PREDICTION ERROR NEURONS IN MOUSE CORTEX ARE MOLECULARLY TARGETABLE CELL TYPES

Sean M. O'Toole[1], Hassana K. Oyibo[1], Georg B. Keller[1,2,*]

[1]*Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland*

[2]*Faculty of Natural Sciences, University of Basel, Basel, Switzerland*

[*]*To whom correspondence should be addressed.*

ABSTRACT

Predictive processing postulates the existence of prediction error neurons in cortex. Functionally, both negative and positive prediction error neurons have been identified in layer 2/3 of visual cortex, but whether they correspond to transcriptionally defined subpopulations is unclear. Here we used the activity-dependent, photoconvertible marker CaMPARI2 to tag neurons in layer 2/3 of visual cortex during stimuli and behaviors designed to trigger prediction errors. We performed single-cell RNA-sequencing on these populations and found that previously annotated *Adamts2* and *Rrad* layer 2/3 cell types were enriched when photolabeling for negative or positive prediction error responses respectively. Finally, we validated these results functionally by designing artificial promoters for use in AAV vectors to express genetically encoded calcium indicators. Thus, positive and negative prediction error responses mapped onto transcriptionally distinct cell types in layer 2/3 that can be targeted using AAV vectors.

**** Dear reader, please note this manuscript is formatted in a standard submission format, and all statistical information is provided in Table S1. ****

INTRODUCTION

Predictive processing is a computational framework that has been proposed to explain cortical function (Bastos et al., 2012; Clark, 2013; Jordan and Rumelhart, 1992; Keller and Mrsic-Flogel, 2018; Koster-Hale and Saxe, 2013; Rao and Ballard, 1999) as well as the symptoms of cortical dysfunction (Griffin and Fletcher, 2017; Palmer et al., 2015). It postulates that cortex maintains an internal representation of the external world by integrating over prediction errors computed between sensory input and predictions of that input. Predictions originate in other cortical areas and are exchanged via long range cortico-cortical

projections (Garner and Keller, 2022; Leinweber et al., 2017). Prediction errors are thought to be computed in a dedicated population of neurons referred to as prediction error neurons. These come in two variants that are referred to as positive and negative prediction error neurons. A positive prediction error signals the unpredicted appearance of a sensory stimulus, and a negative prediction error signals an absence or cessation of a predicted sensory stimulus. While the computational ideas of predictive processing are well established, physiological evidence for the different cell types postulated to exist in cortex began to emerge only in the past decade. A central tenet in the experimental approach to testing the framework has been the fact that one of the best predictors of sensory input is self-generated movement. Almost all change to visual input is the direct consequence of self-generated eye, head, or body movements. Using movement as an experimental proxy for movement-related predictions, a mismatch between movement and resultant sensory feedback can be used to probe for sensorimotor prediction errors. In sensory cortices, sensorimotor prediction error responses were first described in human primary visual cortex (V1) (Stanley and Miall, 2007), and later also found across a range of species and cortical areas (Audette et al., 2021; Ayaz et al., 2019; Eliades and Wang, 2008; Heindorf et al., 2018; Keller and Hahnloser, 2009). At the same time, evidence began to emerge that in layer 2/3 (L2/3) of cortex, positive and negative prediction error neurons constitute functionally distinct sets of neurons (Attinger et al., 2017; Audette et al., 2021; Jordan and Keller, 2020; Keller et al., 2012; Zmarz and Keller, 2016). PPE neurons are thought to be excited by sensory input and inhibited by movement related predictions, while NPE neurons are driven by movement related predictions and inhibited by sensory input.

In L2/3 of mouse V1 three types of excitatory neurons have been identified, each named after one of the differentially expressed genes: *Adamts2*, *Agmat*, and *Rrad* (Tasic et al., 2018). These L2/3 excitatory cell types correspond to layer 2 (L2) and layer 3 (L3) transcriptional classes of excitatory neurons found in human cortex (Hodge et al., 2019). In somatosensory cortex, it has been demonstrated that the L2/3 *Rrad* type is largely driven by feedforward sensory input (Condylis et al., 2022), and in inhibitory neurons it has recently been found that a specific transcriptomic axis predicts how a neuron is functionally modulated by locomotion (Bugeon et al., 2022). Thus, we speculated that positive and negative prediction error neurons can be mapped onto transcriptionally defined neuron types within L2/3 of V1. Experimental and translational progress will hinge on determining whether these computationally identified neuron types can be selectively targeted for transgene expression; whether this is the case is still unclear.

RESULTS

Our strategy to search for transcriptional markers of prediction error neurons in L2/3 of V1 was to first tag neurons by photoconversion that exhibit prediction error responses with a fluorescent label in vivo dissociate the tissue, sort the cells by fluorescence, and single-cell sequence different photoconversion groups separately (**Figure 1A**). To do this, we used CaMPARI2, an engineered fluorescent protein that can be photoconverted from green to red by illumination with violet light only when intracellular calcium levels are high (Moeyaert et al., 2018), which has been demonstrated to distinguish neurons of differing activity levels in vivo (Trojanowski et al., 2021). By timing photoconversion light to coincide with a particular stimulus, one can photo-tag neurons that exhibit calcium responses to that stimulus. All mice first received injections of AAV2/1-hSyn-NES-his-CaMPARI2-WPRE and a cranial window implant bilaterally in V1. Once CaMPARI2 was expressed (21 days after injection), all mice underwent the same visuomotor paradigm in a virtual reality environment. Mice were head-fixed on a spherical treadmill surrounded by a toroidal screen (**Figure 1A**). Over the course of 90 minutes, mice first experienced a closed-loop session in which locomotion was coupled to visual flow feedback presented on the toroidal screen. During this session, to evoke negative prediction error signals, we introduced visuomotor mismatches in the form of brief (1 s) halts of visual flow at random times. The closed-loop session was followed by a dark session. Finally, to evoke positive prediction error prediction error signals, we presented full-field drifting gratings with randomized onsets, durations, and orientations. Throughout all sessions, mice were free to run on the treadmill. In the first group of mice, we directed a violet (405 nm) laser at V1 for 1 s coincident with visuomotor mismatches. In the second group of mice, the laser was directed at V1 triggered by running onsets in the dark session. In the third group of mice, the laser was directed at V1 coincident with the onset of randomized (in orientation, direction, duration, and interstimulus interval) grating presentations when mice were stationary (**Figure 1B**).

This photoconversion paradigm reliably converted a subset of CaMPARI2 expressing neurons from green to red (**Figure 1C**). To validate the targeting of photoconversion to stimulus responsive neurons, we conducted a separate set of experiments in which we measured functional responses with CaMPARI2 as well as the red to green fluorescence ratio after photoconversion. The green fluorescence of CaMPARI2 decreases when it binds to calcium, and thus CaMPARI2 can also be used as a calcium indicator. We measured calcium responses in individual neurons to visuomotor mismatches (**Figure 1D**), running onsets, and grating onsets after photoconversion during visuomotor mismatch events. We compared

the level of photoconversion to the calcium responses to visuomotor mismatch in the same neurons (**Figure 1E**). There was a significant negative correlation ($r = -0.25$; $p = 0.02$) between calcium response to visuomotor mismatch and photoconversion. Note, the correlation is negative because the decrease in fluorescence corresponds to an increase in calcium levels. We found no evidence of a correlation between photoconversion and calcium responses to grating ($r = 0.075$, $p = 0.63$) or running onsets ($r = 0.02$, $p = 0.087$) (**Figure 1F**). Thus, photoconversion during visuomotor mismatch preferentially tagged visuomotor mismatch responsive neurons.

To test for differential enrichment of cell types during visuomotor behaviors we photoconverted with one of the three different stimulus paradigms (**Figure 1B**). We then sacrificed the mice, removed the brain and microdissected layer 1 (L1) through layer 4 (L4) of the photoconverted regions of V1. The tissue was dissociated into a single-cell suspension and subsequently sorted, using fluorescence-activated cell sorting (FACS), into groups of either low, intermediate, or high photoconversion (**Figure S1**). All sorted cellular particles in the three photoconversion groups were then used for single-cell RNA sequencing.

Single-cell RNA-sequencing was performed using the 10x Genomics Chromium platform. In total, we sequenced 65 002 cells across all conditions. All samples were sequenced to an average read depth of 264 million reads, which yielded a median of 3376 unique molecular identifiers (UMIs) per cell (see Methods). Given the comparatively low read depth of our data, we used a bioinformatic strategy to leverage the power published reference dataset from V1 with higher read depth (Tasic et al., 2018) for cell classification. We aligned the two datasets against each other (Welch et al., 2019) and assigned group identities to all cells using a weighted nearest neighbor algorithm (see Methods; **Figure 2A**). To confirm appropriate group assignment, we quantified the expression of marker genes for different cell types in the different groups (**Figure S2**). The distribution of assigned cell type identity for layer 2/3 types largely conformed to the expression pattern of the marker genes (**Figures 2B, 2C and 2D**). We restricted all further analysis to the group of L2/3 excitatory neurons. Within this group we assigned each cell to one of the three L2/3 excitatory cell types previously described: *Adamts2*, *Agmat* or *Rrad* (Tasic et al., 2018). To confirm the accuracy of the cell type assignment, we examined genes that are differentially expressed between these three types and found their expression patterns to be comparable (**Figures S2D and S2E**).

We hypothesized that the L2/3 excitatory cell types might have different functional response properties. In the case of the L2/3 *Rrad* type, it has been found in barrel cortex of mice that *Baz1a* positive neurons (one of the marker genes of the *Rrad* type) are highly responsive to sensory stimuli (Condylis et al., 2022). To test whether each of these cell types had distinct functional response properties, we investigated whether there was a differential enrichment of the three L2/3 excitatory cell types (*Adamts2*, *Agmat* or *Rrad*) across stimulus paradigms (visuomotor mismatch, running onsets in darkness, or grating onsets) and photoconversion state (low, medium, or high). We found that the L2/3 *Adamts2* type was significantly overrepresented in the high visuomotor mismatch photoconversion group (**Figures 2E and 2F**). The *Agmat* type was underrepresented in both high visuomotor mismatch and low grating photoconversion groups (**Figure 2G**), and the L2/3 *Rrad* type was significantly overrepresented in the high grating photoconversion group (**Figure 2H**). These results suggest that on average visuomotor mismatch responsive neurons are enriched in the L2/3 *Adamts2* cell type, and that, consistent with previous findings (Condylis et al., 2022), neurons responding to feed-forward sensory stimuli are enriched in the L2/3 *Rrad* cell type. The remaining L2/3 *Agmat* type appears to represent an intermediate population or potentially a grouping with an as yet unidentified function.

To confirm the functional identity of these cell types, we designed artificial promoters (AP) for use in AAV vectors to target the expression of a genetically encoded calcium indicator to each one of these cell types using a previously described strategy (Jüttner et al., 2019). Artificial promoters were designed based on 2 kb segments surrounding the transcription start site of selected marker genes corresponding to the three cell types (**Figure 3A**). To test for expression, we first used these promoters to drive GFP expression and examined the expression profiles visually. The successful candidate for the L2/3 *Adamts2* cell type was a promoter designed off of the *Adamts2* gene (AP.Adamts2.1), for the L2/3 *Agmat* cell type, one based on the *Agmat* gene (AP.Agmat.1), and for the L2/3 *Rrad* cell type, one based on the *Baz1a* gene (AP.Baz1a.1). We initially tried labelling the L2/3 *Rrad* cell type with an AP based on the Rrad gene, however, expression was too dim and extremely sparse (**Figure S3**)**.** Each of these three artificial promoters resulted in a different expression profile in cortex (**Figures 3B, 3C and 3D**). AP.Adamts2.1 drove expression in superficial L2/3, in superficial L5, and in L6 (**Figures 3B and 3F**). AP.Agmat.1 drove expression throughout L2/3, and exhibited very little expression in other layers (**Figures 3C and 3F**). AP.Baz1a.1 drove expression in L2/3, L4, and L5 (**Figures 3D and 3F**). For each of these promoters we sequenced FACS sorted cells with bulk RNA-sequencing to confirm marker gene expression in the targeted neurons (**Figure S4**). Also, for AP.Baz1a.1 and AP.Adamts2.1, we noticed a large number of

neurons labelled in L1, indicating that these promoters also induce transgene expression in L1. Using RNA-sequencing and histology experiments, we confirmed that indeed a subset of the labelled cells were inhibitory neurons. To prevent expression in inhibitory neurons, we designed Cre dependent variants of the AP.Baz1a.1 and AP.Adamts2.1 vectors and used them in an Emx1-Cre mouse line to restrict expression to excitatory cortical neurons.

Using these promoters, we could then target expression of a GCaMP variant to the three L2/3 excitatory cell types separately to confirm their functional specificity. We did this in three groups of mice that received injections of either AAV2/1-AP.Adamts2.1-DIO-GcaMP8s, AAV2/1-AP.Agmat.1-GcaMP7f, or AAV2/1-AP.Baz1a.1-DIO-GcaMP8s in V1. To measure functional responses in these neurons using two-photon calcium imaging, mice underwent the same visuomotor paradigm (closed-loop, running in darkness, randomized grating presentations) as during photoconversion experiments with the exception that the mice were also exposed to an open-loop session during which the visual flow the mice had self-generated in the preceding closed-loop session was replayed. Thus, we asked whether the response profiles of these cell types behaved in a way similar to the observations found in the single-cell sequencing data analysis.

