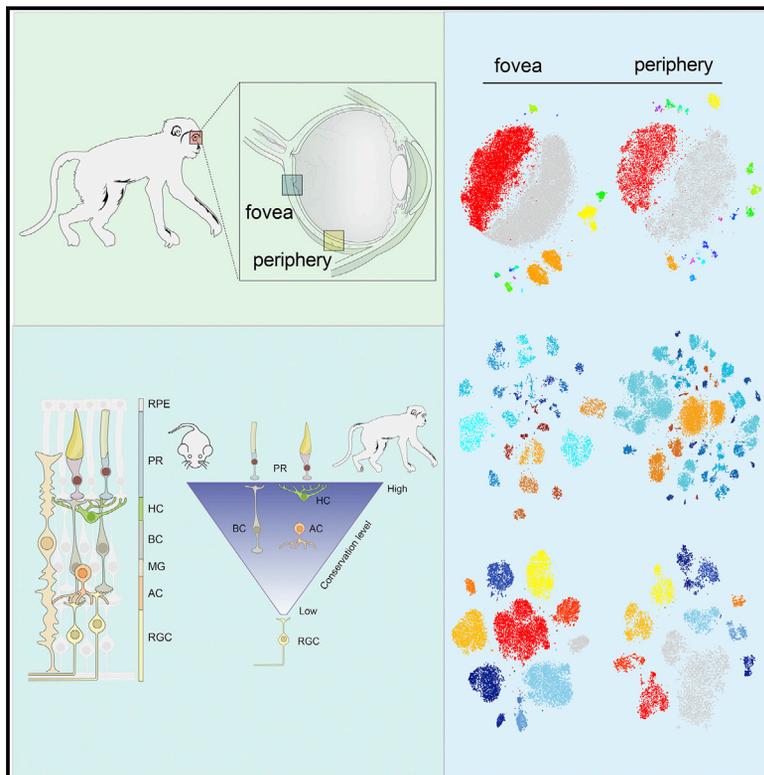


Molecular Classification and Comparative Taxonomics of Foveal and Peripheral Cells in Primate Retina

Graphical Abstract



Authors

Yi-Rong Peng, Karthik Shekhar, Wenjun Yan, ..., Michael Tri. H. Do, Aviv Regev, Joshua R. Sanes

Correspondence

sanesj@mcb.harvard.edu

In Brief

Single-cell-based analysis provides a comprehensive molecular and cellular taxonomy of the primate retina.

Highlights

- Macaque fovea and peripheral retina each contain >65 cell types
- Most types correspond between regions but differ in proportions and gene expression
- Greater conservation of interneuron than ganglion cell types between macaque and mouse
- Cell-type- and region-specific expression of genes implicated in human blindness

Resource

Molecular Classification and Comparative Taxonomics of Foveal and Peripheral Cells in Primate Retina

Yi-Rong Peng,^{1,6} Karthik Shekhar,^{2,6} Wenjun Yan,¹ Dustin Herrmann,¹ Anna Sappington,² Gregory S. Bryman,³ Tavé van Zyl,⁴ Michael Tri. H. Do,³ Aviv Regev,^{2,5} and Joshua R. Sanes^{1,7,*}

¹Center for Brain Science and Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138, USA

²Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA

³F.M. Kirby Neurobiology Center and Department of Neurology, Boston Children's Hospital Harvard Medical School, Boston, MA 02115, USA

⁴Department of Ophthalmology, Harvard Medical School, Boston, MA 02114, USA

⁵Howard Hughes Medical Institute, Koch Institute of Integrative Cancer Research, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02140, USA

⁶These authors contributed equally

⁷Lead Contact

*Correspondence: sanesj@mcb.harvard.edu

<https://doi.org/10.1016/j.cell.2019.01.004>

SUMMARY

High-acuity vision in primates, including humans, is mediated by a small central retinal region called the fovea. As more accessible organisms lack a fovea, its specialized function and its dysfunction in ocular diseases remain poorly understood. We used 165,000 single-cell RNA-seq profiles to generate comprehensive cellular taxonomies of macaque fovea and peripheral retina. More than 80% of >60 cell types match between the two regions but exhibit substantial differences in proportions and gene expression, some of which we relate to functional differences. Comparison of macaque retinal types with those of mice reveals that interneuron types are tightly conserved. In contrast, projection neuron types and programs diverge, despite exhibiting conserved transcription factor codes. Key macaque types are conserved in humans, allowing mapping of cell-type and region-specific expression of >190 genes associated with 7 human retinal diseases. Our work provides a framework for comparative single-cell analysis across tissue regions and species.

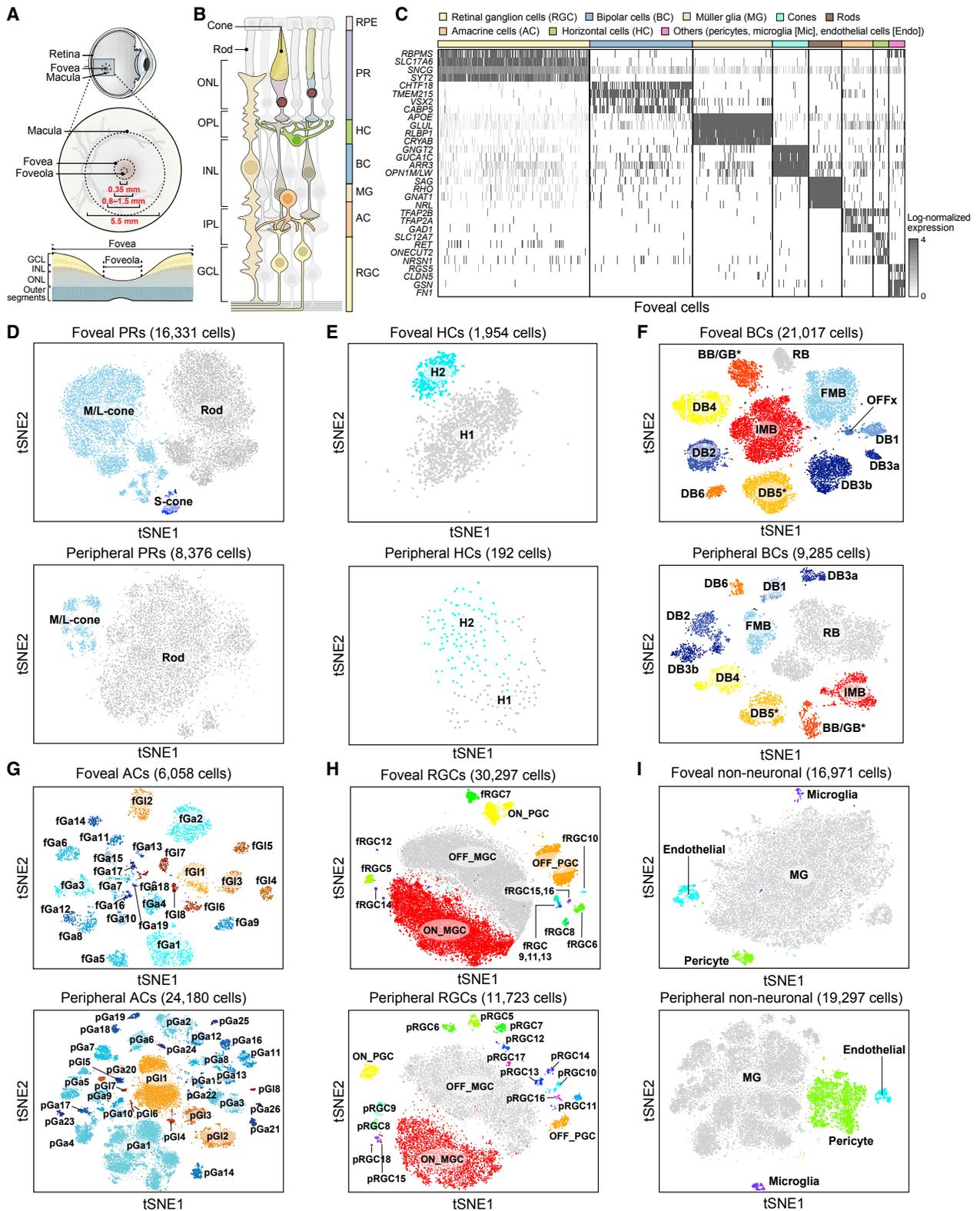
INTRODUCTION

Most primates, including humans, see objects clearly only when they look straight at them, so their image falls on a small central region of the retina called the fovea (Figure 1A). This is because the fovea, which may be the only primate-specific structure in the mammalian central nervous system, mediates high acuity vision. Although it occupies <1% of the retinal surface, it provides ~50% of the input to primary visual cortex (Bringmann et al., 2018; Provis et al., 2013). Diseases that selectively affect this region, such as macular degeneration, are leading causes of human blindness in the developed world (Bourne et al.,

2013). These important aspects of human retinal function and dysfunction cannot be satisfactorily studied in more accessible model organisms, including mice, because they lack a fovea.

Basic aspects of retinal structure and function are shared by all retinal regions in all vertebrate species (Figure 1B). Photoreceptors (PRs; rods and cones) convert light into electrical signals, which are transferred to interneurons (HCs, horizontal cells; BCs, bipolar cells; ACs, amacrine cells). Interneurons process the information and deliver it to retinal ganglion cells (RGCs), which send axons through the optic nerve to the rest of the brain (Masland, 2012). Selective connectivity among multiple AC, BC, and RGC types, along with type-specific intrinsic differences, make each RGC type responsive to specific visual features (Gollisch and Meister, 2010). Despite this conserved plan, however, the fovea and peripheral retina differ in many ways (Bringmann et al., 2018). For example, many RGCs in the fovea are supplied by a single PR, which maximizes acuity, whereas peripheral RGCs are supplied by dozens to hundreds of PRs, enhancing sensitivity (Figure S1A). In addition, the fovea is rich in cones, which mediate color vision in well-lit conditions, whereas peripheral retina is dominated by rods, which function in dim light, further enhancing sensitivity (night vision). Cell types in the two regions also exhibit distinct morphologies and physiological functions (Figure S1A) (Greff, 1900; Kolb and Marshak, 2003; Sinha et al., 2017). Because molecular information is lacking for most primate retinal cell types and genetic access is limited, the underpinnings of these differences are little explored.

Lack of molecular knowledge has also precluded assessment of the conservation of retinal cell types between primates and more widely used model organisms. Identification of orthologous cell types is challenging in the absence of such knowledge. The five retinal neuronal classes (PRs, HCs, BCs, ACs, and RGCs) are divided into >120 types in mice and ~60 in primates, but the extent of their correspondence is unclear (e.g., Sanes and Masland, 2015; Masri et al., 2019). Molecular characterization could provide a lingua franca to enable mapping, informing relationships between common model organisms and primates.



(legend on next page)

To address these issues, we generated an atlas of >165,000 single-cell RNA sequencing (scRNA-seq) profiles from the fovea and peripheral retina of adult crab-eating macaques (*Macaca fascicularis*), a widely used primate model in vision studies (Figure S1B). We identified and molecularly characterized >60 cell types in each region, and validated many of them by pairing molecular markers with cell morphology. We then compared the fovea and periphery, showing that most types are shared by the two regions, but exhibit marked region-specific differences in proportions and expression programs. Next, we compared macaque and mouse retinal cell types. We found a tight correspondence between mouse and macaque PRs, BCs, and ACs, but a divergence in RGC types. Key macaque types and molecular features are also conserved in marmosets and humans. Finally, we used our atlas to determine the cell- and region-specific expression of >190 genes implicated in 7 complex diseases that cause human blindness, demonstrating striking patterns of region- and cell-type-specific expression. Overall, our study provides a general framework for using molecularly derived taxonomies of cell types to understand regional and species specializations in the nervous system and other organs.

RESULTS

Comprehensive Molecular Taxonomy of the Primate Retina Using scRNA-Seq

We generated cell atlases of foveal and peripheral regions of *M. fascicularis* retina using droplet-based scRNA-seq (Zheng et al., 2017) (STAR Methods), yielding high-quality cell profiles from 92,628 foveal and 73,053 peripheral cells (Table S1). For foveal samples, cells were dissociated from 0.5–1.5 mm diameter pieces and profiled without further processing. Because ~80% of total cells in peripheral retina are rod PRs, we depleted rods (CD73+) or enriched RGCs (CD90+) from peripheral samples before profiling (Figure S1B and STAR Methods). In parallel, we assembled a retina-specific transcriptome, which substantially improved the mapping of scRNA-seq reads compared to existing references (Figures S1C–S1E), emphasizing the importance of high-quality tissue-specific transcriptome for scRNA-seq analyses.

To maximize our ability to distinguish cell types consistently, we first grouped cells into the six major retinal classes (RGCs,

BCs, PRs, ACs, HCs, and non-neuronal cells) based on known class-specific gene signatures (Figures 1C and S1F and Table S1) and then iteratively clustered cells within each class separately, building on our earlier methods (Shekhar et al., 2016). Altogether, we distinguished 64 foveal (3 PR, 2 HC, 12 BC, 27 AC, 16 RGC, and 4 non-neuronal) and 71 peripheral (2 PR, 2 HC, 11 BC, 34 AC, 18 RGC, and 4 non-neuronal) clusters (Figures 1D–1I); all clusters were found in all animals (Figures S1G–S1I). As noted below, a few clusters contain >1 cell type. Non-neuronal cells were Müller glia, pericytes, endothelial cells, and microglia (Figure 1I). There are no oligodendrocytes in healthy retina, and astrocytes were not recovered in our samples. Using differential expression analysis, we identified molecular markers for each cluster and used this information to assign clusters to individual cell types. We first describe the types and then return to differences between fovea and periphery.

L and M Cones Are Distinguished Only by Expression of Opsin Paralog

Retinas of trichromatic primates, including macaques and humans, contain rods plus three cone types that preferentially detect long (L; red), medium (M; green), and short (S; blue) wavelengths, because of their expression of L, M and S opsins (*OPN1LW*, *OPN1MW*, *OPN1SW*) respectively. S cones were readily distinguished, whereas M and L cones mapped to a single cluster (Figure 1D). *OPN1MW* and *OPN1LW* have ~98% identical coding sequences (Onishi et al., 2002) and are not distinguished in the macaque reference genome. We therefore retrieved exons that distinguish the two opsins from ~70% of the M/L cones and used them to evaluate M- and L-specific single-nucleotide polymorphisms (SNPs) (Figures S2A–S2C and S2E). Of these, >96% from both fovea and periphery exhibited exclusive L or M SNPs (Figure S2B, S2C, and S2E).

Remarkably, no genes other than *OPN1MW* or *OPN1LW* were differentially expressed (DE) between M and L cones (Figures 2A and S2D), supporting a model in which stochastic choice between the neighboring *OPN1MW* and *OPN1LW* genes diversifies otherwise identical cones (Wang et al., 1999). In contrast, many genes distinguished M/L from S cones and rods and S cones from rods (Figures 2B, 2C, and S2F).

Figure 1. Single-Cell Profiling of Peripheral and Foveal Cells from Macaque Retina

(A) (Top) Sketch of a primate eye showing position of fovea and macula. (Middle) Central region indicating diameters of the foveola (the foveal pit), fovea, and macula. (Bottom) Sketch of a section through macaque fovea, showing foveal pit (Foveola) and displacement of ganglion cell layer (GCL) and inner nuclear layer (INL) cells.

(B) Sketch of peripheral retina showing its major cell classes—photoreceptors (PRs), horizontal cells (HCs), bipolar cells (BCs), amacrine cells (ACs), retinal ganglion cells (RGCs) and Müller glia (MG), outer and inner plexiform (synaptic) layers (OPL and IPL), outer and inner nuclear layers (ONL and INL), and ganglion cell layer (GCL).

(C) Expression patterns of class-specific marker genes (rows; Table S1) in single foveal cells (columns). Cells are grouped by their class (color bar, top). Plot shows randomly selected 10% of total cells. These signatures were used to separate foveal and peripheral cells into classes.

(D–I) Visualization of foveal (top panel) and peripheral (bottom panel) PRs (D), HCs (E), BCs (F), ACs (G), RGCs (H), and non-neuronal cells (I) using t-distributed stochastic neighbor embedding (t-SNE), a 2D non-linear transformation of high-dimensional data that assigns proximal x-y coordinates to cells (points) with similar expression profiles. Individual cells are colored by their cluster assignments. Cluster labels, corresponding to post hoc assigned types, are indicated. Although substructure is visible in (D) and (I), this reflected batch effects, and we were unable to detect subtypes by reanalysis. Because of their low density in the peripheral retina, only three S cones were observed among the peripheral PRs, but these were not sufficient to form a separate cluster from M/L cones. Foveal (f) and peripheral (p) AC clusters are divided into GABAergic (Ga) and Glycinergic (Gl) groups and then numbered from largest to smallest cell number within each group. For RGCs, OFF and ON MGCs and PGCs are labeled; clusters 5+ numbered from largest to smallest in the fovea (f) and periphery (p), respectively. See also Figure S1 and Table S1.

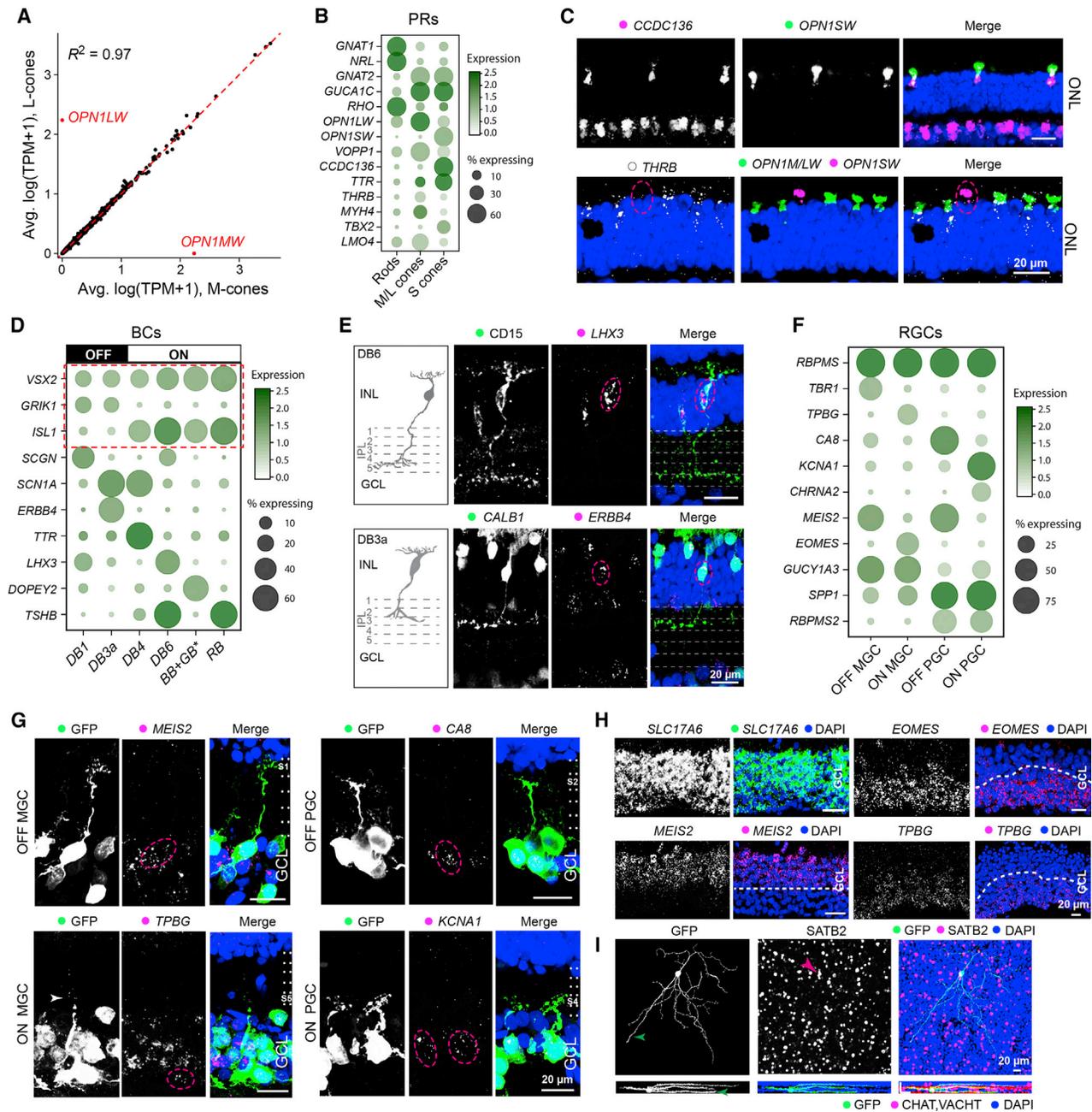


Figure 2. Matching scRNA-Seq Clusters to Neuronal Types of the Primate Retina

(A) Comparison of average transcriptional profiles of foveal M cones and L cones. Each dot corresponds to a gene. No genes other than *OPN1MW* and *OPN1LW* differ significantly in expression levels (>1.2 -fold at $p < 0.01$, MAST test) between the two cone types.

(B) Dot plot showing expression of genes (rows) that distinguish PR types (columns) common to both the fovea and the periphery. The size of each circle is proportional to the percentage of cells expressing the marker (≥ 1 UMI), and its intensity depicts the average transcript count within expressing cells.

(C) Validation of S-cone-specific gene *CCDC136* (upper) and M/L-cone-specific gene *THR8* (lower) by double and triple FISH with *OPN1SW* (S opsin) and *OPN1MW/LW* (M/L opsin) in the peripheral retina. Circle highlights an S cone.

(D) Gene expression patterns of type-enriched markers for selected BC types. Red box highlights a pan-BC, a pan-ON-BC, and a pan-OFF-BC marker. See Figure S2G and Table S2 for lists.

(E) Validation of new markers for two BC types in peripheral retina by FISH combined with immunostaining. (Top) DB6 (circle) cells, known to be CD15-positive, also express *LHX3*. (Bottom) DB3a cell (circle) is *CALB1+ERBB4+*. IPL sublaminae (S1–S5) are demarcated by dashed lines. Sketches redrawn from (Tsukamoto and Omi, 2015, 2016).

(F) Expression patterns of genes selectively enriched among foveal ON and OFF MGCs and PGCs.

(G) Validation of markers for MGCs and PGCs from (F) combining FISH with viral labeling (GFP) to show RGC morphology in the fovea.

(legend continued on next page)

Molecular Classification and Validation of Interneuronal (BC, HC, AC) Types

BCs, the major excitatory interneurons of the retina, are divided into those that receive predominant input from cones or rods; cone BCs are further divided into those excited (ON BCs) or inhibited (OFF BCs) by light (Euler et al., 2014). 12 BC types (1 rod, 5 OFF cone, and 6 ON cone) have been previously reported in macaque peripheral retina based on morphology and immunohistochemistry (Grünert et al., 1994; Tsukamoto and Omi, 2015, 2016).

We identified 12 foveal BC and 11 peripheral BC clusters; 10 mapped 1:1 to known types (Table S2), and supervised methods separated another cluster into two known types (Figure S2H). The additional foveal cluster appears to be a previously unreported OFF type, which we call OFFx (Figure S2I). We identified markers for each BC type and validated several of them (Figures 2D, 2E, S2G, and S2J–S2O).

