bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Unified classification of mouse retinal ganglion cells using function, morphology, and gene expression

Jillian Goetz[1,*], Zachary F. Jessen[1,2,3,*], Anne Jacobi[4,5], Adam Mani[1,6], Sam Cooler[2,7], Devon Greer[1,2], Sabah Kadri[8,9], Jeremy Segal[9], Karthik Shekhar[5,10], Joshua Sanes[5], Gregory W. Schwartz[1,11,12]

*. equal contribution.
1. Department of Ophthalmology, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA.
2. Northwestern University Interdepartmental Neuroscience Program, Northwestern University, Evanston, IL, USA.
3. Medical Scientist Training Program, Northwestern University, Chicago, IL, USA.
4. F.M. Kirby Neurobiology Center, Department of Neurology, Boston Children's Hospital, Harvard Medical School, Boston, MA 02115, USA.
5. Center for Brain Science and Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA, USA.
6. Current address: Department of Neuroscience, Brown University, Providence, RI, USA.
7. Current address: Department of Neurosurgery, Stanford University, Stanford, CA, USA.
8. Current address: Department of Pathology and Department of Preventive Medicine, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA.
9. Department of Pathology, Pritzker School of Medicine, University of Chicago, Chicago, IL, USA.
10. Department of Chemical and Biomolecular Engineering, University of California Berkeley, Berkeley, CA, USA.
11. Department of Physiology, Feinberg School of Medicine, Northwestern University, Chicago, IL, USA.
12. Department of Neurobiology, Weinberg College of Arts and Sciences, Northwestern University, Evanston, IL, USA.

Abstract
Classification and characterization of neuronal types is critical for understanding their function and dysfunction. Neuronal classification schemes typically rely on measurements of electrophysiological, morphological and molecular features, but aligning these data sets has been challenging. Here, we present a unified classification of retinal ganglion cells (RGCs), the sole retinal output neurons. We used visually-evoked responses to classify 1777 mouse RGCs into 42 types. We also obtained morphological or transcriptomic data from subsets and used these measurements to align the functional classification to publicly available morphological and transcriptomic data sets. We created an online database that allows users to browse or download the data and to classify RGCs from their light responses using a machine-learning algorithm. This work provides a resource for studies of RGCs, their upstream circuits in the retina, and their projections in the brain, and establishes a framework for future efforts in neuronal classification and open data distribution.

Introduction
A major goal in biology is the establishment of a comprehensive atlas of cell types. Many large-scale efforts are underway to classify cells in different tissues (Hodge et al., 2019; Regev et al., 2017; Wilbrey-Clark et al., 2020; Yuste et al., 2020). In the central nervous system (CNS), classification efforts have relied mainly on three types of information: functional, morphological, and molecular. Functional classification involves the physiological properties of neurons, typically measured by electrophysiological recordings. Morphological classification uses the dendritic and axonal structures of neurons, typically measured by light microscopy and sometimes by electron microscopy (EM). Molecular classification relies on genomic readouts, particularly gene expression patterns (transcriptomics) where recent advances in single-cell RNAseq and spatial transcriptomics have made

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

large-scale measurements feasible (Close et al., 2021; Yuste et al., 2020). It has become increasingly clear that these different classification methods offer complementary information and that a comprehensive classification of cell types needs to unify all three modalities (Scala et al., 2020; Zeng and Sanes, 2017).

The mammalian retina is especially well suited to provide a template for integrating functional, morphological, and genetic classification for three reasons. First, many retinal cell types exhibit regular spacing, called a mosaic, which ensures smooth and complete sampling of visual space (Bleckert et al., 2014; Kay et al., 2012; Reese and Keeley, 2015; Rockhill et al., 2000; Rousso et al., 2016; Wässle et al., 1981). This property means that experimentalists can sample from a sub-region of the retina and be assured that they will find cells of each type. Moreover, mosaics establish an independent metric to assess whether a set of cells comprises an authentic type. Second, because the retina responds to light *ex vivo*, functional measurements of retinal neurons include both intrinsic biophysical properties and response properties to visual stimuli. Light responses depend on the entire upstream synaptic network, creating a rich data set. Finally, our knowledge of the morphology of retinal neurons, particularly in the mouse, is unparalleled among tissues of the mammalian CNS (Bae et al., 2018; Hoon et al., 2014; Sanes and Masland, 2015).

Here, we present a unified functional, morphological, and genetic classification of mouse retinal ganglion cells (RGCs), the output cells of the retina. We collected detailed functional data from 1777 RGCs, and also obtained morphological or transcriptomic data from subsets of them. We then used these doubly-characterized cells to align the functional classification with publicly available large-scale datasets of RGC morphology (381 RGCs reconstructed from EM sections; Bae et al., 2018) and gene expression (35,699 single-RGC transcriptomes(Bae et al., 2018; Tran et al., 2019), thereby generating a unified atlas. Comparison of the three datasets reveals that close relationships of types by one criterion predicts, albeit imperfectly, close relationships by the other criteria.

Finally, we provide two tools that make the data useful to the community and suggest formats for cross-modal analyses of other populations. First, we devised a machine-learning classifier that allows researchers to infer an RGC's functional type from a small and standardized set of spike measurements. Second, we curated the data in the form of a continuously updated, open-access library (rgctypes.org) from which researchers may browse single cell- or cell type-level data and download functional, morphological, and transcriptomic data sets.

Results

RGCs have traditionally been classified by physiological, morphological, and molecular criteria. Recent studies have used high-throughput methods to categorize mouse RGCs at large scale using all three criteria: optical imaging of visually evoked responses (Baden et al., 2016); reconstruction from serial electron microscopic sections; (Bae et al., 2018); and transcriptomic analysis of single RGCs (Tran et al., 2019). Our goal was to unify these dimensions into a single schema that was as complete as possible in representing all known RGC types in the mouse. We made our measurements in one cell at a time, allowing us to perform functional classification online followed by recovery of the same cells for morphological or transcriptomic measurements (**Figure 1**).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

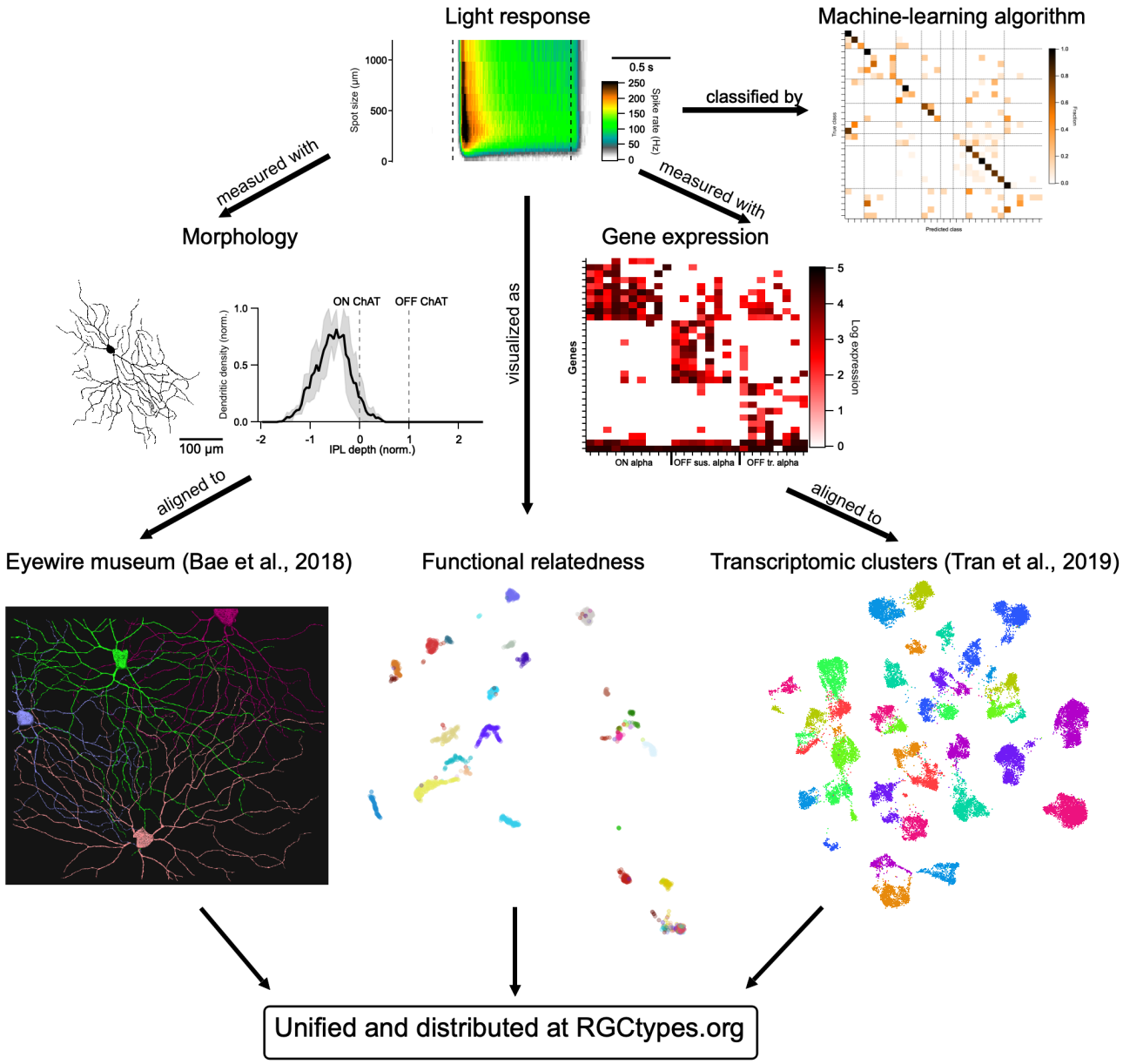

Figure 1. Schematic of approach.
Light responses to a standard stimulus set were measured in 1777 RGCs with subsets measured for morphology or gene expression. Functional data on light responses was visualized using UMAP (see **Methods**) and classified with machine-learning. Morphological and transcriptomic measurements] were aligned to published datasets.

Functional classification of RGCs.

We began with physiological characterization, using a rapid and standardized light stimulus protocol for functional measurements. Experiments were performed in dark-adapted *ex vivo* preparations of the mouse retina where capacitive spikes from RGCs were recorded with cell-attached electrodes. Standard light stimuli presented to every RGC were rod-saturating (~200 isomerizations/rod/s) spots (λ = 450 nm) from darkness with diameters ranging from 30 – 1200 µm, centered on the receptive field (RF) of each individual cell. We presented additional stimuli to subsets of RGCs to test for specific forms of feature selectivity. Moving bars tested direction selectivity (DS), flashed bars and drifting gratings tested orientation selectivity (OS), and contrast series tested contrast suppression (**Figure S1**).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Our standard stimulus paradigm differed from the full-field “checkerboard” white noise and “chirp” stimuli used in previous studies (Baden et al., 2016; Farrow and Masland, 2011; Jouty et al., 2018). Three considerations drove this stimulus choice. First, maintenance of a consistent light-adaptation state was essential because many aspects of RGC light responses change with luminance and with light adaptation (Tikidji-Hamburyan et al., 2015; Wienbar and Schwartz, 2018). High background light is unavoidable in two-photon imaging experiments due to excitation from the laser, limiting the time of stable light responses, especially in preparations lacking the retinal pigment epithelium (Euler et al., 2019). Use of patch electrodes allowed us to make our measurements from darkness. Second, precise localization of stimuli with respect to the RF center cannot be achieved with full-field stimuli but turns out to be critical as shown below. Indeed, many RGCs that respond well to small stimuli in their RF center failed to respond to any full-field stimulus (Jacoby and Schwartz, 2017). Finally, to facilitate standardization in the field, we wanted our stimulus to be simple and rapid and to correspond to those commonly used by others. For example, many previous studies have used RF-centered spots of different sizes, enabling retrospective comparisons. (Jacoby and Schwartz, 2017; Johnson et al., 2018; Krieger et al., 2017; Marco et al., 2013; Rousso et al., 2016)

We constructed a decision tree for functional classification based only on the responses to spots of light with varying size (**Figure S2**). Response properties for additional stimuli are listed as confirmatory tests, and in the cases of DS and OS RGCs, they were needed to subdivide types based on direction or orientation preference. Our functional classification separates RGCs into 42 types: 34 from the responses to spots, and the remainder from direction or orientation preference. We organized the RGC types in 8 functional groups: ON sustained, OFF sustained, ON transient, OFF transient, ON OS, DS, ON-OFF small RF, and Suppressed-by-Contrast (SbC)/Other. These groups were chosen as a starting point based on previous work; a formal, quantitative grouping by functional relatedness is presented below.