The AP.Adamts2.1 neurons responded robustly to visuomotor mismatch as well as to running onsets (in darkness, closed-loop or open-loop sessions) and decreased their fluorescence following grating onsets (**Figures 4A, 4B and 4C**). While AP.Agmat.1 neurons responded primarily to running onsets, AP.Baz1a.1 neurons exhibited a robust grating onset response. The visuomotor mismatch responses were significantly larger in the population of AP.Adamts2.1 neurons than they were in the AP.Baz1a.1 and AP.Agmat.1 populations (**Figure 4A**). Conversely, the grating onset response was stronger in the AP.Baz1a.1 population than in the other two labelled populations (**Figure 4C**). Consistent with an opposing influence of visual flow and motor related input on negative prediction error and positive prediction error neurons (Attinger et al., 2017; Jordan and Keller, 2020), we found that the activity of AP.Baz1a.1 neurons was positively correlated with visual flow speed, while that of AP.Adamts2.1 neurons was negatively correlated (**Figure 4D**). For correlations with running speed, this was reversed in that AP.Adamts2.1 neurons exhibited the strongest correlations, and AP.Baz1a.1 neurons the weakest correlations (**Figure 4E**). The activity of AP.Agmat.1 neurons was uncorrelated with visual flow speed, and exhibited an intermediate correlation with running speed.

DISCUSSION

In our interpretation, we have made the assumption that visually responsive excitatory neurons in L2/3 are positive prediction error neurons, and that visuomotor mismatch responsive neurons are negative prediction error neurons. It should be noted, that exhibiting visual or visuomotor mismatch responses does not mean a neuron functions as a positive prediction error or negative prediction error neuron. However, our interpretation is based on previous work demonstrating that in L2/3, visual and visuomotor mismatch responsive neurons function as comparators between top-down and bottom-up input (Jordan and Keller, 2020). Based on our findings, the L2/3 *Adamts2* population is enriched for negative prediction error neurons, and that of L2/3 *Rrad* is enriched for positive prediction error neurons. Interestingly, we find that the third type of excitatory L2/3 neuron, the L2/3 *Agmat* cell type, does not exhibit responses consistent with visuomotor prediction errors. One possibility is that this population functions to maintain an internal representation, by integrating over negative prediction error and positive prediction error responses, a function that is typically speculated to be performed by L5 neurons (Heindorf and Keller, 2022). Another possibility is that these neurons compute prediction errors that are not visuomotor. The populations of neurons that respond e.g. to spatio-visual prediction errors, is separate from that responding to visuomotor prediction errors (Fiser et al., 2016). Lastly, it is possible that these neurons perform a function outside of the canonical circuit for predictive processing. Identifying how the three different types of L2/3 excitatory neurons interact with each other and the rest of the cortical circuit, will likely be critical for making progress in our understanding of the canonical cortical circuit.

Our results demonstrate that prediction error responsive neurons have distinct transcriptional identities and can be targeted virally. Given that predictive processing has become one of the leading theories of cortical function (Bastos et al., 2012; Clark, 2013; Jordan and Rumelhart, 1992; Keller and Mrsic-Flogel, 2018; Koster-Hale and Saxe, 2013; Rao and Ballard, 1999) with the capacity to explain both neurophysiological results, as well as dysfunctions of cortex that are thought to result in diseases like autism (Lawson et al., 2014, 2017; Sinha et al., 2014) and schizophrenia (Corlett et al., 2009, 2009; Fletcher and Frith, 2009; Frith et al., 2000), it will be essential to have specific genetic access to the functionally identified neuronal subpopulations postulated by this framework. Experimental tests of different variants of predictive processing circuit models will hinge on cell type specific manipulations and recordings of neuronal activity (Keller and Mrsic-Flogel, 2018). In parallel, it is likely that any

208 treatment of disorders involving prediction error computations, will require cell type specific
209 interventions. The identification of *Adamts2* as a marker of neurons with negative prediction error
210 responses and that of *Baz1a* as a marker of neurons with positive prediction error responses in L2/3 of
211 mouse cortex is a first step in this direction.

212

FIGURES

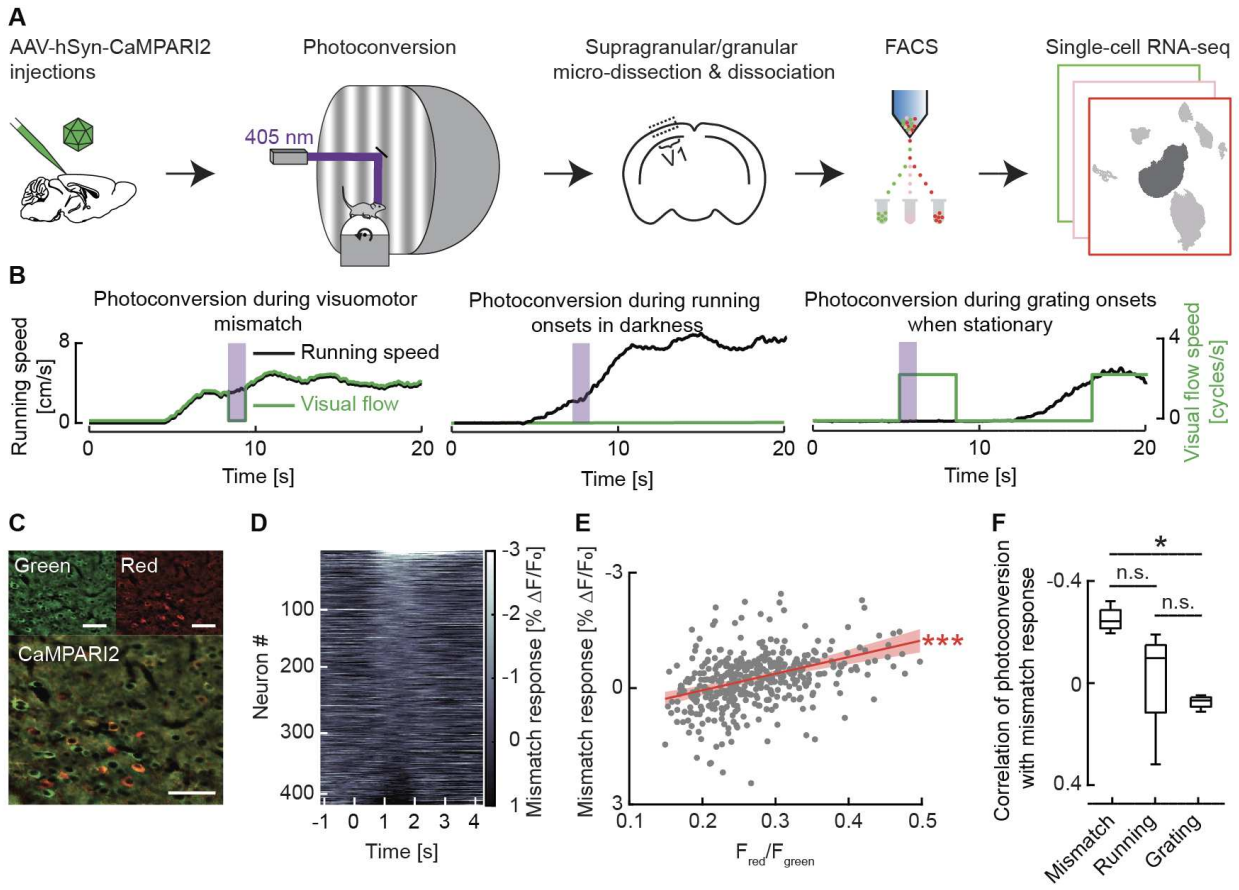


Figure 1. Tagging of functionally identified neurons using CaMPARI2.

(**A**) Schematic of approach: C57BL/6 mice that express CaMAPRI2 in V1 were head-fixed on a spherical treadmill in a virtual reality environment. A 405 nm laser was directed at V1 through a cranial window to trigger photoconversion of CaMPARI2. Tissue from granular and supragranular V1 was dissociated, sorted into different photoconversion groups using FACS (low photoconversion is green, intermediate photoconversion is pink, high photoconversion is red), and single-cell sequenced with the 10x genomics platform.

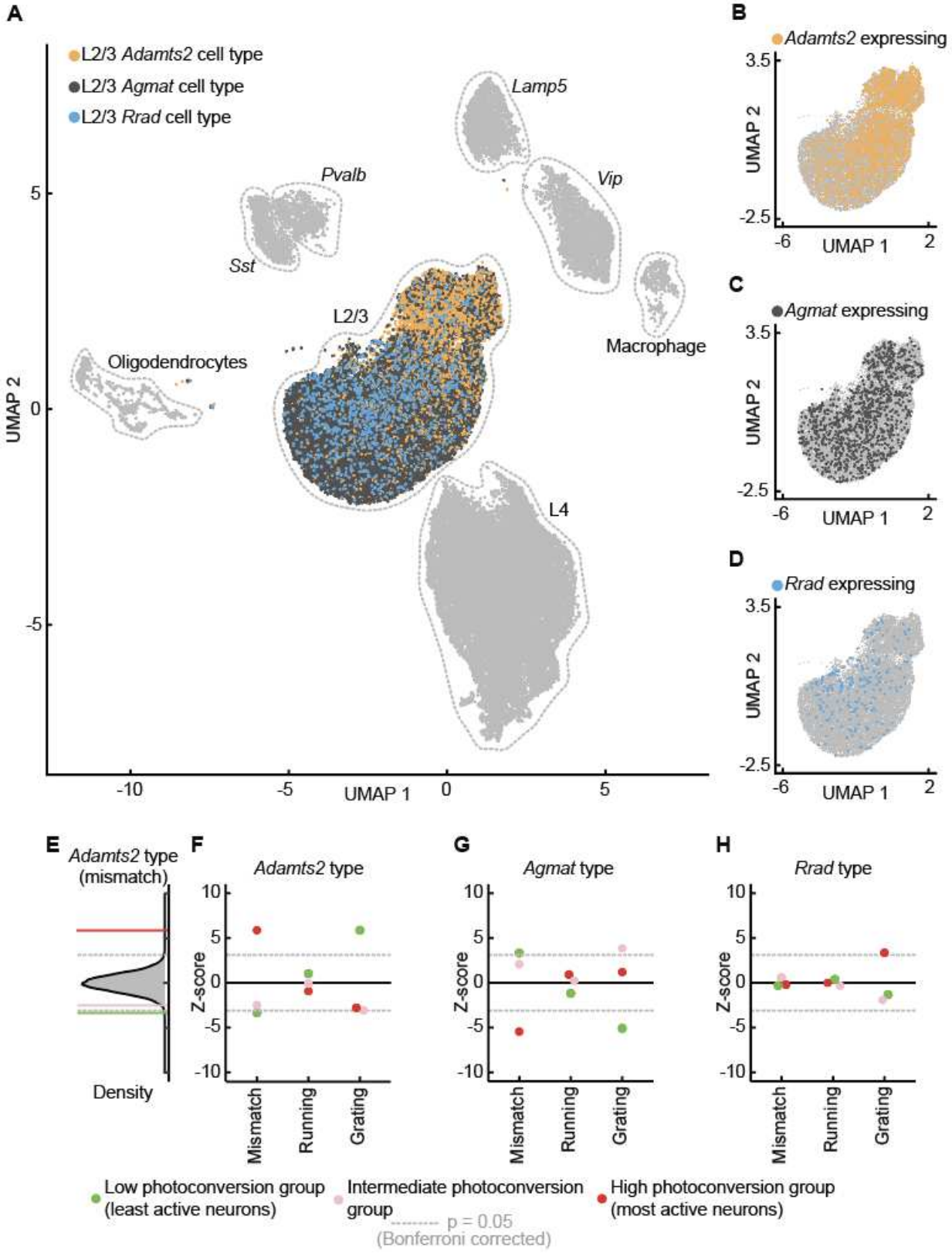
(**B**) Three separate groups of mice were used. In the first group (left), photoconversion (purple shading) was triggered on visuomotor mismatch events. In the second group of mice (middle), photoconversion was triggered on running onsets in darkness. In the third group of mice (right), photoconversion was triggered on onsets of full-field drifting grating stimuli while the mouse was stationary.

(**C**) After photoconversion during visuomotor mismatch, only a subset of neurons in L2/3 that express CaMPARI2 (top left panel) were positive for the red CaMPARI2 variant (top right subpanel). Bottom image is a merge of the green and red images on top. Scale bar in all panels 25 μm.

(**D**) Average CaMPARI2 responses in the green channel to visuomotor mismatch for 417 neurons recorded in 4 mice, sorted by strength of visuomotor mismatch response. Note, increases in calcium result in a decrease in fluorescence.

(**E**) Scatter plot of the average fluorescence response to visuomotor mismatch plotted against the level of photoconversion per neuron, in mice that underwent photoconversion during visuomotor mismatch. The red line is an estimate of a fit to the data using linear regression ($r = -0.043$, intercept = 0.0091), asterisks indicate the significance of the fit. Shading indicates the 95 % confidence bounds for the linear fit. Here and elsewhere, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, n.s.: not significant.

(**F**) Correlation of the population vector of the visuomotor mismatch, running onset, and grating response with photoconversion during mismatch for all mice. When photoconverting during visuomotor mismatch, photoconversion level and calcium response during visuomotor mismatch events were inversely correlated. This was not the case for either running or grating onset responses.



241

Figure 2. Single-cell sequencing of CaMPARI2 labelled cells revealed differential enrichment of L2/3 excitatory cell types in different visuomotor conditions.

(**A**) Uniform manifold approximation and projection (UMAP) of neurons isolated from the superficial cortical layers (L1, L2/3 and L4), based on the iNMF vectors calculated from a shared projection of our data and a reference dataset (Tasic et al., 2018). Major cell groups are labelled, and inferred cell type identity of L2/3 cells are color coded.

(**B**) The L2/3 cluster with cells expressing one or more UMIs corresponding to *Adamts2* highlighted.

(**C**) As in **B**, but for *Agmat*.

(**D**) As in **B**, but for *Rrad*.

