A chromatic feature detector in the retina signals visual context changes

Larissa Höfling[1,2], Klaudia P. Szatko[1,2,8], Christian Behrens[1], Yongrong Qiu[1,2], David A. Klindt[7], Zachary Jessen[4], Gregory W. Schwartz[4], Matthias Bethge[2,3], Philipp Berens[1,2,3], Katrin Franke[1,9], Alexander S. Ecker[5,6], and Thomas Euler[1,2,✉]

[1]Institute for Ophthalmic Research, University of Tübingen, Tübingen, Germany
[2]Centre for Integrative Neuroscience, University of Tübingen, Tübingen, Germany
[3]Tübingen AI Center, University of Tübingen, Tübingen, Germany
[4]Feinberg School of Medicine, Department of Ophthalmology, Northwestern University, Chicago, IL, USA
[5]Institute of Computer Science and Campus Institute Data Science, University of Göttingen, Göttingen, Germany
[6]Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany
[7]Norwegian University of Science and Technology, Trondheim, Norway
[8]Current affiliation: National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA
[9]Current affiliation: Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA

The retina transforms patterns of light into visual feature representations supporting behaviour. These representations are distributed across various types of retinal ganglion cells (RGCs), whose spatial and temporal tuning properties have been extensively studied in many model organisms, including the mouse. However, it has been difficult to link the potentially nonlinear retinal transformations of natural visual inputs to specific ethological purposes. Here, we discover a novel selectivity to chromatic contrast in an RGC type that allows the detection of transitions of the horizon across a retinal region. We trained a convolutional neural network (CNN) model on large-scale functional recordings of RGC responses to natural mouse movies, and then used this model to search *in silico* for stimuli that maximally excite distinct types of RGCs. This procedure predicted centre colour-opponency in transient Suppressed-by-Contrast RGCs (tSbC), a cell type whose function is being debated. We confirmed experimentally that these cells indeed responded very selectively to Green-OFF, UV-ON contrasts, which we found to be characteristic of transitions from ground to sky in the visual scene, as might be elicited by head- or eye-movements across the horizon. Because tSbCs reliably detected these transitions, we suggest a role for this RGC type in providing contextual information (i.e. sky or ground) necessary for the selection of appropriate behavioural responses to other stimuli, such as looming objects. Our work showcases how a combination of experiments with natural stimuli and computational modelling allows discovering novel types of stimulus selectivity and identifying their potential ethological relevance.

retina | colour vision | computational modelling

Correspondence:
thomas.euler@cin.uni-tuebingen.de

Introduction

Sensory systems evolved to generate representations of an animal's natural environment useful for survival and procreation (1). These environments are complex and high-dimensional, and different features are relevant for different species (reviewed in (2)). As a consequence, the representations are adapted to an animal's needs: features of the world relevant for the animal are represented with enhanced precision, whereas less important features are discarded. Sensory processing is thus best understood within the context of the environment an animal evolved in and that it interacts with (reviewed in (3, 4)).

The visual system is well-suited for studying sensory processing, as the first features are already extracted at its front-end, the retina (reviewed in (2, 5)). In the mouse, this experimentally well-accessible tissue gives rise to more than 30 parallel feature channels (6–9), represented by different types of retinal ganglion cells (RGCs), whose axons send information to numerous visual centres in the brain (10). Some of these channels encode basic features, such as luminance changes and motion, that are only combined in downstream areas to support a range of behaviours such as cricket hunting in mice (11). Other channels directly extract specific features from natural scenes necessary for specific behaviours. For instance, transient OFF-α cells trigger freezing or escape behaviour in response to looming stimuli (12–14).

For many RGC types, however, we lack understanding of the features they encode and how these link to behaviour (15). One reason for this is that the synthetic stimuli commonly used to study retinal processing fail to drive retinal circuits "properly" and, hence, cannot uncover critical response properties triggered in natural environments. Colour, for example, is a salient feature in nature, and the mouse visual system dedicates intricate circuitry to the processing of chromatic information (16–21). Studies using synthetic stimuli have revealed nonlinear and centre-surround interactions between colour channels, but it is not clear how these are engaged in retinal processing of natural environments.

Indeed, stimuli capturing the statistics of natural environments have revealed a larger complexity in retinal spatial nonlinearities than had been previously described based on simpler synthetic stimuli (23). Such nonlinearities, crucial for the encoding of natural stimuli, cannot be captured by Linear-Nonlinear (LN) models of retinal processing, and several improvements over LN models have been proposed for the identification of receptive fields (RF) (reviewed in (24)). In recent years, however, convolutional neural network (CNN) models have become the state-of-the-art approach for predictive modelling of visual processing, both in the retina (25–27), as well as in higher visual areas (28–31). In the cortex, two recent studies took the CNN modelling approach

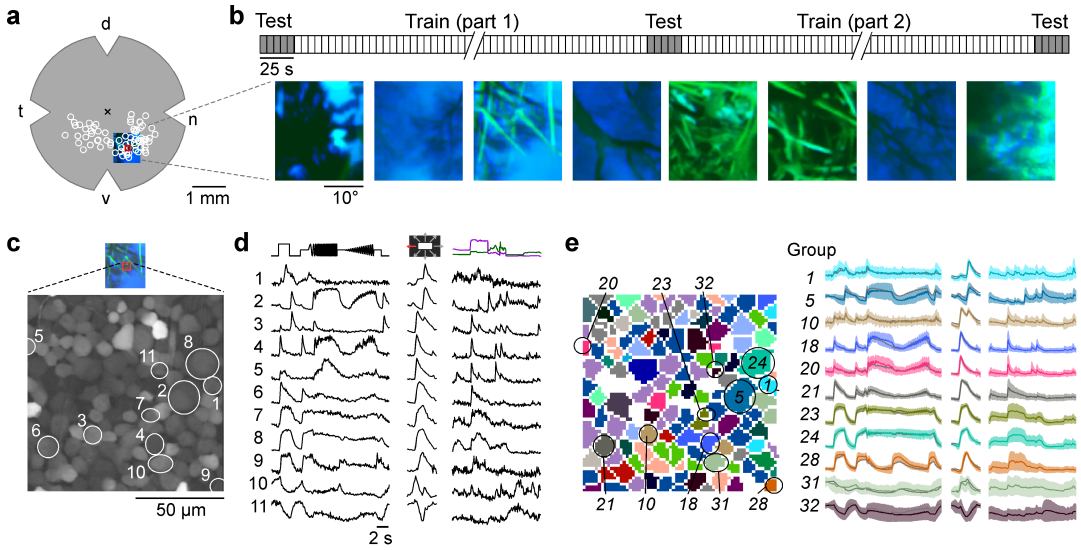

Fig. 1. Mouse RGCs display diverse responses to a natural movie stimulus (a) Illustration of a flat-mounted retina, with recording fields (white circles) and stimulus area centred on the red recording field indicated (cross marks optic disc; d, dorsal; v, ventral; t, temporal; n, nasal). **(b)** Natural movie stimulus structure (top) and example frames (bottom). The stimulus consisted of 5-s clips taken from UV-green footage recorded outside (22), with 3 repeats of a 5-clip test sequence (highlighted in grey) and a 108-clip training sequence (see Methods). **(c)** Representative recording field (bottom; marked by red square in (a)) showing somata of ganglion cell layer (GCL) cells loaded with Ca^{2+} indicator OGB-1. **(d)** Ca^{2+} responses of exemplary RGCs (indicated by circles in (c)) to chirp (left), moving bar (centre), and natural movie (right) stimulus. **(e)** Same recording field as in (c) but with cells colour-coded by functional RGC group (left; see Methods and (7)) and group responses (coloured, mean ± SD across cells; trace of example cells in (d) overlaid in black).

further, beyond response prediction, by probing the networks for stimuli that would maximally excite the modelled neurons (32, 33). The resulting *maximally exciting inputs* (MEIs) represent a nonlinear generalisation of linear RFs and were more complex and diverse than expected based on previous results obtained with synthetic stimuli and linear methods. Leveraging the power of this approach, another study highlighted the ethological relevance of colour by uncovering a state-dependent shift in chromatic preference of V1 neurons, a shift that could facilitate the detection of aerial predators against a UV-bright sky (34).

Here, we combined the power of CNN-based modelling with large-scale recordings from RGCs to investigate colour processing in the mouse retina under natural stimulus conditions. Since mouse photoreceptors are sensitive to green and UV light (35), we recorded RGC responses to stimuli capturing the chromatic composition of natural mouse environments in these two chromatic channels. A model-guided search for MEIs in chromatic stimulus space predicted a novel type of chromatic tuning in transient Suppressed-by-Contrast (tSbC) RGCs, a type whose function is being debated vigorously (36–38).

A detailed *in-silico* characterisation followed up by experimental validation *ex-vivo* confirmed this cell type's pronounced and unique selectivity for full-field changes from green-dominated to UV-dominated scenes, a type of visual input that matches the scene statistics of horizon transitions (22, 39, 40). We therefore suggest a role for tSbC RGCs in detecting horizon transitions, a feature that may support important functions, such as indicating context switches in the visual field (ground vs. sky). These findings showcase the utility of our approach for generating hypotheses about the ethological relevance of sensory representations.

Results

Here, we investigated colour processing in the mouse retina under natural stimulus conditions. To this end, we trained a CNN model on RGC responses to a movie covering both achromatic and chromatic contrasts occurring naturally in the mouse environment, and then performed a model-guided search for stimuli that maximise the responses of RGCs.

Mouse RGCs display diverse responses to a natural movie stimulus. Using two-photon population Ca^{2+} imaging, we recorded responses from 8,388 cells (in 72 recording fields across 32 retinae) in the ganglion cell layer (GCL) of the isolated mouse retina (Fig. 1a) to a range of visual stimuli. Since complex interactions between colour channels have been mostly reported in the ventral retina and opsin-transitional zone, we focused our recordings on these regions (19, 20). The stimuli included two achromatic synthetic stimuli – a contrast and frequency modulation ("chirp" stimulus) and a bright-on-dark bar moving in eight directions ("moving bar", MB) – to identify the functional cell type (see below), as well as a dichromatic natural movie (Fig. 1b,c,d). The latter was composed of footage recorded outside in the field using a camera that captured the spectral bands (UV and green; (22)) to which mouse photoreceptors are sensitive ($\lambda^S_{peak} = 360, \lambda^M_{peak} = 510$ nm for S- and M-cones, respectively (35)). We used 113 different movie clips, each lasting 5 s, that were displayed in pseudo-random order. Five of these constituted the test set and were repeated three times: at the beginning, in the middle and at the end of the movie presentation, thereby allowing to assess the reliability of neuronal responses across the recording (Fig. 1b, top).

The responses elicited by the synthetic stimuli and the natural movie were diverse, displaying ON (Fig. 1d, rows 4-

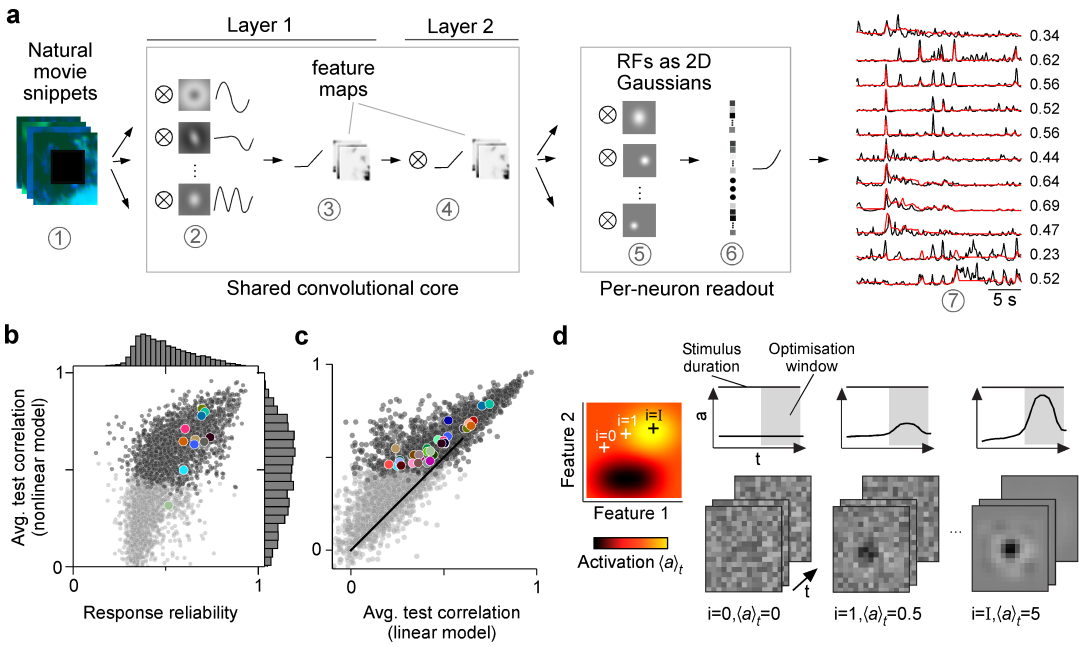

Fig. 2. CNN model captures diverse tuning of RGC groups and predicts MEIs (a) Illustration of the CNN model and its output. The model takes natural movie clips as input (1), performs convolutions with 3D space-time separable filters (2) followed by a nonlinear activation function (3) in two consecutive layers (2-4) within its core, and feeds the output of its core into a per-neuron readout. For each RGC, the readout convolves the feature maps with a learned RF modelled as a 2D Gaussian (5), and finally feeds a weighted sum of the resulting vector through a softplus nonlinearity (6) to yield the firing rate prediction for that RGC (7). Numbers indicate averaged single-trial test set correlation between predicted (red) and recorded (black) responses. **(b)** Test set correlation between model prediction and neural response (averaged across three repetitions) as a function of response reliability (see Methods) for N=3,527 RGCs. Coloured dots correspond to example cells shown in Fig. 1c-e. Dots in darker grey correspond to the N=1,947 RGCs that passed the model test correlation and movie response quality criterion (see Methods and Suppl. Fig. S1c). **(c)** Test set correlation (as in (b)) of model vs. test set correlation of a linearised version of the model (for details, see Methods). Coloured dots correspond to RGC groups 1-32 (7). Dark and light grey dots as in (b). **(d)** Illustration of model-guided search for maximally exciting inputs (MEIs). The trained model captures neural tuning to stimulus features (far left). Starting from a randomly initialised input (2nd from left; a 3D tensor in space and time; only one colour channel illustrated here), the model follows the gradient along the tuning surface (far left) to iteratively update the input until it arrives at the stimulus (bottom right) that maximises the model neuron's activation within an optimisation time window (0.66 s, grey box, top right).

9), ON-OFF (row 3) and OFF (rows 1 and 2), as well as sustained and transient characteristics (e.g. rows 8 and 4, respectively). Some responses were suppressed by temporal contrast (generally, rows 10, 11; at high contrast and frequency, row 9). A total of 6,984 GCL cells passed our response quality criteria (Methods); 3,527 cells could be assigned to one of 32 previously characterised functional RGC groups (7) based on their responses to the chirp and moving bar stimuli using our recently developed classifier (Fig. 1e; Suppl. Fig. S1a) (41). Cells assigned to any of groups 33-46 were considered displaced amacrine cells and were not analysed in this study (for detailed filtering pipeline, see Suppl. Fig. S1c).

CNN model captures diverse tuning of RGC groups and predicts MEIs. We trained a CNN model on the RGCs' movie responses (Fig. 2a) and evaluated model performance as the correlation between predicted and trial-averaged measured test responses, $C(\hat{r}^{(n)}, \langle r^{(n)} \rangle_i)$ (Fig. 2b). This metric can be interpreted as an estimate of the achieved fraction of the maximally achievable correlation (see Methods). The mean correlation per RGC group ranged from 0.32 (G_{14}) to 0.79 (G_{24}) (Suppl. Fig. S1b) and achieved an average of 0.48 (for all N=3,527 cells passing filtering steps 1-3, Suppl. Fig. S1c). This established that the model generally predicted RGC responses to the natural movie stimulus very well. We also tested the performance of our nonlinear model against a linearised version (Methods) and found that the nonlinear CNN model achieved a higher test set correlation for all RGC groups (average correlation linear model: 0.38; G_{14}: 0.2, G_{24}: 0.65, Fig. 2c).

To identify stimuli that optimally drive the potentially nonlinear RGCs, we next synthesised maximally exciting inputs (MEIs; (32)). We used gradient ascent on a randomly initialised, contrast- and range-constrained input to find the stimulus that maximised the mean activation of a given model neuron within a short time window (0.66 s; Methods; Fig. 2d). The fixed contrast budget was shared across the two colour channels.

CNN model predicts centre colour-opponency in RGC group G_{28}. The resulting MEIs were short, dichromatic movie clips; their spatial, temporal, and chromatic properties and interactions thereof are best appreciated in lower-dimensional visualisations (Fig. 3a–c; more example MEIs in Suppl. Fig. S2).

