The logic of single-cell projections from visual cortex

Yunyun Han[1,2,3,*], Justus M Kebschull[4,5,*], Robert AA Campbell[3,*], Devon Cowan[3], Fabia Imhof[3], Anthony M Zador[5,†], Thomas D Mrsic-Flogel[3,6,†]

[1]Department of Neurobiology, School of Basic Medicine and Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[2]Institute for Brain Research, Collaborative Innovation Center for Brain Science, Huazhong University of Science and Technology, Wuhan, China

[3]Biozentrum, University of Basel, 4056 Basel, Switzerland

[4]Watson School of Biological Sciences, Cold Spring Harbor, NY 11724, USA

[5]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

[6]Sainsbury Wellcome Centre, University College London, London, UK

*These authors contributed equally to the work as first authors.

[†]These authors contributed equally to the work as senior authors.

Correspondence and requests should be addressed to Anthony Zador, zador@cshl.edu, or Thomas Mrsic-Flogel, t.mrsic-flogel@ucl.ac.uk

Neocortical areas communicate via extensive axonal projections, but the logic of information transfer is unresolved because the projections of individual neurons have not been systematically characterized. It is unknown whether individual neurons send projections only to single cortical areas, or instead distribute signals across multiple targets. Here we determined the projection patterns of 591 individual neurons in mouse primary visual cortex (V1) using whole-brain fluorescence-based axonal tracing and high-throughput DNA sequencing of genetically barcoded neurons (MAPseq). Projections were highly diverse and divergent, collectively targeting at least 18 cortical and subcortical areas. Most neurons target multiple cortical areas, often in non-random combinations, suggesting the existence of sub-classes of intracortical projection neurons. Thus the dominant mode of intracortical information transfer is not based on "one neuron – one target area" mapping. Instead, signals carried by individual cortical neurons are shared across subsets of target areas, and thus concurrently contribute to multiple functional pathways.

While the inputs received by a neuron drive its activity, its axonal projections determine its impact on other neurons. The axons of excitatory projection neurons residing in cortical layers 2/3, 5 and 6 of the neocortex are the main conduit by which signals are exchanged between cortical areas[1]. To date, no study has systematically investigated the principles by which individual neurons in any region of the mammalian neocortex distribute information to their targets. This knowledge is fundamental for establishing the logic of inter-areal communication and for constraining hypotheses about neural function and identification of putative sub-classes of neurons. Anatomical studies in macaque, cat and mouse, largely based on retrograde tracing methods, suggest an abundance of intracortical projection neurons in sensory neocortex whose axons appear to innervate single

target areas[2–6], raising the possibility that information may be distributed via ensembles of dedicated pathways that are functionally tailored to each target[6–12]. For example, neurons in mouse primary visual cortex (V1) that innervate the posteromedial (PM) or anterolateral (AL) area appear to match the spatial and temporal frequency preference of these target areas[7,13,14]. Similarly, neurons in the mouse primary somatosensory cortex projecting to either primary motor cortex or the secondary somatosensory area comprise largely non-overlapping populations with distinct physiological and functional properties[6,9,10]. These findings indicate that dedicated lines — specialized subpopulations of neurons that preferentially target a single downstream area (Fig. 1**a**, top) — may represent a fundamental mode of cortico-cortical communication. Alternatively, intra-cortically projecting neurons could broadcast to multiple targets[4,5,15–19], either randomly (Fig. 1**a**, middle), or by targeting specific sets of areas (Fig. 1**a**, bottom). These three models of cortical architecture have different implications for inter-areal communication underlying sensory processing in hierarchical networks. We therefore set out to distinguish among them, using two anterograde anatomical approaches, whole-brain fluorescence-based axonal tracing and MAPseq, to map the long-range axonal projection patterns of individual neurons in mouse primary visual cortex (V1), an area that distributes visual information to multiple cortical and subcortical targets[20–22].

Fluorescence-based single neuron tracing

We first traced single-neuron projections using whole-brain fluorescence-based axonal reconstructions. We used single-cell electroporation of a GFP-encoding plasmid to label up to six layer 2/3 cells in the right visual cortex of each mouse. After allowing 3-10 days for GFP expression we imaged the axonal projections of the labeled neurons by whole-brain serial two-photon tomography with 1x1x10 μm resolution[23,24] (Fig. 1**b**). We then traced each fluorescently-labeled cell (Fig. 1**c**,**d**; n = 71) and registered each brain to the Allen Reference Atlas[25] (Fig. 1**e**,**f**). To assess axonal

labelling with GFP, we electroporated neurons labelled retrogradely from the ipsilateral striatum, and in all cases observed axonal terminations therein (n = 9/9 cells; Extended Data Fig. 1), indicating a low false negative rate of filling axon collaterals to distal targets of V1 neurons. Nonetheless, to minimize any possible contribution of incomplete axonal filling, we excluded the contingent of reconstructed V1 neurons whose axon collaterals beyond V1 terminated abruptly without arborizing (n = 28; Extended Data Fig. 2; Supplemental Note 1), although the results below are robust to inclusion of these cells (Extended Data Fig. 2**e**). We did not exclude neurons with abrupt terminations of contralaterally projecting branches (compare ref[6]), instead restricting our analysis to ipsilaterally-projecting axons.