In most recordings, retinal orientation and cell locations were noted to determine whether classification varied based on retinal position. Response patterns within some RGC types have been shown to vary as a function of retinal position in photopic conditions (Joesch and Meister, 2016; Warwick et al., 2018), likely because of a pronounced cone opsin gradient along the dorsoventral axis (Nadal-Nicolás et al., 2020). In our dark-adapted preparation, however, where much of the light response was initiated in rods (Grimes et al., 2014), response variation across retinal position was minimal. Functional types were found relatively uniformly across retinal locations with four exceptions: we found a dorsal position bias for sSbC EW27 and both ONOS SmRF types, and we found a ventral bias for ON small OFF large RGCs (**Figure S3**).

Data from 37 RGC types are presented in 3 ways in **Figure 2**. (Only 37 of the 42 types are illustrated because DS RGCs with different directional preferences did not differ from each other in their responses to light spots.) We first measured the response polarity and kinetics of light responses with a 200 μm light spot centered on the RF (marked “a” in the first panel, which shows an ON alpha RGC). This allowed us to assign cells according to response polarity (ON, OFF, ON-OFF or Suppressed-by-Contrast) and as having sustained or transient responses to luminance changes. Second, to assess how each RGC type’s response varied with stimulus size, we measured the total ON and OFF spike responses for spots of 12 sizes from well below the RF center diameter of the smallest RGC (30 μm) to a size that reached the far RF surround (1200 μm) (marked “b” in the ON alpha RGC panel). This information was critical in separating many types. For example, despite similar responses to the 200 μm spots, ON-OFF DS, HD1, HD2, and UHD all had different response profiles of their ON and OFF responses across spot size. For some RGC types, even the overall polarity of the light response depended on spot size. For example, HD2 RGCs are ON-OFF for small spots and ON for large spots (Jacoby and Schwartz, 2017) and the ON small OFF large RGC switches polarity entirely with spot size as its name suggests. Finally, we combined information about response amplitude and kinetics as a function of spot size into a single plot using a heatmap of firing rate over time for each spot size (marked “c” in the ON alpha RGC panel).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

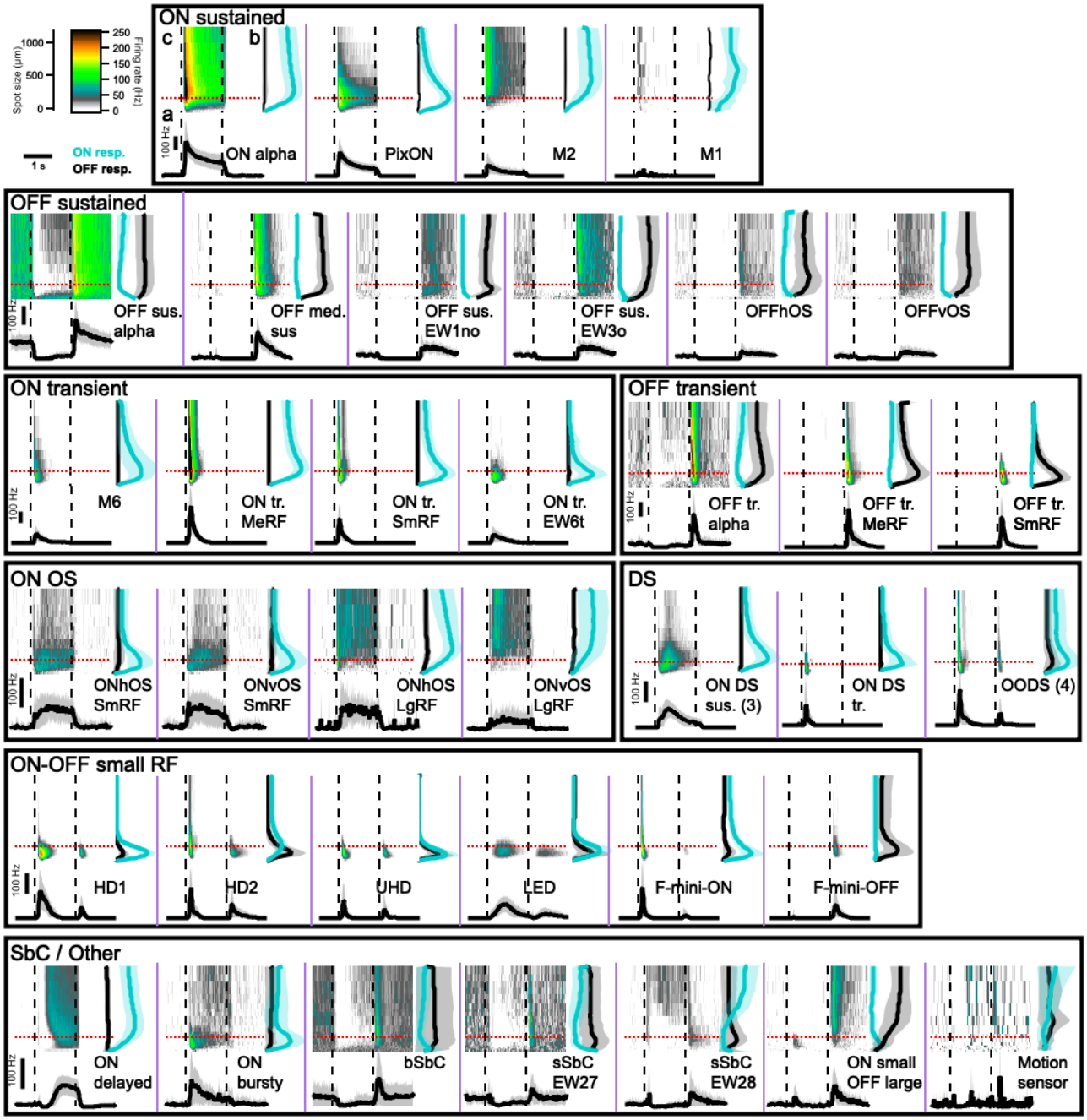

Figure 2. Functional diversity of mouse RGCs.
Each panel (separated by purple lines) contains 3 graphs showing the light response of an RGC type to flashed spots of light (200 R*/rod/s) from darkness. The top left graph (marked 'c' in ON alpha panel) is a heatmap of average firing rate over time (x-axis) for spots from 30 – 1200 µm (y-axis). Dashed lines show the time of spot onset and offset. The top right graph (marked 'b' in ON alpha panel) shows the total spike count during flash onset (cyan) and offset (black) for each spot size. The solid lines indicate mean across cells and the shaded regions indicate standard deviation (s.d.). The bottom graph (marked 'a' in ON alpha panel) shows peristimulus time histogram (PSTH) plots averaging the response of each cell type to 200 µm spots, indicated in upper plots by red dotted lines. Scale bars in the upper left region are shared across all graphs. Separate scale bars for the y-axis of the PSTH plots are provided within each boxed group of cells and apply within that box. Abbreviations for cell types: sus. = sustained; tr. = transient; med. = medium; EW = Eyewire (named based on the Eyewire museum); OS = orientation selective; h = horizontal; v = vertical; DS = direction selective; SmRF = small receptive field; MeRF = medium receptive field; LgRF = large receptive field; HD = high definition; UHD = ultra high definition; LED = local edge detector; (b,s)SbC = (bursty, sustained) suppressed-by-contrast.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Automated functional classification

To validate our functional classification (**Figure S2)**, we implemented a machine-learning classifier to assign RGCs to types based on a feature set comprising spike responses to spots of varying size (see **Methods**). Since the responses to moving bars and drifting gratings were not included in the feature set, we collapsed DS and OS cells across direction and orientation, respectively. We also removed 3 sparsely sampled RGC types with insufficient data to train the classifier (M1, OFF sus. EW3o, and Motion sensor). Together, our filtered dataset comprised 1766 RGCs across 31 types, which we split into a training set (n=1362 RGCs) and a testing set (n=404 RGCs). Other than ensuring that at least 5 cells of each type were represented in the test set, separation of training and test sets was random.

Following training, the performance of the classifier was evaluated on the test set (**Figure 3)**. For each cell, the classifier outputs the probability of membership in each RGC type. Thus, the algorithm provides both a “best guess” and a confidence rating for each prediction. An advantage of probabilistic scoring is that the classifier predictions can easily be updated to include complementary sources of information (e.g. prior probabilities based on stratification depth in the inner plexiform layer, IPL, or labelling in a transgenic line) via Bayes’ rule (MacKay and Mac, 2003)

Without thresholding the probability scores, classification accuracy was 64% overall, exceeded 80% for 9 RGC types, and was 100% for 2 types (**Figures 3B,D**). The correct RGC type was among the top three choices of the classifier 86% of the time (**Figure 3A, inset**), suggesting that additional information (functional, structural or molecular), could be used to refine its predictions. The distribution of classification accuracy was bimodal with 7 types falling within 1 s.d. of the chance level of 1/31 (3.2%) and the other 24 having a median accuracy of 74% (**Figure 3B**). Overall accuracy scaled linearly with unclassified fraction as we increased the classification margin, i.e. minimum probability score at which cells are assigned a type label (**Figure 3A**). Cells with maximal class probabilities below the classification margin are considered “unclassified”. Increasing the classification margin to 0.32 achieved an accuracy of 80% across the whole data set with 31% of cells unclassified (**Figure 3E**). The most significant limitation of our classifier was the size of the training set. RGC types with >40 examples in the training set performed much better (median 79% correct) than those with fewer (median 7% correct)(**Figures 3B,C**). Thus, we expect classifier performance to improve steadily as we continue to collect more data, particularly from rare RGC types. Updated results, newly trained versions, and tutorials for formatting data and running it through the classifier will continue to be made available at rgctypes.org.

Functional relatedness of RGC types

To visualize the relationships between functional RGC types, we used uniform manifold approximation and projection (UMAP)(Becht et al., 2018; McInnes et al., 2018)(**Figure 4**). The UMAP algorithm assigned each cell to a point in 2D space based only on its response to spots of varying size (the data in **Figure 2**) with closely-related cells projecting to nearby locations in this space. We did not include the moving bar or drifting gratings responses as input to the UMAP algorithm because they were not measured for every RGC. Therefore, DS RGCs with different direction preferences and OS RGCs with different orientation preferences were grouped together in this representation. Most RGC types formed clear clusters in UMAP space with a few exceptions for types that were sampled sparsely in our dataset (e.g. ON bursty, sSbC EW27)(**Figure 4A**).

Using distances between the centroids of the RGC types in UMAP space, we constructed a dendrogram representing functional similarity among RGCs within this limited stimulus space (**Figures 4B,C**). As expected, ON and OFF RGC types generally formed distinct clades. In other respects, however, this grouping differed substantially from our historical grouping (**Figure 2**). For example, “sustained” and “transient” types were extensively mixed, indicating that this distinction is not as strong an organizing property as polarity. Interestingly, a major distinction was based on the presence of strong surround suppression (**Figures 4B,C**). The RGC types in this group were either completely silent or nearly so for the largest spot size (see **Figure 2**).