(**E**) Enrichment of the L2/3 *Adamts2* cell type in the high (purple), intermediate (pink), and low (green) photoconversion groups for the visuomotor mismatch labelling experiments, relative to the distribution observed in random samples (z-score distribution shown in black, dashed gray lines are the Bonferroni corrected p=0.05 levels). The L2/3 *Adamts2* cell type is significantly enriched in the high mismatch visuomotor photoconversion group.

(**F**) Enrichment for the L2/3 *Adamts2* cell type across photoconversion and labelling conditions. Z-score values are displayed for visuomotor mismatch (mismatch; as in **E**), running onset (running), and grating onset when stationary (grating) labelling experiments.

(**G**) As in **F**, but for the L2/3 *Agmat* cell type.

(**H**) As in **F**, but for the L2/3 *Rrad* cell type.

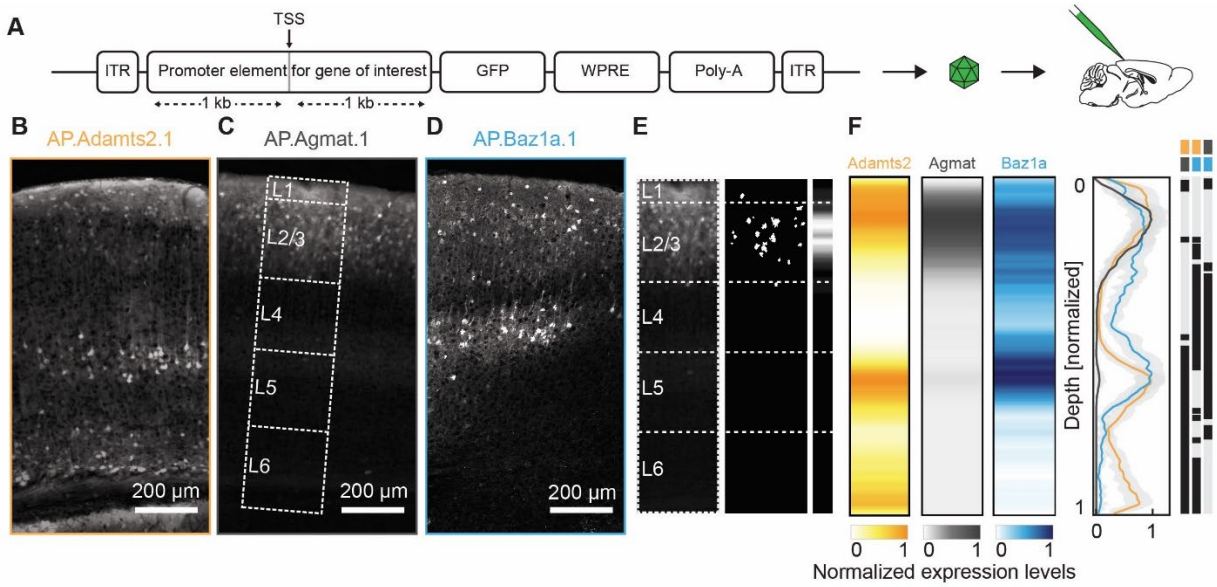


Figure 3. Artificial promoters exhibit differential expression patterns in cortex.

(**A**) Artificial promoter design strategy. The 2 kb segment of the genomic sequence flanking the transcription start site (TSS), for the selected marker genes, was placed into an AAV vector to drive GFP expression.

(**B**) Example expression pattern from an AAV2/1-AP.Adamts2.1-GFP injection in V1.

(**C**) As in **B**, but for AAV2/1-AP.Agmat.1-GFP.

(**D**) As in **B**, but for AAV2/1-AP.Baz1a.1-GFP.

(**E**) Histology images were converted to expression profiles as a function of cortical depth (see Methods). The example section is shown as an inset in **C**.

(**F**) Left: Average expression profiles as a function of cortical depth for the three different artificial promoters. Right: Comparison between the expression profiles. Significant differences are marked as black bars to the right of the plot, and the respective comparisons are labelled by the pair of line segments at the top.

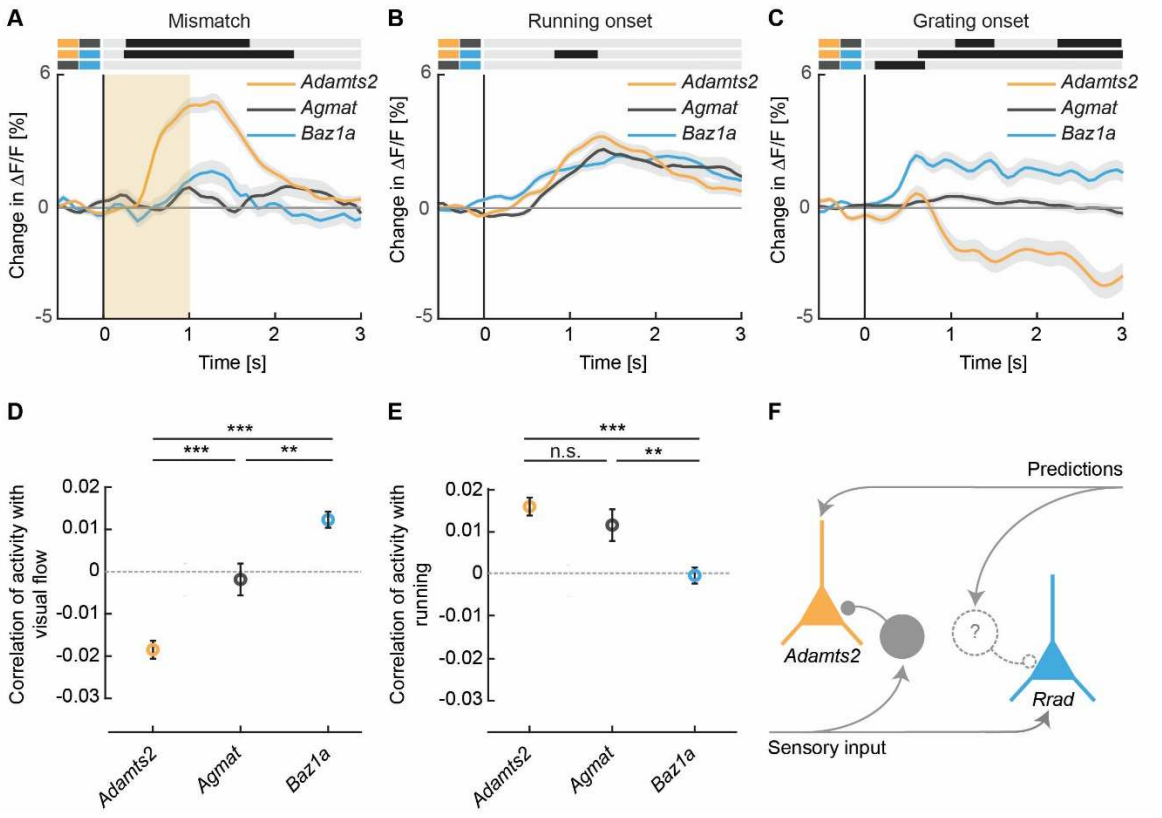


Figure 4. L2/3 neurons targeted by artificial promoters exhibited differential visuomotor responses.

(**A**) Average response to visuomotor mismatch in the AP.Adamts2.1 labelled population (orange) was stronger than in the AP.Agmat.1 (dark gray), or the AP.Baz1a.1 (blue) populations. Orange shading indicates the duration of visuomotor mismatch; grey shading indicates the SEM. Response curves were compared bin-by-bin ($p < 0.05$ indicated by black bars above traces). Each of the three comparisons is denoted by a pair of line segments to the left, corresponding in color to the data being plotted.

(**B**) As in **A**, but for the response to running onsets.

(**C**) As in **A**, but for the response to the onset of randomized grating.

(**D**) Average correlation of neuronal activity with visual flow during open-loop sessions for L2/3 *Adamts2*, *Agmat*, and *Baz1a* neurons. Error bars are SEM over neurons. Here and elsewhere, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$, n.s.: not significant.

(**E**) As in **D**, but for average correlation of neuronal activity with running during open-loop sessions. Error bars are SEM over neurons.

(**F**) Schematic of a model circuit that would explain the response properties of *Adamts2* and *Rrad* L2/3 excitatory neuron types. Triangles, indicate L2/3 excitatory neuron types; circles, indicate inhibitory interneurons. L2/3 *Adamts2* neurons respond robustly to visuomotor mismatch, and their activity is correlated with running, and negatively correlated with visual flow. This is consistent with the functional signature of a negative prediction error type. L2/3 *Rrad* neurons have visually driven responses consistent with a positive prediction error type. Previous work has demonstrated that negative

297 prediction error neurons are bottom-up inhibited (Attinger et al., 2017), while a potential source of top-
298 down driven inhibition for the positive prediction error neuron remains speculative.

SUPPLEMENTARY FIGURES

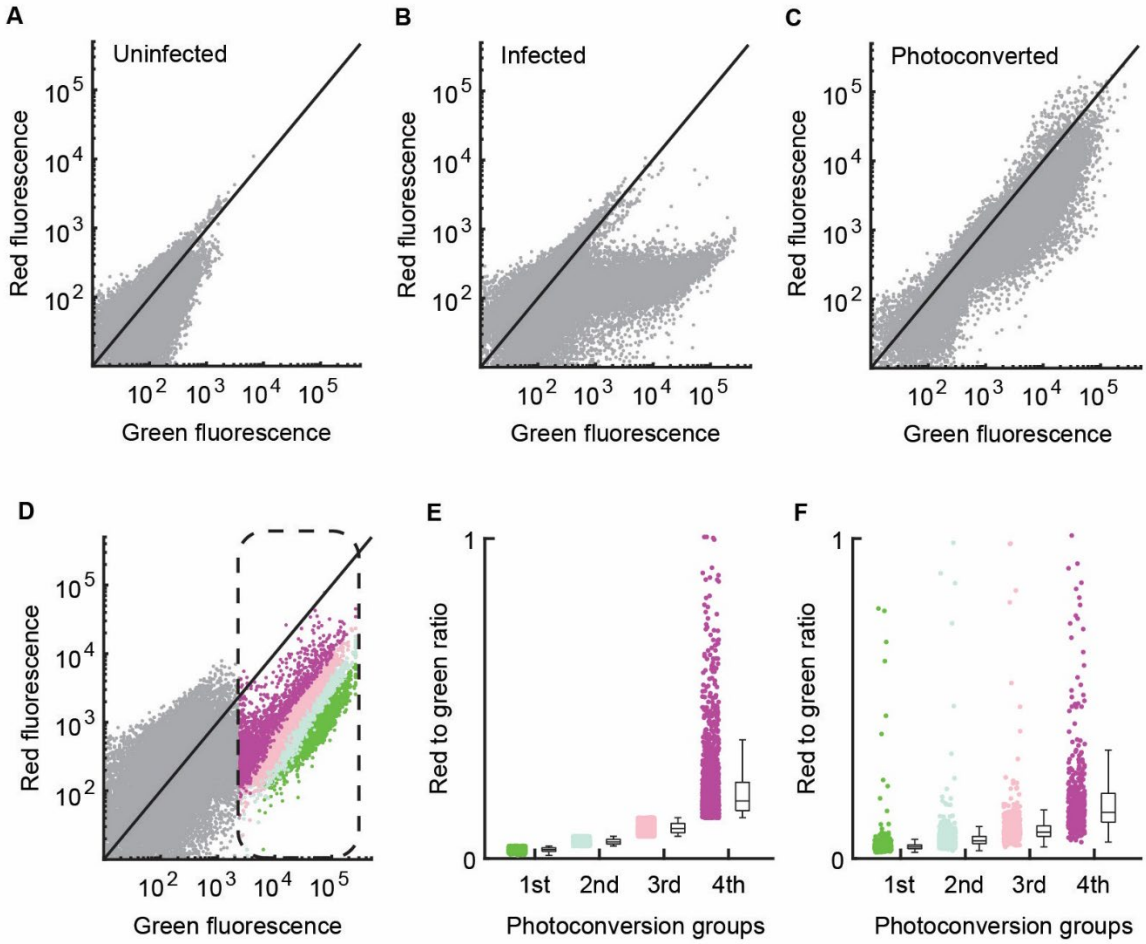


Figure S1. FACS sorting of CaMPARI2 infected cells that underwent photoconversion.

(**A**) FACS plot of an example dissociated piece of cortical tissue without CaMPARI2 expression. Each dot corresponds to one putative cell. Cells were selected for size, filtered for doublets, and high Draq7 (dead) cells were excluded.

(**B**) As in **A**, but for an infected cortical preparation that expressed CaMPARI2 without photoconversion.

(**C**) As in **B**, with photoconversion.

(**D**) An example FACS plot illustrating ratio-metric sorting. FACS gates were set to exclude background fluorescence and to sort by red to green ratio. Cells were sorted into equally sized bins corresponding to the first (green), second (pale blue), third and fourth quantiles of the red to green fluorescence ratio distribution. The middle quartiles were later combined for single-cell analysis.

(**E**) The distribution of red to green ratio values for cells shown in **D**, for the different photoconversion groups separately. A small number of cells had ratio values larger than 1 (not shown).

(**F**) To test for stability of the FACS sorting, we measured the red to green ratio in the photoconversion groups shown in **E** separately by passing them through the FACS sorter again. Shown is the distribution of red to green ratios measured in the second measurement.