HCs and ACs are the main sources of inhibitory input to retinal circuitry. We identified two clusters of HCs in both fovea and periphery (Figure 1E) and mapped them to distinct HC subsets in tissue. We also validated *SPP1* (osteopontin) as a pan-HC marker (Figure S5D). We identified 27 foveal and 34 peripheral AC clusters, including both GABAergic and glycinergic subclasses (Figures 1G, S4A, and S5E and Table S2). We assigned six clusters to known mouse types based on their expression of orthologous markers (Figures S4A and S5E). Neuropeptides were expressed in subsets of GABAergic AC types (Table S2).

Markers for Midget and Parasol RGCs

Morphological studies have distinguished ~18 RGC types in primate retina (Dacey, 2004; Masri et al., 2019). ON and OFF Midget RGCs (MGCs)—the smallest and most abundant—comprise ≥85% of foveal RGCs, larger ON and OFF parasol RGCs (PGCs) comprise ~10%, small bistratified RGCs (SBCs) account for ~5%, and ~11 types with larger dendritic arbors account for the remaining ~1%. MGCs and PGCs also account for a large majority of peripheral RGCs (Dacey, 2004).

We identified 18 peripheral and 16 foveal RGC clusters (Figure 1H). Based on their abundance, we tentatively identified a pair of clusters as ON and OFF MGCs (~85% of RGCs in fovea and 83% in periphery) and a second pair as ON and OFF PGCs (11% in fovea and 4.8% in periphery). Markers for these types include *TBR1* (OFF MGC), *TPBG* (ON MGC), *CHRNA2* (ON PGC), and *CA8* (OFF PGC). *SPP1* and *RBPM2* are expressed by both PGC types and *GUCY1A3* by both MGC types (Figures 2F and S3C–S3O). We confirmed the identities of these four RGC types by combining fluorescence *in situ* hybridization (FISH) with viral labeling to visualize morphology (Figures 2G, S3B, S3P, and S3Q). Interestingly, RGC somata expressing transcription factors (TFs) selective for ON and OFF MGCs and PGCs were segregated into outer and inner halves of GCL, suggesting a mo-

lecular basis for the laminar organization of physiologically distinct subclasses (Perry and Silveira, 1988) (Figures 2H and S3P). We were unable to further partition MGCs and PGCs by supervised methods.

Three peripheral RGC clusters expressed *OPN4* (melanopsin), a marker of intrinsically photosensitive RGCs (ipRGCs). We identified markers for the remaining clusters, including *SATB2* as a marker for SBCs (Figures 2I and S3A). To our knowledge, these are the first molecular markers for the major RGC types of the primate retina.

Most Peripheral and Foveal Retinal Cell Types Match 1:1

To relate foveal and peripheral clusters based on gene expression, we devised a multi-class learning framework that associated each foveal cell with a peripheral identity and then examined the extent to which clusters across the two regions corresponded (STAR Methods).

Most foveal clusters (77%) mapped 1:1 with peripheral clusters, including 10/12 BC, 2/2 HC, 20/27 AC, and 14/16 RGC clusters (Figures 3A–3D). These matches were statistically significant as quantified by the adjusted Rand index (ARI), a measure of similarity between data clusterings. Six foveal clusters (RGC clusters fRGC11,14 and AC clusters fGa3,7,13 and fGl8) could be partitioned because they each mapped to two closely related peripheral clusters. Most instances of multi-mapping were for AC clusters, likely due to a biased frequency distribution in the peripheral samples, which were not optimized for AC recovery (Figure S4B).

Quantifying the compositional similarity between the fovea and the periphery suggested that differences decrease from outer to inner retina. Values of the Jensen-Shannon divergence (JSD)—a distance measure in frequency space ranging from 0 (identical cell-type composition) to 1 (completely region specific) were: PRs (0.21), BCs (0.16), ACs (0.06), and RGCs (0.035) (STAR Methods). There were numerous differences in relative proportions for 1:1 matched types (Figures 3E–3H and S4B and Table S2), some consistent with previous reports (Grünert and Martin, 1991; Wässle et al., 1994). Six RGC types were >4-fold enriched in foveal compared to peripheral samples (Figure 3G), six GABAergic AC types appeared to be periphery specific (Figure 3C), and the OFFx BC type appeared to be fovea specific (Figure S2I). Taken together, these results suggest that foveal and peripheral neuronal circuitry draw upon a similar, but not identical, “parts list” of cell types.

Programs for Phototransduction and GABAergic Neurotransmission Differ between Matching Types in the Fovea and Periphery

We next asked whether corresponding foveal and peripheral cell types exhibited DE genes (Figures 4 and S4). We found a median of 17 ± 8 genes significantly DE between corresponding types

(H) Somata of ON MGCs (*TPBG*+ and *EOMES*+) and OFF MGCs/PGCs (*MEIS2*+) are localized to the inner and outer halves (divided by dashed lines) of the ganglion cell layer, respectively, proximal to the fovea.

(I) *SATB2*-positive SBC labeled by GFP-expressing virus in the peripheral retina. Arrowheads indicate the axon (green) and soma (red). Bottom panels are rotations to show bistratified dendritic lamination.

Scale bar is 20 μ m. DAPI staining is blue in (C), (E), and (G–I).

See also Figures S2 and S3 and Table S2.

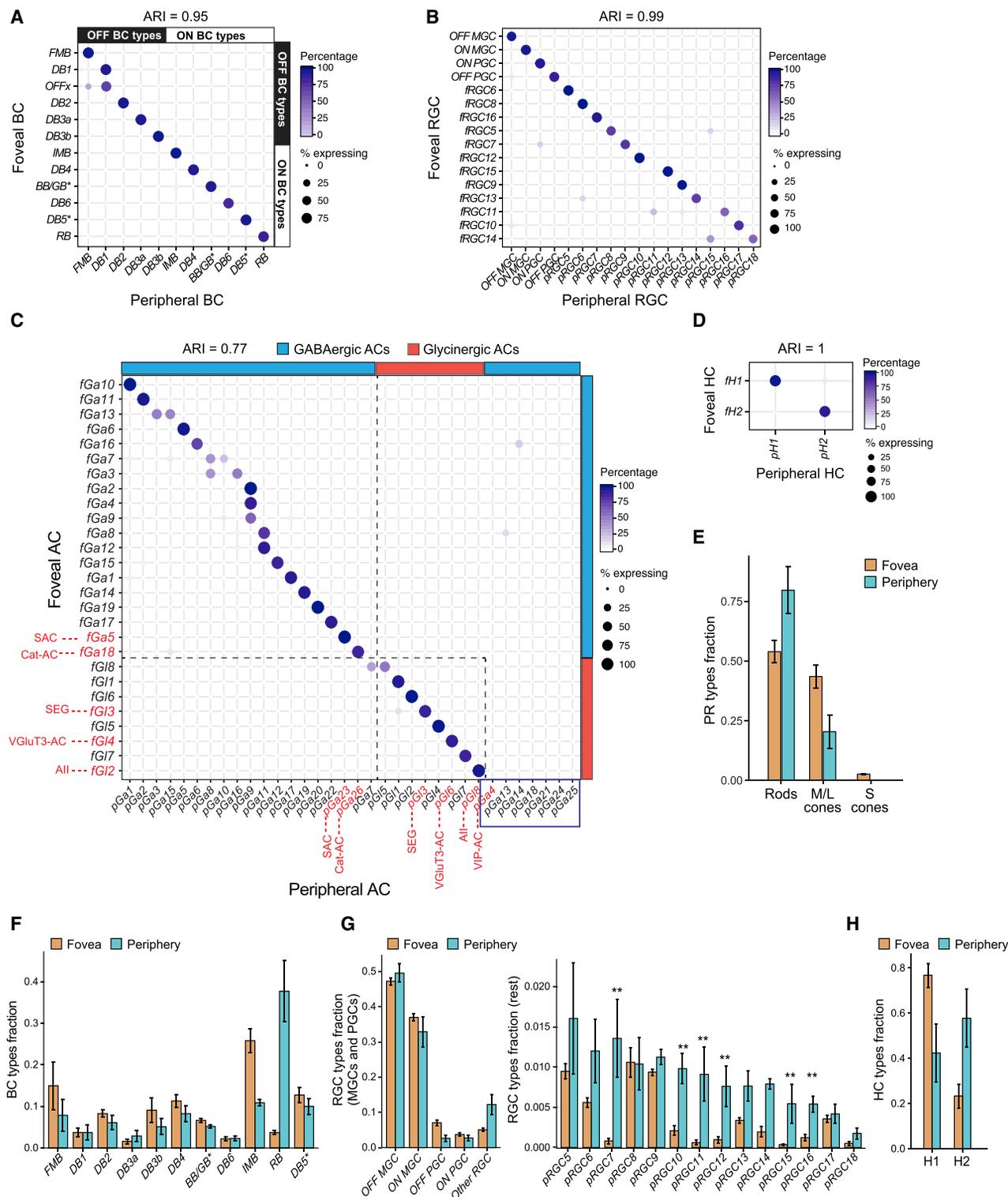


Figure 3. Correspondence between Foveal and Peripheral Clusters

(A–D) Transcriptional correspondence between foveal and peripheral clusters, summarized as “confusion matrices.” Circles and colors indicate the percentage of cells of a given foveal cluster (row) assigned to a corresponding peripheral cluster (column) by the classification algorithm trained exclusively on peripheral cells. (A) BCs. (B) RGCs. (C) ACs. (D) HCs. In (A), bars at the top and right border mark ON and OFF BC subgroups. In (C), bars mark GABAergic and glycinergic AC subsets; key known types are labeled in red, and types that might be periphery specific are highlighted by blue box in (C). [Figure S4A](#) provides molecular markers for AC clusters. In (A–D), the extent of 1:1 cluster matches are quantified by values of the ARI, which range from close to 0 (random) to 1 (perfect 1:1 match).

(legend continued on next page)

(Table S3) and with more differences for RGCs and non-neuronal cells (median 130 genes/type) than for interneurons (HCs, BCs, and ACs; median 13 genes/type) (Figure 4A).

One intriguing difference is in *GNGT1*, which encodes the gamma subunit of transducin, a G protein involved in phototransduction. Light responses are slower in macaque foveal than peripheral cones (Sinha et al., 2017), and homozygous deletion of *Gngt1* accelerates rod responses in mice (Kolesnikov et al., 2011). *GNGT1* is selectively expressed by rod PRs in mice and macaques (Figure 4B) but is also expressed at higher levels in foveal than peripheral cones, as verified by FISH and immunohistochemistry (Figure 4C; see below). It is thus tempting to speculate that the enrichment of *GNGT1* in macaque foveal cones contributes to the kinetic differences—possibly by allowing transducin to compete more effectively with other molecules that deactivate the visual pigment (Doan et al., 2009). Conversely, a cone-specific phototransduction component, *GNB3*, is also expressed at substantial levels by foveal rods (Figure 4B).

Another prominent difference is that several genes involved in GABAergic neurotransmission are expressed at higher levels in peripheral types than in their foveal counterparts. They include the GABA_A receptor subunit *GABRB2* in PRs (Figures 4B and 4C), the GABA synthetic enzyme *GAD2* in rod BCs (Figure 4D) (Lassová et al., 2010), and several GABA receptor subunits in multiple RGC types (Figures 4F–4H and S4C). Thus, the paucity of inhibitory drive to foveal MGCs observed physiologically (Sinha et al., 2017) may reflect both the absence of several GABAergic AC types from fovea (Figure 3C) and decreased expression of GABAergic components in shared cell types.

We also observed >100 DE genes between foveal and peripheral MGs, including *SPP1* enriched in peripheral MGs and *CYP26A1* in foveal MGs (Figures 4I–4K and S4D). These differences provide potential substrates for differences between their interactions with neurons in the two regions (Bringmann et al., 2018; Vecino et al., 2016).

Tight Molecular Correspondence between Macaque and Mouse BC and AC Types

While the conservation of genes across species has been extensively studied, we know less about the extent to which cell types are conserved. scRNA-seq enables systematic comparison of cell atlases across species by leveraging gene orthology (Marioni and Arendt, 2017). To exploit this opportunity, we adapted the multi-class classification framework used for fovea-periphery comparisons to compare macaque cells with those previously collected from mice (Macosko et al., 2015; Shekhar et al., 2016).

As expected from the conserved retinal architecture, there was a 1:1 match between cell classes in the two species (Figure S5A). In addition, there was also a tight correspondence be-

tween mouse and macaque PR, HC, BC, and AC types. Mice are dichromats, with only M and S cone opsins, and many cones express both opsins (Euler et al., 2014). Mouse M and S cones most closely resembled macaque M/L and S cones, respectively (Figures 5A and S5B). Mouse HCs were more transcriptionally similar to macaque H1, which they resemble morphologically, than to H2 (Wässle et al., 2000) (Figures 5B, S5C, and S5D and Table S2). For ACs, there was a striking correspondence between macaque and mouse based on comparison to our previously reported (but underpowered) estimate of 21 molecularly distinct mouse AC clusters (Macosko et al., 2015) (Figure 5C), with clear macaque equivalents of several well-studied mouse types (Figures 5D and S5E and Table S2). A more comprehensive ongoing study has revealed >60 mouse AC types and increased the 1:1 matches with macaque (data not shown).

For BCs, nine peripheral macaque types mapped preferentially to mouse types (Figure 5E; eight were 1:1), and expression patterns of type-enriched orthologs were conserved between the two species (Figure 5F). As noted below, we were unable to identify clear murine equivalents of the most abundant RGC types, MGCs, and PGCs. However, the BCs that selectively innervate MGCs (flat and invaginating midget bipolars [FMBs and IMBs]) correspond to mouse BC1 and 7, and the BCs that provide strongest innervation to PGCs (DB2, 3a, and 4) (Tsukamoto and Omi, 2015, 2016) correspond 1:1 to mouse BC4, 3a, and 5a. It will be interesting to ask whether postsynaptic targets of these BC types share features with MGCs and PGCs.

Divergence in Expression Programs between Macaque and Mouse RGC Types

In contrast to BCs and ACs, mouse and monkey RGC types differ greatly in both number (>40 in mice; Baden et al., 2016; Bae et al., 2018; versus <20 in macaque) and distribution (no single mouse type accounts for >10% of all RGCs, whereas MGCs account for >80% of macaque RGCs). Moreover, using a recently published scRNA-seq atlas of early postnatal mouse RGCs (Rheume et al., 2018), we found few clear matches between the species (ARI = 0.16), one being broadly conserved melanopsin-positive intrinsically photosensitive RGCs (Figures 5G and 5I); macaque RGC types exhibit a similar lack of correspondence with adult mouse RGC types (data not shown). Although limited sampling might reduce our ability to detect conservation, we found no clear mouse equivalents of the abundant macaque MGCs (~35,000 cells).

Nonetheless, we found similar patterns of TF expression across macaque and mouse RGC clusters. For example, *Tbr1*, *Eomes*, *Satb2*, *Tbx20*, and *Foxp2*, several of which regulate type-specific features of mouse RGCs (Liu et al., 2018; Mao et al., 2014; Peng et al., 2017), are also expressed selectively by restricted macaque types (Figure 5H). Among them, *Satb2*

Empirical ARI values were highly significant for all classes, as compared to null ARI values (mean ± SD) from random associations. BCs: $10^{-5} \pm 6 \times 10^{-4}$, RGCs: $-2 \times 10^{-5} \pm 6 \times 10^{-4}$, ACs: $-8.4 \times 10^{-6} \pm 3 \times 10^{-4}$, HCs: $3 \times 10^{-6} \pm 2 \times 10^{-4}$.

(E–H) Comparison of cell-type proportions between the fovea and the periphery (mean ± SD, computed across biological replicates). (E) PRs. (F) BCs. (G) RGCs. (H) HCs. Foveal type OFFx and DB1 are grouped together as “DB1” due to their transcriptional similarity. To facilitate direct comparison of RGCs, each foveal type is assigned a peripheral identity from (B). For ACs, see Figure S4A. Supervised analysis split fRGC11 and fRGC14 into two types each. RGC types underrepresented in the fovea are marked ** in (G).

See also Figure S4.

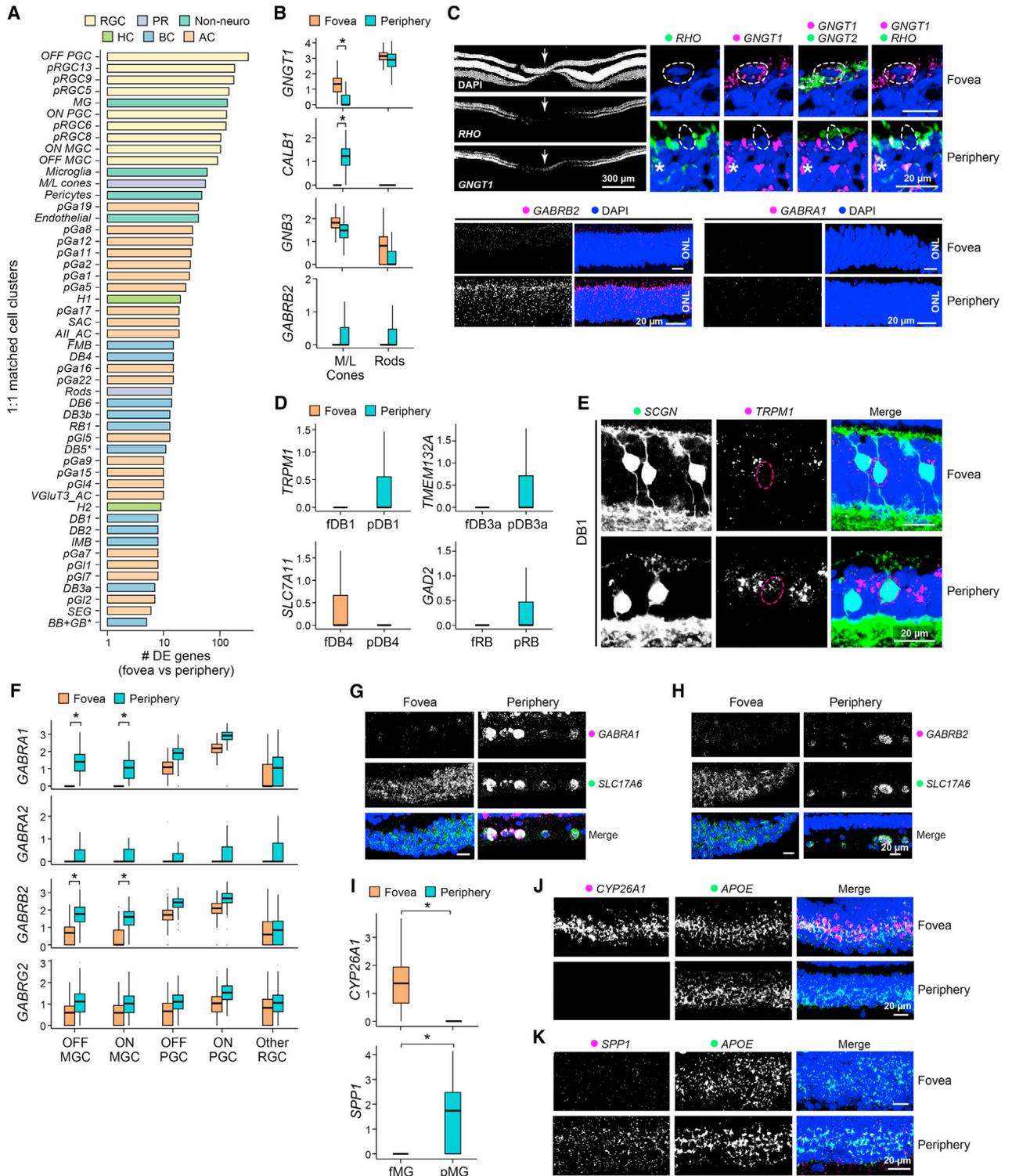


Figure 4. Differences in Gene Expression between Foveal and Peripheral Cell Types

(A) Bar plot showing the number of DE genes per matched cluster between fovea and periphery (log-fold change > 2, $p < 10^{-5}$, MAST test). Bars are labeled based on the corresponding peripheral cluster (except known types), colored by cell class and arranged in decreasing order. Only clusters with ≥ 50 cells in both the fovea and periphery are shown.

(legend continued on next page)

is expressed by bistratified RGCs in both macaques (SBCs; Figure 2I) and mice (ON-OFF direction-selective ganglion cells [oDSGCs]; Peng et al., 2017). Moreover, patterns of TF co-expression were similar in both species (Figure S5F), suggesting a model wherein subfamilies of RGC types are specified by TF codes but with substantial divergence of the regulatory targets of those codes.

Conservation of Cell Types and Markers across Primates

A main value of studies in non-human primates is that their properties are likely to be shared with humans. To ask whether this is true of retina, we performed preliminary analyses on humans, as well as a second non-human primate, the common marmoset (*Callithrix jacchus*), which is increasingly used because of its small size and genetic accessibility (Mitchell et al., 2014).

We found strong conservation of key molecular and cellular features across macaque, human, and marmoset retinas. For example, markers of MGCs, PGCs, S cones, and HC types were conserved in marmoset (Figures 6A, S6A, and S6B), and the laminar segregation of foveal OFF and ON RGCs was apparent in both marmoset and human (Figures 6B and S6D). Distinctions between corresponding peripheral and foveal cell types were also conserved, such as the enrichment of *GNGT1* in foveal cones (Figure S6C) and enrichment of *CYP26A1* and *SPP1* in foveal and peripheral MGs, respectively (Figures 6C, 6D, 7E, and S6E). Moreover, we profiled 2,383 BCs from human peripheral retina and found close correspondence (ARI = 0.74) between human and macaque types (Figures 6E, 6F, and S6F–S6H). Together with previous morphological studies (Chan et al., 2001; Haverkamp et al., 2003; Liao et al., 2016), these results reveal strong conservation of cell types among primates.

Cell-Type-Specific Expression of Human-Retinal-Disease Genes

Finally, encouraged by the correspondence between human and macaque cell types, we assessed the expression of genes that have been implicated in seven diseases associated with disabling vision loss: retinitis pigmentosa, cone-rod dystrophy, diabetic macular edema and retinopathy, congenital stationary night blindness, primary open-angle glaucoma, primary-angle closure glaucoma, and age-related macular degeneration. We focused on ~200 genes for which association has been demonstrated by loss- or gain-of-function mutations or by genome-wide association studies (GWAS) (Farrar et al., 2017; Fritsche

et al., 2016; Graham et al., 2018; Wiggs and Pasquale, 2017; Zeitz et al., 2015).

For each gene, we calculated an enrichment score for each macaque foveal and peripheral cell class. Aggregating these scores by disease groups (Figure 7A) shows that, in general, expression was highest in cell classes primarily affected by the diseases: PRs for retinitis pigmentosa and rod-cone dystrophy, PRs and BCs for congenital stationary night blindness, RGCs for glaucoma, and non-neuronal cells for diabetic macular edema and retinopathy. As expected, many genes implicated in macular degeneration were expressed in PRs and non-neuronal cells, but several were also highly expressed in other cell classes, especially ACs, a class whose role in disease etiology and progression remains to be explored.

