scientific data

Check for updates

OPEN

DATA DESCRIPTOR

A single-nucleus RNA sequencing atlas of the postnatal retina of the shark *Scyliorhinus canicula*

Nicolás Vidal-Vázquez[1,2,9], Ismael Hernández-Núñez[1,9], Pablo Carballo-Pacoret[1,3], Sarah Salisbury[3], Paula R. Villamayor[3,4], Francisca Hervas-Sotomayor[5,8], Xuefei Yuan[5], Francesco Lamanna[5], Céline Schneider[5], Julia Schmidt[5], Sylvie Mazan[6], Henrik Kaessmann[5], Fátima Adrio[1], Diego Robledo[3,7], Antón Barreiro-Iglesias[1,2] & Eva Candal[1,2] ✉

The retina, whose basic cellular structure is highly conserved across vertebrates, constitutes an accessible system for studying the central nervous system. In recent years, single-cell RNA sequencing studies have uncovered cellular diversity in the retina of a variety of species, providing new insights on retinal evolution and development. However, similar data in cartilaginous fishes, the sister group to all other extant jawed vertebrates, are still lacking. Here, we present a single-nucleus RNA sequencing atlas of the postnatal retina of the catshark *Scyliorhinus canicula*, consisting of the expression profiles for 17,438 individual cells from three female, juvenile catshark specimens. Unsupervised clustering revealed 22 distinct cell types comprising all major retinal cell classes, as well as retinal progenitor cells (whose presence reflects the persistence of proliferative activity in postnatal stages in sharks) and oligodendrocytes. Thus, our dataset serves as a foundation for further studies on the development and function of the catshark retina. Moreover, integration of our atlas with data from other species will allow for a better understanding of vertebrate retinal evolution.

Background & Summary

The neural retina shows a remarkable degree of conservation in its cellular structure in all extant vertebrates[1–3]. This basic plan consists of five neuronal cell classes and one glial cell class, whose somata are arrayed in three nuclear layers, interspersed with two plexiform layers, where synapses occur (Fig. 1a). The outer nuclear layer (ONL) contains photoreceptors (PRs), the light-sensitive cells of the retina, which can usually be classified in two morphologically and functionally distinct major types, cones and rods, responsible for photopic and scotopic vision, respectively. In turn, the inner nuclear layer (INL) hosts three types of interneurons, namely horizontal cells (HCs), bipolar cells (BCs) and amacrine cells (ACs), which receive, integrate, modulate and transmit the signals coming from photoreceptors to retinal ganglion cells (RGCs). RGCs are located in the ganglion cell layer (GCL) and project their axons through the optic nerve to the visual processing centres of the brain. Besides, the retina contains a major glial cell class, Müller glia (MG), a type of radial glial cells whose nuclei are located within the INL. Other glial cell types, such as oligodendrocytes, microglia or astrocytes, may be present in the innermost layers of the retina, but, unlike MG, these are not derived from the optic cup; instead, they originate in other parts of the brain and migrate into the eye through the optic nerve[4].

[1]Departamento de Bioloxía Funcional, Facultade de Bioloxía, Universidade de Santiago de Compostela, 15782, Santiago de Compostela, Spain. [2]Aquatic One Health Research Center (ARCUS), Universidade de Santiago de Compostela, 15782, Santiago de Compostela, Spain. [3]The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Edinburgh, EH25 9RG, UK. [4]Departamento de Zooloxía, Xenética e Antropoloxía Física, Facultade de Veterinaria, Universidade de Santiago de Compostela, 27002, Lugo, Spain. [5]Center for Molecular Biology (ZMBH), DKFZ-ZMBH Alliance, Heidelberg University, Heidelberg, Germany. [6]CNRS, Sorbonne Université, Biologie Intégrative des Organismes Marins, UMR7232-BIOM, Banyuls-sur-Mer, France. [7]Departamento de Zooloxía, Xenética e Antropoloxía Física, CIBUS, Facultade de Bioloxía, Universidade de Santiago de Compostela, 15782, Santiago de Compostela, Spain. [8]Present address: INRAE, LPGP, Rennes, France. [9]These authors contributed equally: Nicolás Vidal-Vázquez, Ismael Hernández-Núñez. ✉e-mail: eva.candal@usc.es

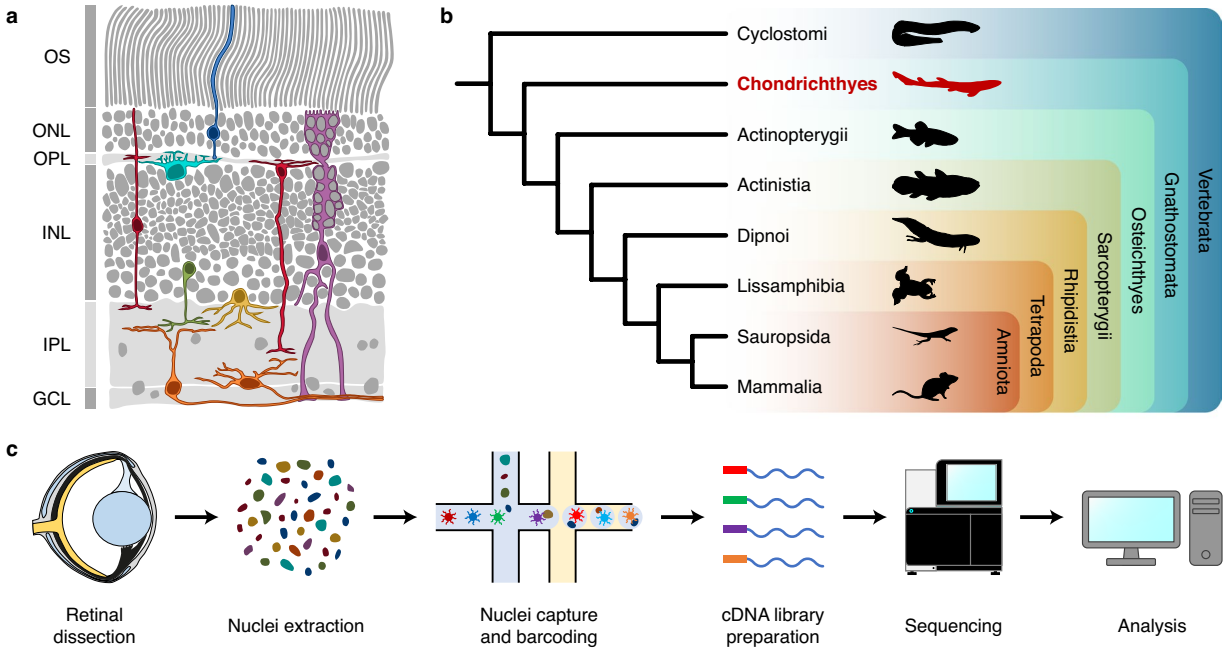

Fig. 1 Background and experimental workflow. (**a**) Structure of the juvenile catshark retina, based on microscopy images kindly provided by Dr J. Francisco-Morcillo; cell drawings based on Neumayer[75]. (**b**) Cladogram showing relationships among major extant vertebrate groups, highlighting the position of chondrichthyans as the sister group to all other gnathostomes. Silhouettes from PhyloPic[95]. (**c**) Experimental workflow. Eye diagram after Collin[45]. GCL, ganglion cell layer; INL, inner nuclear layer; IPL, inner plexiform layer; ONL, outer nuclear layer; OPL, outer plexiform layer; OS, photoreceptor outer segments.

Most of these cell classes can be subdivided into a number of cell types with diverse morphological, functional and molecular properties[5–7]. In recent years, high-throughput single-cell RNA sequencing (scRNA-seq) studies have provided unprecedented insights into molecular cell type diversity in the mature and developing vertebrate retina[6,8–10]. Although most studies have focused on the retina of mammals[11–35] (particularly that of rodents[12–18] and primates[19–35]), cellular atlases of the chick[13,36], brown anole lizard[11], zebrafish[9,13,37–43] and sea lamprey[11,44] retinas have also been generated to date, providing new resources to study the evolution and development of vertebrate retinal cell types. However, scRNA-seq studies on the retina of cartilaginous fishes are still lacking.