We analysed the ipsilateral projection patterns of 38 pyramidal neurons in layer 2/3, including 31 neurons in area V1 (Fig. 1**g**, Extended Data Fig. 3 and 4) and 7 neurons in nearby higher visual areas (Extended Data Fig. 5). Inspection of individual axonal arbors of V1 neurons revealed a high degree of projectional diversity with respect to the number and identity of target areas (Fig. 1**g**, Extended Data Fig. 3 and 4), which is obscured in bulk projection data[20,21] (Fig. 1**g**, top left).

Almost all layer 2/3 cells projected out of V1 (Fig. 1**h;** 97%, 30/31) to one or more of 18 target areas in the telencephalon (Fig. 1**i**), typically innervating nearby cortical areas but occasionally also projecting to anterior cingulate cortex, striatum (Extended Data Fig. 1) and amygdala. To mitigate errors arising both from technical noise in atlas registration and from subject-to-subject variability in the boundaries between brain areas, we excluded low-confidence "buffer zones" of 100 μm around the area boundaries from analysis, and define as a "target" only those areas that received over 1 mm of axonal input from an individual cell (see Methods). Eighty-five percent of all projection patterns appeared only once, highlighting the diversity of long-range projections.

The majority of reconstructed layer 2/3 projection neurons sent axon collaterals to more than one target area (77%, 23/30), with some targeting up to seven areas (Fig. 1**j**). Although individual neurons innervated different target areas with different axonal densities, and thus might influence the computations in one area more than another, we found that a large fraction of broadcasting cells innervated more than one target with comparable strengths (Fig. 1**k**). Moreover, the total length of axon scaled with the number of target areas (average length per brain area = 4.6 ± 2.2 mm), such that the innervation density per target was, on average, similar irrespective of how many targets an axon innervated (Extended Data Fig. 6**a**,**b**). The innervation in higher visual areas was most dense in layers 2/3 and 5, consistent with recent reports[26,27], often recapitulating the pattern of lateral axonal projections of layer 2/3 cells within V1 (Extended Data Fig. 6**c**-**h**).

Posteromedial (PM), posterolateral (P), postrhinal (POR) and lateromedial (LM) visual areas were the most common targets of V1 neurons (Fig. 1**l**). Even when the analysis was restricted to neurons that projected to at least one of six nearby cortical visual areas (LI, LM, AL, PM, AM, RL), we found that half projected to two or more of these areas (Extended Data Fig. 7**a**-**e**). The fraction of input provided by dedicated projection neurons to any area comprised no more than 25% of the total (Fig. 1**m**), and most target areas received no dedicated input. These conclusions were robust to changes in the size of the border exclusion zone between neighbouring areas and the minimum projection strength in the target area (Extended Data Fig. 7**f**-**h**). Similar to projections from V1, all seven reconstructed neurons whose cell bodies resided in nearby higher visual areas also projected to more than one target area (Extended Data Fig. 5). Our results thus reveal that most layer 2/3 neurons distribute information to multiple areas, rather than projecting to single targets.

Interestingly, cell body location within V1 was predictive of projection target for some recipient areas (Extended Fig. 8). Given the retinotopic organization of V1, this suggest that visual information from different parts of visual field may be preferentially distributed to specific target areas, consistent with recent findings[28].

High-throughput MAPseq tracing

We next investigated whether broadcasting cells choose their cortical target areas independently, or whether they target specific subsets of areas. While targeting different combinations of areas distinguishes individual V1 projection neurons (Fig. 1), their classification into putative sub-types requires a demonstration of higher-order projectional structure within the population. We define higher-order structure in terms of the connection patterns predicted by the per-neuron (first order) probability of projecting to each target. For example, if the probability of any given neuron projecting to area A is 0.5 and the probability of projecting to area B is also 0.5 then we would expect $P(A \cap B)=P(A)*P(B)=0.25$ of all neurons to project to both A and B if the decisions to target these areas are independent. Significant deviations from this expectation would indicate organization into non-random projection motifs. Probing for high order structure requires large datasets, because, if a sample size of N neurons is required to estimate the first order probabilities, then a sample size of N^2 is needed to estimate pairwise probabilities with comparable accuracy. Although single neuron reconstruction provides unrivaled spatial resolution, despite increases in throughput for data acquisition[17,29], the tracing of axons remains highly labor intensive.

