MBE OXFORD

Widespread Genetic Signals of Visual System Adaptation in Deepwater Cichlid Fishes

Julia I. Camacho García [iD],[1,]* Milan Malinsky [iD],[2] Domino A. Joyce [iD],[3] M. Emília Santos [iD],[4] Grégoire Vernaz [iD],[5,6,9] Maxon J. Ngochera [iD],[7] Hannes Svardal [iD][1,8,]*

[1]Evolutionary Ecology Group, Department of Biology, University of Antwerp, Antwerp, Belgium
[2]Institute of Ecology and Evolution, Department of Biology, University of Bern, Bern, Switzerland
[3]Evolutionary and Ecological Genomics Group, University of Hull, Hull, UK
[4]Department of Zoology, University of Cambridge, Cambridge, UK
[5]Wellcome/Cancer Research UK Gurdon Institute, University of Cambridge, Cambridge, UK
[6]Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, UK
[7]Department of Fisheries Headquarters, P.O. Box 593, Lilongwe, Malawi
[8]Naturalis Biodiversity Center, Leiden, The Netherlands
[9]Present address: Zoological Institute, University of Basel, Basel, Switzerland
Corresponding authors: E-mails: julia.camachogarcia@uantwerpen.be; hannes.svardal@uantwerpen.be.
Associate editor: Belinda Chang

Abstract

The light environment exerts a profound selection pressure on the visual system, driving morphological and molecular adaptations that may also contribute to species diversification. Here, we investigate the evolution and genetic basis of visual system diversification in deepwater cichlid fishes of the genus *Diplotaxodon*. We find that *Diplotaxodon* exhibit the greatest eye size variation among Lake Malawi cichlids and that this variation is largely uncoupled from phylogeny, with various nonsister species sharing similar eye sizes. Using a combination of genome-wide association analysis across nine *Diplotaxodon* species, haplotype-based selection scans, and transcriptome analysis, we uncover consistent and widespread signatures of evolution in visual pathways, centered on green-sensitive opsins and throughout the phototransduction cascade, suggesting coordinated evolution of eye size and visual molecular pathways. Our findings underscore the role of visual system diversification in niche specialization within deepwater habitats and offer new insights into visual system evolution within this extraordinary cichlid radiation.

Keywords: cichlid, visual system, ecological adaptation, evolutionary genomics

Introduction

Light is an important driver of evolutionary diversification of the visual sensory system, shaping species diversity through morphological and physiological changes (Cronin et al. 2014). In aquatic environments, where light attenuates exponentially with depth, this relationship can be specially pronounced, as organisms must compensate for extreme reductions in light intensity and shifts in spectral composition (Warrant and Locket 2004; Warrant and Johnsen 2013). One of the most conspicuous modifications to the visual system is an increase in eye size—a trait observed in many deepwater fishes and cephalopods, as well as in nocturnal vertebrates and insects (Warrant and Locket 2004; Land and Nilsson 2012). Larger eyes can capture more light by accommodating larger pupil apertures and lenses, thereby enhancing visual sensitivity in dimly lit environments (Land and Nilsson 2012).

Beyond eye size, visual adaptations occur across multiple levels of biological organization, including changes to lens properties, visual pigments (composed of a light-absorbing chromophore and an opsin protein; Bowmaker 2008), photoreceptor composition (rods and cones), and the phototransduction cascade (Carleton et al. 2020). These features often form suites of adaptations that broadly align with an organism's visual ecology. For example, in deepwater and nocturnal

fishes, morphological features like enlarged eyes are often accompanied by other structural and molecular specializations, such as rod-only or rod-dominated retinas with opsins tuned to blue-green wavelengths, or elongated photoreceptors (Carleton et al. 2020; de Busserolles et al. 2020). Some of these mechanisms have been extensively studied in teleost fishes, providing valuable insights into how visual systems diversify in response to ecological pressures (Carleton et al. 2020; Musilova et al. 2021). Among these, cichlid fishes (family Cichlidae) have been particularly powerful as a model system due to their extraordinary visual diversity (Carleton 2009; Carleton et al. 2016; Carleton and Yourick 2020).

Cichlids, a species-rich group of freshwater fishes, strongly rely on vision for navigation, feeding, and mate choice—key factors in their rapid diversification (McGee et al. 2020; Santos et al. 2023). Differences in ecology and light environment have been linked to extensive visual pigment variation in cichlids (reviewed in Carleton 2009; Carleton and Yourick 2020), with depth-related visual adaptation playing crucial roles in their adaptive radiations in lakes Tanganyika (Ronco et al. 2021; Ricci et al. 2022, 2023), Victoria (Seehausen et al. 2008), Malawi (Hahn et al. 2017; Malinsky et al. 2018; Blumer et al. 2025), Masoko (Malinsky et al. 2015), and Barombi Mbo (Musilova,

Indermaur et al. 2019) and contributing to ecological divergence and reproductive isolation among incipient species (Terai et al. 2006; Seehausen et al. 2008; Malinsky et al. 2015; Takuno et al. 2019). However, visual diversification beyond opsin tuning remains understudied, and the extent to which other components of cichlid visual systems have diversified remains largely unexplored. Here, we address this question by examining broad-scale visual adaptations in deepwater cichlids from Lake Malawi.

Lake Malawi, home to a cichlid radiation of over 800 extant species (Konings 2016), has deepwater habitats that extend from approximately 50 m to the anoxic layer at about 230 m (Eccles 1974). Of the seven lineages that make up this extraordinary radiation—usually referred to as "eco-morphological" clades (Malinsky et al. 2018)—two are adapted to the deepwater environment: a diverse group of ~150 benthic species (the "deep benthic" clade), and the genus *Diplotaxodon* (Trewavas 1935), comprising ~27 estimated species (nine described) of open-water planktivores and piscivores (Turner et al. 2004; Konings 2016; Stauffer et al. 2018). Previously, a radiation-wide scan for adaptive protein evolution (Malinsky et al. 2018), identified key visual genes (rhodopsin, green-sensitive opsins, and several other phototransduction genes) as outliers, with many of them exhibiting signals of parallel adaptation between the deep benthic clade and *Diplotaxodon*. Within *Diplotaxodon*, catch-depth records and marked variation in eye size suggest that diversifying selection may have contributed to differential visual adaptations. Notably, species restricted to greater depths (deeper than 50–80 m) tend to exhibit larger eyes than those occupying broader depth ranges that include shallower waters (20 to 30 m) (Turner et al. 2004). This raises the possibility of habitat partitioning and warrants the hypothesis that their visual systems have undergone adaptive evolution in response to ecological specialization. This idea is further supported by a restriction-site associated DNA sequencing (RADseq) study (Hahn et al. 2017) that identified candidate regions under divergent selection across four *Diplotaxodon* species, including genes related to visual perception (e.g. rhodopsin).

In this study, we investigate the evolution and genetic underpinnings of visual diversification in *Diplotaxodon* by integrating morphological and genomic data. We focus on eye size as a key trait of interest, given its high variability across cichlids and within this genus, and its general association with light availability and, more roughly, depth. While genetic variation linked to eye size may directly contribute to differences in this trait, it may also more broadly reflect adaptation to (visual) ecological niches beyond eye size itself. To contextualize eye size variation in *Diplotaxodon*, we examine morphological data from 1,334 individuals spanning 251 species across the Lake Malawi cichlid radiation. Performing genome-wide association (GWA) analysis for eye size, we identify many vision-related genes covering a broad functional spectrum. Enrichments for nonsynonymous substitutions, signals of natural selection, and cross-species differential expression of these genes support our hypothesis that visual adaptation contributed to diversification in *Diplotaxodon*. Furthermore, transcriptome analysis of eyes revealed expression differences in oxygen transport-related genes across species exhibiting differences in eye size, supporting depth-related adaptations beyond visual tuning. Overall, our study provides new insights into the genetic and phenotypic mechanisms underlying visual specializations in one of the most extreme environments into which cichlids have diversified.

Results

Large-Eye Size Variation Within and Across *Diplotaxodon* Species

To assess eye size variation within the Lake Malawi cichlid adaptive radiation, we measured eye diameter across 1,334 individuals of six eco-morphological clades (supplementary Data S1 and S2, Supplementary Material online). As expected, eye size scaled positively with body standard length (Fig. 1a), although this relationship was not equally strong across eco-morphological groups (supplementary fig. S1a, Supplementary Material online). Therefore, we used linear regression to adjust for body size within each clade, revealing that the deepwater pelagic clade of *Diplotaxodon* stood out as having the largest within-group variation in eye size (supplementary fig. S1b and c, Supplementary Material online; mean residual *Diplotaxodon* = 0.17, other groups = from 0.05 to 0.1).

Parallel Evolution of Eye Size in *Diplotaxodon*

Given uncertainties regarding species assignments and relationships within *Diplotoxodon* (Turner et al. 2004; Malinsky et al. 2018), we used whole-genome data of 79 *Diplotaxodon* individuals from 13 species to provide evolutionary context for investigating differences in eye size. Constructing a neighbor joining (NJ) tree based on pairwise genetic distances (Fig. 1b; supplementary fig. S2, Supplementary Material online), we found that all individuals assigned to the same species based on morphological examination formed monophyletic groups, except for our two sequenced *D.* "holochromis", which clustered separately from each other and will be treated as distinct taxonomic entities (*D.* "holochromis 1" and *D.* "holochromis 2") in the following.