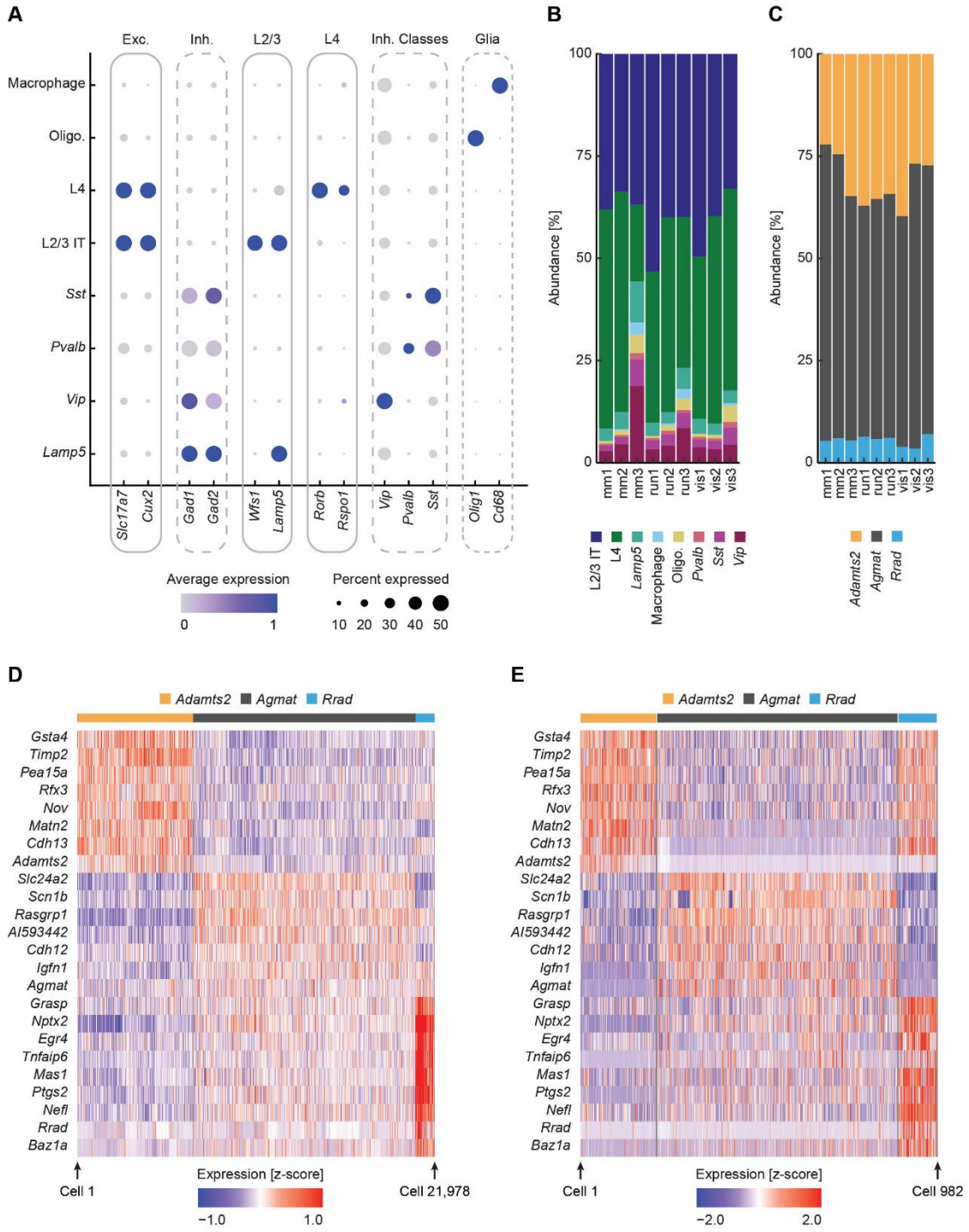


Figure S2. Cell type assignment accuracy.

(**A**) Dot plot for a selected number of marker genes for each of the major cell groups (with a cumulative population size greater than 500 cells across all samples). Dot size represents the percentage of cells

expressing each marker while the color represents the average normalized expression value. Exc.: excitatory neurons, Inh.: inhibitory neurons, Oligo.: Oligodendrocytes, IT: intertelencephalic.

(**B**) Abundance of each major cell group as shown in **A**, across the three photoconversion groups and the three photoconversion types. mm: Mismatch; run: Running onsets; vis: Grating onsets; 1: Low photoconversion group; 2: Intermediate photoconversion group; 3: High photoconversion group.

(**C**) As in **B**, but for the three L2/3 excitatory types.

(**D**) Heatmap for a select number of genes enriched in each of the three L2/3 excitatory cell types. There were 21 978 L2/3 excitatory neurons examined in this study (7113 *Adamts2*, 13 688 *Agmat*, 1177 *Rrad*, and 64 unassigned neurons, the latter are not shown in the heatmap). Color values correspond to expression levels measured as a z-score. To prevent aliasing, data are smoothed across cells with a window size of 20.

(**E**) As in **D**, for all L2/3 V1 excitatory neurons from a reference dataset (Tasic et al., 2018).

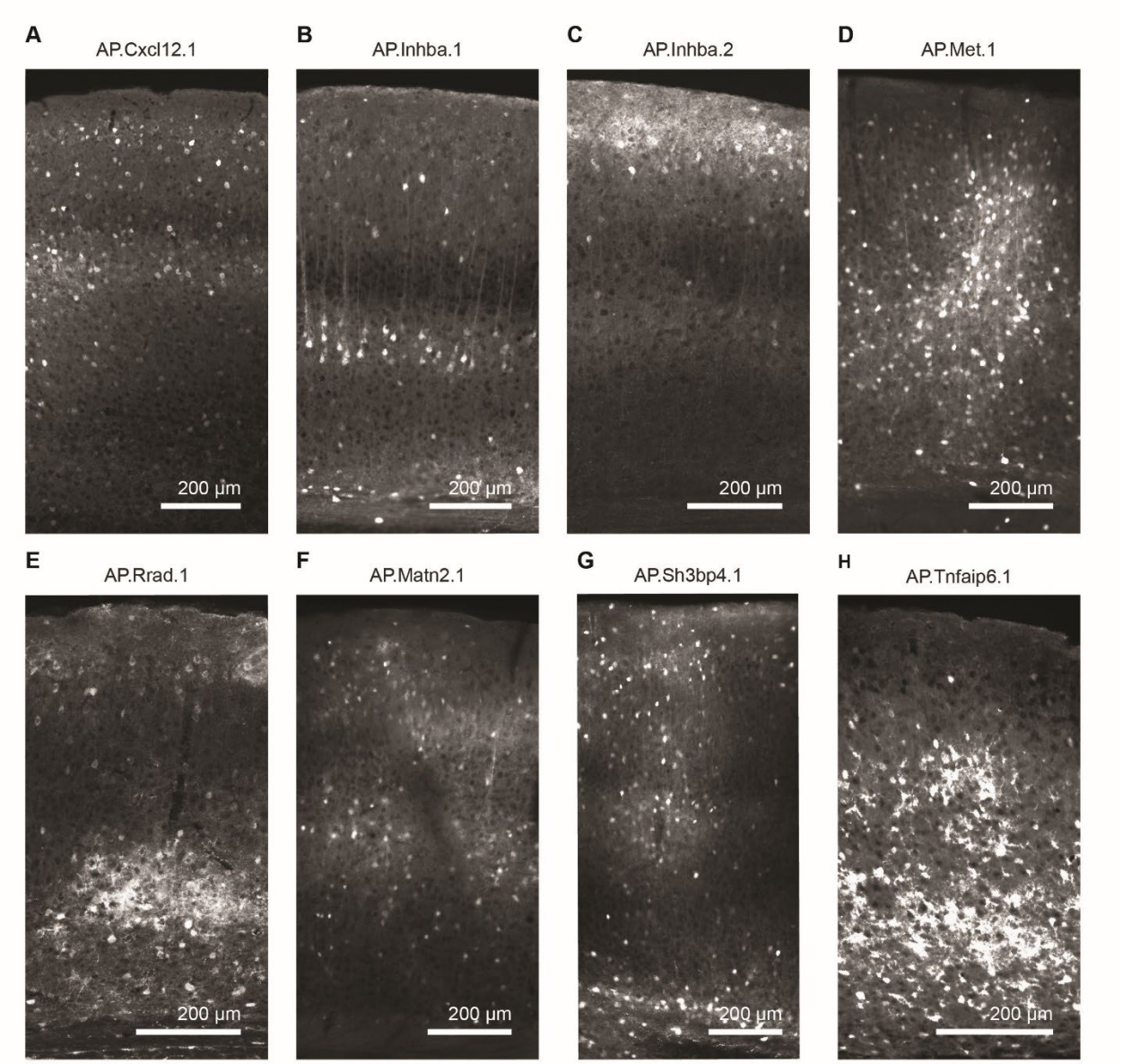


Figure S3. Some of the failed attempts of promoter designs.

(**A**) Expression pattern of an AAV2/1-AP.Cxcl2.1-GFP with an artificial promoter designed based on the 2 kb surrounding the transcription start site of *Cxcl12*, a putative maker of the L2/3 *Agmat* cell type.

(**B**) As in **A**, but for an artificial promoter designed based on the 2 kb surrounding the first transcription start site (relative to the + strand of chromosome 13) of the *Inhba* gene, a putative marker of the L2/3 *Adamts2* cell type.

(**C**) As in **A**, but for an artificial promoter designed based on the 2 kb surrounding the second transcription start site (relative to the + strand of chromosome 13) of the *Inhba* gene, a putative marker of the L2/3 *Adamts2* cell type.

(**D**) As in **A**, but for an artificial promoter designed based on the 2 kb surrounding the first transcription start site (relative to the + strand of chromosome 6) of the *Met* gene, a putative marker of the L2/3 *Adamts2* cell type.

(**E**) As in **A**, but for an artificial promoter designed based on the 2 kb surrounding the transcription start site of the *Rrad* gene, a putative marker of the L2/3 *Rrad* cell type.

(**F**) As in **A**, but for an artificial promoter designed based on the 2 kb surrounding the transcription start site of the *Matn2* gene, a putative marker of the L2/3 *Adamts2* cell type.

(**G**) As in **A**, but for an artificial promoter designed based on the 2 kb surrounding the transcription start site of the *Sh3bp4* gene, putative marker of the L2/3 *Adamts2* cell type.

(**H**) As in **A**, but for an artificial promoter designed based on the 2 kb surrounding the transcription start site of the *Tnfaip6* gene, putative marker of the L2/3 *Rrad* cell type.

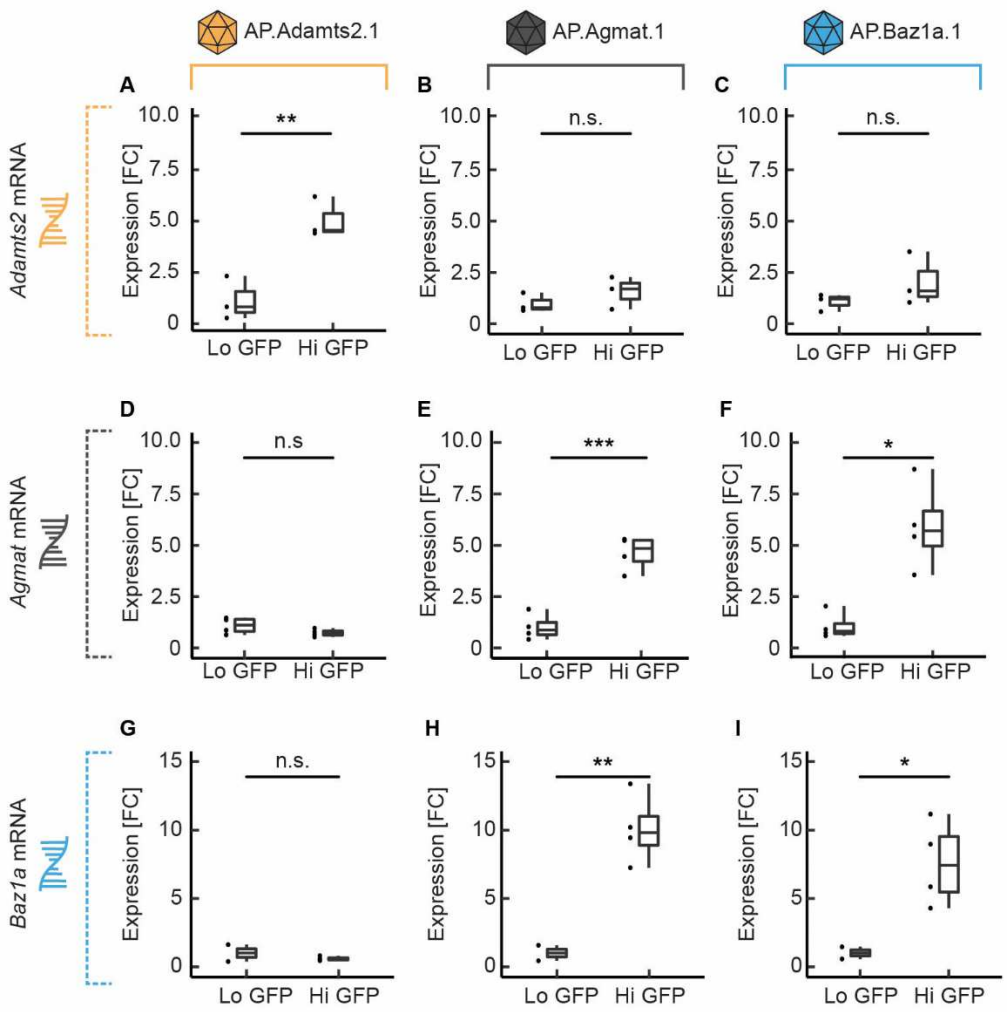


Figure S4. Differential expression of marker genes for artificial promoter infected neurons.

(**A**) Bulk RNA-sequencing data for *Adamts2* expression in populations of L1, L2/3 and L4 cortical neurons infected with an AAV2/1-AP.Adamts2.1-GFP, for high (Hi) and low (Lo) GFP expression level populations separately. The low and high GFP groups constitute the lower and upper thirds of the GFP fluorescence distribution (cells with no expression were excluded). Fold-change (FC) values were normalized to the average expression in the low GFP group. Note all samples were collected in pairs, however, in some cases library preparation failed for one of the two samples thus sample sizes were unequal (here and in other figure panels of this figure).