Patterns of expression for several genes are noteworthy. At least three genes associated with diabetic macular edema, which, as the name implies, selectively affects the macula, were expressed at higher levels in certain foveal non-neuronal types than their peripheral counterparts, such as *PDGFB* and *EDN1* in endothelial cells (Figure 7C) (Graham et al., 2018). Likewise, the macular degeneration susceptibility gene *VTN* (Fritsche et al., 2016) was enriched in foveal compared to peripheral rods and cones (Figure 7B). *HTRA1*, a major susceptibility gene for macular degeneration (Fritsche et al., 2016), which has been implicated in dysfunction of retinal pigment epithelium, is also expressed at high levels in macaque HCs, but not in mouse HCs (Figure S7). Glaucoma-associated genes *CYP1B1* and *CYP26A1* were enriched in a GABAergic AC type and foveal MG, respectively (Figures 7B, 7E, and S7). Several other glaucoma-associated genes were enriched in specific RGC subsets: *MAP3K1* by MGCs and PGCs, *SIX6* by MGCs, and *GAS7* in multiple non-MGC and PGC types (Figure 7D). These region- and cell-type-selective expression patterns can inform future analyses of genetic variants associated with disease risk. Interestingly, expression patterns of many human-retinal-disease genes were not conserved in mice, highlighting the limitations of using a mouse atlas as a reference (Figure S7).

DISCUSSION

We used scRNA-seq to generate a cell atlas of the primate retina and used it as a foundation to address several questions about retinal structure and function. First, to explore regional differences within a tissue, we compared the cellular compositions

(B) Box-and-whisker plots show examples of gene expression differences foveal and peripheral PRs. P values were calculated using the MAST package. Black horizontal line, median; bars, interquartile range; vertical lines, minimum and maximum. All differences between fovea and periphery in (B), (D), (F), and (I) are statistically significant with $p < 10^{-19}$. Stars indicate >2-fold changes based on means.

(C) *In situ* validation for (B). (Upper) Foveal but not peripheral cones express *GNGT1*. Arrow indicates foveal center. Circles highlight cones; asterisks highlight rods (*RHO+*). (Lower) Peripheral PRs express higher *GABRB2* than foveal PRs; neither type expresses *GABRA1*.

(D) Same as (B), for BCs.

(E) *In situ* validation for (D). *TRPM1* is expressed in peripheral but not foveal DB1. Both are SCGN+.

(F) Same as (B), for RGCs.

(G and H) *In situ* validation for (F). *GABRA1* (G) and *GABRB2* (H) expression in foveal and peripheral RGCs (*SLC17A6+*).

(I) Same as (B), for MG.

(J and K) *In situ* validation for (I). *CYP26A1* (J) and *SPP1* (K) are selectively expressed by foveal and peripheral MG respectively. *APOE* is a pan-MG marker.

Scale bars, 20 and 300 μm as indicated. DAPI staining is blue in (C), (E), (G), (H), (J) and (K).

See also Figure S4 and Table S3.

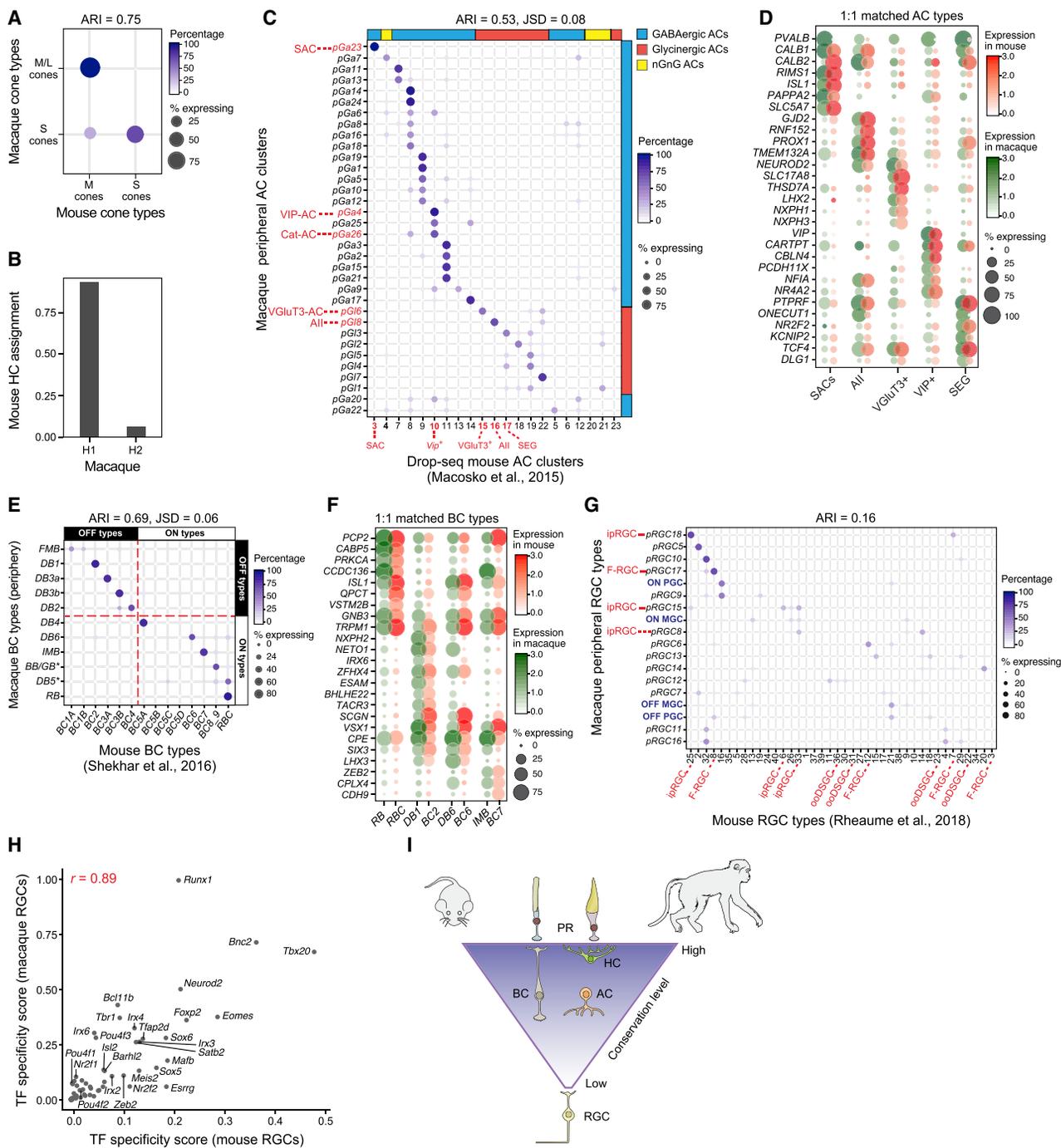


Figure 5. Conservation of Retinal Cell Types between Mouse and Macaque

(A) Transcriptional correspondence between macaque (rows) and mouse cone types (from Macosko et al., 2015) (columns). Only mouse cones expressing M or S opsin, but not both, were used for this comparison.

(B) Supervised classification shows that HCs are more closely related to macaque H1 than to H2 cells.

(C) Transcriptional correspondence between macaque peripheral AC clusters and mouse AC clusters (from Macosko et al., 2015). 1:1 mapping of multiple macaque AC clusters reflects the incomplete resolution of AC types in published mouse data. Known types that map 1:1 are indicated (red).

(D) Example of orthologous gene expression patterns in matched macaque-mouse AC types.

(E) Transcriptional correspondence between macaque peripheral BC types and mouse BC types (from Shekhar et al., 2016). 9 out of 11 macaque BC types map preferentially to a single mouse type. Crossed red lines highlight correspondence of OFF and ON groups.

(F) Example of orthologous gene expression patterns in matched macaque-mouse BC types.

(legend continued on next page)

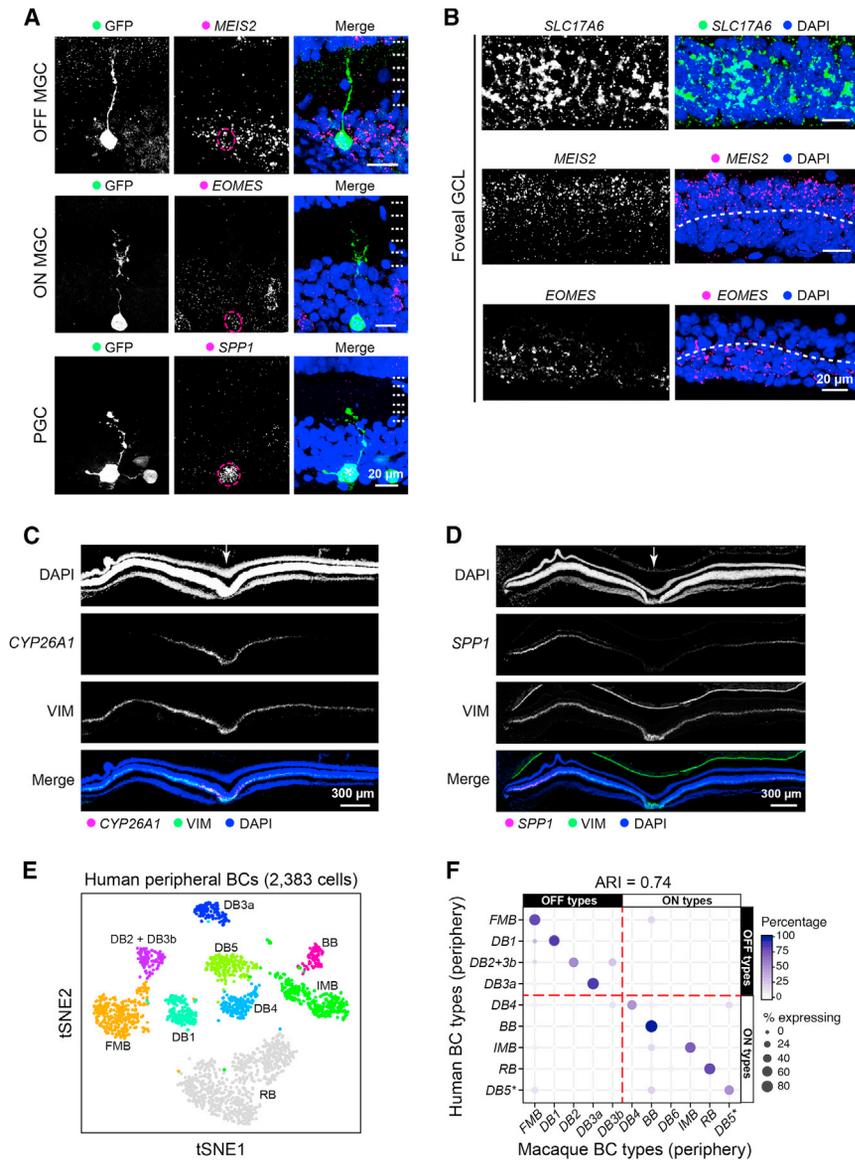


Figure 6. Conservation in Marmosets and Humans

(A) Biolistic labeling combined with FISH shows *MEIS2*+ OFF-MGs (top), *EOMES*+ ON-MGs (middle), and *SPP1*+ PGCs (bottom) in the marmoset fovea. (B) The exclusive expression of *MEIS2* and *EOMES* in the marmoset foveal GCL layer. (C and D) *CYP26A1* (C) and *SPP1* (D) are selectively expressed by foveal and peripheral Müller glia, respectively, in marmoset retinas. Arrows indicate the center of the fovea. (E) t-SNE visualization of human peripheral BCs. Representation as in Figures 1D–1I. Clusters are labeled based on their correspondence to macaque BC types (F). (F) Transcriptional correspondence between human (rows) and macaque (columns) peripheral BC types. Representation as in Figure 3A. Each human BC cluster is labeled here and in (E), retrospectively, based on its most similar macaque BC. Arrows indicate the center of fovea. Scale bars, 20 μ m and 300 μ m as indicated. DAPI staining is blue in (A–D). See also Figure S6.

Primate Retinal Cell Atlas

We detected >70 cell types in peripheral retina and >65 in the fovea. Although extremely rare retinal cell types may have gone undetected, we believe our atlas is reasonably complete. The number of cells needed to generate a comprehensive atlas depends on multiple factors, including the number and the relative frequency of types and the extent to which types differ from each other (Shekhar et al., 2016). We analyzed >165,000 cells and, for peripheral retina, depleted abundant rods from some samples and enriched RGCs in others. These strategies enabled us to detect very rare cell types, such as ipRGCs, which comprise

<0.002% of all retinal cells (Liao et al., 2016). Complete compendia of cell types from comparably complex tissues will likely require similar cell numbers and/or enrichment procedures. One product of the atlas is a set of molecular markers that can be used to mark specific cell types. Among these, the ON and OFF MGCs and PGCs are particularly noteworthy. These four types comprise ~95% of foveal and ~85% of peripheral

of macaque fovea and peripheral retina. Second, to assess evolutionary specializations across species, we compared macaques (which have a fovea) and mice (which do not) and also asked whether foveal specializations in macaque are conserved in two other primates, human and marmoset. Finally, we analyzed the expression of nearly 200 genes that have been associated with blinding diseases.

(G) Transcriptional correspondence between macaque peripheral RGC clusters and mouse RGC clusters (from Rheaume et al., 2018). Select mouse and macaque types are indicated. (H) Despite poor 1:1 correspondence of RGC types across the two species (G), transcription factors (TFs) that exhibit restricted expression among subgroups are similar, based on high Pearson correlation coefficient (0.89) of their specificity scores (STAR Methods) between macaque and mouse. (I) Schematic showing decreasing molecular conservation of types within cell classes from outer to inner retina.

For (A), (C), (E), and (G), null ARI values (mean \pm SD) for a random model are PRs: $1.7 \times 10^{-4} \pm 9 \times 10^{-3}$, ACs: $2 \times 10^{-5} \pm 10^{-3}$, BCs: $3 \times 10^{-4} \pm 9 \times 10^{-4}$, RGCs: $2 \times 10^{-6} \pm 7 \times 10^{-5}$. Cross-species compositional similarity for BCs and ACs was quantified using the JSD metric. To avoid enrichment biases (Figure S4B), only macaque ACs from CD73- samples were considered. See also Figure S5.

See also Figure S5.

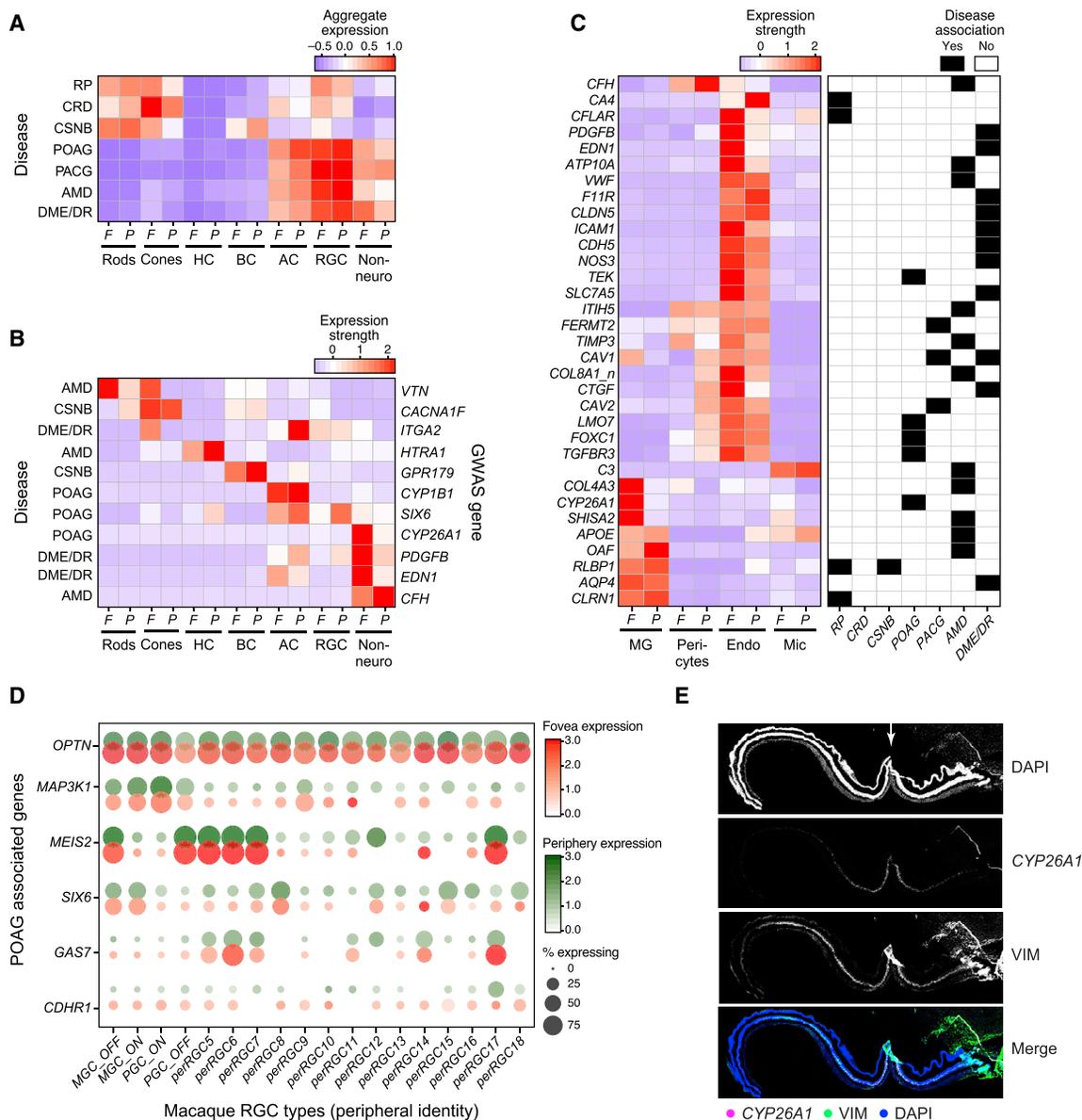


Figure 7. Cell-Type- and Region-Specific Expression Patterns of Human-Retinal-Disease-Associated Genes

(A) Aggregated expression of disease-associated genes in foveal and peripheral cell classes. Disease groups are RP, retinitis pigmentosa; CRD, cone rod dystrophy; CSNB, congenital stationary night blindness; POAG, primary open-angle glaucoma; PACG, primary angle-closure glaucoma; AMD, age-related macular degeneration; and DME/DR, diabetic macular edema and retinopathy.

(B) Expression patterns of a subset of retinal-disease-associated genes by cell class (columns), as in (A). The primary disease associated with each gene is indicated on the right. See Figure S7 for full list.

(C) Expression patterns of specific retinal-disease associated genes (rows) by non-neuronal types in the fovea and the periphery (columns).

(D) Expression of glaucoma associated genes in RGC types in the fovea and the periphery.

(E) The POAG susceptible gene, *CYP26A1*, is selectively expressed by human foveal Müller glia.

Arrow indicates the center of fovea. Scale bars, 300 μm. See also Figure S7.

RGCs; they therefore transmit the vast majority of information from eye to brain. Yet, to date, there have been no means to identify them other than by laborious and low-throughput physiological or morphological methods. The availability of markers now enables a variety of studies on these indispensable cell types.

We also explored differences between L and M cones. These cell types transmit different spectral information, but the question of how they do so has been controversial. If M and L cones connect selectively to different postsynaptic cell types, as some studies suggest (Lee et al., 2012), one would expect them to express different recognition molecules,

implying transcriptional differences that extend beyond the opsin genes. If wiring is non-selective, as indicated by other studies (Wool et al., 2018), no non-opsin expression differences would be needed, implying that our ability to distinguish green from red is experience dependent rather than innate. Our results are consistent with the latter view. However, it is possible that subtle distinctions between L and M cones might have gone undetected, or critical differences might be restricted to early stages, when cones form connections (Hoshino et al., 2017).

Fovea and Periphery

The fovea mediates most high-acuity vision and much chromatic vision in primates. Indeed, one view of the peripheral retina is that at moderate to high light levels, its main job is to aid in directing eye movements to focus salient images on the fovea. Numerous physiological and morphological differences between fovea and periphery have been documented (see Introduction), but lack of molecular information has made it difficult to elucidate the extent to which differences stem from cell types unique to one region, from altered compositions across the same set of types, or from region-specific molecular programs within shared types.

Our results suggest that all three factors are involved, but to different extents. A few cell types are unique to or highly enriched in each region, and these may underlie certain functional differences between them. The OFFx bipolar is particularly interesting in this regard, and the absence of some foveal GABAergic AC types may contribute to the low level of foveal inhibition (Sinha et al., 2017). Nevertheless, our results suggest that the dramatic specializations of foveal circuitry and function arise mostly from regional differences in proportion and gene expression within shared types rather than from region-specific types. The differences in proportions (as measured by JSD) appear to be higher at the sensory and preprocessing layers (PRs and BCs) than for downstream processors and feature detectors (ACs and RGCs). Some of the differentially expressed genes, such as components of phototransduction and neurotransmitter pathways, may underlie documented differences in visual responses mediated by the fovea and periphery (Sinha et al., 2017).

Macaques and Mice

It has been clear for over a century that the major retinal cell classes are conserved across vertebrates (Cajal, 1893). Far less is known about the conservation of types within classes. Comparison of macaque and mouse allowed us to address this issue. Conservation was stronger for types within outer retinal classes (PRs, BCs, and ACs) than for RGCs. Most notably, MGCs, the major RGC types in primates, lack clear counterparts in mice. Similarly, the molecular distinctions between the fovea and the periphery are higher for RGCs than for PRs and interneurons (Figure 4A). These patterns suggest that the outer retina may comprise a conserved set of information processors, with adaptation to species- and region-specific visual-processing requirements beginning at the level of RGCs and then continuing centrally.

This is not to say that macaque and mouse RGC types are unrelated. There may be an underlying shared gene regulatory program that specifies RGC diversity based on conserved patterns

of type-specific TF expression and co-expression. These factors may specify groups that diversify differentially in the two species, raising the possibility that cross-species differences in homologous regions can be explained by rewiring of downstream programs specified by a conserved set of TFs.

In addition to the paucity of 1:1 matches, the number of RGC types is far lower in macaques than in mice. This difference may seem paradoxical, given the greater visual capabilities of primates. One possible explanation is that primates may rely less than mice on a broad array of retinal feature detectors (e.g., direction-selective RGCs) and more on complex processing in the cortex. Thus, primate RGCs, especially MGCs, may be designed to transmit a relatively unprocessed image with little loss of information to the cortex, where more sophisticated and flexible computations are performed. Presumably what is lost in speed is more than compensated by finesse.

Retinal Diseases

Many genes have been implicated in the pathogenesis of retinal diseases—some based on rare but highly penetrant mutations and others as common variants that confer susceptibility (Farrar et al., 2017; Fritsche et al., 2016; Graham et al., 2018; Wiggs and Pasquale, 2017; Zeitz et al., 2015). We assessed expression of nearly 200 such genes and found that many are selectively expressed in particular retinal cell classes, in particular types within classes, and in foveal or peripheral cohorts of particular types. For example, *PDGFB* and *EDN1*, implicated in diabetic macular edema (Graham et al., 2018), are expressed at significantly higher levels by foveal than peripheral endothelial cells, suggesting a basis for the increased susceptibility of the macula in this disease. Interestingly, several of the patterns we document in macaque are not conserved in mice, emphasizing the value of using primate models for investigating human retinal dysfunction. Together, these results provide both technical and conceptual foundations for investigations aiming to understand cellular mechanisms underlying blinding retinal diseases.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODELS AND SUBJECT DETAILS
 - Tissue Procurement
- METHOD DETAILS
 - RNA-sequencing
 - Histological methods
 - Computational Methods
- QUANTIFICATION AND STATISTICAL ANALYSIS
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures, four tables, and one methods file and can be found with this article online at <https://doi.org/10.1016/j.cell.2019.01.004>.