Across all *Diplotaxodon* individuals, eye diameter varied between 7% and 13% of standard length (Fig. 1b). Interestingly, we found that this variation in relative eye size was largely decoupled from the phylogeny. Specifically, Pagel's lambda (λ), a measure of phylogenetic signal with respect to a Brownian model of trait evolution (Pagel 1999), was close to zero (Pagel's $\lambda = 6.6 \times 10^{-5}$), indicating a large degree of independence of eye size from the species evolutionary relationships. Given the low sample size for some species and a lack of a fully resolved phylogeny, we interpret this value with caution. However, it is clear that various nonsister lineages share similar eye sizes across the phylogeny, for example, *D. macrops*, *D.* "bigeye black dorsal", and *D. apogon* have a "large-eye" phenotype and *D. limnothrissa*, and *D.* "holochromis 1/2" have a "small-eye" phenotype (Fig. 1b; note that despite their small relative eye size, we will generally not consider the predatory *D. longimaxilla* and *D. greenwoodi* in the latter category as their much larger size and distinctive general body shape set them apart from the rest). Moreover, the correlation between genetic distance and eye size differences is relatively low ($R^2 = 0.15$, $P = 4.7 \times 10^{-42}$), much lower than, for example, the correlation between genetic distance and body shape differences ($R^2 = 0.55$, $P = 5.8 \times 10^{-201}$) (supplementary fig. S3, Supplementary Material online). Overall, the decoupling of this trait from phylogeny and genetic distance suggests that eye size has evolved in parallel in distinct lineages and creates an opportunity for investigating which genetic variants are linked to variation in the trait.
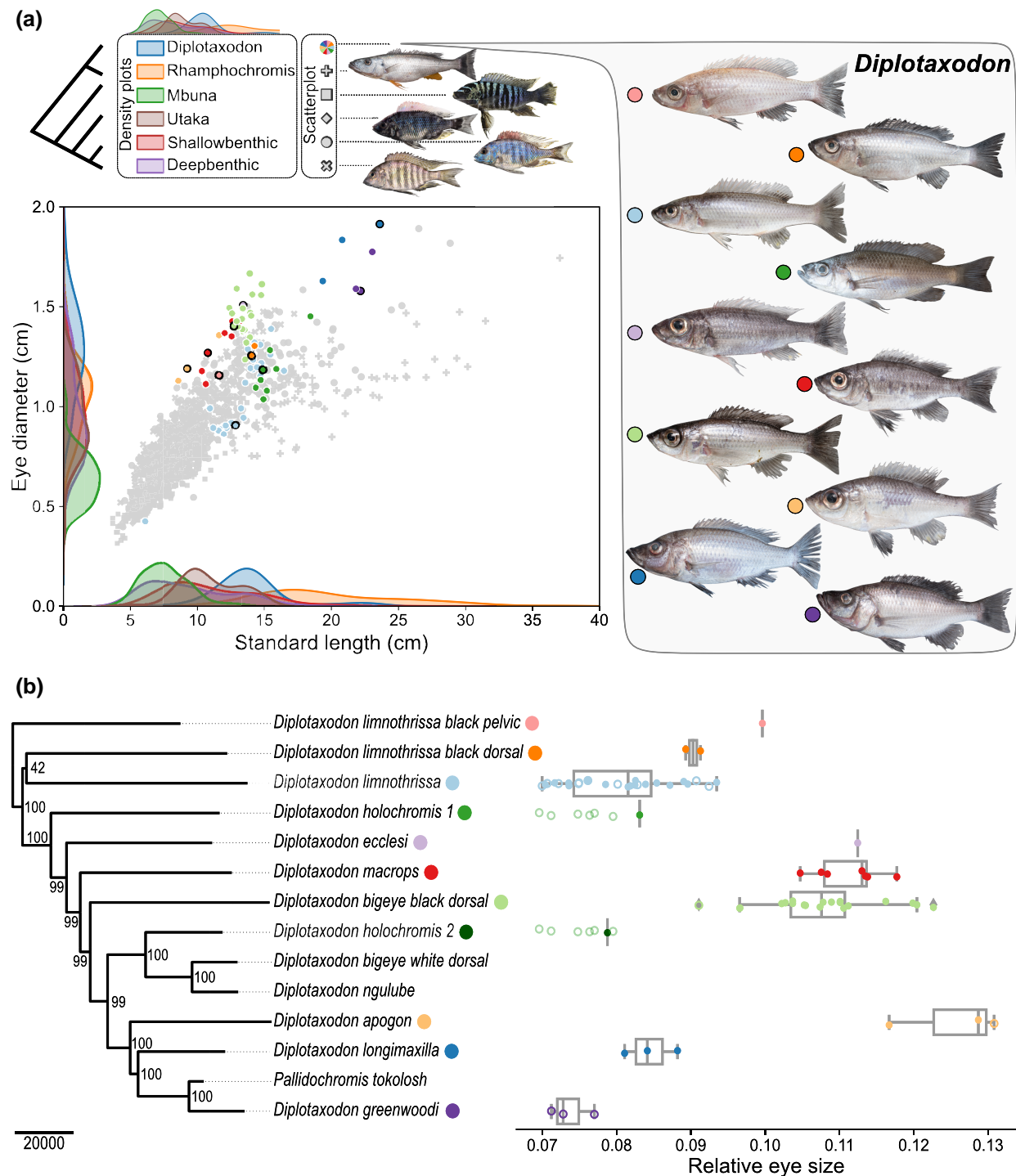
Fig. 1. Eye size variation and genetic relationships in *Diplotaxodon*. a) Eye diameter versus standard length for 10 *Diplotaxodon* species ($N = 75$, colored circles) and members of the other five major eco-morphological clades of the Malawi cichlid radiation (+: Rhamphochromis, $N = 62$; □: Mbuna, $N = 563$; ○: Shallow benthic, $N = 401$; ✕: Deep benthic, $N = 170$; ◇: Utaka, $N = 63$; grey markers). Kernel density estimate plots along the axes show the distributions of eye diameter and standard length per clade. Above, representative specimens from each clade (not to scale) and a diagram illustrating their phylogenetic relationships are displayed. The representative species are: *Rhamphochromis ferox* (Rhamphochromis), *Maylandia zebra* (Mbuna), *Otopharynx tetrastigma* (Shallow benthic), *Placidochromis hennydaviesae* (Deep benthic), *Copadichromis quadrimaculatus* (Utaka). For *Diplotaxodon*, one representative specimen per species is shown, with the corresponding data points outlined in black on the scatterplot. b) Neighbor joining (NJ) phylogeny based on whole-genome sequences of 13 *Diplotaxodon* species, including the monotypic genus *Pallidochromis* nested within the *Diplotaxodon* clade. Node labels show bootstrap support based on resampling distance matrices in 100 kb windows along the genome. On the right, relative eye sizes (eye diameter divided by standard length) of the individuals in (a); filled circles represent sequenced samples included in the genome-wide association analysis for eye size ($N = 53$), while empty circles represent phenotyped specimens without sequences, assigned to a species based on morphological inspection. The boxes indicate the three quartile values of the distribution, and the whiskers are the minimum and maximum values. Values outside 1.5× the interquartile range are represented as diamonds. Given that the two sequenced *D.* "holochromis" individuals are not monophyletic in the NJ tree, it is not possible to confidently assign nonsequenced samples morphologically identified as *D.* "holochromis" to either group; however, to show eye size variation, these samples are included along both species ("holochromis 1/2") (duplicated data points = green empty circles with increased transparency).
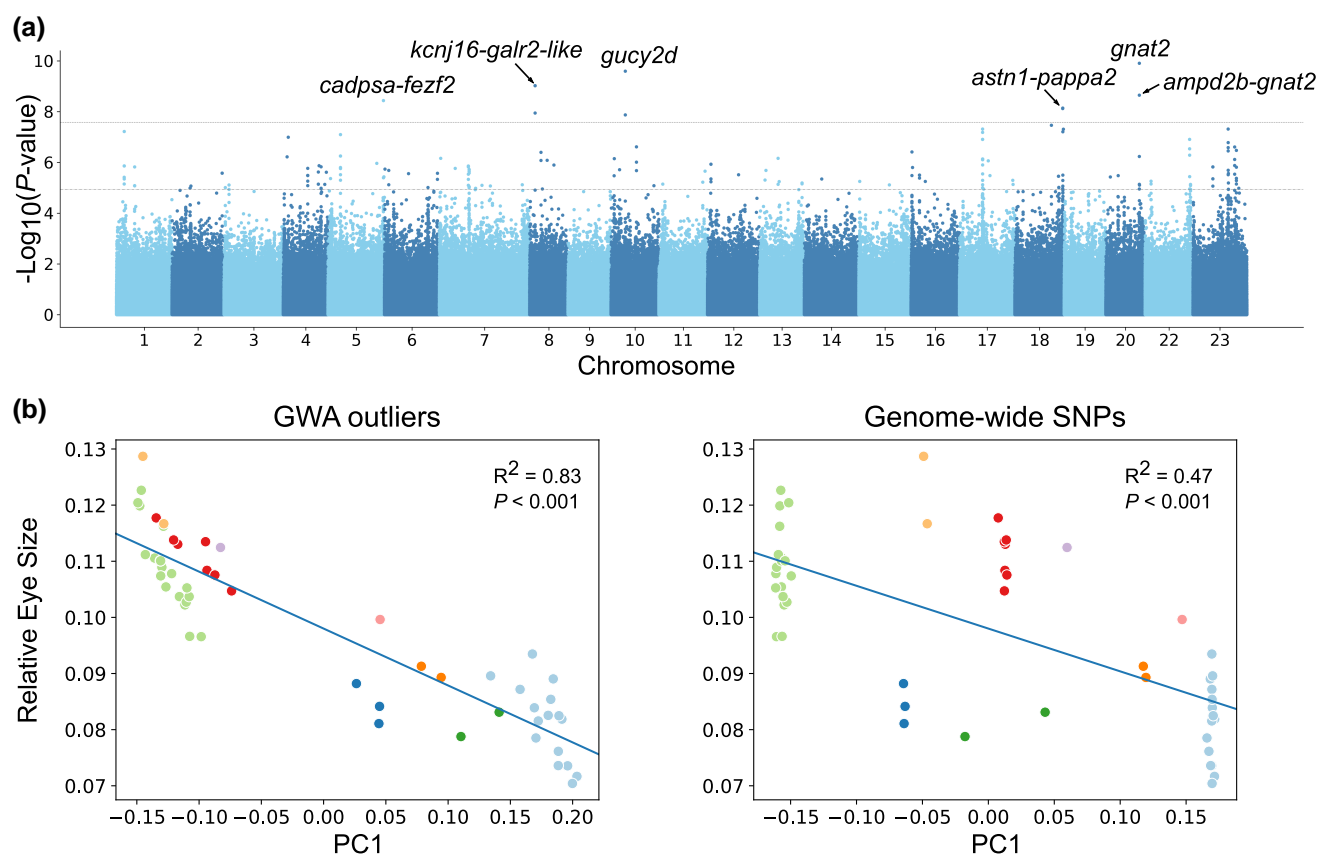
Fig. 2. Genome-wide association analysis for eye size in *Diplotaxodon*. a) Manhattan plot showing genome-wide associations for relative eye size across nine *Diplotaxodon* species (N = 53). Top horizontal line: Bonferroni-corrected threshold at *FWER* < 0.05; bottom horizontal line (dashed): cutoff 0.01% most significant associations. Genes annotated to Bonferroni-significant SNPs are shown as a single name (if exonic) or two names separated by a dash (intergenic); b) Principal component (PC) 1 against relative eye size for PC analyses summarizing genetic variation on GWA outliers (190 SNPs; left) and genome-wide (MAF > 5%, 377,273 SNPs; right). In both plots, a regression line and R-squared of relative eye size ~ PC1 are shown. Species color keys as in Fig. 1.

Genome-Wide Significant Eye Size Associations

To investigate genetic variation associated with the observed differences in eye size, we performed a GWA analysis for relative eye size, focusing on the *Diplotaxodon* individuals with both whole-genome sequences and phenotypic data (filled circles in Fig. 1b, N = 53). Despite the moderate sample size, we identified 10 single nucleotide polymorphisms (SNPs) associated with eye size with genome-wide significance (Bonferroni FWER < 0.05) on five chromosomes (Fig. 2a).

The two strongest associations, on chromosomes 10 and 20, corresponded to nonsynonymous mutations in the genes *gucy2d* ($P = 2.57 \times 10^{-10}$) and *gnat2* ($P = 1.26 \times 10^{-10}$) (Fig. 2a), both coding for proteins of the cone photoreceptors in the retina, where they play essential roles in phototransduction—the process by which photons of light are converted into electrical signals (Lamb and Curtin 2022). Among the remaining eight Bonferroni-significant SNPs, there was an additional nonsynonymous mutation mapping to *gucy2d*. The rest were located in intergenic regions nearby the genes *kcnj16*, LOC113027815 (*galr2-like*), *ampd2b*, *gnat2*, *cadpsa*, *fezf2*, *pappa2*, and *astn1*. These genes participate in a range of functions, including neurotransmission and neurogenesis (*galr2*, *fezf2*, *cadpsa*, *astn1*), energy metabolism (*ampd2b*), and homeostasis (*kcnj16*) (Stelzer et al. 2016; The UniProt Consortium 2023). Beyond the conservative Bonferroni threshold, there were several other association peaks that might correspond to biologically relevant signals (Fig. 2a).

Therefore, to include these in further analyses we considered as outlier SNPs the 0.01% most significant associations, corresponding to 190 SNPs with an eye size association of $P < 0.00001$ ("GWA outliers").

The linear mixed model used to estimate GWA accounts for population stratification (Price et al. 2010; Zhou and Stephens 2012). To verify that the GWA signals were not driven by residual stratification in our highly structured sample set, we used principal component (PC) analysis to compare genome-wide sample genetic variation to that among GWA outlier SNPs. We found that GWA outliers separated samples along PC1 by eye size to a larger degree than would be expected by phylogeny alone (Fig. 2b). Specifically, *D.* "holochromis 1" and *D.* "holochromis 2" were closer to samples of *D. limnothrissa* spp. (collectively referred to as *D. limnothrissa* complex) than to their sister species, which formed a separate cluster. A linear regression confirmed that PC1 from GWA outlier SNPs explain variation in eye size to a much greater extent ($R^2 = 0.83$, $P < 0.0001$) than PC1 from genome-wide variation ($R^2 = 0.47$, $P < 0.0001$) (Fig. 2b).

Eye Size-Associated Loci Harbor Vision-Related Genes

To gain insight into the biological functions of the genetic variants associated with eye size, we examined whether genes mapping to 50 kb regions centered on each of the GWA

outliers were enriched for specific gene ontology (GO) terms. We found 104 genes annotated to these regions, and significant enrichment of 30 GO terms in the "biological process" category (Fisher's exact test, $P < 0.05$; supplementary table S3, Supplementary Material online), with the terms "cone photoresponse recovery" (GO:0036368, $P = 0.0006$), "detection of visible light" (GO:0009584, $P = 0.003$) and "visual perception" (GO:0007601, $P = 0.004$) as the most highly significant. Among the genes in these categories was *gnat2*, which corresponds to the strongest GWA (Fig. 2a), together with *arr3a*, *cnga3a*, *grk7a*, and *rcvrn3* (supplementary table S3, Supplementary Material online). Like *gnat2*, these genes are integral to the phototransduction cascade, contributing to the activation, termination, and modulation of the visual response (Pugh and Lamb 2000; Lamb and Curtin 2022).

Given the high representation of visual transduction genes within GWA outlier loci, we followed up on this result with a manual examination of the functional annotation of GWA hits (see Methods), revealing three additional phototransduction genes: *gucy2d* (reported above as Bonferroni-significant, Fig. 2a), *gngt2*, and *pde6h*-like. Lowering the GWA significance cutoff from 0.01% to 0.025% (475 SNPs, $P < 5 \times 10^{-5}$) further revealed the presence of the cone arrestin *arr3b*, rhodopsin (*rh1*), and the green-sensitive opsin *rh2aβ* within GWA loci. However, as these genes do not meet the established 0.01% cutoff, they are noted here but not considered in downstream analyses. Besides phototransduction, we found other examples of genes implicated in visual functions annotated to GWA loci, namely eye development (*pawr*; Acharya et al. 2011), retina regeneration (*ascl1a*; Džulová et al. 2022), and photoreceptor synaptic structure (*pappaa*; Miller et al. 2018) (see supplementary Data S5, Supplementary Material online for the full annotation of GWA outlier SNPs).

Excess of Nonsynonymous Variants Among GWA Outliers

Focusing on the 190 (0.01% most significant) GWA outlier SNPs associated with eye size, we observed that 34 of these were amino acid changing, which was 9.6 times higher than expected by chance (empirical permutation $P < 0.001$; supplementary fig. S4, Supplementary Material online). Of these nonsynonymous SNPs, 12 mapped to 5 of the phototransduction genes reported above, while the remaining 22 were located in other genes, some of which are potentially relevant to depth adaptation and muscle function (Table 1). The presence of genes with nonvisual functions among the candidates may reflect broader species-level divergence related to their habitat depths and/or ecological niches. This suggests that we might have identified causal SNPs with an adaptive function.

Following the discovery of an excess of nonsynonymous variation, we next investigated whether evolutionary change happened in the big-eyed or small-eyed species by inspecting the derived allele frequencies (AF) at these nonsynonymous sites. Overall, across the 34 nonsynonymous GWA outliers, there was no consistent pattern of whether the derived allele was present at high frequency in big or small-eyed species

Table 1 Genes annotated to GWA outliers putatively relevant for deepwater adaptation and skeletal muscle function

Gene symbol	Description	PT	Role(s)
gnat2	G protein subunit alpha transducin 2	Yes	Involved in the activation of phototransduction in cone photoreceptors. Stimulates the coupling of the opsin and cGMP-phosphodiesterase during visual impulses[a]. Human orthologs associated with achromatopsia and cone dystrophy[b].
gngt2	G protein subunit gamma transducin 2	Yes	Involved in the activation of phototransduction in cone photoreceptors. Required for GTPase activity, replacement of GDP by GTP, and G protein-effector interaction[c].
LOC113021214 (pde6h-like)	Retinal cone rhodopsin-sensitive cGMP 3′,5'-cyclic phosphodiesterase subunit gamma-like	Yes	Involved in the activation of phototransduction in cone photoreceptors. Mediates the transmission and amplification of the visual signal by catalyzing the conversion of cGMP to GMP, causing the nucleotide-gated channels to close[a]. Human orthologs associated with retinal cone dystrophy and achromatopsia[b].
cnga3a	Cyclic nucleotide-gated channel subunit alpha 3a	Yes	Involved in the activation of phototransduction in cone photoreceptors. Regulates the flow of cations into the cone, modifying its electrical charge and generating an electrical signal (interpreted as vision in the brain)[a].
gucy2d	Retinal guanylyl cyclase 2-like	Yes	Regulates phototransduction by mediating the synthesis of cGMP from GTP in rods and cones[a]. Human orthologs associated with Leber congenital amaurosis and cone-rod dystrophy[b].
grk7a	G protein-coupled receptor kinase 7a	Yes	Involved in the termination of phototransduction in cone photoreceptors by deactivating cone opsins via phosphorylation[a].
arr3a	Arrestin 3a, retinal (X-arrestin)	Yes	Involved in the termination of phototransduction in cone photoreceptors by binding to the phosphorylated opsin[a].
rcvrn3	Recoverin-like	Yes	Modulates light sensitivity of the photoreceptors in dark and dim conditions. In low light levels, inhibits GRK activity, prolonging opsin activation[a]. Facilitates the detection of change and motion in bright light[b].
si:ch1073-13h15.3	Putative all-trans-retinol 13,14-reductase (retsat)	No	May play a role in the metabolism of vitamin A (the visual pigment chromophore) in zebrafish[b].