Given their phylogenetic position as the sister group to all other extant gnathostomes (Fig. 1b), chondrichthyans (sharks, rays and chimaeras) constitute a particularly interesting group to study retinal evolution[45–47]. Among cartilaginous fishes, the shark *Scyliorhinus canicula* (Linnaeus, 1758), also known as the small-spotted catshark or the lesser-spotted dogfish, stands as a suitable model for experimental studies, owing to its abundance, relatively small size and accessible oviparous development[46]. Furthermore, the catshark retina shows persistent cell proliferation in postnatal stages[48,49] (similar to other fishes[50]), making *S. canicula* an interesting species for the study of postnatal retinal neurogenesis.

In this study, we generated a single-nucleus RNA sequencing (snRNA-seq) atlas of the postnatal retina of *S. canicula* (Fig. 1c). Early juvenile specimens were chosen to allow profiling of both mature and progenitor cell types, since previous work has shown that cell proliferation decreases as the animal grows, with mitotic activity being virtually absent in sexually mature adults[48,51]. Thus, we used three retinas from three female, juvenile catsharks to generate a dataset consisting of 17,438 high quality nuclei. Unsupervised clustering revealed 22 cell types representing all major retinal cell classes, as well as retinal progenitor cells and oligodendrocytes. This constitutes, to the best of our knowledge, the first single-cell transcriptomic study of the retina of a chondrichthyan, providing a groundwork for comparative studies on the evolution of both retinal cell type diversity and retinal neurogenesis.

Methods

Animals. Three female, juvenile specimens of *S. canicula*, with a total length of 10.5 to 11.1 cm (Table 1) were kindly provided by the Aquarium Finisterrae in A Coruña (Spain) and kept in artificial seawater tanks under standard conditions of temperature (15-16 °C), pH (7.5–8.5) and salinity (35 g/L). All procedures were performed in accordance with the guidelines for animal experimentation established by the European Union and the Spanish government and were approved by the Bioethics Committee of the University of Santiago de Compostela (license number 15004/2022/001).

Retina sampling. Animals were deeply anaesthetised with 0.5% tricaine methanesulfonate (MS-222; Thermo Fisher Scientific, 118000500) in seawater. The animals were removed from water, the eyes enucleated,

Sample	Total length (cm)	Sex
Sample 1	11.0	Female
Sample 2	10.5	Female
Sample 3	11.1	Female

Table 1. Sample information.

and the retinas (one from each specimen, n = 3) were dissected out under a stereomicroscope. Retinas were immediately put in Eppendorf tubes and frozen in liquid nitrogen. Retinal samples were kept at -80 °C until they were further processed for nuclei isolation and RNA sequencing.

Nuclei isolation and snRNA-seq data generation. Retina nuclei were extracted following a published protocol[52] with small modifications for use in non-standard vertebrate species[53,54]. Briefly, the frozen retinas were homogenised using a micropestle in 400 µL ice-cold homogenisation buffer (250 mM sucrose, 25 mM KCl, 5 mM $MgCl_2$, 10 mM Tris-HCl [pH 8], 0.1% IGEPAL, 1 µM dithiothreitol [DTT], 0.4 U/µL Murine RNase Inhibitor [New England BioLabs, M0314L], and 0.2 U/µL SUPERase-In [Ambion, AM2694]). The homogenates were triturated gently using a p1000 tip for 10 times, incubated on ice for 5 min and then centrifuged at 100 g for 1 min at 4 °C to pellet any unlysed tissue chunks. The supernatant was transferred into another 1.5 mL Eppendorf tube and centrifuged at 400 g for 4 min at 4 °C to collect nuclei. The nuclei were washed twice in 400 µL homogenisation buffer and strained using a 40 µm Flowmi strainer (Sigma, BAH136800040) during the second wash step to remove nuclei aggregates. The final nuclei pellet was resuspended in 30–50 µL Nuclei Buffer (10x Genomics, PN-2000207). To estimate the nuclei concentration, nuclei aliquots were diluted in phosphate-buffered-saline (PBS) with Hoechst and propidium iodide (PI) DNA dyes and counted on Countless II FL Automated Cell Counter (Thermo Fisher Scientific, RRID: SCR_020236). Around 15,000 nuclei were used as input for the snRNA-seq experiment. The Chromium Next GEM Single Cell 3' Reagent Kits v3.1 (PN-1000121, PN-1000120, and PN-1000213) were used to make snRNA-seq libraries. Libraries were quantified on a Qubit Fluorometer (Thermo Fisher Scientific; RRID: SCR_018095) and quality checked on a Fragment Analyzer (Agilent; RRID: SCR_019417). Libraries were sequenced on NextSeq550 (Illumina; RRID: SCR_016381; 28 cycles for Read 1, 56 cycles for Read 2, 8 cycles for i7 index).

Genome indexing and read alignment. Genome indexing and library mapping was performed with STAR (v2.7.10a)[55,56]. The *S. canicula* genome assembly sScyCan 1.1[57] (GCF_902713615.1; *GCF_902713615.1_sScyCan1.1_genomic.fna*) and its associated annotation in GFF format (*GCF_902713615.1_sScyCan1.1_genomic.gff*) were downloaded from the National Center for Biotechnology Information (NCBI). The mitochondria (NC_001950.1) annotations within the GFF file were manually edited to convert "CDS" annotations to "exon" annotations and to convert all annotations of "tRNA" and "rRNA" to "gene" annotation. This was done to ensure mitochondrial gene annotations were consistent with those of the nuclear genes. The GFF annotation file was converted to GTF format using gffread (v0.10.1)[58]. The genome and its annotation (GTF) were indexed using STAR (--runMode genomeGenerate). Each library was then mapped against the genome with the 10x V3 cell barcode whitelist (*3M-february-2018.txt*) and using standard parameters for single cell libraries (--soloMultiMappers Unique --soloBarcodeReadLength 28 --soloType CB_UMI_Simple --soloUMIlen 12 --soloCBwhitelist 3M-february-2018.txt --soloFeatures GeneFull --clipAdapterType CellRanger4 --outFilterScoreMin 20 --soloCBmatchWLtype 1MM_multi_Nbase_pseudocounts --soloUMIfiltering MultiGeneUMI_CR --soloUMIdedup 1MM_CR --readFilesCommand zcat --outSAMtype BAM Unsorted). The raw (unfiltered) files (*genes.tsv*, *barcodes.tsv* and *matrix.mtx*) generated for each sample were then used for downstream analysis. On average, there were 274 million reads per sample with 94% of reads with valid barcodes, and a 45% saturation. A summary of the STAR output statistics for each sample can be found in Table 2.

Bioinformatic quality control. Samples were then analysed in an R (v4) environment using Seurat (v4.3.0)[59]. We created Seurat objects for each library after removing nuclei with less than 200 features and features occurring in fewer than three nuclei. Nuclei where mitochondrial (mtDNA) features accounted for 10% or more of their total unique molecular identifiers (UMIs) were removed before removing all mtDNA features. After sub-setting the Seurat object into individual samples, upper and lower thresholds for UMI and feature counts per nuclei were then applied individually to each sample based on knee plot visualisation. In sample 1, nuclei with more than 1,000 but less than 20,000 UMIs and more than 500 but less than 5,500 features were retained. In sample 2, nuclei with more than 500 but less than 24,000 UMIs and more than 500 but less than 6,500 features were retained. Finally, in sample 3, nuclei with more than 750 but less than 20,000 UMIs and more than 300 but less than 7,000 features were retained.