ACKNOWLEDGMENTS

We thank R. Born, A. Liu, F. Rieke, and J. Wiggs for advice; D. Juric for human tissue, and L. Gaffney, A. Hupalowska, and E. Martersteck for assistance. This work was supported by grants from HHMI, the Klarman Cell Observatory, the NIH (MH105960, EY025840, EY028633, EY025555, and EY028625), and the BrightFocus Foundation (M2014055).

AUTHOR CONTRIBUTIONS

Y.-R.P., K.S., A.R., and J.R.S. conceived the study and wrote the manuscript with input from all authors. K.S., W.Y., and A.S. performed bioinformatic analysis with guidance from A.R. Y.-R.P. and D.H. performed molecular and histological experiments with guidance from M.T.H.D. and J.R.S. T.v.Z. and G.S.B. procured and prepared tissue and developed methods. M.T.H.D. provided guidance on study design and analysis.

COMPETING INTERESTS

A.R. is a cofounder and an equity holder in Celsius Therapeutics and an SAB member in Syros Pharmaceuticals and Thermo Fisher Scientific.

Received: September 26, 2018

Revised: November 8, 2018

Accepted: December 31, 2018

Published: January 31, 2019

REFERENCES

- Baden, T., Berens, P., Franke, K., Román Rosón, M., Bethge, M., and Euler, T. (2016). The functional diversity of retinal ganglion cells in the mouse. *Nature* **529**, 345–350.
- Bae, J.A., Mu, S., Kim, J.S., Turner, N.L., Tartavull, I., Kemnitz, N., Jordan, C.S., Norton, A.D., Silversmith, W.M., Prentki, R., et al. (2018). Digital Museum of Retinal Ganglion Cells with Dense Anatomy and Physiology. *Cell* **173**, 1293–1306.
- Blanks, J.C., and Johnson, L.V. (1984). Specific binding of peanut lectin to a class of retinal photoreceptor cells. A species comparison. *Invest. Ophthalmol. Vis. Sci.* **25**, 546–557.
- Bourne, R.R., Stevens, G.A., White, R.A., Smith, J.L., Flaxman, S.R., Price, H., Jonas, J.B., Keeffe, J., Leasher, J., Naidoo, K., et al.; Vision Loss Expert Group (2013). Causes of vision loss worldwide, 1990–2010: a systematic analysis. *Lancet Glob. Health* **1**, e339–e349.
- Bringmann, A., Syrbe, S., Görner, K., Kacza, J., Francke, M., Wiedemann, P., and Reichenbach, A. (2018). The primate fovea: Structure, function and development. *Prog. Retin. Eye Res.* **66**, 49–84.
- Brooks, M.J., Rajasimha, H.K., Roger, J.E., and Swaroop, A. (2011). Next-generation sequencing facilitates quantitative analysis of wild-type and *Nrl*(^{-/-}) retinal transcriptomes. *Mol. Vis.* **17**, 3034–3054.
- Butler, A., Hoffman, P., Smibert, P., Papalex, E., and Satija, R. (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420.
- Cajal, S.R. (1893). La rétine des vertébrés. *Cellule* **9**, 121–255.
- Chan, T.L., Martin, P.R., Clunas, N., and Grünert, U. (2001). Bipolar cell diversity in the primate retina: morphologic and immunocytochemical analysis of a new world monkey, the marmoset *Callithrix jacchus*. *J. Comp. Neurol.* **437**, 219–239.
- Chen, T., and Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (AMC)*, pp. 785–794.
- Cherry, T.J., Wang, S., Bormuth, I., Schwab, M., Olson, J., and Cepko, C.L. (2011). NeuroD factors regulate cell fate and neurite stratification in the developing retina. *J. Neurosci.* **31**, 7365–7379.
- Dacey, D.M. (2004). Origins of perception: retinal ganglion cell diversity and the creation of parallel visual pathways. In *The Cognitive Neurosciences, Third Edition*, M.S. Gazzaniga, ed. (Cambridge: MIT Press), pp. 281–301.
- Diamond, J.S. (2017). Inhibitory Interneurons in the Retina: Types, Circuitry, and Function. *Annu. Rev. Vis. Sci.* **3**, 1–24.
- Doan, T., Azevedo, A.W., Hurley, J.B., and Rieke, F. (2009). Arrestin competition influences the kinetics and variability of the single-photon responses of mammalian rod photoreceptors. *J. Neurosci.* **29**, 11867–11879.
- Euler, T., Haverkamp, S., Schubert, T., and Baden, T. (2014). Retinal bipolar cells: elementary building blocks of vision. *Nat. Rev. Neurosci.* **15**, 507–519.
- Farrar, G.J., Carrigan, M., Dockery, A., Millington-Ward, S., Palfi, A., Chadderton, N., Humphries, M., Kiang, A.S., Kenna, P.F., and Humphries, P. (2017). Toward an elucidation of the molecular genetics of inherited retinal degenerations. *Hum. Mol. Genet.* **26** (R1), R2–R11.
- Finak, G., McDavid, A., Yajima, M., Deng, J., Gersuk, V., Shalek, A.K., Slichter, C.K., Miller, H.W., McElrath, M.J., Prlic, M., et al. (2015). MAST: a flexible statistical framework for assessing transcriptional changes and characterizing heterogeneity in single-cell RNA sequencing data. *Genome Biol.* **16**, 278.
- Fortunato, S., and Barthélemy, M. (2007). Resolution limit in community detection. *Proc. Natl. Acad. Sci. USA* **104**, 36–41.
- Fritsche, L.G., Igl, W., Bailey, J.N., Grassmann, F., Sengupta, S., Bragg-Gresham, J.L., Burdon, K.P., Hebbaring, S.J., Wen, C., Gorski, M., et al. (2016). A large genome-wide association study of age-related macular degeneration highlights contributions of rare and common variants. *Nat. Genet.* **48**, 134–143.
- Gollisch, T., and Meister, M. (2010). Eye smarter than scientists believed: neural computations in circuits of the retina. *Neuron* **65**, 150–164.
- Graham, P.S., Kaidonis, G., Abhary, S., Gillies, M.C., Daniell, M., Essex, R.W., Chang, J.H., Lake, S.R., Pal, B., Jenkins, A.J., et al. (2018). Genome-wide association studies for diabetic macular edema and proliferative diabetic retinopathy. *BMC Med. Genet.* **19**, 71.
- Greiff, R. (1900). Die mikroskopische anatomie des sehner ven und der netzhaut. In *Handbuch der gesamten Augenheilkunde* **1**.
- Grünert, U., and Martin, P.R. (1991). Rod bipolar cells in the macaque monkey retina: immunoreactivity and connectivity. *J. Neurosci.* **11**, 2742–2758.
- Grünert, U., Martin, P.R., and Wässle, H. (1994). Immunocytochemical analysis of bipolar cells in the macaque monkey retina. *J. Comp. Neurol.* **348**, 607–627.
- Haghverdi, L., Lun, A.T.L., Morgan, M.D., and Marioni, J.C. (2018). Batch effects in single-cell RNA-sequencing data are corrected by matching mutual nearest neighbors. *Nat. Biotechnol.* **36**, 421–427.
- Haverkamp, S., and Wässle, H. (2004). Characterization of an amacrine cell type of the mammalian retina immunoreactive for vesicular glutamate transporter 3. *J. Comp. Neurol.* **468**, 251–263.
- Haverkamp, S., Haeseleer, F., and Hendrickson, A. (2003). A comparison of immunocytochemical markers to identify bipolar cell types in human and monkey retina. *Vis. Neurosci.* **20**, 589–600.
- Hendrickson, A., Yan, Y.H., Erickson, A., Possin, D., and Pow, D. (2007). Expression patterns of calretinin, calbindin and parvalbumin and their colocalization in neurons during development of Macaca monkey retina. *Exp. Eye Res.* **85**, 587–601.
- Hofer, H., Carroll, J., Neitz, J., Neitz, M., and Williams, D.R. (2005). Organization of the human trichromatic cone mosaic. *J. Neurosci.* **25**, 9669–9679.
- Hoshino, A., Ratnapriya, R., Brooks, M.J., Chaitankar, V., Wilken, M.S., Zhang, C., Starostik, M.R., Gieser, L., La Torre, A., Nishio, M., et al. (2017). Molecular Anatomy of the Developing Human Retina. *Dev. Cell* **43**, 763–779.
- Johnson, T.V., and Martin, K.R. (2008). Development and characterization of an adult retinal explant organotypic tissue culture system as an in vitro intraocular stem cell transplantation model. *Invest. Ophthalmol. Vis. Sci.* **49**, 3503–3512.
- Kay, J.N., Voinescu, P.E., Chu, M.W., and Sanes, J.R. (2011). Neurod6 expression defines new retinal amacrine cell subtypes and regulates their fate. *Nat. Neurosci.* **14**, 965–972.

- Kent, W.J. (2002). BLAT—the BLAST-like alignment tool. *Genome Res.* *12*, 656–664.
- Klump, K.E., Zhang, A.J., Wu, S.M., and Marshak, D.W. (2009). Parvalbumin-immunoreactive amacrine cells of macaque retina. *Vis. Neurosci.* *26*, 287–296.
- Kolb, H., and Marshak, D. (2003). The midget pathways of the primate retina. *Doc. Ophthalmol.* *106*, 67–81.
- Kolesnikov, A.V., Rikimaru, L., Hennig, A.K., Lukasiewicz, P.D., Fliesler, S.J., Govardovskii, V.I., Kefalov, V.J., and Kisselev, O.G. (2011). G-protein beta-gamma-complex is crucial for efficient signal amplification in vision. *J. Neurosci.* *31*, 8067–8077.
- Krishnaswamy, A., Yamagata, M., Duan, X., Hong, Y.K., and Sanes, J.R. (2015). Sidekick 2 directs formation of a retinal circuit that detects differential motion. *Nature* *524*, 466–470.
- Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., and Weirauch, M.T. (2018). The Human Transcription Factors. *Cell* *172*, 650–665.
- Lammerding-Köppel, M., Thier, P., and Koehler, W. (1991). Morphology and mosaics of VIP-like immunoreactive neurons in the retina of the rhesus monkey. *J. Comp. Neurol.* *312*, 251–263.
- Lassová, L., Fina, M., Sulaiman, P., and Vardi, N. (2010). Immunocytochemical evidence that monkey rod bipolar cells use GABA. *Eur. J. Neurosci.* *31*, 685–696.
- Lee, B.B., Shapley, R.M., Hawken, M.J., and Sun, H. (2012). Spatial distributions of cone inputs to cells of the parvocellular pathway investigated with cone-isolating gratings. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* *29*, A223–A232.
- Liao, H.W., Ren, X., Peterson, B.B., Marshak, D.W., Yau, K.W., Gamin, P.D., and Dacey, D.M. (2016). Melanopsin-expressing ganglion cells on macaque and human retinas form two morphologically distinct populations. *J. Comp. Neurol.* *524*, 2845–2872.
- Linderman, G.C., Rachh, M., Hoskins, J.G., Steinerberger, S., and Kluger, Y. (2017). Efficient Algorithms for t-distributed Stochastic Neighborhood Embedding. *arXiv*, arXiv:1712.09005, <https://arxiv.org/abs/1712.09005>.
- Liu, J., Reggiani, J.D.S., Laboulaye, M.A., Pandey, S., Chen, B., Rubenstein, J.L.R., Krishnaswamy, A., and Sanes, J.R. (2018). Tbr1 instructs laminar patterning of retinal ganglion cell dendrites. *Nat. Neurosci.* *21*, 659–670.
- Macosko, E.Z., Basu, A., Satija, R., Nemes, J., Shekhar, K., Goldman, M., Tirosh, I., Bialas, A.R., Kamitaki, N., Martersteck, E.M., et al. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* *161*, 1202–1214.
- Majumdar, S., Wässle, H., Jusuf, P.R., and Haverkamp, S. (2008). Mirror-symmetrical populations of wide-field amacrine cells of the macaque monkey retina. *J. Comp. Neurol.* *508*, 13–27.
- Mao, C.A., Li, H., Zhang, Z., Kiyama, T., Panda, S., Hattar, S., Ribelayga, C.P., Mills, S.L., and Wang, S.W. (2014). T-box transcription regulator Tbr2 is essential for the formation and maintenance of Opn4/melanopsin-expressing intrinsically photosensitive retinal ganglion cells. *J. Neurosci.* *34*, 13083–13095.
- Mariani, A.P. (1984). Bipolar cells in monkey retina selective for the cones likely to be blue-sensitive. *Nature* *308*, 184–186.
- Mariani, A.P., and Hokoc, J.N. (1988). Two types of tyrosine hydroxylase-immunoreactive amacrine cell in the rhesus monkey retina. *J. Comp. Neurol.* *276*, 81–91.
- Marioni, J.C., and Arendt, D. (2017). How Single-Cell Genomics Is Changing Evolutionary and Developmental Biology. *Annu. Rev. Cell Dev. Biol.* *33*, 537–553.
- Masland, R.H. (2012). The neuronal organization of the retina. *Neuron* *76*, 266–280.
- Masri, R.A., Percival, K.A., Koizumi, A., Martin, P.R., and Grünert, U. (2019). Survey of retinal ganglion cell morphology in marmoset. *J. Comp. Neurol.* *527*, 236–258.
- Meyer-Franke, A., Kaplan, M.R., Pfrieger, F.W., and Barres, B.A. (1995). Characterization of the signaling interactions that promote the survival and growth of developing retinal ganglion cells in culture. *Neuron* *15*, 805–819.
- Mills, S.L., and Massey, S.C. (1999). All amacrine cells limit scotopic acuity in central macaque retina: A confocal analysis of calretinin labeling. *J. Comp. Neurol.* *411*, 19–34.
- Mitchell, J.F., Reynolds, J.H., and Miller, C.T. (2014). Active vision in marmosets: a model system for visual neuroscience. *J. Neurosci.* *34*, 1183–1194.
- Nathans, J., Thomas, D., and Hogness, D.S. (1986). Molecular genetics of human color vision: the genes encoding blue, green, and red pigments. *Science* *232*, 193–202.
- Ng, L., Hurley, J.B., Dierks, B., Srinivas, M., Saltó, C., Vennström, B., Reh, T.A., and Forrest, D. (2001). A thyroid hormone receptor that is required for the development of green cone photoreceptors. *Nat. Genet.* *27*, 94–98.
- Onishi, A., Koike, S., Ida-Hosonuma, M., Imai, H., Shichida, Y., Takenaka, O., Hanazawa, A., Komatsu, H., Mikami, A., Goto, S., et al. (2002). Variations in long- and middle-wavelength-sensitive opsin gene loci in crab-eating monkeys. *Vision Res.* *42*, 281–292.
- Pandey, S., Shekhar, K., Regev, A., and Schier, A.F. (2018). Comprehensive Identification and Spatial Mapping of Habenular Neuronal Types Using Single-Cell RNA-Seq. *Curr. Biol.* *28*, 1052–1065.
- Park, S.J., Borghuis, B.G., Rahmani, P., Zeng, Q., Kim, I.J., and Demb, J.B. (2015). Function and Circuitry of VIP+ Interneurons in the Mouse Retina. *J. Neurosci.* *35*, 10685–10700.
- Peng, Y.R., Tran, N.M., Krishnaswamy, A., Kostadinov, D., Martersteck, E.M., and Sanes, J.R. (2017). Satb1 Regulates Contactin 5 to Pattern Dendrites of a Mammalian Retinal Ganglion Cell. *Neuron* *95*, 869–883.
- Perry, V.H., and Silveira, L.C. (1988). Functional lamination in the ganglion cell layer of the macaque's retina. *Neuroscience* *25*, 217–223.
- Perte, M., Kim, D., Perte, G.M., Leek, J.T., and Salzberg, S.L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* *11*, 1650–1667.
- Provis, J.M., Dubis, A.M., Maddess, T., and Carroll, J. (2013). Adaptation of the central retina for high acuity vision: cones, the fovea and the avascular zone. *Prog. Retin. Eye Res.* *35*, 63–81.
- Rheume, B.A., Jereen, A., Bolisetty, M., Sajid, M.S., Yang, Y., Renna, K., Sun, L., Robson, P., and Trakhtenberg, E.F. (2018). Single cell transcriptome profiling of retinal ganglion cells identifies cellular subtypes. *Nat. Commun.* *9*, 2759.
- Rodríguez, A.R., de Sevilla Müller, L.P., and Brecha, N.C. (2014). The RNA binding protein RBPMS is a selective marker of ganglion cells in the mammalian retina. *J. Comp. Neurol.* *522*, 1411–1443.
- Roorda, A., Metha, A.B., Lennie, P., and Williams, D.R. (2001). Packing arrangement of the three cone classes in primate retina. *Vision Res.* *41*, 1291–1306.
- Rosvall, M., and Bergstrom, C.T. (2008). Maps of random walks on complex networks reveal community structure. *Proc. Natl. Acad. Sci. USA* *105*, 1118–1123.
- Rouso, D.L., Qiao, M., Kagan, R.D., Yamagata, M., Palmiter, R.D., and Sanes, J.R. (2016). Two Pairs of ON and OFF Retinal Ganglion Cells Are Defined by Intersectional Patterns of Transcription Factor Expression. *Cell Rep.* *15*, 1930–1944.
- Sanes, J.R., and Masland, R.H. (2015). The types of retinal ganglion cells: current status and implications for neuronal classification. *Annu. Rev. Neurosci.* *38*, 221–246.
- Shekhar, K., Lapan, S.W., Whitney, I.E., Tran, N.M., Macosko, E.Z., Kowalczyk, M., Adiconis, X., Levin, J.Z., Nemes, J., Goldman, M., et al. (2016). Comprehensive Classification of Retinal Bipolar Neurons by Single-Cell Transcriptomics. *Cell* *166*, 1308–1323.
- Sinha, R., Hoon, M., Baudin, J., Okawa, H., Wong, R.O., and Rieke, F. (2017). Cellular and Circuit Mechanisms Shaping the Perceptual Properties of the Primate Fovea. *Cell* *168*, 413–426.

- Stella, S.L., Jr., Li, S., Sabatini, A., Vila, A., and Brecha, N.C. (2008). Comparison of the ontogeny of the vesicular glutamate transporter 3 (VGLUT3) with VGLUT1 and VGLUT2 in the rat retina. *Brain Res.* 1215, 20–29.
- Sweeney, N.T., James, K.N., Nistorica, A., Lorig-Roach, R.M., and Feldheim, D.A. (2019). Expression of transcription factors divides retinal ganglion cells into distinct classes. *J. Comp. Neurol.* 527, 225–235.
- Tsukamoto, Y., and Omi, N. (2015). OFF bipolar cells in macaque retina: type-specific connectivity in the outer and inner synaptic layers. *Front. Neuroanat.* 9, 122.
- Tsukamoto, Y., and Omi, N. (2016). ON Bipolar Cells in Macaque Retina: Type-Specific Synaptic Connectivity with Special Reference to OFF Counterparts. *Front. Neuroanat.* 10, 104.
- Vecino, E., Rodríguez, F.D., Ruzafa, N., Pereiro, X., and Sharma, S.C. (2016). Glia-neuron interactions in the mammalian retina. *Prog. Retin. Eye Res.* 51, 1–40.
- Wang, Y., Smallwood, P.M., Cowan, M., Blesh, D., Lawler, A., and Nathans, J. (1999). Mutually exclusive expression of human red and green visual pigment-reporter transgenes occurs at high frequency in murine cone photoreceptors. *Proc. Natl. Acad. Sci. USA* 96, 5251–5256.
- Wässle, H., Grünert, U., Martin, P.R., and Boycott, B.B. (1994). Immunocytochemical characterization and spatial distribution of midget bipolar cells in the macaque monkey retina. *Vision Res.* 34, 561–579.
- Wässle, H., Dacey, D.M., Haun, T., Haverkamp, S., Grünert, U., and Boycott, B.B. (2000). The mosaic of horizontal cells in the macaque monkey retina: with a comment on biplexiform ganglion cells. *Vis. Neurosci.* 17, 591–608.
- Wiggs, J.L., and Pasquale, L.R. (2017). Genetics of glaucoma. *Hum. Mol. Genet.* 26 (R1), R21–R27.
- Wikler, K.C., Williams, R.W., and Rakic, P. (1990). Photoreceptor mosaic: number and distribution of rods and cones in the rhesus monkey retina. *J. Comp. Neurol.* 297, 499–508.
- Wool, L.E., Crook, J.D., Troy, J.B., Packer, O.S., Zaidi, Q., and Dacey, D.M. (2018). Nonselective Wiring Accounts for Red-Green Opponency in Midget Ganglion Cells of the Primate Retina. *J. Neurosci.* 38, 1520–1540.
- Yamada, E.S., Dmitrieva, N., Keyser, K.T., Lindstrom, J.M., Hersh, L.B., and Marshak, D.W. (2003). Synaptic connections of starburst amacrine cells and localization of acetylcholine receptors in primate retinas. *J. Comp. Neurol.* 461, 76–90.
- Zeitl, C., Robson, A.G., and Audo, I. (2015). Congenital stationary night blindness: an analysis and update of genotype-phenotype correlations and pathogenic mechanisms. *Prog. Retin. Eye Res.* 45, 58–110.
- Zhang, D.Q., Zhou, T.R., and McMahon, D.G. (2007). Functional heterogeneity of retinal dopaminergic neurons underlying their multiple roles in vision. *J. Neurosci.* 27, 692–699.
- Zheng, G.X., Terry, J.M., Belgrader, P., Ryvkin, P., Bent, Z.W., Wilson, R., Ziraldo, S.B., Wheeler, T.D., McDermott, G.P., Zhu, J., et al. (2017). Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* 8, 14049.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Chicken polyclonal anti-GFP	Abcam	CAT#ab13970; RRID:AB_300798
Rabbit polyclonal anti-GFP	Millipore	Millipore Cat# AB3080P, RRID: AB_2630379
Mouse polyclonal anti-Satb2	Abcam	Cat# ab51502; RRID: AB_882455
Goat polyclonal anti-Choline Acetyltransferase	Millipore	CAT#AB144P; RRID: AB_11214092
Goat polyclonal anti-Vesicular Acetylcholine Transporter	Millipore	CAT#ABN100; RRID: AB_2630394
Guinea polyclonal anti-Rbpms	PhosphoSolutions	CAT# 1832-RBPMS; RRID: AB_2492226
Rabbit polyclonal anti-Rbpms	Abcam	CAT# 194213; RRID: AB_648803
Rabbit polyclonal anti-Calbindin	Swant	CAT# CB38a
Mouse monoclonal anti-Calretinin	Millipore	CAT# MAB1568; RRID: AB_94259
Goat polyclonal anti-osteopontin (SPP1)	R&D Systems	Cat# AF808; RRID: AB_2194992
Rabbit polyclonal anti-RPKCa	Sigma	CAT# P4334; RRID: AB_477345
Mouse monoclonal anti-PKCa	Abcam	CAT# ab31; RRID: AB_303507
Rabbit polyclonal anti-Secretagoin	BioVendor	CAT# RD181120100; RRID: AB_10720931
Mouse anti-Human CD15	BD Biosciences	CAT# 559045; RRID: AB_397181
anti-human monoclonal CD90 microbeads	MACS Miltenyi Biotec	CAT# 130-096-253
anti-mouse IgG1 microbeads	MACS Miltenyi Biotec	CAT# 130-047-102; RRID: AB_244355
Mouse anti-Human CD73 Clone AD2	BD Biosciences	CAT# 550256; RRID: AB_393560
Chemicals, Peptides, and Recombinant Proteins		
AMES' Medium	Sigma	CAT# A1420
Papain	Worthington	CAT# LS003126
Ovomucoid	Worthington	CAT#LS003087
Large cell column	MACS Miltenyi Biotec	CAT# 130042202
AzuraQuant cDNA synthesis kit	Azura	CAT# AZ-1995;
Recombinant Human BDNF	PepruTech	CAT# AF-450-02
Recombinant Human CNTF	PepruTech	CAT# AF-450-13
Critical Commercial Assays		
Chromium Single Cell 3' Library & Gel Bead Kit v2, 16 rxns	10X Genomics	CAT# 120237
Chromium Single Cell A Chip Kit, 16 rxns	10X Genomics	CAT# 1000009
Chromium i7 Multiplex Kit 96 rxns	10X Genomics	CAT# 120262
Deposited Data		
Raw data files for scRNA-seq of macaque retinal cells	This study	GEO:GSE118480
Raw data files for Tru-Seq RNA-sequencing	This study	SRA: PRJNA514024
Raw data files for scRNA-seq of human retinal cells	This study	DUOS-000105
Recombinant DNA		
AAVrg-CAG-GFP(Addgene) (titer: 1e12)	Addgene	CAT#37825
Software and Algorithms		
StringTie v1.3.3		https://ccb.jhu.edu/software/stringtie/
Hisat2		https://ccb.jhu.edu/software/hisat2/index.shtml

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Cell Ranger v2.1.0	10X genomics	https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest
IGV		https://www.broadinstitute.org/igv/ ; RRID:SCR_011793
ImageJ		http://ezp-prod1.hul.harvard.edu/login?url=http://rsb.info.nih.gov/2fij%2findex.html ; RRID:SCR_003070
Pairwise stitching ImageJ plugin		https://imagej.net/Image_Stitching
Segmentation Simple Neurite Tracer ImageJ plugin		https://imagej.net/Simple_Neurite_Tracer
Interactive Stack Rotation ImageJ plugin	Stephan Saafeld	https://imagej.net/Interactive_Stack_Rotation
Zen	Zeiss	RRID:SCR_013672 https://www.zeiss.com/microscopy/en_us/products/microscope-software/zen.html

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Dr. Joshua R. Sanes (sanesj@mcb.harvard.edu).

