. https://doi.org/10.1016/S0959-4388(99)80062-2

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, 26(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

Ranwez, V., Chantret, N., & Delsuc, F. (2021). Aligning protein-coding nucleotide sequences with MACSE. *Methods in Molecular Biology*, 2231, 51–70.

Ranwez, V., Harispe, S., Delsuc, F., & Douzery, E. (2011). MACSE: Multiple alignment of coding sequences accounting for frameshifts and stop codons. *PLoS ONE*, 6(9), e22594. https://doi.org/10.1371/journal.pone.0022594

Reddy, S., Kimball, R. T., Pandey, A., Hosner, P. A., Braun, M. J., Hackett, S. J., Han, K.-L., Harshman, J., Huddleston, C. J., Kingston, S., Marks, B. D., Miglia, K. J., Moore, W. S., Sheldon, F. H., Witt, C. C., Yuri, T., & Braun, E. L. (2017). Why do phylogenomic data sets yield conflicting trees? Data type influences the avian tree of life more than taxon sampling. *Systematic Biology*, 66(5), 857–879. https://doi.org/10.1093/sysbio/syx041

Rhie, A., McCarthy, S. A., Fedrigo, O., Damas, J., Formenti, G., Koren, S., Uliano-Silva, M., Chow, W., Fungtammasan, A., Kim, J., Lee, C., Ko, B. J., Chaisson, M., Gedman, G. L., Cantin, L. J., Thibaud-Nissen, F., Haggerty, L., Bista, I., & Smith, M., ... Jarvis, E. D. (2021). Towards complete and error-free genome assemblies of all vertebrate species. *Nature*, 592(7856), 737–746.

Rosel, P. E., & Block, B. A. (1996). Mitochondrial control region variability and global population structure in the swordfish, Xiphias gladius. *Marine Biology*, 125(1), 11–22. https://doi.org/10.1007/BF00350756

Sboner, A., Mu, X. J., Greenbaum, D., Auerbach, R. K., & Gerstein, M. B. (2011). The real cost of sequencing: Higher than you think! *Genome Biology*, 12(8), 125. https://doi.org/10.1186/gb-2011-12-8-125

Schott, R. K., Panesar, B., Card, D. C., Preston, M., Castoe, T. A., & Chang, B. S. W. (2017). Targeted capture of complete coding regions across divergent species. *Genome Biology and Evolution*, 9(2), 398–414. https://doi.org/10.1093/gbe/evx005

Schwarze, K., Buchanan, J., Fermont, J. M., Dreau, H., Tilley, M. W., Taylor, J. M., Antoniou, P., Knight, S. J. L., Camps, C., Pentony, M. M., Kvikstad, E. M., Harris, S., Popitsch, N., Pagnamenta, A. T., Schuh, A., Taylor, J. C., & Wordsworth, S. (2020). The complete costs of genome sequencing: A microcosting study in cancer and rare diseases from a single center in the United Kingdom. *Genetics in Medicine*, 22(1), 85–94. https://doi.org/10.1038/s41436-019-0618-7

Silva, S. M., Agne, C. E., Aleixo, A., & Bonatto, S. L. (2018). Phylogeny and systematics of Chiroxiphia and Antilophia manakins (Aves, Pipridae). *Molecular Phylogenetics and Evolution*, 127, 706–711. https://doi.org/10.1016/j.ympev.2018.06.016

Smit, A., Hubley, R., & Green, P. (2013–2015). *RepeatMasker Open-4.0.* Retrieved from http://www.repeatmasker.org

Smith, M. D., Wertheim, J. O., Weaver, S., Murrell, B., Scheffler, K., & Kosakovsky Pond, S. L. (2015). Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection. *Molecular Biology and Evolution*, 32(5), 1342–1353. https://doi.org/10.1093/molbev/msv022

Sugihara, T., Nagata, T., Mason, B., Koyanagi, M., & Terakita, A. (2016). Absorption characteristics of vertebrate non-visual Opsin, Opn3. *PLoS ONE*, 11(8), e0161215. https://doi.org/10.1371/journal.pone.0161215

Tamashiro, R. A., White, N. D., Braun, M. J., Faircloth, B. C., Braun, E. L., & Kimball, R. T. (2019). What are the roles of taxon sampling and model fit in tests of cyto-nuclear discordance using avian mitogenomic

data? *Molecular Phylogenetics and Evolution*, 130, 132–142. https://doi.org/10.1016/j.ympev.2018.10.008

Tomonari, S., Migita, K., Takagi, A., Noji, S., & Ohuchi, H. (2008). Expression patterns of the opsin 5-related genes in the developing chicken retina. *Developmental Dynamics*, 237(7), 1910–1922.