Samples were then merged into a single Seurat object before splitting samples again into individual sample datasets. This was done to ensure that the same features were considered across samples. Each sample was then independently clustered and assessed for the presence of doublets. For each sample, counts were normalised using the "NormalizeData" function prior to calculating cell cycle scores using the "CellCycleScoring" function (for the list of genes used, see Data Records). The "v2" SCTransform version with the glmGamPoi method (v1.9.0)[60] was used to normalise RNA counts separately for each sample, regressing out scores for the S and G2M cell cycle stages. Linear dimension reduction was conducted independently for each sample using the "RunPCA" function with 50 principal components (PCs). After consulting Elbowplots for each sample, a Uniform Manifold Approximation and Projection (UMAP) using 20 PCs was run separately for each sample and

Sample	Sample 1	Sample 2	Sample 3
Number of reads	229,861,905	249,344,842	343,563,083
Reads with valid barcodes	0.960049	0.932741	0.917902
Sequencing saturation	0.502799	0.521001	0.338812
Q30 bases in CB + UMI	0.957978	0.95938	0.961753
Q30 bases in RNA read	0.89873	0.90096	0.888598
Reads mapped to genome: unique + multiple	0.88851	0.846251	0.647372
Reads mapped to genome: unique	0.781012	0.711432	0.518868
Reads mapped to GeneFull: unique + multiple GeneFull	0.693977	0.471452	0.354986
Reads mapped to GeneFull: unique GeneFull	0.642748	0.424422	0.314171
Estimated number of cells	3,028	1,074	2,521
Unique reads in cells mapped to GeneFull	62,372,962	34,713,678	39,214,416
Fraction of unique reads in cells	0.422171	0.328021	0.363307
Mean reads per cell	20,598	32,321	15,555
Median reads per cell	15,254	21,413	11,331
UMIs in cells	30,760,018	16,433,368	25,334,077
Mean UMIs per cell	10,158	15,301	10,049
Median UMIs per cell	7,593	10,144	7,340
Mean GeneFull per cell	3,508	4,872	3,988
Median GeneFull per cell	3,170	4,285	3,540
Total GeneFull Detected	23,271	23,076	23,449

Table 2. Summary statistics for STAR outputs for each sample. Note that raw files were used for downstream analysis (which includes all barcodes, not just those estimated to be viable cells by STAR). Therefore, the statistics relating to STAR-assigned cells are not necessarily relevant but are provided for the reader's interest.

Group	Abbreviation	Clusters
Rod photoreceptors	PR_Rod	PR_Rod1, PR_Rod2, PR_Rod3
Photoreceptor precursors	PR_Precursor	PR_Precursor
Horizontal cells	HC	HC
Bipolar cells	BC	BC1, BC2, BC3
GABAergic amacrine cells	AC_GABA	AC_GABA1, AC_GABA2, AC_GABA3, AC_GABA4
Cholinergic amacrine cells	AC_ACh	AC_ACh
Glycinergic amacrine cells	AC_Gly	AC_Gly1, AC_Gly2, AC_Gly3
Retinal ganglion cells	RGC	RGC1, RGC2
Müller glia	MG	MG
Retinal progenitor cells	RPC	RPC1, RPC2
Oligodendrocytes	OL	OL

Table 3. Class/subclass-level cluster groups for differential expression analysis.

the "FindNeighbours" function was applied using 20 PCs, before using the "FindClusters" function with a resolution of 0.5. DoubletFinder (v2.0.3)[61] was then applied independently to each sample selecting pK values with the highest associated mean-variance normalised bimodality coefficient (BCmvn) value. We assumed a 4% doublet formation rate (based on the Chromium instrument specifications) and adjusted for homotypic doublets.

Clustering and differential gene expression analyses. Samples were integrated using 2,000 features and anchors that were identified with the "rpca" reduction method and the "FindIntegrationAnchors" function. A principal component analysis (PCA) was rerun on the integrated dataset using 50 PCs, and 30 PCs were used for subsequent UMAP generation and clustering with a resolution of 0.5. Markers for each cluster were assessed using the logistic regression method and the FindAllMarkers function on the "SCT" assay and "data" slot, using sample ID as a latent variable to help reduce batch effects among samples. We used a pseudocount of 0.001, set a p-value threshold of 0.01, and only considered genes that were upregulated, expressed in at least 25% of all nuclei (in either of the compared groups), and demonstrated a threshold of 0.25 X difference (log-scale) between the two compared groups. Three clusters that were mostly composed of cells from a single sample and did not show differential expression of typical retinal cell marker genes were removed (see below; for a list of differentially expressed genes in the removed clusters, see Data Records), and all remaining cells were re-clustered using the same parameters. The same differential expression analysis was applied to the finalised clusters grouped together based on common diagnostic marker genes (see below for details) to identify marker genes for retinal cell classes composed of several clusters (see Table 3).

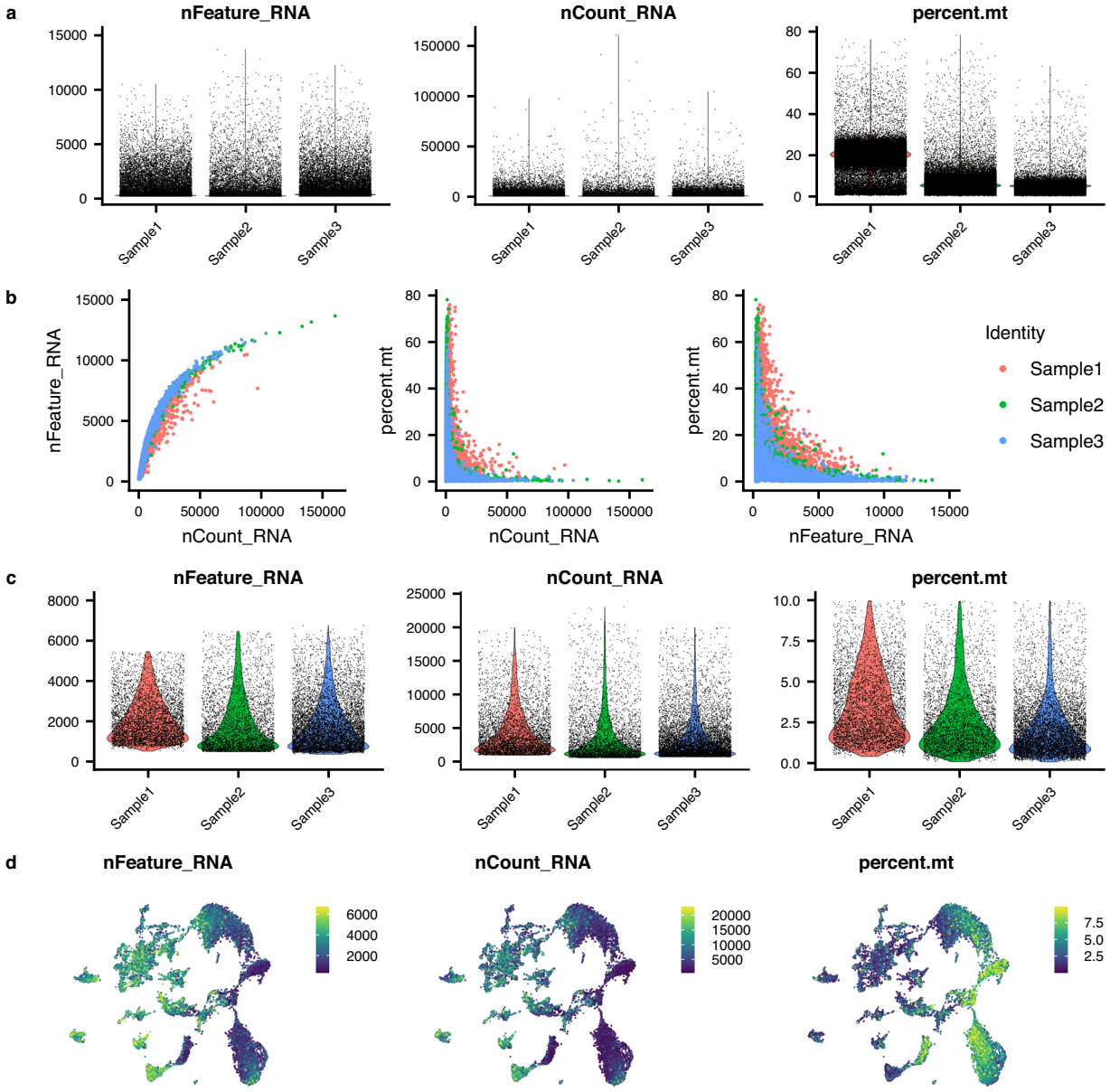

Fig. 2 Quality control for snRNA-seq data. (**a**) Violin plots showing: left, the number of detected unique genes ("nFeature_RNA"); middle, the number of unique molecular identifiers (UMIs, "nCount_RNA"); and right, percentage of mitochondrial genes ("percent.mt") in each cell from each sample before quality control. (**b**) Scatter plots showing the relationship between UMI counts and gene counts (left), UMI counts and percentage of mitochondrial genes (middle), together with gene counts and percentage of mitochondrial genes (right) in the three samples before quality control. (**c**) Violin plots showing gene counts (left), UMI counts (middle) and percentage of mitochondrial genes (right) in the final dataset. Note that all mitochondrial features were removed from the analyses after quality control (see Methods). (**d**) Gene counts (left), UMI counts (middle) and percentage of mitochondrial genes (right) mapped onto the UMAP.