(continued)

Table 1 Continued

Gene symbol	Description	PT	Role(s)
			Involved in apocarotenoid metabolism in birds, mediating SWS2 spectral tuning[d]. Changes in vitamin A signaling during development mediated by *retsat* could be implicated in changes in eye size and abnormal eye morphology in Atlantic haddock[e]. Implicated in hypoxia adaptation in Qinghai-Tibet Plateau mammals[f].
trhde.2	Thyrotropin releasing hormone degrading enzyme tandem duplicate 2	No	Part of the thyroid hormone (TH) pathway, where it cleaves and inactivates the neuropeptide thyrotropin-releasing hormone. TH plays a role in photoreceptor development[g] and visual plasticity by modulating opsin expression and chromophore exchange[h,i]. The gene trhde is involved in high-altitude adaptation in Ethiopian sheep[j].
LOC113022044 (*igfn1*-like)	Immunoglobulin-like and fibronectin type III domain containing protein 1	No	Predicted to be involved in homophilic cell adhesion, retina layer formation and synapse assembly[b]. Human ortholog associated with polypoidal choroidal vasculopathy, a disease affecting the layer of blood vessels in the choroid[k].
cahz	Carbonic anhydrase-like	No	Involved in carbon dioxide transport and hypotonic salinity response. Contributes to acid–base homeostasis, fluid balance, and the formation of aqueous humor in the eye, influencing its water content. May enhance tissue oxygen delivery in teleosts[l]. Suggested to play a role in oxygenation of the inner retina, potentially contributing to the adaptive evolution of vision in teleost fish[m].
pappaa	Pappalysin 1 Pregnancy-associated plasma protein aa	No	Encodes a metalloendopeptidase known to activate insulin-like growth factor 1 (IGF1)[b]. Regulates photoreceptor synaptic structure and function of cone synapses to transmit light-offset information. Pappaa mutant zebrafish have impaired responses to light offset[n].
ttn.2	Titin, tandem duplicate 2	No	Structural component of muscle involved in heart contraction, myofibril assembly and striated muscle tissue development[c].
neb	Nebulin	No	Involved in muscle fiber development. May play a role in maintaining the structural integrity of sarcomeres and the membrane system associated with the myofibrils[c].
pvalb*	Parvalbumin	No	Parvalbumins are calcium ion-binding proteins involved in muscle relaxation after contraction. Expressed in other tissues, where its function is not well understood[o,p].

It is noted if the genes are involved in the phototransduction molecular pathway ("PT") or not. Genes with nonsynonymous GWA outliers are highlighted in bold.

*pvalb is the generic name used to refer to genes encoding parvalbumin, of which multiple copies are present in teleosts (Mukherjee et al. 2021). Several copies are present in the annotation of GWA outliers, here collectively referred to as *pvalb* for simplicity.

References: [a]Pugh and Lamb (2000); [b]Stelzer et al. (2016); [c]The UniProt Consortium (2023); [d]Toomey et al. (2016); [e]Lie et al. (2019); [f]Xu et al. (2021); [g]Ng et al. (2001); [h]Karagic et al. (2018); [i]Volkov et al. (2020); [j]Edea et al. (2019); [k]Wen et al. (2018); [l]Rummer and Brauner (2011); [m]Damsgaard et al. (2020); [n]Miller et al. (2018); [o]Mukherjee et al. (2021); [p]Permyakov and Uversky (2022).

(chi-square $P = 0.8$). However, when considering only phototransduction genes, we observed a striking difference: in 11 out of the 12 nonsynonymous GWA outliers mapping to phototransduction genes the derived allele was fixed or nearly fixed in the big-eyed species, but absent or at low frequency in small-eyed species (chi-square $P < 0.001$; supplementary table S4, Supplementary Material online). The opposite pattern was found for genes with putative skeletal muscle functions for which seven out of eight genes showed high-frequency derived alleles in small-eyed species (chi-square $P = 0.01$).

Overall, our findings confirm that eye size-linked genetic variants are enriched for functional alleles and that qualitatively different molecular functions have been under adaptive evolution in big-eyed compared to small-eyed *Diplotaxodon*.

Recent Divergent Selection Around Candidate Genes for Depth Adaptation

To further test for recent divergent selection between small and large-eyed *Diplotaxodon*, we applied the cross-population extended haplotype homozygosity (XP-EHH)

statistic (Sabeti et al. 2007) to the two representative species with the largest sample sizes—the small-eyed species *D. limnothrissa* ($N = 33$) and the large-eyed species *D.* "bigeye black dorsal" ($N = 20$). We focused specifically on genes harboring nonsynonymous GWA outliers, which we defined as candidate genes. We considered XP-EHH scores within 15 kb from nonsynonymous GWA outliers as evidence of selection if they ranked in the top 5% (quantile > 0.95) or bottom 5% (quantile < 0.05) of the genome-wide distribution of XP-EHH scores in 30 kb windows. Based on this definition, we found signals of selection around 11 of the 17 examined genes (Fig. 3).

Negative XP-EHH scores are suggestive of positive selection in the big-eyed species (Fig. 3a–e) while positive scores indicate selection in the small-eyed species (Fig. 3f–h). The strongest signal for *D.* "bigeye black dorsal" (XP-EHH score −6.8; smaller than the most negative score in any of the 10,000 genome-wide control windows; Fig. 3i) was detected downstream of the gene *cahz* (Fig. 3a and i), which encodes a carbonic anhydrase, an enzyme with a role in tissue oxygen delivery in teleosts (Rummer and Brauner 2011) (Table 1). For *D. limnothrissa*, the strongest
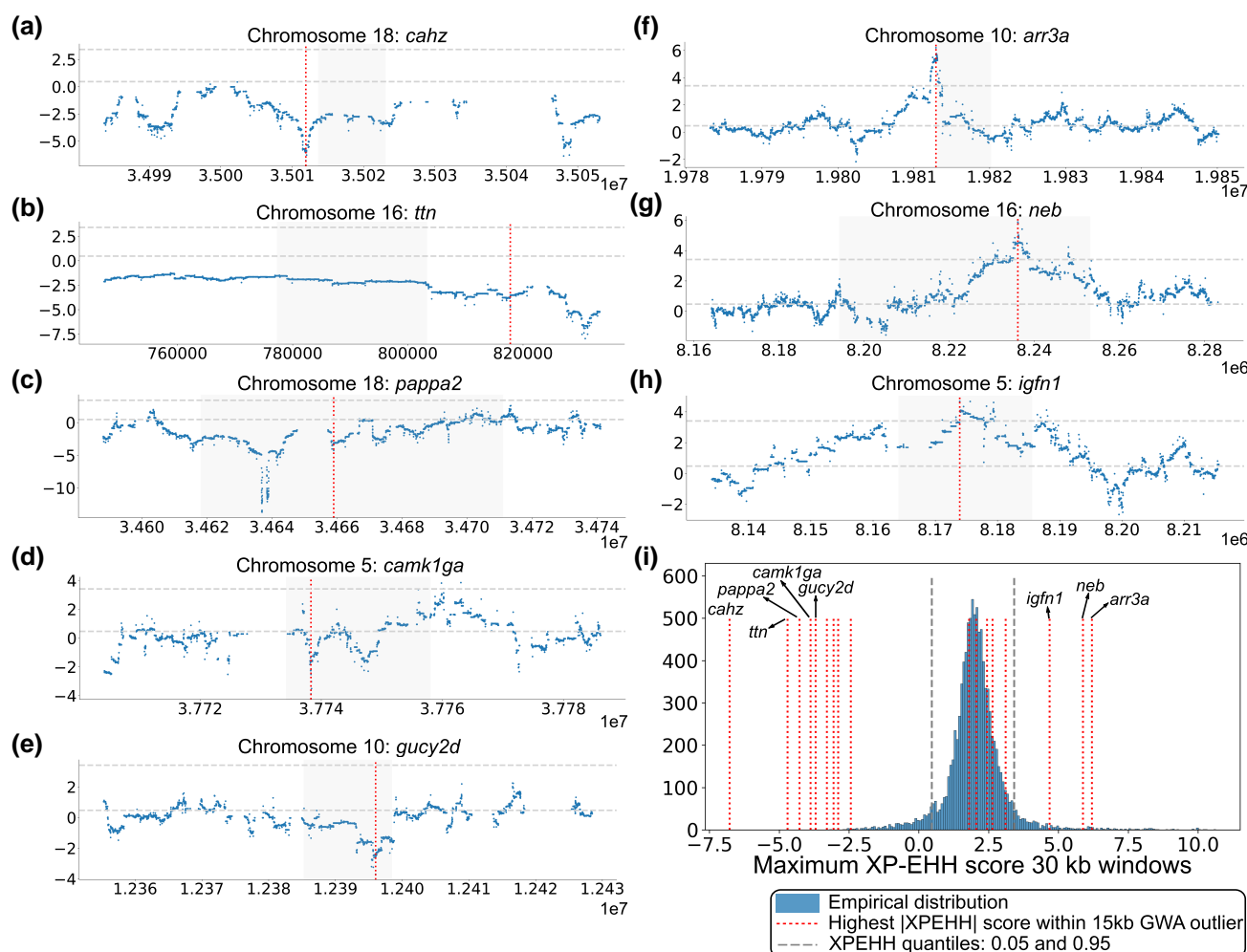
Fig. 3. Evidence of divergent selection in candidate genes between *D. limnothrissa* (*N* = 33) and *D.* "bigeye black dorsal" (*N* = 20). Extreme negative (empirical genome-wide quantile < 0.05; panels a–e) and positive (quantile > 0.95; panels f–h) XP-EHH scores (y-axis) within 15 kb from nonsynonymous GWA outliers are marked with red vertical lines. For negative outliers, only the top five genes are shown. Horizontal dashed lines indicate the 0.05 and 0.95 quantiles (see panel i). Grey shading highlights the coding region of each gene; i) Empirical distribution of maximum XP-EHH scores in random 30 kb windows across the genome from 10,000 iterations, with red vertical lines marking the highest XP-EHH score within 15 kb from nonsynonymous GWA outliers for eye size.

selection signal (score 6.2, 0.99 quantile) was found immediately upstream of the gene *arr3a* ([Fig. 3f and i]). This gene is expressed in cone photoreceptor cells and codes for arrestin 3a, which mediates the termination of phototransduction ([Pugh and Lamb 2000]) ([Table 1]). While our initial outlier detection analysis was restricted to windows containing nonsynonymous GWA outliers, in some instances XP-EHH peaks were present in the same genes but outside of these windows. For example, we found two peaks of extreme negative scores (XP-EHH < −12; more negative than any control window; [Fig. 3c]), suggesting positive selection in the big-eyed species, within the coding region of *pappa2* situated 22 kb away from the nonsynonymous GWA outlier. Another particularly interesting pattern was present around the gene *ttn*, for which negative XP-EHH scores span sites across > 50 kb suggesting that selection acted rapidly or on multiple targets, or that the region lacks recombination ([Fig. 3b]).

Selective Sweep in the Promoter Region of *arr3a*

To complement the cross-population selection scan (XP-EHH), we used an additional statistic, the integrated haplotype score (iHS), which provides evidence of recent or on-going selective sweeps within species ([Voight et al. 2006]). This analysis revealed a strong signal of selection in *D.* "bigeye black dorsal" immediately upstream of the transcription start site (TSS) of the vision-related *arr3a*, almost perfectly aligned with the XP-EHH peak around this gene ([Fig. 4a]). The overlap between the two statistics is seemingly contradictory, as they indicate selection acting in big-eyed species (iHS) and small-eyed species (XP-EHH). Such a pattern could be explained if this region had experienced historical selection in both populations, differing in nature and timing. For example, high XP-EHH scores could be reflecting an older selection event in *D. limnothrissa*, while the iHS signal could be pointing at a recent or partial selective sweep in *D.* "bigeye black dorsal", with different alleles being favored in each species as expected under disruptive ecological selection ([Maynard 1962]). Consistent with this, we found a haplotype of approximately one kilobase containing derived alleles at intermediate frequency in *D.* "bigeye black dorsal", with some of the alleles also present in other species ([Fig. 4b and c]). Most derived alleles in this region occur exclusively in heterozygous state ([Fig. 4c]). To confirm that this pattern was not caused by misalignment of an unresolved
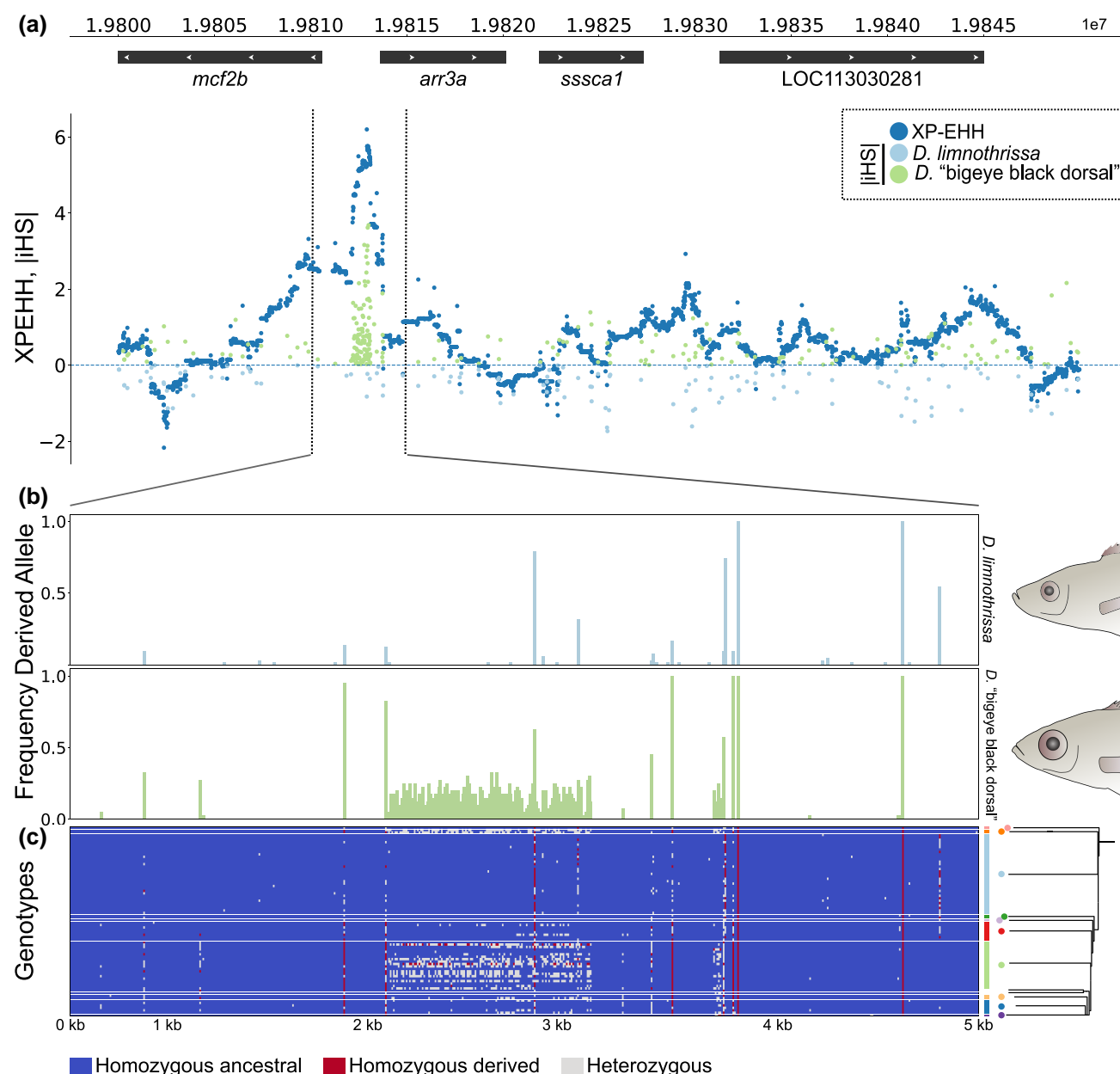
Fig. 4. Evidence of selective sweep in the putative promoter region of the vision gene *arr3a*. a) Overlap of cross-population extended haplotype homozygosity (XP-EHH, dark blue dots) and absolute integrated haplotype scores (iHS) in *D.* "bigeye black dorsal" (green, positive axis) and in *D. limnothrissa* (light blue, negative axis) in a 50 kb region containing the *arr3a* locus. b–c) Allele frequencies (b) in *D. limnothrissa* ($N = 33$) and *D.* "bigeye black dorsal" ($N = 20$) and genotypes (c) in 10 *Diplotaxodon* species across a 5 kb region spanning the XP-EHH and iHS signal. Genotypes are sorted by phylogeny.

duplication, we checked coverage and mapping quality (MQ), concluding that no duplication was present (supplementary fig. S5, Supplementary Material online). We therefore suggest that this pattern is either the result of a recent, ongoing selective sweep or caused by a combination of balancing selection and suppressed recombination. In either case, our findings point to complex, ongoing selection on the promoter region of *arr3a*, likely affecting the expression of this key gene within the phototransduction pathway.