(**B**) As in **A**, except for the AP.Agmat.1 promoter.

(**C**) As in **A**, except for the AP.Baz1a.1 promoter.

(**D-F**) As in **A-C**, except for *Agmat* expression.

(**G-I**) As in **A-C**, except for *Baz1a* expression.

MATERIALS AND METHODS

Mice and surgical procedures. All animal procedures were approved by and carried out in accordance with guidelines of the Veterinary Department of the Canton Basel-Stadt, Switzerland. For most experiments we used C57BL/6J mice (Charles River). For the experiments designed to restrict transgene expression to excitatory neurons we used Emx-Cre knock-in mice (Gorski et al., 2002). For all surgical procedures, mice were anesthetized with a mixture of Fentanyl (0.05 mg/kg; Actavis), Midazolam (5.0 mg/kg; Dormicum, 603 Roche) and Medetomidine (0.5 mg/kg; Domitor, Orion). In all mice used for functional imaging, a 4 mm or 5 mm craniotomy was made over right V1 and 4 injections of approximately 250 nl each were made centered on 2.5 mm lateral and 0.5 mm anterior of lambda. For CaMPARI2 single-cell sequencing experiments two 4 mm windows were implanted at the same coordinates bilaterally with two injections in each hemisphere. Circular glass cover slips were glued (Ultragel, Pattex) in place to close the craniotomy, and the exposed skull covered in Histoacryl (Braun). To enable head-fixation, a custom-made titanium head-plate was attached to the skull with dental cement (Heraeus Kulzer).

Acclimatization to experimental setups. Prior to the start of photoconversion and imaging experiments, mice were trained on the experimental setups for 3 days in the case of photoconversion experiments, and 6 days for functional imaging experiments. Each training session lasted for 1 hour during which the mouse was head-fixed on the spherical treadmill and experienced closed-loop visual flow feedback projected onto a toroidal screen surrounding the mouse. The rotation of the treadmill was restricted to rotation around the transverse axis.

Photoconversion. Photoconversion experiments were performed on mice that robustly expressed CaMPARI2 throughout left and right V1. Out of a total of 34 mice, 12 (4 for each condition) were selected for photoconversion experiments based on CaMPARI2 expression levels, vasculature, absence of any signs of infection or noticeable bone regrowth. For photoconversion experiments, mice were head-fixed on a spherical treadmill surrounded by a toroidal screen. The screen covered an area corresponding to 240 degrees horizontally and 100 degrees vertically relative to the mouse's field of view. All mice were exposed to the same sequence of visuomotor sessions. First, mice experienced 30 min of closed-loop coupling between running and visual flow feedback in a virtual corridor. To induce visuomotor mismatches, we briefly (1 s) halted the visual flow at random intervals (on average every 15 s, +/- 10 s). Next, mice experienced 30 min of a dark session during which all light sources were either

398 turned off or shielded. Finally, mice were presented with full-field drifting sinusoidal grating stimuli (2.2
399 Hz, 0.04 cycles per degree), designed to be minimally predictable, for 3 to 8 s of randomly selected
400 directions (0°, 45°, 90°, 135°, 180°, 225°, or 270°), spaced by 2 s to 6 s gray screen presentations, for a
401 total of 30 min. Mice were free to run throughout all sessions.

402 For photoconversion, we used a 405 nm laser (Obis, Coherent). The laser was directed bilaterally at V1
403 through the cranial windows for a duration of 1 s for each trigger. The laser power was 100 mW with a
404 diameter of 3 mm (FWHM) at the surface of the brain, similar to that used in a previous study (Moeyaert
405 et al., 2018). The laser was directed between a blank position and the cranial windows using a two axis
406 galvo scan head. During the 1 s stimulation, the beam was alternated between left and right V1 at a
407 frequency of 60 Hz. To minimize visible changes in ambient light, we shielded the entire light path up to
408 the head-bar of the mouse. All mice received between 30 and 50 photostimulation pulses of 1 s during
409 the 30 min photoconversion session. In instances in which mice did not receive 50 pulses, the
410 photoconversion session was extended until the mouse received 50 pulses. In instances in which mice
411 received 50 pulses before the end of the session, stimulation was stopped for that session. For the
412 visuomotor mismatch photoconversion events, we triggered the laser only on mismatch events during
413 which the running speed of the mouse was above 2 cm/s. For photoconversion during running onsets in
414 darkness, running onsets were detected in real-time as events with increases in running speed of at least
415 1.4 cm/s over a span of 500 ms, and a running speed in the window -1500 ms to -500 ms before the
416 event that did not exceed 1 cm/s. For photoconversion during grating onsets, the laser was triggered at
417 the grating onset if the running speed of the mouse was below 0.5 cm/s. Following photoconversion
418 experiments, all mice were immediately anesthetized, examined on an epifluorescence microscope for
419 evidence of successful photoconversion and sacrificed for single-cell RNA-sequencing.

420 Due to sample processing limitations, only two session types were collected at a time. In order to
421 balance potential batch effects, each session type was collected twice in tandem with one of the other
422 session types. Thus, we executed the single-cell sequencing photoconversion experiment three times
423 (mismatch with visual; mismatch with running; running with visual). So, for all photoconversion
424 experiments, mice were split into 3 groups of 2. In the first group of 2 mice, photoconversion was
425 triggered concurrent with visuomotor mismatch stimuli in the closed-loop session. In the second group
426 of 2 mice, photoconversion was triggered at the onset of running in the dark session. In the third group
427 of 2 mice, photoconversion was triggered at the onset of grating stimuli in instances when the mice

were stationary. After each experiment photoconversion was confirmed by examining the cranial window with an epifluorescence microscope. This was repeated twice for a total of 12 animals.

Solutions. We used artificial cerebrospinal fluid (ACSF) for the generation of single-cell suspensions as previously described (Hempel et al., 2007; Tasic et al., 2018), with some modifications. We used *N*-methyl-d-glucamine (NMDG) based ACSF to improve cell viability (Ting et al., 2014). The NMDG ACSF formulation, based on a previous publication (Tasic et al., 2018), consisted of NMDG (96 mM), HCl (96 mM), $NaHCO_3$ (25 mM), glucose (25 mM), HEPES (20 mM), N-acetylcysteine (12 mM), $MgSO_4$ (10 mM), sodium l-ascorbate (5 mM), myo-inositol (3 mM), sodium pyruvate (3 mM), KCl (2.5 mM), thiourea (2 mM), NaH_2PO_4 (1.25 mM), $CaCl_2$ (0.5 mM), and taurine (0.01 mM). All ACSF stocks were kept at pH 7.35 and 315 mOsm. Additionally, all solutions were supplemented with 13.2 mM trehalose to enhance cell survival during the dissociation and sorting (Saxena et al., 2012). Solutions were supplemented with a transcriptional blocker, actinomycin-d (1 µg/ml, Sigma, cat. no. A1410), to reduce dissociation related expression changes (Wu et al., 2017). In addition, the ion channel blockers TTX (0.1 µM, Tocris,), DNQX (200 µM, Merck/Sigma), and APV (50 µM, Tocris) were added to prevent any activity related changes and reduce excitotoxicity (Hempel et al., 2007). The ACSF used for slicing and cooling the brain was pre-chilled to an icy slush, and bubbled with carbogen gas (95 % O_2, 5 % CO_2) throughout the experiment.

Generation of single-cell suspensions. The procedure to generate a single-cell suspension was identical for both single-cell and bulk RNA-sequencing experiments. Brain extraction and cellular dissociation for all experiments followed a standard protocol (Hempel et al., 2007) with some modifications. Mice were anesthetized and decapitated. To remove the brain, the skull was cut medio-laterally directly anterior of the cerebellum, along the lateral edge of the dorsal skull, and medio-laterally above the olfactory bulb. The bone flap was then opened, and the brain removed with a spatula. With a scalpel, the brain was cut, coronally, at 1 mm anterior of bregma to form a flat surface, the anterior most portion of the brain was discarded. On this surface, the brain was then glued to a cutting platform next to an agarose block (or embedded in low-melt agarose for single-cell sequencing experiments) and rapidly immersed in ice-cold carbogenated ACSF. Brain sections (300 µm) were cut on a vibratome (7000smz-2, Campden Instruments) with a ceramic blade at an oscillation frequency of 80 Hz and a cutting speed of 200 µm/s. Sections within and around V1 were collected with a plastic Pasteur pipette and placed into a bubbled chamber filled with room-temperature ACSF and allowed to recover for at least 15 min. Subsequently, samples were placed into 30 ml of ACSF with 1 mg/ml pronase (Protease Type XIV, Sigma), along with 1

mg collagenase (Collagenase from Clostridium histolyticum, Sigma), as previously described (O'Toole et al., 2017) and digested for 75 min. Then slices were placed in ACSF (without protease/collagenase) and allowed to recover for 15 min. After that, the slices were transferred to a 30 cm petri dish containing cold carbogenated ACSF supplemented with 1 % Fetal Bovine Serum (FBS). Supragranular layers of V1 were microdissected under an epifluorescent microscope. Microdissected V1 pieces were then transferred to 1.5 ml RNase free tubes. These tissue pieces were triturated with a series of fire-polished glass Pasteur pipettes, with inner diameters of 600 µm, 300 µm, and 150 µm, that had been coated in ACSF with 1% FBS (Hempel et al., 2007). Single-cell suspensions were then strained through nylon filters (Celltrics filters, sysmex) and placed into tubes (12 × 75 mm Round Bottom Polystyrene Test Tubes, Fisher) along with 2 µl of live/dead cell stain (DRAQ7, BioLegend) per ml and then FACS sorted.

FACS. For FACS sorting, we used either a BD FACSAria III cell sorter or a Sony MA900 cell sorter. To enrich for cellular particles, we used a series of gates based on particle size and complexity, excluding doublets and particles that stained positive for DRAQ7, a marker of dead cells. For ratiometric FACS, we ran a series of control experiments to determine the range in which GFP and RFP expression could be observed. An uninfected control was used to determine the cut off gate for viral expression (**Figure S1**). In some cases, cell samples were examined under a fluorescence microscope to check for fluorescence. At the beginning of each experiment, we sorted at least 1000 cells above the expression threshold to estimate the fluorescence gates that would bin approximately 25% of the neurons into one of four photoconversion groups (the middle two were later combined for analysis purposes), based on their red to green fluorescence ratio. Using these gates, we then sorted 10 000 neurons per photoconversion group. For each photoconversion experiment, the single cell suspensions obtained from two mice were sorted (serially) to reach a threshold of 40 000 cells. Between mice, gates were recalibrated to match the photoconversion distribution of the current sample. We confirmed sorting reliability by resorting a sorted sample and found that the difference between photoconversion groups were maintained (**Figure S1F**). All samples were sorted directly into a PBS/BSA solution such that the final concentration was 0.04 % BSA. Immediately after sorting, samples were used for Gel Bead in Emulsions and cDNA generation.

For FACS sorting of the bulk RNA-sequencing experiments we followed a similar protocol. Prior to the experiment, mice were injected at 6 to 8 weeks of age, and brains were extracted 4 weeks later to generate single-cell suspensions. We used a series of gates based on particle size, excluding doublets and particles that stained positive for DRAQ7. As before, a pilot experiment was used to determine a

cutoff value for the absence of viral expression by examining an uninfected sample. Then, for infected samples, 1000 cellular particles were measured to determine the distribution of fluorescence values. Gates were established that captured four categories of fluorescence: particles with no expression as well as low, medium, and high expressing particles each corresponding to a third of the distribution of particles above the baseline cutoff for viral expression. We restricted analysis to cells expressing low (just above the expression threshold) and high amounts of GFP. These dynamically drawn gates were used to reduce batch effects caused by infection efficiency and differences in viral titer. 1000 cell particles were sorted for each group directly into lysis buffer (Norgen). Following sorting, samples were either stored at -80°C or immediately used for RNA purification.

Sequencing library preparation. RNA for bulk RNA-sequencing samples was purified with the Single Cell RNA Purification Kit (Norgen), and columns were treated with a DNase to remove genomic DNA (Norgen). The isolation was performed according to the manufacturer's instructions. Prior to cDNA generation, samples were measured with the Qubit PicoGreen Assay (to assess RNA quality). Samples were then processed with the SmartSeq2 protocol (Picelli et al., 2014). Prior to sequencing, library size distribution and concentration were assessed with the bioanalyzer High Sensitivity DNA assay. For all single-cell RNA-sequencing experiments, single-cell suspensions were immediately processed with the 10x genomics chromium instrument after FACS sorting. GEM formation, reverse transcription, barcoding, cDNA amplification, and purification were all done in accordance with the manufacturer's recommendations. Both v2 and v3 versions of the kit (10x Genomics, single cell RNA-sequencing 3') were used. Single-cell libraries were sequenced on the NextSeq at 75 cycles with a paired end protocol. Single-cell samples were sequenced to an average read depth of 264 million reads per sample with a median count of 3 376 UMIs per cell. Bulk RNA-sequenced samples had a read depth of 15 to 20 million reads. The AP.Adamts2.1-GFP and AP.Baz1a.1-GFP samples were sequenced on the HiSeq 2500 (50 bp, single-end), while the AP.Agmat.1-GFP samples were sequenced on the NovaSeq (2x 50 bp, paired-end). The change in sequencing instrument and protocol was a consequence of logistical changes made at the institute.