Cluster annotation. To assign a cell class identity to each cluster, we elaborated a list of marker genes reported in previous studies on other vertebrate species[11–15,18–21,24,26–31,33,36,38,44]. Putative orthologues in the *S. canicula* genome were identified by comparing the sequence of the mouse gene (genome assembly GRCm39, GCF_000001635.27)[62] with the *S. canicula* genome assembly sScyCan1.1 (GCF_902713615.1)[57] using Blastn[63]. For marker genes absent from the mouse genome, either the human (genome assembly GRCh38.p14, GCF_000001405.40)[64] or zebrafish (genome assembly GRCz11, GCF_000002035.6)[65] genes were used. All marker gene sequences were obtained from GenBank[66]. Marker genes present in the *S. canicula* genome where then visualised in our dataset using the "FeaturePlot" and "DotPlot" functions of Seurat. This, coupled with the examination of the top differentially expressed genes in each cluster (see above), allowed us to assign all clusters in the final object to known retinal cell classes (see Technical Validation). In cell classes comprising more than a single cluster, a number was added to the cluster name, according to the number of cells in each cluster.

Sample	Sample 1	Sample 2	Sample 3	Total
Unfiltered barcodes	69,723	52,932	64,138	186,793
Filtered nuclei	4,869	4,130	8,439	17,438
Filtered features	22,472	22,458	23,014	23,489
Mean UMIs per cell	4,527	3,716	3,486	3,831
Median UMIs per cell	3,333	2,234	2,359	2,572
Mean features per cell	2,153	1,885	1,852	1,944
Median features per cell	1,892	1,434	1,502	1,604
Mean percent of mitochondrial features	3.55	2.66	2.13	2.65
Median percent of mitochondrial features	3.06	2.05	1.64	2.05

Table 4. Final cell quantification and sequencing statistics. "Unfiltered barcodes" refers to the number of barcodes after importing the STAR files to Seurat, excluding those with less than 200 UMIs. All other parameters refer to the final atlas.

Gene nomenclature. For the sake of readability, we decided to change the symbols of genes annotated with a LOC number in the main text and figures of the present article. We replaced LOC numbers for symbols based on the identification of their protein products, adding an "-*l*" at the end for those annotated as "-like", except for uncharacterised genes, which were left with their LOC number. Supplementary Table 1 shows the correspondence between the symbols used in the text and the sScyCan1.1 annotation, which is the one used in the dataset.

It should also be noted that many *S. canicula* genes follow the nomenclature used in zebrafish, including a letter or number at the end of their symbol. Since teleosts experienced an additional whole genome duplication after diverging from cartilaginous fishes[67], many gene paralogues present in zebrafish are absent from the catshark genome. Therefore, current gene symbols do not accurately reflect the existence of paralogues in *S. canicula*, nor their correspondence to specific zebrafish paralogues.

Data Records

Raw sequencing data for each sample (*.fastq* files) have been uploaded to the NCBI Sequence Read Archive (SRA)[68] under BioProject accession number PRJNA1056918[69]. For each sample, STAR raw output files, i.e. expression matrices for each gene in each cell (*matrix.mtx*), barcodes (*barcodes.tsv*) and genes (*features.tsv*), have been uploaded to the Gene Expression Omnibus (GEO)[70,71] under accession number GSE282457[72]. The final annotated Seurat object (*Scyliorhinus_retina.rds*) is available on Figshare[73].

Two *.xlsx* files (*Differential_expression_Clusters.xlsx* and *Differential_expression_Cell_classes.xlsx*) containing the list of differentially expressed genes in each cluster and each cell class (see Table 3) in the final atlas have been uploaded to Figshare[73]. For each gene, the p-value ("p_val"), adjusted p-value ("p_val_adj"), average $\log_2$ fold change ("avg_log2FC"), the proportion of cells expressing that gene in the present cluster ("pct.1"), the proportion of cells expressing said gene in the rest of the dataset ("pct.2"), the gene symbol and the full name of its protein product are provided. Please note that protein product names are provided for the sake of readability, and do not accurately reflect the isoform of each feature. An additional file (*Differential_expression_Removed_clusters.xlsx*) provides the same information for the two clusters removed during clustering (see Methods). Furthermore, a table containing the list of genes used in cell cycle scoring (*Cell_cycle_scoring_genes.xlsx*) can also be found on Figshare[73].

Technical Validation

Quality control. Retinas from three similarly sized, female, juvenile *Scyliorhinus canicula* specimens (Table 1) were used to generate the snRNA-seq dataset. To ensure that all the obtained barcodes correspond to viable nuclei, we established selection criteria based on feature counts, UMI counts and expression of mitochondrial genes (Fig. 2). Prior to filtering, the three samples showed a similar number of barcodes (Table 4), features and UMIs, with the proportion of mitochondrial transcripts being higher in sample 1 and lower in sample 3 (Fig. 2a). As expected, there was a positive correlation between the number of detected genes and the number of UMIs, whereas a negative correlation was observed between the expression levels of mitochondrial genes and both gene counts and UMI counts, with no major differences between samples (Fig. 2b). Quality control filtering was performed by removing barcodes with mitochondrial transcripts representing more than 10% of the total counts, and upper and lower thresholds for UMIs and number of unique genes were defined for each sample (see Methods). After filtering, initial clustering analyses revealed three clusters that were mostly composed of cells from a single sample that did not show expression of typical retinal cell marker genes (see Data Records). Since these nuclei most likely correspond to either low quality cells or ambient RNA, they were removed from the dataset. The final object was composed of 17,438 cells containing 23,489 total features (unique genes), with an average of 1,944 features per cell (Table 4). There were no major differences in the number of unique genes, UMI count (number of transcripts) or mitochondrial feature levels between samples (Fig. 2c,d). However, sample 3 yielded a higher number of cells (Table 4), which can be attributed to minor technical differences in retinal dissection and/or library preparation.