Differential Gene Expression Between Big and Small-Eyed *Diplotaxodon*

While nonsynonymous variants were overrepresented among GWA outliers, the vast majority (169 of 190) were noncoding,

potentially reflecting gene regulatory changes contributing to trait evolution (Hoekstra and Coyne 2007; Wray 2007). Therefore, to complement the GWA analyses, we examined gene expression differences across whole-eye transcriptomes between small-eyed (*D. limnothrissa* complex; $N = 5$) and big-eyed (*D.* "macrops black dorsal" + *D.* "macrops off-shore"; $N = 5$) species. With this approach, we targeted not only genes expressed in the retina, such as phototransduction genes, but also genes expressed in other ocular tissues (e.g. lens, choroid, vitreous, blood vessels), capturing additional functions that could be relevant for depth adaptation.

Out of 20,299 analyzed genes, 4,648 were differentially expressed (DE) (adjusted $P < 0.05$) between the two groups. Among the 40 most highly DE genes (absolute log2FC > 4.5, corresponding to a fold change in expression > 23;
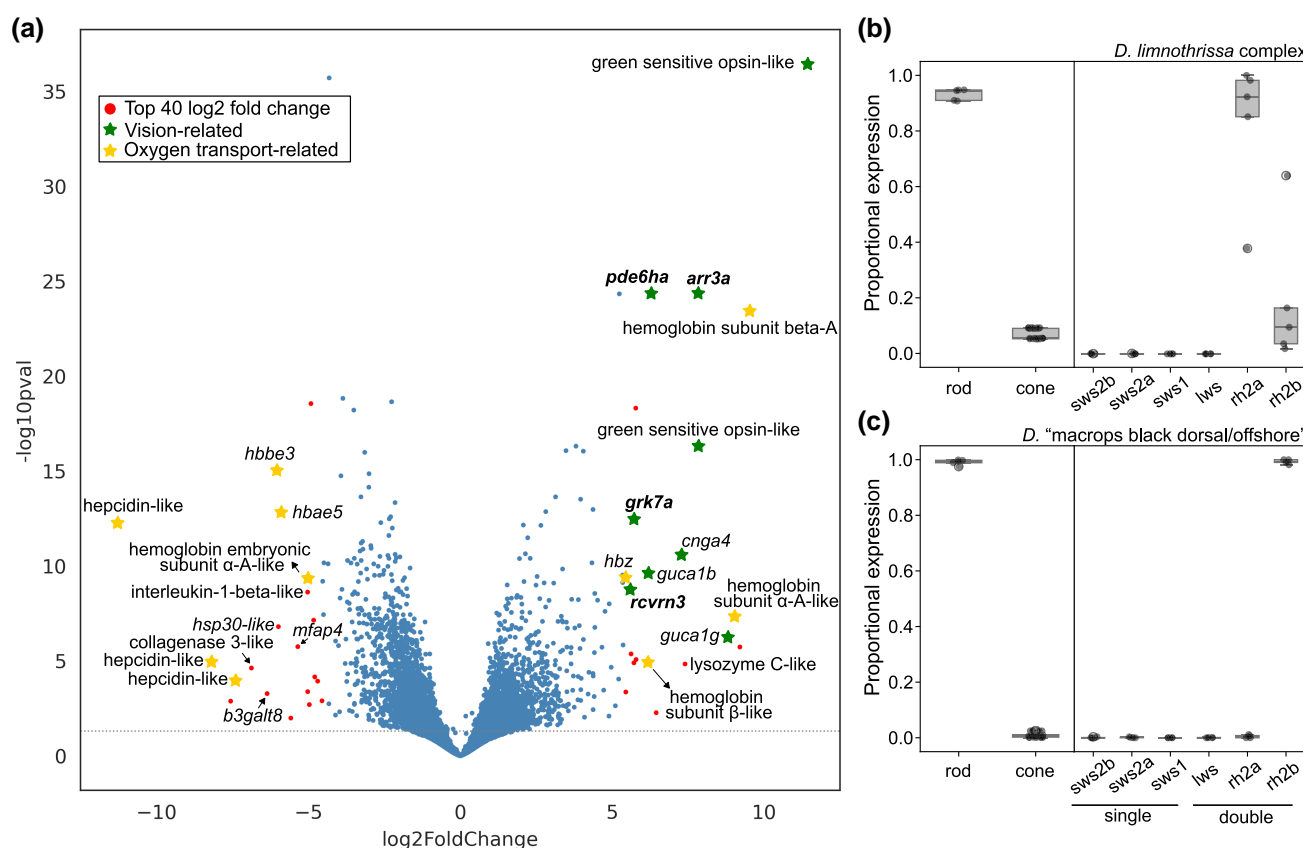
Fig. 5. Differential gene expression analysis between small and large-eyed *Diplotaxodon* and opsin expression profiles. a) Gene expression differences (log2 fold change) across whole-eye transcriptomes of small-eyed species of the *D.* "limnothrissa complex" ($N = 5$) and big-eyed species of *D.* "macrops black dorsal/offshore" ($N = 5$) plotted against the -log10 P-value (Wald test, adjusted P-value after Benjamini–Hochberg correction). Each dot represents a gene, and the horizontal line indicates the $P = 0.05$ cutoff. The top 40 most differentially expressed genes (adjusted $P < 0.05$, absolute log2FC > 4.5) are marked with red dots and colored stars. Vision-related genes (based on the *A. calliptera* annotation, fAstCal1.2), are highlighted with green stars, and oxygen transport genes with yellow stars. Candidate genes that overlap with the genomic analysis are indicated in bold; b, c) Opsin gene expression in both groups, shown as rod opsin (*rh1*) and cone opsin expression relative to total opsin expression (left panel), and individual cone opsin expression relative to total cone opsin expression (single: *sws1*, *sws2a*, *sws2b*, and double cones: *rh2a*, *rh2b*, *lws*). Expression of the two *rh2a* copies is combined into a single value. Boxplots show the first and third quartiles (boxes), the median (central line), and the minimum and maximum values (whiskers). Outliers (values outside 1.5× the interquartile range) are shown as diamonds.

supplementary table S5, Supplementary Material online) there were two green-sensitive opsins and six additional genes of the phototransduction pathway (Fig. 5a). Four of these eight genes overlapped with the GWA candidates and, interestingly, all of them were overexpressed in the small-eyed species relative to the big-eyed species. Of these, *arr3a*, the promoter region of which we have shown to be under putative recent selection (Fig. 4), ranks sixth in the list of genes overexpressed in *D. limnothrissa* (supplementary table S5, Supplementary Material online). This further supports the idea that expression changes in *arr3a* play a role in ecological specialization between small and big-eyed *Diplotaxodon*. Besides vision-related genes, we found seven highly DE hemoglobin subunit genes (Fig. 5a; supplementary table S5, Supplementary Material online) and three hepcidin genes that were underexpressed in the *D. limnothrissa* group relative to the big-eyed group (Fig. 5a). Hepcidin mediates the uptake and release of iron—an essential component of hemoglobin and myoglobin—and has been shown to be downregulated in response to hypoxia (Nemeth and Ganz 2006).

Expression Profile of Visual Opsin Genes

The subset of visual opsin genes expressed by a species— referred to as their "visual palette"—is crucial for determining spectral sensitivity and often correlates with ecological factors such as the light environment and foraging and mating behaviors (Carleton et al. 2016). To characterize the visual opsin repertoire used by small and big-eyed *Diplotaxodon*, we examined opsin gene expression values normalized as transcript per million (TPM) from the eye transcriptomes.

Rhodopsin (*rh1*) expression accounted for the majority of opsin gene expression in both groups (Fig. 5b and c). A high rod:cone opsin gene expression ratio is expected due to the much higher density of rods relative to cones in cichlid retinas (Braekevelt et al. 1998), but the extreme prevalence of rhodopsin expression observed (median [range]: 94.5% [90.8% to 94.8%] and 99.3% [97.5% to 99.9%] of total opsin expression in *D. limnothrissa* and *D.* "macrops black dorsal/offshore", respectively) is likely reflecting evolutionary adaptation to low light in the deepwater environment shared by both species groups (Carleton et al. 2020). Consistent with this, cone opsin expression was dominated by green-sensitive opsins (*rh2*), while expression of long (*lws*) and short-wavelength sensitive (*sws1*, *sws2a*, *sws2b*) opsins was virtually absent in both species (TPM < 1, accounting for $< 0.1\%$ of total cone opsin expression; Fig. 5b and c; supplementary fig. S6, Supplementary Material online). While the small-eyed species expressed both *rh2a* and *rh2b*

(90.6% [37.2% to 98.2%] and 9.4% [1.8% to 62.8%] of total cone opsin expression, respectively), the repertoire appeared more restricted in the big-eyed species, with only *rh2b* accounting for 99% [98.2% to 100%] of cone opsin expression (Fig. 5c). These differences in relative green opsin expression between small and big-eyed *Diplotaxodon* suggest fine-tuning of vision to their respective depth ranges within otherwise similar "deepwater visual palettes".

Discussion

Vision is essential for feeding, recognizing conspecifics, avoiding predators, and identifying potential mates in many organisms (Land and Nilsson 2012; Cronin et al. 2014). As a consequence, depth-related gradients of light attenuation are commonly associated with niche diversification involving both morphological and molecular adaptations of the visual system (Land and Nilsson 2012). In this study, we examined eye size variation and associated molecular changes to explore some of the adaptations that allowed the genus *Diplotaxodon* of the Malawi cichlid adaptive radiation to diversify in deepwater habitats.

Beyond the general trend for large eyes in *Diplotaxodon*—a phenotype common in deepwater fishes (Warrant and Locket 2004)—we found that *Diplotaxodon* also exhibit significant variability in relative eye size, exceeding that of any other ecomorphological clade within the radiation. This variation occurs both within and between species. Intraspecific variation may reflect phenotypic plasticity or differences between sexes and ontogeny (Meuthen et al. 2018; Svanbäck and Johansson 2019). However, much of this variability is interspecific, consistent with taxonomic reviews of the genus highlighting eye size as a distinguishing feature between some species (Turner et al. 2004).

Eye size is a complex trait shaped by multiple ecological and evolutionary forces, including diet, habitat complexity, developmental constraints and phylogenetic history (Niven and Laughlin 2008; de Busserolles et al. 2013; Lisney et al. 2020). *Diplotaxodon* are open-water zooplanktivores and predators of small fish, differing little in diet or general habitat preference aside from their presumed depth distributions (Turner et al. 2004). Although the challenges in accessing deepwater habitats limit detailed knowledge of their specific ecologies, trawl surveys and artisanal fishery data suggest that *Diplotaxodon* species with smaller eyes occupy broader depth ranges (~20 to 220 m) compared to a restriction to greater depth (~50 to 200 m) in species with larger eyes (Turner et al. 2004; Genner and Turner 2012), except for possible nocturnal migrations to the surface (Thompson et al. 1996). The steep decline in light intensity with depth likely results in a sharp contrast in the light environments experienced by species across even relatively small depth differences, suggesting that interspecific differences in eye size may, at least in part, reflect distinct visual demands associated with depth. The repeated evolution of eye size in several nonsister lineages, as revealed by our inferred genetic relationships, further supports this hypothesis. We therefore investigated eye size variation as a potential proxy for molecular (visual) adaptations in this group.

The lack of strong phylogenetic signal in eye size variation is important because it increases the power to distinguish functional genetic variation responsible for shared molecular adaptation from linked variation and neutral species divergence. And indeed, despite a relatively small sample size, we found genome-wide significant associations with eye size on five chromosomes, with a range of additional evidence (e.g. overrepresentation of amino acid changing mutations) suggesting that we were able to identify functional genetic variation.

Although none of the identified variants are known to directly control eye size, some could still play an indirect role in eye development. Notably, one of the 10 Bonferroni-significant SNPs fell in an intergenic region adjacent to the gene *fezf2*. This gene encodes a transcription factor part of a regulatory network involved in early brain patterning, during which the regions that give rise to the forebrain and the eye field compete for developmental territory (Sylvester et al. 2010; Patil et al. 2021). Divergence in this network has been shown to underlie differences in forebrain structure between rocky and sand-dwelling cichlids, with potential implications for behavior and sensory processing (Sylvester et al. 2010; Patil et al. 2021). These findings raise the possibility of an effect of *fezf2* in eye size, potentially through its role in early neural patterning, though establishing a causal relationship would require experimental validation.