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

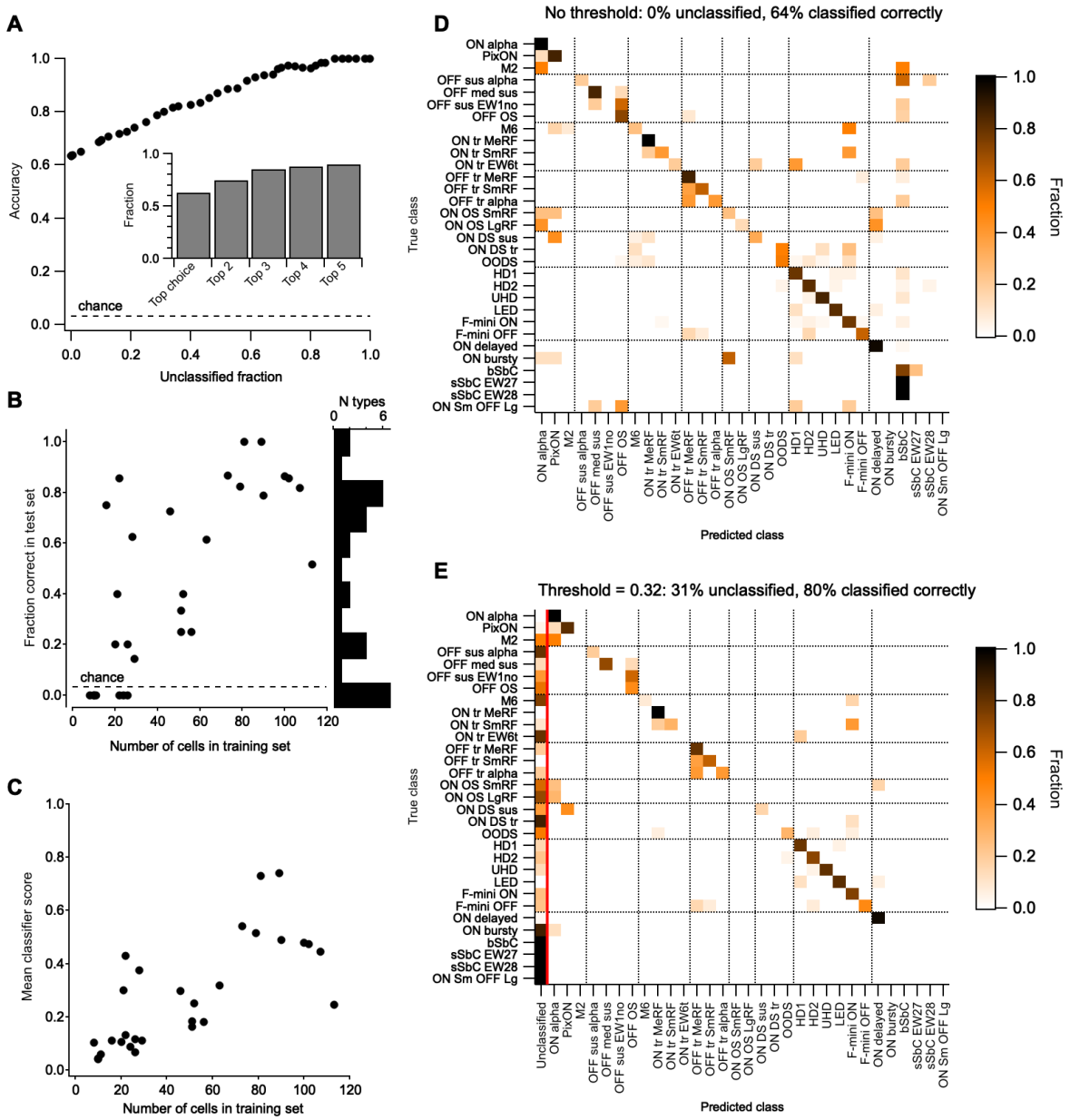

Figure 3. Functional classification from spot responses.
(**A**) Overall model accuracy (y-axis) as a function of the fraction of unclassified cells in the test cells (x-axis), which increases with the classification margin. The dashed line represents the expected accuracy of a random classifier. *Inset*, fraction of instances when the correct choice was present among the top 1-5 probability scores in the classifier output.
(**B**) Fraction of test cells of each type classified correctly versus the number of cells of that type in the training set. Histogram at the right shows the distribution of classifier accuracy across RGC types.
(**C**) Mean classification score for test cells of each type versus the number of cells of that type in the training set.
(**D**) Confusion matrix (row normalized) for the classifier with no explicit classification margin set. Dotted lines separate RGC groups as in Figure 2.
(**E**) Confusion matrix (row normalized) for the classifier with a classification margin of 0.62. The fraction of unclassified cells of each type is shown in the first column. Remaining entries in the matrix only consider classified cells.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

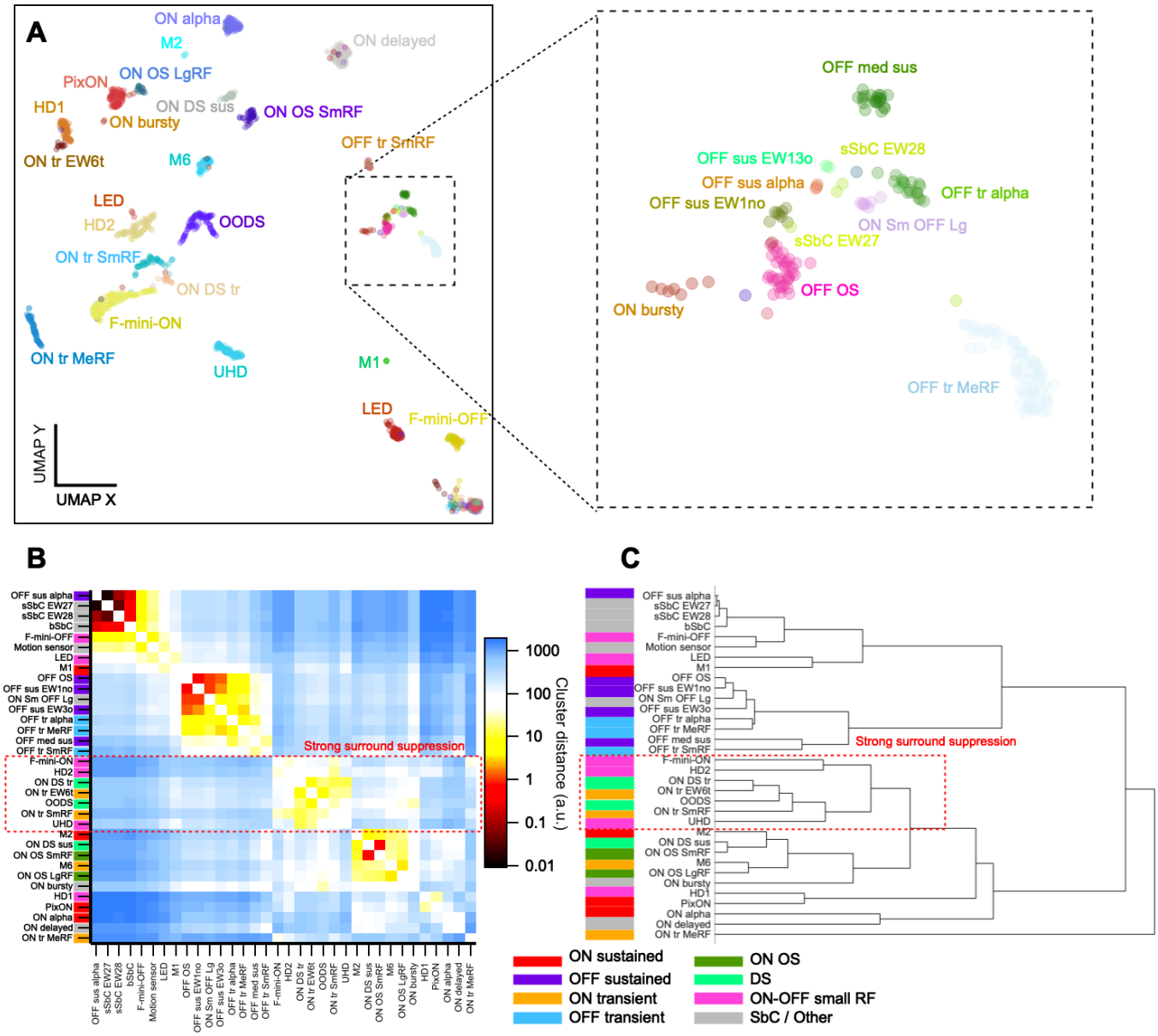

Figure 4. Visualization of functional relationships among RGCs.
(**A**) UMAP projection of 1777 RGCs colored by assigned functional type. Inset shows magnified view of boxed region.
(**B**) Heatmap showing pairwise distances among RGC types based on their centroids in UMAP space. Types (rows and columns) are ordered based on a dendrogram in panel **C**. Groups from Fig. 2 are shown as colors along the left as indicated in the legend below. Red dashed lines indicate a group of types with strong surround suppression.
(**C**) Dendrogram computed using agglomerative hierarchical clustering showing a hierarchy of functional similarity among RGC types. The strong surround suppression group is highlighted.

Alignment of functional and morphological classification

After we recorded visually evoked responses from RGCs, we filled 180 with Neurobiotin to measure *en face* morphology and stratification patterns within the IPL (**Figure 5**). Patterns were matched with the morphological types defined in the Eyewire museum (Bae et al., 2018). Despite some limitations mentioned below, the

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

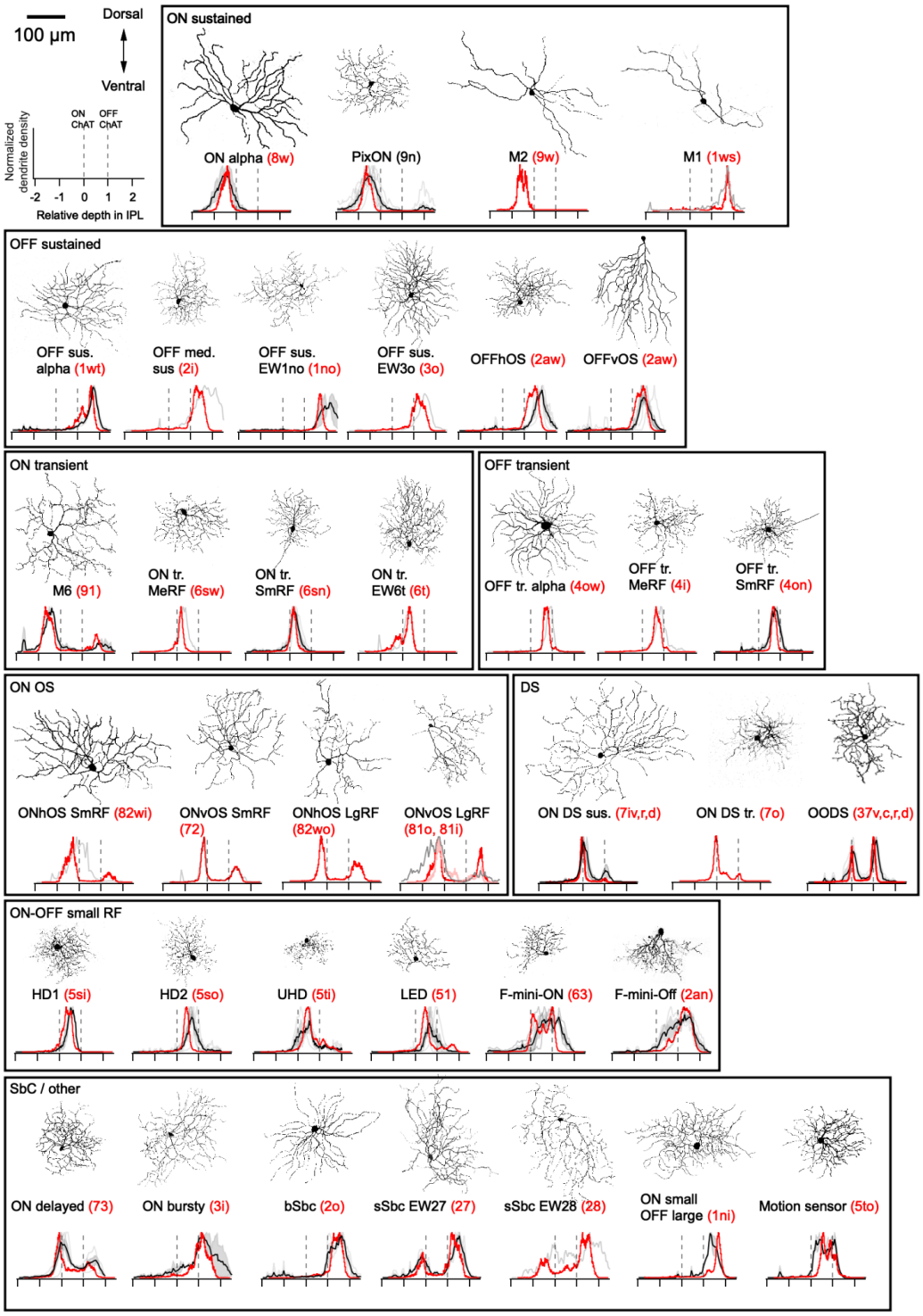

Figure 5. Morphological diversity of mouse RGCs

Each panel shows a representative *en face* image of an RGC of the specified type (name in black) and the matching type in the Eyewire museum (name in red). Below each panel is a graph of normalized dendritic stratification in the IPL. Mean stratification profile from the Eyewire museum is shown in red. Individual cells from the present dataset are shown in thin grey lines with their mean in black and their s.e.m. in shaded bars, where appropriate. Dotted lines on the stratification plots indicate the ON and OFF choline acetyltransferase (ChAT) bands. Scale bar and arrows for dorsal and ventral directions on the retina apply to all images.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

morphology of each functionally defined RGC type was consistent with that of a single morphologically-defined type in the Eyewire museum (comparisons available at rgctypes.org). In some cases, details of the morphology were instructive in separating otherwise similar types. For example, despite their similar stratification profiles, the ON transient MeRF and ON transient SmRF types were matched to Eyewire types 6sw and 6sn, respectively, based on differences in dendritic area (SmRF smaller), tortuosity (SmRF more tortuous), and branch angle (SmRF more sharp angles) (**Figure S4**).

A central conclusion of the work of (Bae et al., 2018) was that stratification patterns in the IPL were extremely stereotyped within RGC types. Thus, an RGC's stratification profile was the dominant factor in specifying its morphological type. After computationally flattening our images with respect to the ChAT bands (Sümbül et al., 2014), we matched each stratification pattern to one of the types in the Eyewire museum (**Figure 5**). The fits were excellent for most types despite the limitations of our lower Z-resolution and potentially different warping due to the different fixation methods (most evident above and below the ChAT bands).