Cluster annotation. Unsupervised clustering of filtered cells revealed 22 clusters (Fig. 3a), all of which were present in generally similar proportions in the three samples (Fig. 3b), with cells from the three retinas showing similar distributions across the UMAP (Fig. 3c), thus confirming the similarity of the samples. To annotate the

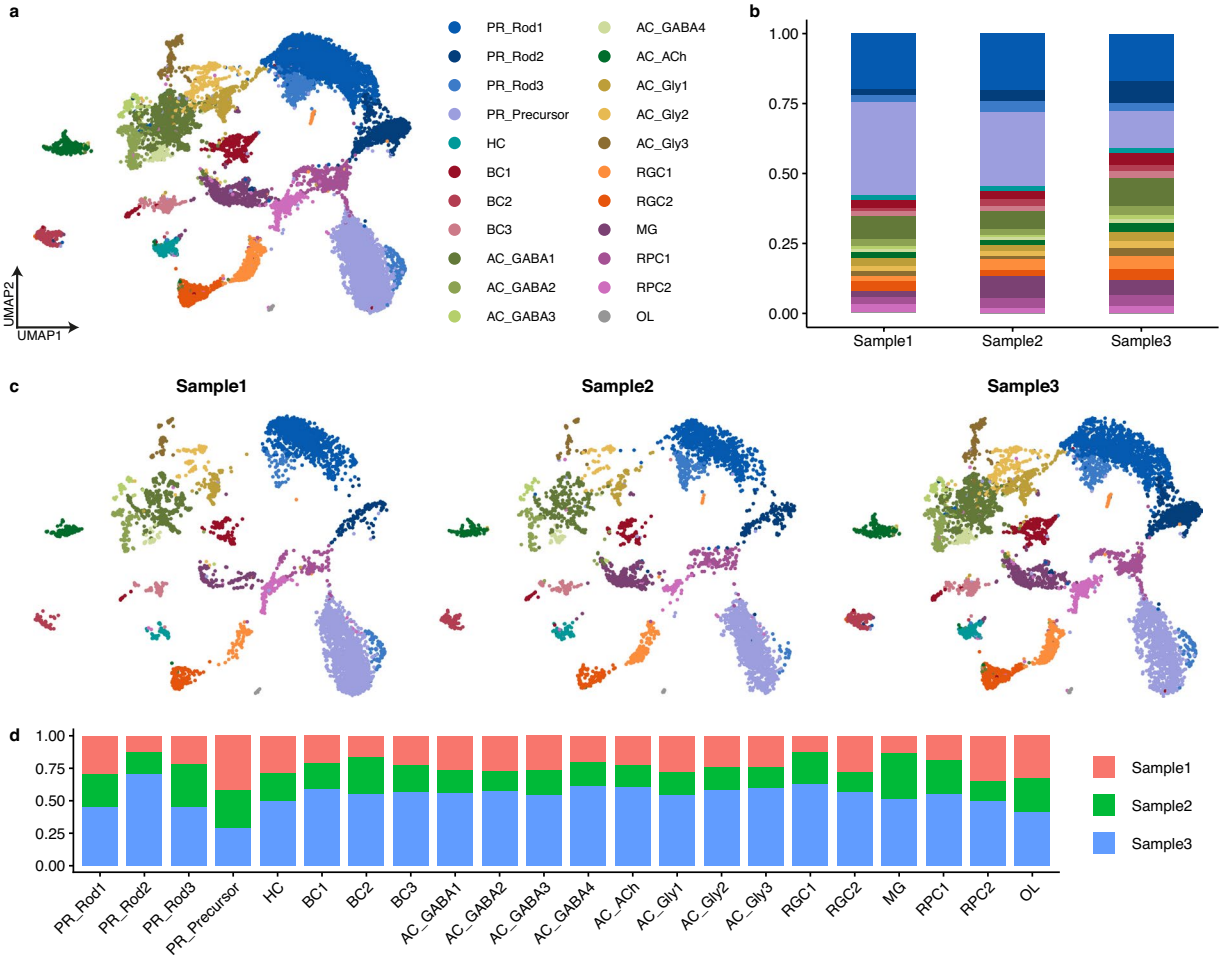

Fig. 3 Clustering of snRNA-seq data reveals cell heterogeneity in the juvenile catshark retina. (**a**) UMAP of retinal cells showing their unbiased assignment to the 22 clusters identified in this study. For abbreviations, see Table 3. (**b**) Barplots showing the proportions of each cluster for each sample. (**c**) Separate UMAPs of cells from each sample, coloured according to their assigned identities. (**d**) Barplots showing the proportions of cells from the three samples in each cluster.

clusters, we analysed the expression levels of known marker genes from other species (see Methods) for the major retinal cell classes or subclasses (Fig. 4a,b). This allowed us to divide cells into the six major retinal cell classes (PRs, HCs, BCs, ACs, RGCs and MG), as well as to identify two clusters of retinal progenitor cells (RPCs) and a single cluster of oligodendrocytes (OLs) (Fig. 3a).

Three clusters could be identified as rod PRs by the expression of genes such as *rho*, *pdc-l* or *pde6a*[74]. Expression of established cone PR marker genes (not shown) was either present in rod clusters or virtually non-existent. This is in accordance with previous observations that *S. canicula* has either scarce cones[75] or a pure-rod retina[49,76] (reviewed in Ferreiro-Galve *et al.*[49]), similar to some other benthic sharks[76,77] and the skate[78]. A fourth cluster that did not show high levels of most rod markers, but showed significant expression of *nr2e3* and *neurod1*, known markers of rod precursors/progenitors[79], was thus annotated as PR precursor cells. Since rod PRs usually comprise a single cell type[36,41,80] (with some exceptions, such as those of the amphibian retina[81]), the four PR clusters identified here could correspond to different states of PR differentiation (since active rod PR generation is known to persist in the postnatal catshark retina[49], with rods showing different states of maturation/differentiation, based on morphology of their outer segments), but further research is needed to confirm this.

Regarding retinal interneurons, a single cluster of HCs (expressing the canonical marker *onecut3a*[36,82]) and 3 clusters of BCs (expressing *vsx1* and *vsx2*[11,14,83], as well as high levels of *isl1* in one of the BC clusters, confirming previous immunohistochemical results[84]) could be identified in our dataset. We also found eight clusters of ACs, most of which showed high levels of *pax6*, confirming previous results[85]. The AC clusters could be further divided into three subclasses according to the expression of small molecule neurotransmitter-associated genes. Thus, we found four clusters of GABAergic ACs expressing *gad1* and/or *gad2*[18,86]; a single cluster of cholinergic (starburst) ACs, which show high levels of *chata* (similar to those described in *Squalus*[87]); and three clusters of glycinergic ACs, which show high levels of glycine transporters *slc6a5* and *slc6a9*[18]. Furthermore, two clusters of RGCs (expressing canonical markers, such as *nefl-l*[15]) could be recognised.