We therefore used a higher throughput strategy, MAPseq[30], to obtain the required number of single neuron projections for higher-order statistical analysis. In a MAPseq experiment, hundreds or thousands of neurons are labeled uniquely with random RNA sequences (barcodes) by a single injection of a library of barcoded Sindbis virus (Supplemental Note 2). The barcodes are expressed

and then actively transported into the axonal processes of each labeled neuron, where they can be read out by high throughput barcode sequencing after dissection of potential target areas. The abundance of each barcode sequence in each area serves as a measure of the projection strength of the corresponding barcode-labeled neuron. MAPseq thus simultaneously maps the projections of all labeled neurons to dissected target areas, and therefore allows in-depth analysis of projections to a smaller set of targets.

We used MAPseq to map the projection patterns of 553 neurons from V1 to six higher visual areas — LI, LM, AL, AM, PM and RL — that can be identified reliably by intrinsic signal imaging *in vivo* and dissected *ex vivo* for barcode sequencing (Fig. 2**a**,**b**; Extended Data Fig. 9; see Methods). To avoid virus spillover from V1 into adjacent areas, we made small focal injections of the MAPseq virus to yield 100-200 traced cells per animal. Consistent with the analysis of fluorescence-based single neuron reconstructions restricted to the six higher visual areas (Fig. 2**c**, left), almost half (44%) of all MAPseq neurons projected to more than one area (Fig. 2**c**, right). Furthermore, the projection patterns obtained by fluorescence-based tracing were statistically indistinguishable from those obtained by MAPseq (using a bootstrap procedure; see Supplemental Note 3), whereas randomly generated neurons with projection strengths sampled from a uniform distribution were markedly different (Fig. 2**d**). Thus the findings from the MAPseq dataset were consistent with those from single neuron tracing.

We first catalogued the diversity of single neuron projection patterns from V1 to six higher visual areas by unsupervised clustering of the MAPseq dataset (k-means clustering with a cosine distance metric). These projectional data were best described by eight clusters (Fig. 2**e**, Extended Data Fig. 10), of which all but one contained cells targeting more than one area. The most common partners for broadcasting neurons were LM and PM, consistent with the fact that a large fraction of neurons

153 targeted these areas and the suggestion of LM[22] and PM as integrative hubs of V1 signals, akin to
154 monkey V2 (Fig. 2**f**).

155 To uncover the existence of non-random projection motifs in the MAPseq dataset, we measured
156 the likelihood of specific bi-, tri- or quadfurcations, by comparing them to expected probabilities
157 of these divergent projections (assuming independence between each projection type; Fig. 3**a**,**b**).
158 This analysis identified six projection motifs that were significantly over- or under-represented
159 after a correction for multiple comparisons (Bonferroni adjustment; Fig. 3**b**,**c**). Together, these six
160 projection motifs represented 73% of all broadcasting cells identified by MAPseq. Therefore the
161 majority of V1 cells projecting to multiple target areas do so in a non-random manner, suggesting
162 that broadcasting motifs reflect several sub-classes of projection neurons for divergent information
163 transfer from V1 to higher visual areas.

164 The most under-represented broadcasting motif was the bifurcation between areas PM and AL
165 (Fig. 3**d**). These two areas exhibit distinct visual response properties[13,14] and receive functionally
166 specialized input from V1[7], consistent with the idea of exclusive projections from V1 into these
167 areas. Moreover, the under-represented population of neurons that do project to both PM and AL
168 was further split into two groups according to projection strength, one that primarily innervates
169 PM and another that primarily innervates AL (Fig. 3**d**). A second under-represented motif is the
170 bifurcation between PM and LM (Fig. 3**e**). In contrast to the PM-AL bifurcation, however, the
171 detected PM-LM projecting neurons do not separate cleanly into two classes. Our findings there-
172 fore provide an anatomical substrate for the previously reported functional dichotomy of areas AL
173 and PM, and suggest that a few 'dedicated' output channels can co-exist with a preponderance of
174 broadcasting cells co-innervating multiple targets.