Alignment of functional and transcriptomic classification

Recent large-scale investigations of single-cell transcriptomes in the retina have identified ~45 molecularly distinct types of postnatal mouse RGCs, comparable to the number of RGC types identified through physiological and morphological analyses (Rheaume et al., 2018; Tran et al., 2019). While some clusters could be matched 1:1 with previously known types based on well-established molecular markers (Sanes and Masland, 2015), approximately 40% of clusters remained unmatched. Moreover, these methods used dissociated tissue, precluding direct harmonization of gene expression with function.

To relate functional to molecular criteria, we used a variant of the Patch-seq technique (Cadwell et al., 2016) in which RGCs were first classified based on their cell-attached light responses and then the cytoplasm was collected for RNA-seq by aspirating the soma with a clean pipette (see **Methods**). We obtained 91 high-quality single RGC transcriptomes (>2000 genes/cell). We used gradient boosted decision trees (Chen and Guestrin, 2016) to match each of our transcriptomes to a cluster in the published adult RGC dataset (Tran et al., 2019)(see **Methods**). Many of our functionally-identified cells matched the transcriptomic clusters with high concordance (**Figure 6A**) providing putative matches to previously unknown clusters. For example, the three types of ON DS sus. cells all aligned to C10 (an uncharacterized type), ON tr. SmRF aligned with C21, corresponding to T-RGC S2 (Liu et al., 2018) and ON delayed (Mani and Schwartz, 2017), previously observed in *CCK-ires-Cre* mice (Jacoby and Schwartz, 2018; Tien et al., 2015) aligned with a cluster (C14), which was distinguished by the expression of the neuropeptide *Cck*.

T5-RGCs share a functional and morphological profile

Alignment of our physiologically characterized types to transcriptomically defined RGC groups (Tran et al., 2019) enables a deeper analysis of the relationships between gene expression of RGCs and their function and morphology. One example is provided by *Tusc5* (also known as *Trarg1*), which we identified as a key marker of a group of 9 mostly unidentified transcriptomic clusters termed T5 RGCs (Tran et al., 2019). Most of these RGCs are labeled by the transgene TYW3, which exhibits insertion-site dependent expression (Laboulaye et al., 2018). T5-RGCs stratify their dendrites between the ChAT bands and possess transient light responses and strong surround suppression. Seven of the T5-RGC clusters have now been assigned to functional RGC types, and they all fit this functional and morphological profile. Further, the matched T5 RGC types comprise 6 of the 11 morphological types that stratified the majority of dendrites between the ChAT bands. Finally, 4 of the 5 Tusc5-negative RGCs with this dendritic stratification have different functional characteristics (**Figure 6B,C**). Other subclasses of RGCs can be queried in this way, with increasing power as additional data is added to rgctypes.org.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

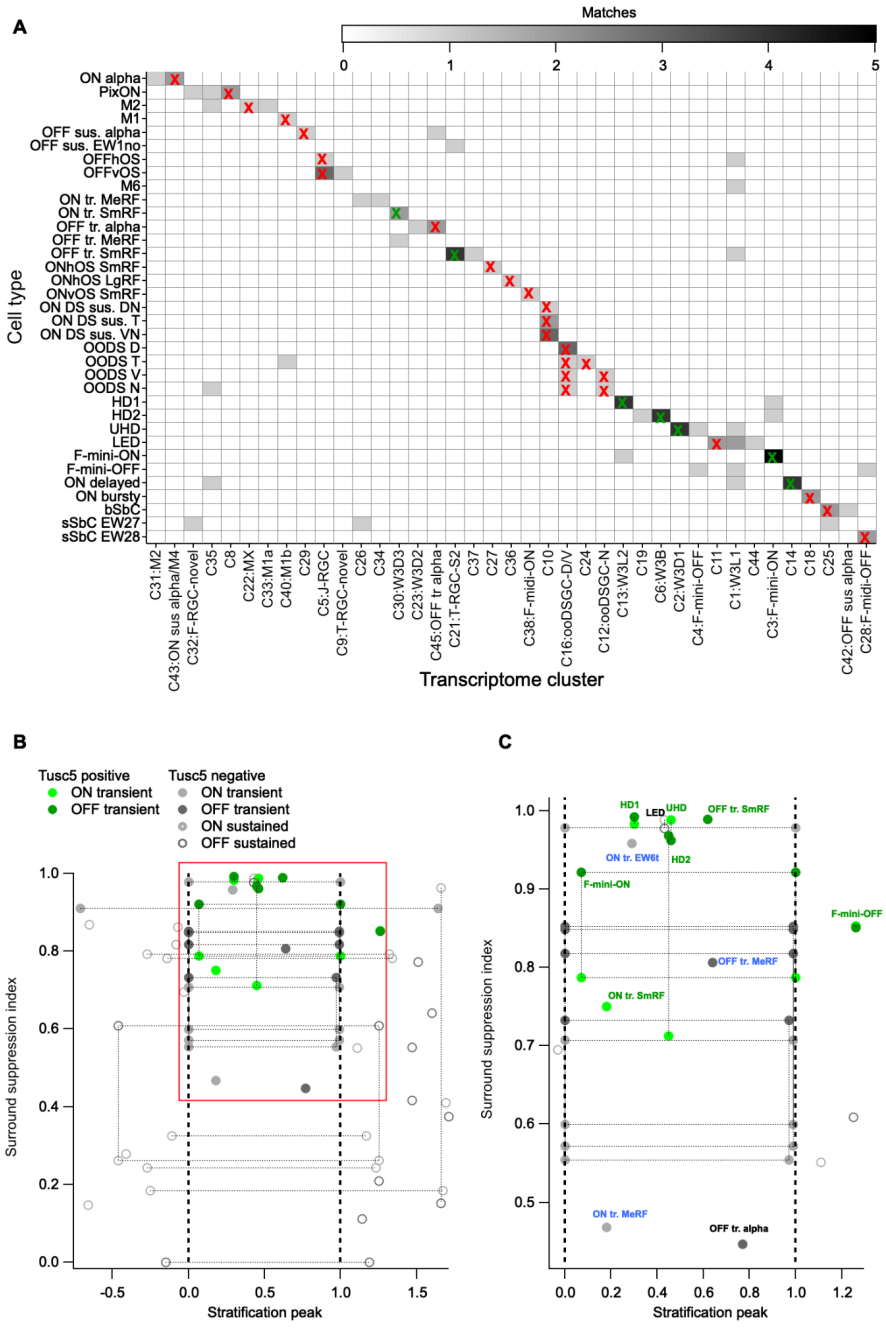

Figure 6. Matches between functional types and transcriptomic clusters.

(**A**) Heatmap showing correspondence between functional types (rows) and transcriptomic clusters reported in Tran et al., 2019 (columns). Matches used in subsequent analyses are indicated by an 'X'. Green Xs indicate matches to transcriptomic clusters identified as T5-RGCs, characterized by the specific expression of gene *Tusc5/Trarg1*, in (Tran et al., 2019).

(**B**) Surround suppression index (see **Methods**) is plotted against the peak stratification depth in the IPL for each RGC type. Both ON responses (light gray) and OFF responses (dark gray) are plotted. Dotted lines connect data points from the same type: horizontal lines for bistratified types and vertical lines for ON-OFF types. Filled symbols indicate transient cells. RGC types aligned to Tusc5-positive transcriptomic clusters are colored green (light for ON; dark for OFF). Dotted lines at 0 and 1 stratification level indicate ON and OFF ChAT bands, respectively.

(**C**) Magnified view of the red boxed region in (**B**). RGC types aligned to Tusc5-positive clusters are labeled with green text; those aligned to Tusc5-negative clusters are labeled with black text; types lacking a transcriptomic match are labeled with blue text. The unlabeled RGC types along the ChAT bands are ON-OFF DS RGCs.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The question of completeness

One way to estimate the completeness of our classification is to record nearly all the RGCs in a small region of the retina and count how many can be assigned to one of our types. We performed such an experiment and stained the tissue with the pan-RGC marker gamma-synuclein (Surgucheva et al., 2008) to confirm RGC identity post-hoc (**Figure S5**). We recorded 55 spiking cells and 25 cells for which we could not elicit spikes with our test stimuli. Of the 55 spiking cells, 48 were successfully identified in the fixed tissue. In the live tissue we had labeled 42 of these cells as RGCs matching one of our types and 6 as spiking amacrine cells. All 48 of these identifications were verified by the gamma-synuclein staining (42/42 gamma-synuclein positive RGCs and 6/6 gamma-synuclein negative spiking amacrine cells). Of the 25 cells for which we could not elicit spikes, 22 were identified in the fixed tissue: 10 were gamma-synuclein negative, presumably non-spiking amacrine cells, and 12 were gamma-synuclein positive, presumably RGCs that we failed to identify. Thus, we identified 78% (42 / 54) of the putative RGCs in this sample. While somewhat less than our estimate of 89% coverage of the types in the Eyewire museum (**Table 1**), it is a conservative estimate because some of the non-responding RGCs were likely damaged during removal of the inner limiting membrane or by the recording procedure and did not spike (e.g. because the axon initial segment was destroyed) but survived enough structurally for gamma-synuclein staining.

Relatedness of functional, morphological, and transcriptomic space

The main goal of our study was to directly relate physiological, structural and molecular definitions of cell type. With these data in hand, we were able to address an additional question: to what extent do relationships among types depend on the criteria used to define them. To this end, we constructed a UMAP embedding based on morphology using the stratification profiles of each cell in the Eyewire museum (Bae et al., 2018)(**Figure 7A**). We then compared nearest neighbors in this space to those in the UMAP embedding of physiological relatedness among RGC types described above (**Figure 4A**) and the UMAP embedding of molecular relatedness generated from the single-cell transcriptomic data (Tran et al., 2019; reproduced in **Figure 7B**).

We then queried the set of alignments between RGC types in these three different spaces (**Figure 7C**). For each RGC type, we computed the fractional overlap among its nearest neighbors in one of these maps with its nearest neighbors in the other maps. We repeated this analysis for neighborhood sizes from 2 to 12 nearest neighbors. To establish statistical significance on this fractional overlap measure, we used the bootstrap approach. We randomly shuffled type identities in each of the maps and recomputed the fractional overlap. Repeating this process 1000 times yielded an empirical null distribution. Fractional overlap values obtained from the real data are reported as z-scores relative to this null distribution with positive values indicating greater overlap in the real data than in the null distribution (**Figure 7D–F**).

Each of the three pairwise comparisons between modalities had a preponderance of positive fractional overlap z-scores, indicating significant alignment of type-type interrelationships above that expected by random chance. Some types, such as the ON alpha, PixON, and M2 RGCs showed very high correspondence between functional and morphological UMAP spaces for a range of small to medium neighborhood sizes (**Figure 7D**). For other types, correspondence was lower and in some cases types had negative overlap scores, meaning that their neighbors in functional space (particularly for large neighborhoods) are more different from those in morphological space than one would expect by chance. The map alignment from function to gene expression (**Figure 7E**) was globally similar to the function-morphology alignment, but with somewhat different RGC types showing high vs. low correspondence. The map alignment between morphology and gene expression (**Figure 7F**), while still weighted toward positive overlap scores, had overall weaker correspondence than the other two alignments.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