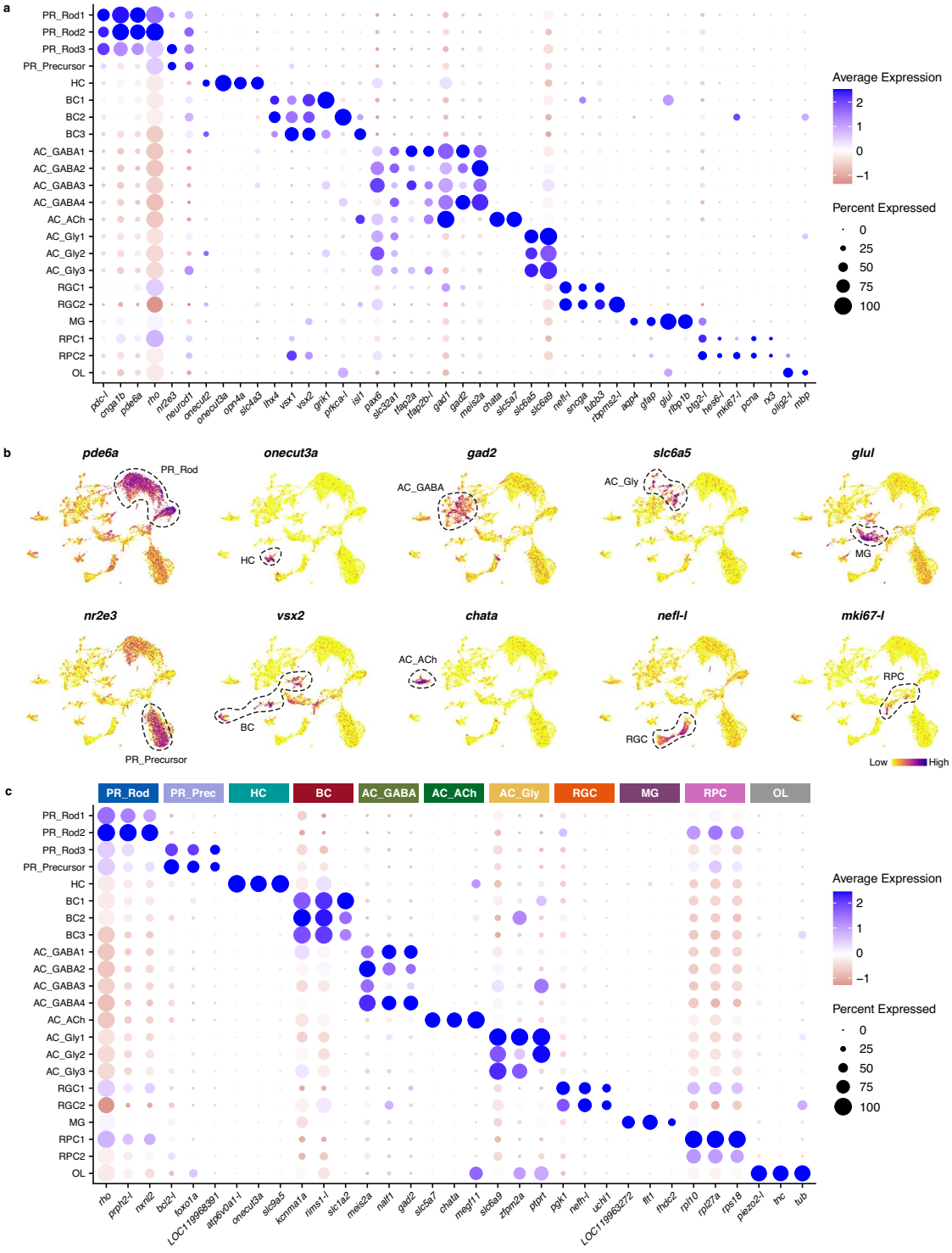

Fig. 4 Expression of marker genes used for cluster annotation. (**a**) Dot plot showing expression levels of marker genes selected for identification of retinal cell classes (cluster annotation). (**b**) UMAPs showing expression of selected canonical marker genes. (**c**) Dot plot showing expression of the top three differentially expressed genes from class-level groupings (see text and Table 3 for details). See Methods and Supplementary Table 1 for information on gene nomenclature.

MG were represented by a single cluster with high levels of *rlbp1b*[11,88] and *glul* (the gene encoding glutamine synthetase, whose expression in *S. canicula* MG has previously been described by means of immunohistochemistry[84,89]). Two clusters of RPCs could also be identified by the expression of cell proliferation markers (*mki67-l*[19,90], *pcna*[48,49,89]) and typical markers of retinal progenitors (*btg2-l*, *rx3*[12]). Finally, a small cluster of OLs was identified by high levels of *olig2-l* and *mbp* expression[13,36] corresponding to cells from the optic nerve head

or the optic nerve fibres layer, where myelinated fibres are present in other elasmobranchs[45]. A remarkable absence from our dataset are microglial cells, which are known to be present in the innermost retinal layers of the postnatal retina of *S. canicula*[91]. This is most likely due to microglia being lost in the nuclei dissociation or quality control filtering processes.

To further validate cluster annotation and find novel markers specific to the *S. canicula* retinal cell classes, we identified marker genes (differential gene expression analyses vs the remaining cells in the dataset), both for individual clusters and for clusters grouped into classes/subclasses (Table 3). Figure 4c shows the top marker genes for each class-level group, including both established and novel marker genes. For a complete list, see Data Records. Overall, the expression of marker genes confirms that our atlas comprises all major cell classes expected in the juvenile catshark retina, although some low abundant cell types (e.g. microglia) could be absent from the dataset. Future re-clustering of each cell class separately will allow further characterisation of cell type heterogeneity within the cell classes identified here. Thus, this dataset will provide a groundwork for studies on cell type diversity in the shark retina, allowing a better understanding of vertebrate retinal evolution and development.

Usage Notes

All the code used to analyse the dataset is available in GitHub (see Code Availability). Raw *.fastq* files are standard Illumina sequencing files for 10x Genomics single-cell RNA sequencing libraries, and as such they can be processed using any typical single-cell analysis software (STAR[55,56], as in our work, or Cell Ranger[92], Kallisto[93], Alevin[94], etc.). STAR output files for each sample can be loaded into Seurat using the "ReadSTARsolo" function, and then merged into a single Seurat object using the "merge" function. Finally, the *Scyliorhinus_retina.rds* file is a Seurat object with the processed dataset that can be loaded into R with "ReadRDS".

Code availability

The code used to process the raw sequencing files and generate all the results presented in this study can be found in https://github.com/Roslin-Aquaculture/SHARK_retina.

Received: 15 April 2024; Accepted: 28 January 2025;
Published online: 07 February 2025

References

1. Baden, T. Vertebrate vision: lessons from non-model species. *Semin. Cell Dev. Biol.* **106**, 1–4 (2020).
2. Kolb, H., Nelson, R., Ahnelt, P. & Cuenca, N. Cellular organization of the vertebrate retina. *Prog. Brain Res.* **131**, 3–26 (2001).
3. Lamb, T. D. Evolution of phototransduction, vertebrate photoreceptors and retina. *Prog. Retin. Eye Res.* **36**, 52–119 (2013).
4. Vecino, E., Rodriguez, F. D., Ruzafa, N., Pereiro, X. & Sharma, S. C. Glia–neuron interactions in the mammalian retina. *Prog. Retin. Eye Res.* **51**, 1–40 (2016).
5. Goetz, J. *et al*. Unified classification of mouse retinal ganglion cells using function, morphology, and gene expression. *Cell Rep.* **40**, 111040 (2022).
6. Shekhar, K. & Sanes, J. R. Generating and using transcriptomically based retinal cell atlases. *Annu. Rev. Vis. Sci.* **7**, 43–72 (2021).
7. Zeng, H. & Sanes, J. R. Neuronal cell-type classification: challenges, opportunities and the path forward. *Nat. Rev. Neurosci.* **18**, 530–546 (2017).
8. Shiau, F., Ruzycki, P. A. & Clark, B. S. A single-cell guide to retinal development: cell fate decisions of multipotent retinal progenitors in scRNA-seq. *Dev. Biol.* **478**, 41–58 (2021).
9. Vöcking, O. & Famulski, J. K. Single cell transcriptome analyses of the developing zebrafish eye— perspectives and applications. *Front. Cell Dev. Biol.* **11**, 1213382 (2023).
10. Zhang, X., Leavey, P., Appel, H., Makrides, N. & Blackshaw, S. Molecular mechanisms controlling vertebrate retinal patterning, neurogenesis, and cell fate specification. *Trends Genet.* **39**, 736–757 (2023).
11. Hahn, J. *et al*. Evolution of neuronal cell classes and types in the vertebrate retina. *Nature* **624**, 415–424 (2023).
12. Clark, B. S. *et al*. Single-cell RNA-seq analysis of retinal development identifies NFI factors as regulating mitotic exit and late-born cell specification. *Neuron* **102**, 1111–1126.e5 (2019).
13. Hoang, T. *et al*. Gene regulatory networks controlling vertebrate retinal regeneration. *Science* **370**, eabb8598 (2020).
14. Macosko, E. Z. *et al*. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* **161**, 1202–1214 (2015).
15. Rheaume, B. A. *et al*. Single cell transcriptome profiling of retinal ganglion cells identifies cellular subtypes. *Nat. Commun.* **9**, 2759 (2018).
16. Shekhar, K. *et al*. Comprehensive classification of retinal bipolar neurons by single-cell transcriptomics. *Cell* **166**, 1308–1323.e30 (2016).
17. Tran, N. M. *et al*. Single-cell profiles of retinal ganglion cells differing in resilience to injury reveal neuroprotective genes. *Neuron* **104**, 1039–1055.e12 (2019).
18. Yan, W. *et al*. Mouse retinal cell atlas: molecular identification of over sixty amacrine cell types. *J. Neurosci.* **40**, 5177–5195 (2020).
19. Cowan, C. S. *et al*. Cell types of the human retina and its organoids at single-cell resolution. *Cell* **182**, 1623–1640.e34 (2020).
20. Gautam, P. *et al*. Multi-species single-cell transcriptomic analysis of ocular compartment regulons. *Nat. Commun.* **12**, 5675 (2021).
21. Hu, F. *et al*. Single-cell RNA-seq reveals the cellular diversity and developmental characteristics of the retinas of an infant and a young child. *Front. Cell Dev. Biol.* **10**, 803466 (2022).
22. Hu, Y. *et al*. Dissecting the transcriptome landscape of the human fetal neural retina and retinal pigment epithelium by single-cell RNA-seq analysis. *PLoS Biol.* **17**, e3000365 (2019).
23. Li, L. *et al*. Single-cell transcriptomic sequencing reveals tissue architecture and deciphers pathological reprogramming during retinal ischemia in *Macaca fascicularis*. *Invest. Opthalmol. Vis. Sci.* **65**, 27 (2024).
24. Liang, Q. *et al*. Single-nuclei RNA-seq on human retinal tissue provides improved transcriptome profiling. *Nat. Commun.* **10**, 5743 (2019).
25. Lu, Y. *et al*. Single-cell analysis of human retina identifies evolutionarily conserved and species-specific mechanisms controlling development. *Dev. Cell* **53**, 473–491.e9 (2020).
26. Lukowski, S. W. *et al*. A single-cell transcriptome atlas of the adult human retina. *EMBO J.* **38**, e100811 (2019).
27. Lyu, Y. *et al*. Implication of specific retinal cell-type involvement and gene expression changes in AMD progression using integrative analysis of single-cell and bulk RNA-seq profiling. *Sci. Rep.* **11**, 15612 (2021).
28. Menon, M. *et al*. Single-cell transcriptomic atlas of the human retina identifies cell types associated with age-related macular degeneration. *Nat. Commun.* **10**, 4902 (2019).
29. Orozco, L. D. *et al*. Integration of eQTL and a single-cell atlas in the human eye identifies causal genes for age-related macular degeneration. *Cell Rep.* **30**, 1246–1259.e6 (2020).