Overall, rather than affecting eye size directly, we believe that many of the identified variants are causally linked to molecular traits of the visual system that co-vary with eye size, potentially because of concerted adaptation to specific light environments. Consistent with this, candidate genes showed broad enrichment for visual function. Previously, Malinsky et al. (2018) identified signals of parallel adaptation between *Diplotaxodon* and the deep benthic group involving genes in the cone phototransduction cascade, suggesting a key role of this pathway in-depth adaptation among Malawi cichlids. Our study builds on these findings by examining eye size-associated genetic variation within *Diplotaxodon*, allowing us to capture signals of adaptation occurring between species rather than across the entire clade. While we identified some of the same candidate genes (green-sensitive opsins, *arr3a*, *gnat2*, *grk7a*) (Malinsky et al. 2018), our analysis further revealed widespread signals of adaptation throughout the phototransduction cascade (Fig. 6). Furthermore, many of these genes showed differential expression between small and big-eyed species and/or signatures of selection. The cone arrestin *arr3a* stood out in our analysis as it harbored nonsynonymous mutations strongly associated with eye size variation, ranked among the most highly DE genes, and showed footprints of selection. Intriguingly, independent signals of selection in small and large-eyed species might point to diversifying selection—the main driver of sympatric speciation (Maynard 1962)—having acted on this gene, a hypothesis that warrants future research. Altogether, we find these results remarkable because they provide independent evidence for concerted adaptive evolutionary tinkering of the phototransduction pathway in the ecological divergence of *Diplotaxodon* species.

The phototransduction cascade is generally functionally conserved across vertebrates, and mutations in orthologs of genes identified in this study are associated with a large family of retinal diseases and vision impairments in humans (Felden et al. 2019; Solaki et al. 2022) and cone degeneration in animal models (Stearns et al. 2007; Naggert et al. 2023). Most of the phototransduction genes identified in our analysis are expressed in cones, which mediate vision in bright light conditions (Yokoyama 2008), and candidate nonsynonymous mutations are predominantly derived in big-eyed species. Together, this might suggest that some of these mutations
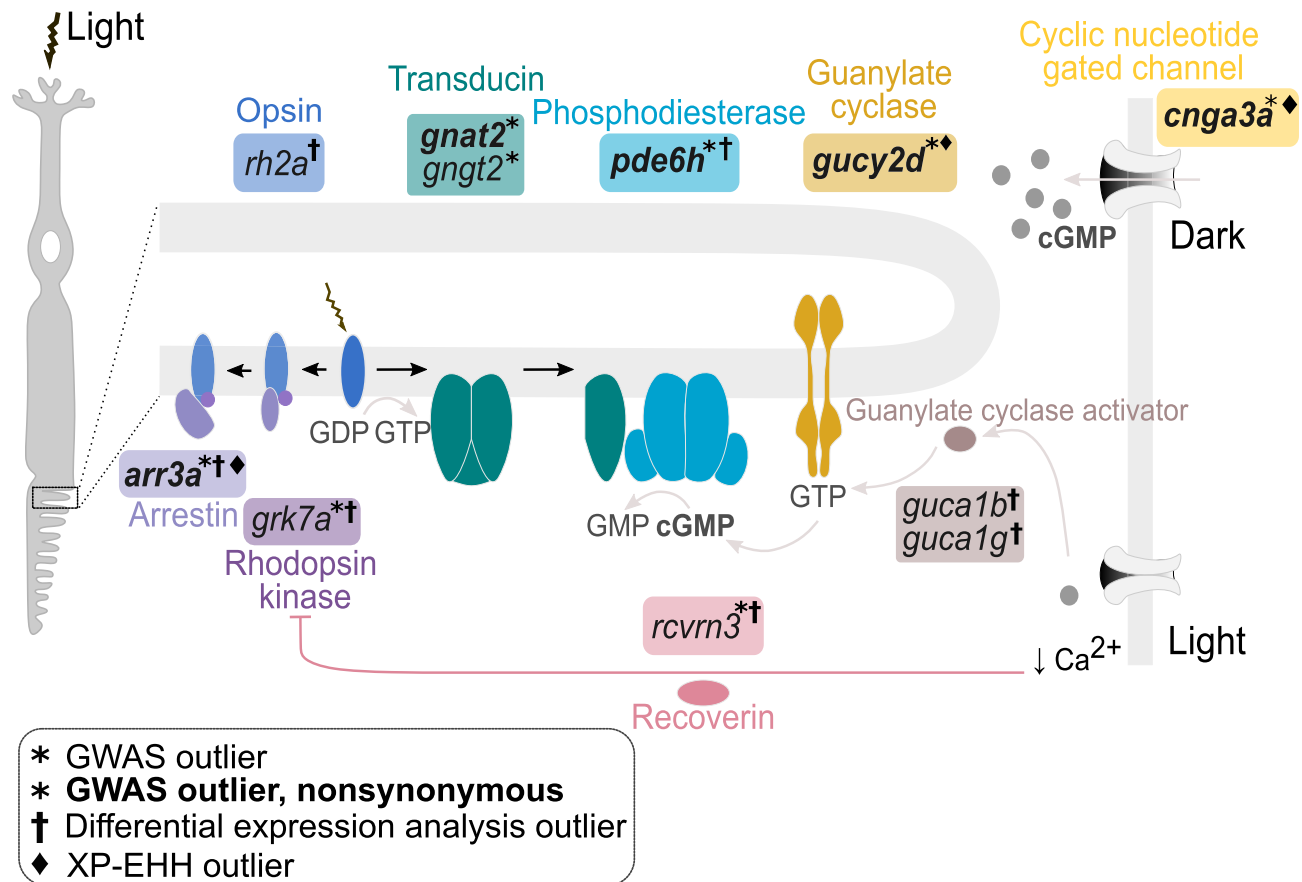
Fig. 6. Schematic of the vertebrate phototransduction cascade. The main components and events are summarized. Briefly, light activates the opsin protein, which stimulates the alpha subunit of transducin, leading to the activation of phosphodiesterase. This enzyme hydrolyzes cyclic guanosine monophosphate (cGMP), lowering its concentration in the cell and causing the cyclic gated nucleotide channels to close. This results in a reduced influx of cations (Na^+, Ca^{2+}), which is translated as an electrical signal. Reduced calcium (Ca^{2+}) concentration activates guanylate cyclase activators, which promote the restoration of cGMP, thus acting as negative feedback. The termination of the visual response involves the phosphorylation of opsin by rhodopsin kinase (which is prevented by recoverin in the dark) and the binding of arrestin. In this study, we have identified genes coding for elements involved in the different stages of phototransduction as candidates for adaptive divergence between deepwater Malawi cichlids (*Diplotaxodon*), combining GWAS for eye size, haplotype-based selection scans, and differential gene expression analysis. These genes are in boxes colored according to the type of protein they encode. This illustration was inspired by Figure 2 of Yau and Hardie 2009, Figure 4 of Fain et al. 2010, and Figure 1 of Musilova et al. 2021.

disrupt functions not required in the depth range of big-eyed species, reflecting relaxed selection on cone function. However, the parallelism in the GWA signal, short evolutionary timescales, and signals of positive selection make it more likely that these mutations are adaptive. The fact that genes in the cone phototransduction cascade are evolving in parallel with the deep benthic clade, showing some of the same amino acid changes (Malinsky et al. 2018), further supports this. Nonsynonymous mutations could provide, for example, an energy-saving mechanism (Moran et al. 2015) or, more likely, confer a direct functional advantage, such as fine-tuning the speed of activation or shutdown of the visual pigment to match species-specific ecologies (e.g. Renninger et al. 2011; Morshedian et al. 2018). Aligning with this, we found strong expression differences in genes that collectively mediate the recovery of the visual pigment after light exposure, including recoverin (*rcvrn3*), rhodopsin kinase (*grk7a*), arrestin (*arr3a*) and guanylyl cyclase activators (*guca1b, guca1g*) (Figs. 4 and 6). We hypothesize that amino acid changes and expression differences in these genes could lead to altered cone response times, potentially mediating the trade-off between cone speed and sensitivity (Warrant and Locket 2004; Land and Nilsson 2012). For example, delayed cone responses in

big-eyed *Diplotaxodon* could increase sensitivity to twilight (mesopic) conditions in their diurnal depth ranges or during nocturnal migrations to the surface (Thompson et al. 1996). It would be interesting to examine whether these species show other adaptations such as photoreceptor cell transmutations, as seen in mesopelagic deep-sea fishes (de Busserolles et al. 2017).

Our gene expression data indicate rod-dominated opsin expression in both small and big-eyed *Diplotaxodon* species, with cone opsin expression restricted to those sensitive to the blue-green region of the spectrum, consistent with patterns seen in organisms inhabiting dim habitats (reviewed in Carleton et al. 2020; de Busserolles et al. 2020; Musilova et al. 2021), including deepwater cichlids (Musilova, Indermaur et al. 2019; Ricci et al. 2023). Rhodopsin (*rh1*) mediates achromatic vision in dim-light conditions (Yokoyama 2008) and has been extensively implicated in deepwater adaptation (Sugawara et al. 2005; Nagai et al. 2011; Malinsky et al. 2015; Musilova, Cortesi et al. 2019; Ricci et al. 2022), also in *Diplotaxodon* (Hahn et al. 2017; Malinsky et al. 2018). The high rhodopsin expression observed could be driven by differences in photoreceptor abundance and influenced by eye size (Ricci et al. 2023). Further investigation, such as retinal

microscopic examination and cell type-specific expression analyses (e.g. scRNAseq; Ogawa and Corbo 2021), is needed to determine whether this pattern reflects a higher number of rods in larger eyes or proportionally higher rod expression. Despite small and big-eyed species exhibiting typical "deepwater palettes", there were substantial differences in the green opsin types expressed, with big-eyed *Diplotaxodon* almost exclusively expressing *rh2b* while small-eyed species expressed a combination of *rh2b* and *rh2aβ*. Visual modeling in the Malawi cichlid *Maylandia zebra* showed that co-expression of *rh2b* and *rh2aβ* can have a significant impact on contrast detection and color discrimination, as opposed to single expression of *rh2b* (Dalton et al. 2014). This suggests that differential opsin (co)expression in *Diplotaxodon* could provide species-specific ecological specializations. It would be interesting to examine how opsin expression varies within their retinas.

Overall, the limited opsin repertoire and consistently lower expression across several other (cone) phototransduction genes in big-eyed species suggest a decreased reliance on cone-based vision. However, the signatures of adaptation in these genes and expression of a cone green opsin, though at very low levels, suggest that color discrimination may be important in *Diplotaxodon*. In many deep-sea and nocturnal fish, green opsins are the only opsins remaining, though the functional relevance of expressing cone opsins in their extremely dim environments remains unclear (Musilova et al. 2021; Musilova and Cortesi 2023). In the deep, clear, open waters of Lake Malawi, predominant wavelengths align with peak absorbance of *rh2* genes (Carleton et al. 2016), possibly explaining the retention of cone function. This raises the question of whether opsin expression differences contribute to assortative-mating-based species isolation in our study system (Seehausen et al. 2008). *Diplotaxodon* species are expected to be fully interfertile. They coexist in an open-water environment void of any physical barriers and some genetically distinct but sympatric species are morphologically indistinguishable except by male breeding dress (Turner et al. 2004). Previous research has suggested a role for male nuptial patterning in species recognition and assortative mating in *Diplotaxodon* (Genner et al. 2007). Studying the link between species-specific opsin expression, spectral sensitivity, and the ability to discriminate nuptial patterning and hue in their light environments could offer meaningful insights into the role of visual adaptation in speciation. Unfortunately, their inaccessible habitat depths make it difficult to delimit their specific depth ranges and preclude behavioral studies.

Our analysis has also shed light on broader patterns of physiological depth adaptation, involving the hemoglobin complex—previously implicated in adaptation to anoxic and deepwater habitats in cichlids including *Diplotaxodon* (Hahn et al. 2017; Malinsky et al. 2018; Vernaz et al. 2022; Ricci et al. 2023)—and related genes with oxygen transport functions such as hepcidin and carbonic anhydrase. Interestingly, two of the candidate genes, *trhde* and *retsat*, have been linked to high-altitude adaptation in mammals, specifically in regulating body temperature and hypoxia response (Table 1; Edea et al. 2019; Xu et al. 2021). This highlights a certain degree of parallelism in selective pressures between deepwater and high-altitude environments, offering potential for comparative studies on parallel evolution. Besides this possible link to high-altitude adaptation, the presence of *trhde* among our candidates is intriguing because of its potential

role in visual plasticity. This gene, which codes for thyrotropin-releasing hormone degrading peptide, is part of the thyroid hormone (TH) signaling pathway, known to influence spectral sensitivity in vertebrates by regulating opsin expression and chromophore shifts in visual pigments (Volkov et al. 2020). TH-mediated visual plasticity in response to light changes has been shown in Midas cichlids (Härer et al. 2017; Karagic et al. 2018, 2022) but its role remains underexplored in East African cichlids. It would be interesting to investigate whether *trhde* contributes to visual spectral tuning, potentially by modulating TH concentration in *Diplotaxodon* and, more generally, in cichlid radiations.

In conclusion, our study reveals that the evolution of eye size is tightly linked to broad visual adaptive tinkering largely involving the phototransduction cascade, uncovering some of the molecular mechanisms that have enabled *Diplotaxodon* to diversify in deepwater environments. While much previous research has focused on visual opsins due to their crucial role in modulating spectral sensitivity and their potential to aid speciation, further research is needed to elucidate the phenotypic and fitness impacts of the identified protein modifications and regulatory changes. Nonetheless, our results offer important insights into depth adaptation within one of the largest extant vertebrate evolutionary radiations, contributing to our understanding of how species radiate into novel environments.

Materials and Methods

Sampling and DNA Sequencing

A total of 1,360 samples were used in the analysis of eye size variation across the Malawi cichlid radiation and the *Diplotaxodon* genomics analyses (supplementary Data S1, Supplementary Material online). They were collected in 2016 and 2017 (with the exception of seven samples, which were collected in 2004) from fishermen along the shores of Lake Malawi, caught by a variety of methods including hook and line fishing, mid-water trawling, gill nets, and nighttime Chirimila, a type of purse seine, or by the Ndunduma research trawler of the Fisheries Research Unit of the Government of Malawi. Specimens were photographed in lefthand side lateral view with a scale bar. Most of these samples have been included in previous studies (Malinsky et al. 2018; Svardal et al. 2020; Blumer et al. 2025). DNA was extracted from fin clips using standard protocols and short reads (100 to 150 bp paired-end) were obtained on Illumina HiSeq platforms as described in Malinsky et al. (2018) and Blumer et al. (2025).

For RNA sequencing, freshly caught, adult individuals were purchased from the Nkhata Bay market in 2018; their eyes immediately dissected after collection and preserved in 5 to 10 mL of RNA later. To minimize the potential influence of diurnal variation in gene expression, all dissections took place in the morning.

Details of the *Diplotaxodon* whole-genome sequences ($N = 79$) and the collection of samples for RNA sequencing ($N = 10$) included in this study can be found in supplementary Supplementary Data S3 and S4, Supplementary Material online. Sample collection, export, and use are covered by permits of the Government of Malawi.