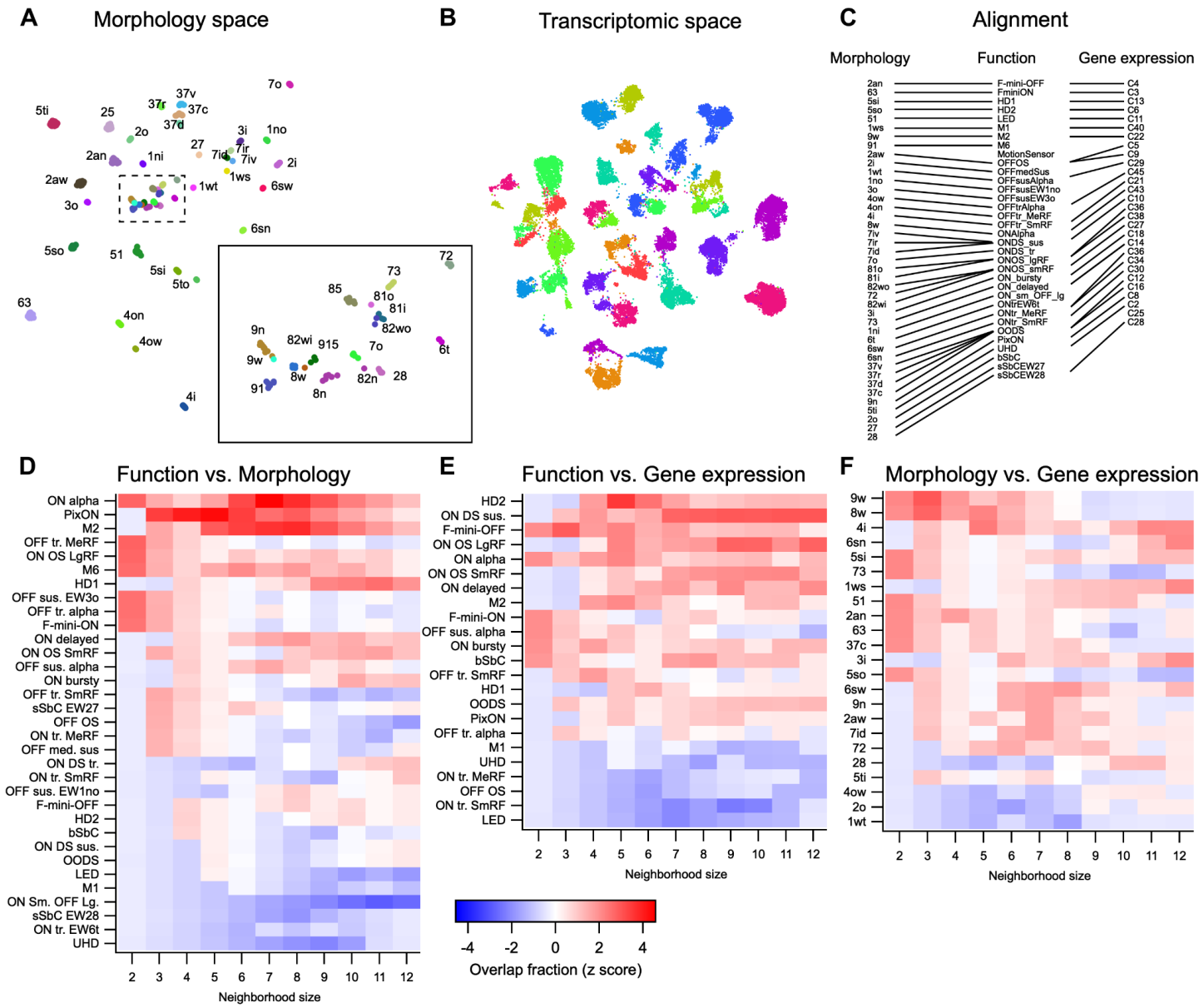

Figure 7. Correspondence between RGC relatedness in functional, morphological, and transcriptomic space.
(**A**) UMAP embedding of RGC morphology constructed from the stratification profiles in the Eyewire museum (Bae et al., 2018). Inset shows boxed region at higher magnification.
(**B**) UMAP embedding of RGC gene expression from Tran et al. (2019). Cluster labels removed for clarity.
(**C**) Alignments between the three classification schemes that we used for subsequent analysis. Lines connect putative corresponding RGC types in each classification schema.
(**D**) Heatmap of fractional overlap z scores for the correspondence between functional and morphological UMAP spaces. Each row shows the overlap score between maps for a particular RGC type across different neighborhood sizes.
(**E, F**) Same as (**D**) but showing alignment between functional and morphological space (**E**) or morphological and gene expression space (**F**).

Integrated web-based RGC compendium

A major goal of this project was to create a resource so that labs around the world can come to a consensus on the classification of mouse RGCs. To that end, we have developed a website, rgctypes.org (**Figure 8**), with a direct pipeline to our database of functional and morphological measurements. Following a curation step and type assignment, every RGC recorded in the Schwartz lab will automatically update to rgctypes.org. Other

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

researchers are invited to submit data for integration as well. Cells can also be reassigned to different types if evidence supports a different assignment. Full datasets are available for download immediately, regardless of publication status. We have also provided a downloadable version of our automated classifier and instructions on how to prepare a data file to obtain a type prediction and confidence score.

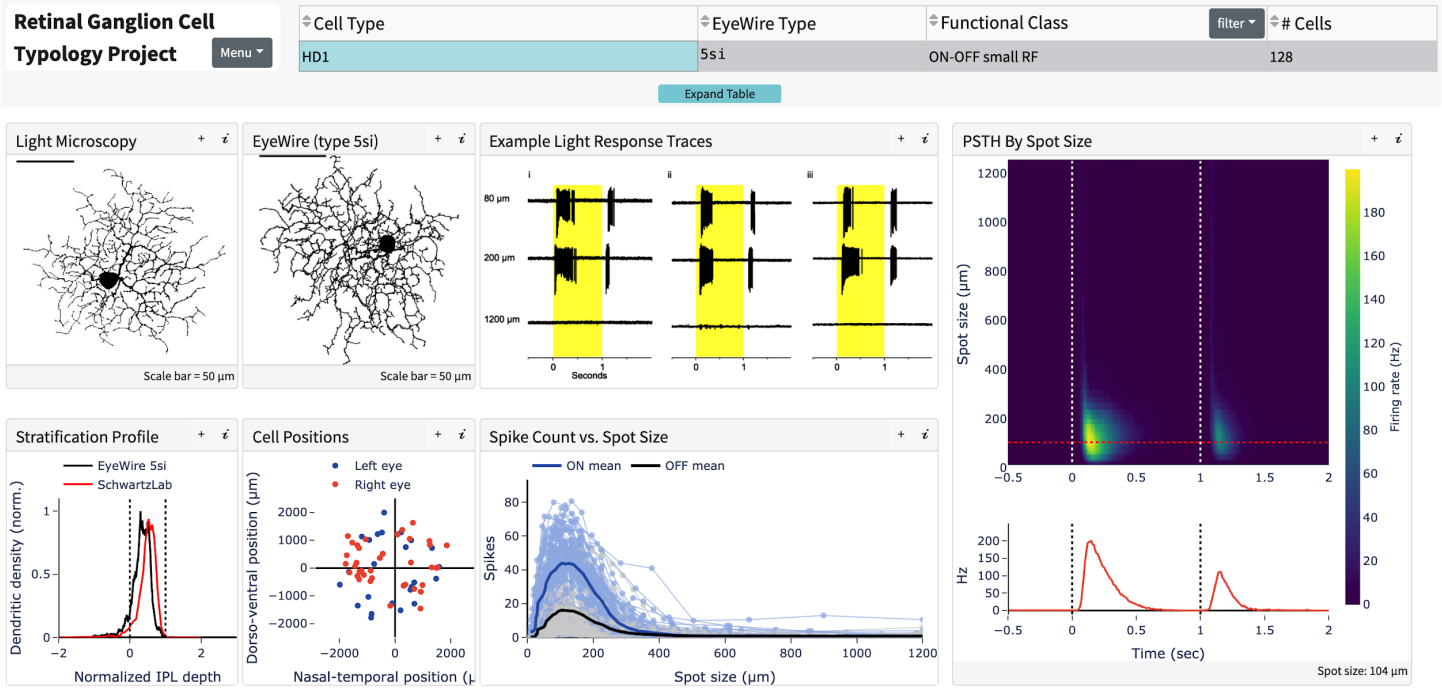

Figure 8. Screenshot from rgctypes.org.

Discussion

We present a resource of physiological, morphological, and transcriptomic data aimed at establishing a comprehensive typology of mouse RGCs. A summary of our classification and its alignment with previous RGC classifications is provided in **Table 1**.

<u>Method for functional classification</u>

We recorded from RGCs one at a time, which allowed us to center stimuli on the receptive field of each cell. This undeniably limits throughput. On the other hand when activities of many RGCs are recorded simultaneously – for example by calcium imaging or with multielectrode arrays – it is not feasible to center stimuli on individual RGCs, so these studies have used a combination of full-field modulation, large moving objects or gratings, and spatiotemporal white noise. These stimulus choices come with a significant cost. Many RGC types, including some of the most numerous types, respond poorly or not at all to full-field stimuli or spatiotemporal white noise(Jacoby and Schwartz, 2017). Other types respond both to small (RF-centered) and large stimuli, but basic response properties depend on spot size. For example, the ON small OFF large RGC would be classified as an OFF cell for full-field stimulation but responds as an ON cell for small spots in its RF center. Surround suppression differentially affects both the total spike count and response kinetics in most RGC types (**Figure 2**)(Wienbar and Schwartz, 2018), providing information that we found necessary to separate otherwise functionally similar types.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

RGC type	Eyewire type	Eyewire density (%)	Baden et al. (2016) Group	Baden et al. (2016) match confidence	Tran et al. (2019) Cluster	Physiology *n*	Morphology *n*	Single-cell RNA-seq *n*	References
ON alpha	8w	1.0	24	high	43	89	10	3	(Krieger et al., 2017)
PixON	9n	2.4	22a,b	low	8	100	24	4	(Johnson et al., 2018)
M2	9w	0.3				11	1	2	(Schmidt and Kofuji, 2009)
M1	1ws	0.5			40	4	1	1	(Berson et al., 2010; Do et al., 2009)
OFF sus. alpha	1wt	0.8	5a,b,c	high	29	26	3	2	(Krieger et al., 2017)
OFF med. sus.	2i	3.1	3	low		22	3	0	
OFF sus. EW1no	1no	2.1	4a	medium		10	3	1	
OFF sus. EW3o	3o	1.0	4b	medium		4	2	0	
OFFhOS	2aw	5.5	1, 14	medium	5	17	9	1	(Nath and Schwartz, 2017)
OFFvOS	2aw	5.5	1, 2, 6, 14	medium	5	29	1	3	(Kim et al., 2008; Nath and Schwartz, 2017)
M6	91	1.8	20	low		58	3	0	(Quattrochi et al., 2019)
ON tr. MeRF	6sw	1.3	18a	low		82	4	2	
ON tr. SmRF	6sn	1.6	18b	low	30	54	5	2	
ON tr. EW6t	6t	0.8	21	low		20	4	0	
OFF tr. alpha	4ow	1.0	8a,b	high	45	22	1	3	(Krieger et al., 2017)
OFF tr. MeRF	4on	3.1	9	low		73	2	1	
OFF tr. SmRF	4i	2.9	9	low	21	28	5	3	
ONhOS SmRF	82wi	0.8	17a,b,c	medium	27	24	3	0	(Nath and Schwartz, 2016)
ONhOS LgRF	82wo	0.5	30	medium	36	11	1	0	
ONvOS SmRF	72	1.3	17a,b,c	medium	38	27	0	1	(Nath and Schwartz, 2016)
ONvOS LgRF	81o, 81i	1.0	30	medium		16	1	0	
ON DS sus. DN	7iv	0.5	25	high	10	11	0	1	(Estevez et al., 2013)
ON DS sus. T	7ir	1.0	26,29	high	10	12	1	2	(Estevez et al., 2013)
ON DS sus. V	7id	0.8	26,29	high	10	13	0	3	(Estevez et al., 2013)
ON DS sus. - direction unknown	7i*		25,26,29	high	10	16	2		
ON DS tr.	7o	1.8	16	medium		22	0	0	(Gauvain and Murphy, 2015)
OODS D	37v	2.4	12a,b,13	high	16	20	1	4	(Kay et al., 2011)
OODS T	37r	2.6	12a,b,13	high	16,24	40	2	4	(Kay et al., 2011)
OODS V	37d	1.6	12a,b,13	high	16,12	21	1	0	(Trenholm et al., 2011)
OODS N	37c	2.4	12a,b,13	high	16,12	21	2	2	(Kay et al., 2011)
OODS - direction unknown	37*		12a,b,13	high		11	1		
HD1	5si	2.6	10,11a,b	low	13	90	8	2	(Jacoby and Schwartz, 2017)
HD2	5so	4.5	10,11a,b	low	6	107	14	5	(Jacoby and Schwartz, 2017)
UHD	5ti	7.9	10,11a,b	low	2	102	8	3	(Jacoby and Schwartz, 2017)
LED	51	5.2	10,11a,b	low	11	79	4	3	(Jacoby and Schwartz, 2017)
F-mini-ON	63	6.6	10,11a,b	low	3	177	5	6	(Cooler and Schwartz, 2020; Rousso et al., 2016)
F-mini-OFF	2an	6.6				63	5	4	(Cooler and Schwartz, 2020; Rousso et al., 2016)
ON delayed	73	1.8	27,28a,b	low	14	161	16	5	(Mani and Schwartz, 2017)
ON bursty	3i	1.6	27	low	18	26	4	2	
bSbC	2o	1.3	32a,b,c	medium	25	16	9	3	
sSbC EW27	27	1.0	31a,b,c,d,e	low		8	4	2	(Jacoby et al., 2015)
sSbC EW28	28	1.3	31a,b,c,d,e	low	28	24	1	1	(Jacoby et al., 2015)
ON Sm. OFF Lg.	1ni	1.3				10	3	0	
Motion sensor	5to	0.8				3	3	0	
unknown	25	5.8							
unknown	82n	1.8							
unknown	85	2.1							
unknown	8n	0.3							
unknown	915	1.0							

Table 1. Mouse RGC types and their alignment to previous classifications.