To analyse the MEIs in terms of these properties, we decomposed them into their spatial and temporal components, separately for green and UV, and parameterised the spatial component as a Difference-of-Gaussians (DoG) (40) (N=1,613 out of 1,947, see Methods). We then located MEIs along the axes in stimulus space corresponding to three properties: centre size, mean temporal frequency, and centre contrast, separately for green and UV (Fig. 3d-f). These MEI properties reflect RGC response properties classically probed

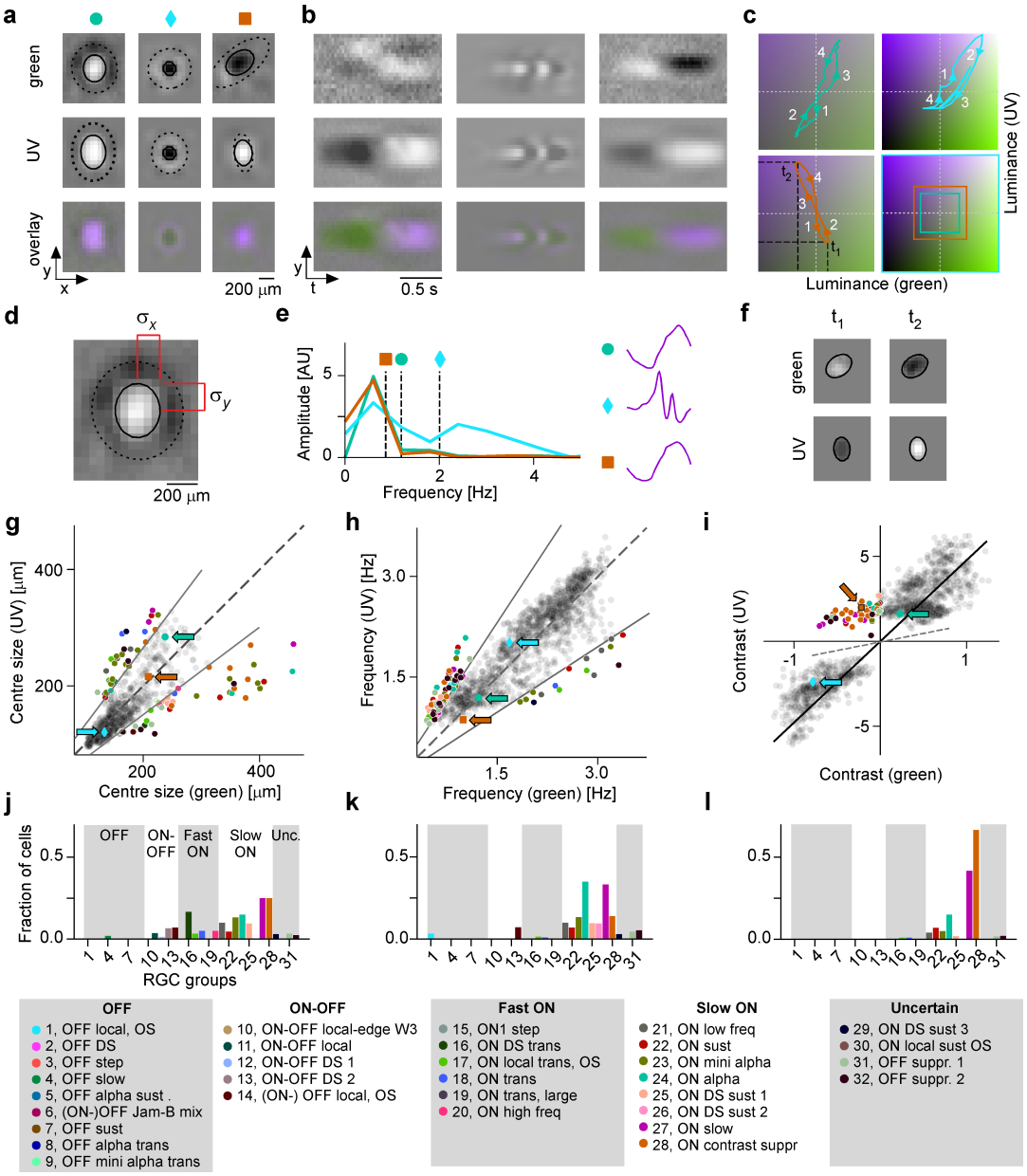

Fig. 3. Spatial, temporal and chromatic properties of MEIs differ between RGC groups (a) Spatial component of three example MEIs for green (top), UV (middle) and overlay (bottom). Solid and dashed circles indicate MEI centre and surround fit, respectively. For display, spatial components s in the two channels were re-scaled to a similar range and displayed on a common grey-scale map ranging from black for $-max(|s|)$ to white for $max(|s|)$, i.e. symmetric about 0 (grey). **(b)** Spatio-temporal (y-t) plot for the three example MEIs (from (a)) at a central vertical slice for green (top), UV (middle) and overlay (bottom). Grey-scale map analogous to (a). **(c)** Trajectories through colour space over time for the centre of the three MEIs. Trajectories start at the origin (grey level); direction of progress indicated by arrow heads. Bottom right: Bounding boxes of the respective trajectory plots. **(d)** Calculation of MEI centre size, defined as $\sigma_x + \sigma_y$, with σ_x and σ_y the s.d. in horizontal and vertical direction, respectively, of the DoG fit to the MEI. **(e)** Calculation of MEI temporal frequency: Temporal components are transformed using Fast Fourier Transform, and MEI frequency is defined as the amplitude-weighted average frequency of the Fourier-transformed temporal component. **(f)** Calculation of centre contrast, which is defined as the difference in intensity at the last two peaks (indicated by t_1 and t_2, respectively, in (c)). For the example cell (orange markers and lines), green intensity decreases, resulting in OFF contrast, and UV intensity increases, resulting in ON contrast. **(g)** Distribution of green and UV MEI centre sizes across N=1,613 cells (example MEIs from (a-c) indicated by arrows; symbols as shown on top of (a)). 95% of MEIs were within an angle of $\pm 8°$ of the diagonal (solid and dashed lines); MEIs outside of this range are coloured by cell type. **(h)** As (g) but for distribution of green and UV MEI temporal frequency. 95% of MEIs were within an angle of $\pm 11.4°$ of the diagonal (solid and dashed lines). **(i)** As (g) but for distribution of green and UV MEI centre contrast. MEI contrast is shifted away from the diagonal (dashed line) towards UV by an angle of 33.2° due to the dominance of UV-sensitive S-opsin in the ventral retina. MEIs at an angle $> 45°$ occupy the upper left, colour-opponent (UV^{ON}-$green^{OFF}$) quadrant. **(j, k)** Fraction of MEIs per cell type that lie outside the angle about the diagonal containing 95% of MEIs for centre size and temporal frequency. Broad RGC response types indicated as in (7). **(l)** Fraction of MEIs per cell type in the upper-left, colour-opponent quadrant for contrast.

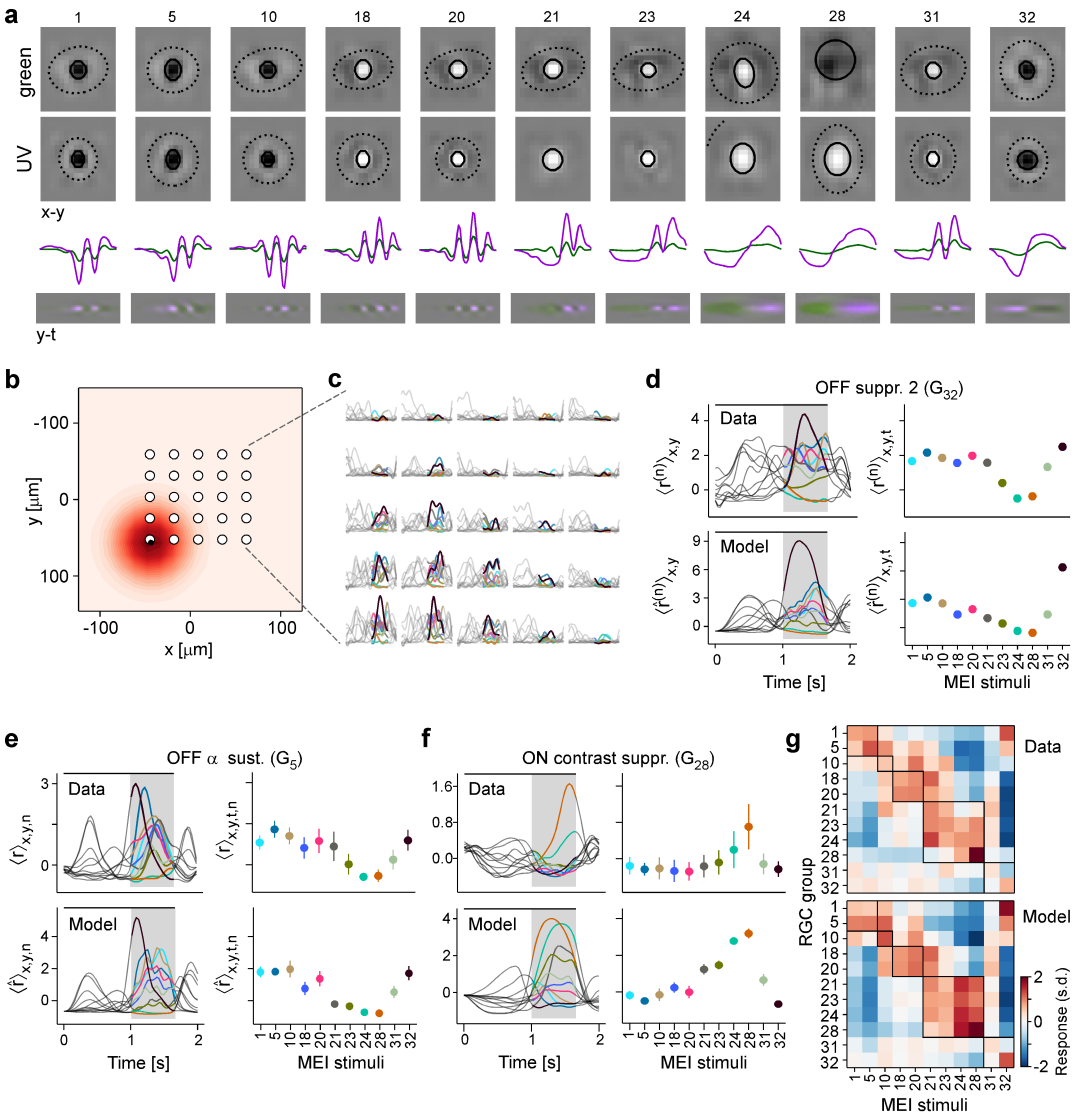

Fig. 4. Experiments confirm MEIs predicted by model (a) MEIs shown during the experiment, with green and UV spatial components (top two rows), as well as green and UV temporal components (third row) and a spatio-temporal visualisation (fourth row). For display, spatial components s in the two channels were re-scaled to a similar range and displayed on a common grey-scale map ranging from black for $-max(|s|)$ to white for $max(|s|)$, i.e. symmetric about 0 (grey). Relative amplitudes of UV and green are shown in the temporal components. **(b)** Illustration of spatial layout of MEI experiment. White circles represent 5×5 grid of positions where MEIs were shown; red shading shows an example RF estimate of a recorded G_{32} RGC, with black dot indicating the RF centre position (Methods). **(c)** Responses of example RGC from (b) to the 11 different MEI stimuli at 25 different positions. **(d)** Recorded (top, $r^{(n)}$) and predicted (bottom, $\hat{r}^{(n)}$) responses to the 11 different MEIs for example RGC n from (b, c). Left: responses are averaged across the indicated dimensions x, y (different MEI locations); black bar indicates MEI stimulus duration (from 0 to 1.66 s), grey rectangle marks optimisation time window (from 1 to 1.66 s). Right: Response to different MEIs, additionally averaged across time (t; within optimisation time window). **(e,f)** Same as in (d), but additionally averaged across all RGCs (n) of G_5 (N=6) (e) and of G_{28} (N=12) (f). Error bars show SD across cells. **(g)** Confusion matrix, each row showing the z-scored response magnitude of one RGC group (averaged across all RGCs of that group) to the MEIs in (a). Confusion matrix for recorded cells (top; "Data") and for model neurons (bottom; "Model"). Black squares highlight broad RGC response types according to (7): OFF cells, ($G_{1,5}$) ON-OFF cells (G_{10}), fast ON cells ($G_{18,20}$), slow ON ($G_{21,23,24}$) and ON contrast suppressed (G_{28}) cells, and OFF suppressed cells ($G_{31,32}$).

with synthetic stimuli, such as spots of different sizes (6), temporal frequency modulations (7), and stimuli of varying chromatic contrast (19, 20). The MEIs predicted by the model reproduce known properties of RGC groups (Fig. 3g-i). For example, sustained ON α RGCs (G_{24}), which are known to prefer large stimuli (7, 36), had MEIs with large centres (G_{24}, N=20 cells: green centre size, mean $\pm$ SD: 195 $\pm 82\,\mu$m; UV centre size $178 \pm 45\,\mu$m; average across all RGC groups: green $148 \pm 42\,\mu$m, UV $141 \pm 42\,\mu$m; see Fig. 3g). The MEI's temporal frequency relates to the temporal frequency preference of an RGC: MEIs of G_{20} and G_{21}, termed ON high frequency and ON low frequency (7), had high and low average temporal frequency, respectively (G_{20}, N=40 cells, green, mean $\pm$ SD: 2.71 ± 0.16 Hz, UV 2.86 ± 0.22 Hz; G_{21}, N=50 cells, green, mean $\pm$ SD: 2.32 ± 0.63 Hz, UV 1.98 ± 0.5 Hz; see Fig. 3h). Finally, the contrast of an MEI reflects what is traditionally called a cell's ON vs. OFF preference: MEIs of ON and OFF RGCs had positive and negative contrasts, respectively (Fig. 3i). An ON-OFF preference can be interpreted as a tuning map with two local optima - one in the OFF- and one in the ON-contrast regime. Consequently, the MEI optimisation resulted in MEIs with ON or OFF contrast, depending on the relative strengths of the two optima and on the initial condi-

tions (Suppl. Fig. S2, G_{10}). Together with the consistency of MEIs within functional RGC groups (Suppl. Fig. S2), these results provided strong evidence that RGCs grouped based on responses to synthetic stimuli (chirp and MB) also form functional groups in natural movie response space.

Our goal was to explore chromatic tuning of RGCs, specifically focusing on regions in chromatic stimulus space where a given stimulus property differs for green and UV. Therefore, for centre size and temporal frequency, we asked, which RGC groups contributed to the 5% MEIs furthest away from the diagonal (i. e. achromatic tuning) (Fig. 3g, h, j, k). These 5% MEIs furthest away from the diagonal were almost exclusively contributed by ON cells; and among these, more so by slow than by fast ON cells. MEI contrast needed to be analysed differently than size and temporal frequency for two reasons. First, due to the dominance of UV-sensitive S-opsin in the ventral retina (16) and the fixed contrast budget shared across channels (Methods), the contrast of most MEIs is strongly shifted towards UV (Fig. 3i). Second, contrast in green and UV can not only vary along positive valued axes (as is the case for size and temporal frequency), but can also take on opposite signs, resulting in colour-opponent stimuli. Whereas most MEIs had the same contrast polarity in both colour channels (i.e. both ON or OFF, Fig. 3c, blue and turquoise trajectories), some MEIs had opposing contrast polarities in UV and green (Fig. 3c, orange trajectory, and Fig. 3i, upper left quadrant). Thus, for contrast we asked which RGC groups contributed to colour-opponent MEIs (i.e. MEIs in the colour-opponent, upper left or lower right quadrant in Fig. 3i). Again, slow ON RGCs made up most of the cells with colour-opponent MEIs. Here, G_{28} stood out: 66% (24/36) of all cells of this group had colour-opponent MEIs (UV^{ON}-$green^{OFF}$), making up 39% of all colour-opponent MEIs, followed by G_{22} that contributed another 15% of colour-opponent MEIs. This was not a form of centre-surround colour-opponency as described before (19), but rather a centre-opponency ("co-extensive" colour-opponent RF; reviewed in (42)), as can be seen in the lower-dimensional visualisations (Fig. 3a,b, right column; 3c, orange trajectory).

In conclusion, our model-guided *in-silico* exploration of chromatic stimulus space revealed a variety of preferred stimuli that captured known properties of RGC groups, and highlighted G_{28} as exhibiting a previously unknown preference for centre colour-opponent, UV^{ON}-$green^{OFF}$ stimuli.

Experiments confirm selectivity for chromatic contrast. Next, we verified experimentally that the MEIs predicted for a given RGC group actually drive cells of that group optimally. To this end, we performed new experiments in which we added to our battery of stimuli a number of MEIs chosen according to the following criteria: We wanted the MEIs to (*i*) span the response space (ON, ON-OFF, OFF, transient, sustained, and contrast-suppressed) and (*ii*) to represent both well-described RGC types, such as α cells (i.e. $G_{5,24}$), as well as poorly understood RGC types, such as suppressed-by-contrast cells ($G_{28,31,32}$) (Fig. 4a). We therefore chose MEIs of RGCs from groups G_1 (OFF local), G_5 (OFF α sustained), G_{10} (ON-OFF local-edge), G_{18} (ON transient), G_{20} (ON high frequency), G_{21} (ON low frequency), G_{23} (ON mini α), G_{24} (sustained ON α), G_{28} (ON contrast suppressed), G_{31} (OFF suppressed 1), and G_{32} (OFF suppressed 2). For simplicity, in the following we refer to the MEI of an RGC belonging to group g as group g's MEI, or MEI g.

We presented these MEIs on a regularly spaced 5×5 grid to achieve approximate centring of stimuli on RGC RFs in the recording field (Fig. 4b,c). For these recordings, we fit models whose readout parameters allowed us to estimate the RGCs' RF locations. We used these RF location estimates to calculate a spatially weighted average of the responses to the MEIs displayed at different locations, weighting the response at each location proportional to the RF strengths at those locations (Fig. 4b, red highlight, and Fig. 4d, top). We then performed the same experiment *in-silico*, confirming that the model accurately predicts responses to the MEIs (Fig. 4d, bottom; Suppl. Fig. S3). These experiments allowed us to evaluate MEI responses at the RGC group level (Fig. 4e–f; Suppl. Fig. S2).

We expected RGCs to show a strong response to their own group's MEI, a weaker response to the MEIs of functionally related groups, and no response to MEIs of groups with different response profiles. Indeed, most RGC groups exhibited their strongest ($G_{5,20,21,28,32}$) or second-strongest ($G_{1,10,23}$) response to their own group's MEI (Fig. 4g, top). Conversely, RGC groups from opposing regions in response space showed no response to each others' MEIs (e.g. $G_{1,5}$ (OFF cells) vs. $G_{21\text{-}28}$ (slow ON cells)). The model's predictions showed a similar pattern (Fig. 4g, bottom), thereby validating the model's ability to generalise to the MEI stimulus regime.

Notably, G_{28} RGCs responded very selectively to their own MEI 28, displaying only weak responses to most other MEIs (Fig. 4f,g, selectivity index G_{28} to MEI 28 $SI_{G_{28}}(28)$ defined as the average difference in response between MEI 28 and all other MEIs in units of standard deviation of the response, mean $\pm$ SD: 2.58 ± 0.76; see Methods). This was in contrast to other RGC groups, such as G_{23} and G_{24}, that responded strongly to MEI 28, but also to other MEIs from the slow ON response regime (Fig. 4g, top; Suppl. Fig. S3, $SI_{G_{23}}(28)$, mean $\pm$ SD: 1.04 ± 0.69, $SI_{G_{24}}(28)$, mean $\pm$ SD: 1.01 ± 0.46).

Chromatic contrast selectivity derives from a warped representation of stimulus space. We hypothesised that the selectivity of G_{28} to MEI 28 derives from a warping of stimulus space that amplifies the distance between MEI 28 and all other MEIs in neuronal response space, compared to their original distances in stimulus space. We tested this hypothesis in the framework of representational similarity analysis (RSA) which allows comparing representations in terms of their geometry (43).