In addition to the two under-represented projectional motifs, we also identified four over-represented motifs, i.e. combinations of target areas receiving more shared input from individual V1 neurons than expected from first-order projection statistics (Fig. 3**f**-**h**). Cells jointly innervating PM and AM were significantly more abundant than expected by chance (Fig. 3**f**). Resolving the projection strengths within this motif revealed two subpopulations of neurons, one innervating PM more than AM, the other innervating the two areas with similar strength. Moreover, neurons bifurcating to LM and AL were also highly over-represented (Fig. 3**g**) and comprised the most abundant class of broadcasting cells (Fig. 3**b**). The most significantly over-represented trifurcation motif was the projection to PM, LM and LI, comprising a relatively homogenous population that projects to LM and PM with similar strengths while slightly less to LI (Fig. 3**h**). Finally, we discovered the over-representation of the PM-AM-RL trifurcation, but it appeared only rarely in our dataset (Fig. 3**b**). These motifs did not arise from false negatives (undetected connections) or false positives (Supplemental Note 4; Extended Data Fig. 2**f**).

These projectional data have implications for the categorization of higher visual areas into putative streams of visual processing in mouse neocortex. Areas AL and PM on the one hand, and LM and LI on the other, have been suggested to belong to dorsal and ventral processing streams in the mouse visual system, respectively[31–33]. Given that these areas receive a high degree of shared input (e.g. LM-PM bifurcation, even if underrepresented, still abundant; AL-LM bifurcation; PM-LM-LI trifurcation), such a distinction is unlikely to originate as a result of segregated V1 input into these areas.

Discussion

In summary, our results reveal some of the principles by which single neurons in one cortical area distribute information to downstream target areas. Almost all layer 2/3 pyramidal cells projected

198 outside of V1, indicating that V1 neurons concurrently engage in local and distal computations.
199 We found that the single neuron projections beyond V1 were highly diverse, innervating up to
200 seven targets, predominantly in specific, non-random combinations (Extended Data Fig. 10**g**,**f**).
201 These results suggest a functional specialization of subpopulations of projection cells beyond 'one
202 neuron – one target area' mapping.

203 The fraction of neurons in V1 that broadcast information to multiple targets is considerably greater
204 than previously documented using retrograde methods[2,5,16]. This difference is unlikely caused by
205 differences in the sensitivity with which these approaches detect the projections patterns of indi-
206 vidual cells. Instead, anterograde tracing maps projections to many or all targets simultaneously,
207 whereas retrograde tracing typically probes only two or three potential target sites at a time. Be-
208 cause the fraction of neurons projecting to any pair of targets selected for retrograde tracing is
209 relatively low (typically <10%), most neurons will not be doubly labeled in any given experiment;
210 only by sampling many potential targets in a single experiment can the true prevalence of broad-
211 casting be uncovered. Indeed, if we simulate double retrograde tracing based on our MAPseq re-
212 sults, the fractions of bifurcating neurons are comparable to those observed using retrograde meth-
213 ods in primates[2,5,16,18] (Supplemental Table 1).

214 We speculate that dedicated projection neurons — which comprise the minority of neurons in V1
215 — convey specialized visual information tailored to their target area, as suggested previously[6–11].
216 Indeed, the most under-represented projection motif from V1, the PM-AL bifurcation, innervates
217 two target areas with distinct preferences for visual features[13,14]. In contrast, we suggest that the
218 majority of cells encode information that is shared and in a form suitable for generating visual
219 representations or multimodal associations across subsets of areas. Indeed, those target areas that

220 are preferentially co-innervated by broadcasting neurons appear to have more similar visual re-
221 sponse properties[13,14]. Broadcasting cells may also coordinate activity among the subset of areas
222 they co-innervate, thus providing a signal that links different processing streams. The divergent
223 nature of signal transmission from a primary sensory cortex to its targets may thus help constrain
224 models of hierarchical sensory processing. The existence of distinct projection motifs that either
225 avoid or favor subsets of target areas suggests the existence of sub-types of intracortical projection
226 neurons and raises the question of how these specific, long-range connectivity patterns are estab-
227 lished during development.

References

1. Harris, K. D. & Shepherd, G. M. G. The neocortical circuit: themes and variations. *Nat. Neurosci.* **18,** 170–181 (2015).

2. Nakamura, H., Gattass, R., Desimone, R. & Ungerleider, L. The modular organization of projections from areas V1 and V2 to areas V4 and TEO in macaques. *J. Neurosci.* **13,** (1993).

3. Segraves, M. & Innocenti, G. Comparison of the distributions of ipsilaterally and contralaterally projecting corticocortical neurons in cat visual cortex using two fluorescent tracers. *J. Neurosci.* **5,** (1985).

4. Rockland, K. S. Collateral branching of long-distance cortical projections in monkey. *J. Comp. Neurol.* **521,** 4112–4123 (2013).