EXPERIMENTAL MODELS AND SUBJECT DETAILS

Tissue Procurement

Non-human primate and human retinas were obtained and used in accordance with the guidelines for the care and use of animals and human subjects at Harvard University and Boston Children’s Hospital, and supplying institutions. All the procedures on non-human primates were approved by the Institutional Animal Care and Use Committees. Acquisition and use of human tissue was approved by the Human Study Subject Committees (DFCI Protocol Number: 13-416 and MEE - NHR Protocol Number 18-034H).

Eyes from male macaques (*Macaca fascicularis*, 3–9 years of age) were kindly provided by institutions including Biomere and Massachusetts General Hospital. Eyes were collected either pre-mortem under deep anesthesia or ≤ 45 min post-mortem. In some cases, whole globes were immediately placed in ice-cold Ames solution (Sigma-Aldrich; equilibrated with 95% O₂/5% CO₂ for all use), where they were stored before experimentation. In others, a rapid hemisection was performed to remove the vitreous and the anterior chamber, and the posterior eyecup was immersed in room-temperature Ames. Retinas were then dissected free and stored in ice-cold Ames solution. Experiments commenced within 8 hours. For access to the macaque samples used in this study, we thank V. Belov, C. Cetrulo, D. Guberski, A. Hall, P. Kovalenko, A. LaRochelle, J. Madsen, M. Nedelman, M. Papisov, and S. Smith.

Eyes from male and female common marmosets (*Callithrix jacchus*, 2–10 years of age) were generously provided by the McGovern Institute for Brain Research (Massachusetts Institute of Technology). Animals were perfused with 4% paraformaldehyde (PFA) under deep anesthesia. Eyes were then collected, post-fixed for 1hr in 4% PFA, and stored in ice-cold Ames solution. We are gratefully to G. Feng, Q. Zhang and C. Wu for access to this tissue.

Macaque and marmoset eyes were collected from animals that had reached the end of unrelated studies at supplying institutions. No ocular or visual abnormalities were noted. Data presented in this manuscript did not covary with any treatment that had been applied to the animals.

The human eye used for sequencing and *in situ* hybridization was collected ~6 hours postmortem from a 74 year-old male. The whole globe was immediately transported back to the lab in a humid chamber. Hemisection was performed to remove the anterior chamber and the posterior pole was recovered in Ames equilibrated with 95% O₂/5% CO₂ before further dissection and dissociation. The donor was confirmed to have no history or clinical evidence of ocular disease or intraocular surgery. We are grateful to Dr. Juric and the Rapid Autopsy Program, Susan Eid Tumor Heterogeneity Initiative, Massachusetts General Hospital for expeditious access to this material, enabling recovery of high quality RNA for sequencing.

Other human eyes used for *in situ* hybridization were provided by the Lions VisionGift (Portland, OR) and the National Disease Research Interchange (NDRI; Philadelphia, PA). They were collected < 11 hr postmortem from a 28 year-old male and a 69 year-old female, hemisected, and fixed overnight in ice-cold 4% PFA following removal of the cornea. No ocular disease was reported in these donors.

METHOD DETAILS

RNA-sequencing

Single Cell Isolation

0.5–1.5 mm diameter foveal tissues centering on the foveal pit were dissected out from four macaque eyes. Foveal samples were digested with 200 units papain (Worthington, LS003126). Foveae M1 and M2 were digested at 37°C; foveae M3 and M4 were

digested at 22°C. Following digestion, retinas were dissociated and triturated into single cell suspensions with 0.04% bovine serum albumin (BSA) in Ames solution. Peripheral retinal pieces were dissected and pooled from all quadrants of the retina. Single cell suspensions were dissociated at 37°C as described for fovea. Dissociated cells were incubated with CD90 microbeads (Miltenyi Biotec, 130-096-253; 1 μ l per 10⁷ cells) to enrich RGCs or with anti-CD73 (BD Biosciences, clone AD2; 5 μ l per 10⁷ cells) followed by anti-mouse IgG1 microbeads (Miltenyi Biotec, 130-047-102; 10 μ l per 10⁷ cells) to deplete rods. Incubations were at room temperature for 10 min. CD90 positive cells or CD73 negative cells were selected via large cell columns through a MiniMACS Separator (Miltenyi Biotec). Single cell suspensions were diluted at a concentration of 500-1800 cells/ μ L in 0.04% BSA/Ames for loading into 10X Chromium Single Cell A Chips. Human cells were dissociated and treated with anti-CD73 as above.

Droplet-Based scRNA-seq

Single cell libraries were prepared using the Chromium 3' v2 platform (10X Genomics, Pleasanton, CA) following the manufacturer's protocol. Briefly, single cells were partitioned into Gel beads in EMulsion (GEMs) in the GemCode instrument followed by cell lysis and barcoded reverse transcription of RNA, amplification, shearing and 5' adaptor and sample index attachment. On average, approximately 10,000 single cells were loaded on each channel and approximately 6,000 cells were recovered. Libraries were sequenced on the Illumina HiSeq 2500 (Paired end reads: Read 1, 26bp, Read 2, 98bp).

Histological methods

Viral Transduction

Macaque or marmoset retinas were divided into four quadrants according to cardinal axes with the macula in the temporal region and mounted on Millicell culture inserts (Millipore) with the ganglion cell layer facing up. Retinas were infected with 5 μ l of AAV retrograde virus encoding GFP directed by the CAG promoter (Addgene) with a titer \sim 10e12. Methods were adapted from (Johnson and Martin, 2008; Meyer-Franke et al., 1995). Transduced retinas were cultured in Neurobasal A medium (Fisher 10888-022) supplemented with 100 units/mL penicillin and 0.1 mg/mL streptomycin (Sigma P4333), 1 mM L-glutamine (Sigma G7513), B27 (Fisher 17504-044), 0.005 mM forskolin (Sigma F6886), 1 mM sodium pyruvate (Fisher 11360-070), 25 ng/mL BDNF (PeproTech 450-02), and 10 ng/mL CNTF (PeproTech 450-13) for four days. Retinas were then fixed in 4% PFA in phosphate-buffered saline (PBS) for 1 hr at 4°C and separated from the culture insert after fixation.

Biolistic Transfection

The biolistics procedure was adapted from a published protocol (Masri et al., 2019). Briefly, gold particles (1.6 μ m diameter; 12.5 mg; Bio-Rad) were coated with pCMV-GFP DNA plasmid (25 μ g). The particles were delivered to retinal cells in whole-mount retinas with GCL facing up using a Helios gene gun (Bio-Rad). The retinal explants were cultured for 3 days followed by fixation as described for virus transduced retinas.

Fluorescent In Situ Hybridization (FISH)

Eyes were fixed in 4% PFA. Marmoset and human eyes were fixed by perfusion and immersion respectively as described under "Tissue Procurement." For macaque eyes, slits or windows were made in the cornea of eyes collected post-mortem, and the globe was immersion-fixed overnight in ice-cold 4% PFA. Probe templates were generated using cDNA derived from adult macaque, marmoset, or human retinas following RNA extraction and reverse transcription with AzuraQuant cDNA Synthesis Kit (Azura, AZ-1995). Antisense probes were generated by PCR using a reverse primer with a T7 sequence adaptor to permit *in vitro* transcription (see Table S4 for primer sequences). DIG rUTP (Roche, 11277073910), DNP rUTP (Perkin Elmer, NEL555001EA), and Fluorescein rUTP (Roche, 11685619910) were used for synthesizing probes for single, double or triple fluorescent *in situ* hybridization (FISH) experiments. Retinas fixed as above were rinsed with PBS, immersed with 30% sucrose overnight at 4°C, embedded in Tissue Freezing Medium (EMS) and cryosectioned at 20 μ m. Details of FISH were described in Shekhar et al. (2016). Briefly, sections were mounted on Superfrost slides (Thermo Scientific), treated with 1.5 μ g/mL of proteinase K (NEB, P8107S), and then post-fixed and treated with acetic anhydride for deacetylation. Probe detection was performed with anti-DIG HRP (1:1000), anti-DNP HRP (1:500), anti-Fluorescein (1:1000) followed by tyramide amplification. Detection of protein epitopes was performed following probe detection. Antibodies were diluted in block consisting of 3% donkey serum (Jackson, 017-000-121), and 0.3% Triton-X in PBS at concentrations of 1:500 (anti-GFP), 1:1000 (anti-Calbindin), 1:2000 (anti-PKC α), and 1:100 (anti-CD15) (See below for antibody information).

Immunohistochemistry

Tissue was fixed as described above. Immunostaining was performed as described in Shekhar et al. (2016) and Peng et al. (2017). Antibodies used were as follows: chick and rabbit anti-GFP (1:500, Abcam; 1:2000, Millipore); mouse anti-Satb2(1:1000, Abcam); goat anti anti-CHX10 (1:300, Santa Cruz); rabbit anti-TFAP2B (1:1000, cell signaling); goat anti-Choline Acetyltransferase (1:500, Millipore); goat anti-VACHT (1:1000, Millipore); rabbit and guinea pig anti-RBPMS (1:1000, Abcam; 1:5000, PhosphoSolutions); rabbit anti-Calbindin (1:2000, Swant); mouse anti-Calretinin (1:5000, Millipore); mouse anti-Gad65/67 (1:1000, Millipore); mouse and rabbit anti-PKC α (1:2000, Abcam; 1:2000, sigma); rabbit anti- Secretagogin (1:10,000; BioVendor); mouse anti-Human CD15 (1:100, BD); rabbit anti-GNGT1 (1:1000, gift from Dr. N. Gautam), mouse anti-7G6 (1:1000, gift from Dr. Peter MacLeish), and chick anti-OPN1M/LW (1:2000, gift from Dr. Jeremy Nathans). Nuclei were labeled with DAPI (1:1000, Invitrogen). Secondary antibodies were conjugated to Alexa Fluor 488, 568, and 647 (Invitrogen) and used at 1:1000. Sections and whole mounts were mounted in ProLong Gold Antifade (Invitrogen).

Image Acquisition, Processing and Analysis

Images were acquired on Zeiss LSM 710 confocal microscopes with 405, 488-515, 568, and 647 lasers, processed using Zeiss ZEN software suites, and analyzed using ImageJ (NIH). Images were acquired with 16X, 40X or 63X oil lens at the resolution of 1024x1024 pixels, a step size of 0.5-1.5 μ m, and 90 μ m pinhole size. ImageJ (NIH) software was used to generate maximum intensity projections and neuronal dendrites were reconstructed with ImageJ plugin simple neurite tracer. Adobe Photoshop CC was used for adjustments to brightness and contrast.

Computational Methods

Assembling a Retina-specific Transcriptome for *M. fascicularis*

We obtained high quality total RNA from macaque retinal tissue (RNA Integrity Number (RIN) score 9.8), and prepared strand-specific libraries using the TruSeq strand-specific Total RNA kit (Illumina Inc.), which was sequenced on the NextSeq 500 system to obtain 45 million 100bp paired end reads. We used StringTie (v1.3.3) (Pertea et al., 2016) to assemble a genome-guided transcriptome, using an available NCBI transcriptome for *M. fascicularis* (annotation release 101; <https://ncbi.nlm.nih.gov/genome/776>) as an initial guide. Briefly, we mapped RNA-seq reads onto the existing transcriptome in a strand-specific manner using the Hisat2 software (with command line options `-dta-rna-strandedness RF`), following published guidelines (Pertea et al., 2016). Next, we used StringTie (with command line option `-rf`) to assemble a new transcriptome based on TruSeq reads. We then reran StringTie (with the command line option `merge`) to obtain an updated transcriptome annotation, which contained modifications of transcript body definitions from the existing NCBI reference as well as novel transcripts supported by the TruSeq reads. While the modified transcripts retained gene names from the original annotation, the novel transcripts were initially named according to Stringtie's naming convention (e.g., MSTRG.5141). We call this the NCBI + TruSeq reference in Figure S1C. There were a few instances where sense and antisense transcripts were predicted for some genes; in these cases we added suffixes “_p” and “_n” against the corresponding gene name to indicate this fact (e.g., *FEZF1_p* and *FEZF1_n*).

To facilitate transcriptional mapping of macaque types to mouse and human types, we used BLAT (Kent, 2002) to associate each macaque transcript with its closest ortholog (reciprocal best matches) in the NCBI mouse (version GRCh38) and human (version GRCh37) transcriptomes, respectively. Encouragingly, this automated procedure matched transcript gene names with similar letter codes (e.g., macaque *TRPM1* with mouse *Trpm1* and human *TRPM1*). We use this mapping to associate each reconstructed transcript with the closest human gene by sequence, in cases where such homology was strong. In some cases, we also found that certain tenuously named loci in the NCBI reference could be associated with homologous mouse and human gene names (e.g., macaque *LOC101864869* could be associated with mouse *Rps27a*, *LOC102132859* could be associated with mouse *Cyp26a1* etc). For many loci, however, we were unable to associate a human gene and for these cases, we retained the naming convention assigned by StringTie (e.g., MSTRG.5141).

Pre-Processing Of 3' Droplet-based scRNA-Seq Data

Sample demultiplexing, alignment to the NCBI+TruSeq reference, quantification and initial quality control (QC) was performed using the Cell Ranger software (version 2.1.0, 10X Genomics) for each sample (i.e., 10X channel) separately. We used the option “`-force-cells 6000`” in Cell Ranger count to obtain a 36,162 genes x 6,000 cells count matrix each sample to deliberately extract a larger number of cell barcodes in the data, as we found that the automatic estimate of Cell Ranger was too conservative, and was unfavorable for small cell-classes like BCs. Here, 6,000 represented a “loose” upper bound on the number of cells that could be recovered given the density of each the cell suspension loaded onto every channel per the manufacturer's estimates. We grouped the count matrices separately for the foveal and the peripheral samples to generate a consolidated matrix for each region. Only cells which expressed > 500 genes were retained for further analysis. Further pruning of cells (low quality cells, doublets) was done for each class separately (see below). Library size, as quantified by the number of genes detected per cell on average, was highly cell-class dependent. Values by class, from highest to lowest, were: RGCs, 4000 genes/cell; ACs, 1500 genes/cell; cones, 1400 genes/cell; non-neuronal cells, 1200 genes/cell; HCs, 1000 genes/cell; rods, 900 genes/cell; and BCs, 750 genes/cell).

Separation of Major Cell Classes

Both the foveal and the peripheral datasets consisted of multiple biological replicates (Figures S1G–S1I). Moreover, depending on the enrichment method, each sample contained widely different distributions of the main cell classes. For example, the foveal samples typically comprised of ~30% RGCs, ~28% BCs, ~8% ACs (Figure S1H), the peripheral CD90+ samples contained ~28% RGCs, 0% BCs, ~50% ACs, and the peripheral CD73- samples contained < 1% RGCs, ~30% BCs and ~10% ACs (Figure S1I). Moreover, we anticipated that the differences in proportions also extends to types within a class (Figure 3), and that foveal and peripheral types likely had molecular differences (Figure 4). This made conventional batch correction using methods such as Surrogate Variable Analysis and ComBat difficult, because they make strict distributional assumptions that are violated here. Furthermore, covariates that can be used to counteract these assumptions are not known *a priori*. We also attempted to perform global batch correction using recently published methods that use Canonical Correlation Analysis (CCA; (Butler et al., 2018)) or Mutual Nearest Neighbor (MNN; (Haghverdi et al., 2018)) matching, but in many cases these did not completely remove batch effects, collapsed distinct types with known molecular markers, or appeared to have varying impact on different cell classes.

To make batch correction more tractable and to avoid extensive biases in the initial stages for the reasons stated above, we first analyzed the foveal and the peripheral datasets separately to stratify the cells by their class. We followed our earlier PCA + Louvain-Jaccard graph-clustering pipeline (Shekhar et al., 2016) without batch correction to obtain a set of transcriptionally distinct clusters.

At this stage, we deliberately set parameters to “overcluster” the data to avoid combining distinct cell classes. Next, we scored each cell for gene signatures of well-known retinal cell classes - Rods, Cones, Bipolar Cells (BC), Horizontal Cells (HC), Amacrine Cells (AC), Retinal Ganglion Cells (RGC), and non-neuronal cells such as Müller Glia (MG) and others (Table S1). Classes with shared markers often scored similarly, but with one higher than the other (e.g., *TFAP2A/2B* and *ONECUT1/2* are expressed in ACs and HCs, with *TFAP2A/2B* being higher in the former and *ONECUT1/2* being higher in the latter). In such cases, we combined these classes and analyzed them as a group in downstream pipelines. Thus, we separated the foveal and peripheral datasets into 5 groups comprising - (i) Rods and Cones, (ii) Amacrine Cells (ACs) and Horizontal Cells (HCs), (iii) Bipolar Cells (BCs), (iv) Retinal Ganglion Cells (RGCs), and (v) Others consisting predominantly of Müller Glia, but also pericytes, endothelial cells and microglia (Figure 1C). Astrocytes are absent from the fovea and are largely confined to the optic fiber layer in the periphery (Vecino et al., 2016). None were recovered from either region in our samples. Encouragingly, cells of the same class were more highly correlated with each other than cells of other classes, supporting this separation (Figure S1F).

Although clusters at this stage exhibited batch effects, ~90% of clusters could be unequivocally assigned to a class. The other clusters comprised low quality cells and “doublets.” These were often characterized by overall low gene count that was tightly distributed around the minimum cutoff value (~500-700 genes/cell), the mixed expression of more than one class- specific scores albeit at lower levels than “pure” clusters, and the absence of expression of non- class markers that were specific. Moreover, these cells typically clustered proximally on the dendrogram or tSNE visualization to another large, bonafide cluster, as observed previously (Shekhar et al., 2016). To avoid discarding cells erroneously, we analyzed these clusters separately (not shown), and confirmed that they were doublets and low-quality cells. Because the excluded cells were heavily biased toward low gene-count cells, we could have excluded them *ab initio* by setting a more stringent threshold, but this would have risked excluding genuinely small, but intact single- cell libraries.

Dimensionality Reduction, Clustering and Visualization For Each Foveal and Peripheral Group

Following the initial clustering step, we analyzed the 6 major classes of cells - RGCs, BCs, ACs, HCs, PRs, and non-neuronal cells for fovea and periphery separately (Figures 1D–1I). HCs and ACs were separated for this analysis. We employed the following computational steps to cluster each of these groups as described below.

1. Normalization: Expression values $E_{i,j}$ for gene i in cell j were calculated following (Shekhar et al., 2016). Briefly, the Cell Ranger reported UMI (i.e., transcript) count value for each gene i in each cell j was divided by the sum of the total UMI counts in cell j to normalize for differences in library size, and then multiplied by M , the median UMI counts for all cells within the group, resulting in Transcripts-per-median ($TPM_{i,j}$) values. $E_{i,j}$ was then calculated as $\log(TPM_{i,j} + 1)$
2. Identification of highly variable genes (HVGs): We first calculated the mean (μ) and coefficient of variation (CV) of transcript counts for each gene in the data. We then computed for each gene, the “excess CV” (or eCV) by subtracting the observed CV from a null model of CV versus μ . This null model is based on a Poisson-Gamma mixture, which was shown to accurately model the null CV versus μ relationships for data containing UMIs for a wide range of 3'-biased protocols (Pandey et al., 2018). We calculated the mean and standard deviation (SD) of the eCV values, and selected genes (HVGs) that had $eCV > \text{mean} + 0.7 \cdot \text{SD}$. This typically selected 700-2500 HVGs for each group.
3. Batch correction: We restricted the expression matrix $E_{i,j}$ to HVGs and used a linear regression model (adapted from the source code of the ‘RegressOut’ function of the R package ‘Seurat’) to remove correlations of expression values associated with three covariates – the animal of origin (Figure S1G), the total number of genes observed per cell, and the expression strength of ribosomal genes within each cell. The choice of covariates was guided by running initial tests on PRs, BCs and RGCs where some prior knowledge of the underlying types exists. We tested different combinations of covariates, and settled on a combination that led to the alignment of known types across replicates (e.g., cones or IMBs), and used this scheme consistently for all cell group analyses. Here, we also tested the CCA-based (Butler et al., 2018) and MNN-based (Haghverdi et al., 2018) batch correction strategies, but found that the performance of these methods depended on the choice of certain tunable parameters, which were different for different groups. The linear regression model, on the other hand, did not contain additional parameters outside of the choice of covariates, and was computationally much faster than the other two methods. In a few cases, some residual batch effects still remained, which we corrected for in a supervised manner (see below). For BCs and RGCs, two groups extensively validated in this study, we additionally confirmed the absence of any artifacts by re-clustering single batches separately and verifying that all the abundant clusters were preserved. We refer to the corrected expression matrix as \tilde{E} . For each of the HVGs, we compute the Pearson correlation coefficient between the values in \tilde{E} and those in \tilde{E} . Encouragingly, most genes had a similar correlation coefficient ($r > 0.97$), but ~30 genes had a significant reduction in correlation ($r < 0.6$). We flagged these “batch specific genes” in our differential- expression analysis of clusters in 5 (see below).
4. Principal component analysis and 2D visualization: We restricted the corrected expression matrix \tilde{E} to HVGs, and values were centered and scaled before input to Principal Component Analysis (PCA), which was implemented using the R package ‘irlba’. Statistically significant PCs were estimated using the Tracy-Widom distribution (Shekhar et al., 2016), which we found to be an excellent approximation to exact values computed using the more exact, but computationally intensive, bootstrap permutation

test (Shekhar et al., 2016). For visualization purposes only, dimensionality was further reduced to 2D using t-distributed stochastic neighbor embedding (t-SNE). We used a recently developed rapidly accelerated implementation (Linderman et al., 2017) of t-SNE setting the “perplexity” parameter to 30. The resulting t-SNE cell coordinates were used to generate the visualizations in Figures 1D–1I.