We considered two different kinds of representations of MEIs. First, a representation ϕ_s in stimulus space, where an MEI's location is defined by a vector holding its pixel values; and second, a representation $\phi_{\nu(g)}$ for each RGC group g,

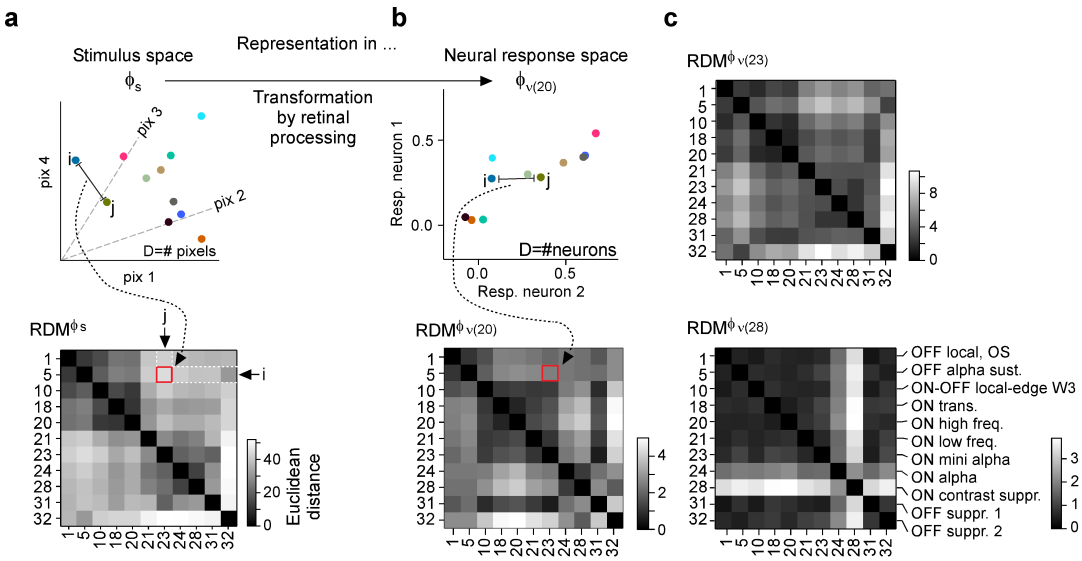

Fig. 5. Chromatic contrast selectivity derives from a warped representation of stimulus space (a) Schematic illustration of representations of MEIs in stimulus space. We consider the flattened spatio-temporal vectors defining each MEI as points in a D-dimensional space, where $D = \#channels \times height \times width \times time = 28,800$ (top). We then measure (dis)similarity between each pair (i, j) of MEIs in stimulus space representation, ϕ_s, as Euclidean distance between their vectors. We summarise the pairwise distances in a *representational dissimilarity matrix* RDM$^{\phi_s}$ (bottom), whose rows and columns are indexed by stimulus. **(b)** Same as (a), but in neural response space of G_{20} (real distances). Each MEI is represented as an N-dimensional vector holding the responses of N RGCs to this MEI, where $N = \#$neurons for a given RGC group (here, N=8). **(c)** Same as RDM in (b) but for groups G_{23} (top, N=40) and G_{28} (bottom, N=12).

where an MEI's location is defined by a response vector of dimension n, with n equal to the number of neurons in RGC group g (Fig. 5a,b, top panels). $\phi_{\nu(g)}$ thus denotes the representation of MEIs in the neuronal response space of RGC group g. Based on these representations, we computed pairwise distances between MEIs and summarised them in representational dissimilarity matrices (RDMs), whose rows and columns are indexed by stimulus (Fig 5a,b, bottom panels; and c; Suppl. Fig. S4a). RDMs capture the geometry of a representation (44). Therefore, we next correlated RDMs to yield a measure $\Gamma(\phi_{\nu(g)}, \phi_s)$ that quantifies the similarity between the geometries of different representations (see Methods). This analysis showed that the representational geometry of G_{28} is substantially warped relative to stimulus space, and more so than for most other RGC groups ($N_{28} = 12$, $\Gamma(\phi_{\nu(28)}, \phi_s) = 0.14$, bootstrapped 95% confidence interval $CI_{95} = [-0.05, 0.20]$, mean $\pm$ SD across all groups: 0.61 ± 0.23; Suppl. Fig. S4b, c, see Methods). This warping sacrifices discriminability in large regions of the stimulus space for increased discriminability of stimuli in the region of stimulus space that contains the centre colour-opponent MEIs 28 (44). We therefore hypothesised that the representation of visual input formed by G_{28} might serve to detect an ethologically relevant, colour-opponent feature from the visual scene. What may be this feature?

Warped representation allows for detection of ground-to-sky transitions. Studies analysing visual scenery from the mouse's perspective have repeatedly found that chromatic contrast changes strongly at the horizon (20, 22, 39, 40). Because the MEI of G_{28} consists of a large, slow and sustained change in luminance from green to UV, we hypothesised that the ethologically relevant feature represented by G_{28} might be a transition of the horizon across its RF, as elicited by head or eye movements. Being able to detect such transitions is useful for multiple reasons: For example, it could contribute to aligning the visual field with the horizon (45, 46) or signal the brain a change in context (i.e. ground vs. sky), which may help interpreting signals in other RGC channels.

To test if G_{28} responds to such a stimulus, we used the transitions between movie clips (*inter-clip transitions*; cf. Fig. 1b) as a proxy for the type of visual input elicited by head or eye movements: ground-to-ground and sky-to-sky transitions for horizontal movements, and ground-to-sky and sky-to-ground transitions for vertical movements crossing the horizon. We then calculated the contrast of these transitions in the green and UV channel and mapped them to the chromatic contrast stimulus space (Fig. 6a). We found that ground-to-ground and sky-to-sky transitions were distributed along the diagonal, whereas the two transitions resembling visual input elicited by vertical movements crossing the horizon fell into the two colour-opponent quadrants: sky-to-ground transitions in the lower right quadrant, and ground-to-sky transitions in the upper left quadrant (Fig. 6a,b). Thus, UV^{ON}-$green^{OFF}$ MEIs 28 share a location in stimulus space with ground-to-sky transitions in terms of chromatic contrast (cf. Fig 3i).

Do G_{28} RGCs indeed respond strongly to ground-to-sky transitions, i.e. to the "naturally occurring version" of their MEIs? To answer this question, we extracted the RGC responses to the inter-clip transitions, thereby mapping out their tuning across chromatic contrasts (Fig. 6c-e; Suppl. Fig. S5a, b), and then averaged the resulting single-cell tuning maps for each RGC group (Fig. 6f,g). G_{28} is most strongly tuned to full-field transitions in the upper left quadrant containing mostly ground-to-sky inter-clip transitions (Fig. 6f) – unlike, for example, non-colour-opponent

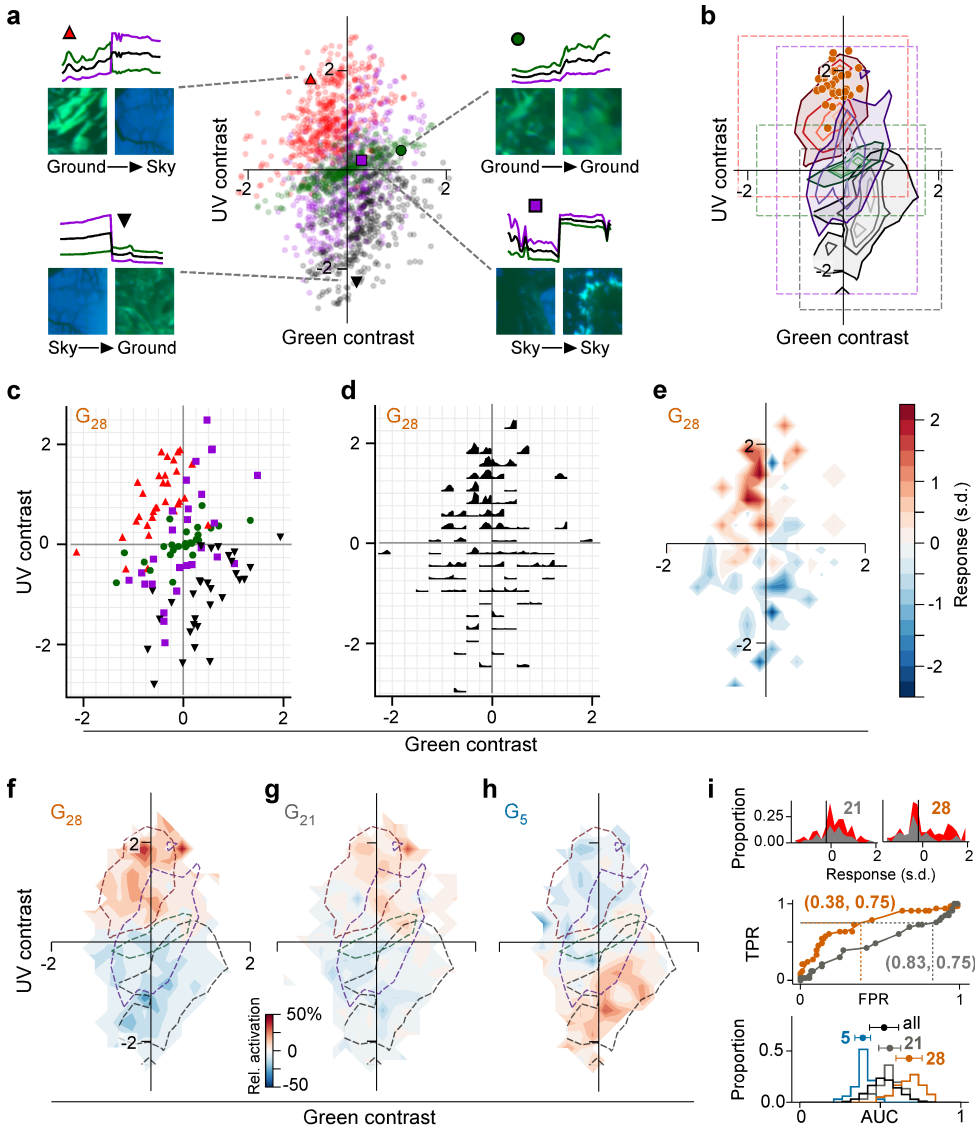

Fig. 6. Chromatic contrast tuning allows detection of ground-to-sky transitions (a) Distribution of green and UV contrasts of all movie inter-clip transitions (centre), separately for the 4 transition types, for each of which an example is shown: ground-to-sky (N=525, top left, red triangle), ground-to-ground (N=494, top right, green disk), sky-to-ground (N=480, bottom left, black downward triangle), and sky-to-sky (N=499, bottom right, purple square). Images show last and first frame of pre- and post-transition clip, respectively. Traces show mean full-field luminance of green and UV channels in the last and first 1 s of the pre- and post-transition clip. Black trace shows luminance averaged across the two colour channels. **(b)** Distributions as in (a), but shown as contours indicating iso-density lines of inter-clip transitions in chromatic contrast space. Density of inter-clip transitions was estimated separately for each type of transition from histograms within 10×10 bins that were equally spaced within the coloured boxes. Four levels of iso-density for each transition type shown, with density levels at 20 % (outermost contour, strongest saturation), 40 %, 60 % and 80 % (innermost contour, weakest saturation) of the maximum density observed per transition: 28 sky-to-ground (black), 75 ground-to-ground (green), 42 sky-to-sky (purple) and 45 ground-to-sky (red) transitions per bin. Orange markers indicate locations of N=36 G_{28} MEIs in chromatic contrast space (cf. Fig. 3i). **(c)** Distribution of green and UV contrasts of N=122 inter-clip transitions seen by an example G_{28} RGC, coloured by transition type. **(d)** Responses of example RGC in the 1 s following an inter-clip transition, averaged across transitions within the bins indicated by the grid. **(e)** Responses shown in (d), transformed into a tuning map by averaging within bins defined by grid in (c, d). Responses in (d,e) are z-scored ($\mu = 0$, $\sigma = 1$). **(f)** Tuning map of G_{28} RGCs (N=78), created by averaging the tuning maps of the individual RGCs (overlaid with outermost contour lines from (b)). **(g,h)** Same as (f) for G_{21} RGCs ((g), N=97) and G^5 ((h), N=33). **(i)** *Top*: Illustration of ROC analysis for two RGCs, a G_{21} (left) and a G_{28} (right). For each RGC, responses to all inter-clip transitions were binned, separately for ground-to-sky (red) and all other transitions (grey). *Middle*: Sliding a threshold d across the response range, classifying all transitions with response $> d$ as ground-to-sky, and registering the false-positive-rate (FPR) and true-positive-rate (TPR) for each threshold yields an ROC curve. Numbers in brackets indicate (FPR, TPR) at the threshold indicated by vertical line in histograms. *Bottom*: We evaluated performance for each cell as area under the ROC curve (AUC), and plotted the distribution across AUC values for all cells (black), G_{21} (grey), G_5 (blue), and G_{28} (orange); AUC mean $\pm$ SD indicated as dots and horizontal lines above histograms.

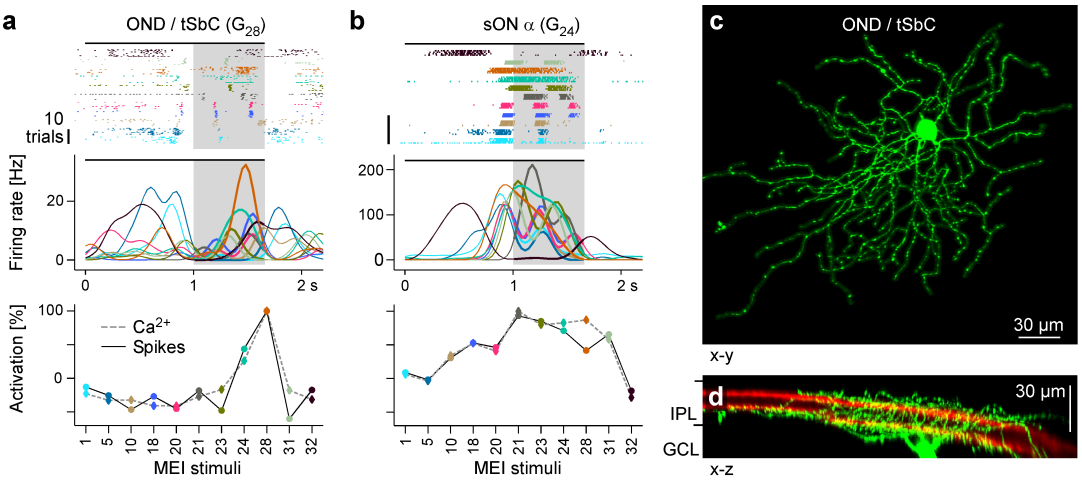

Fig. 7. Electrical single-cell recordings of responses to MEI stimuli confirm chromatic selectivity of tSbC RGCs. (a) Spiking activity (top, raster plot; middle, firing rate) of a OND RGC in response to different MEI stimuli (black bar indicates MEI stimulus duration; grey rectangle marks optimisation time window, from 1 to 1.66 s). Bottom: Activation relative to mean as a function of MEI stimulus, averaged across cells (solid line, from electrical recordings, N=4; dashed line, from Ca^{2+} imaging, N=11 cells). Colours as in Fig. 4. **(b)** Like (a) but for a sustained ON α cell (G_{24}; N=4 cells, both for electrical and Ca^{2+} recordings). **(c)** ON delayed (OND/tSbC, G_{28}) RGC (green) from (a) dye-loaded by patch pipette after cell-attached electrophysiology recording (z-projection; x-y plane). **(d)** Cell from (c, green) as side-projection (x-z), showing dendritic stratification pattern relative to choline-acetyltransferase (ChAT) amacrine cells (tdTomato, red) within the inner plexiform layer (IPL).

reference RGC groups from the slow ON and OFF response regime (Fig. 6g,h).

Could a downstream visual area detect ground-to-sky transitions based on input from G_{28} RGCs? To answer this question, we performed a linear detection analysis for each RGC by sliding a threshold across its responses to the inter-clip transitions, classifying all transitions that elicited an above-threshold response as ground-to-sky, and evaluating false-positive and true-positive rates (FPR and TPR, respectively) for each threshold (Fig. 6i for ventral cells; for dorsal cells, see Suppl. Fig. S6h). Plotting the resulting TPRs for all thresholds as a function of FPRs yields a receiver operating characteristic (ROC) curve (47) (Fig. 6i, middle). The area under this curve (AUC) can be used as a measure of detection performance: it is equivalent to the probability that a given RGC will respond more strongly to a ground-to-sky transition than to any other type of transition. Indeed, G_{28} RGCs achieved the highest AUC on average (Fig. 6i, bottom, and Suppl. Fig. S5c; G_{28}, mean $\pm$ SD AUC (N=78 cells): 0.68 ± 0.08; all other groups (N=3,449): 0.52 ± 0.09, $\Delta\text{AUC} = 0.16$, bootstrapped 95% confidence interval $\text{CI}_{95} = [0.14, 0.18]$, Cohen's $d = 1.82$, two-sample permutation test G_{28} vs. all other groups (see Methods): $p = 0$ with 100,000 permutations; next-best performing G_{24} (N=29): 0.61 ± 0.09, $\Delta\text{AUC} = 0.06$, bootstrapped 95% confidence interval $\text{CI}_{95} = [0.03, 0.1]$, Cohen's $d = 0.78$, two-sample permutation test G_{28} vs. G_{24}: $p = 0.00043$ with 100,000 permutations; for statistics and AUC distributions for dorsal cells, see Suppl. Fig. S6h legend). So far, we focused only on cells in the ventral retina. However, the horizon may also appear in the lower visual field, that is, on the dorsal retina, where RGCs receive weaker UV input (19). Therefore, we recorded additional fields in the dorsal retina (Suppl. Fig. S6a). Also here, G_{28} displayed the strongest tuning to ground-to-sky transitions among all dorsal RGCs (Suppl. Fig. S6b-g).

Hence, a downstream area, reading out from a single RGC group, would achieve the best performance in detecting ground-to-sky transitions if it based its decisions on inputs from G_{28} RGCs.

G_{28} corresponds to the transient Suppressed-by-Contrast RGC type. Next, we sought to identify which RGC type G_{28} corresponds to. In addition to its unique centre colour-opponency, the responses of G_{28} displayed a pronounced transient suppression to temporal contrast modulations (cf. chirp response in Fig. 1e). Therefore, we hypothesised that G_{28} corresponds to the transient Suppressed-by-Contrast (tSbC) RGC type (37, 38, 48), which is one of three retinal SbC RGC types identified so far and is also referred to as ON delayed (OND) cell because of its delayed response onset (49).

To test this hypothesis, we performed cell-attached electrophysiology recordings (Fig. 7) targeting tSbC/OND cells (N=4), identified by their responses to spots of multiple sizes (6), and later confirmed by their distinctive morphology ((49); type 73 in (8)) (Fig. 7c,d). We recorded spikes in these RGCs while presenting the MEI stimuli (Fig. 7a, top). Just like G_{28} RGCs in the Ca^{2+} imaging, tSbC/OND cells exhibited a pronounced selectivity for MEI 28, and were suppressed by most other MEIs (Fig. 7a, middle and bottom). Notably, the characteristic delayed response onset was visible in both the Ca^{2+} (Fig. 4f, top) and electrical (Fig. 7a) responses but was not predicted by the model (Fig. 4f, bottom). As a control, we also recorded MEI responses of a different, well-characterised RGC type, sustained (s) ON α (G_{24}; (50)) (Fig. 7b, top; N=4). Again, the electrical recordings of the cells' MEI responses yielded virtually the same results as the Ca^{2+} imaging (Fig. 7b, middle and bottom; cf. Suppl. Fig. S3). Crucially, sON α cells were not selective for MEI 28. These results suggest that the tSbC/OND RGC type signals horizon transitions to downstream visual centres in the brain.

Discussion

We combined large-scale recordings of RGC responses to natural movie stimulation with CNN-based modelling to investigate colour processing in the mouse retina. By searching the stimulus space *in silico* to identify *most exciting inputs* (MEIs), we found a novel type of chromatic tuning in a distinct RGC type, the tSbC cells. We revealed this RGC type's pronounced and unique selectivity for full-field changes from green-dominated to UV-dominated scenes, a stimulus that matches the statistics of ground-to-sky horizon transitions in natural scenes. Therefore, we suggest that tSbC cells may signal horizon transitions within their RF. Beyond our focus on tSbC cells, our study demonstrates the utility of an *in silico* approach for generating and testing hypotheses about the ethological relevance of sensory representations.