Comparisons to previously defined RGC types

Our 42 RGC types appear to include all 28 types previously identified functionally (referenced in **Table 1** and at rgctypes.org) as well as 14 “novel” types that have not, to our knowledge, been defined previously. Remarkably, most of these types can be distinguished based on their response patterns to spots of varying

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

size. The total is close to previous estimates (Baden et al., 2016; Bae et al., 2018; Rheaume et al., 2018; Tran et al., 2019), supporting the view that mouse RGC classification is approaching completion. Many of the "novel" types had certainly been encountered in previous studies, but we list them as such here based on our belief that they had not been identified separately as distinct functional types (e.g. multiple types had been grouped into "ON transient" and "OFF transient" categories). The novel types include several sets of functionally similar RGCs (ON tr. MeRF / ON tr. SmRF / ON tr. EW6t, OFF tr. MeRF / OFF tr. SmRF, OFF med. sus / OFF sus. EW1no / OFF sus. EW3o), all of which match 1:1 to morphological types and many to transcriptomic types.

Why did the retina evolve entire populations of RGCs that vary only subtly in function? Of many possible answers, we believe the most likely is that functionally similar types would reveal profound differences under stimulus conditions beyond those in our simple battery. A striking example is Eyewire type 25. This type is abundant (5.8% of the population), and forms a convincing and statistically validated dendritic mosaic(Bae et al., 2018), yet we were unable to find its match in thousands of recordings. A natural hypothesis is that this RGC type does not respond to our standard test stimuli, so it was consistently passed over. Supporting this idea, the calcium responses for type '25' in the Eyewire museum are weak (~1% ΔF/F with low signal-to-noise ratio as opposed to some RGC types which reached 20% ΔF/F). Similarly, we failed to find a clear "trigger feature" for several RGC types (e.g. ON bursty, Motion sensor, sSbC EW27). Responses of these cells to flashed spots were inconsistent. For simplicity and reproducibility, our study omitted the vast space of light stimuli that may have differentiated these cell types, including high luminance, variations in color, and complex forms of motion.

Direction selective (DS) and orientation selective (OS) RGCs represent a substantial fraction of the RGCs in the mouse retina (14/42; 33% of types). We identified ON-OFF DS RGCs preferring all four cardinal directions (dorsal, ventral, nasal, temporal), ON DS sustained types preferring three different directions, and one ON DS transient type encountered infrequently and with a wide distribution of preferred directions (**Figure S6**). While there is broad agreement that there are four ON-OFF DS RGC types in the mouse, there is not as strong a consensus about ON DS RGCs. Some studies have reported three types (Estevez et al., 2013) while another reported four (Sabbah et al., 2017). It remains unclear whether this discrepancy is due to one of the ON DS RGC types being transient and the other three being sustained. One study reported a functionally and morphologically distinct ON DS RGC that projects to superior colliculus (SC) and not to the medial terminal nucleus (MTN) or nucleus of the optic tract (NOT) of the accessory optic system(Gauvain and Murphy, 2015). This SC-projecting ON DS RGC had transient responses and more balanced ON and OFF dendritic strata than the MTN-projecting types, consistent with type '7o' in the Eyewire museum. While the previous study on these cells did not report the distribution of their preferred directions (Gauvain and Murphy, 2015), calcium responses for Eyewire type '7o' consistently preferred a nearly nasal direction on the retina (Bae et al., 2018). Our sample of ON DS sustained RGCs had a distribution of preferred directions with three clusters, separated by ~120 degrees, but the sparsely sampled ON DS transient RGCs had inconsistent direction preference (**Figure S6**), and we have so far been unable to reconstruct its morphology. Thus, we have provisionally assigned the ON DS trans. RGC to Eyewire type '7o', but it is the match in which we have the least confidence. A more focused study on ON DS RGCs will be needed to resolve this final issue in the classification of DS RGC types.

OS RGCs, described long ago in other species (Levick, 1967; Maturana and Frenk, 1963), were only recently identified in the mouse (Nath and Schwartz, 2016, 2017). OFF OS RGCs were separated into horizontal- and vertical-preferring types based on their physiology, and the vertical-preferring type tended to have ventrally directed dendrites, while horizontally-preferring cells had a less consistent asymmetry (Nath and Schwartz, 2017). The Eyewire data did not have a corresponding type consisting only of cells with strong ventrally directed dendrites, although they note that type '2aw,' with its similar range in dendritic asymmetry, has a much higher coverage factor than the other types and likely corresponds to at least two RGC types that were not

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

separable based on morphology alone (Bae et al., 2018). Given these facts and the corresponding stratification patterns between these types, we are confident in the categorization of both OFFhOS and OFFvOS RGCs as Eyewire type '2aw'.

ON OS RGCs were also classified into horizontal- and vertical-preferring types when they were reported in mouse (Nath and Schwartz, 2016), but here we further subdivide each group into separate "Small RF" and "Large RF" types based on the spot size to which they respond optimally and their degree of surround suppression. All four ON OS RGC types are among the largest in the retina in terms of dendritic span, so their morphology is captured incompletely in the Eyewire dataset. Nonetheless, we have been able to assign each of these functional OS RGC types to its most likely matching morphological type, in one case combining two Eyewire types ('81i' and '81o').

We identified three RGC types as suppressed-by-contrast (SbC), and a fourth, the ON delayed RGC, has been classified as an SbC RGC under some conditions (Jacoby and Schwartz, 2018; Tien et al., 2015). The RGC type we originally identified as the sustained SbC (Jacoby et al., 2015) has now been split into two types (EW27 and EW28) based on both physiological and morphological criteria. The bursty SbC (bSbC) RGC is distinguished from the sustained SbC types by its much higher baseline firing rate, more transient suppression, and monostratified morphology. Overall, our data underscores the fact that, like the other three polarities (ON, OFF, and ON-OFF), SbC is a response class composed of multiple RGC types (Jacoby and Schwartz, 2018).

Relationships between morphology, function, and gene expression

Having matched functional, morphological, and transcriptomic information for most RGC types, we were able to assess the relationships among these properties. Three surprising trends emerged in comparing function to morphology. First, there are many exceptions to the rule that RGCs with dendrites in the outer half of the IPL have OFF responses. The M1 ipRGC was a well-known exception, because it receives ectopic synapses from ON bipolar cells in the outer IPL (Dumitrescu et al., 2009), but it is far from the only exception to this rule. All four ON OS RGC types, the ON delayed, the M6, and both sSbC types have OFF dendrites but lack OFF spike responses. Additionally, the OFF OS RGCs and the F-mini-ON RGC receive OFF input via gap junctions but lack OFF bipolar cell input under any stimulus condition we have tested (Cooler and Schwartz, 2020; Nath and Schwartz, 2017). An important caveat is that stimuli beyond our test set could reveal OFF responses, perhaps in bright conditions (Pearson and Kerschensteiner, 2015; Tikidji-Hamburyan et al., 2015).

Second, the dendritic area of an RGC has often been associated with the size of its RF center. While this association has a strong basis in the anatomy of the vertical excitatory pathways of the retina, there are a number of exceptions in our data set. For example, "Small RF" and "Large RF" ON OS RGC types do not differ appreciably in dendritic area, and M6 RGCs have smaller RFs than ON delayed RGCs despite substantially larger dendritic area. Differential influences of inhibition and disinhibition likely explain some of these effects (Mani and Schwartz, 2017; Wienbar and Schwartz, 2018).

Finally, RGCs with dendrites near the inner and outer margins of the IPL are typically assumed to have more sustained light responses while those stratifying near the middle of the IPL are assumed to be more transient (Awatramani and Slaughter, 2000; Roska and Werblin, 2001). This association has gained support from large-scale measurements of the kinetics of glutamate release from bipolar cells throughout the IPL (Franke et al., 2017; Marvin et al., 2013). While our data generally fit this trend (**Figure 6B,C**), there were two notable exceptions. The M6 RGC is transient despite stratification at both margins of the IPL, and the LED RGC is sustained despite stratification near the middle of the IPL (Jacoby and Schwartz, 2017).

The literature linking gene expression in particular RGC types to their morphology and function has been more fragmentary because the lack of known matches has prevented a wide view. We found that expression of the gene *Tusc5* is strongly associated with a particular physiological profile (transient light responses and strong surround suppression) and a morphological profile (stratification between the ChAT bands)(**Figure 6B,C**). As

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

more information about each of the RGC types becomes available, including their projection patterns in the brain, we expect more insights into the molecular determinants of RGC wiring patterns both within the retina and to the brain. Future studies may also link biophysical properties of RGCs to the expression of ion channels.

Limitations of the dataset and future directions

Several limitations of our dataset suggest directions for future work. First, our stimuli were limited to a single wavelength distribution, a small range of scotopic to mesopic luminance, and a simple set of artificial patterns (spots, gratings, and moving bars). These stimulus choices meant that we could not explore how RGC responses differed over the parameters of luminance or wavelength. More generally, RGCs evolved not for selectivity to the artificial parameterized stimuli we presented but to detect behaviorally relevant features of natural scenes. Second, while centering the stimulus for each RGC was important for measuring the spatial features of its response, this step complicates the recovery of locally-complete RGC mosaics. Therefore, a future step in RGC typology alignment will be needed to match our types with those in large-scale recordings using either calcium imaging or multi-electrode arrays. Finally, our morphological alignment to the Eyewire dataset was largely qualitative; the limited number of cells in both datasets made quantitative analysis of morphological parameters impractical. With additional cells, a future classification algorithm might make RGC type predictions from morphology alone, and since our functional classification algorithm produces a posterior probability for each class, functional and morphological information could be incorporated seamlessly into a single prediction. Similarly, our improving understanding of the gene expression profiles of each RGC type could enable more accurate composite predictions from the expression of a few key genes plus functional and/or morphological data.

Web-based resource

Standardization in the definitions of RGC types among different research groups is essential to support studies on retinal computation, circuit connectivity, and disease pathology. Additionally, there is rapidly expanding interest in the projection patterns of different RGC types throughout the brain (Dhande et al., 2015; Johnson et al., 2021; Martersteck et al., 2017), which similarly relies on standardized type definitions. For these reasons, we created an open online resource at rgctypes.org where users can search and download full datasets, use our classification algorithm, and contribute their own data to this effort. By unifying the separate functional, morphological, and molecular RGC classification schemas, this resource will allow researchers to connect data across experimental modalities. For example, a set of RGCs labeled by their projection to a certain brain region could be classified by gene expression, and our alignment between transcriptomics and function would provide insights into the functional input to that brain region without additional measurements of light responses. Or the complement of RGC types in a new transgenic mouse line could be measured by confocal microscopy, and our alignment between morphology and function could help generate hypotheses about the functional deficits that might exist if this RGC population were ablated. We expect rgctypes.org to play a central role in the fields of retinal neurobiology and vision science moving forward and, more broadly, to serve as a template for data sharing and collaboration that is applicable to neuronal classification projects throughout the CNS.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Methods

Animals. Wild-type mice (C57/Bl6 - JAX 000664) of either sex were dark-adapted overnight and sacrificed according to standards provided by Northwestern University Center for Comparative Medicine. Retinal tissue was isolated under infrared illumination (900 nm) with the aid of night-vision goggles and IR dissection scope attachments (BE Meyers). Retinal orientation was identified using scleral landmarks (Wei et al., 2010), and preserved using relieving cuts in cardinal directions, with the largest cut at the ventral retina. Retinas were mounted on 12mm poly-D-lysine coated glass affixed to a recording dish with grease, with the ganglion cell layer up. Oxygenation was maintained by superfusing the dish with carbogenated Ames medium (US Biological, A1372-25) warmed to 32˚C.

Visual stimuli. RGC types were identified via cell-attached capacitive spike train responses to light stimuli as previously described (Jacoby and Schwartz, 2017; Jacoby et al., 2015; Mani and Schwartz, 2017; Nath and Schwartz, 2016, 2017). Briefly, stimuli were presented using a custom designed light-projector (DLP LightCrafter; Texas Instruments) at a frame rate of 60 Hz. Spatial stimuli patterns were generated on a 912x1140-pixel digital projector using blue (450nm) LEDs focused on the photoreceptor layer. Neutral density filters (Thorlabs) were used to attenuate the light intensity of stimuli to 200 rhodopsin isomerizations per rod per second (R*/rod/s) from darkness.