5. Graph-based clustering: To partition the data into clusters of transcriptionally related cells, we used unsupervised clustering based on the Louvain algorithm with the Jaccard correction (Shekhar et al., 2016). We first constructed a k -nearest neighbor graph (kNN) using the Euclidean distance metric between pairs of cells in the reduced dimension space of the significant PCs, as implemented in the R package ‘RANN’. $k = 30$ was used for all groups, except peripheral HCs where $k = 15$ was used because of the low cell count.
6. Cluster refinement: As the Louvain method suffers from a well-known “resolution” limit (Fortunato and Barthélemy, 2007), we refined the set of initial clusters as well as corrected for potential overclustering due to residual batch effects in three steps as follows. *Step 1* and *Step 2* were performed for each of the initial clusters in (5) separately. In *Step 1*, each cluster was sub-clustered using the more sensitive Infomap community detection algorithm (Rosvall and Bergstrom, 2008) on a new k -nearest neighbor graph ($k = 5 \cdot \log(\text{cluster size})$). However, this often ended up overclustering the data, as noted before (Shekhar et al., 2016). In *Step 2*, the subclusters within each cluster were assessed for differential expression, and iteratively merged as follows. We built a dendrogram on the subclusters (hierarchical agglomerative clustering with Euclidean distance, complete linkage), and tested the closest three pairs of leaves of the dendrogram for differential gene expression using the MAST test (Finak et al., 2015), corresponding to the three pairs of the most transcriptionally similar subclusters. If any pair did not show sufficient differential expression (minimum 5 DE genes enriched both ways, fold-change > 1.4 , p value < 0.001 , excluding mitochondrial transcripts and ribosomal protein genes, or “batch genes” defined above), the corresponding subclusters were merged. If any of the three pairs of subclusters were merged, we rebuilt the dendrogram using the new subclusters, and repeated this procedure until convergence. *Step 1* and *Step 2* resulted in the splitting of some of the original Louvain clusters. In *Step 3* we built a dendrogram on the new set of clusters and iteratively merged (as in *Step 2*, except with more stringent parameters) until convergence. During this final dendrogram-based merging round, we elected to merge transcriptionally proximal clusters using a more stringent rule. In addition to the criteria described above (≥ 5 DE genes enriched both ways, fold-change > 1.4 , p value < 0.001 , excluding mitochondrial and ribosomal, or batch-specific genes), we required that the two clusters being tested be distinguished by ON/OFF expression (i.e., not just differences in expression levels) of at least one gene in each of them. More specifically, we required that each of the two clusters show at least one gene that was expressed in at least 4-fold higher proportion of cells the other cluster (e.g., $\geq 80\%$ of cells in cluster 1 and $\leq 20\%$ of cells in cluster 2, and $\geq 60\%$ of cells in cluster 2 and $\leq 15\%$ of cells in cluster 2).
7. Removing doublets and contaminants: Although dimensionality reduction and clustering were performed separately for each cell class after an initial segregation, we noted that the data was not entirely free of contaminants. For example, we found small PR-like or AC-like clusters when analyzing BCs in the fovea or periphery. These were likely doublets as these clusters also expressed BC-markers, and comprised < 1 -2% of the cells in the dataset. While such clusters appeared more BC-like in the initial stage, when all cell classes were analyzed together, their contaminant-like features became more apparent when analyzed only with pure BCs. These contaminant-like clusters were removed manually, and we verified that the final clusters for each cell-class expressed at most negligible levels of markers of other cell-classes.

Distinguishing M versus L Cones

The transcript sequences of *OPN1MW* and *OPN1LW* are extremely similar across Old World Monkeys and humans (Nathans et al., 1986; Onishi et al., 2002). In the crab-eating macaque, there are only 26 single nucleotide polymorphism (SNPs) over 1,095 nucleotides of the full-length cDNA sequences of *OPN1MW* and *OPN1LW*, located between exon 2 and exon 5 (Onishi et al., 2002). This made it difficult to unambiguously assign a short read (98bp) aligning to this locus to either *OPN1MW* and *OPN1LW*, as SNPs could not be distinguished from sequencing errors. Thus, cells in the foveal cone cluster that expressed transcripts aligning to this locus, but not the S-opsin *OPN1SW*, were initially tagged as “M/L cones.” However, these SNPs could be robustly distinguished from sequencing errors when aggregating across all the reads aligning to this locus (Figure S2A), and their positions were consistent with previous studies (Onishi et al., 2002). For each M/L cone cell, we then counted the *OPN1MW*- and *OPN1LW*-specific SNPs (Figure S2B). $\sim 70\%$ of M/L cones in the data possessed reads that had at least one transcript covering one of the 26 SNPs. Cells with exclusively *OPN1MW*- and *OPN1LW*-specific SNP counts were labeled M and L-cones, respectively, and the remainder were called “mixed” (Figures S2D and S2E).

Differential Expression Analysis

Differential expression (DE) tests for a gene between a pair of clusters or between a cluster and the remaining clusters were performed using the ‘MAST’ package (Finak et al., 2015), which implements a two-part ‘hurdle’ model to control for both technical quality and cross-animal variation. P values were corrected for multiple hypotheses by controlling the false discovery rate (FDR) using the R function ‘p.adjust’.

We sought markers for a cluster by comparing it to other clusters only within the same class for DE genes, since class specific markers are well-defined (Table S1, Figures 1C). To obtain highly specific markers for histological validation, we ranked significantly enriched (> 1.5 -fold difference, FDR adjusted $p < 10^{-6}$) markers in increasing order of the proportion of background cells where they

were detected (Table S3). Differential gene expression patterns were visualized either in dotplots (e.g., Figure 2B) or box-and-whisker plots (e.g., Figure 4B).

To identify fovea or periphery specific DE genes for cell types that had 1:1 matches across the two regions, we collected the corresponding cells from each region and applied MAST as before, while controlling for library size differences. In choosing candidates to pursue for validation, we also avoided weakly enriched genes that could be explained by large differences in abundance of a different cell class in the same region. For example, *RHO* and *GNAT1* were weakly enriched in peripheral RGCs, BCs and ACs, which likely reflects contamination from the predominant rods in the periphery.

Transcriptional Mapping of Foveal Types to Peripheral Types within a Class

We compared the consistency between foveal and peripheral clusters for HCs, BCs, ACs and RGCs within each cell class using a multi-class classification approach, as described before (Shekhar et al., 2016). For PRs and non-neuronal cells, the fovea versus periphery mapping was straightforward as the types could be easily identified based on known markers, and their transcriptional identities could be matched between the two regions.

For BCs, RGCs, ACs and HCs, we trained a multi-class classifier using Random Forests (RF) as well as the recently published Xgboost algorithm, which uses gradient boosted trees (Chen and Guestrin, 2016) using the R packages ‘randomForest’ and ‘xgboost’, respectively. In each case, we trained the classifier on the peripheral cluster labels, using as features the common HVGs identified in the foveal and peripheral datasets as above (cluster-specific markers were not favored in any way). The training was performed on 50% of the cells in peripheral datasets and its performance was tested on the remaining, “held-out” dataset. (The foveal data was not used for training.) For each of the 4 classes (HCs, BCs, ACs and RGCs), the maximum error rate for types in the held-out data was < 1% (not shown), suggesting that the type-specific transcriptional signatures were robust, learnable, and not prone to overfitting.

The trained classifier was then used to assign each foveal cell a peripheral identity, in a manner that was agnostic to its foveal cluster identity. For each cell, we computed the “margin,” i.e the fraction of decision trees that vote for the “winner” type (majority vote). We excluded cells which fell into the following criterion,

$$\text{margin}(\text{cell}) < \min\left(0.6, \frac{5}{\# \text{ of types}}\right)$$

This was done to buffer against potential misclassifications due to a marginal majority.

For all of the cell classes tested, < 4% of cells were excluded due to this criterion. The congruences of the foveal cluster identities and their peripheral matches were visualized as confusion matrices for each class (Figures 3A–3D). A mapping between a foveal cluster *F* and a peripheral cluster *P* was considered 1:1 if > 85% of cells from *F* mapped to *P*, and if < 5% of cells from every other peripheral cluster mapped to *F*. Results from Xgboost and RF were extremely comparable, although the former was much more computationally efficient. The confusion matrices (Figures 3A–3D) shown are equally-weighted aggregated results of Xgboost and RF.

To quantify the extent of 1:1 mapping between foveal and peripheral clusters we computed the Adjusted Rand Index (ARI), a measure of the similarity between two data clusterings, between the foveal cluster labels and the Xgboost/RF–assigned cluster labels using the R package “mclust.” The significance of the computed ARI values were evaluated by comparing against “null” ARI values computed by randomly permuting the cluster labels across cells. Mean and standard-deviation of null ARI values were computed by averaging 10,000 trials.

Comparison of Cluster Frequencies between the Fovea and the Periphery

We computed frequency differences of 1:1 matched types between the fovea and the periphery, and assessed their significance using a two sample t test. For foveal types that were a mixture of multiple peripheral types (e.g., fRGC11), we assigned each foveal cell its matched peripheral cluster identity prior to comparison (see previous section). A fovea versus periphery frequency comparison could not be meaningfully done for ACs, as CD90 non-uniformly labels ACs. We believe that this phenomenon likely explains the poor correlation between foveal and peripheral AC type frequencies ($r = 0.12$); we could therefore not make a fair comparison of AC types between the regions. Interestingly, although the CD73- samples contributed only ~10% of all peripheral ACs, the AC type frequencies in these samples correlated much better with the foveal population ($r = 0.79$; Figure S4B), and suggested that a few peripheral types (e.g., VIP+ ACs) were conspicuously underrepresented in the foveal samples (Figure 3C).

We quantified the compositional similarity between foveal and peripheral clusterings for each cell class by computing the Jensen-Shannon divergence (JSD), as follows. Let $P = \{p_1, p_2, \dots, p_n\}$ and $Q = \{q_1, q_2, \dots, q_n\}$ be the peripheral and foveal frequencies of n clusters of a cell class *C*, each of which sums to unity. Here n is the total number of clusters which is a sum of the 1:1 matched clusters, and the clusters unique to either region. Then the JSD between the fovea and the periphery for cell class *C* is computed as,

$$\text{JSD}(P|Q) = \frac{1}{2} D(P|M) + \frac{1}{2} D(Q|M)$$

Here, $M = \frac{1}{2}(P + Q)$, and $D(P|Q)$ is the Kullback-Leibler (KL) divergence between probability distributions *P* and *Q* as follows,

$$D(P|Q) = \sum_{i=1}^n p_i \log \frac{p_i}{q_i}$$

Note that while $D(P | Q) \neq D(Q | P)$, $JSD(P | Q)$ is symmetric in that $JSD(P | Q) = JSD(Q | P)$.

In comparing compositional similarity between foveal and peripheral cell classes, we used average frequencies across all replicates for foveal PRs, BCs, HCs, ACs and RGCs, but only BCs, HCs, ACs and RGCs in the periphery. This is because our collection procedure (Figure S1B) methods selected against peripheral PRs. We therefore used a published estimate of peripheral PR composition: 95% rods, 4.75% M/L cones and 0.25% S cones (Roorda et al., 2001; Wikler et al., 1990). In addition, we considered only the CD73- samples for ACs because the CD90+ samples differentially excluded some AC types. For peripheral ACs, we only used their composition in the CD73- samples (see above). We expect rods to be underrepresented in our peripheral samples due to our collection procedure.

Transcriptional Mapping of Cell Types across Species

Mouse. We transcriptionally related macaque cones, HC, BC, AC, and RGC types to their mouse counterparts from earlier publications using the same approach that related foveal to peripheral types, with minor modifications (see below). We conducted the following mappings:

1. 11 macaque peripheral BC clusters (BB/GB* was considered as a single cluster and then subclustered post hoc) were mapped to 15 mouse BC types published in our earlier study that identified these types molecularly and matched them to morphology (Shekhar et al., 2016)
2. 34 Macaque peripheral AC clusters were mapped to 21 mouse AC clusters identified in our earlier whole-retina study (Macosko et al., 2015), which represents an incomplete molecular characterization of ACs, estimated to contain > 60 types based on morphological (Diamond, 2017) and molecular (data not shown) diversity.
3. Macaque peripheral H1 and H2 were transcriptionally related to cells from a single cluster of mouse HCs in our previous whole retina study (Macosko et al., 2015) to determine whether mouse HCs were molecularly more similar to one of the two types.
4. 18 macaque peripheral RGC clusters were mapped to a recently published dataset of postnatal day 5 (P5) mouse RGCs, which identified 40 distinct molecular clusters (Rheaume et al., 2018).
5. As we did not sample peripheral S-cones, we related macaque foveal M/L and S cones to mouse cones from our previous full retina data (Macosko et al., 2015). We computationally separated mouse cones into those that exclusively expressed *Opn1sw* or *Opn1mw* to enrich for pure S-cones and pure M-cones, respectively.

In all cases except HCs, we trained the classifier on the mouse types, and mapped the macaque cells on to these types. In each case 1-5 above, we trained RF/Xgboost classifiers on the training data (mouse for Cones, BCs, ACs, RGCs and macaque for HC) using as features the common set of highly variable genes (HVGs) in both the mouse and the macaque datasets among the 1:1 orthologs. Restricting the features to the common set of HVG orthologs was important for the performance of the classifier; in our initial tests, we found that using HVG orthologs identified only in one of the two species resulted in mappings that were more diffuse, likely due to genes that expressed in one species at high levels but, were completely absent in the other due to transcriptional rewiring. However, we emphasize that no manual curation of the feature set was performed to favor or disfavor certain genes based on their cell-type specific expression. We trained the classifier using 80% of the cells in the relevant mouse dataset, and validated it on the “held-out” cells in that dataset, which showed excellent performance in all cases. Next, we applied the classifier on the test set (macaque Cones, BCs, ACs, RGCs and mouse HCs), and assigned each cell an identity in the other species. As above, cluster identity of cells in the test set was not used to guide the mapping. Mappings of cell types across species were visualized using confusion matrices for Cones, ACs, BCs and RGCs (Figures 5A and 5C–5E) and a barplot for HCs (Figure 5B). A mapping between a test cluster C_{test} and a training cluster C_{train} was considered 1:1 if > 75% of cells from C_{test} mapped to C_{train} , and if < 5% of cells from every other test cluster mapped to C_{train} .

Human. We related human BC clusters to macaque BC clusters using an approach similar to that described above for mouse. We trained a classifier on the macaque peripheral BC data, applied it to map each human BC to a macaque cluster, and visualized the resulting confusion matrix to identify 1:1 matches (Figure 6F).

Conservation of Transcription Factor (TF) Codes in Mouse and Macaque RGCs

We downloaded a curated list of human and mouse transcription factors from the TFCat Transcription Factor Database Catalog (www.tfcats.org), CIS-BP database (<http://cisbp.cccb.utoronto.ca/>) and (Lambert et al., 2018), and used these to assemble a corresponding list of macaque transcription factors from our orthology list (see above). We filtered this list in two ways. **First**, we restricted the list to those TFs that had 1:1 mouse and macaque orthologs. **Second**, we only considered those TFs that were expressed (> 0 UMIs/cell) in > 25% of the cells in at least one cluster.

For each of the remaining TFs, we computed a TF specificity score (TFSS) among RGC types separately in macaque and mice. Specifically, for a transcription factor t , we computed a vector \vec{p}_t of the fraction of cells expressing t in every RGC cluster. For example, $\vec{p}_t = (0.7, 0, 0.3, \dots)$ implies that 70% of cells in RGC cluster 1, 0% of cells in cluster 2, 30% in cluster 3, and so on, express transcription factor t . We normalized \vec{p}_t to sum to 1, and computed the Shannon index of the resulting normalized vector,

$$TFSS(t) = - \sum_{clust} p_{t,j} \ln p_{t,j}$$

We computed $TFSS_{mouse}(t)$ and $TFSS_{macaque}(t)$ for all filtered transcription factors t ($n = 93$), and compared them (Figure 5G). Note that constitutively expressed TFs like *Pou4f1/Six6* have low values on the plot, because of their low Shannon indices.

To compute TF co-expression scores and compare across species, we further filtered the 1:1 orthologous TF pairs to only include those TFs that are expressed in at least 15% of clusters in both species. For each species, we then binarized TF expression in RGC clusters as a matrix \mathbf{T} , where $\mathbf{T}(t, c) = 1$ if TF t is expressed in at least 25% of cells in cluster c . For every TF t , we computed its probability of expression in types, $\alpha(t)$, in any species as,

$$\alpha(t) = \frac{1}{N_c} \sum_c \mathbf{T}(t, c)$$

Here, N_c refers to the number of clusters. The observed probability of co-expression of TFs t_1 and t_2 , $A(t_1, t_2)$, can be computed as,

$$A(t_1, t_2) = \frac{1}{N_c} \sum_c \mathbf{T}(t_1, c) * \mathbf{T}(t_2, c)$$

We compute the TF co-expression scores (TFCS) in either species as,

$$TFCS(t_1, t_2) = \log\left(\frac{\text{Observed co-expression probability}}{\text{Expected co-expression probability}}\right) = \log\left(\frac{A(t_1, t_2)}{\alpha(t_1)\alpha(t_2)}\right)$$

Figure S5G compares $TFCS_{mouse}(t_1, t_2)$ and $TFCS_{macaque}(t_1, t_2)$.

Evaluating Cell-Type Specific Expression of Disease Associated Loci

We assembled lists of genome-wide associated study genes for prominent retinal diseases that cause blindness – primary open angle glaucoma (POAG), primary angle closure glaucoma (PACG), cone-rod dystrophy (CRD), retinitis pigmentosa (RP), congenital stationary night blindness (CSNB), age-related macular degeneration (AMD) and diabetic macular edema (DME) from (Farrar et al., 2017; Fritsche et al., 2016; Graham et al., 2018; Wiggs and Pasquale, 2017; Zeitz et al., 2015).

To facilitate 1:1 comparison between the fovea and the periphery, we assigned each foveal cell a peripheral type identity using the RF/Xgboost mapping described above. We only considered those genes that were expressed in at least 20% of cells of at least one cell type, which resulted in 196 genes for further consideration. We note that this does not necessarily rule out the possibility of these genes turning on in specific cell types during disease pathogenesis, just as the expression of a gene in a cell type at baseline does not necessarily imply vulnerability in disease.

We visualized expression patterns of the disease genes in our atlas using one of three visualizations:

1. Dotplots, which show cell-type resolution (as in Figure 7D),
2. Heatmaps showing relative expression of individual genes in major cell classes (Figures 7B and S7), and
3. Heatmap showing aggregated expression of groups of genes corresponding to individual diseases in major cell classes (Figure 7A). While 2. and 3. lack cell-type resolution, they do show statistical enrichment of genes and gene groups in specific cell classes, and differences in fovea and periphery, if they exist.

For each of the fovea and periphery, we summarized the expression of all disease genes (see above) across all cell types, by computing two matrices \mathbf{E} and \mathbf{F} . For gene g and cell-type c , $\mathbf{F}(g, c)$ was the fraction of cells in cell type c that had non-zero expression (> 0 UMIs) of g , and $\mathbf{E}(g, c)$ was the average number of transcripts in cells with non-zero expression. To remove outliers, we capped values in \mathbf{E} to the 99.5th quantile of expression. We then computed a matrix of gene expression scores for all genes across all cell types as the product of these two matrices,

$$\tilde{\mathbf{E}} = \mathbf{E}\mathbf{F}$$

For (2), we computed the expression strength $\mathbf{S}(g, C)$ of each gene g in a cell class C (one of Rods, Cones, HCs, BCs, ACs, RGCs, and non-neuronal cells)

$$\mathbf{S}(g, C) = \max_{c \in C} \tilde{\mathbf{E}}(g, c)$$

Here, we computed the maximum, rather than the average or aggregate, to highlight specific instances of highly cell-type specific expression of genes in a class. We combined $\mathbf{S}_{fovea}(g, C)$ and $\mathbf{S}_{periphery}(g, C)$, and visualized them as a single heatmap, where we z-scored the rows (corresponding to genes) to highlight relative expression strengths $\tilde{\mathbf{S}}_{comb}(g, C)$.

Specific rows are highlighted in Figure 6I. Here, the initial filtering of removing genes that were not expressed in at least 20% of cells in any cell type avoided spurious patterns.

For (3), we subset $\tilde{\mathbf{S}}_{comb}(g, C)$ by disease and computed the mean relative expression strengths within each cell type group for each disease. This aggregate expression strength was visualized in Figure 7A.

In Figure S7, we visualized the relative expression strengths of genes grouped by disease as in Figure 7B, but after subdividing non-neuronal cells into MGs, Pericytes, Endothelial cells and Microglia to highlight differential expression among these types.

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical methods and packages used for performing various types analysis in this work are cited where appropriate in the [STAR Methods](#) text. All the analysis was performing using the R statistical language in the RStudio IDE. Expression patterns of genes across cell clusters are shown in dotplots (e.g., [Figure 2B](#)), which simultaneously depict the fraction of cells in a cell cluster/type (column) that express a particular (row) based on the size of the dot, and the average number of transcripts in cluster based on the color scale, as indicated in the legends. In box and whisker plots depicting gene expression differences foveal and peripheral types, black horizontal line represents the median expression value, colored bars represent the interquartile range, and vertical lines span the minimum and maximum values. Stars indicate > 2-fold changes based on mean expression values.

The concordance between two sets of clusters as assessed by the multi-class classification approach (see METHOD DETAILS) are visualized using a “Confusion Matrix” (e.g., foveal BCs versus peripheral BCs in [Figure 3A](#), or macaque BCs versus mouse BCs in [Figure 5E](#)). The concordance between the two data clusterings is quantified using the Adjusted Rand Index (ARI).