30. Peng, Y.-R. *et al*. Molecular classification and comparative taxonomics of foveal and peripheral cells in primate retina. *Cell* **176**, 1222–1237.e22 (2019).
31. Voigt, A. P. *et al*. Molecular characterization of foveal versus peripheral human retina by single-cell RNA sequencing. *Exp. Eye Res.* **184**, 234–242 (2019).
32. Wang, S. *et al*. Deciphering primate retinal aging at single-cell resolution. *Protein Cell* **12**, 889–898 (2021).
33. Yan, W. *et al*. Cell atlas of the human fovea and peripheral retina. *Sci. Rep.* **10**, 9802 (2020).
34. Yi, W. *et al*. A single-cell transcriptome atlas of the aging human and macaque retina. *Natl. Sci. Rev.* **8**, nwaa179 (2020).
35. Zhang, L. *et al*. Evolutionary and developmental specialization of foveal cell types in the marmoset. *Proc. Natl. Acad. Sci. USA* **121**, e2313820121 (2024).
36. Yamagata, M., Yan, W. & Sanes, J. R. A cell atlas of the chick retina based on single-cell transcriptomics. *eLife* **10**, e63907 (2021).
37. Celotto, L. *et al*. Single cell RNA sequencing unravels the transcriptional network underlying zebrafish retina regeneration. *eLife* **12**, RP86507 (2023).
38. Emmerich, K. *et al*. Large-scale CRISPR screen reveals context-specific genetic regulation of retinal ganglion cell regeneration. *Development* **151**, dev.202754 (2024).
39. Kölsch, Y. *et al*. Molecular classification of zebrafish retinal ganglion cells links genes to cell types to behavior. *Neuron* **109**, 645–662. e9 (2021).
40. Lyu, P. *et al*. Common and divergent gene regulatory networks control injury-induced and developmental neurogenesis in zebrafish retina. *Nat. Commun.* **14**, 8477 (2023).
41. Ogawa, Y. & Corbo, J. C. Partitioning of gene expression among zebrafish photoreceptor subtypes. *Sci. Rep.* **11**, 17340 (2021).
42. Santhanam, A., Shihabeddin, E., Wei, H., Wu, J. & O'Brien, J. Molecular basis of retinal remodeling in a zebrafish model of retinitis pigmentosa. *Cell. Mol. Life Sci.* **80**, 362 (2023).
43. Xu, B. *et al*. Unifying developmental programs for embryonic and post-embryonic neurogenesis in the zebrafish retina. *Development* **147**, dev185660 (2020).
44. Wang, J. *et al*. Molecular characterization of the sea lamprey retina illuminates the evolutionary origin of retinal cell types. *Nat. Commun.* **15**, 10761 (2024).
45. Collin, S. P. Scene through the eyes of an apex predator: a comparative analysis of the shark visual system. *Clin. Exp. Optom.* **101**, 624–640 (2018).
46. Coolen, M. *et al*. The dogfish *Scyliorhinus canicula*: a reference in jawed vertebrates. *Cold Spring Harb. Protoc.* **3**, pdb.emo111 (2008).
47. Hart, N. S. Vision in sharks and rays: opsin diversity and colour vision. *Semin. Cell Dev. Biol.* **106**, 12–19 (2020).
48. Hernández-Núñez, I. *et al*. Loss of active neurogenesis in the adult shark retina. *Front. Cell Dev. Biol.* **9**, 628721 (2021).
49. Ferreiro-Galve, S., Rodríguez-Moldes, I., Anadón, R. & Candal, E. Patterns of cell proliferation and rod photoreceptor differentiation in shark retinas. *J. Chem. Neuroanat.* **39**, 1–14 (2010).
50. Stenkamp, D. L. Neurogenesis in the fish retina. *Int. Rev. Cytol.* **259**, 173–224 (2007).
51. Rodríguez-Arrizabalaga, M., Hernández-Núñez, I., Candal, E. & Barreiro-Iglesias, A. Use of vivo-morpholinos for gene knockdown in the postnatal shark retina. *Exp. Eye Res.* **226**, 109333 (2023).
52. Krishnaswami, S. R. *et al*. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat. Protoc.* **11**, 499–524 (2016).
53. Murat, F. *et al*. The molecular evolution of spermatogenesis across mammals. *Nature* **613**, 308–316 (2023).
54. Sepp, M. *et al*. Cellular development and evolution of the mammalian cerebellum. *Nature* **625**, 788–796 (2024).
55. Dobin, A. *et al*. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
56. Kaminow, B., Yunusov, D. & Dobin, A. STARsolo: accurate, fast and versatile mapping/quantification of single-cell and single-nucleus RNA-seq data. Preprint at https://www.biorxiv.org/content/10.1101/2021.05.05.442755v1 (2021).
57. Genome assembly sScyCan1.1. *NCBI Genome assemby database* https://identifiers.org/refseq.gcf:GCF_902713615.1 (2020).
58. Pertea, G. & Pertea, M. GFF Utilities: GffRead and GffCompare. *F1000Res.* **9**, 304 (2020).
59. Hao, Y. *et al*. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).
60. Ahlmann-Eltze, C. & Huber, W. glmGamPoi: fitting Gamma-Poisson generalized linear models on single cell count data. *Bioinformatics* **36**, 5701–5702 (2021).
61. McGinnis, C. S., Murrow, L. M. & Gartner, Z. J. DoubletFinder: doublet detection in single-cell RNA sequencing data using artificial nearest neighbors. *Cell Syst.* **8**(329-337), e4 (2019).
62. Genome assembly GRCm39. *NCBI Genome assembly database* https://identifiers.org/refseq.gcf:GCF_000001635.27 (2020).
63. Camacho, C. *et al*. BLAST + : architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
64. Genome assembly GRCh38.p14. *NCBI Genome assembly database* https://identifiers.org/refseq.gcf:GCF_000001405.40 (2022).
65. Genome assembly GRCz11. *NCBI Genome assembly database* https://identifiers.org/refseq.gcf:GCF_000002035.6 (2017).
66. Benson, D. A. *et al*. GenBank. *Nucleic Acids Res.* **41**, D36–D42 (2012).
67. Glasauer, S. M. K. & Neuhauss, S. C. F. Whole-genome duplication in teleost fishes and its evolutionary consequences. *Mol. Genet. Genomics* **289**, 1045–1060 (2014).
68. Sayers, E. W. *et al*. Database resources of the national center for biotechnology information. *Nucleic Acids Res.* **50**, D20–D26 (2022).
69. *Scyliorhinus canicula* retina single-nuclei RNA sequencing, *NCBI Sequence Read Archive*, https://identifiers.org/insdc.sra:SRP480047 (2023).
70. Barrett, T. *et al*. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* **41**, D991–D995 (2013).
71. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002).
72. Vidal-Vázquez, N. *et al*. A single-nucleus RNA sequencing atlas of the postnatal retina of the shark *Scyliorhinus canicula*. *GEO* http://identifiers.org/geo:GSE282457 (2024).
73. Vidal-Vázquez, N. *et al*. A single-nucleus RNA sequencing atlas of the postnatal retina of the shark *Scyliorhinus canicula*. *Figshare* https://doi.org/10.6084/m9.figshare.27900822 (2025).
74. Lamb, T. D. Evolution of the genes mediating phototransduction in rod and cone photoreceptors. *Prog. Retin. Eye Res.* **76**, 100823 (2020).
75. Neumayer, L. Der feinere Bau der Selachier-Retina. *Arch. Mikr. Anat.* **48**, 83–111 (1896).
76. Bozzano, A., Murgia, R., Vallerga, S., Hirano, J. & Archer, S. The photoreceptor system in the retinae of two dogfishes, *Scyliorhinus canicula* and *Galeus melastomus*: possible relationship with depth distribution and predatory lifestyle. *J. Fish Biol.* **59**, 1258–1278 (2001).
77. Schieber, N. L., Collin, S. P. & Hart, N. S. Comparative retinal anatomy in four species of elasmobranch. *J. Morphol.* **273**, 423–440 (2012).
78. Dowling, J. E. & Ripps, H. On the duplex nature of the skate retina. *J. Exp. Zool.* **256**, 55–65 (1990).
79. Nelson, S. M., Frey, R. A., Wardwell, S. L. & Stenkamp, D. L. The developmental sequence of gene expression within the rod photoreceptor lineage in embryonic zebrafish. *Dev. Dyn.* **237**, 2903–2917 (2008).
80. Baden, T. & Osorio, D. The retinal basis of vertebrate color vision. *Annu. Rev. Vis. Sci.* **5**, 177–200 (2019).
81. Yovanovich, C. A. M. *et al*. The dual rod system of amphibians supports colour discrimination at the absolute visual threshold. *Philos. Trans. R. Soc. B-Biol.* **372**, 20160066 (2017).
82. Kreplova, M. *et al*. Dose-dependent regulation of horizontal cell fate by Onecut family of transcription factors. *PLoS One* **15**, e0237403 (2020).