The receptive field (RF) centers of individual RGCs were determined by monitoring their relative light responses to horizontal and vertical bars (200 x 40 μm, or 100 x 40 μm in the case of cells with high surround suppression) flashed at 30 μm intervals at 11 locations along each axis. Subsequent stimuli were presented at the RF center. For generic light steps, a spot of 200 μm diameter was presented for 1 s, with cell-attached responses recorded for at least 0.5 s pre-stimulus and 1s post-stimulus. For spots of multiple sizes, spots with diameters from 30-1200 μm (on a logarithmic scale) were presented in pseudorandom order, with similarly timed epochs. Direction preference of direction-selective (DS) RGCs was determined by moving bar stimuli, consisting of a rectangular bar (600 x 200 μm) passing through the receptive field center at 1000 μm/s (ON-OFF DS RGCs) or 500 μm/s (ON DS RGCs). Flashed bar stimuli for testing orientation selectivity were 800 x 50 μm and presented at 12 different orientations (Nath and Schwartz, 2016). Drifting gratings and contrast series were presented from a background luminance of 1000 R*/rod/s following protocols from previous studies (Jacoby et al., 2015; Nath and Schwartz, 2017)

Automated classification. RGC type labels were assigned manually by two of the authors (JG and GWS) according to the decision tree in **Figure S2**. Cells were then randomly assigned to a training set (~80% per type) and testing set (~20% but at least 5 cells per type). The multi-class classification problem was broken down into a series of binary ones using the error-correcting output code (ECOC) scheme, such that a series of classifiers each learns to discern different combinations of RGC types. Each binary learner in the ECOC scheme was trained using Ada-boosted decision trees (Hastie et al., 2009) with initial weights set to enforce a uniform prior probability of each RGC type.

Individual trees were trained by performing elastic net logistic regression on a random subset of firing rates from peristimulus time histogram (PSTH) vs. spot size for feature reduction and choosing the threshold that minimized class uncertainty (Friedman et al., 2010; Schneider et al., 2015). Since not all PSTHs were recorded over the same time and spot size ranges, we imputed missing data using a nearest neighbor approach. Poorly sampled points were penalized in both random selection and regression: for time points the penalty was inversely proportional to their frequency of occurrence across cells (since all PSTHs were binned with the same Δt); for spot sizes we aimed to account for the nonlinearity of responses in the penalty with the following formula:

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$$penalty^{-1}(s) \propto max\ (log(MSE(\cdot))_{+}) - log(MSE(s))_{+},$$

where MSE is the mean across cells of the squared error between the chosen spot size, *s*, and the nearest recorded spot size, and $(\cdot)_{+}$ denotes positive rectification.

To extract probabilities from classification scores, we next used an independent training fold to train an isotonic regression model that transformed each binary learner score into a probability, again enforcing a uniform prior using sample weighting (Zadrozny and Elkan, 2002). The probabilities from each binary learner were then coupled to obtain a probability for each class (Zadrozny, 2002). Thus for our final results the training set was divided into one fold for training the Adaboost models and one fold for training the isotonic regression models. We used three-fold cross validation to train a Bayesian optimization model for hyperparameter tuning. Table 3 lists the hyperparameters we optimized and their final values. The classifier is available for use at rgctypes.org, and the source code is available at https://github.com/zfj1/rgc-classifier.

Parameter	Algorithm level	Optimized value	Optimization range
Number of features	Elastic net	54	5 to 100
Number of folds	Elastic net	6	2 to 10
Alpha	Elastic net	.326	0.0 to 1.0
Number of lambda values	Elastic net	6	5 to 50
Number of repetitions	Decision node	5	5 to 20
Minimum size	Decision node	13	5 to 100
Maximum depth	Decision tree	6	2 to 8
Minimum tree count	Adaboost forest	76	20 to 100
Maximum tree count	Adaboost forest	95	25 to 100
Stopping criterion	Adaboost forest	1.44% improvement over last 44 trees	1% to 50% improvement over last 10 to 50 trees
Ensemble size	ECOC	96	32 to 100
Probability of ensemble membership	ECOC	22.2% in positive class, 38.2% in negative class, 60.4% null	10% to 90% in positive/negative class

Table 2. Hyperparameters for the automated RGC classifier.

Imaging. A subset of recorded RGCs were injected with Neurobiotin (Vector Laboratories, SP-1150, ~3% w/v and ~280 mOsm in potassium aspartate internal solution) using patch pipettes. Retinas were then fixed in 4% paraformaldehyde for 15 minutes at 25˚C, washed three times with PBS, and incubated for 1 hour in blocking solution (1X PBS with 3% normal donkey serum, 0.05% sodium azide, 0.5% Triton X-100) including streptavidin conjugated to a fluorophore (Alexa Fluor-488 or Alexa Fluor-568). Next, retinas were incubated again in blocking solution with primary antibody against choline acetyltransferase (ChAT; Millipore, AB144P, goat anti-ChAT, 1:1000) for 5 nights at 4˚C. Retinas were then rinsed in PBS three times at no less than 1 hour per wash before incubation overnight at 4˚C with streptavidin (Jackson, 016-600-084) and secondary antibody (Donkey anti-Goat 647, Fisher, A11055). Retinas were then rinsed again in PBS three times at no less than 1 hour per wash before mounting on slides with Fluoromount.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

RGCs filled with AlexaFluor were imaged immediately using two-photon microscopy (920 nm, MaiTai HP; SpectraPhysics) under a 60× water-immersion objective (Olympus LUMPLan FLN 60×/1.00 numerical aperture). A 520–540 nm band-pass filter was used to collect emission. After immunohistochemistry, confocal imaging was performed at the Center for Advanced Microscopy at Northwestern University Feinberg School of Medicine generously supported by NCI CCSG P30 CA060553 awarded to the Robert H Lurie Comprehensive Cancer Center. Dendrites were traced in Fiji, and analysis to flatten the retina relative to the ChAT bands was performed as described in (Sümbül et al., 2014).

Single-cell transcriptomics

Library generation. Following physiological recording, a subset of RGCs was isolated for single-cell transcriptome sequencing. First, the area surrounding cells of interest was cleaned of nearby cells and visible debris by aspiration through a large (3-4um inner diameter) patch pipette. Cells were then aspirated using a freshly flame-pulled patch pipette (2.5 inner diameter) and placed into a 5 µl of lysis Buffer TCL (Qiagen, 1031576) + 1% 2-mercaptoethanol (Millipore-Sigma, 63689) before being flash-frozen on dry ice.

We generated RNA-Seq libraries using a modified Smart-seq2 method (Picelli et al., 2014) with the following minor changes: Before reverse transcription, RNA was purified using 2.2X SPRI-beads (Beckman Coulter, A3987) followed by 3 wash steps with 80% EtOH, elution in 4 µl of RT primer mix and denatured at 72 °C for 3 min. Six µl of the first-strand reaction mix, containing 0.1 µl SuperScript II reverse transcriptase (200 U/µl, Invitrogen), 0.25 µl RNAse inhibitor (40 U/µl, Clontech), 2 µl Superscript II First-Strand Buffer (5x, Invitrogen), 0.1 µl $MgCl_2$ (100 mM, Sigma), 0.1 µl TSO (100 µM) and 3.45 µl Trehalose (1M), were added to each sample. Reverse transcription was carried out at 50°C for 90 min followed by inactivation at 85 °C for 5 min. After PCR preamplification, product was purified using a 0.8X of AMPure XP beads (Beckman Coulter), with the final elution in 12 µl of EB solution (Qiagen). For tagmentation the Nextera DNA Sample Preparation kit (FC-131-1096, Illumina) was used and final PCR was performed as follows: 72 °C 3 min, 95 °C 30 s, then 12 cycles of (95 °C 10 s, 55 °C 30 s, 72 °C 1 min), 72°C 5min. Purification was done with a 0.9X of AMPure XP beads. Libraries were diluted to a final concentration of 2 nM, pooled and sequenced on Next-Seq(Mid), 75bp paired end.

Alignment and quantification of scRNA-cell transcriptomic libraries. Gene expression levels were quantified using RNA-seq by Expectation Maximization (RSEM) (Li and Dewey, 2011). Under the hood, Bowtie 2 (Langmead and Salzberg, 2012) was used to map paired-end reads to a mouse transcriptome index (mm10/GRCm38 UCSC build). RSeQC (Wang et al., Bioinformatics, 2012) was used to quantify quality metrics for the alignment results. We only considered cells where the read alignment rate to the genome and transcriptome exceeded 85% and 35% respectively, and the total number of transcriptome-mapped reads was less than 350,000. RSEM yielded an expression matrix (genes x samples) of transcript per million counts (TPM), which were log-transformed after the addition of 1 to avoid zeros. Overall 91 RGCs, each of which carried a functional type label, were selected for further analysis.

Matching gene-expression clusters to cell types. To map each of the 91 RGC transcriptomes to a molecular cluster in Tran et al., 2019 we used the XGboost algorithm (Chen and Guestrin, 2016), as implemented in the R package xgboost. Briefly, we trained and validated an xgboost multi-class classifier on the atlas of 35,699 RGCs subdivided into 45 molecularly distinct groups (C1-C45). Around 50% of the data was used for training and the remaining 50% was held out and used for validation. We optimized hyperparameters (e.g. tree depth, number of features, class-specific weights) to achieve a validation set accuracy of >90% across each of the 45 transcriptomic classes. This trained classifier was then used to assign a cluster label for each of the 105 transcriptomes profiled in this study. We assigned a transcriptomic label to each RGC if a minimum of 15% of trees in the forest voted on the majority decision. This choice of voting margin was >6x higher than the random

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

threshold of 2.3%, based on the fact that there are 45 classes. The correspondences between functional and transcriptomic labels were visualized as confusion matrices.

UMAP, agglomerative hierarchical clustering, and cross-modality neighborhood comparisons. The functional input data to the UMAP algorithm was a linearized version of the full matrix of the PSTH for each cell across spot sizes (as in **Figure 2**). We used a MATLAB implementation of UMAP (https://www.mathworks.com/matlabcentral/fileexchange/71902) supervised by the RGC type labels for the data set of 1777 cells. Agglomerative hierarchical clustering of RGC type centroids in UMAP space (**Figure 4B, C**) used the MATLAB 'linkage' function with the median (weighted center of mass) method. The input to the UMAP algorithm for morphology was the unnormalized stratification profile for each RGC from the Eyewire museum (381 cells) supervised by the labels in the museum. Although no attempt was made to capture details of the *en face* morphological characteristics of each cell, the unnormalized stratification data allowed the algorithm to use information about total dendritic length. The input to the UMAP algorithm for transcriptomic space was a vector of gene expression values for RGC-type-selective genes from the published dataset (~35,699 cells) as described in (Tran et al., 2019)**.**

We measured similarities between the three UMAP spaces (function, morphology, and genetics) by comparing nearest neighbors between spaces (Supplemental Fig. 5). For each RGC type in which we established a match between the two spades being compared, we measured the fractional overlap between the neighest neighbors in the first space and those in the second space (matching types / neighborhood size). The analysis was repeated for neighborhood sizes from 2 - 12. To assess the statistics of the measured overlap values, we created a bootstrap distribution by randomly shuffling the cluster identities in one of the spaces. Data in **Figure 7D–F** are z scores with respect to this bootstrap distribution which was Gaussian.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

Awatramani, G.B., and Slaughter, M.M. (2000). Origin of transient and sustained responses in ganglion cells of the retina. J. Neurosci. *20*, 7087–7095.

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.e19.

Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I.W.H., Ng, L.G., Ginhoux, F., and Newell, E.W. (2018). Dimensionality reduction for visualizing single-cell data using UMAP. Nat. Biotechnol.

Berson, D.M., Castrucci, A.M., and Provencio, I. (2010). Morphology and mosaics of melanopsin-expressing retinal ganglion cell types in mice. J. Comp. Neurol. *518*, 2405–2422.

Bleckert, A., Schwartz, G.W., Turner, M.H., Rieke, F., and Wong, R.O.L. (2014). Visual space is represented by nonmatching topographies of distinct mouse retinal ganglion cell types. Curr. Biol. *24*, 310–315.

Cadwell, C.R., Palasantza, A., Jiang, X., Berens, P., Deng, Q., Yilmaz, M., Reimer, J., Shen, S., Bethge, M., Tolias, K.F., et al. (2016). Electrophysiological, transcriptomic and morphologic profiling of single neurons using Patch-seq. Nat. Biotechnol. *34*, 199–203.

Chen, T., and Guestrin, C. (2016). XGBoost. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

Close, J.L., Long, B.R., and Zeng, H. (2021). Spatially resolved transcriptomics in neuroscience. Nat. Methods *18*, 23–25.

Cooler, S., and Schwartz, G.W. (2020). An offset ON–OFF receptive field is created by gap junctions between distinct types of retinal ganglion cells. Nat. Neurosci. 1–11.

Dhande, O.S., Stafford, B.K., Lim, J.-H.A., and Huberman, A.D. (2015). Contributions of Retinal Ganglion Cells to Subcortical Visual Processing and Behaviors. Annu Rev Vis Sci *1*, 291–328.

Do, M.T.H., Kang, S.H., Xue, T., Zhong, H., Liao, H.-W., Bergles, D.E., and Yau, K.-W. (2009). Photon capture and signalling by melanopsin retinal ganglion cells. Nature *457*, 281–287.