Circuit mechanisms for colour-opponency in tSbC RGCs. Most previous studies of colour-opponency in the mouse retina have identified sparse populations of colour-opponent RGCs that have not been systematically assigned to a particular functional type (19, 20). The only studies that have examined the mechanisms of colour-opponency in identified mouse RGC types showed a centre-surround organisation, with RF centre and surround having different chromatic preferences (17, 51); ((52), but see (53)), whereas we found that tSbC RGCs respond to spatially co-extensive colour-opponent stimuli. While centre-surround opponency has been attributed to the opsin gradient (51) and rod contributions in the outer retina (17, 19), the circuitry for spatially co-extensive opponency in mice remains unknown. It seems unlikely, though, that the opsin gradient plays a major role in the cell's colour opponency, because both ventral and dorsal tSbC cells preferentially responded to full-field green-to-UV transitions. In primates, spatially co-extensive colour-opponency in small bistratified RGCs is thought to arise from the selective wiring of S-ON and M/L-OFF bipolar cells onto the inner and outer dendritic strata, respectively ((54), but see (55)). A similar wiring pattern seems unlikely for tSbC RGCs, since their inner dendrites do not co-stratify with the S-ON (type 9) bipolar cells, nor do their outer dendrites co-stratify with the candidate M-OFF bipolar cell (type 1) (56). The large RF centres of the tSbC cells, extending well beyond their dendritic fields, come from a non-canonical circuit, in which tonic inhibition onto the RGC via $GABA_B$ receptors is relieved via serial inhibition from different amacrine cells using $GABA_C$ receptors (36). An intriguing possibility is that a colour-selective amacrine cell is part of this circuit, perhaps supporting chromatically tuned disinhibition in the absence of selective wiring from the aforementioned cone-selective bipolar cells onto the RGC.

A new functional role for tSbC RGCs. Suppressed-by-contrast responses have been recorded along the early visual pathway in dorsal lateral geniculate nucleus (dLGN), superior colliculus (SC), and primary visual cortex (V1) (57–59), with their function still being debated (60). In the retina, three types of SbC RGCs have so far been identified (reviewed in (49)), among them the tSbC cell (36–38). Despite their relatively recent discovery, tSbC RGCs have been suggested to play a role in several different visual computations. The first report of their light responses in mice connected them to the SbC RGCs previously discovered in rabbit, cat, and macaque, and suggested a role in signalling self-generated stimuli, perhaps for saccade suppression (37). Aided by a new intersectional transgenic line to selectively label tSbC RGCs (38), their projections were traced to areas in SC, v- and dLGN, and nucleus of the optic tract (NOT). The latter stabilises horizontal eye movements; however, as the medial terminal nucleus (MTN), which serves stabilisation of vertical eye movements, lacks tSbC innervation, it is unclear whether and how these RGCs contribute to gaze stabilisation. A retinal study identified the circuit mechanisms responsible for some of the unique spatial and temporal response properties of tSbC cells and suggested a possible role in defocus detection to drive emmetropization in growing eyes and accommodation in adults (36, 61). Here, we identified another potential role for these RGCs in vision based on the chromatic properties of their RFs: signalling visual context changes (see also next section). These different possible functional roles are not mutually exclusive, and might even be complementary in some cases, highlighting the difficulty in assigning single features to distinct RGC types (15). In particular, the centre colour-opponency that we discovered in tSbC RGCs could serve to enhance their role in defocus detection by adding a directional signal (myopic vs. hyperopic) based on the chromatic aberration of lens and cornea (62). Future studies may test these theories by manipulating these cells *in vivo* using the new transgenic tSbC mouse line (38).

Behavioural relevance of horizon detection. The horizon is a prominent landmark in visual space: it bisects the visual field into two regions, ground and sky. Visual stimuli carry different meaning depending on where they occur relative to the horizon, and context-specific processing of visual inputs is necessary for selecting appropriate behavioural responses (reviewed in (63)). For example, it is sensible to assume that a looming stimulus above the horizon is a predator, the appropriate response to which would be avoidance (that is, escape or freezing). A similar stimulus below the horizon, however, is more likely to be harmless or even prey. To allow for time-critical perceptual decisions – predator or prey – and corresponding behavioural response selection – avoidance or approach – it might be useful to share circuitry for the joint extraction and processing of contextual and stimulus information. Notably, VGluT3-expressing amacrine cells (a "hub" for distributing information about motion) represent a shared element in upstream circuitry, providing opposite-sign input to tSbC and to RGCs implicated in triggering avoidance behaviour, such as tOFF α (12, 50) and W3 cells (64). In downstream circuitry, SbC inputs have been found to converge with "conventional" RGC inputs onto targets in dLGN and NOT; whether tSbC axons specifically converge with tOFF α or W3 axons remains to be tested. Such convergence may allow "flagging" the activity of these RGCs with their local context (sky/threat or ground/no threat).

***In-silico* approaches to linking neural tuning and function.** The modelling of retinal responses to natural stimuli has advanced our understanding of the complexity of retinal processing in recent years. As suggested in a recent review, it is helpful to consider the contributions of different studies in terms of one of three perspectives on the retinal encoding of natural scenes: The circuit perspective ("how?"), the normative perspective ("why?"), and the coding perspective ("what?") (65). For example, an *in-silico* dissection of a CNN model of the retina offered explanations on how the surprisingly complex retinal computations, such as motion reversal, omitted stimulus response, and polarity reversal, emerge from simpler computations within retinal circuits (25, 26). Taking on the normative perspective, anatomically constrained deep CNNs trained on image recognition suggested a dependency between the complexity of retinal representations and the computational power of downstream cortical networks: Whereas a computationally powerful cortex, as found in primates, can deal with faithful, linear representations of visual inputs, a simpler cortical circuitry, as found in mice, pushes feature extraction upstream, into the retina (66, 67). However, the full potential of CNN models as tools for understanding sensory processing goes beyond reproducing effects that are already described in the literature. Reaping this potential for a take at the coding perspective ("what?"), Goldin and colleagues (68) performed a search in static stimulus space for locally optimal inputs for RGCs and found that the selectivity for positive or negative contrast in a subset of cells is context-dependent. These cells signal absolute contrast, a feature one level of abstraction and complexity above response polarity ("classical" ON vs. OFF).

Here, we developed an approach that allows investigating the complexity of retinal processing from all three perspectives: A global search for most exciting mouse RGC inputs in dynamic, chromatic stimulus space answers the question of *what* it is that retinal neurons encode. Classifying individual RGCs into types then allows to dissect the circuitry and to explain *how* specific retinal computations are implemented. And finally, interpreting the abstract features extracted by the retina against the backdrop of natural stimulus space points to *why* these features might be behaviourally relevant.

Methods

Animals and tissue preparation. All imaging experiments were conducted at the University of Tübingen; the corresponding animal procedures were approved by the governmental review board (Regierungspräsidium Tübingen, Baden-Württemberg, Konrad-Adenauer-Str. 20, 72072 Tübingen, Germany) and performed according to the laws governing animal experimentation issued by the German Government. All electrophysiological experiments were conducted at Northwestern University; the corresponding animal procedures were performed according to standards provided by Northwestern University Center for Comparative Medicine and approved by the Institutional Animal Care and Use Committee (IACUC).

For all imaging experiments, we used 4- to 15-week-old C57Bl/6 J mice (n=23; JAX 000664) of either sex (10 male, 13 female). These animals were housed under a standard 12 h day/night rhythm at 22° and 55% humidity. On the day of the recording experiment, animals were dark-adapted for at least 1 h, then anaesthetised with isoflurane (Baxter) and killed by cervical dislocation. All following procedures were carried out under very dim red (> 650 nm) light. The eyes were enucleated and hemisected in carboxygenated (95% O_2, 5% CO_2) artificial cerebrospinal fluid (ACSF) solution containing (in mM): 125 NaCl, 2.5 KCl, 2 $CaCl_2$, 1 $MgCl_2$, 1.25 NaH_2PO_4, 26 $NaHCO_3$, 20 glucose, and 0.5 L-glutamine at pH 7.4. Next, the retinae were bulk-electroporated with the fluorescent Ca^{2+} indicator Oregon–Green BAPTA-1 (OGB-1), as described earlier (69). In brief, the dissected retina was flat-mounted onto an Anodisc (#13, 0.2 µm pore size, GE Healthcare) with the RGCs facing up, and placed between a pair of 4-mm horizontal plate electrodes (CUY700P4E/L, Nepagene/Xceltis). A 10-μl drop of 5 mM OGB-1 (hexapotassium salt; Life Technologies) in ACSF was suspended from the upper electrode and lowered onto the retina. Next, nine pulses ($\approx$ 9.2 V, 100 ms pulse width, at 1 Hz) from a pulse generator/wideband amplifier combination (TGP110 and WA301, Thurlby handar/Farnell) were applied. Finally, the tissue was placed into the microscope's recording chamber, where it was perfused with carboxygenated ACSF (at $\approx$ 36° C) and left to recover for $\geq$ 30 min before recordings started. To visualise vessels and damaged cells in the red fluorescence channel, the ACSF contained $\approx$ 0.1 μM Sulforhodamine-101 (SR101, Invitrogen) (70). All procedures were carried out under dim red (> 650 nm) light.

For electrophysiology experiments, we used ChAT-Cre (JAX 006410) x Ai14 (JAX 007914) mice on a C57Bl/6J background (n=2, male, aged 27 and 30 weeks). Mice were housed with siblings in groups up to 4, fed normal mouse chow and maintained on a 12:12 h light/dark cycle. Before the experiment, mice were dark-adapted overnight and sacrificed by cervical dislocation. 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 (71), and preserved using relieving cuts in cardinal directions, with the largest cut at the dorsal retina. Retinas were mounted on 12mm poly-D-lysine coated glass affixed to a recording dish with grease, with the GCL up. Oxygenation was maintained by superfusing the dish with carboxygenated Ames medium (US Biological, A1372-25) warmed to 32 °C. For cell-attached single cell recordings, we used Symphony software (https://symphony-das.github.io/) with custom extensions (https://github.com/Schwartz-AlaLaurila-Labs/sa-labs-extension).

Owing to the exploratory nature of our study, we did not use randomisation and blinding. No statistical methods were used to predetermine sample size.

Two-photon calcium imaging. We used a MOM-type two-photon microscope (designed by W. Denk; purchased from Sutter Instruments) (70, 72), which was equipped with a mode-locked Ti:Sapphire laser (MaiTai-HP DeepSee, Newport Spectra-Physics) tuned to 927 nm, two fluorescence detection channels for OGB-1 (HQ 510/84, AHF/Chroma) and SR101 (HQ 630/60, AHF), and a water immersion objective (CF175 $LWD \times 16$/0.8W, DIC N2, Nikon, Germany). Image acquisition was performed with custom-made software (ScanM by M. Müller and T.E.) running under IGOR Pro 6.3 for Windows (Wavemetrics), taking time-lapsed 64×64 pixel image scans ($\approx (100\ \mu\text{m})^2$ at 7.8125 Hz (Fig. 1c). For simplicity, we refer to such a time-lapsed scan of a local population of GCL cells as a "recording". For documenting the position of the recording fields, the retina under the microscope was oriented such that the most ventral edge pointed always towards the experimenter. In addition, higher resolution images (512×512 pixel) were acquired and recording field positions relative to the optic nerve were routinely logged.

Data preprocessing. Ca^{2+} traces were extracted for individual ROIs as described previously (7, 19). Extracted traces $\boldsymbol{r}_{raw}$ were then detrended to remove slow drifts in the recorded signal that were unrelated to changes in the neural response. First, a smoothed version of the traces, $\boldsymbol{r}_{smooth}$, was calculated by applying a Savitzky-Golay filter of 3rd polynomial order and a window length of 60 s using the SciPy implementation `scipy.signal.savgol_filter`. This smoothed version was then subtracted from the raw traces to yield the detrended traces.

$$\boldsymbol{r}_{detrend} = \boldsymbol{r}_{raw} - \boldsymbol{r}_{smooth}$$

To make traces non-negative ($\boldsymbol{r}_{nn}$), we then clipped all values smaller than the 2.5^{th} percentile, $\eta_{2.5}$, to that value, and then subtracted $\eta_{2.5}$ from the detrended traces:

$$\boldsymbol{r}_{nn} = \boldsymbol{r}_{detrend} - \eta_{2.5}$$

This procedure (i.e. clipping to, and subtracting $\eta_{2.5}$) was more robust than simply subtracting the minimum.

Finally, traces were then divided by the standard deviation within the time window before stimulus start at t_0:

$$\boldsymbol{r} := \boldsymbol{r}_{final} = \frac{\boldsymbol{r}_{nn}}{SD(\boldsymbol{r}_{nn_{[:t_0]}})}$$

For training the model on movie response, we estimated firing rates from the detrended Ca^{2+} traces using the package C2S (`https://github.com/lucastheis/c2s`, Theis et al. (73)).

Inclusion criteria. We applied a sequence of quality filtering steps to recorded cells before analysis illustrated in Suppl. Fig. S1c. As a first step, we applied a general response quality criterion, defined as a sufficiently reliable response to the Moving bar stimulus (as quantified by a quality index $QI_{MB} > 0.6$), *or* a sufficiently reliable response to the chirp stimulus (as quantified by a quality index $QI_{chirp} > 0.35$). The quality index is defined as in ref.(7):

$$QI = \frac{\text{Var}[\langle \boldsymbol{r} \rangle_i]_t}{\langle \text{Var}[\boldsymbol{r}]_t \rangle_i}$$

where $\boldsymbol{r}$ is the T by I response matrix (time samples by stimulus repetitions) and $\langle\rangle_x$ and $\text{Var}[]_x$ denote the mean and variance across the indicated dimension x, respectively.

The second and third step made sure only cells were included that were assigned to a ganglion cell group (i.e., group index between 1 and 32) with sufficient confidence. Confidence is defined as the probability assigned to the predicted class by the random forest classifier (see (41)), and the threshold was set at ≥ 0.25.

The fourth step made sure only cells with a sufficient model prediction performance, defined as an average single-trial test set correlation of $\langle C(\hat{r}^{(n)}, r_i^{(n)}) \rangle_i > .3$, were included.

All cells passing steps 1-3 were included in the horizon detection analysis (Fig. 6); all cells passing steps 1-4 were included in the MEI analysis (Fig. 3); the "red" cells passing steps 1-4 were included in the MEI validation analysis (Fig. 4). In the process of analysing MEIs, we fitted DoGs to their green and UV spatial component (see Methods section Concentric anisotropic 2D Difference-of-Gaussians fit). For the analysis of MEI properties (temporal frequency, centre size, chromatic contrast), we only included cells with a sufficient DoG goodness-of-fit, determined as a value of the cost function of $< .11$ for both green and UV on the resulting DoG fit. This threshold was determined by visual inspection of the DoG fits and led to the inclusion of 1613 out of 1947 RGCs in the MEI property analysis.

Visual stimulation. For light stimulation (imaging experiments), we projected the image generated by a digital light processing (DLP) projector (lightcrafter DPM-FE4500MKIIF, EKB Technologies Ltd) through the objective onto the tissue. The lightcrafter featured a light-guide port to couple in external, band-pass filtered UV and green LEDs (light-emitting diodes) (green: 576 BP 10, F37-576; UV: 387 BP 11, F39-387; both AHF/Chroma) (74). To optimise spectral separation of mouse M- and S-opsins, LEDs were band-pass filtered (390/576 dual-band, F59-003, AHF/Chroma). LEDs were synchronised with the microscope's scan retrace. Stimulator intensity (as photoisomerization rate, $10^3\ P^*s^{-1}$ per cone) was calibrated to range from ≈ 0.5 (black image) to ≈ 20 for M- and S-opsins, respectively. Additionally, we estimated a steady illumination component of $\approx 10^4\ P^*s^{-1}$ per cone to be present during the recordings because of two photon excitation of photopigments (70, 72). Before data acquisition, the retina was adapted to the light stimulation by presenting a binary noise stimulus (20×15 matrix, $(40\,\mu\text{m})^2$ pixels, balanced random sequence) at 5 Hz for 5 min to the tissue.

For electrophysiology experiments, stimuli were presented using a digital projector (DPM-FE4500MKII, EKB Technologies Ltd) at a frame rate of 60 Hz and a spatial res-

olution of 1140 × 912 pixels (1.3 μm per pixel) focused on the photoreceptor layer. Neutral density filters (Thorlabs), a multi-band pass filter (69000x, Chroma), and a custom LED controller circuit were used to attenuate the light intensity of stimuli either to match that of the Ca^{2+} imaging experiments (for MEI presentation) or to range from ≈ 0-200 $P^* s^{-1}$ per rod (for cell identification). Stimuli were presented using Symphony software (https://symphony-das.github.io/) with custom extensions (https://github.com/Schwartz-AlaLaurila-Labs/sa-labs-extension).

Identifying retinal ganglion cell types. To functionally identify RGC groups in the Ca^{2+} imaging experiments, we used our default "fingerprinting" stimuli, as described earlier (7). These stimuli included a full-field (700 μm in diameter) chirp stimulus, and a 300 × 1,000 μm bright bar moving at 1,000 μm $\cdot s^{-1}$ in eight directions across the recording field (with the shorter edge leading; Fig. 1b).

The procedure and rationale for identifying cells in the electrophysiological recordings is presented in ref. (6). Cells with responses that qualitatively matched that of the OND and ON α types were included in the study. Following recording, cells were filled with AlexaFluor-488 by patch pipette and imaged under a two-photon microscope. Dendrites were traced in Fiji (NIH) using the SNT plugin (75). Dendritic arbours were computationally flattened using a custom MATLAB tool (https://doi.org/10.5281/zenodo.6578530) based on the method in ref. (76) to further confirm their identity as morphological type 73 from ref. (8).

Mouse natural movies. The natural movie stimulus consisted of clips of natural scenes recording outside in the field with a specialised, calibrated camera (22). This camera featured a fish-eye lens, and two spectral channels, UV (band-pass filter F37-424, AHF, >90% transmission at 350–419 nm) and green (F47-510, >90%, 470–550 nm, AHF), approximating the spectral sensitivities of mouse opsins (35). In mice, eye movements often serve to stabilise the image on the retina during head movements (45). Therefore, the camera was also stabilised by mounting it on a gimbal. As a result, the horizon bisected the camera's visual field.

A *mouse cam movie* frame contained a circular field of view (FOV) of 180° corresponding to 437 pixels along the diameter. To minimise the influence of potential chromatic and spatial aberrations introduced by the lenses, we focused on image cut-outs (crops; $30° \times 26°$, equivalent to 72×64 pixels in size) from upper and lower visual field, centred at $[28°, 56°]$ and $[-42°, -31°]$, respectively, relative to the horizon (for details, see (22)). Our *stimulus movie* consisted of 113 movie clips, each 150 frames (= 5 s) long. 108 clips were randomly reordered for each recording and split into two 54 clips-long *training sequences*. The remaining 5 clips formed a fixed *test sequence* that was presented before, in between, and after the training sequences (Fig. 1b). To keep intensity changes at clip transitions small, we only used clips with mean intensities between 0.04 and 0.22 (for intensities in $[0, 1]$). For display during the experiments, intensities were then mapped to the range covered by the stimulator, i.e. $[0, 255]$.

Convolutional neural network model of the retina. We trained a convolutional neural network (CNN) model to predict responses of RGCs to a dichromatic natural movie. The CNN model consisted of two modules, a convolutional core that was shared between all neurons, and a readout that was specific for each neuron (77).

The core module was modelled as a two-layer convolutional neural network with 16 feature channels in each layer. Both layers consisted of space-time separable 3D convolutional kernels followed by a batch normalisation layer and an ELU nonlinearity. In the first layer, sixteen $2 \times 11 \times 11 \times 21$ (c=#channels × h=height × w=width × t=#frames) kernels were applied as valid convolution; in the second layer, sixteen $16 \times 5 \times 5 \times 11$ kernels were applied with zero padding along the spatial dimensions. We parameterised the temporal kernels as Fourier series and added one time stretching parameter per recording to account for inter-experimental variability affecting the speed of retinal processing (78).