Alignment and Variant Calling

Sequencing reads were aligned to the high-quality reference genome of a closely related Malawi cichlid, *Astatotilapia*

calliptera (<0.3% sequence divergence, fAstCal1.2, GCA_900246225.3) using the mem algorithm of bwa v0.7.17. Variants were called and filtered using bcftools v1.9. Sites were excluded if they showed poor read mapping, unusual depth of coverage, biased allele balance in heterozygotes, or an excess of missing genotypes. Specifically, sites with a root mean square MQ below 50.0 and with a fraction of MQ zero reads greater than 0.10 were filtered to remove low-confidence calls. Sites were further excluded if more than 20% of the genotypes were missing, if the test of excess heterozygosity as implemented in bcftools showed a value below 0.2, if the Mann–Whitney U test of MQ versus strand bias was highly significant ($P < 0.001$), or if allele balance in heterozygous calls was significantly skewed ($P < 0.01$). Additionally, sites with abnormally variable read depths were excluded using a depth deviation metric, defined as the average across all samples of the deviation of read depth at a focal site from the genome-wide mean coverage of each individual, expressed in units of standard deviation. Sites with a depth deviation > 1.5 were filtered out. Biallelic SNPs passing all filtering criteria were phased using Eagle v2.4.1 (Loh et al. 2016) with "--Kpbwt = 40,000" and "--expectIBDcM = 0.5" to suit the cichlid model (Blumer et al. 2025). Missing genotypes were imputed (default). The variant calling workflow was performed on a larger sample set of 1,465 Lake Malawi cichlids, most of which are presented in Blumer et al. (2025), and then down sampled to only include the *Diplotaxodon* individuals used in this study (supplementary Data S3, Supplementary Material online).

Diplotaxodon Genetic Relationships
Pairwise genetic differences were calculated using custom Python code available at https://github.com/feilchenfeldt/pypopgen3. A neighbor-joining (NJ) tree was constructed using scikit-bio v0.5.6. The tree was rooted using an ancestral sequence reconstructed from whole-genome alignments between the *A. calliptera* reference genome and outgroup species (data and methods available in Blumer et al. 2025).

Eye Size and Body Shape Measurements
Morphological data were obtained from coordinates of 18 homologous landmark points from digital photographs following Theis et al. (2014) (see supplementary fig. S1, Supplementary Material online therein), using tpsDig2 v2.26 (Rohlf 2004). The landmarks surrounding the eye were used for the eye size analysis. To remove the effect of body size on eye size variation, for the GWA analysis we defined eye size as the ratio of horizontal eye diameter to standard length (supplementary fig. S7, Supplementary Material online). Eye size measurements were obtained for 75 *Diplotaxodon* specimens and for 1,259 specimens of the remaining five ecomorphological clades of the radiation included in this study (supplementary Data S1 and S2, Supplementary Material online).

A general measure of body shape, used in the calculation of the correlation between *Diplotaxodon* body shape differences and genetic distance (supplementary fig. S3, Supplementary Material online), was obtained by summarizing gross body morphology from coordinates of all 18 landmarks (supplementary fig. S7, Supplementary Material online) using principal component analysis (PCA) in MorphoJ v1.07a. To remove variation due to size, orientation, and position of the

fish, we first applied a Procrustes superimposition of the data. Subsequently, PCA was applied to the Procrustes coordinates to summarize shape variation. The first principal component (PC1), accounting for 54% of the variation, was used as a representative measure of body shape.

Phylogenetic Signal
As a measure of phylogenetic signal for eye size in *Diplotaxodon*, we estimated Pagel's lambda (λ) using the R package phytools (Revell 2012) v.0.7.70 (function "phylosig").

GWA Analysis
The GWA analysis for eye size was conducted by fitting a univariate linear mixed model using the software GEMMA v0.98 (Zhou and Stephens 2012). Sites with minor allele frequency (MAF) < 5% were excluded. A centered relatedness matrix (option "-gk 1") was used to account for population structure. After MAF filtering, 1,900,366 SNPs were analyzed. The *P*-values reported correspond to the likelihood ratio test computed by GEMMA. A total of 53 specimens of 9 *Diplotaxodon* species were included in this analysis. These were all specimens for which both phenotypes and sequences were available (supplementary Data S3, Supplementary Material online). Note that sample sizes differed between species (see supplementary table S1, Supplementary Material online).

Principal Component Analysis
PCAs to summarize genetic variation in GWA outliers (top-ranking 0.01% SNPs, $N = 190$) and genome-wide were conducted using PLINK v1.9. For the PCA of genome-wide variation, SNPs with MAF < 0.05 were excluded using bcftools v1.14 "--min-af 0.05:minor" and SNPs were linkage-disequilibrium pruned using the option "--indep-pairwise 50 10 0.1" in PLINK. The genome-wide PCA was performed on the remaining 377,273 SNPs. The linear regression model of PC1 on eye size was fitted by ordinary least squares (OLS) using the Python module statsmodels v0.14.0 (supplementary table S2, Supplementary Material online).

Identification of GWA Outliers and Candidate Genes
GWA outliers were identified based on the top 0.01% most significant associations (190 SNPs, $P < 0.0001$), corresponding to the 99.99th percentile of the $-\log_{10}(P\text{-value})$ distribution. This threshold was chosen to increase the number of variants for downstream analyses while retaining the strongest signals. Additional thresholds, such as the 0.025% outlier level, and exploratory searches for vision-related genes were conducted to provide broader context but all downstream analyses were based on the SNPs passing the 0.01% cutoff, referred to as GWA outliers throughout the text. Candidate genes were identified as those containing nonsynonymous SNPs among the GWA outliers (17 genes, Table 1).

Annotation of GWA Outlier SNPs
The SNP annotation was performed using snpEff v5.1 (Cingolani, Platts, et al. 2012) using a prebuilt database for the *Astatotilapia calliptera* reference genome fAstCal1.2 (GCA_900246225.3; GenBank assembly). Throughout the text, genes are referred to as annotated by snpEff

(supplementary Data S5, Supplementary Material online). Whenever a gene name/symbol was missing but an Ensembl gene identifier (ENSACLG) was present instead, we manually searched for a match in the RefSeq genome assembly GCF_900246225.1 using the NCBI gene database (https://www.ncbi.nlm.nih.gov/gene/) and the NCBI Genome Data Viewer (https://www.ncbi.nlm.nih.gov/genome/gdv/browser/genome/?id=GCF_900246225.1).

GO Enrichment Analysis

The GO enrichment analysis was performed using the R package topGO v2.38.1 (Alexa and Rahnenfuhrer 2019) from the Bioconductor project (www.bioconductor.org).

As test gene set, we included all genes with transcription start sites (TSS) located within ± 25 kb of a GWA outlier SNP. This window was chosen to capture genes directly harboring outlier variants in their gene bodies, as well as those that may be cis-regulated by nearby variants due to their proximity to target promoters (Kleinjan and van Heyningen 2005). The background gene set was generated by extracting the protein-coding genes from the annotated GWA callset using the SnpSift (Cingolani, Patel et al. 2012) "extractFields" option. To map the genes to GO terms, we used the zebrafish annotation package (org.Dr.eg.db) from Bioconductor. Prior to this, Ensembl gene identifiers had been converted to their "gene symbol", when possible, to maximize the number of genes that were mapped to GO terms, using the Ensembl BioMart data mining tool (Ensembl Release 108, fAstCal1.2 genes) and g:Convert, provided by the web server g:Profiler (Raudvere et al. 2019). Despite implementing this approach, some genes lacked correspondence to zebrafish genes (e.g. paralogs of duplicated genes) and were thus not included in the analysis. To check for potentially relevant genes that could have been excluded this way, we manually inspected the annotation of GWA outliers for *Danio rerio* genes in the GO terms by browsing the AmiGO 2 database (https://amigo.geneontology.org/amigo/landing).

A total of 11,623 feasible genes could be mapped to terms in the Biological Process ("BP") ontology. The minimum number of genes mapping to a GO category was set to 5 (option "nodeSize = 5"). Statistical overrepresentation of GO terms was calculated by performing Fisher's exact test, using the weight algorithm (object class "weightCount"). This method improves explanatory power over the classic GO scoring algorithm by taking into account the local dependencies between GO terms that arise as a consequence of the hierarchical structure of the GO graph (Alexa et al. 2006). The estimated *P*-values (supplementary table S3, Supplementary Material online) are reported without further multiple testing correction, as suggested by the authors (Alexa and Rahnenfuhrer 2019).

Testing for Overrepresentation of Functional Annotations Among GWA Outliers

To test whether any of the functional effects predicted by snpEff was overrepresented in the genomic annotation of GWA outliers, we compared the observed number of GWA outliers per annotation effect to an empirical distribution obtained by 1,000 draws of randomized sets of SNPs, matched on MAF, as described in the following.

First, using bcftools v1.14, we extracted MAF per position for the entire callset used in the GWAS (control set) and in a reduced subset of this containing only the GWA outliers (test set, 190 SNPs). SNPs in the control set were grouped into discrete nonoverlapping MAF bins of width 0.05. Then, in each iteration, a set of 190 random SNPs matching the test set on MAF distribution was drawn from the control set. The number of SNPs per annotation effect in each random set of SNPs was counted to generate a genome-wide distribution of effects (note that one SNP can have more than one annotation). The empirical *P*-value was calculated as the proportion of permutations in which the number of SNPs with an annotation effect was larger than the observed number in the test set (supplementary fig. S4, Supplementary Material online).

The functional annotation categories tested for overrepresentation were: "3-prime UTR variant", "Downstream gene variant", "Upstream gene variant", "Intergenic region", "Intron variant", "Synonymous variant", "Missense variant" and "Splice region variant and intron variant".

Patterns of Allelic Segregation at Nonsynonymous GWA Outliers

To gain an overview of which alleles at nonsynonymous GWA outlier sites might have been under selection in small and big-eyed *Diplotaxodon* species, we inspected allelic segregation of each outlier across the phylogeny. For this analysis, we used a larger data set with all sequenced samples ($N = 79$), including those for which no eye size data was available and that had thus not been included in the GWAS (supplementary Data S3, Supplementary Material online).

The data was phased so that the reference allele corresponded to the ancestral state and the alternative allele to the derived state (see "Alignment and variant calling"). For each nonsynonymous site, we calculated allele frequency (AF) per species when at least two samples were available. Then, species for which the derived allele was dominant (AF > 0.5) at each site were grouped based on whether the group consisted of small-eyed species or big-eyed species, assigning *D. apogon*, *D. macrops* and D. "bigeye black dorsal" as big-eyed species and *D. limnothrissa*, *D.* "limnothrissa black dorsal", *D.* "holochromis 1" and *D.* "holochromis 2" as small-eyed species (supplementary table S4, Supplementary Material online). We observed that while most species consistently clustered within their eye size group, *D. longimaxilla* and *D.* "limnothrissa black dorsal" were found within both groups, not following a clear pattern, and were therefore ignored. We counted the number of sites for which the derived allele was dominant in each group. Following the observation that the derived allele was almost consistently dominant or fixed in the big-eyed species group for the phototransduction genes (11 out of 12 sites; supplementary table S4, Supplementary Material online), we tested whether this pattern of segregation was not independent of the species group by performing a chi-square test. This was repeated for genes with putative skeletal muscle functions.

Signals of Selection in Candidate Genes

To detect signatures of selection around candidate genes, i.e. genes annotated to nonsynonymous GWA outliers, we applied two haplotype-based tests: the cross-population extended haplotype homozygosity statistic, XP-EHH (Sabeti et al. 2007), and the iHS (Voight et al. 2006). Both statistics are based on the EHH concept (Sabeti et al. 2002), which measures the length

of a haplotype around a focal SNP, and work best for relatively recent sweeps, before the long, selected haplotypes are broken down by recombination. The tests differ in their power to detect complete or partial sweeps. Specifically, XP-EHH compares EHH between two populations and is used to identify loci under divergent selection, in which the selected haplotype is near or has reached fixation in one population but remains polymorphic in the other. iHS assesses EHH within a population and is used to detect very recent or partial selective sweeps, which appear as large differences in haplotype length of the ancestral and derived allele at the locus under selection. This analysis was performed using the R package rehh v3.2.2 (Gautier et al. 2017) on the two species with largest sample sizes representative of "small" and "large" eye *Diplotaxodon*: *D. limnothrissa* ($N = 33$) and *D.* "bigeye black dorsal" ($N = 20$), respectively.

We computed XP-EHH along the genome and assessed whether XP-EHH scores around the candidate genes represent outliers in the genome-wide distribution of XP-EHH scores. To do this, we computed the distribution of maximum XP-EHH scores (most positive or negative values) in 10,000 randomly selected 30 kb windows along the genome. Based on this distribution, we then ranked the maximum XP-EHH scores in 30 kb windows centered around nonsynonymous GWA outliers. Windows with top-ranking XP-EHH scores (quantile > 0.95 or < 0.05) were considered to have been under recent divergent selection.

To search for additional signals of recent selective sweeps within each population (*D. limnothrissa* and *D.* "bigeye black dorsal"), we computed iHS along the genome using rehh v3.2.270 with default parameters. As we were interested in selection on sites contributing to the divergence between small and big-eyed species, we visually inspected the candidate regions (defined as above) for peaks of absolute iHS scores present for one species but not the other.

Coverage of Mapped Reads Around arr3a Locus

To check for a potential duplication of the gene *arr3a*, we inspected the coverage of reads mapped to the reference genome using samtools v1.14 (Li et al. 2009). We estimated the median coverage in the *arr3a* gene region (chr10:19,813,620–19,820,182) per sample and normalized it by the median genome-wide coverage.

RNA Extraction, Library Preparation and Sequencing

All procedures were conducted on ice unless otherwise specified. Eye samples stored in RNAlater were thawed from −80 °C and transferred to 5 mL of Trizol. Each sample was then homogenized using a handheld homogenizer and 1 mL of the homogenate was chilled in Trizol and allowed to rest for 5 min. Next, 200 μL of chloroform (ThermoFisher Scientific) was added and the samples were vigorously shaken for 15 s, briefly vortexed, and incubated at room temperature for 15 min. The samples were then centrifuged at $300 \times g$ for 20 min at 4 °C. The supernatant was carefully transferred to a fresh tube and further processed using the Direct-Zol RNA Purification Kit (Zymo) according to the manufacturer's instructions. The quality and quantity of the extracted total RNA were assessed using Qubit (RNA BR assay, Agilent) and Tapestation (Agilent). Total RNA extracted from each eye was submitted individually for sequencing. All libraries were prepared, quality-controlled, and sequenced by Novogene Corporation (China) using the Illumina NovaSeq 6000 platform to generate paired-end reads of 150 base pairs (bp). Raw sequencing reads have been deposited in the Sequence Read Archive (SRA) under BioProject accession number PRJNA1271605 (supplementary Data S4, Supplementary Material online).