Single-cell RNA-seq analysis. Initial processing, up until the construction of the cell-barcode count matrices, of the single-cell sequencing data was done with Cell Ranger software package. Reads were mapped against a custom genome constructed from the Genome Reference Consortium Mouse Build 39 with a version 104 GTF file. Viral expression was accounted for by amending the CaMPARI2 plasmid

sequence (Addgene) to the end of the FASTA file and the GTF file was edited to include the regions present on the viral mRNA. GTF and FASTA edits were performed in python and the final genome build was constructed with Cell Ranger. All subsequent analyses on the raw cell feature barcode matrices were done in R.

Expression data was imported with the SCATER package (McCarthy et al., 2017). Cells were distinguished from empty droplets with DropletUTILS (Lun et al., 2019), using a FDR of 0.05 % and a lower limit for cell detection that was calculated on a per dataset basis or set to the median UMI count for droplets greater than 10 UMIs. Cells with a mitochondrial read percentage greater than 20 % as well as cells with read counts that were more than three standard deviations below the sample average were excluded. Feature selection and UMI correction were done with SCtransform from the Seurat package (Hafemeister and Satija, 2019; Satija et al., 2015), initially 2000 features were selected and mitochondrial percentage, sample effects, 10x chemistry (v2 vs v3), and total feature counts were regressed out. Selected features were filtered to remove any features that were differentially expressed between groups of cells separated by read count or mitochondrial percentage, thus reducing the impact of cell quality on the analysis. The final corrected UMI matrix was then used for alignment with a reference dataset (Tasic et al., 2018).

To align cells from this study with the reference dataset (Tasic et al., 2018), excluding L5 and L6 neurons, we used the Liger package (Welch et al., 2019) with a k-value of 12 and lambda-value of 10. These parameters were chosen based on how well groups were separated on a UMAP plot, as well as the expression of various markers normally enriched in each group (**Figure S2**). Then, iNMF factors were used to assign each cell from this study to a group defined in the reference dataset (L4, L2/3, SST etc.) using a weighted nearest neighbor algorithm, where the largest group sum for the of the inverse of the distance for the 10 closest reference cells was used to assign group identity. Cells whose distance to the closest cell in the reference dataset exceeded 0.002 in the iNMF factor space were removed from the dataset. Additionally, we excluded cell types that contained less than 500 cells across the entire dataset. Furthermore, the initially assigned *Pvalb* group contained a large number of cells expressing excitatory neuron markers, graphical clustering was used to subset this group and isolate *Pvalb* positive neurons that were negative for excitatory markers. L2/3 subtypes (*Adamts2*, *Rrad*, and *Agmat*) were determined separately within the L2/3 group with scMAP (Kiselev et al., 2018) using the scmapCluster function; 250 features were used with no similarity cutoff. These parameters allowed a demonstrably accurate

548 assignment of L2/3 neurons to their corresponding cell types from the reference dataset (Tasic et al.,
549 2018) types (**Figure S2**).

550 To analyze how the proportion of L2/3 types changed as a function of photoconversion (**Figures 2E, 2F,**
551 **2G and 2H**), we estimated the fraction of each L2/3 type (within each experiment or condition) by
552 pooling the number of cells collected across photoconversion conditions, correcting for differences in
553 cell number and then calculating the proportion of each cell type in these pooled datasets. These
554 estimates allowed us to simulate a randomized distribution of proportion values for each L2/3 cell type
555 in each experiment. Then, experimentally obtained proportion values were compared to a distribution
556 obtained by bootstrap sampling from the combined population across photoconversion conditions.

557 **Artificial promoter design.** Artificial promoters for selective expression in L2/3 excitatory neuron types
558 were designed based on the ATG proximal sequence design strategy described previously (Jüttner et al.,
559 2019). Artificial promoters contained a sequence of 2 kb surrounding the transcription start site of the
560 endogenous marker gene taken from the UCSC genome browser (GRCm38/mm10). For the
561 AP.Adamts2.1 promoter, the 2 kb sequence originally contained a Kozak sequence (corresponding to the
562 translation start site of the *Adamts2* gene). To enhance transgene expression, the endogenous start
563 codon ('ATG') of the Kozak sequence was changed to a stop codon ('TGA').

564 **Virus production.** Artificial promoter constructs were synthesized and subcloned into an AAV vector
565 backbone plasmid by a commercial supplier (Genscript), placed upstream of GFP for histology and gene
566 expression experiments, and upstream of GCaMP7f (Dana et al., 2019) (modified from Addgene plasmid
567 #104495), or a conditional version of GCaMP8s (Zhang et al., 2021) for calcium imaging experiments.
568 AAV vectors were produced by the FMI vector core following standard protocols. In brief, HEK293T cells
569 were grown to 80 % confluence and co-transfected with an AAV transgene plasmid, an AAV helper
570 plasmid, and a plasmid containing the Rep/Cap proteins (Addgene Plasmid #112862) using
571 polyethylenimine (Polysciences). Cells were harvested and pelleted in a benchtop centrifuge 60 h after
572 transfection and the supernatant discarded. Cell pellets were either stored at -80 °C or immediately
573 processed. Cell pellets were resuspended in 15 ml lysis buffer (150 mM NaCl, 20 mM Tris, pH 8.0).
574 Solutions went through 4 freeze-thaw cycles alternating between liquid nitrogen and a 37°C water bath
575 (15 min each) interrupted by brief periods of mixing on a Vortex shaker. Subsequently, $MgCl_2$ was added
576 to a final concentration of 1 mM, followed by the addition of Turbonuclease (bpsBioscience) at 250 U/nl.
577 Then samples were incubated at 37 °C for 15 min. Virus solutions were spun down and the pellet was

discarded. The supernatant was transferred into a prepared iodixanol OptiPrep (Sigma) gradient (6 ml 17 %, 6 ml 25 %, 5 ml 40 % and 5 ml 60 %) contained within an OptiSeal tube (Beckman,). Gradients were spun at 242 000 g at 16 °C and the viral fraction was harvested from the 40 % fraction with a 21-gauge needle. The viral fraction was further purified with Amicon 100K columns (Merck) according to the manufacturer's instructions. Viral aliquots were stored at -80 °C.

Titers of AAV vectors were determined by qPCR. First, extra-capsid DNA was digested with DNase, 10 U/µl (Roche) for 15 min at 37 °C, and inactivated for 10 min at 95 °C. Viral capsids were then digested with proteinase K at a concentration of 100 ng/ml (Macherey Nagel) for 15 min at 37 °C and inactivated at 95 °C for 10 min. All AAVs were measured relative to standards with primers targeted to the ITRs (forward: 5'-GGAACCCCTAGTGATGGAGTT, reverse: 5'-CGGCCTCAGTGAGCGA) with the SYBR Green PCR Master Mix (ThermoFisher). qPCR was performed on the StepOnePlus real-time PCR system. Then qPCR curves were analyzed with the Step One v2.3 software from Agilent.

Histology. To examine the viral expression pattern for artificial promoter constructs, mice were transcardially perfused with phosphate buffered saline (PBS) until the liver was exsanguinated. Mice were then perfused with 4% paraformaldehyde (PFA) in PBS. Brains were removed and placed into cold (4 °C) PFA and fixed overnight. The following day, brains were transferred to cold PBS and washed 3 times for 20 min (4 °C). Brains were sectioned with either a vibratome (Leica VT1000 S) or a cryostat (CryoStar NX70). For cryosections, brains were cryoprotected with 3 overnight washes in 30 % sucrose in PBS. Afterwards brains were equilibrated in Tissue-Tek O.C.T. compound (Biosytems) and then flash frozen in OCT with liquid nitrogen. Brains were sectioned at 20 µm and adhered to slides (FisherScientific). For vibratome sections, brains were sliced in room temperature PBS at 50 µm, stored in a 48 well plate, and then mounted on glass slides. Sections were mounted in VectaShield HardSet with DAPI (VectorLabs). Most slices also underwent a staining protocol prior to mounting. Sections were washed 4 times in PBS with 0.1 % Triton (PBST) for 10 min. Then the sections were blocked for 2 h at room temperature in PBST with 5 % neutral goat serum. Thereafter, the sections were incubated overnight at 4 °C with chicken anti-GFP (ThermoFisher). Then sections were washed in PBST 4 times for 10 min. Afterwards, the sections were stained with secondary antibodies (1:500) along with NeuroTrace 640/660 (ThermoFisher) which was diluted (1:300), followed by a final washing step (4 times 10 min PBST). Images were acquired on a Zeiss Axio Scan.Z1 slide scanner and in some cases on an Axio Imager Z2 with a LSM 700 scan head.

Histology analysis. To determine the expression profile of viral vectors as a function of cortical depth, we used images corresponding to 200 µm in width and spanning the entire depth of cortex. Image processing was done in FIJI where images were converted to 8-bit and then enhanced for local contrast (CLAHE plugin, blocksize = 30, histogram = 256, maximum = 3). The images were then binarized with a cutoff value corresponding to the 92.5th percentile of pixel intensities (calculated for each image separately) and filtered for cellular particles. Binarized images were converted to depth profiles by averaging in the horizontal dimension. All vectors were then scaled to an equal length and normalized to peak. Finally, vectors were averaged across groups, normalized to peak, and passed through a Savitzky-Golay filter.

Bulk RNA-seq analysis. To examine the differences in expression between neurons infected with different artificial promoters, a custom genome was constructed that excluded all DNA segments used in the artificial promoter viruses. This was done to increase the accuracy of the read counts for the marker genes. A custom GTF file was used to include the viral genomes. As with the single-cell RNA-sequencing analysis, the custom genome was based on the Genome Reference Consortium Mouse Build 39 with a version 104 GTF file. Differences in sequencing method (paired vs single-end) were taken into account for the trimming and mapping steps. Reads were trimmed with the Cutadapt package (trim sequence: CTGTCTCTTATA). Gene counts were calculated with HTSeq count. Libraries were corrected for batch effects related to sequencing runs. Count values were normalized to library depth as counts per million (cpm). Genes with less than 5 cpm on average, across datasets, were excluded. Statistics were performed on fold change values.

Two-photon calcium imaging. Calcium imaging was confined to L2/3 of V1 and was performed as previously described (Leinweber et al., 2014). A modified Thorlabs Bergamo II microscope with a 16x, 0.8 NA objective (Nikon N16XLWD-PF) was used for imaging. The illumination source was a tunable femtosecond laser (Insight, Spectra Physics or Chameleon, Coherent) tuned to 930 nm for the GCaMP imaging experiments, to 950 nm for the functional CaMPARI2 experiments, and to 1040 nm for the simultaneous imaging of red and green channels to characterize the extent of photoconversion. Emission light was band-pass filtered using a 525/50 nm or 607/70 nm filter (Semrock), and signals were detected with a GaAsP photomultiplier (Hamamatsu, H7422). PMT signals were amplified (Femto, DHCPCA-100), digitized at 800 MHz (National Instruments, NI5772), and band-pass filtered at 80 MHz with a digital Fourier-transform filter on a field-programmable gate array (National Instruments, PXIe-

7965). The microscope scan head was a 12 kHz resonant scanner (Cambridge Technology). Image acquisition was at a resolution of 750 pixels by 400 pixels (60 Hz frame rate) with a field of view of 375 µm by 300 µm. Four layers, separated by 15 µm were imaged sequentially with a piezo-electric linear actuator (Physik Instrumente, P-726), and the effective frame rate per layer was 15 Hz.

Two-photon image analysis. Image analysis was performed as previously described (Keller et al., 2012). In brief, acquired data were full-frame registered to correct for brain motion. In datasets with weaker signal, registration was performed on a running average over 5 to 10 frames. Raw fluorescence traces for individual neurons were calculated as the average signal in a region of interest (ROI) that was manually selected based on mean and maximum fluorescence images. To correct for slow drift in fluorescence, the raw fluorescence trace was first corrected with an 8^{th} percentile filter over a window of 66 s (1000 frames) as described previously (Dombeck et al., 2007), and F_0 was defined as the median fluorescence over the entire trace. To calculate the average response traces, we first calculated the average event-triggered fluorescence trace for each neuron. The responses of all neurons were averaged across triggers and baseline subtracted. The baseline window for the CaMPARI2 analysis was 1.33 s (20 frames) and for the GCaMP analysis it was 0.66 s (10 frames). We used different analysis windows for the two indicators due to the differences in signal to noise levels. To quantify the difference of average calcium responses, for GCaMP based analyses (**Figure 4**), as a function of time, we performed a two-sample t-test for every bin of the calcium trace (15 Hz) and marked bins as significantly different for $p < 0.001$. For visual clarity, we removed isolated selected bins, such that a significant bin was only marked if at least two neighboring bins were also significant.