In the readout, we modelled each cell's spatial receptive field (RF) as a 2D isotropic Gaussian, parameterised as $\mathcal{N}(\mu_x, \mu_y; \sigma)$. We then modelled the neural response as an affine function of the core feature maps weighted by the spatial RF, followed by a softplus nonlinearity.

For the linearised version of the model, the architecture was exactly the same except for the fact that there was no ELU nonlinearity after both convolutional layers.

Model training and evaluation. We trained our network by minimising the Poisson loss

$$\sum_{i=1}^{N} \left(\hat{\boldsymbol{r}}^{(n)} - \boldsymbol{r}^{(n)} \log \hat{\boldsymbol{r}}^{(n)} \right)$$

where N is the number of neurons, $\boldsymbol{r}^{(n)}$ is the measured and $\hat{\boldsymbol{r}}^{(n)}$ the predicted firing rate of neuron n for an input of duration t=50 frames. We followed the training schedule of Lurz et al. (2021)(79). Specifically, we used early stopping (80) on the correlation between predicted and measured neuronal responses on the validation set, which consisted of 15 out of the 108 movie clips. If the correlation failed to increase during any 5 consecutive passes through the entire training set (epochs), we stopped the training and restored the model to the best performing model over the course of training. We went through 4 cycles of early stopping, restoring the model to the best performing, and continuing training, each time reducing the initial learning rate of 0.01 by a learning rate decay factor of 0.3. Network parameters were iteratively optimised via stochastic gradient descent using the Adam optimiser (81) with a batch size of 32 and a chunk size (number of frames for each element in the batch) of 50. For all analyses and MEI generation, we used an ensemble of models as described in ref. (34). Briefly, we trained 5 instances of the same model initialised with different random seeds. Inputs to the ensemble model were passed to each member and the

final ensemble model prediction was obtained by averaging the outputs of the 5 members. For ease of notation, we thus redefine $\hat{\boldsymbol{r}}^{(n)}$ to be the *ensemble* model prediction.

After training, we evaluated model performance for each modelled neuron n as the correlation to the mean, i.e. the correlation between predicted response $\hat{r}^{(n)}$ and measured response $r^{(n)}$ to the held-out test sequence, the latter averaged across 3 repetitions $i = \{1,2,3\}$: $C(\hat{r}^{(n)}, \langle r_i^{(n)} \rangle_i)$. Unlike the single-trial correlation $C(\hat{r}^{(n)}, r_i^{(n)})$ which is always limited to values < 1 by inherent neuronal noise, a perfect model can in theory achieve a value of 1 for the correlation to the mean, in the limit of infinitely many repetitions when the sample average $\langle r^{(n)_i} \rangle_i$ is a perfect estimate of the true underlying response $\rho^{(n)}$. The observed correlation to the mean can thus be interpreted as an estimate of the fraction of the maximally achievable correlation achieved by our model. For deciding which cells to exclude from analysis, we used average single-trial correlation ($\langle C(\hat{r}^{(n)}, r_i^{(n)}) \rangle_i$) since this measure reflects both model performance as well as reliability of the neuronal response to the movie stimulus for neuron n (see also Methods section on Inclusion criteria).

Synthesising MEIs. We synthesised maximally exciting inputs for RGCs as described previously (32). Formally, for each model neuron n we wanted to find

$$\boldsymbol{x}^{*(n)} = \arg\max_{\boldsymbol{x}} \langle \hat{\boldsymbol{r}}^{(n)}(\boldsymbol{x})_{30:50} \rangle_t, \tag{1}$$

i.e. the input $\boldsymbol{x}^{*(n)}$ where the model neuron's response $\langle \hat{\boldsymbol{r}}(\boldsymbol{x})_{30:50} \rangle_t$, averaged across frames 30 to 50, attains a maximum, subject to norm and range constraints (see below). To this end, we randomly initialised an input $\boldsymbol{x}_0^{(n)} \in \mathcal{R}^{c \times w \times h \times t}$ of duration t=50 frames with Gaussian white noise, and then iteratively updated $x_i^{(n)}$ according to the gradient of the model neuron's response:

$$\boldsymbol{x}_{i+1}^{(n)} = \boldsymbol{x}_i^{(n)} + \lambda \frac{\delta}{\delta \boldsymbol{x}_i^{(n)}} \langle \hat{\boldsymbol{r}}^{(n)}(\boldsymbol{x}_i^{(n)})_{30:50} \rangle_t, \tag{2}$$

where $\lambda = 10$ was the learning rate. The optimisation was performed using Stochastic Gradient Descent (SGD), and was subject to a norm and a range constraint. The norm constraint was applied jointly across both channels and ensured that the L2 norm of each MEI did not exceed a fixed budget b of 30. The norm-constrained MEI $\tilde{\boldsymbol{x}}_i^{(n)}$ was calculated at each iteration as

$$\tilde{\boldsymbol{x}}_i^{(n)} = \frac{b}{||\boldsymbol{x}_i^{(n)}||_2} \times \boldsymbol{x}_i^{(n)} \tag{3}$$

The range constraint was defined and applied for each colour channel separately and ensured that the range of the MEI values stayed within the range covered by the training movie. This was achieved by clipping values of the MEI exceeding the range covered by the training movie to the minimum or maximum value. Optimisation was run for at least 100 iterations, and then stopped when the number of iterations reached 1000, or when it had converged (whichever occurred first). Convergence was defined as 10 consecutive iterations with a change in model neuron activation of less than 0.001; model neuron activations ranged from ≈ 1 to ≈ 10. We denote the resulting MEI for neuron n as $\boldsymbol{x}^{*(n)}$.

Analysing MEIs. We analysed MEIs to quantify their spatial, temporal, and chromatic properties.

Spatial and temporal components of MEIs. For each colour channel c, we decomposed the spatiotemporal MEIs into a spatial component and a temporal component (henceforth spatial component and temporal component) by singular value decomposition:

$$U, S, V = \text{svd}(\boldsymbol{x}_c^{*(n)})$$

with $x_c^{*(n)} \in \mathcal{R}^{50 \times 288}$ for $c \in$ [green, UV] is the MEI of neuron n in a given colour channel with its spatial dimension (18x16=288) flattened out. As a result, any spatiotemporal dependencies are removed and we only analyse spatial and temporal properties separately. The following procedures were carried out in the same manner for the green and the UV component of the MEI, and we drop the color channel index c for ease of notation. The temporal component is then defined as the first left singular vector, $U_{:1}$, and the spatial component is defined as the first right singular vector, $V_{:1}^T$, reshaped to the original dimensions 18×16.

Concentric anisotropic 2D Difference-of-Gaussians fit. We modelled the spatial component as concentric anisotropic Difference-of-Gaussians (DoG) using the nonlinear least-squares solver `scipy.optimize.least_squares` with soft-L1 loss function (40). The DoGs were parameterized by a location (μ_x, μ_y) shared between centre and surround, amplitudes A^c, A^s, variances (σ_x^c, σ_y^c), (σ_x^s, σ_y^s), and rotation angles θ^c, θ^s separately for centre and surround:

$$\text{DoG} = G^c - G^s$$

with

$$\begin{aligned} G^c(x,y) = \mathrm{A}^c \exp(&-f^c(x-\mu_x)^2 \\ &+2g^c(y-\mu_y)(x-\mu_x) \\ &+h^c(y-\mu_y)^2) \end{aligned}$$

and

$$\begin{aligned} f^c &= \frac{\cos^2\theta^c}{2\sigma_x^c} + \frac{\sin^2\theta^c}{2\sigma_y^c}, \\ g^c &= \frac{\sin 2\theta^c}{4\sigma_y^c} - \frac{\sin 2\theta^c}{4\sigma_y^c}, \\ h^c &= \frac{\sin^2\theta^c}{2\sigma_x^c} + \frac{\cos^2\theta^c}{2\sigma_y^c}, \end{aligned}$$

and likewise for G^s. We initialised (μ_x, μ_y) in the following way: Since we set the model readout's location parameters to (0, 0) for all model neurons when generating their MEIs, we also expected the MEIs to be centred at (0, 0), as well. Hence, we determined the location of the minimum and the maximum value of the MEI; whichever was closer to the centre (0,0) provided the initial values for the parameters (μ_x, μ_y). Starting from there, we then first fit a single Gaussian to the MEI, and took the resulting parameters as initial parameters for the DoG fit. This was a constrained optimisation problem, with lower and upper bounds on all parameters; in particular, such that the location parameter would not exceed the canvas of the MEI, and such that the variance would be strictly positive.

MEI properties.

Centre size We defined the diameter of the centre of the MEI in the horizontal and the vertical orientation, respectively, as $d_x^c = 2\sigma_x^c$ and $d_y^c = 2\sigma_y^c$. The centre size was calculated as $\frac{1}{2}(d_x^c + d_y^c)$. We then estimated a contour outlining the MEI centre as the line that is defined by all points at which the 2D centre Gaussian G^c attains the value $G^c(x, y)$ with $(x, y) = (\mu_x + \sigma_x^c, \mu_y + \sigma_y^c)$. The centre mask m was then defined as a binary matrix with all pixels within the convex hull of this contour being 1 and all other pixels set to 0. This mask is used for calculating centre chromatic contrast (see below).

Temporal frequency To estimate temporal frequency of the MEIs, we estimated the power spectrum of the temporal components using a Fast Fourier Transform after attenuating high frequency noise by filtering with a 5th order low-pass Butterworth filter with cutoff frequency 10 Hz. We then estimated the mean frequency of the temporal component by calculating an average of the frequency components, each weighted with its relative power.

Contrast The contrast of the MEIs in the two channels, $\gamma(\boldsymbol{x}_c^{*(n)})$ for $c \in$ [green, UV], was defined as the difference between the mean value within the centre mask m at the two last peaks of the temporal component of the MEI in the UV channel at time points t_2 and t_1:

$$\gamma(\boldsymbol{x}_c^{*(n)}) = (\boldsymbol{x}_c^{*(n)} \odot m)(t_2) - (\boldsymbol{x}_c^{*(n)} \odot m)(t_1),$$

where $\odot$ denotes the element-wise multiplication of the MEI and the binary mask. (see Fig. 3 f). The peaks were found with the function `scipy.signal.find_peaks`, and the peaks found for the UV channel were used to calculate contrast both in the green and the UV channel.

Validating MEIs experimentally.

Generating MEI stimuli. In order to test experimentally whether the model correctly predicts which stimuli would maximally excite RGCs of different RGC groups, we performed a new set of experiments (numbers indicated in red in Suppl. Fig. S1c) where we complemented our stimulus set with MEI stimuli. For the MEI stimuli, we selected 11 RGCs, chosen to span the responses space and to represent both well-described and poorly understood RGC groups, for which we generated MEIs at different positions on a 5×5 grid (spanning $110\mu m$ in vertical and horizontal direction). We decomposed the MEIs as described above, and reconstructed MEIs as rank 1 tensors by taking the outer product of the spatial and temporal components:

$$\bar{\boldsymbol{x}}^* = S_{11} U_{:1} \otimes V_{:1}^T$$

We used this rank 1 reconstruction of the MEIs as stimuli, since the model's convolutional kernels were space-time-separable, hence could not capture RGC selectivity to spatio-temporally varying stimulus properties, and so any MEI components of rank > 1 were assumed to be spurious. The MEI stimuli, lasting 50 frames (1.66 s) were padded with 10 frames (.34 s) of inter-stimulus grey, and were randomly interleaved. With 11 stimuli, presented at 25 positions and lasting 2 s each, the total stimulus duration was $11 \times 25 \times 2\ s = 550\ s$. Since the model operated on a z-scored (0 mean, 1 SD) version of the movie, MEIs as predicted by the model lived in the same space and had to be transformed back to the stimulator range ([0, 255]) before being used as stimuli in an experiment by scaling with the movie's SD and adding the movie's mean. The MEIs' green channel was then displayed with the green LED, and the UV channel was displayed with the UV LED. For experiments at Northwestern University, an additional transform was necessary to achieve the same levels of photoreceptor activation (photoisomerization rates) for M- and S-cones with different LEDs. To ensure proper chromatic scaling between the different experimental apparatuses with different spectral profiles, we described the relative activation of M- and S-cones by the green and UV LEDs in the stimulation setup used in the two photon imaging experiments (setup A) by a matrix

$$\mathbf{A} = \begin{bmatrix} a_{mg} & a_{sg} \\ a_{mu} & a_{su} \end{bmatrix} = \begin{bmatrix} 1 & 0.19 \\ 0 & 1 \end{bmatrix},$$

and the relative activation of M- and S-cones by the stimulation setup used in the patch-clamp experiments (setup B) by a matrix

$$\mathbf{B} = \begin{bmatrix} b_{mg} & b_{sg} \\ b_{mu} & b_{su} \end{bmatrix} = \begin{bmatrix} 1 & 0.9 \\ 0.035 & 1 \end{bmatrix},$$

where diagonal entries describe the activation of M-cones by the green LED, and of S-cones by the UV LED, and entries in the off-diagonal describe the cross-activation (i.e., M-cones by UV-LED and S-cones by green LED). The activation of M-cones and S-cones $\mathbf{e}^{\mathbf{T}} = (e_m, e_s)$ by a stimulus $\boldsymbol{x} \in \mathcal{R}^{2\times 1}$ displayed on a given stimulation setup was approximated as $\mathbf{e} = \mathbf{A}\boldsymbol{x}$ (82). Hence, a stimulus $\boldsymbol{x}'$ displayed on setup B, defined as $\boldsymbol{x}' = \mathbf{B}^{-1}\mathbf{A}\boldsymbol{x}$, will achieve the same photoreceptor activation as stimulus $\boldsymbol{x}$ displayed on setup A. Since the solution exceeded the valid range of the stimulator ([0, 255]), we added an offset and multiplied with a scalar factor to ensure all stimuli were within the valid range.

Analysing RGC responses to MEI stimuli. We wanted to evaluate the responses of RGCs to the MEI stimuli in a spatially resolved fashion, i.e. weighting responses to MEIs displayed at different locations proportional to the strength of the RGCs RF at that location. In order to be able to meaningfully compare MEI responses between RGCs and across groups, for each RGC, we first centred and scaled the responses to zero mean and a standard deviation of 1. Then, for each RGC n, we computed a spatial average of its responses, weighting its responses at each spatial location (x, y) proportional to the Gaussian density $\mathcal{N}_{\boldsymbol{\mu_n},\sigma_n}(x, y)$, where the parameters of the Gaussian $\boldsymbol{\mu_n} = (\mu_x, \mu_y), \sigma_n$ were the model's estimated readout parameters for neuron n (Fig. 4 b, c, d left):

$$\langle \boldsymbol{r}^{(n)} \rangle_{x,y} = \sum_{x'=1}^{5} \sum_{y'=1}^{5} \boldsymbol{r}^{(n)}_{x',y'} \cdot \mathcal{N}_{\boldsymbol{\mu_n},\sigma_n}(x', y')$$

where $\boldsymbol{r}^{(n)}_{x',y'} \in \mathcal{R}^{11\times 60}$ is the 60 frames (2 s) long response of neuron n to an MEI at position $(x, y) = (x', y')$, resampled from the recording frame rate of 7.81 Hz to 30 Hz. We then averaged $\langle \boldsymbol{r}^{(i)} \rangle_{x,y}$ across time in the optimisation time window, i.e. frames 30-50, to get a scalar response $\tilde{r}^{(n)} = \langle \boldsymbol{r}^{(n)} \rangle_{x,y,t}$ for each MEI stimulus (Fig. 4, d).

Selectivity index. To quantify the selectivity of the response $\tilde{r}^{(n)}(\boldsymbol{x_i^*})$ of an RGC n to an MEI $\boldsymbol{x_i^*}$, we defined a selectivity index as follows. First, we standardised the responses $\tilde{r}^{(n)}$ across all MEIs by subtracting the mean and dividing by the standard deviation. The selectivity index of RGC group $\mathrm{G_g}$ to MEI $\boldsymbol{x_i^*}$ was then defined as

$$\mathrm{SI}_g(\boldsymbol{x_i^*}) = \langle \tilde{r}^{(n)}(\boldsymbol{x_i^*}) - \frac{1}{10} \sum_{j=1}^{11} \delta_{ij} \tilde{r}^{(n)}(\boldsymbol{x_j^*}) \rangle_n,$$

where δ_{ij} is the Kronecker delta. In words, the SI is the difference (in units of SD response) between the response to the MEI of interest ($\boldsymbol{x_i^*}$) and the mean response to all other (10) MEIs, $\frac{1}{10}\sum_{j=1}^{11} \delta_{ij} \tilde{r}^{(n)}(\boldsymbol{x_j^*})$, averaged across all cells n belonging to the group of interest $\mathrm{G_g}$.

Representational similarity analysis. To compare how RGC groups transform visual input into neural representations, and how these representations differ between groups, we performed representational similarity analysis (RSA) on the RGC group responses to the MEI stimuli (43). First, we considered representations ϕ_k of MEIs in different spaces. In stimulus space, MEIs are represented by the vectors holding their pixel values,

$$\phi_s(\boldsymbol{x}^*) \in \mathcal{R}^d, \qquad \textbf{(4)}$$

where $d = c \times h \times w \times t$. Second, a group of n neurons spans an n-dimensional neural response space, where each MEI is represented as the n-dimensional response vector,

$$\phi_\nu(\boldsymbol{x}^*) = (\tilde{r}^{(1)}, \tilde{r}^{(2)}, ..., \tilde{r}^{(n)}), \qquad \textbf{(5)}$$

where $\tilde{r}^{(n)}$ as defined above.

We then calculated the pairwise Euclidean distances between MEIs in these different representations, and summarised them in representational dissimilarity matrices (RDMs):

$$\mathrm{RDM}^{\phi_k}_{i,j} = ||(\phi_k(\boldsymbol{x}^{*(i)}) - \phi_k(\boldsymbol{x}^{*(j)}))||_2$$

The entry $\mathrm{RDM}^{\phi_k}_{i,j}$ thus tells us how (dis)similar MEIs $\boldsymbol{x}^{*(i)}$ and $\boldsymbol{x}^{*(j)}$ are under the representation ϕ_k (Fig. 5 a-c and Suppl. Fig. S4a). To quantify how (dis)similar MEI representations ϕ_k and ϕ_l are from each other, we then define a measure $\Gamma(\phi_k, \phi_l)$ (Suppl. Fig. S4b, c) as

$$\Gamma(\phi_k, \phi_l) = C(triu(\mathrm{RDM}^{\phi_k}), triu(\mathrm{RDM}^{\phi_l})), \qquad \textbf{(6)}$$

where C is the correlation coefficient, and $triu(\mathrm{RDM}^{\phi_k})$ is the upper triangle of RDM^{ϕ_k}, since RDMs are symmetric about the diagonal. The correlation coefficient can in theory take values between -1 and 1, but in practice, we did not observe any anti-correlated RDMs.

Detection performance analysis. To test the performance of individual RGCs of different groups in detecting the target class of inter-clip transitions (ground-to-sky) from all other classes of inter-clip transitions, we performed a receiver operating characteristic (ROC) analysis (47). For each RGC, we calculated its response to an inter-clip transition occurring at time t_0 as the baseline-subtracted average response within 1 second following the transition, i.e. $\frac{1}{T}\sum_{t=0}^{T} r(t) - r(t_0)$, with T=30 frames at 30 Hz. For all n=40 equally spaced thresholds within the response range of a RGC, we then calculated the true positive rate (TPR) and false positive rate (FPR) of a hypothetical classifier classifying all transitions eliciting an above-threshold response as a positive, and all other transitions as negative. Plotting the TPR as a function of FPR yields an ROC curve, the area under which (AUC) is equivalent to the probability that the RGC will respond more strongly to a randomly chosen inter-clip transition of the target class than to a randomly chosen inter-clip transition of a different class. The AUC thus is a measure of performance for RGCs in this detection task.