83. Passini, M. A., Levine, E. M., Canger, A. K., Raymond, P. A. & Schechter, N. *Vsx*-1 and *Vsx*-2: differential expression of two *Paired*-like homeobox genes during zebrafish and goldfish retinogenesis. *J. Comp. Neurol.* **388**, 495–505 (1997).
84. Bejarano-Escobar, R. *et al*. Retinal histogenesis and cell differentiation in an elasmobranch species, the small-spotted catshark *Scyliorhinus canicula*. *J. Anat.* **220**, 318–335 (2012).
85. Ferreiro-Galve, S., Rodríguez-Moldes, I. & Candal, E. Pax6 expression during retinogenesis in sharks: comparison with markers of cell proliferation and neuronal differentiation. *J. Exp. Zool. B Mol. Dev. Evol.* **318**, 91–108 (2012).
86. Ferreiro-Galve, S., Candal, E., Carrera, I., Anadón, R. & Rodríguez-Moldes, I. Early development of GABAergic cells of the retina in sharks: an immunohistochemical study with GABA and GAD antibodies. *J. Chem. Neuroanat.* **36**, 6–16 (2008).
87. Brandon, C. Cholinergic amacrine neurons of the dogfish retina. *Vis. Neurosci.* **6**, 553–562 (1991).
88. Saari, J. C. & Crabb, J. W. Focus on molecules: cellular retinaldehyde-binding protein (CRALBP). *Exp. Eye Res.* **81**, 245–246 (2005).
89. Sánchez-Farías, N. & Candal, E. Identification of radial glia progenitors in the developing and adult retina of sharks. *Front. Neuroanat.* **10**, 65 (2016).
90. Kee, N., Sivalingam, S., Boonstra, R. & Wojtowicz, J. M. The utility of Ki-67 and BrdU as proliferative markers of adult neurogenesis. *J. Neurosci. Methods* **115**, 97–105 (2002).
91. Bejarano-Escobar, R., Blasco, M., Durán, A. C., Martín-Partido, G. & Francisco-Morcillo, J. Chronotopographical distribution patterns of cell death and of lectin-positive macrophages/microglial cells during the visual system ontogeny of the small-spotted catshark *Scyliorhinus canicula*. *J. Anat.* **223**, 171–184 (2013).
92. Zheng, G. X. Y. *et al*. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).
93. Bray, N. L., Pimentel, H., Melsted, P. & Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
94. Srivastava, A., Malik, L., Smith, T., Sudbery, I. & Patro, R. Alevin efficiently estimates accurate gene abundances from dscRNA-seq data. *Genome Biol.* **20**, 65 (2019).
95. Keesey, M. T. *PhyloPic* https://www.phylopic.org/ (2024).

Acknowledgements

This work was supported by grant ED431C 2021/18 funded by Xunta de Galicia to E.C., and grant PID2020-115121GB-I00 funded by MICIU/AEI/10.13039/501100011033 to A.B.-I. N.V.-V. was supported by grant FPU21/03076 funded by Ministerio de Ciencia, Innovación y Universidades. D.R. was supported by the Axencia Galega the Innovación (GAIN, Xunta de Galicia) as part of the Oportunius programme, and by BBSRC Institute Strategic Grants to the Roslin Institute (BBS/E/20002172, BBS/E/D/30002275, BBS/E/D/10002070 and BBS/E/RL/230002 A). S.S. gratefully acknowledges an NSERC PDF award. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (VerteBrain to H.K., grant agreement no. 101019268). This work has made use of the resources provided by the Edinburgh Compute and Data Facility (ECDF) (http://www.ecdf.ed.ac.uk/), and by the Galician Supercomputing Center (CESGA). The supercomputer FinisTerrae III has been funded by the NextGeneration EU 2021 Plan de Recuperación, Transformación y Resiliencia (ICT2021-006904), by the Programa Operativo Plurirregional de España 2014–2020 of the European Regional Development Fund (ERDF) (ICTS-2019-02-CESGA-3), and by the Programa Estatal de Fomento de la Investigación Científica y Técnica de Excelencia del Plan Estatal de Investigación Científica y Técnica y de Innovación 2013–2016 (1) Subprograma estatal de infraestructuras científicas y técnicas y equipamiento of ERDF (CESG15-DE-3114).

Author contributions

Conceptualisation: N.V.-V., I.H.-N., A.B.-I., E.C. Data acquisition: I.H.-N., P.C.-P., S.S., P.R.V., F.H.-S., X.Y., C.S., J.S., S.M. (provided early access to the catshark genome), H.K. (supervision), D.R. Data analysis: N.V.-V., P.C.-P., S.S., P.R.V., F.L., D.R., E.C. Data interpretation: N.V.-V., P.C.-P., A.B.I., E.C. Writing (original draft): N.V.-V., P.C.-P., S.S., D.R., A.B.-I., E.C. Writing (review and editing): N.V.-V., I.H.-N., P.C.-P., S.S., P.R.V., S.M., H.K., F.A., D.R., A.B.-I., E.C. Funding acquisition: H.K., D.R., A.B.-I., E.C. Project supervision: E.C.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41597-025-04547-2.

Correspondence and requests for materials should be addressed to E.C.