658

659 **Key Resources Table**

REAGENT or RESOURCE	Source	Identifier
Bacterial and Virus Strains		
AAV2/1-AP.Adamts2.1-GFP-WPRE (3.0E+10 GC/ml)	FMI vector core	vector.fmi.ch
AAV2/1-AP.Agmat.1-GFP-WPRE (7.3E+10 GC/ml)	FMI vector core	vector.fmi.ch
AAV2/1-AP.Baz1a.1-GFP-WPRE (3.6E+12 GC/ml)	FMI vector core	vector.fmi.ch
AAV2/1-AP.Adamts2.1-DIO-GCaMP8s-WPRE (1.8E+14 GC/ml)	FMI vector core	vector.fmi.ch
AAV2/1-AP.Agmat.1-GCaMP7f-WPRE (2.08E+12 GC/ml)	FMI vector core	vector.fmi.ch
AAV2/1-AP.Baz1a.1-DIO-GCaMP8s-WPRE (1.1E+14 GC/ml)	FMI vector core	vector.fmi.ch
AAV2/1-hSyn-NES-CaMPARI2-WPRE (1.4E+12 GC/ml)	Addgene	101060-AAV1
Rep/Cap 2.1 plasmid	Addgene	Plasmid # 112862
pHGTI-adeno1 plasmid	Laboratory of Constance Cepko	N/A
Chemicals, Peptides, and Recombinant Proteins		
Fentanyl citrate	Actavis	CAS 990-73-8
Midazolam (Dormicum)	Roche	CAS 59467-96-8
Medetomidine (Domitor)	Orion Pharma	CAS 86347-14-0
Ropivacaine	Presenius Kabi	CAS 132112-35-7
Lidocaine	Bichsel	CAS 137-58-6
Buprenorphine	Reckitt Benckiser Healthcare	CAS 52485-79-7
Ophthalmic gel (Humigel)	Virbac	N/A
Flumazenil (Anexate)	Roche	CAS 78755-81-4
Atipamezole (Antisedan)	Orion Pharma	CAS 104054-27-5
N-Butyl-2-cyanoacrylate (Histoacryl)	Braun	CAS 6606-65-1
Dental cement (Paladur)	Heraeus Kulzer	CAS 9066-86-8
Dulbecco's Modified Eagle Medium	ThermoFisher	CAS 31330095
PEI	Polysciences	CAS 23966-2
HEK293T cells	Sigma	CAS 12022001-1VL
Turbonuclease	Sigma	CAS T4330-50KU
OptiPrep Density Gradient Medium	Sigma	CAS 92339-11-2
Amicon Ultra-15 Centrifugal Filter Concentrator	Sigma	CAS UFC910024
DNase I	Sigma	CAS 4716728001
Proteinase K	Macherey Nagel	CAS 740506
SYBR Green PCR Master Mix	ThermoFisher	CAS 4309155
fwd ITR primer, 5'-GGAACCCCTAGTGATGGAGTT-3'	Microsynth	N/A
rev ITR primer, 5'-CGGCCTCAGTGAGCGA-3'	Microsynth	N/A
ColorFrost Plus Microscope Slides	Fisher	CAS 12-550-18
VECTASHIELD HardSet Antifade Mounting Medium	Vector labs	CAS H-1400-10
DL-AP5	Tocris	CAS 0105
Tetrodotoxin citrate	Tocris	CAS 1069
DNQX	Sigma	CAS D0540
Actinomycin-d	Sigma	CAS A1410
Trehalose	Sigma	CAS 6138-23-4
DRAQ7 Dye	ThermoFisher	CAS D15106
sterile single-pack CellTrics filters	Sysmex	CAS 04-004-2326
Norgen Single Cell RNA Purification Kit	Norgen	CAS 51800
RNase-Free DNase I Kit	Norgen	CAS 25710
Deposited data		
All data and code used to generate manuscript figures	This paper	data.fmi.ch
Experimental Models: Organisms/Strains		

Mus musculus: C57BL/6J	Charles River	N/A
Mus musculus: B6.129S2-Emx1tm1(cre)Krj/J	Jackson laboratories (Gorski et al., 2002)	RRID:IMSR_JAX:005628
Software and algorithms		
MATLAB (2020b)	MathWorks	RRID: SCR_001622
LabVIEW	National Instruments	RRID:SCR_014325
Two-photon acquisition software	Keller laboratory	sourceforge.net/p/iris-scanning/
Image data processing software	Keller laboratory	sourceforge.net/p/iris-scanning/calliope
Python	python.org	RRID:SCR_008394
Panda3D	panda3d.org	RRID:SCR_021216
FlowJo_V10	Flowjo.com	RRID:SCR_008520
R (4.1.3)	r-project.org	RRID:SCR_001905
CellRanger (3.1.0)	10xgenomics.com	RRID:SCR_017344
FIJI (1.53f51)	imagej.net/software/fiji	RRID:SCR_002285
HTSeq (0.13.5)	https://htseq.readthedocs.io/en/release_0.11.1/count.html	RRID:SCR_011867
Cutadapt (3.0)	https://cutadapt.readthedocs.io/en/stable/	RRID:SCR_011841
R Libraries		
Scater (1.22.0)	bioconductor.org	RRID:SCR_015954
DropletUtils (1.14.2)	bioconductor.org	No RRID available
Liger (2.0.1)	bioconductor.org	RRID:SCR_018100
Seurat (4.1.0)	bioconductor.org	RRID:SCR_007322
Limma (3.50.3)	bioconductor.org	RRID:SCR_010943
SVA (3.35.2)	bioconductor.org	RRID:SCR_012836
EdgeR (3.36.0)	bioconductor.org	RRID:SCR_012802
Other		
Virtual reality and two-photon setup	(Leinweber et al., 2014, 2017)	DOI: 10.3791/50885, 10.1016/j.neuron.2017.08.036
OBIS LX 405 nm 250 mW Laser	Coherent	SKU 1303774
Titanium headplate	FMI/ETHZ workshop	N/A
Dental drill	Meisinger	N/A

660

661

Table S1. Statistics

All values are rounded to two significant decimals, except values smaller than 10^{-3}. The tests used, were two-sample independent t-test (t-test 2), one way ANOVA with a post-hoc TUKEY, and Monte Carlo simulations. Sample sizes were estimated based on previous experiments performed in the same behavioral paradigms (Attinger et al., 2017; Heindorf et al., 2018).

Reference	Description	Test	N_1 (mice)	N_2 (mice)	Unit	P-value/CI
Figure 1E	Linear regression (intercept)	F-test	417 (4)	N/A	Neurons	$<10^{-5}$
	Linear regression (slope)	F-test	417 (4)	N/A	Neurons	$<10^{-5}$
Figure 1F	Correlations comparisons					
	ANOVA (all conditions)	One-way	417 (4)	N/A	Neurons	0.021
	Mismatch (N1) vs Running (N2)	TUKEY	417 (4)	417 (4)	Neurons	0.087
	Running (N1) vs Grating (N2)	TUKEY	417 (4)	417 (4)	Neurons	0.62
	Mismatch (N1) vs Grating N2)	TUKEY	417 (4)	417 (4)	Neurons	0.021
Figure 2F	*Adamts2* population enrichment					
	Mismatch – Low photoconversion	Monte Carlo	267 (4)	N/A	Neurons	$<10^{-3}$
	Mismatch – Int photoconversion	Monte Carlo	556 (4)	N/A	Neurons	0.013
	Mismatch – High photoconversion	Monte Carlo	475 (4)	N/A	Neurons	$<10^{-5}$
	Running – Low photoconversion	Monte Carlo	1402 (4)	N/A	Neurons	0.3
	Running – Int photoconversion	Monte Carlo	1608 (4)	N/A	Neurons	0.94
	Running – High photoconversion	Monte Carlo	462 (4)	N/A	Neurons	0.35
	Grating – Low photoconversion	Monte Carlo	1148 (4)	N/A	Neurons	$<10^{-5}$
	Grating – Int photoconversion	Monte Carlo	1051 (4)	N/A	Neurons	0.0021
	Grating – High photoconversion	Monte Carlo	144 (3)	N/A	Neurons	0.0053
Figure 2G	Agmat population enrichment					
	Mismatch – Low photoconversion	Monte Carlo	856 (4)	N/A	Neurons	$<10^{-3}$
	Mismatch – Int photoconversion	Monte Carlo	1682 (4)	N/A	Neurons	0.038
	Mismatch – High photoconversion	Monte Carlo	768 (4)	N/A	Neurons	$<10^{-5}$
	Running – Low photoconversion	Monte Carlo	2175 (4)	N/A	Neurons	0.24
	Running – Int photoconversion	Monte Carlo	2619 (4)	N/A	Neurons	0.82
	Running – High photoconversion	Monte Carlo	885 (4)	N/A	Neurons	0.36
	Grating – Low photoconversion	Monte Carlo	1624 (4)	N/A	Neurons	$<10^{-5}$
	Grating – Int photoconversion	Monte Carlo	2731 (4)	N/A	Neurons	$<10^{-3}$
	Grating – High photoconversion	Monte Carlo	348 (3)	N/A	Neurons	0.23
Figure 2H	*Rrad* population enrichment					
	Mismatch – Low photoconversion	Monte Carlo	66 (4)	N/A	Neurons	0.74
	Mismatch – Int photoconversion	Monte Carlo	142 (4)	N/A	Neurons	0.56
	Mismatch – High photoconversion	Monte Carlo	74 (4)	N/A	Neurons	0.83
	Running – Low photoconversion	Monte Carlo	246 (4)	N/A	Neurons	0.7
	Running – Int photoconversion	Monte Carlo	264 (4)	N/A	Neurons	0.75
	Running – High photoconversion	Monte Carlo	103 (4)	N/A	Neurons	1.0
	Grating – Low photoconversion	Monte Carlo	109 (4)	N/A	Neurons	0.18
	Grating – Int photoconversion	Monte Carlo	136 (4)	N/A	Neurons	0.052
	Grating – High photoconversion	Monte Carlo	37 (3)	N/A	Neurons	$<10^{-3}$
Figure 4D	Correlation with visual flow					
	ANOVA (all conditions)	One-way	2268 (16)	N/A	Neurons	$<10^{-5}$
	Adamts2 (N1) vs *Agmat* (N2)	TUKEY	1016 (5)	538 (6)	Neurons	$<10^{-3}$
	Adamts2 (N_1) vs *Baz1a* (N_2)	TUKEY	1016 (5)	714 (5)	Neurons	$<10^{-5}$

	Agmat (N_1) vs *Baz1a* (N_2)	TUKEY	538 (6)	714 (5)	Neurons	0.0029
Figure 4E	Correlation with running					
	ANOVA (all conditions)	One-way	2268 (16)	N/A	Neurons	$<10^{-5}$
	Adamts2 (N1) vs Agmat (N2)	TUKEY	538 (5)	1016 (6)	Neurons	0.3
	Adamts2 (N_1) vs *Baz1a* (N_2)	TUKEY	538 (5)	714 (5)	Neurons	$<10^{-5}$
	Agmat (N_1) vs *Baz1a* (N_2)	TUKEY	1016 (6)	714 (5)	Neurons	0.0062

Reference	Description	Test	N_1 (mice)	N_2 (mice)	Unit	P-value/CI
Figure S4	Low exp. (N_1) vs High exp. (N_2)					
	AP.Adamts2.1 – *Adamts2*	t-test 2	3	3	Mice	0.0096
	AP.Adamts2.1 – *Agmat*	t-test 2	3	3	Mice	0.18
	AP.Adamts2.1 – *Baz1a*	t-test 2	3	3	Mice	063
	Ap.Agmat.1 – *Adamts2*	t-test 2	4	4	Mice	0.45
	Ap.Agmat.1 – *Agmat*	t-test 2	4	4	Mice	$<10^{-3}$
	Ap.Agmat.1 – *Baz1a*	t-test 2	4	4	Mice	0.015
	Ap.Baz1a.1 – *Adamts2*	t-test 2	2	4	Mice	0.63
	Ap.Baz1a.1 – *Agmat*	t-test 2	2	4	Mice	0.0032
	Ap.Baz1a.1 – *Baz1a*	t-test 2	2	4	Mice	0.02

Table S2. Number of mice per experiment

The table below lists the number of mice in each experiment.

Dataset	Number of mice
Mismatch-photoconversion, functional imaging (Figures 1D - 1F)	**4**
Mismatch-photoconversion, single-cell (Figures 2, S2)	**4**
Running-photoconversion, single-cell (Figures 2, S2)	**4**
Gratings-photoconversion, single-cell (Figures 2, S2)	**4**
AP.Adamts2.1-GFP, histology (Figures 3B and 3F)	**4**
AP.Agmat.1-GFP, histology (Figures 3C, 3E and 3F)	**2**
AP.Baz1a.1-GFP, histology (Figures 3D and 3F)	**2**
AP.Adamts2.1-GFP, bulk RNA-seq (Figures S4A, S4D and S4G)	**3**
AP.Agmat.1-GFP, bulk RNA-seq (Figures S4B, S4E and S4H)	**4**
AP.Baz1a.1-GFP, bulk RNA-seq (Figures S4C, S4F and S4I)	**4**
AP.Adamts2.1-GCaMP (Figures 4A - 4E, S5)	**5**
AP.Agmat.1-GCaMP (Figures 4A - 4E, S5)	**6**
AP.Baz1a.1-GCaMP (Figures 4A - 4E, S5)	**5**
Ef1α-GCaMP (Figures S5A and S5B)	**16**

Data and code availability

Software for controlling the two-photon microscope and preprocessing of calcium imaging data is available on https://sourceforge.net/projects/iris-scanning/. Raw data and code to generate all figures of this manuscript are available on https://data.fmi.ch/PublicationSupplementRepo/.

676 **Acknowledgements**

677 We thank all the members of the Keller lab for discussion and support as well as Tingjia Lu and Daniela
678 Gerosa for vector production**.** We thank Sirisha Aluri, Sebastien Smallwood and Hubertus Kohler for help
679 with the single-cell sequencing experiments as well as Koshika Yadava for help with the histology. This
680 project has received funding from the Swiss National Science Foundation, the Novartis Research
681 Foundation, and the European Research Council (ERC) under the European Union's Horizon 2020
682 research and innovation programme (grant agreement No. 865617).

683 **Author contributions**

684 SO designed and performed the experiments and analyzed the data. HO designed and performed
685 experiments. S.O. and G.K. wrote the manuscript.

686 **Declaration of Interests**

687 The authors declare no competing financial interests.

688

REFERENCES

Attinger, A., Wang, B., and Keller, G.B. (2017). Visuomotor Coupling Shapes the Functional Development of Mouse Visual Cortex. Cell *169*, 1291-1302.e14. https://doi.org/10.1016/j.cell.2017.05.023.

Audette, N.J., Zhou, W., and Schneider, D.M. (2021). Temporally precise movement-based predictions in the mouse auditory cortex.

Ayaz, A., Stäuble, A., Hamada, M., Wulf, M.A., Saleem, A.B., and Helmchen, F. (2019). Layer-specific integration of locomotion and sensory information in mouse barrel cortex. Nature Communications *10*, 1–14. https://doi.org/10.1038/s41467-019-10564-8.