Statistical analysis.

Permutation test. We wanted to test how likely the difference in AUC observed for different RGC groups are to occur under the null hypothesis that the underlying distributions they are sampled from are equal. To this end, we performed a permutation test. We generated a null distribution for our test statistic, the absolute difference in AUC values ΔAUC, by shuffling the RGC group labels of the two groups of interest (e.g. $\mathrm{G_{28}}$ and $\mathrm{G_{24}}$) and calculating the test statistic with shuffled labels 100,000 times. We then obtained a p-value for ΔAUC observed with true labels as the proportion of entries in the null distribution larger than ΔAUC.

Bootstrapped confidence intervals. We bootstrapped confidence intervals for ΔAUC (Fig. 6 and Suppl. Fig. S6) and for $\Gamma(\phi_s, \phi_{\nu_g})$ (Fig. 5 and Suppl. Fig. S4). For ΔAUC, we generated a bootstrapped distribution by sampling 100 times with replacement from the AUC values of the two groups that were being compared and calculated ΔAUC. We then estimated the 95 % confidence interval for ΔAUC as the interval defined by the 2.5^{th} and 97.5^{th} percentile of the bootstrapped distribution of ΔAUC.

For $\Gamma(\phi_s, \phi_{\nu_g})$, we generated a bootstrapped distribution by sampling 100 times with replacement from the MEI responses of RGC group g and then calculating $\text{RDM}^{\phi_{\nu_g}}$ and $\Gamma(\phi_s, \phi_{\nu_g})$ for each sample. We then estimated the 95 % confidence interval for $\Gamma(\phi_s, \phi_{\nu_g})$ as the interval defined by the 2.5^{th} and 97.5^{th} percentile of the bootstrapped distribution of $\Gamma(\phi_s, \phi_{\nu_g})$.

Estimating effect size. The effect size of difference in AUC observed for different RGC groups l and k, ΔAUC (Fig. 6 and Suppl. Fig. S6), was estimated as Cohen's d (83, 84):

$$\frac{|m_k - m_l|}{s},$$

with

$$s = \sqrt{\frac{(N_k - 1)s_k^2 + (N_l - 1)s_l^2}{N_k + N_l - 2}}$$

and m_k and s_k the sample mean and standard deviation, respectively, of the AUC observed for the N_k RGCs of group k.

Estimating linear correlation. Wherever the linear correlation between two paired samples x and y of size N was calculated (for evaluating model performance, Fig. 2, Suppl. Figs. S1, S3; and similarity between representations, Fig. 5, Suppl. Fig. S4), we used Pearson's correlation coefficient:

$$C_{xy} = \frac{\sum_i^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i^N (x_i - \bar{x})^2}\sqrt{\sum_i^N (y_i - \bar{y})^2}}$$

Data availability

The data and the movie stimulus will be made available at `https://retinal-functomics.net` upon journal publication.

Code availability

Custom analysis and model training code will be made available at `https://github.com/eulerlab` upon journal publication.

Author contributions

L. H.: Conceptualisation, methodology, software, validation, formal analysis, data curation, writing (original draft), visualisation **K. P. S.**: validation, investigation, data curation **C. B.**: methodology, software, formal analysis **Y. Q.**: resources, software **D. A. K.**: methodology, software, writing (review and editing) **Z. J.**: methodology, validation, investigation, visualisation, writing (original draft) **G. W. S.**: methodology, validation, investigation, writing (original draft), supervision, funding acquisition **M. B.**: conceptualisation, writing (review and editing), supervision, funding acquisition **P. B.**: conceptualisation, writing (review and editing), supervision, funding acquisition **K. F.**: conceptualisation, writing (review and editing), supervision, funding acquisition **A. S. E.**: conceptualisation, writing (review and editing), supervision, funding acquisition **T. E.**: conceptualisation, writing (original draft), visualisation, supervision, funding acquisition, project administration.

ACKNOWLEDGEMENTS

We thank Jonathan Oesterle and Dominic Gonschorek for feedback on the manuscript, Jan Lause for statistical consulting, and Merle Harrer for general assistance. We also thank all members of the Sinz lab for regular discussions on the project. This work was supported by the German Research Foundation (DFG; CRC 1233 "Robust Vision: Inference Principles and Neural Mechanisms", project number 276693517 to P.B., M.B., T.E., K.F.; Heisenberg Professorship, BE5601/8-1 to P.B.; Excellence Cluster EXC 2064, project number 390727645 to P.B.), the Federal Ministry of Education and Research (FKZ 01IS18039A to P.B.), National Institutes of Health (NIH; NEI EY031029, NEI EY031329 to G.W.S.; NEI F30EY031565, NIGMS T32GM008152 to Z.J.), and the European Research Council (ERC) under the European Union's Horizon Europe research and innovation programme (A.S.E, grant agreement No. 101041669).

Bibliography

1. J Y Lettvin, H R Maturana, W S McCulloch, and W H Pitts. What the Frog's Eye Tells the Frog's Brain. *Proceedings of the IRE*, 47(11):1940–1959, 1959.
2. Tom Baden, Thomas Euler, and Philipp Berens. Understanding the retinal basis of vision across species. *Nature Reviews Neuroscience*, 21(1):5–20, 2020. ISSN 14710048. doi: 10.1038/s41583-019-0242-1.
3. Maxwell H. Turner, Luis Gonzalo Sanchez Giraldo, Odelia Schwartz, and Fred Rieke. Stimulus- and goal-oriented frameworks for understanding natural vision. *Nature Neuroscience*, 22(1):15–24, 2019. ISSN 15461726. doi: 10.1038/s41593-018-0284-0.
4. Eero P Simoncelli and Bruno A Olshausen. Natural Image Statistics and Neural Representation. *Annual Reviews Neurosciences*, 24:1193–1216, 2001.
5. Daniel Kerschensteiner. Feature Detection by Retinal Ganglion Cells. *Annual Review of Vision Science*, 8(1):1–35, 9 2022. ISSN 2374-4642. doi: 10.1146/annurev-vision-100419-112009.
6. Jillian Goetz, Zachary F. Jessen, Anne Jacobi, Adam Mani, Sam Cooler, Devon Greer, Sabah Kadri, Jeremy Segal, Karthik Shekhar, Joshua R. Sanes, and Gregory W. Schwartz. Unified classification of mouse retinal ganglion cells using function, morphology, and gene expression. *Cell Reports*, 40(2):111040, 7 2022. ISSN 22111247. doi: 10.1016/j.celrep.2022.111040.
7. Tom Baden, Philipp Berens, Katrin Franke, Miroslav Román Rosón, Matthias Bethge, and Thomas Euler. The functional diversity of retinal ganglion cells in the mouse. *Nature*, 529(7586):345–350, 2016. ISSN 15206041. doi: 10.1038/nature16468.
8. J. Alexander Bae, Shang Mu, Jinseop S. Kim, Nicholas L. Turner, Ignacio Tartavull, Nico Kemnitz, Chris S. Jordan, Alex D. Norton, William M. Silversmith, Rachel Prentki, Marissa Sorek, Celia David, Devon L. Jones, Doug Bland, Amy L.R. Sterling, Jungman Park, Kevin L. Briggman, and H. Sebastian Seung. Digital Museum of Retinal Ganglion Cells with Dense Anatomy and Physiology. *Cell*, 173(5):1293–1306, 2018. ISSN 10974172. doi: 10.1016/j.cell.2018.04.040.
9. Bruce A. Rheaume, Amyeo Jereen, Mohan Bolisetty, Muhammad S. Sajid, Yue Yang, Kathleen Renna, Lili Sun, Paul Robson, and Ephraim F. Trakhtenberg. Single cell transcriptome profiling of retinal ganglion cells identifies cellular subtypes. *Nature Communications*, 9(1), 2018. ISSN 20411723. doi: 10.1038/s41467-018-05134-3.
10. Emily M. Martersteck, Karla E. Hirokawa, Mariah Evarts, Amy Bernard, Xin Duan, Yang Li, Lydia Ng, Seung W. Oh, Benjamin Ouellette, Joshua J. Royall, Michelle Stoecklin, Quanxin Wang, Hongkui Zeng, Joshua R. Sanes, and Julie A. Harris. Diverse Central Projection Patterns of Retinal Ganglion Cells. *Cell Reports*, 18(8):2058–2072, 2 2017. ISSN 22111247. doi: 10.1016/j.celrep.2017.01.075.
11. Keith P Johnson, Michael J Fitzpatrick, Lei Zhao, Bing Wang, Sean McCracken, Philip R Williams, and Daniel Kerschensteiner. Cell-type-specific binocular vision guides predation in mice. *Neuron*, 3 2021. ISSN 0896-6273. doi: 10.1016/j.neuron.2021.03.010.
12. Thomas A. Münch, Rava Azeredo Da Silveira, Sandra Siegert, Tim James Viney, Gautam B. Awatramani, and Botond Roska. Approach sensitivity in the retina processed by a multifunctional neural circuit. *Nature Neuroscience*, 12(10):1308–1316, 2009. ISSN 10976256. doi: 10.1038/nn.2389.
13. Melis Yilmaz and Markus Meister. Rapid innate defensive responses of mice to looming visual stimuli. *Current Biology*, 23(20):2011–2015, 2013. ISSN 09609822. doi: 10.1016/j.cub.2013.08.015.

14. T. Kim, N. Shen, J. C. Hsiang, K. P. Johnson, and D. Kerschensteiner. Dendritic and parallel processing of visual threats in the retina control defensive responses. *Science Advances*, 6(47):1–12, 2020. ISSN 23752548. doi: 10.1126/sciadv.abc9920.
15. Gregory W. Schwartz and David Swygart. Circuits for Feature Selectivity in the Inner Retina. In *The Senses: A Comprehensive Reference*, pages 275–292. Elsevier, 1 2020. doi: 10.1016/b978-0-12-809324-5.24186-2.
16. A. Szél, P. Röhlich, A. R. Caffé, B. Juliusson, G. Aguirre, and T. Van Veen. Unique topographic separation of two spectral classes of cones in the mouse retina. *Journal of Comparative Neurology*, 325(3):327–342, 1992. ISSN 0021-9967. doi: 10.1002/cne.903250302.
17. Maximilian Joesch and Markus Meister. A neuronal circuit for colour vision based on rod-cone opponency. *Nature*, 532(7598):236–239, 2016. ISSN 14764687. doi: 10.1038/nature17158.
18. Tom Baden, Timm Schubert, Le Chang, Tao Wei, Mariana Zaichuk, Bernd Wissinger, and Thomas Euler. A tale of two retinal domains: Near-Optimal sampling of achromatic contrasts in natural scenes through asymmetric photoreceptor distribution. *Neuron*, 80(5):1206–1217, 2013. ISSN 08966273. doi: 10.1016/j.neuron.2013.09.030.
19. Klaudia P. Szatko, Maria M. Korympidou, Yanli Ran, Philipp Berens, Deniz Dalkara, Timm Schubert, Thomas Euler, and Katrin Franke. Neural circuits in the mouse retina support color vision in the upper visual field. *Nature Communications*, 11(1), 2020. ISSN 20411723. doi: 10.1038/s41467-020-17113-8.
20. Mohammad Hossein Khani and Tim Gollisch. Linear and nonlinear chromatic integration in the mouse retina. *Nature Communications*, 12(1):1900, 12 2021. ISSN 2041-1723. doi: 10.1038/s41467-021-22042-1.
21. Josh W. Mouland, Abigail Pienaar, Christopher Williams, Alex J. Watson, Robert J. Lucas, and Timothy M. Brown. Extensive cone-dependent spectral opponency within a discrete zone of the lateral geniculate nucleus supporting mouse color vision. *Current Biology*, 31(15):3391–3400, 2021. ISSN 18790445. doi: 10.1016/j.cub.2021.05.024.
22. Yongrong Qiu, Zhijian Zhao, David Klindt, Magdalena Kautzky, Klaudia P. Szatko, Frank Schaeffel, Katharina Rifai, Katrin Franke, Laura Busse, and Thomas Euler. Natural environment statistics in the upper and lower visual field are reflected in mouse retinal specializations. *Current Biology*, 31(15):3233–3247, 2021. ISSN 18790445. doi: 10.1016/j.cub.2021.05.017.
23. Dimokratis Karamanlis and Tim Gollisch. Nonlinear spatial integration underlies the diversity of retinal ganglion cell responses to natural images. *Journal of Neuroscience*, 41(15):3479–3498, 4 2021. ISSN 15292401. doi: 10.1523/JNEUROSCI.3075-20.2021.
24. Tatyana O. Sharpee. Computational identification of receptive fields. *Annual Review of Neuroscience*, 36:103–120, 2013. ISSN 0147006X. doi: 10.1146/annurev-neuro-062012-170253.
25. Niru Maheswaranathan, Lane T. McIntosh, Hidenori Tanaka, Satchel Grant, David B. Kastner, Josh B. Melander, Aran Nayebi, Luke Brezovec, Julia Wang, Surya Ganguli, and Stephen A. Baccus. The dynamic neural code of the retina for natural scenes. *bioRxiv*, page 340943, 2018. ISSN 2692-8205.
26. Hidenori Tanaka, Aran Nayebi, Niru Maheswaranathan, Lane T. McIntosh, Stephen A. Baccus, and Surya Ganguli. From deep learning to mechanistic understanding in neuroscience: the structure of retinal prediction. *Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS 2019)*, (33):1–11, 2019.
27. Eleanor Batty, Josh Merel, Nora Brackbill, Alexander Heitman, Alexander Sher, Alan Litke, E. J. Chichilnisky, and Liam Paninski. Multilayer Recurrent Network Models of Primate Retinal Ganglion Cell Responses. *Proceedings of the 5th International Conference on Learning Representations*, (5), 2017.
28. Daniel L.K. Yamins, Ha Hong, Charles F Cadieu, Ethan A Solomon, Darren Seibert, and James J. DiCarlo. Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 111(23):8619–8624, 2014. ISSN 10916490. doi: 10.1073/pnas.1403112111.
29. Santiago A Cadena, George H Denfield, Edgar Y Walker, Leon A Gatys, Andreas S Tolias, Matthias Bethge, and Alexander S Ecker. Deep convolutional models improve predictions of macaque V1 responses to natural images. *PLoS Computational Biology*, 15(4):1–28, 2019. ISSN 15537358. doi: 10.1371/journal.pcbi.1006897.
30. Alexander S Ecker, Fabian H Sinz, Emmanouil Froudarakis, Paul G Fahey, Santiago A Cadena, Edgar Y Walker, Erick Cobos, Jacob Reimer, Andreas S Tolias, and Matthias Bethge. A rotation-equivariant convolutional neural network model of primary visual cortex. *Proceedings of the 7th International Conference on Learning Representations*, (7), 2019.
31. Ivan Ustyuzhaninov, Max F Burg, Santiago A Cadena, Jiakun Fu, Taliah Muhammad, Kayla Ponder, Emmanouil Froudarakis, Zhiwei Ding, Matthias Bethge, Andreas S Tolias, and Alexander S Ecker. Digital twin reveals combinatorial code of non-linear computations in the mouse primary visual cortex. *bioRxiv*, page 2022.02.10.479884, 2 2022. doi: 10.1101/2022.02.10.479884.
32. Edgar Y. Walker, Fabian H. Sinz, Erick Cobos, Taliah Muhammad, Emmanouil Froudarakis, Paul G. Fahey, Alexander S. Ecker, Jacob Reimer, Xaq Pitkow, and Andreas S. Tolias. Inception loops discover what excites neurons most using deep predictive models. *Nature Neuroscience*, 22(December), 2019. ISSN 15461726. doi: 10.1038/s41593-019-0517-x.
33. Pouya Bashivan, Kohitij Kar, and James J. DiCarlo. Neural population control via deep image synthesis. *Science*, 364(6439), 2019. ISSN 10959203. doi: 10.1126/science.aav9436.
34. Katrin Franke, Konstantin F Willeke, Kayla Ponder, Mario Galdamez, Na Zhou, Taliah Muhammad, Saumil Patel, Emmanouil Froudarakis, Jacob Reimer, Fabian H Sinz, and Andreas S Tolias. State-dependent pupil dilation rapidly shifts visual feature selectivity. *Nature*, 9 2022. ISSN 0028-0836. doi: 10.1038/s41586-022-05270-3.
35. Gerald H. Jacobs, Gary A. Williams, and John A. Fenwick. Influence of cone pigment coexpression on spectral sensitivity and color vision in the mouse. *Vision Research*, 44(14):1615–1622, 2004. ISSN 00426989. doi: 10.1016/j.visres.2004.01.016.
36. Adam Mani and Gregory W. Schwartz. Circuit Mechanisms of a Retinal Ganglion Cell with Stimulus-Dependent Response Latency and Activation Beyond Its Dendrites. *Current Biology*, 27(4):471–482, 2017. ISSN 09609822. doi: 10.1016/j.cub.2016.12.033.
37. Nai Wen Tien, James T. Pearson, Charles R. Heller, Jay Demas, and Daniel Kerschensteiner. Genetically identified suppressed-by-contrast retinal ganglion cells reliably signal self-generated visual stimuli. *Journal of Neuroscience*, 35(30):10815–10820, 2015. ISSN 15292401. doi: 10.1523/JNEUROSCI.1521-15.2015.
38. Nai-Wen Tien, Carmela Vitale, Tudor C. Badea, and Daniel Kerschensteiner. Layer-Specific Developmentally Precise Axon Targeting of Transient Suppressed-by-Contrast Retinal Ganglion Cells (tSbC RGCs). *The Journal of Neuroscience*, 42(38):2332–21, 9 2022. ISSN 0270-6474. doi: 10.1523/jneurosci.2332-21.2022.
39. Luca Abballe and Hiroki Asari. Natural image statistics for mouse vision. *Plos One*, 17(1): e0262763, 2022. doi: 10.1371/journal.pone.0262763.
40. Divyansh Gupta, Wiktor Mlynarski, Olga Symonova, Jan Svaton, and Maximilian Joesch. Panoramic visual statistics shape retina-wide organization of receptive fields. *bioRxiv*, pages 1–31, 2022.
41. Yongrong Qiu, David A Klindt, Klaudia P Szatko, Dominic Gonschorek, Larissa Hoefling, Timm Schubert, Laura Busse, Matthias Bethge, and Thomas Euler. Efficient coding of natural scenes improves neural system identification. *bioRxiv*, 2022.
42. Gregory William Schwartz. Chapter 16 - Color processing. In Gregory William Schwartz, editor, *Retinal Computation*, pages 288–317. Academic Press, 2021. ISBN 978-0-12-819896-4. doi: https://doi.org/10.1016/B978-0-12-819896-4.00017-2.
43. Nikolaus Kriegeskorte, Marieke Mur, and Peter Bandettini. Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, 2(NOV):1–28, 2008. ISSN 16625137. doi: 10.3389/neuro.06.004.2008.
44. Nikolaus Kriegeskorte and Xue Xin Wei. Neural tuning and representational geometry. *Nature Reviews Neuroscience*, 22(11):703–718, 2021. ISSN 14710048. doi: 10.1038/s41583-021-00502-3.
45. Arne F. Meyer, John O'Keefe, and Jasper Poort. Two Distinct Types of Eye-Head Coupling in Freely Moving Mice. *Current Biology*, 30(11):2116–2130, 2020. ISSN 18790445. doi: 10.1016/j.cub.2020.04.042.
46. Philip R.L. Parker, Elliott T.T. Abe, Emmalyn S.P. Leonard, Dylan M. Martins, and Cristopher M. Niell. Joint coding of visual input and eye/head position in V1 of freely moving mice. *Neuron*, 9 2022. ISSN 08966273. doi: 10.1016/j.neuron.2022.08.029.
47. Tom Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, 6 2006. ISSN 01678655. doi: 10.1016/j.patrec.2005.10.010.
48. Nai Wen Tien, Tahnbee Kim, and Daniel Kerschensteiner. Target-Specific Glycinergic Transmission from VGluT3-Expressing Amacrine Cells Shapes Suppressive Contrast Responses in the Retina. *Cell Reports*, 15(7):1369–1375, 5 2016. ISSN 22111247. doi: 10.1016/j.celrep.2016.04.025.
49. Jason Jacoby and Gregory William Schwartz. Typology and Circuitry of Suppressed-by-Contrast Retinal Ganglion Cells. *Frontiers in Cellular Neuroscience*, 12(August):1–7, 2018. ISSN 16625102. doi: 10.3389/fncel.2018.00269.
50. Brenna Krieger, Mu Qiao, David L. Rousso, Joshua R. Sanes, and Markus Meister. Four alpha ganglion cell types in mouse retina: Function, structure, and molecular signatures. *PLoS ONE*, 12(7):1–21, 2017. ISSN 19326203. doi: 10.1371/journal.pone.0180091.
51. Le Chang, Tobias Breuninger, and Thomas Euler. Chromatic Coding from Cone-type Unselective Circuits in the Mouse Retina. *Neuron*, 77(3):559–571, 2013. ISSN 08966273. doi: 10.1016/j.neuron.2012.12.012.
52. Maureen E. Stabio, Shai Sabbah, Lauren E. Quattrochi, Marissa C. Ilardi, P. Michelle Fogerson, Megan L. Leyrer, Min Tae Kim, Inkyu Kim, Matthew Schiel, Jordan M. Renna, Kevin L. Briggman, and David M. Berson. The M5 Cell: A Color-Opponent Intrinsically Photosensitive Retinal Ganglion Cell. *Neuron*, 97(1):150–163, 2018. ISSN 10974199. doi: 10.1016/j.neuron.2017.11.030.
53. Takuma Sonoda, Yudai Okabe, and Tiffany M. Schmidt. Overlapping morphological and functional properties between M4 and M5 intrinsically photosensitive retinal ganglion cells. *Journal of Comparative Neurology*, 528(6):1028–1040, 4 2020. ISSN 10969861. doi: 10.1002/cne.24806.
54. Dennis M. Dacey and Barry B. Lee. The 'blue-on' opponent pathway in primate retina originates from a distinct bistratified ganglion cell type. *Nature*, 367(6465):731–735, 1994. ISSN 00280836. doi: 10.1038/367731a0.
55. Greg D. Field, Alexander Sher, Jeffrey L. Gauthier, Martin Greschner, Jonathon Shlens, Alan M. Litke, and E. J. Chichilnisky. Spatial properties and functional organization of small bistratified ganglion cells in primate retina. *Journal of Neuroscience*, 27(48):13261–13272, 11 2007. ISSN 02706474. doi: 10.1523/JNEUROSCI.3437-07.2007.
56. Christian Behrens, Timm Schubert, Silke Haverkamp, Thomas Euler, and Philipp Berens. Connectivity map of bipolar cells and photoreceptors in the mouse retina. *eLife*, 5 (NOVEMBER2016):1–20, 2016. ISSN 2050084X. doi: 10.7554/eLife.20041.
57. Cristopher M. Niell and Michael P. Stryker. Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. *Neuron*, 65(4):472–479, 2010. ISSN 08966273. doi: 10.1016/j.neuron.2010.01.033.
58. Denise M. Piscopo, Rana N. El-Danaf, Andrew D. Huberman, and Cristopher M. Niell. Diverse visual features encoded in mouse lateral geniculate nucleus. *Journal of Neuroscience*, 33(11):4642–4656, 3 2013. ISSN 02706474. doi: 10.1523/JNEUROSCI.5187-12.2013.
59. Shinya Ito, David A. Feldheim, and Alan M. Litke. Segregation of visual response properties in the mouse superior colliculus and their modulation during locomotion. *Journal of Neuroscience*, 37(35):8428–8443, 8 2017. ISSN 15292401. doi: 10.1523/JNEUROSCI.3689-16.2017.
60. Richard H. Masland and Paul R. Martin. The unsolved mystery of vision. *Current Biology*, 17(15):R577–R582, 8 2007. ISSN 09609822. doi: 10.1016/j.cub.2007.05.040.
61. Tom Baden, Frank Schaeffel, and Philipp Berens. Visual Neuroscience: A Retinal Ganglion Cell to Report Image Focus? *Current Biology*, 27(4):R139–R141, 2 2017. ISSN 09609822. doi: 10.1016/j.cub.2016.12.022.
62. Timothy J. Gawne and Thomas T. Norton. An opponent dual-detector spectral drive model of emmetropization. *Vision Research*, 173:7–20, 8 2020. ISSN 18785646. doi: 10.1016/j.visres.2020.03.011.
63. Dominic A. Evans, A. Vanessa Stempel, Ruben Vale, and Tiago Branco. Cognitive Control of Escape Behaviour. *Trends in Cognitive Sciences*, 23(4):334–348, 4 2019. ISSN 1879307X. doi: 10.1016/j.tics.2019.01.012.
64. Y. Zhang, I.-J. Kim, J. R. Sanes, and M. Meister. The most numerous ganglion cell type of the mouse retina is a selective feature detector. *Proceedings of the National Academy of Sciences*, 109(36):E2391–E2398, 2012. ISSN 0027-8424. doi: 10.1073/pnas.1211547109.