Toomey, M. B., Lind, O., Frederiksen, R., & Curley, R. W. (2016). Complementary shifts in photoreceptor spectral tuning unlock the full adaptive potential of ultraviolet vision in birds, *eLife*, 5, e15675.

Umino, Y., Guo, Y., Chen, C. K., Pasquale, R., & Solessio, E. (2019). Rod photoresponse kinetics limit temporal contrast sensitivity in mesopic vision. *Journal of Neuroscience*, 39(16), 3041–3056. https://doi.org/10.1523/JNEUROSCI.1404-18.2019

Umino, Y., Herrmann, R., Chen, C. K., Barlow, R. B., Arshavsky, V. Y., & Solessio, E. (2012). The relationship between slow photoresponse recovery rate and temporal resolution of vision. *Journal of Neuroscience*, 32(41), 14364–14373. https://doi.org/10.1523/JNEUROSCI.1296-12.2012

van Hazel, I., Dungan, S. Z., Hauser, F. E., Morrow, J. M., Endler, J. A., & Chang, B. S. W. (2016). A comparative study of rhodopsin function in the great bowerbird (Ptilonorhynchus nuchalis): Spectral tuning and light-activated kinetics. *Protein Science*, 25(7), 1308–1318.

White, N. D. (2021a). Decoy sequences (Version v1) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.4734556

White, N. D. (2021b). *PTC-guided-assembly v1.1.* Available from https://doi.org/10.5281/zenodo.4734544

White, N. D., & Braun, M. J. (2019). Extracting phylogenetic signal from phylogenomic data: Higher-level relationships of the nightbirds (Strisores). *Molecular Phylogenetics and Evolution*, 141, 106611. https://doi.org/10.1016/j.ympev.2019.106611

Wilkie, S. E., Robinson, P. R., Cronin, T. W., Poopalasundaram, S., Bowmaker, J. K., & Hunt, D. M. (2000). Spectral tuning of avian violet- and ultraviolet-sensitive visual pigments. *Biochemistry*, 39(27), 7895–7901. https://doi.org/10.1021/bi992776m

Wu, Y., Hadly, E. A., Teng, W., Hao, Y., Liang, W., Liu, Y., & Wang, H. (2016). Retinal transcriptome sequencing sheds light on the adaptation to nocturnal and diurnal lifestyles in raptors. *Scientific Reports*, 6, 33578. https://doi.org/10.1038/srep33578

Wu, Y., Wang, H., & Hadly, E. A. (2017). Invasion of ancestral mammals into dim-light environments inferred from adaptive evolution of the phototransduction genes. *Scientific Reports*, 7, 46542. https://doi.org/10.1038/srep46542

Yamashita, T. (2020). Unexpected molecular diversity of vertebrate non-visual opsin Opn5. *Biophysical Reviews*, 12, 333–338. https://doi.org/10.1007/s12551-020-00654-z

Yates, A., Akanni, W., Amode, M. R., Barrell, D., Billis, K., Carvalho-Silva, D., Cummins, C., Clapham, P., Fitzgerald, S., Gil, L., Girón, C. G., Gordon, L., Hourlier, T., Hunt, S. E., Janacek, S. H., Johnson, N., Juettemann, T., Keenan, S., Lavidas, I., ... Flicek, P. (2016). Ensembl 2016. *Nucleic Acids Research*, 44(D1), D710–D716. https://doi.org/10.1093/nar/gkv1157

Yokoyama, S. (2000). Molecular evolution of vertebrate visual pigments. *Progress in Retinal and Eye Research*, 19(4), 385–419. https://doi.org/10.1016/S1350-9462(00)00002-1

Yokoyama, S., Radlwimmer, F. B., & Blow, N. S. (2000). Ultraviolet pigments in birds evolved from violet pigments by a single amino acid change. *Proceedings of the National Acadmey of Sciences*, 97(13), 7366–7371. https://doi.org/10.1073/pnas.97.13.7366

Yokoyama, S., Yang, H., & Starmer, W. T. (2008). Molecular basis of spectral tuning in the red- and green-sensitive (M/LWS) pigments in vertebrates. *Genetics*, 179(4), 2037–2043. https://doi.org/10.1534/genetics.108.090449

Zele, A. J., & Cao, D. (2015). Vision under mesopic and scotopic illumination. *Frontiers in Psychology*, 5, 1594. https://doi.org/10.3389/fpsyg.2014.01594

Zhang, G., Li, C., Li, Q., Li, B., Larkin, D. M., Lee, C., Storz, J. F., Antunes, A., Greenwold, M. J., Meredith, R. W., Odeen, A., Cui, J., Zhou, Q., Xu, L., Pan, H., Wang, Z., Jin, L., Zhang, P., Hu, H., ... Avian Genome Consortium (2014). Comparative genomics reveals insights into avian genome evolution and adaptation. *Science*, *346*(6215), 1311–1320.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.