Differential expression of genes between clusters were evaluated using the MAST test ([Finak et al., 2015](#)). We employed multiple hypothesis correction wherever significance was evaluated across multiple statistical tests, only considering genes more significant than False-discovery rate (FDR) threshold of 0.001. We provide FDR corrected p values for specific instances of fovea versus periphery differential expression reported in [Figure 4](#).

DATA AND SOFTWARE AVAILABILITY

The accession number for the following data files reported in this paper is GEO: GSE118480 – Raw and processed data files for the macaque single-cell RNA-seq experiments. The accession number for the following data files reported in this paper is SRA: PRJNA514024 – (1) Raw data files for the long-read bulk RNA-seq data from Tru-Seq libraries, and (2) gene transfer format (gtf) file summarizing the transcriptome for *M. fascicularis* used in this study ([Methods S1](#)). The accession number for the following data files reported in this paper is DUOS-000105 – Raw and processed data files for the human single-cell RNA-seq experiments. The macaque single cell data can be visualized in the Broad Institute’s Single Cell Portal at https://portals.broadinstitute.org/single_cell. Instructions are provided in supplementary materials.

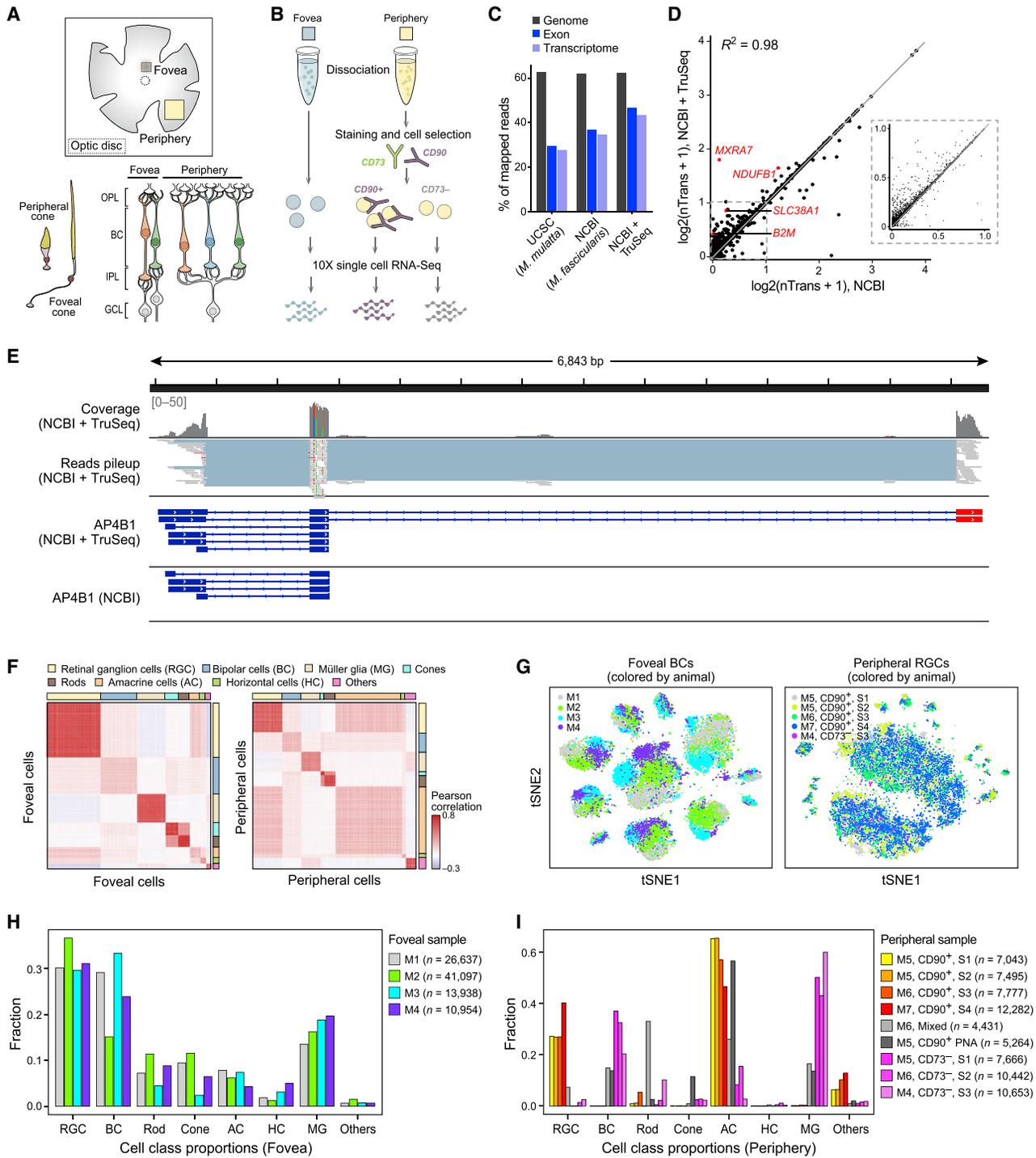


Figure S1. Foveal Specializations, Experimental Design and Data Quality, Related to Figure 1

A. (Top) The position of fovea and optic disc in flat mounted retina. (Lower left) Sketch of foveal and peripheral cones: foveal cones are longer and more slender than peripheral cones, and bear longer axons. Adapted from (Greeff, 1900). (Lower right) Foveal and peripheral retinal circuits: In the fovea, one cone provides input to two midget BCs, each of which innervates one midget RGC for a cone:midget RGC ratio of 1:2. In the periphery, multiple cones provide input to each midget BC, and multiple midget BCs innervate one midget RGC, for a cone:midget RGC ratio of $\geq 10:1$. Thus the degree of convergence of cones onto interneurons and RGCs differs by an order of magnitude between the fovea and periphery. In addition, peripheral MGCs receive input indirectly (via synapses from rod BCs to All ACs to cone BCs) from hundreds of rods; whereas foveal MGCs receive little if any rod input. Thus, the total difference in PR:MGC convergence is even greater between regions. Adapted from (Kolb and Marshak, 2003).

(legend continued on next page)

B. scRNA-seq workflow. Cells were dissociated from < 1.5 mm-diameter foveal samples and collected without further processing. Because the peripheral retina is dominated by rod PRs (~80% of cells), we used magnetic columns to deplete rods (CD73+) or enrich RGCs (CD90+).

C. Mapping rates of scRNA-seq reads to the genome, exonic and transcriptomic (exonic with splice-junction constraints) regions using three different transcriptomic references – UCSC (University of California Santa Cruz Genome Browser reference for *Macaca mulatta*, a related but different species), NCBI (the publicly available reference for *Macaca fascicularis*) and NCBI + TruSeq (the retina specific transcriptome assembled in this study using the NCBI reference as a starting point). Note that the new transcriptome we used increased the percentage of exonic reads that could be assigned to a gene by ~25% compared to the NCBI reference for *Macaca fascicularis* (from 37% to 47%) and by ~68% compared to the NCBI reference for *Macaca mulatta* (from 28% to 47%).

D. Comparison of expression levels ($\log_2(\text{number of transcripts} + 1)$) of common genes between the NCBI and NCBI + TruSeq references, exhibits a high degree of correlation. A greater number of gene loci exhibit increased expression levels due to higher mapping rate (panel C) as evident by the fact that the majority of points lie above the red line ($x = y$). This is particularly true for genes expressed at low levels, as shown in the inset. A small number of loci were mapped less well in the improved transcriptome because new transcripts could not be assigned unambiguously to a single gene.

E. Example of improved gene-body definition in the assembled transcriptome of *AP4B1* as visualized using the Integrated Genomics Viewer (IGV). The lower panel shows the gene-body definitions for the NCBI and the NCBI + TruSeq references. In this example, the NCBI+TruSeq reference includes a distal 3' exon that is absent from the NCBI reference. The middle panel shows the pile-up of individual reads from a sample 10X run mapped to the expanded locus and the upper panels show the read coverage. Blue shading connects portions of a read that spans a splice junction. The coverage plot shows that a large proportion of reads mapped to 3' exons present in the NCBI + TruSeq reference (highlighted in red) but absent from the NCBI reference.

F. Heatmap of Pearson correlation coefficients between each pair of 92,628 foveal cells (left) or 73,053 peripheral cells (right) (rows and columns) ordered by cell class (annotated as color bars along row and column). "Other" cells are comprised of pericytes, vascular endothelial cells and microglia.

G. Examples illustrating the lack of strong batch effects. tSNE visualization of foveal BCs (left, as in [Figure 1F](#)) and peripheral RGCs (right, as in [Figure 1H](#)), which are now colored by their animal of origin (M1-7), shows the representation of all animals across clusters.

H., I. Proportions of major cell classes in foveal (H) and peripheral (I) samples colored by experimental batches. Peripheral samples (I) colored using palettes corresponding to one of four processing methods prior to collection: (1) CD90+: RGCs were enriched on a magnetic column using beads conjugated with antibodies to the RGC class marker, Thy1 (CD90). Cells that did not bind were discarded, and bound cells were eluted and used. (2) CD73-: Rod photoreceptors were depleted by passage over a magnetic column containing beads conjugated with antibodies to the rod-specific marker CD73. Unbound cells were used. (3) Mixed: In this experiment CD90+ cells (magnetic column selection) and non-enriched cells were mixed 1:1 prior to use. (4) CD90+PNA: In this experiment, PNA ([Blanks and Johnson, 1984](#)) was added to enrich for cones. M4 here corresponds to the same animal. Each experimental batch corresponds to an independent animal, denoted M1-M7. The entire fovea was dissociated, and cells were collected for unbiased sampling of all types. Sequencing batches, typically containing 3000-5000 cells each, are aggregated within experiments as they did not exhibit any appreciable batch effects. Cell numbers in each experimental batch are indicated in parentheses.

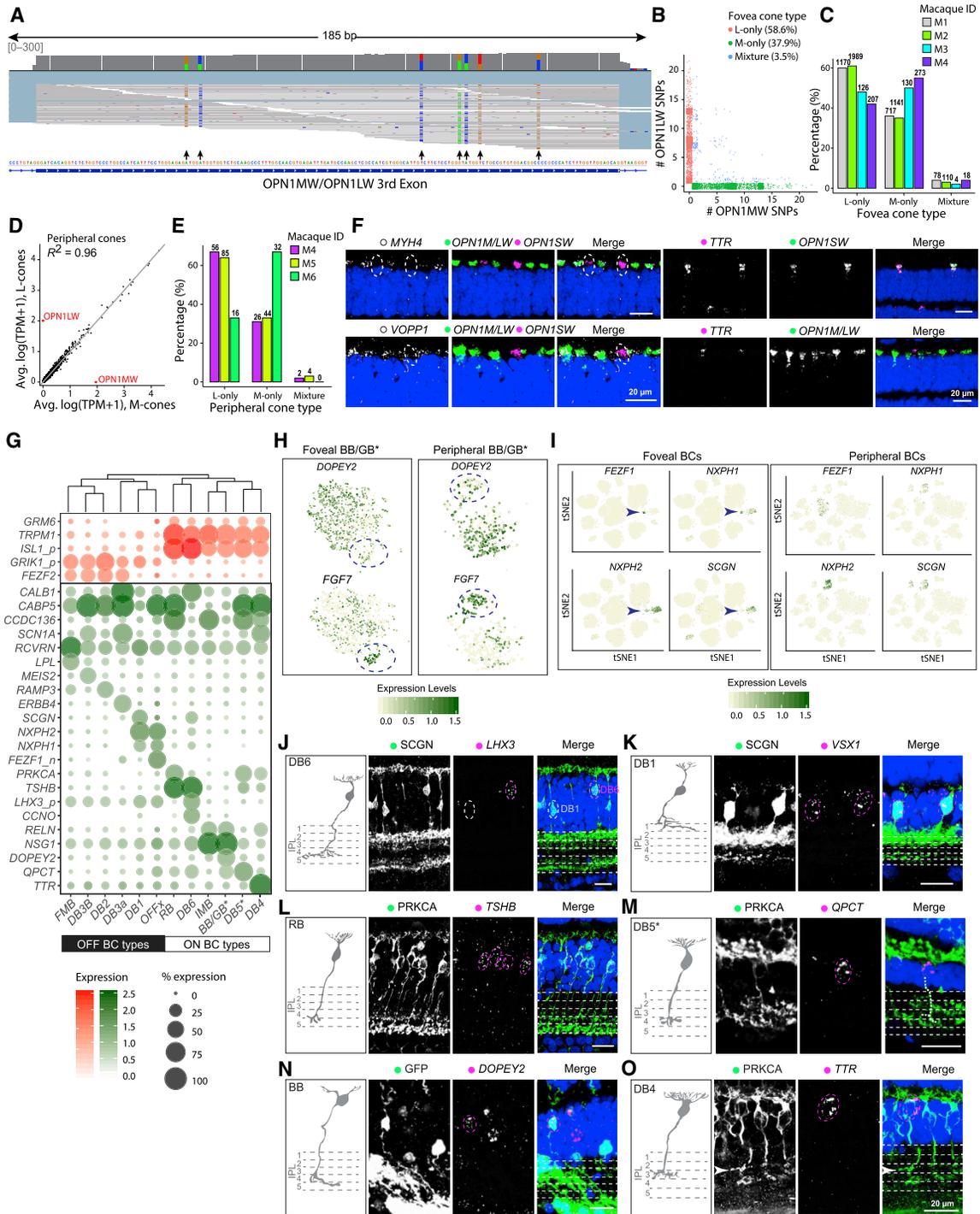


Figure S2. Characterization of Cone and BC Types, Related to Figure 2

A. The reference macaque genome contains a single gene locus named “*OPN1MW/OPN1LW*” and, because of their high sequence similarity revealed from cloned cDNAs (98%; (Onishi et al., 2002)), our NCBI+TruSeq reference did not distinguish *OPN1MW* from *OPN1LW*. We therefore analyzed single-nucleotide polymorphisms (SNPs) to distinguish *OPN1MW*- and *OPN1LW*-specific expression in individual cones. An Integrated Genomics Viewer (IGV) screenshot of reads from a representative sample aligning to exon 3 of the loci shows that SNPs (arrows) are distinguishable from random sequencing errors at known locations (Onishi et al., 2002) based on their higher biallelic frequency.

B. Counts of *OPN1LW*-specific SNPs versus *OPN1MW*-specific SNPs for every M, L cone (dots) show that most M/L cones in our data exclusively express either *OPN1MW* or *OPN1LW*. The remaining cells exhibit a mixed identity but only 0.8% show appreciable expression of both genes. It is generally believed that all M/L

(legend continued on next page)

cones express either M or L opsin but not both. However, experimental evidence for this conclusion is lacking; the most authoritative studies, based on spectrophotometry of individual human cones in live retinae, cannot exclude the possibility that up to 5% of M/L express both opsins (Hofer et al., 2005).

C. Proportions of M, L and mixed M/L cones in the fovea.

D. Comparison of average transcriptional profiles of peripheral M-cones and L-cones identified through analysis of SNPs, as in Figure 2A. As in the fovea, no genes other than *OPN1MW* and *OPN1LW* differ significantly in expression levels (> 1.2 fold at $p < 0.01$, MAST test) between the two cone types. Dashed red line, $y = x$. As noted in the main text, the question of specificity in connections of M and L with synaptic partners has been controversial. If connectivity is selective, it is likely that M and L cones express different recognition molecules, implying transcriptional differences between the two that extend beyond the opsin genes. In addition to some physiological evidence for specificity, there is a report of possible structural differences between the terminals of M and L cones (Lee et al., 2012). If wiring is nonselective no additional expression differences between the two cone types would be needed. Consistent with the idea, several physiological studies detect little or no selectivity (Wool et al., 2018) in chromatic inputs to MGC surrounds. A prevalent model is that the choice between expression of M and L opsin is stochastic, mediated by regulatory elements at the M/L locus (Wang et al., 1999). In this model no additional transcriptional differences would be predicted. Our results are consistent with this view. However, we cannot draw firm conclusions because it is possible that transcriptional differences are present during development but not maintained into adulthood.

E. Proportions of M, L and mixed M/L cones in the peripheral retina.

F. Additional FISH validations of markers that distinguish M/L-cones (*MYH4* and *VOPP1*) from S-cones (*TTR*) (see Figure 2B). Circles denote location of S cones. Scale bar is 20 μm . DAPI staining is in blue.

G. Gene expression patterns (rows) of broad markers (top panel) and type-enriched markers (bottom panel) for foveal/peripheral BC types (columns) ordered using hierarchical clustering (see dendrogram on top). Representation as in Figure 2D.

H. *FGF7* and *DOPEY2* are enriched in distinct subpopulations within a single cluster that appears to contain both Blue Bipolar and Giant Bipolar types (BB/GB* cluster). Data were extracted from the full foveal and peripheral BC datasets and visualized by tSNE, as in Figure 1F. As *DOPEY2* (hi) cells correspond to the BB type (Figure S2N), we speculate that the *DOPEY2* (low), *FGF7*(hi) (dashed circles) subpopulation corresponds to the GB type. These assignments are consistent with their relative frequencies (Table S2). Colors represent expression levels. Table S2 also includes references to prior studies that document type-specific expression of BC markers in primates.

I. Expression of *FEZF1*, *NXPH1*, *NXPH2*, and *SCGN* in foveal (left) and peripheral BCs (right) visualized by tSNE as in Figs. 1F. Coexpression of *FEZF1*, *NXPH1* and *NXPH2* mark OFFx BCs in fovea (arrows), but they are not coexpressed in periphery, supporting the claim that OFFx may be a fovea-specific type. *SCGN* is expressed by DB1 and DB6 (see panels G, J, K) but not by OFFx.

J. Secretogogin (*SCGN*) labels two BC types in the fovea, DB6 (magenta circle), which is also *LHX3+* (see Figs. 2D,E) and DB1 (white circle), which is *LHX3-*. K. DB1 (*SCGN+*) is also *VSX1+*.

L., *TSHB* (L) and *QPCT* (M) label two distinct types of PRKCA+ BCs with distinct morphologies. Based on their abundance, *TSHB+* cells are rod bipolars (RB) and *QPCT+* cells could be DB5* (Table S2). The axon of DB5* was traced and is represented by a white dashed line. In contrast, PRKCA is a selective rod bipolar marker in mice (Shekhar et al., 2016).

N. A viral labeled BB cell is *DOPEY2* positive. BB type has a distinguishable bifurcated dendritic morphology (Mariani, 1984).

O. *TTR* is expressed by PRKCA+ DB4. Arrowheads indicate the DB4 axon terminal.

In J-O, sketches were redrawn from (Tsukamoto and Omi, 2016). Scale bar, 20 μm . DAPI staining is in blue.

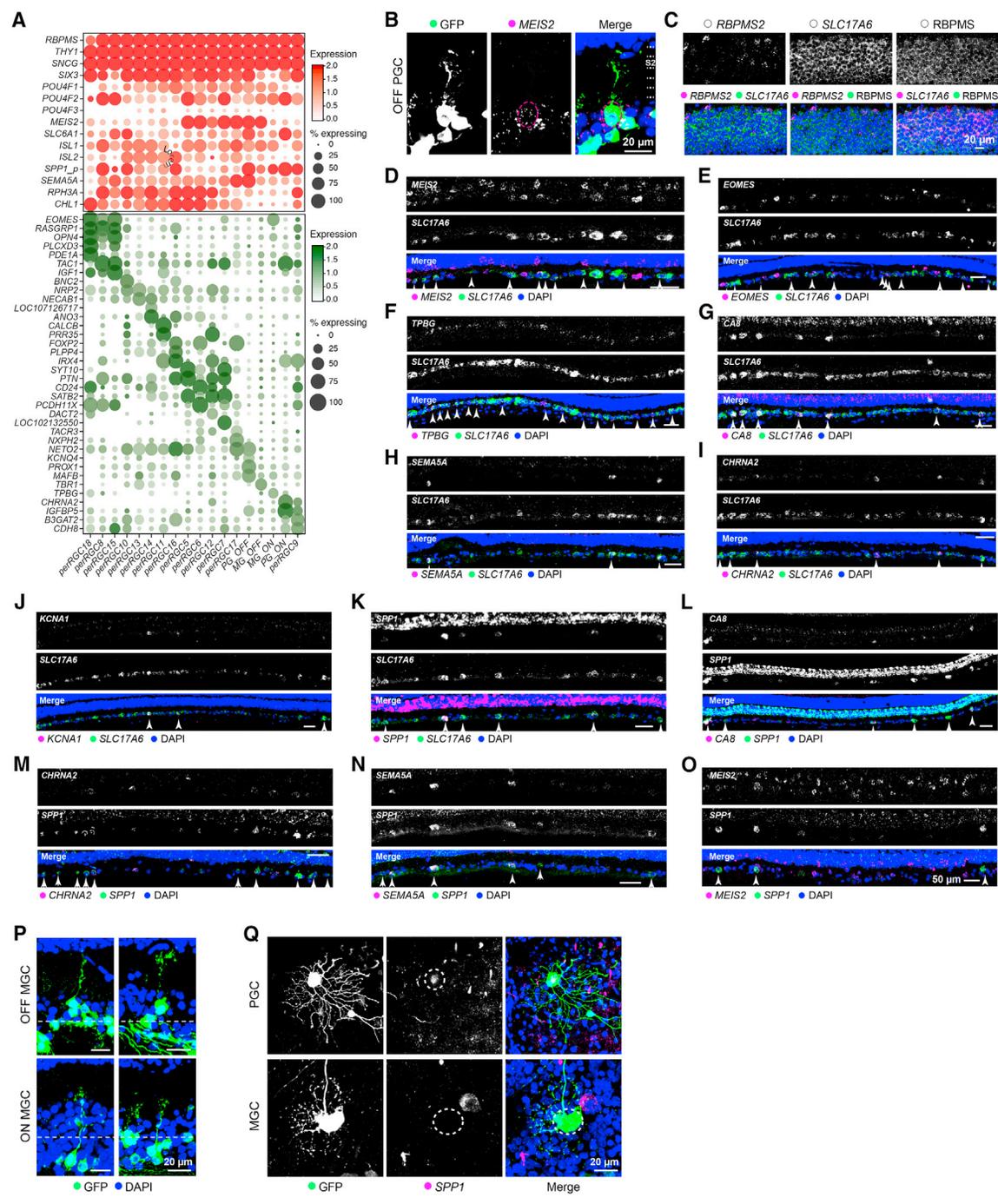


Figure S3. Histological Validation of Markers for RGCs, Related to Figure 2

A. Gene expression patterns (rows) of broad (top panel) and type-enriched markers (bottom panel) for peripheral RGC clusters (columns) ordered by transcriptional similarity. Representation as in Figure 2F. See Figure 3B for transcriptional mapping of foveal and peripheral RGC clusters.

B. FISH combined with viral labeling shows *MEIS2* expression by an OFF PGC with dendrites that arborize in the S1/S2 sublaminae of the IPL.

C. Triple-labeling of a foveal section shows that all foveal RGCs are *SLC17A6*⁺ and *RBPMS*⁺ (immunohistochemistry) but only a minority are *RBPMS2*⁺, consistent with its expression by PGCs but not MGCs (panel A). *SLC17A6* encodes VGlut2, which is a pan-RGC marker in rodents, but is not expressed by other cells in the ganglion cell layer (Stella et al., 2008). *RBPMS* is a selective marker for RGCs in both rodents and macaques (Rodriguez et al., 2014) but to our knowledge expression of its homolog, *Rbpms2*, has not been reported in mouse retina.