Bastos, A.M., Usrey, W.M., Adams, R.A., Mangun, G.R., Fries, P., and Friston, K.J. (2012). Canonical Microcircuits for Predictive Coding. Neuron *76*, 695–711. https://doi.org/10.1016/j.neuron.2012.10.038.

Bugeon, S., Duffield, J., Dipoppa, M., Ritoux, A., Prankerd, I., Nicoloutsopoulos, D., Orme, D., Shinn, M., Peng, H., Forrest, H., et al. (2022). A transcriptomic axis predicts state modulation of cortical interneurons. Nature 1–9. https://doi.org/10.1038/s41586-022-04915-7.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav Brain Sci *36*, 181–204. https://doi.org/10.1017/S0140525X12000477.

Condylis, C., Ghanbari, A., Manjrekar, N., Bistrong, K., Yao, S., Yao, Z., Nguyen, T.N., Zeng, H., Tasic, B., and Chen, J.L. (2022). Dense functional and molecular readout of a circuit hub in sensory cortex. Science *375*, eabl5981. https://doi.org/10.1126/science.abl5981.

Corlett, P.R., Frith, C.D., and Fletcher, P.C. (2009). From drugs to deprivation: a Bayesian framework for understanding models of psychosis. Psychopharmacology *206*, 515–530. https://doi.org/10.1007/s00213-009-1561-0.

Dana, H., Sun, Y., Mohar, B., Hulse, B.K., Kerlin, A.M., Hasseman, J.P., Tsegaye, G., Tsang, A., Wong, A., Patel, R., et al. (2019). High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. Nat Methods *16*, 649–657. https://doi.org/10.1038/s41592-019-0435-6.

Dombeck, D.A., Khabbaz, A.N., Collman, F., Adelman, T.L., and Tank, D.W. (2007). Imaging Large-Scale Neural Activity with Cellular Resolution in Awake, Mobile Mice. Neuron *56*, 43–57. https://doi.org/10.1016/j.neuron.2007.08.003.

Eliades, S.J., and Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature *453*, 1102–1106. https://doi.org/10.1038/nature06910.

Fiser, A., Mahringer, D., Oyibo, H.K., Petersen, A.V., Leinweber, M., and Keller, G.B. (2016). Experience-dependent spatial expectations in mouse visual cortex. Nature Neuroscience *19*, 1658–1664. https://doi.org/10.1038/nn.4385.

Fletcher, P.C., and Frith, C.D. (2009). Perceiving is believing: a Bayesian approach to explaining the positive symptoms of schizophrenia. Nat Rev Neurosci *10*, 48–58. https://doi.org/10.1038/nrn2536.

723 Frith, C.D., Blakemore, S.-J., and Wolpert, D.M. (2000). Explaining the symptoms of schizophrenia:
724 Abnormalities in the awareness of action. Brain Research Reviews *31*, 357–363.
725 https://doi.org/10.1016/S0165-0173(99)00052-1.

726 Garner, A.R., and Keller, G.B. (2022). A cortical circuit for audio-visual predictions. Nat Neurosci *25*, 98–
727 105. https://doi.org/10.1038/s41593-021-00974-7.

728 Gorski, J.A., Talley, T., Qiu, M., Puelles, L., Rubenstein, J.L.R., and Jones, K.R. (2002). Cortical Excitatory
729 Neurons and Glia, But Not GABAergic Neurons, Are Produced in the Emx1-Expressing Lineage. J.
730 Neurosci. *22*, 6309–6314. https://doi.org/10.1523/JNEUROSCI.22-15-06309.2002.

731 Griffin, J.D., and Fletcher, P.C. (2017). Predictive Processing, Source Monitoring, and Psychosis. Annu Rev
732 Clin Psychol *13*, 265–289. https://doi.org/10.1146/annurev-clinpsy-032816-045145.

733 Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell RNA-seq
734 data using regularized negative binomial regression. Genome Biology *20*, 296.
735 https://doi.org/10.1186/s13059-019-1874-1.

736 Heindorf, M., and Keller, G.B. (2022). Reduction of layer 5 mediated long-range cortical communication
737 by antipsychotic drugs. 2022.01.31.478462. https://doi.org/10.1101/2022.01.31.478462.

738 Heindorf, M., Arber, S., and Keller, G.B. (2018). Mouse Motor Cortex Coordinates the Behavioral
739 Response to Unpredicted Sensory Feedback. Neuron *99*, 1040-1054.e5.
740 https://doi.org/10.1016/j.neuron.2018.07.046.

741 Hempel, C.M., Sugino, K., and Nelson, S.B. (2007). A manual method for the purification of fluorescently
742 labeled neurons from the mammalian brain. Nature Protocols *2*, 2924–2929.
743 https://doi.org/10.1038/nprot.2007.416.

744 Hodge, R.D., Bakken, T.E., Miller, J.A., Smith, K.A., Barkan, E.R., Graybuck, L.T., Close, J.L., Long, B.,
745 Johansen, N., Penn, O., et al. (2019). Conserved cell types with divergent features in human versus
746 mouse cortex. Nature *573*, 61–68. https://doi.org/10.1038/s41586-019-1506-7.

747 Jordan, M.I., and Rumelhart, D.E. (1992). Forward models: Supervised learning with a distal teacher.
748 Cognitive Science *16*, 307–354. https://doi.org/10.1016/0364-0213(92)90036-T.

749 Jordan, R., and Keller, G.B. (2020). Opposing Influence of Top-down and Bottom-up Input on Excitatory
750 Layer 2/3 Neurons in Mouse Primary Visual Cortex. Neuron *108*, 1194-1206.e5.
751 https://doi.org/10.1016/j.neuron.2020.09.024.

752 Jüttner, J., Szabo, A., Gross-Scherf, B., Morikawa, R.K., Rompani, S.B., Hantz, P., Szikra, T., Esposti, F.,
753 Cowan, C.S., Bharioke, A., et al. (2019). Targeting neuronal and glial cell types with synthetic promoter
754 AAVs in mice, non-human primates and humans. Nature Neuroscience *22*, 1345–1356.
755 https://doi.org/10.1038/s41593-019-0431-2.

756 Keller, G.B., and Hahnloser, R.H.R. (2009). Neural processing of auditory feedback during vocal practice
757 in a songbird. Nature *457*, 187–190. https://doi.org/10.1038/nature07467.

Keller, G.B., and Mrsic-Flogel, T.D. (2018). Predictive Processing: A Canonical Cortical Computation. Neuron *100*, 424–435. https://doi.org/10.1016/j.neuron.2018.10.003.

Keller, G.B., Bonhoeffer, T., and Hübener, M. (2012). Sensorimotor Mismatch Signals in Primary Visual Cortex of the Behaving Mouse. Neuron *74*, 809–815. https://doi.org/10.1016/j.neuron.2012.03.040.

Kiselev, V.Y., Yiu, A., and Hemberg, M. (2018). scmap: projection of single-cell RNA-seq data across data sets. Nat Methods *15*, 359–362. https://doi.org/10.1038/nmeth.4644.

Koster-Hale, J., and Saxe, R. (2013). Theory of Mind: A Neural Prediction Problem. Neuron *79*, 836–848. https://doi.org/10.1016/j.neuron.2013.08.020.

Lawson, R.P., Rees, G., and Friston, K.J. (2014). An aberrant precision account of autism. Frontiers in Human Neuroscience *8*. .

Lawson, R.P., Mathys, C., and Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. Nat Neurosci *20*, 1293–1299. https://doi.org/10.1038/nn.4615.

Leinweber, M., Zmarz, P., Buchmann, P., Argast, P., Hübener, M., Bonhoeffer, T., and Keller, G.B. (2014). Two-photon Calcium Imaging in Mice Navigating a Virtual Reality Environment. JoVE (Journal of Visualized Experiments) e50885. https://doi.org/10.3791/50885.

Leinweber, M., Ward, D.R., Sobczak, J.M., Attinger, A., and Keller, G.B. (2017). A Sensorimotor Circuit in Mouse Cortex for Visual Flow Predictions. Neuron *95*, 1420-1432.e5. https://doi.org/10.1016/j.neuron.2017.08.036.

Lun, A.T.L., Riesenfeld, S., Andrews, T., Dao, T.P., Gomes, T., participants in the 1st Human Cell Atlas Jamboree, and Marioni, J.C. (2019). EmptyDrops: distinguishing cells from empty droplets in droplet-based single-cell RNA sequencing data. Genome Biol *20*, 63. https://doi.org/10.1186/s13059-019-1662-y.

McCarthy, D.J., Campbell, K.R., Lun, A.T.L., and Wills, Q.F. (2017). Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics *33*, 1179–1186. https://doi.org/10.1093/bioinformatics/btw777.

Moeyaert, B., Holt, G., Madangopal, R., Perez-Alvarez, A., Fearey, B.C., Trojanowski, N.F., Ledderose, J., Zolnik, T.A., Das, A., Patel, D., et al. (2018). Improved methods for marking active neuron populations. Nat Commun *9*, 4440. https://doi.org/10.1038/s41467-018-06935-2.

O'Toole, S.M., Ferrer, M.M., Mekonnen, J., Zhang, H., Shima, Y., Ladle, D.R., and Nelson, S.B. (2017). Dicer maintains the identity and function of proprioceptive sensory neurons. Journal of Neurophysiology *117*, 1057–1069. https://doi.org/10.1152/jn.00763.2016.

Palmer, C.J., Seth, A.K., and Hohwy, J. (2015). The felt presence of other minds: Predictive processing, counterfactual predictions, and mentalising in autism. Consciousness and Cognition *36*, 376–389. https://doi.org/10.1016/j.concog.2015.04.007.

Picelli, S., Faridani, O.R., Björklund, Å.K., Winberg, G., Sagasser, S., and Sandberg, R. (2014). Full-length RNA-seq from single cells using Smart-seq2. Nature Protocols *9*, 171–181. https://doi.org/10.1038/nprot.2014.006.

Rao, R.P.N., and Ballard, D.H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci *2*, 79–87. https://doi.org/10.1038/4580.

Rupprecht, P., Carta, S., Hoffmann, A., Echizen, M., Blot, A., Kwan, A.C., Dan, Y., Hofer, S.B., Kitamura, K., Helmchen, F., et al. (2021). A database and deep learning toolbox for noise-optimized, generalized spike inference from calcium imaging. Nat Neurosci *24*, 1324–1337. https://doi.org/10.1038/s41593-021-00895-5.

Satija, R., Farrell, J.A., Gennert, D., Schier, A.F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. Nat Biotechnol *33*, 495–502. https://doi.org/10.1038/nbt.3192.

Saxena, A., Wagatsuma, A., Noro, Y., Kuji, T., Asaka-Oba, A., Watahiki, A., Gurnot, C., Fagiolini, M., Hensch, T.K., and Carninci, P. (2012). Trehalose-enhanced isolation of neuronal sub-types from adult mouse brain. BioTechniques *52*, 381–385. https://doi.org/10.2144/0000113878.

Sinha, P., Kjelgaard, M.M., Gandhi, T.K., Tsourides, K., Cardinaux, A.L., Pantazis, D., Diamond, S.P., and Held, R.M. (2014). Autism as a disorder of prediction. Proceedings of the National Academy of Sciences *111*, 15220–15225. https://doi.org/10.1073/pnas.1416797111.

Stanley, J., and Miall, R.C. (2007). Functional activation in parieto-premotor and visual areas dependent on congruency between hand movement and visual stimuli during motor-visual priming. NeuroImage *34*, 290–299. https://doi.org/10.1016/j.neuroimage.2006.08.043.

Tasic, B., Yao, Z., Graybuck, L.T., Smith, K.A., Nguyen, T.N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M.N., Viswanathan, S., et al. (2018). Shared and distinct transcriptomic cell types across neocortical areas. Nature *563*, 72–78. https://doi.org/10.1038/s41586-018-0654-5.

Ting, J.T., Daigle, T.L., Chen, Q., and Feng, G. (2014). Acute brain slice methods for adult and aging animals: Application of targeted patch clamp analysis and optogenetics. Methods in Molecular Biology *1183*, 221–242. https://doi.org/10.1007/978-1-4939-1096-0_14.

Trojanowski, N.F., Bottorff, J., and Turrigiano, G.G. (2021). Activity labeling in vivo using CaMPARI2 reveals intrinsic and synaptic differences between neurons with high and low firing rate set points. Neuron *109*, 663-676.e5. https://doi.org/10.1016/j.neuron.2020.11.027.

Welch, J.D., Kozareva, V., Ferreira, A., Vanderburg, C., Martin, C., and Macosko, E.Z. (2019). Single-Cell Multi-omic Integration Compares and Contrasts Features of Brain Cell Identity. Cell *177*, 1873-1887.e17. https://doi.org/10.1016/j.cell.2019.05.006.

Wu, Y.E., Pan, L., Zuo, Y., Li, X., and Hong, W. (2017). Detecting Activated Cell Populations Using Single-Cell RNA-Seq. Neuron *96*, 313-329.e6. https://doi.org/10.1016/j.neuron.2017.09.026.

826 Zhang, Y., Rózsa, M., Liang, Y., Bushey, D., Wei, Z., Zheng, J., Reep, D., Broussard, G.J., Tsang, A., Tsegaye,
827 G., et al. (2021). Fast and sensitive GCaMP calcium indicators for imaging neural populations.
828 2021.11.08.467793. https://doi.org/10.1101/2021.11.08.467793.

829 Zmarz, P., and Keller, G.B. (2016). Mismatch Receptive Fields in Mouse Visual Cortex. Neuron *92*, 766–
830 772. https://doi.org/10.1016/j.neuron.2016.09.057.

831

832