65. Dimokratis Karamanlis, Helene Marianne Schreyer, and Tim Gollisch. Retinal Encoding of Natural Scenes. *Annual Review of Vision Science*, 8(1), 9 2022. ISSN 2374-4642. doi: 10.1146/annurev-vision-100820-114239.
66. Samuel A. Ocko, Jack Lindsey, Surya Ganguli, and Stephane Deny. The emergence of multiple retinal cell types through efficient coding of natural movies. *Advances in Neural Information Processing Systems*, 2018-Decem(NeurIPS):9389–9400, 2018. ISSN 10495258.
67. Jack Lindsey, Samuel A. Ocko, Surya Ganguli, and Stephane Deny. A Unified Theory of Early Visual Representations from Retina to Cortex through Anatomically Constrained Deep CNNs. *International Conference on Learning Representations*, pages 1–17, 2019.
68. Matías A Goldin, Baptiste Lefebvre, Samuele Virgili, Mathieu Kim Pham Van Cang, Alexander Ecker, Thierry Mora, Ulisse Ferrari, and Olivier Marre. Context-dependent selectivity to natural images in the retina. *Nature Communications*, 13(1), 2022. ISSN 20411723. doi: 10.1038/s41467-022-33242-8.
69. Kevin L. Briggman and Thomas Euler. Bulk electroporation and population calcium imaging in the adult mammalian retina. *Journal of Neurophysiology*, 105(5):2601–2609, 2011. ISSN 00223077. doi: 10.1152/jn.00722.2010.
70. Thomas Euler, Susanne E Hausselt, David J Margolis, Tobias Breuninger, Xavier Castell, Peter B Detwiler, and Winfried Denk. Eyecup scope-optical recordings of light stimulus-evoked fluorescence signals in the retina. *Pflugers Archiv European Journal of Physiology*, 457(6):1393–1414, 2009. ISSN 00316768. doi: 10.1007/s00424-008-0603-5.
71. Wei Wei, Justin Elstrott, and Marla B. Feller. Two-photon targeted recording of GFP-expressing neurons for light responses and live-cell imaging in the mouse retina. *Nature Protocols*, 5(7):1347–1352, 7 2010. ISSN 17502799. doi: 10.1038/nprot.2010.106.
72. Thomas Euler, Katrin Franke, and Tom Baden. Studying a Light Sensor with Light: Multiphoton Imaging in the Retina. In Espen Hartveit, editor, *Multiphoton Microscopy*, chapter 10, pages 225–250. Humana Press, 2019. ISBN 9781493997015. doi: 10.1002/9781118696736.ch22.
73. Lucas Theis, Philipp Berens, Emmanouil Froudarakis, Jacob Reimer, Miroslav Román Rosón, Tom Baden, Thomas Euler, Andreas S. Tolias, and Matthias Bethge. Benchmarking Spike Rate Inference in Population Calcium Imaging. *Neuron*, 90(3):471–482, 2016. ISSN 10974199. doi: 10.1016/j.neuron.2016.04.014.
74. Katrin Franke, André Maia Chagas, Zhijian Zhao, Maxime J.Y. Zimmermann, Philipp Bartel, Yongrong Qiu, Klaudia P. Szatko, Tom Baden, and Thomas Euler. An arbitrary-spectrum spatial visual stimulator for vision research. *eLife*, 8, 9 2019. ISSN 2050084X. doi: 10.7554/eLife.48779.
75. Cameron Arshadi, Ulrik Günther, Mark Eddison, Kyle I.S. Harrington, and Tiago A. Ferreira. SNT: a unifying toolbox for quantification of neuronal anatomy. *Nature Methods*, 18(4): 374–377, 4 2021. ISSN 15487105. doi: 10.1038/s41592-021-01105-7.
76. Uygar Sümbül, Sen Song, Kyle McCulloch, Michael Becker, Bin Lin, Joshua R. Sanes, Richard H. Masland, and H. Sebastian Seung. A genetic and computational approach to structurally classify neuronal types. *Nature Communications*, 5, 2014. ISSN 20411723. doi: 10.1038/ncomms4512.
77. David A. Klindt, Alexander S. Ecker, Thomas Euler, and Matthias Bethge. Neural system identification for large populations separating "what" and "where". *31st Conference on Neural Information Processing Systems*, (Nips):4–6, 2017. ISSN 10495258.
78. Zhijian Zhao, David A. Klindt, André Maia Chagas, Klaudia P. Szatko, Luke Rogerson, Dario A. Protti, Christian Behrens, Deniz Dalkara, Timm Schubert, Matthias Bethge, Katrin Franke, Philipp Berens, Alexander S. Ecker, and Thomas Euler. The temporal structure of the inner retina at a single glance. *Scientific Reports*, 10(1):4399, 12 2020. ISSN 2045-2322. doi: 10.1038/s41598-020-60214-z.
79. Konstantin-Klemens Lurz, Mohammad Bashiri, Konstantin Willeke, Akshay K Jagadish, Eric Wang, Edgar Y Walker, Santiago A Cadena, Taliah Muhammad, Erick Cobos, Andreas S Tolias, Alexander S Ecker, and Fabian H Sinz. Generalization in data-driven models of primary visual cortex. *Proceedings of the 9th International Conference on Learning Representations*, (9), 2021.
80. Lutz Prechelt. Early stopping - But when? In *Neural Networks:Tricks of the Trade. Lecture Notes in Computer Science*, volume 1524, pages 55–69. 1998. ISBN 9783642352881. doi: 10.1007/978-3-642-35289-8{_}5.
81. Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. *Proceedings of the 3rd International Conference on Learning Representations*, (3):1–15, 2015. ISSN 09252312. doi: http://doi.acm.org.ezproxy.lib.ucf.edu/10.1145/1830483.1830503.
82. Matthias P. Christenson, S. Navid Mousavi, Elie Oriol, Sarah L. Heath, and Rudy Behnia. Exploiting colour space geometry for visual stimulus design across animals. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 377(1862), 10 2022. ISSN 14712970. doi: 10.1098/rstb.2021.0280.
83. Jacob Cohen. *Statistical Power Analysis for the Behavioral Sciences*. Routledge, 5 1988. doi: 10.4324/9780203771587.
84. Jean-Christophe Goulet-Pelletier and Denis Cousineau. A review of effect sizes and their confidence intervals, Part I: The Cohen's d family. *The Quantitative Methods for Psychology*, 14(4):242–265, 12 2018. ISSN 2292-1354. doi: 10.20982/tqmp.14.4.p242.

Supplementary information

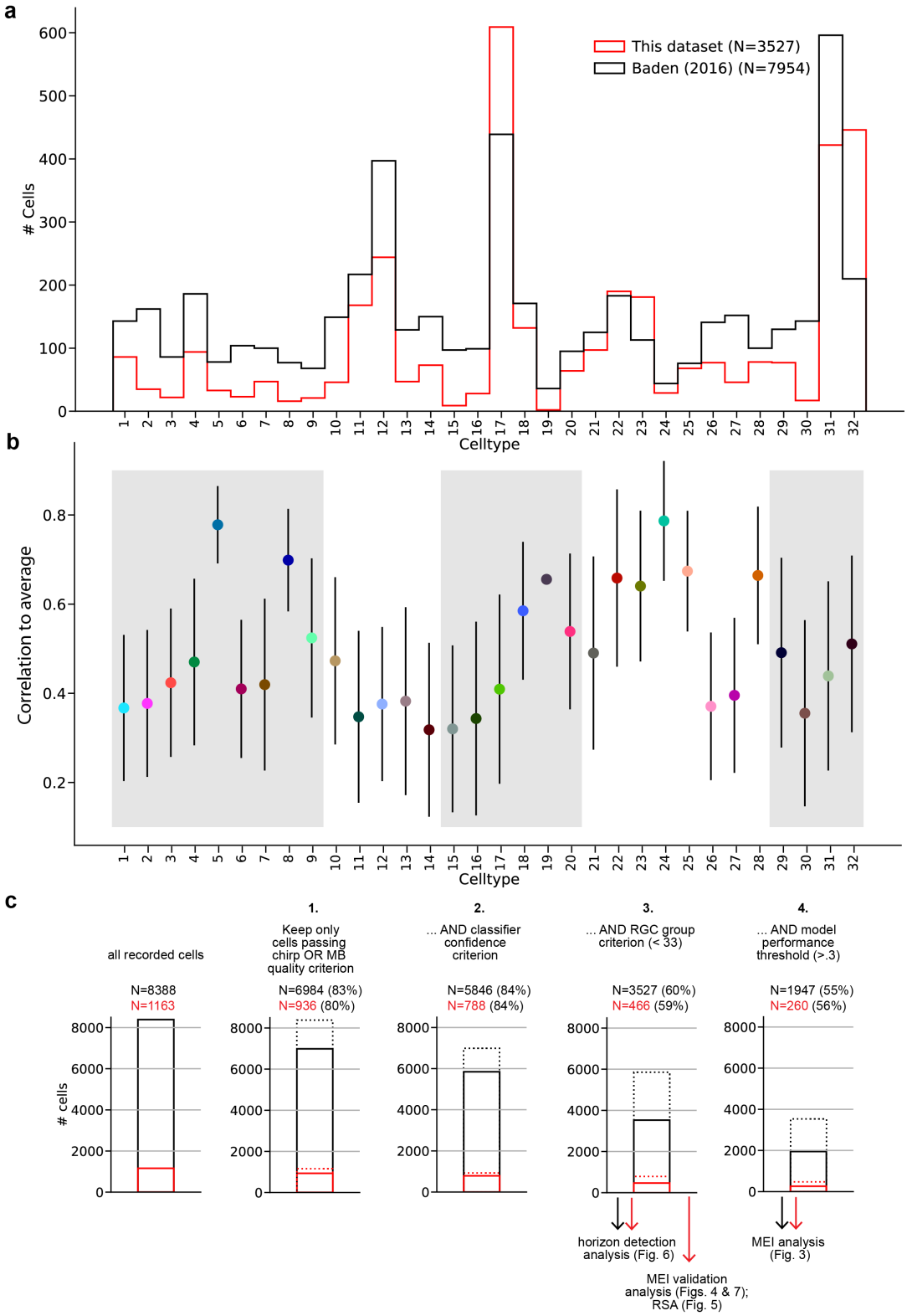

Fig. S1. Supplementary Figure to Figure 1. (a) Distribution across cell types for this dataset, and for the dataset described in Baden et al. (7) which was the basis for our classifier (22). **(b)** Mean $\pm$ SD of model performance, evaluated as correlation between model prediction and RGC response on the 25 s long test sequence, averaged across 3 repetitions of the test sequence, for each cell type. **(c)** Response quality, RGC group assignment and model performance filtering pipeline showing the consecutive steps and the fraction of cells remaining after each step. Black bars and numbers indicate cells from all experiments (i.e. all RGCs for which we recorded chirp, MB and movie responses), red bars and numbers indicate the subset of cells recorded in the MEI validation experiments (i.e. those RGCs for which we *additionally* recorded MEI stimuli responses). Dotted bars indicate the number of cells before the current filtering step. The filtering steps were as follows: **1.** Keep only cells that pass the chirp OR MB quality criterion ($\mathrm{QI}_{MB} > .6$ OR $\mathrm{QI}_{chirp} > .35$). **2.** Keep only cells that the classifier assigns to a group with confidence $\geq .25$. **3.** Keep only cells assigned to a ganglion cell group (groups 1-32; groups 33-46 are amacrine cell groups); **4.** Keep only cells with sufficiently high model performance ($\langle C(\hat{r}^{(n)}, r_i^{(n)})\rangle_i > .3$). All cells passing steps 1-3 were included in the horizon detection analysis (Fig. 6); all cells passing steps 1-4 were included in the MEI analysis (Fig. 3); the "red" cells passing steps 1-4 were included in the MEI validation analysis (Fig. 4). All quality criteria are described in the Methods section.

Fig. S2. Supplementary Figure to Figure 3. Example MEIs for example cell types. Rows in each panel as in Fig. 4a.