Dumitrescu, O.N., Pucci, F.G., Wong, K.Y., and Berson, D.M. (2009). Ectopic retinal ON bipolar cell synapses in the OFF inner plexiform layer: Contacts with dopaminergic amacrine cells and melanopsin ganglion cells. J. Comp. Neurol. *517*, 226–244.

Estevez, M., Quattrochi, L., Dhande, O., Kim, I., Firman, T., Eldanaf, R., Huberman, A., and Berson, D. (2013). Form and function of the three ON-type direction-selective retinal ganglion cells in the Hoxd10 mouse. Invest. Ophthalmol. Vis. Sci. *54*, 1298–1298.

Euler, T., Franke, K., and Baden, T. (2019). Studying a Light Sensor with Light: Multiphoton Imaging in the Retina. In Multiphoton Microscopy, E. Hartveit, ed. (New York, NY: Springer New York), pp. 225–250.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Farrow, K., and Masland, R.H. (2011). Physiological clustering of visual channels in the mouse retina. J. Neurophysiol. *105*, 1516–1530.

Franke, K., Berens, P., Schubert, T., Bethge, M., Euler, T., and Baden, T. (2017). Inhibition decorrelates visual feature representations in the inner retina. Nature *542*, 439–444.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. J. Stat. Softw. *33*, 1–22.

Gauvain, G., and Murphy, G.J. (2015). Projection-Specific Characteristics of Retinal Input to the Brain. Journal of Neuroscience *35*, 6575–6583.

Grimes, W.N., Schwartz, G.W., and Rieke, F. (2014). The synaptic and circuit mechanisms underlying a change in spatial encoding in the retina. Neuron *82*, 460–473.

Hastie, T., Rosset, S., Zhu, J., and Zou, H. (2009). Multi-class adaboost. Stat. Interface *2*, 349–360.

Hodge, R.D., Bakken, T.E., Miller, J.A., Smith, K.A., Barkan, E.R., Graybuck, L.T., Close, J.L., Long, B., Johansen, N., Penn, O., et al. (2019). Conserved cell types with divergent features in human versus mouse cortex. Nature *573*, 61–68.

Hoon, M., Okawa, H., Della Santina, L., and Wong, R.O.L. (2014). Functional architecture of the retina: development and disease. Prog. Retin. Eye Res. *42*, 44–84.

Jacoby, J., and Schwartz, G.W. (2017). Three Small-Receptive-Field Ganglion Cells in the Mouse Retina Are Distinctly Tuned to Size, Speed, and Object Motion. Journal of Neuroscience *37*, 610–625.

Jacoby, J., and Schwartz, G.W. (2018). Typology and Circuitry of Suppressed-by-Contrast Retinal Ganglion Cells. Front. Cell. Neurosci. *12*, 269.

Jacoby, J., Zhu, Y., DeVries, S.H., and Schwartz, G.W. (2015). An Amacrine Cell Circuit for Signaling Steady Illumination in the Retina. Cell Rep. *13*, 2663–2670.

Joesch, M., and Meister, M. (2016). A neuronal circuit for colour vision based on rod–cone opponency. Nature *532*, 236–239.

Johnson, K.P., Zhao, L., and Kerschensteiner, D. (2018). A Pixel-Encoder Retinal Ganglion Cell with Spatially Offset Excitatory and Inhibitory Receptive Fields. Cell Rep. *22*, 1462–1472.

Johnson, K.P., Fitzpatrick, M.J., Zhao, L., Wang, B., McCracken, S., Williams, P.R., and Kerschensteiner, D. (2021). Cell-type-specific binocular vision guides predation in mice. Neuron *109*, 1527–1539.e4.

Jouty, J., Hilgen, G., Sernagor, E., and Hennig, M.H. (2018). Non-parametric Physiological Classification of Retinal Ganglion Cells in the Mouse Retina. Front. Cell. Neurosci. *12*, 481.

Kay, J.N., De la Huerta, I., Kim, I.-J., Zhang, Y., Yamagata, M., Chu, M.W., Meister, M., and Sanes, J.R. (2011). Retinal ganglion cells with distinct directional preferences differ in molecular identity, structure, and central projections. J. Neurosci. *31*, 7753–7762.

Kay, J.N., Chu, M.W., and Sanes, J.R. (2012). MEGF10 and MEGF11 mediate homotypic interactions required

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

for mosaic spacing of retinal neurons. Nature *483*, 465–469.

Kim, I.-J., Zhang, Y., Yamagata, M., Meister, M., and Sanes, J.R. (2008). Molecular identification of a retinal cell type that responds to upward motion. Nature *452*, 478–482.

Krieger, B., Qiao, M., Rousso, D.L., Sanes, J.R., and Meister, M. (2017). Four alpha ganglion cell types in mouse retina: Function, structure, and molecular signatures. PLoS One *12*, e0180091.

Laboulaye, M.A., Duan, X., Qiao, M., Whitney, I.E., and Sanes, J.R. (2018). Mapping Transgene Insertion Sites Reveals Complex Interactions Between Mouse Transgenes and Neighboring Endogenous Genes. Front. Mol. Neurosci. *11*, 385.

Levick, W.R. (1967). Receptive fields and trigger features of ganglion cells in the visual streak of the rabbit's retina. J. Physiol. *188*, 285–307.

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.

MacKay, D.J.C., and Mac, D.J. (2003). Information Theory, Inference and Learning Algorithms (Cambridge University Press).

Mani, A., and Schwartz, G.W. (2017). Circuit Mechanisms of a Retinal Ganglion Cell with Stimulus-Dependent Response Latency and Activation Beyond Its Dendrites. Curr. Biol. *27*, 471–482.

Marco, S.D., Di Marco, S., Protti, D.A., and Solomon, S.G. (2013). Excitatory and inhibitory contributions to receptive fields of alpha-like retinal ganglion cells in mouse. Journal of Neurophysiology *110*, 1426–1440.

Martersteck, E.M., Hirokawa, K.E., Evarts, M., Bernard, A., Duan, X., Li, Y., Ng, L., Oh, S.W., Ouellette, B., Royall, J.J., et al. (2017). Diverse Central Projection Patterns of Retinal Ganglion Cells. Cell Rep. *18*, 2058–2072.

Marvin, J.S., Looger, L.L., and Demb, J.B. (2013). Two-photon imaging of nonlinear glutamate release dynamics at bipolar cell synapses in the mouse retina. Journal of.

Maturana, H.R., and Frenk, S. (1963). DIRECTIONAL MOVEMENT AND HORIZONTAL EDGE DETECTORS IN THE PIGEON RETINA. Science *142*, 977–979.

McInnes, L., Healy, J., and Melville, J. (2018). UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction.

Nadal-Nicolás, F.M., Kunze, V.P., Ball, J.M., Peng, B.T., Krishnan, A., Zhou, G., Dong, L., and Li, W. (2020). True S-cones are concentrated in the ventral mouse retina and wired for color detection in the upper visual field. Elife *9*.

Nath, A., and Schwartz, G.W. (2016). Cardinal Orientation Selectivity Is Represented by Two Distinct Ganglion Cell Types in Mouse Retina. Journal of Neuroscience *36*, 3208–3221.

Nath, A., and Schwartz, G.W. (2017). Electrical synapses convey orientation selectivity in the mouse retina. Nat. Commun. *8*, 2025.

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Pearson, J.T., and Kerschensteiner, D. (2015). Ambient illumination switches contrast preference of specific retinal processing streams. J. Neurophysiol. *114*, 540–550.

Quattrochi, L.E., Stabio, M.E., Kim, I., Ilardi, M.C., Michelle Fogerson, P., Leyrer, M.L., and Berson, D.M. (2019). The M6 cell: A small-field bistratified photosensitive retinal ganglion cell. J. Comp. Neurol. *527*, 297–311.

Reese, B.E., and Keeley, P.W. (2015). Design principles and developmental mechanisms underlying retinal mosaics. Biol. Rev. Camb. Philos. Soc. *90*, 854–876.

Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., et al. (2017). The Human Cell Atlas. Elife *6*.

Rheaume, 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.

Rockhill, R.L., Euler, T., and Masland, R.H. (2000). Spatial order within but not between types of retinal neurons. Proc. Natl. Acad. Sci. U. S. A. *97*, 2303–2307.

Roska, B., and Werblin, F. (2001). Vertical interactions across ten parallel, stacked representations in the mammalian retina. Nature *410*, 583–587.

Rousso, 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.

Sabbah, S., Gemmer, J.A., Bhatia-Lin, A., Manoff, G., Castro, G., Siegel, J.K., Jeffery, N., and Berson, D.M. (2017). A retinal code for motion along the gravitational and body axes. Nature *546*, 492–497.

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.

Scala, F., Kobak, D., Bernabucci, M., Bernaerts, Y., Cadwell, C.R., Castro, J.R., Hartmanis, L., Jiang, X., Laturnus, S., Miranda, E., et al. (2020). Phenotypic variation of transcriptomic cell types in mouse motor cortex. Nature.

Schmidt, T.M., and Kofuji, P. (2009). Functional and morphological differences among intrinsically photosensitive retinal ganglion cells. J. Neurosci. *29*, 476–482.

Schneider, M., Hirsch, S., Weber, B., Székely, G., and Menze, B.H. (2015). Joint 3-D vessel segmentation and centerline extraction using oblique Hough forests with steerable filters. Med. Image Anal. *19*, 220–249.

Sümbül, U., Song, S., McCulloch, K., Becker, M., Lin, B., Sanes, J.R., Masland, R.H., and Sebastian Seung, H. (2014). A genetic and computational approach to structurally classify neuronal types. Nature Communications *5*.

Surgucheva, I., Weisman, A.D., Goldberg, J.L., Shnyra, A., and Surguchov, A. (2008). Gamma-synuclein as a marker of retinal ganglion cells. Mol. Vis. *14*, 1540–1548.

Tien, N.-W., Pearson, J.T., Heller, C.R., Demas, J., and Kerschensteiner, D. (2015). Genetically Identified

bioRxiv preprint doi: https://doi.org/10.1101/2021.06.10.447922; this version posted June 11, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Suppressed-by-Contrast Retinal Ganglion Cells Reliably Signal Self-Generated Visual Stimuli. J. Neurosci. *35*, 10815–10820.

Tikidji-Hamburyan, A., Reinhard, K., Seitter, H., Hovhannisyan, A., Procyk, C.A., Allen, A.E., Schenk, M., Lucas, R.J., and Münch, T.A. (2015). Retinal output changes qualitatively with every change in ambient illuminance. Nat. Neurosci. *18*, 66–74.

Tran, N.M., Shekhar, K., Whitney, I.E., Jacobi, A., Benhar, I., Hong, G., Yan, W., Adiconis, X., Arnold, M.E., Lee, J.M., et al. (2019). Single-Cell Profiles of Retinal Ganglion Cells Differing in Resilience to Injury Reveal Neuroprotective Genes. Neuron *104*, 1039–1055.e12.

Trenholm, S., Johnson, K., Li, X., Smith, R.G., and Awatramani, G.B. (2011). Parallel mechanisms encode direction in the retina. Neuron *71*, 683–694.

Warwick, R.A., Kaushansky, N., Sarid, N., Golan, A., and Rivlin-Etzion, M. (2018). Inhomogeneous Encoding of the Visual Field in the Mouse Retina. Curr. Biol.

Wässle, H., Peichl, L., and Boycott, B.B. (1981). Dendritic territories of cat retinal ganglion cells. Nature *292*, 344–345.

Wei, W., Elstrott, J., and Feller, M.B. (2010). Two-photon targeted recording of GFP-expressing neurons for light responses and live-cell imaging in the mouse retina. Nat. Protoc. *5*, 1347–1352.

Wienbar, S., and Schwartz, G.W. (2018). The dynamic receptive fields of retinal ganglion cells. Progress in Retinal and Eye Research *67*, 102–117.

Wilbrey-Clark, A., Roberts, K., and Teichmann, S.A. (2020). Cell Atlas technologies and insights into tissue architecture. Biochem. J *477*, 1427–1442.

Yuste, R., Hawrylycz, M., Aalling, N., Aguilar-Valles, A., Arendt, D., Armañanzas, R., Ascoli, G.A., Bielza, C., Bokharaie, V., Bergmann, T.B., et al. (2020). A community-based transcriptomics classification and nomenclature of neocortical cell types. Nat. Neurosci. *23*, 1456–1468.

Zadrozny, B. (2002). Reducing multiclass to binary by coupling probability estimates. In Advances in Neural Information Processing Systems 14, T.G. Dietterich, S. Becker, and Z. Ghahramani, eds. (MIT Press), pp. 1041–1048.

Zadrozny, B., and Elkan, C. (2002). Transforming classifier scores into accurate multiclass probability estimates. In Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, (New York, NY, USA: Association for Computing Machinery), pp. 694–699.

Zeng, H., and Sanes, J.R. (2017). Neuronal cell-type classification: challenges, opportunities and the path forward. Nat. Rev. Neurosci. *18*, 530–546.