Differential Gene Expression Analysis

The raw reads from RNA-seq were mapped against the *Astatotilapia calliptera* reference genome (NCBI RefSeq accession GCF_900246225.1) using STAR v2.7.10b (Dobin et al. 2013) and the number of read counts within gene coding sequences (CDS) summarized using htseq-count v2.0.4 (Anders et al. 2015). Differential expression analysis was conducted using the R package DESeq2 v1.42.0 (Love et al. 2014). Before the analysis, the read count table was filtered to retain only protein-coding genes, and genes with low read counts were excluded by requiring at least five counts in a minimum of three samples. After filtering, 20,299 genes out of 26,070 protein-coding genes in the count dataset were included in the differential expression analysis. DE genes between the *D. limnothrissa* and *D.* "macrops black dorsal/offshore" species groups were identified as those with a Wald test (default) adjusted *P*-value (Benjamini–Hochberg method) below the cutoff FDR < 0.05. To facilitate the ranking of DE genes by the log2 fold change, we applied the apeglm method (Zhu et al. 2019) for effect size shrinkage.

Opsin Gene Expression

To examine the expression profile of visual opsin genes, we normalized the raw read counts outputted by htseq-count (see above) as transcripts per million (TPM) and identified opsin genes by searching for orthologs to the *Oreochromis niloticus* (supplementary table S6, Supplementary Material online). Using the normalized values, we calculated cone and rod opsin expression relative to total opsin gene expression and individual cone opsin expression relative to total cone opsin expression. Due to the high degree of sequence similarity between the two copies of *RH2A*, i.e. *RH2Aα* and *RH2Aβ*, ($> 90\%$) and given that the purpose of this analysis was to obtain a general overview rather than an in-depth characterization opsin expression patterns, the expression values of reads mapping to either copy were summed up and collectively expressed as *RH2A*.

Supplementary Material

Supplementary material is available at *Molecular Biology and Evolution* online.

differential gene expression analysis, and to Fabrizia Ronco for advice on phylogenetic signal analysis. Finally, we thank Karen Carleton for her early suggestions during email exchanges, which helped shape some of the ideas later presented in the discussion.

Author Contributions

J.C.G. and H.S. conceived the study. M.M. collected geometric morphometrics data. G.V. and M.E.S. collected samples for RNA-seq and M.E.S. performed RNA extractions. D.J. contributed financial support for RNA-seq. H.S. produced the original variant calling set and the neighbor-joining phylogenies. J.C.G. conducted all analyses, with supervision from H.S. and input from M.M. J.C.G. and H.S. wrote the initial manuscript draft with contributions from M.M. All authors commented on and approved the final manuscript.

Competing Interests

The authors declare no competing interests.

Data Availability

Raw whole-genome sequencing reads for *Diplotaxodon* samples published in previous studies are available in the European Nucleotide Archive (ENA) under BioProjects PRJEB1254, PRJEB15289, and PRJEB21343. RNA-seq reads are available under BioProject PRJNA1271605. Corresponding BioSample accessions are listed in supplementary Data S3 and S4, Supplementary Material online. Variant call format (VCF) files are available on Zenodo: https://doi.org/10.5281/zenodo.15576996. The code and additional data associated with this study are available on GitHub (https://github.com/j-camacho/Diplotaxodon_deepwater_adaptation) and archived at Zenodo: https://doi.org/10.5281/zenodo.15602855. We note that the genetic material and sequences used in this study are subject to an Access and Benefit Sharing (ABS) agreement with the Government of Malawi. Any person who wishes to use this data for any form of commercial purpose must first enter into a commercial licensing and benefit-sharing agreement with the government of Malawi.

References

Acharya M, Huang LJ, Fleisch VC, Allison WT, Walter MA. A complex regulatory network of transcription factors critical for ocular development and disease. *Hum Mol Genet*. 2011:20(8):1610–1624. https://doi.org/10.1093/hmg/ddr038.

Alexa A, Rahnenfuhrer J. *topGO: Enrichment analysis for gene ontology*. R package version 2.38.1. 2019.

Alexa A, Rahnenführer J, Lengauer T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*. 2006:22(13):1600–1607. https://doi.org/10.1093/bioinformatics/btl140.

Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015:31(2):166–169. https://doi.org/10.1093/bioinformatics/btu638.

Blumer LM, Burskaia V, Artiushin I, Saha J, Garcia JC, Elkin J, Fischer B, Zhou C, Gresham S, Malinsky M, *et al*. Introgression dynamics of sex-linked chromosomal inversions shape the Malawi cichlid adaptive radiation. *Science*. 2025:388(6752):eadr9961. https://doi.org/10.1126/science.adr9961.

Bowmaker JK. Evolution of vertebrate visual pigments. *Vision Res*. 2008:48(20):2022–2041. https://doi.org/10.1016/j.visres.2008.03.025.

Braekevelt CR, Smith SA, Smith BJ. Photoreceptor fine structure in *Oreochromis niloticus* L. (Cichlidae; Teleostei) in light- and dark-adaptation. *Anat Rec*. 1998:252(3):453–461. https://doi.org/10.1002/(SICI)1097-0185(199811)252:3<453::AID-AR13>3.0.CO;2-3.

Carleton K. Cichlid fish visual systems: mechanisms of spectral tuning. *Integr Zool*. 2009:4(1):75–86. https://doi.org/10.1111/j.1749-4877.2008.00137.x.

Carleton KL, Dalton BE, Escobar-Camacho D, Nandamuri SP. Proximate and ultimate causes of variable visual sensitivities: insights from cichlid fish radiations. *Genesis*. 2016:54(6):299–325. https://doi.org/10.1002/dvg.22940.

Carleton KL, Escobar-Camacho D, Stieb SM, Cortesi F, Marshall NJ. Seeing the rainbow: mechanisms underlying spectral sensitivity in teleost fishes. *J Exp Biol*. 2020:223(8):jeb193334. https://doi.org/10.1242/jeb.193334.

Carleton KL, Yourick MR. Axes of visual adaptation in the ecologically diverse family Cichlidae. *Semin Cell Dev Biol*. 2020:106:43–52. https://doi.org/10.1016/j.semcdb.2020.04.015.

Cingolani P, Patel VM, Coon M, Nguyen T, Land SJ, Ruden DM, Lu X. Using *Drosophila melanogaster* as a model for genotoxic chemical mutational studies with a new program, SnpSift. *Front Genet*. 2012:3:35. https://doi.org/10.3389/fgene.2012.00035.

Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)*. 2012:6(2):80–92. https://doi.org/10.4161/fly.19695.

Cronin TW, Johnsen S, Marshall NJ, Warrant EJ. *Visual ecology*. Princeton (NJ): Princeton University Press; 2014.

Dalton BE, Loew ER, Cronin TW, Carleton KL. Spectral tuning by opsin coexpression in retinal regions that view different parts of the visual field. *Proc R Soc B Biol Sci*. 2014:281:20141980. http://doi.org/10.1098/rspb.2014.1980.

Damsgaard C, Lauridsen H, Harter TS, Kwan GT, Thomsen JS, Funder AM, Supuran CT, Tresguerres M, Matthews PG, Brauner CJ. A novel acidification mechanism for greatly enhanced oxygen supply to the fish retina. *eLife*. 2020:9:e58995. https://doi.org/10.7554/eLife.58995.

de Busserolles F, Cortesi F, Helvik JV, Davies WIL, Templin RM, Sullivan RKP, Michell CT, Mountford JK, Collin SP, Irigoien X, *et al*. Pushing the limits of photoreception in twilight conditions: the rod-like cone retina of the deep-sea pearlsides. *Sci Adv*. 2017:3(11):eaao4709. https://doi.org/10.1126/sciadv.aao4709.

de Busserolles F, Fitzpatrick JL, Paxton JR, Marshall NJ, Collin SP. Eye-size variability in deep-sea lanternfishes (myctophidae): an ecological and phylogenetic study. *PLoS One*. 2013:8(3):e58519. https://doi.org/10.1371/journal.pone.0058519.

de Busserolles F, Fogg L, Cortesi F, Marshall J. The exceptional diversity of visual adaptations in deep-sea teleost fishes. *Semin Cell Dev Biol*. 2020:106:20–30. https://doi.org/10.1016/j.semcdb.2020.05.027.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013:29(1):15–21. https://doi.org/10.1093/bioinformatics/bts635.

Džulová D, Lawless D, Pinton GG, Renner NA, Schorderet DF. Incomplete recovery of zebrafish retina following cryoinjury. *Cells*. 2022:11(8):1373. https://doi.org/10.3390/cells11081373.

Eccles DH. An outline of the physical limnology of Lake Malawi (Lake Nyasa). *Limnol Oceanogr*. 1974:19(5):730–742. https://doi.org/10.4319/lo.1974.19.5.0730.

Edea Z, Dadi H, Dessie T, Kim KS. Genomic signatures of high-altitude adaptation in Ethiopian sheep populations. *Genes Genomics*. 2019:41(8):973–981. https://doi.org/10.1007/s13258-019-00820-y.

Fain GL, Hardie R, Laughlin SB. Phototransduction and the evolution of photoreceptors. *Curr Biol*. 2010:20(3):R114–R124. https://doi.org/10.1016/j.cub.2009.12.006.

Felden J, Baumann B, Ali M, Audo I, Ayuso C, Bocquet B, Casteels I, Garcia-Sandoval B, Jacobson SG, Jurklies B, *et al*. Mutation spectrum and clinical investigation of achromatopsia patients with mutations in the GNAT2 gene. *Hum Mutat*. 2019:40(8):1145–1155. https://doi.org/10.1002/humu.23768.

Gautier M, Klassmann A, Vitalis R. Rehh 2.0: a reimplementation of the R package rehh to detect positive selection from haplotype structure. *Mol Ecol Resour*. 2017:17(1):78–90. https://doi.org/10.1111/1755-0998.12634.

Genner MJ, Nichols P, Carvalho GR, Robinson RL, Shaw PW, Turner GF. Reproductive isolation among deep-water cichlid fishes of Lake Malawi differing in monochromatic male breeding dress. *Mol Ecol*. 2007:16(3):651–662. https://doi.org/10.1111/j.1365-294X.2006.03173.x.

Genner MJ, Turner GF. Ancient hybridization and phenotypic novelty within lake Malawi's cichlid fish radiation. *Mol Biol Evol*. 2012:29(1):195–206. https://doi.org/10.1093/molbev/msr183.

Hahn C, Genner MJ, Turner GF, Joyce DA. The genomic basis of cichlid fish adaptation within the deepwater 'twilight zone' of Lake Malawi. *Evol Lett*. 2017:1(4):184–198. https://doi.org/10.1002/evl3.20.

Härer A, Torres-Dowdall J, Meyer A. Rapid adaptation to a novel light environment: the importance of ontogeny and phenotypic plasticity in shaping the visual system of Nicaraguan Midas cichlid fish (*Amphilophus citrinellus* spp). *Mol Ecol*. 2017:26(20):5582–5593. https://doi.org/10.1111/mec.14289.

Hoekstra HE, Coyne JA. The locus of evolution: evo devo and the genetics of adaptation. *Evolution*. 2007:61(5):995–1016. https://doi.org/10.1111/j.1558-5646.2007.00105.x.

Karagic N, Härer A, Meyer A, Torres-Dowdall J. Heterochronic opsin expression due to early light deprivation results in drastically shifted visual sensitivity in a cichlid fish: possible role of thyroid hormone signaling. *J Exp Zool*. 2018:330(4):202–214. https://doi.org/10.1002/jez.b.22806.

Karagic N, Härer A, Meyer A, Torres-Dowdall J. Thyroid hormone tinkering elicits integrated phenotypic changes potentially explaining rapid adaptation of color vision in cichlid fish. *Evolution*. 2022:76(4):837–845. https://doi.org/10.1111/evo.14455.

Kleinjan DA, van Heyningen V. Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am J Hum Genet*. 2005:76(1):8–32. https://doi.org/10.1086/426833.

Konings A. *Malawi Cichlids in their natural habitat*. El Paso (USA): Cichlid Press; 2016.

Lamb TD, Curtin J. Photoreceptor physiology and evolution: cellular and molecular basis of rod and cone phototransduction. *J Physiol*. 2022:600(21):4585–4601. https://doi.org/10.1113/JP282058.

Land MF, Nilsson D. *Animal eyes*. 2nd ed. Oxford: Oxford University Press; 2012.

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

Lie KK, Meier S, Sørhus E, Edvardsen RB, Karlsen Ø, Olsvik PA. Offshore crude oil disrupts retinoid signaling and eye development in larval Atlantic haddock. *Front Mar Sci*. 2019:6:368. https://doi.org/10.3389/fmars.2019.00368.

Lisney TJ, Collin SP, Kelley JL. The effect of ecological factors on eye morphology in the western rainbowfish, *Melanotaenia australis*. *J Exp Biol*. 2020:223(10):jeb223644. https://doi.org/10.1242/jeb.223644.

Loh PR, Danecek P, Palamara PF, Fuchsberger C, Reshef YA, Finucane HK, Schoenherr S, Forer L, McCarthy S, Abecasis GR, *et al.* Reference-based phasing using the haplotype reference consortium panel. *Nat Genet*. 2016:48(11):1443–1448. https://doi.org/10.1038/ng.3679.

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014:15(12):550–521. https://doi.org/10.1186/s13059-014-0550-8.

Malinsky M, Challis RJ, Tyers AM, Schiffels S, Terai Y, Ngatunga BP, Miska EA, Durbin R, Genner MJ, Turner GF. Genomic islands of speciation separate cichlid ecomorphs in an East African crater lake. *Science*. 2015:350(6267):1493–1498. https://doi.org/10.1126/science.aac9927.

Malinsky M, Svardal H, Tyers AM, Miska EA, Genner MJ, Turner GF, Durbin R. Whole-genome sequences of Malawi cichlids reveal multiple radiations interconnected by gene flow. Nat. *Ecol Evol*. 2018:2:1940–1955. https://doi.org/10.1038/s41559-018-0717-x.

Maynard SJ. Disruptive selection, polymorphism and sympatric speciation. *Nature*. 1962:195(4836):60–62. https://doi.org/10.1038/195060a0.

McGee MD, Borstein SR, Meier JI, Marques DA, Mwaiko S, Taabu A, Kishe MA, O'Meara B, Bruggmann R, Excoffier L, *et al.* The ecological and genomic basis of explosive adaptive radiation. *Nature*. 2020:586(7827):75–79. https://doi.org/10.1038/s41586-020-2652-7.

Meuthen D, Baldauf SA, Bakker TCM, Thünken T. Neglected patterns of variation in phenotypic plasticity: age- and sex-specific antipredator plasticity in a cichlid fish. *Am Nat*. 2018:191(4):475–490. https://doi.org/10.1086/696264.

Miller AH, Howe HB, Krause BM, Friedle SA, Banks MI, Perkins BD, Wolman MA. Pregnancy-associated plasma protein-aa regulates photoreceptor synaptic development to mediate visually guided behavior. *J Neurosci*. 2018:38(22):5220–5236. https://doi.org/10.1523/JNEUROSCI.0061-18.2018.