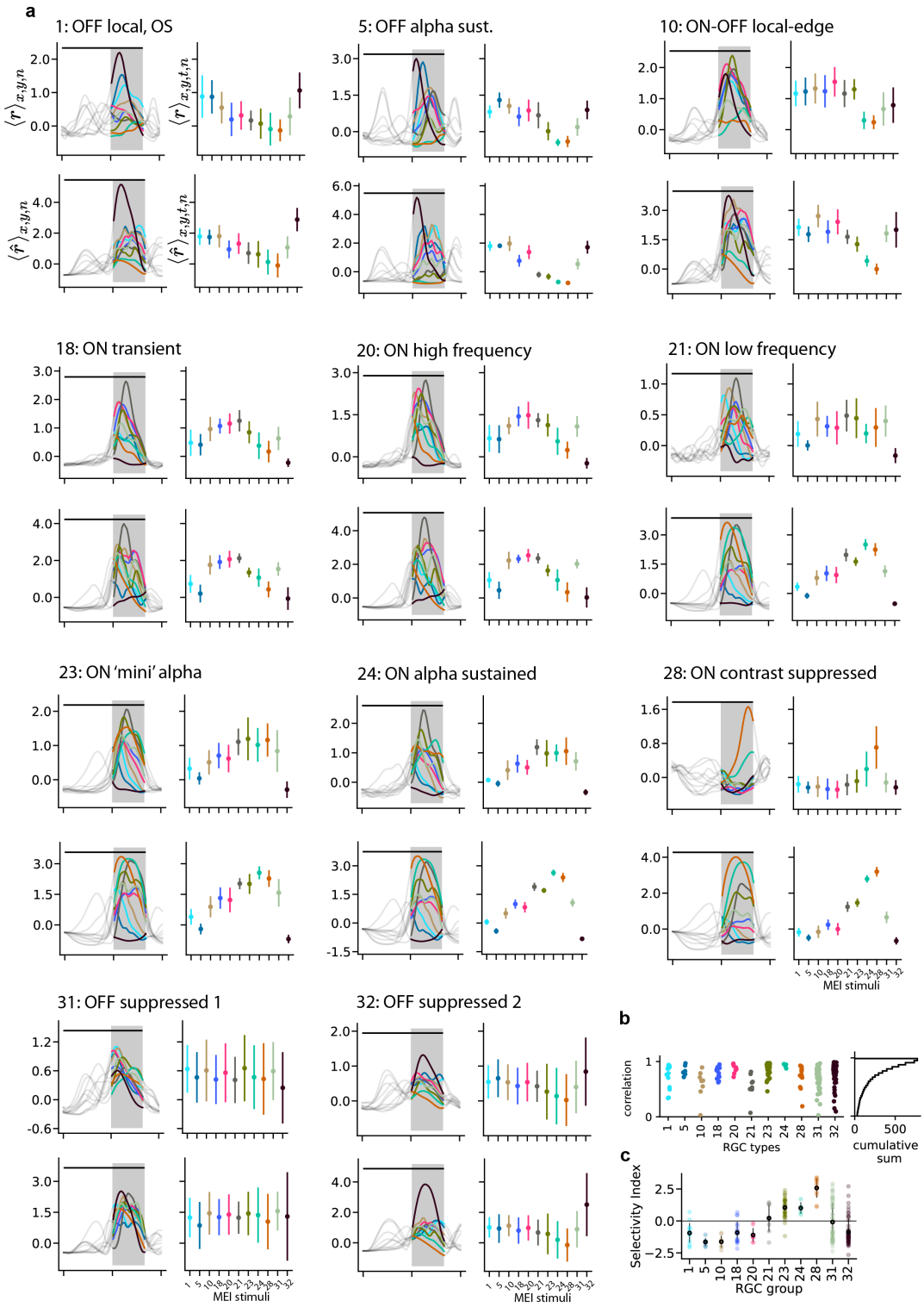

Fig. S3. Supplementary Figure to Figure 4. (a) Recorded (top, r) and predicted (bottom, $\hat{r}$) responses to the 11 different MEIs for all example cell types. Left: responses are averaged across the indicated dimensions x, y, n: different MEI locations (x, y) and RGCs in a group (n); black bar indicates stimulus duration (from 0 to 1.66 s), grey rectangle marks optimisation time window (from 1 to 1.66 s). Right: Responses to different MEIs, additionally averaged across time (t) within the optimisation time window. Error bars indicated SD across cells. **(b)** Correlation between the measured and predicted response magnitudes to the MEI stimuli per example cell type. Cumulative histogram is across all N=788 cells; 50% of cells have a correlation between measured and predicted response magnitude of ≥ 0.8. **(c)** Mean $\pm$ SD of selectivity index (see Methods) for the example cell groups, indicating the difference in response to MEI 28 vs. the average response to all other MEIs in units of standard deviation of the response.

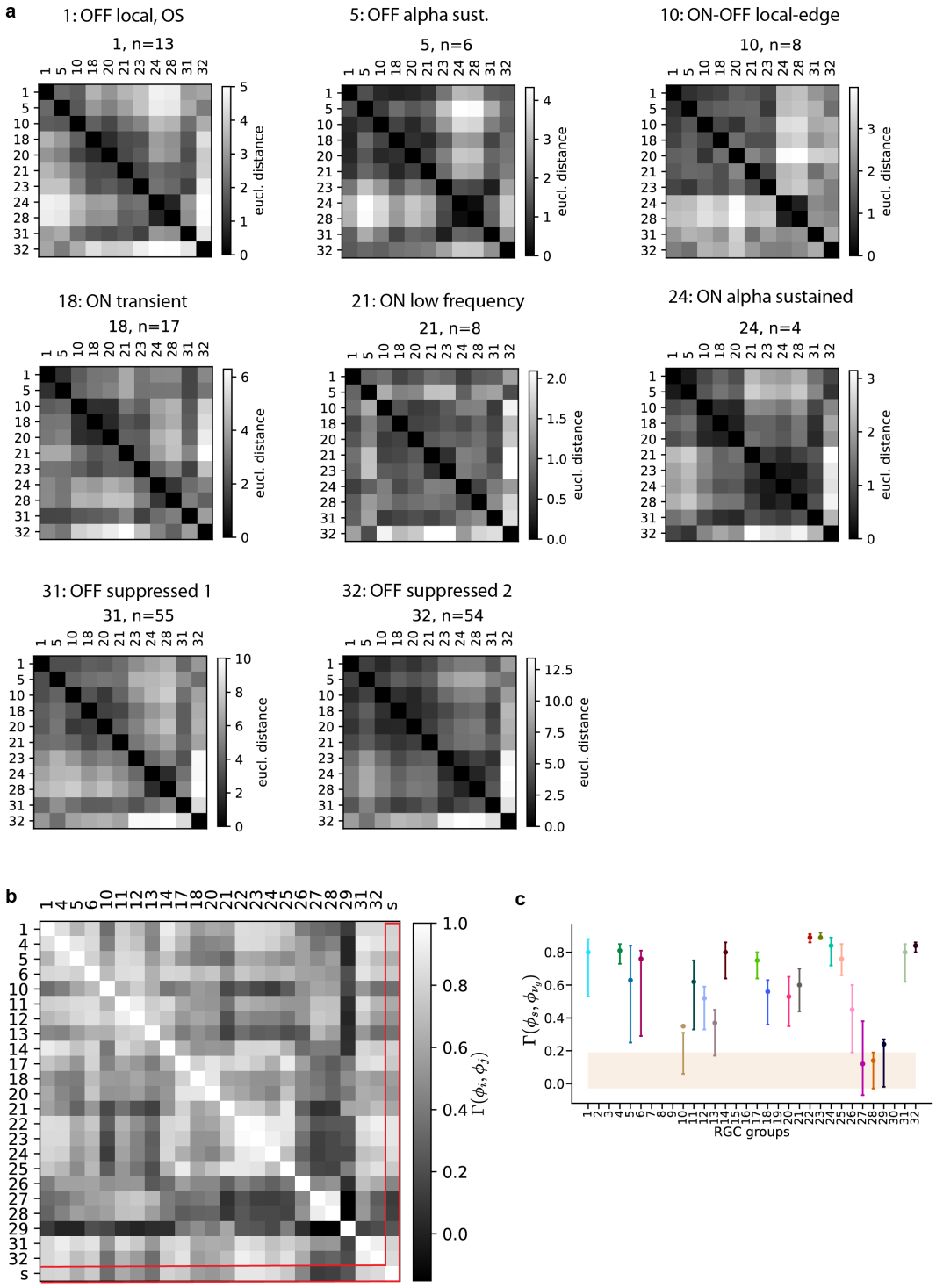

Fig. S4. Supplementary Figure to Figure 5. (a) Representational dissimilarity matrices of MEIs in neural response space of example cell types that are not already shown in the main figure. **(b)** 2nd-order dissimilarity matrix for all RGC groups with at least 4 cells, where entry $[i, j]$ indicates the similarity between representational spaces i and j as quantified by $\Gamma(\phi_i, \phi_j)$ (see Methods). Column highlighted in red corresponds to $\Gamma(\phi_s, \phi_{\nu_g})$ for all g. **(c)** $\Gamma(\phi_s, \phi_{\nu_g})$ for all RGC groups with at least 4 cells. Bars represent bootstrapped 95 % confidence interval; $\mathrm{CI}_{95}(G_{28})$ additionally indicated by orange box. Numbers of cells per group ranged from 4 (G_6, G_{13}, G_{24}, G_{27}) to 67 (G_{17}).

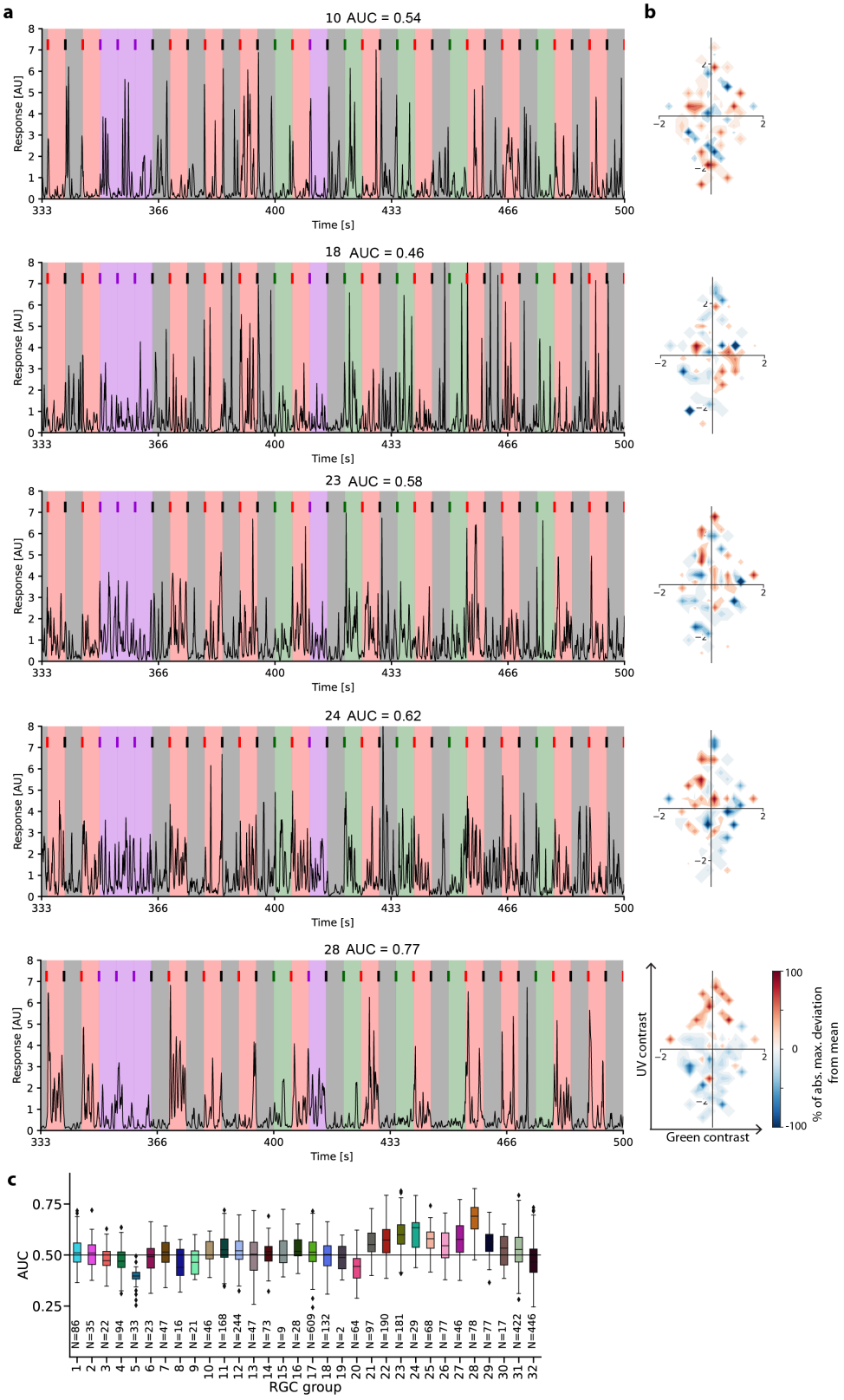

Fig. S5. Supplementary Figure 1 to Figure 6. (a) Traces of example cells of different cell groups (G_{10}, G_{18}, G_{23}, G_{24}, G_{28}) from a single recording field, responding to 33 (of 122) inter-clip transitions. Inter-clip transitions are colour-coded by transition type (red: ground-to-sky, purple: sky-to-sky, green: ground-to-ground, black: sky-to-ground. **(b)** The resulting tuning maps in chromatic contrast space. **(c)** Boxplot of AUC distributions per cell type. The box extends from the first quartile (Q_1) to the third quartile (Q_3) of the data; the line within a box indicates the median. The whiskers extend to the most extreme points still within [$Q_1 - 1.5 \times$ IQR, $Q_3 + 1.5 \times$ IQR], IQR = inter-quartile range. Diamonds indicate points outside this range. All elements of the plot (upper and lower boundaries of the box, median line, whiskers, diamonds) correspond to actual observations in the data. Numbers of RGCs for each type are indicated in the plot.

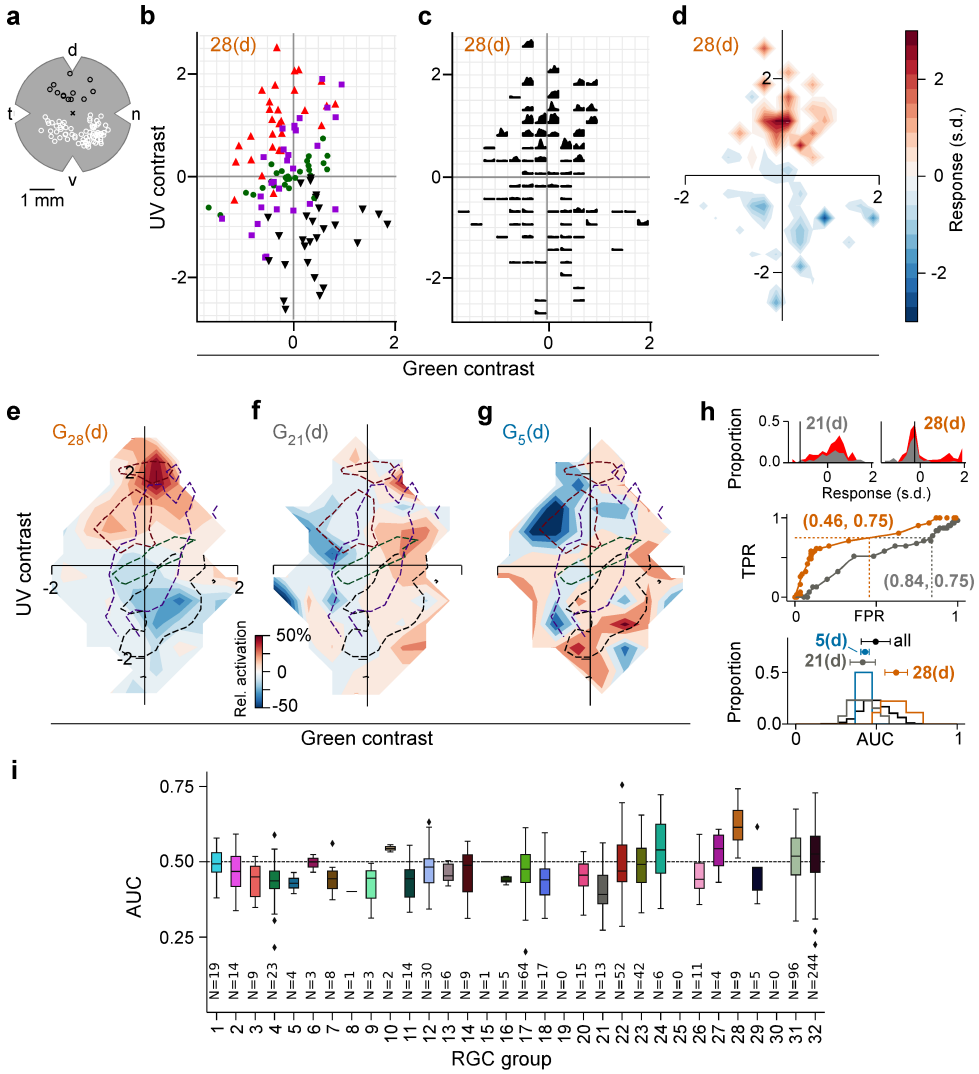

Fig. S6. Supplementary Figure 2 (dorsal data) to Figure 6. Like Fig. 6c-i but for cells recorded in the dorsal retina. **(a)** Illustration of a flat-mounted retina, with recording fields in the dorsal (black circles) and ventral (white circles) retina (cross marks optic disc; d, dorsal; v, ventral; t, temporal; n, nasal). **(b)** Distribution of green and UV contrasts of N=122 inter-clip transitions seen by a dorsal group 28 (G_{28}) RGC, coloured by transition type (red triangle, ground-to-sky; green disk, ground-to-ground; black downward triangle, sky-to-ground; purple square, sky-to-sky). **(c)** Responses of example RGC in the 1 s following a inter-clip transition, averaged across transitions within the bins indicated by the grid. **(d)** Responses shown in (b), transformed into a tuning map by averaging within bins defined by grid shown in (a,b). Responses in (c) and (d) are z-scored ($\mu = 0$, $\sigma = 1$). **(e)** Tuning map of N=9 dorsal G_{28} RGCs, created by averaging the tuning maps of the individual RGCs. **(f)** Same as (d) for N=13 G_{21} RGCs. **(g)** Same as (d) for N=4 G_5 RGCs. **(h)** *Top*: Illustration of ROC analysis for two dorsal RGCs, a G_{21} (left) and a G_{28} (right). For each RGC, responses were binned to all inter-clip transitions, separately for ground-to-sky (red) and all other transitions (grey). *Middle*: Sliding a threshold d across the response range, classifying all transitions with response $> d$ as ground-to-sky, and registering the false-positive-rate (FPR) and true-positive-rate (TPR) for each threshold yields an ROC curve (middle). Numbers in brackets indicate (FPR, TPR) at the threshold indicated by black vertical line in histogram plots. *Bottom*: We evaluated performance for each cell as the area under the ROC curve (AUC), and plotted the distribution across AUC values for all cells (black), for G_5 (blue), for G_{21} (grey), and for G_{28} (orange). Among the dorsal RGCs, G_{28} RGCs achieved the highest AUC on average (mean $\pm$ SD AUC, G_{28} (N=9 cells): 0.62 ± 0.07; all other groups (N=720): 0.49 ± 0.09, ΔAUC $= 0.13$, bootstrapped 95% confidence interval $\text{CI}_{95} = [0.08, 0.18]$, Cohen's $d = 1.45$, two-sample permutation test G_{28} vs. all other groups (see Methods): $p = 0$ with 100,000 permutations; next-best performing G_{24} (N=6): 0.54 ± 0.12, ΔAUC $= 0.08$, bootstrapped 95% confidence interval $\text{CI}_{95} = [0.01, 0.18]$, Cohen's $d = 0.87$; two-sided t-test G_{28} vs. G_{24}: $p = .15$ with 100,000 permutations (not significant)). AUC mean $\pm$ SD indicated as dots and horizontal lines above histograms. **(i)** Boxplot of AUC distributions per cell type. The box extends from the first quartile (Q_1) to the third quartile (Q_3) of the data; the line within a box indicates the median. The whiskers extend to the most extreme points still within [$Q_1 - 1.5 \times$ IQR, $Q_3 + 1.5 \times$ IQR], IQR = inter-quartile range. Diamonds indicate points outside this range. All elements of the plot (upper and lower boundaries of the box, median line, whiskers, diamonds) correspond to actual observations in the data. Numbers of RGCs for each type are indicated in the plot.