D.-K., Double FISH shows that *MEIS2* (D), *EOMES* (E), *TPBG* (F), *CA8* (G), *SEMA5A* (H), *CHRNA2* (I), *KCNA1* (J), and *SPP1* (K) colocalize with RGCs (*SLC17A6*⁺) in peripheral retina. Arrowheads indicate double positive cells.

(legend continued on next page)

L.-O. Double FISH shows that *CA8* (L), *CHRNA2* (M), and *SEMA5A* (N) label subsets of *SPP1*-positive RGCs. In contrast, only a small fraction of *MEIS2*+ RGCs are *SPP1*+ (O); these are OFF PGCs (see panel B). Arrowheads indicate double positive cells; arrows indicate *SPP1* single positive cells.

P. Viral labeling of foveal RGCs shows that the somata of OFF-MGCs (top) are localized at the outer half of the GCL while those of ON-MGCs are located at the inner half (bottom).

Q. Viral labeling combined with immunostaining shows that a PGC expresses *SPP1* (upper) but not a MGC (lower). Somata positions of labeled RGCs are circled in the middle panel. Cells are from the temporal peripheral retina.

Scale bars is 20 μm (B, C, P, Q) and 50 μm (D-O).

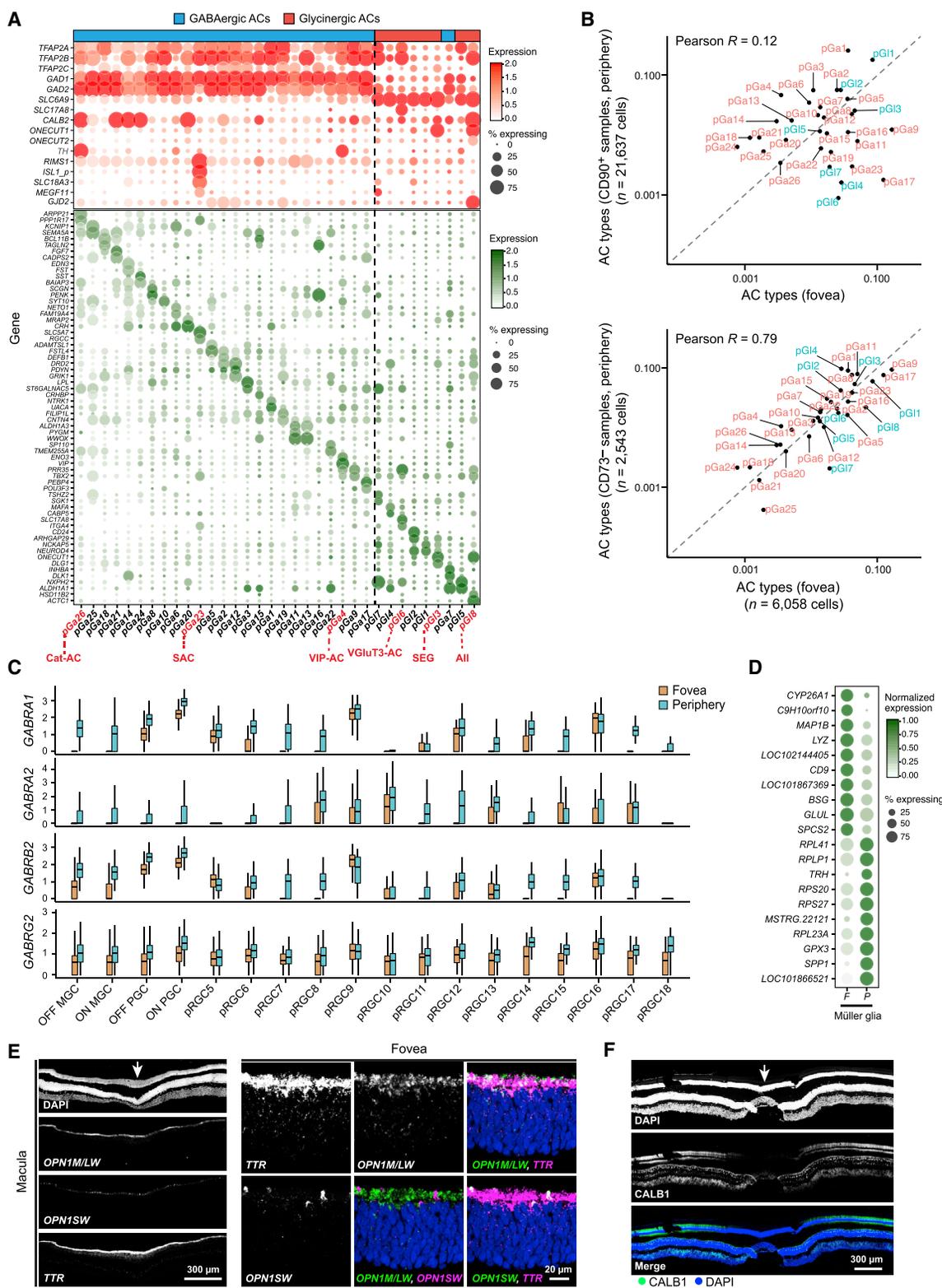


Figure S4. Differences in Proportions of Cell Types and Gene Expression between the Fovea and Peripheral Retina, Related to Figures 3, 4
 A. Gene expression patterns (rows) of broad markers (top panel) and type-enriched markers (bottom panel) for 34 peripheral AC cluster (columns) ordered by transcriptional similarity, and their approximate segregation is highlighted by a black dashed line. Representation as in Figure S2G. Annotation bar on the top highlights GABAergic (n = 26) and Glycinergic (n = 8) types. Key known types of ACs including Starburst AC (SAC), All ACs, Excitatory (VGlut3+) AC and

(legend continued on next page)

Vasoactive Intestinal Peptide expressing (VIP+) AC, catecholaminergic (Cat+) and SEG AC are highlighted below. Rodent AC types have been described in detail (Diamond, 2017; Haverkamp and Wässle, 2004; Kay et al., 2011; Krishnaswamy et al., 2015; Park et al., 2015; Zhang et al., 2007) as have several primate equivalents (Klump et al., 2009; Lammerding-Köppel et al., 1991; Majumdar et al., 2008; Mariani and Hokoc, 1988; Mills and Massey, 1999; Yamada et al., 2003).

B. Comparison of frequencies of matched AC types between the foveal (x axis) and peripheral datasets (y axis). To facilitate comparison, we assigned each foveal AC a peripheral identity based on Figure 3C and then compare frequencies of 1:1 matched types. The peripheral ACs are subdivided into those originating from CD90+ samples (upper) and CD73- samples (lower). Values are plotted on a ln-ln scale to highlight low frequency types. The poor correlation of the AC types in the CD90+ sample with the fovea suggests a biased labeling of AC types by this marker. Hence, we only used AC cluster frequencies in the CD73- samples, which show a much higher correlation to the foveal frequencies, to compute compositional similarities between the fovea and the periphery.

C. Expanded version of Figure 4F, showing higher expression levels of GABRA receptor subunits *GABRA1/2*, *GABRB2* and *GABRG2* in the majority peripheral RGC compared to foveal RGC types. For 1:1 comparison each foveal RGC is assigned a peripheral identity based on Figure 3B.

D. Gene expression differences between foveal and peripheral MGs. The average number of transcripts per expressing cells (colors) are normalized so as to sum up to 1 along each row to highlight relative differences.

E. Co-labeling of *TTR* with both *OPN1LW* and *OPN1SW* in the fovea suggests that it is expressed by both foveal M/L and S cones.

F. Foveal cones were CALB1-negative, as documented in (Hendrickson et al., 2007).

Scale bars, 20 μm in E, and 300 μm in E,F). Arrows indicate the center of fovea

In the bar and whisker plots in C, black horizontal line, median; bars, interquartile range; vertical lines, minimum and maximum.

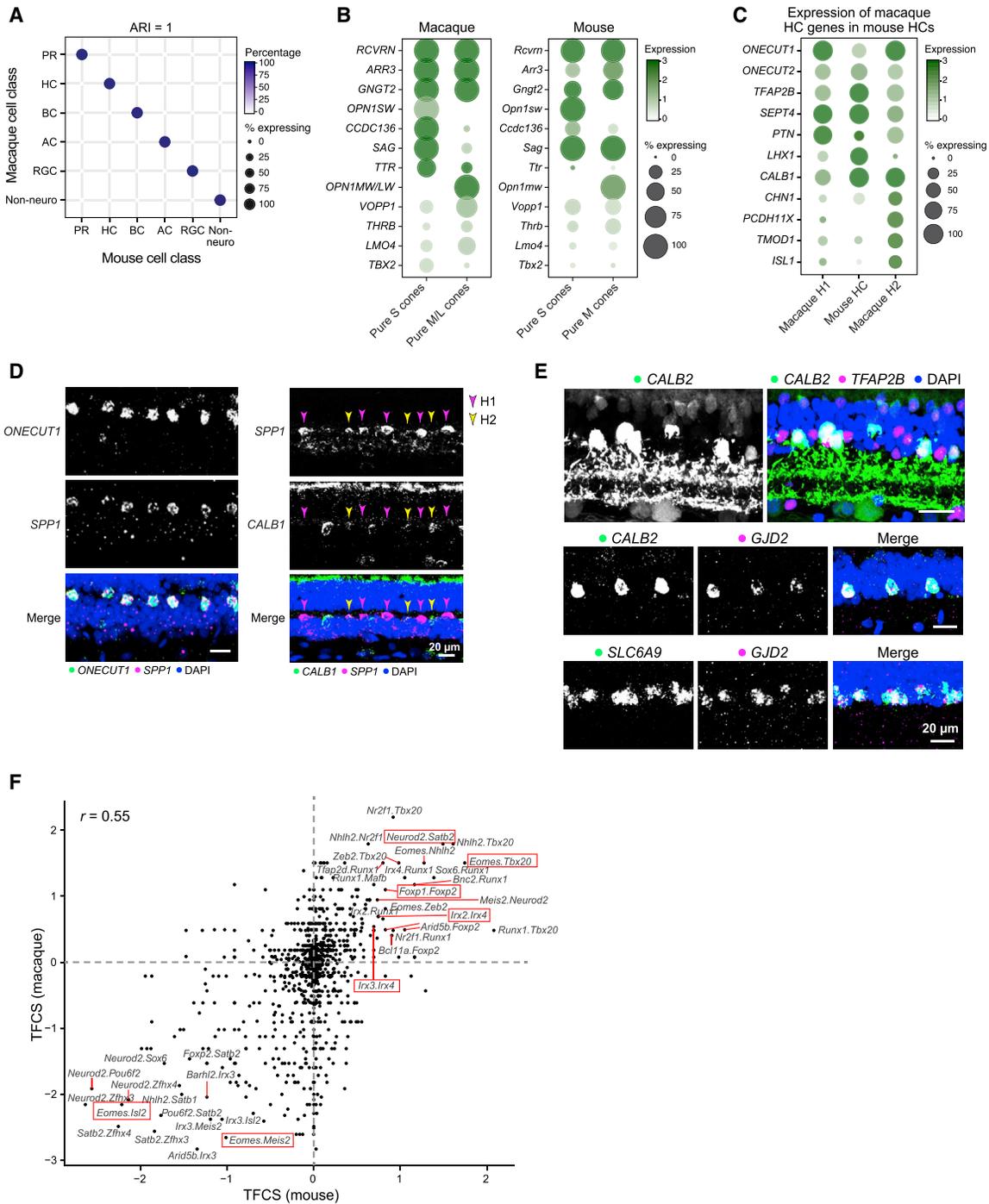


Figure S5. Comparison of Mouse and Macaque Cell Types, Related to Figure 5

A. Transcriptional correspondence between macaque (rows) and mouse (columns) cell classes. All types are included within each class. Representation as in Figure 5A

B. Mouse S cones express higher levels of *Ccdc136* compared to M cones, the mouse ortholog of macaque S cone marker *CCDC136* (Figure 2B). Other S cone or M/L cone markers in macaque, however, are not shared in mice. *CCDC136* locus is located near the *OPN1SW* locus in both mouse and human (~16 kb separation on human chromosome 7 and mouse chromosome 6) suggesting the possibility of transcriptional co-regulation, possibly by *NRL* (Brooks et al., 2011). *THRB*, selectively expressed by *OPN1MW*-expressing mouse cones and M/L macaque cones, has been implicated in differentiation of this cone type (Ng et al., 2001).

C. Mouse HCs are transcriptionally more similar to macaque H1 (Figure 5B), and do not express markers that are specific to macaque H2.

D. *SPP1* is a macaque HC marker gene and H1 and H2 are distinguished by the expression of *CALB1*. Double FISH of *SPP1* and *CALB1* in a peripheral retina section labels both H1 (*SPP1*+*CALB1*-, magenta arrows) and H2 (*SPP1*+*CALB1*+, orange arrows).

(legend continued on next page)

E. Validations of conserved All markers: (Top) All ACs, identifiable by their unique morphology, are *CALB2+TFAP2B+*. Double FISH shows that coexpression of All marker *GJD2* (Macosko et al., 2015) with *CALB2* (middle) and *SLC6A9* (bottom).

F. Comparison of transcription factor co-expression scores (TFCS, see STAR Methods) of pairs of transcription factors for macaque (y axis) and mouse (x axis) RGC types. Dashed red lines denote the x and the y axes. The relative predominance of points in the (+,+) and (-,-) quadrants compared to the (+,-) and (-,+) quadrants suggests that synergistic and antagonistic relationships between TFs are largely preserved across the species. Red boxes highlight several examples in which TF pairs show similar patterns of co-expression or mutually exclusive expression in both species. Expression of transcription factors in mouse RGC subsets is documented in (Cherry et al., 2011; Liu et al., 2018; Mao et al., 2014; Peng et al., 2017; Rousso et al., 2016; Sweeney et al., 2019).

Scale bar, 20 μm

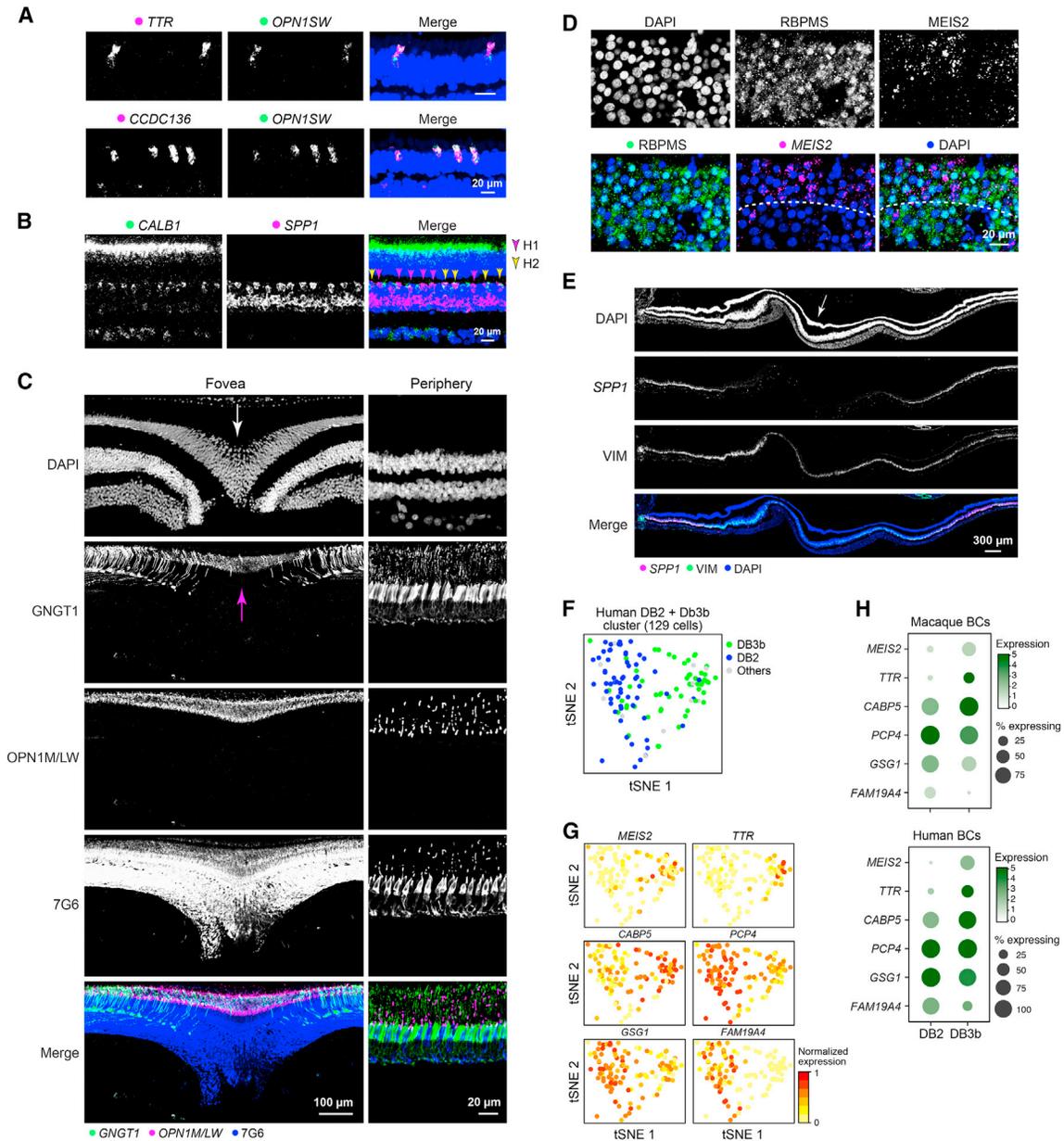


Figure S6. Key Molecular Features Identified from the Macaque Retinal Atlas Are Conserved in Marmosets and Humans, Related to Figure 6

A. *TTR* and *CCDC136* positive S cones in marmoset

B. H1 (*SPP1*+*CALB1*-) and H2 (*SPP1*+*CALB1*+) HC in the marmoset peripheral retina.

C. Foveal cones are GNGT1 positive in the marmoset. Staining marmoset foveal and peripheral retinal tissues with antibodies against GNGT1, OPN1LW, and 7G6 (cone arrestin) shows that GNGT1 is specifically expressed by foveal cones and localized at the outer segments (indicated by the expression of OPN1LW) of foveal cones. Arrows indicate the center of fovea.

D. The expression of *MEIS2* in the outer half of the human foveal GCL layer.

E. *SPP1* is expressed by peripheral Müller glia (*VIM*) respectively in human retina. Arrows indicate the center of fovea.

F. tSNE plot of cells from human bipolar cluster DB2+3b, color coded by the macaque peripheral bipolar type that each human cell mapped to. Gray indicating cells mapped to clusters other than DB2 and DB3b.

G. Expression of DB2 and DB3b genes in the DB2+3b cluster.

H. Expression of the same gene sets as G in macaque and human peripheral bipolar cells

Scale bar, 20 μ m (A, B, C, D), 100 μ m (C), and 300 μ m (E). DAPI staining is in blue.

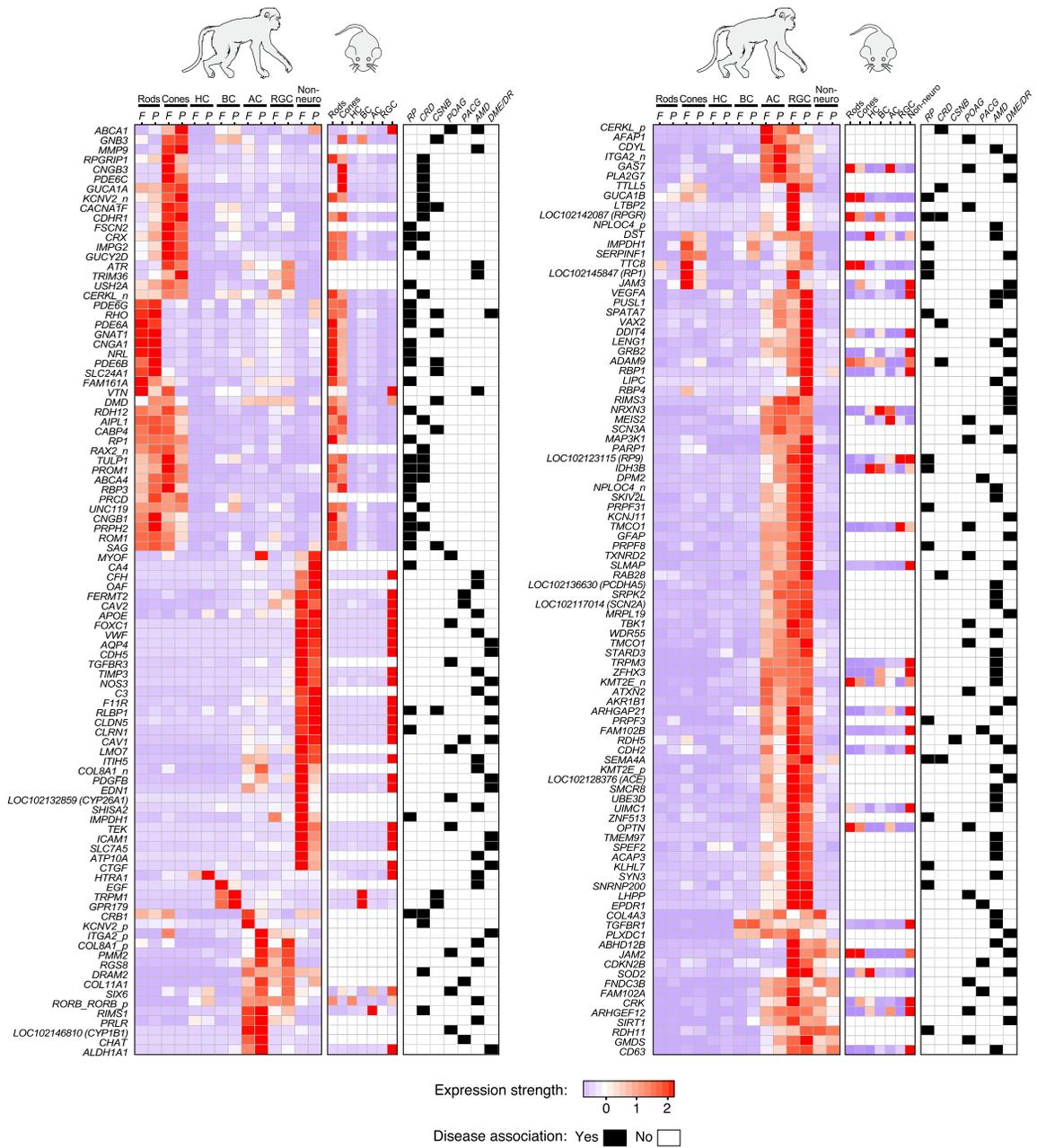


Figure S7. Expression Patterns of Retinal-Disease Associated Genes across Major Cell Classes in Macaque Fovea and Periphery, as Well as Mouse Retina, Related to Figure 7

Red-blue heatmaps show expression patterns of individual retinal-disease associated genes (rows) by cell classes (columns), for macaque and mouse respectively, as Figure 7B. For each gene, associated retinal diseases are shown in white-black heatmaps. For genes in the macaque transcriptome that have been mapped, but not characterized (e.g. *LOC102123115*), their predicted human orthologs are indicated in parenthesis.