Moran D, Softley R, Warrant EJ. The energetic cost of vision and the evolution of eyeless Mexican cavefish. *Sci Adv*. 2015:1(8):e1500363. https://doi.org/10.1126/sciadv.1500363.

Morshedian A, Woodruff ML, Fain GL. Role of recoverin in rod photoreceptor light adaptation. *J Physiol*. 2018:596(8):1513–1526. https://doi.org/10.1113/JP275779.

Mukherjee S, Bartoš O, Zdě Nková K, Hanák P, Horká P, Musilova Z. Evolution of the parvalbumin genes in teleost fishes after the whole-genome duplication. *Fishes*. 2021:6(4):70. https://doi.org/10.3390/fishes6040070.

Musilova Z, Cortesi F. The evolution of the green-light-sensitive visual opsin genes (RH2) in teleost fishes. *Vision Res*. 2023:206:108204. https://doi.org/10.1016/j.visres.2023.108204.

Musilova Z, Cortesi F, Matschiner M, Davies WIL, Patel JS, Stieb SM, De Busserolles F, Malmstrøm M, Tørresen OK, Brown CJ, *et al.* Vision using multiple distinct rod opsins in deep-sea fishes. *Science*. 2019:364(6440):588–592. https://doi.org/10.1126/science.aav4632.

Musilova Z, Indermaur A, Bitja-Nyom AR, Omelchenko D, Kłodawska M, Albergati L, Remišová K, Salzburger W. Evolution of the visual sensory system in cichlid fishes from crater lake Barombi Mbo in Cameroon. *Mol Ecol*. 2019:28(23):5010–5031. https://doi.org/10.1111/mec.15217.

Musilova Z, Salzburger W, Cortesi F. The visual opsin gene repertoires of teleost fishes: evolution, ecology, and function. *Annu Rev Cell Dev Biol*. 2021:37(1):441–468. https://doi.org/10.1146/annurev-cellbio-120219-024915.

Nagai H, Terai Y, Sugawara T, Imai H, Nishihara H, Hori M, Okada N. Reverse evolution in RH1 for adaptation of cichlids to water depth in Lake Tanganyika. *Mol Biol Evol*. 2011:28(6):1769–1776. https://doi.org/10.1093/molbev/msq344.

Naggert ASEN, Collin GB, Wang J, Krebs MP, Chang B. A mouse model of cone photoreceptor function loss (cpfl9) with degeneration due to a mutation in Gucy2e. *Front Mol Neurosci*. 2022:15:1080136. https://doi.org/10.3389/fnmol.2022.1080136.

Nemeth E, Ganz T. Regulation of iron metabolism by hepcidin. *Annu Rev Nutr*. 2006:26(1):323–342. https://doi.org/10.1146/annurev.nutr.26.061505.111303.

Ng L, Hurley JB, Dierks B, Srinivas M, Saltó C, Vennström B, Reh TA, Forrest D. A thyroid hormone receptor that is required for the development of green cone photoreceptors. *Nat Genet*. 2001:27(1):94–98. https://doi.org/10.1038/83829.

Niven JE, Laughlin SB. Energy limitation as a selective pressure on the evolution of sensory systems. *J Exp Biol*. 2008:211(11):1792–1804. https://doi.org/10.1242/jeb.017574.

Ogawa Y, Corbo JC. Partitioning of gene expression among zebrafish photoreceptor subtypes. *Sci Rep*. 2021:11(1):17340. https://doi.org/10.1038/s41598-021-96837-z.

Pagel M. Inferring the historical patterns of biological evolution. *Nature*. 1999:401(6756):877–884. https://doi.org/10.1038/44766.

Patil C, Sylvester JB, Abdilleh K, Norsworthy MW, Pottin K, Malinsky M, Bloomquist RF, Johnson ZV, McGrath PT, Streelman JT.

Genome-enabled discovery of evolutionary divergence in brains and behavior. *Sci Rep.* 2021:11(1):13016. https://doi.org/10.1038/s41598-021-92385-8.

Permyakov EA, Uversky VN. What is parvalbumin for? *Biomolecules.* 2022:12(5):656. https://doi.org/10.3390/biom12050656.

Price AL, Zaitlen NA, Reich D, Patterson N. New approaches to population stratification in genome-wide association studies. *Nat Rev Genet.* 2010:11(7):459–463. https://doi.org/10.1038/nrg2813.

Pugh EN, Lamb TD. Phototransduction in vertebrate rods and cones: molecular mechanisms of amplification, recovery and light adaptation. In: *Handbook of biological physics.* Vol. 3. Amsterdam: North-Holland; 2000. p. 183–255.

Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J. G:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* 2019:47(W1):W191–W198. https://doi.org/10.1093/nar/gkz369.

Renninger SL, Gesemann M, Neuhauss SCF. Cone arrestin confers cone vision of high temporal resolution in zebrafish larvae. *Eur J Neurosci.* 2011:33(4):658–667. https://doi.org/10.1111/j.1460-9568.2010.07574.x.

Revell LJ. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol.* 2012:3(2):217–223. https://doi.org/10.1111/j.2041-210X.2011.00169.x.

Ricci V, Ronco F, Boileau N, Salzburger W. Visual opsin gene expression evolution in the adaptive radiation of cichlid fishes of Lake Tanganyika. *Sci Adv.* 2023:9(36):eadg6568. https://doi.org/10.1126/sciadv.adg6568.

Ricci V, Ronco F, Musilova Z, Salzburger W. Molecular evolution and depth-related adaptations of rhodopsin in the adaptive radiation of cichlid fishes in Lake Tanganyika. *Mol Ecol.* 2022:31(10):2882–2897. https://doi.org/10.1111/mec.16429.

Rohlf F. *Tpsdig, digitize landmarks and outlines, version 2.0.* New York: Department of Ecology and Evolution, State University of New York at Stony Brook; 2004.

Ronco F, Matschiner M, Böhne A, Boila A, Büscher HH, El Taher A, Indermaur A, Malinsky M, Ricci V, Kahmen A, et al. Drivers and dynamics of a massive adaptive radiation in cichlid fishes. *Nature.* 2021:589(7840):7681. https://doi.org/10.1038/s41586-020-2930-4.

Rummer JL, Brauner CJ. Plasma-accessible carbonic anhydrase at the tissue of a teleost fish may greatly enhance oxygen delivery: in vitro evidence in rainbow trout, *Oncorhynchus mykiss. J Exp Biol.* 2011:214(14):2319–2328. https://doi.org/10.1242/jeb.054049.

Sabeti PC, Reich DE, Higgins JM, Levine HZP, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ, et al. Detecting recent positive selection in the human genome from haplotype structure. *Nature.* 2002:419(6909):832–837. https://doi.org/10.1038/nature01140.

Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, XieX, Byrne EH, McCarroll SA, Gaudet R, et al. Genome-wide detection and characterization of positive selection in human populations. *Nature.* 2007:449(7164):913–918. https://doi.org/10.1038/nature06250.

Santos ME, Lopes JF, Kratochwil CF. East African cichlid fishes. *EvoDevo.* 2023:14(1):1. https://doi.org/10.1186/s13227-022-00205-5.

Seehausen O, Terai Y, Magalhaes IS, Carleton KL, Mrosso HDJ, Miyagi R, Van Der Sluijs I, Schneider MV, Maan ME, Tachida H, et al. Speciation through sensory drive in cichlid fish. *Nature.* 2008:455(7213):620–626. https://doi.org/10.1038/nature07285.

Solaki M, Baumann B, Reuter P, Andreasson S, Audo I, Ayuso C, Balousha G, Benedicenti F, Birch D, Bitoun P, et al. Comprehensive variant spectrum of the CNGA3 gene in patients affected by achromatopsia. *Hum Mutat.* 2022:43(7):832–858. https://doi.org/10.1002/humu.24371.

Stauffer JR, Phiri TB, Konings AF. Description of two deep-water fishes of the genus Diplotaxodon (Teleostei: Cichlidae) from Lake Malaŵi, Africa. *Proc. Biol. Soc. Wash.* 2018:131(1):90–100. https://doi.org/10.2988/PBSW-D-17-00004.

Stearns G, Evangelista M, Fadool JM, Brockerhoff SE. A mutation in the cone-specific pde6 gene causes rapid cone photoreceptor

degeneration in zebrafish. *J Neurosci.* 2007:27(50):13866–13874. https://doi.org/10.1523/JNEUROSCI.3136-07.2007.

Stelzer G, Rosen N, Plaschkes I, Zimmerman S, Twik M, Fishilevich S, Iny Stein T, Nudel R, Lieder I, Mazor Y, et al. The GeneCards suite: from gene data mining to disease gene sequence analyses. *Curr Protoc Bioinform.* 2016:54(1):1–30. https://doi.org/10.1002/cpbi.5.

Sugawara T, Terai Y, Imai H, Turner GF, Koblmüller S, Sturmbauer C, Shichida Y, Okada N. Parallelism of amino acid changes at the RH1 affecting spectral sensitivity among deep-water cichlids from Lakes Tanganyika and Malawi. *Proc Natl Acad Sci U S A.* 2005:102(15):5448–5453. https://doi.org/10.1073/pnas.0405302102.

Svanbäck R, Johansson F. Predation selects for smaller eye size in a vertebrate: effects of environmental conditions and sex. *Proc R Soc B Biol Sci.* 2019:286(1897):20182625. https://doi.org/10.1098/rspb.2018.2625.

Svardal H, Quah FX, Malinsky M, Ngatunga BP, Miska EA, Salzburger W, Genner MJ, Turner GF, Durbin R. Ancestral hybridisation facilitated species diversification in the Lake Malawi cichlid fish adaptive radiation. *Mol Biol Evol.* 2020:37(4):1100–1113. https://doi.org/10.1093/molbev/msz294.

Sylvester JB, Rich CA, Loh Y-HE, van Staaden MJ, Fraser GJ, Streelman JT. Brain diversity evolves via differences in patterning. *Proc Natl Acad Sci U S A.* 2010:107(21):9718–9723. https://doi.org/10.1073/pnas.1000395107.

Takuno S, Miyagi R, Onami J, Takahashi-Kariyazono S, Sato A, Tichy H, Nikaido M, Aibara M, Mizoiri S, Mrosso HDJ, et al. Patterns of genomic differentiation between two Lake Victoria cichlid species, *Haplochromis pyrrhocephalus* and *H. sp. 'macula'. BMC Evol Biol.* 2019:19(1):68–17. https://doi.org/10.1186/s12862-019-1387-2.

Terai Y, Seehausen O, Sasaki T, Takahashi K, Mizoiri S. Divergent selection on opsins drives incipient speciation in Lake Victoria cichlids. *PLoS Biol.* 2006:4(12):e433. https://doi.org/10.1371/journal.pbio.0040433.

Theis A, Ronco F, Indermaur A, Salzburger W, Egger B. Adaptive divergence between lake and stream populations of an East African cichlid fish. *Mol Ecol.* 2014:23(21):5304–5322. https://doi.org/10.1111/mec.12939.

The UniProt Consortium. UniProt: the universal protein knowledgebase in 2023. *Nucleic Acids Res.* 2023:51(D1):D523–D531. https://doi.org/10.1093/nar/gkac1052.

Thompson AB, Allison EH, Ngatunga BP. Distribution and breeding biology of offshore cichlids in Lake Malawi/Niassa. *Environ Biol Fishes.* 1996:47(3):235–254. https://doi.org/10.1007/BF00000496.

Toomey MB, Lind O, Frederiksen R, Curley RW, Riedl KM, Wilby D, Schwartz SJ, Witt CC, Harrison EH, Roberts NW, et al. Complementary shifts in photoreceptor spectral tuning unlock the full adaptive potential of ultraviolet vision in birds. *eLife.* 2016:5:e15675. https://doi.org/10.7554/eLife.15675.

Turner GF, Robinson RL, Shaw PW, Carvalho GR, Snoeks J. Identification and biology of Diplotaxodon, Rhamphochromis and Pallidochromis. In: *The cichlid diversity of Lake Malawi/Nyasa/Niassa: identification.* El Paso (TX): Cichlid Press; 2004. p. 198–251.

Vernaz G, Hudson AG, Santos ME, Fischer B, Carruthers M, Shechonge AH, Gabagambi NP, Tyers AM, Ngatunga BP, Malinsky M, et al. Epigenetic divergence during early stages of speciation in an African crater lake cichlid fish. *Nat Ecol Evol.* 2022:6:1940–1951. https://doi.org/10.1038/s41559-022-01894-w.

Voight BF, Kudaravalli S, Wen X, Pritchard JK. A map of recent positive selection in the human genome. *PLOS Biol.* 2006:4(3):e72. https://doi.org/10.1371/journal.pbio.0040072.

Volkov LI, Kim-Han JS, Saunders LM, Poria D, Hughes AEO, Kefalov VJ, Parichy DM, Corbo JC. Thyroid hormone receptors mediate two distinct mechanisms of long-wavelength vision. *Proc Natl Acad Sci U S A.* 2020:117(26):15262–15269. https://doi.org/10.1073/pnas.1920086117.

Warrant EJ, Johnsen S. Vision and the light environment. *Curr Biol.* 2013:23(22):R990–R994. https://doi.org/10.1016/j.cub.2013.10.019.

Warrant EJ, Locket NA. Vision in the deep sea. *Biol Rev Camb Philos Soc*. 2004:79(3):671–712. https://doi.org/10.1017/S1464793103006420.

Wen X, Liu Y, Yan Q, Liang M, Tang M, Liu R, Pan J, Liu Q, Chen T, Guo S, *et al*. Association of IGFN1 variant with polypoidal choroidal vasculopathy. *J Gene Med*. 2018:20(2-3):e3007. https://doi.org/10.1002/jgm.3007.

Wray GA. The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet*. 2007:8(3):206–216. https://doi.org/10.1038/nrg2063.

Xu D, Yang C, Shen Q, Pan S, Liu Z, Tongzuo Z, Zhou X, Lei M, Chen P, Yang H, *et al*. A single mutation underlying phenotypic convergence for hypoxia adaptation on the Qinghai-Tibetan Plateau. *Cell Res*. 2021:31(9):1032–1035. https://doi.org/10.1038/s41422-021-00517-6.

Yau K-W, Hardie RC. Phototransduction motifs and variations. *Cell*. 2009:139(2):246–264. https://doi.org/10.1016/j.cell.2009.09.029.

Yokoyama S. Evolution of dim-light and color vision pigments. *Annu Rev Genomics Hum Genet*. 2008:9(1):259–282. https://doi.org/10.1146/annurev.genom.9.081307.164228.

Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet*. 2012:44(7):821–824. https://doi.org/10.1038/ng.2310.

Zhu A, Ibrahim JG, Love MI. Heavy-tailed prior distributions for sequence count data: removing the noise and preserving large differences. *Bioinformatics*. 2019:35(12):2084–2092. doi:10.1093/bioinformatics/bty895.