5. Sincich, L. C. & Horton, J. C. Independent Projection Streams from Macaque Striate Cortex to the Second Visual Area and Middle Temporal Area. *J. Neurosci.* **23,** 5684–5692 (2003).

6. Yamashita, T. *et al.* Membrane Potential Dynamics of Neocortical Projection Neurons Driving Target-Specific Signals. *Neuron* **80,** 1477–1490 (2013).

7. Glickfeld, L. L., Andermann, M. L., Bonin, V. & Reid, R. C. Cortico-cortical projections in mouse visual cortex are functionally target specific. *Nat. Neurosci.* **16,** 219–226 (2013).

8. Sato, T. R. & Svoboda, K. The Functional Properties of Barrel Cortex Neurons Projecting to the Primary Motor Cortex. *J. Neurosci.* **30,** 4256–4260 (2010).

9. Chen, J. L., Carta, S., Soldado-Magraner, J., Schneider, B. L. & Helmchen, F. Behaviour-dependent recruitment of long-range projection neurons in somatosensory cortex. *Nature*

499, 336–340 (2013).

10. Yamashita, T. & Petersen, C. C. H. Target-specific membrane potential dynamics of neocortical projection neurons during goal-directed behavior. *Elife* **5,** 2221 (2016).

11. Movshon, J. A. & Newsome, W. T. Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. *J. Neurosci.* **16,** 7733–7741 (1996).

12. Nassi, J. J. & Callaway, E. M. Parallel processing strategies of the primate visual system. *Nat. Rev. Neurosci.* **10,** 360–372 (2009).

13. Andermann, M. L., Kerlin, A. M., Roumis, D. K., Glickfeld, L. L. & Reid, R. C. Functional specialization of mouse higher visual cortical areas. *Neuron* **72,** 1025–1039 (2011).

14. Marshel, J. H. H., Garrett, M. E. E., Nauhaus, I. & Callaway, E. M. M. Functional Specialization of Seven Mouse Visual Cortical Areas. *Neuron* **72,** 1040–1054 (2011).

15. Massé, I. O., Régnier, P. & Boire, D. in *Axons and Brain Architecture* 93–116 (2016). doi:10.1016/B978-0-12-801393-9.00005-0

16. Bullier, J. & Kennedy, H. Axonal bifurcation in the visual system. *Trends Neurosci.* **10,** 205–210 (1987).

17. Economo, M. N. *et al.* A platform for brain-wide imaging and reconstruction of individual neurons. *Elife* **5,** 13 (2016).

18. Weisenhorn, D. M. V., Ilung, R. B. & Spatz, W. B. Morphology and connections of neurons in area 17 projecting to the extrastriate areas mt and 19DM and to the superior colliculus in the monkey Callithrix jacchus. *J. Comp. Neurol.* **362,** 233–255 (1995).

19. Ding, S.-L., Van Hoesen, G. & Rockland, K. S. Inferior parietal lobule projections to the presubiculum and neighboring ventromedial temporal cortical areas. *J. Comp. Neurol.* **425,**

510–530 (2000).

20. Zingg, B. *et al.* Neural networks of the mouse neocortex. *Cell* **156,** 1096–1111 (2014).

21. Oh, S. W. *et al.* A mesoscale connectome of the mouse brain. *Nature* **508,** 207–214 (2014).

22. Wang, Q. & Burkhalter, A. Area map of mouse visual cortex. *J. Comp. Neurol.* **502,** 339–357 (2007).

23. Ragan, T. *et al.* Serial two-photon tomography for automated ex vivo mouse brain imaging. *Nat. Methods* **9,** 255–258 (2012).

24. Osten, P. & Margrie, T. W. Mapping brain circuitry with a light microscope. *Nat. Methods* **10,** 515–523 (2013).

25. Lein, E. S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445,** 168–176 (2007).

26. D'Souza, R. D., Meier, A. M., Bista, P., Wang, Q. & Burkhalter, A. Recruitment of inhibition and excitation across mouse visual cortex depends on the hierarchy of interconnecting areas. *Elife* **5,** e19332 (2016).

27. Yang, W., Carrasquillo, Y., Hooks, B. M., Nerbonne, J. M. & Burkhalter, A. Distinct Balance of Excitation and Inhibition in an Interareal Feedforward and Feedback Circuit of Mouse Visual Cortex. *J. Neurosci.* **33,** 17373–17384 (2013).

28. Zhuang, J. *et al.* An extended retinotopic map of mouse cortex. *Elife* **6,** e18372 (2017).

29. Gong, H. *et al.* High-throughput dual-colour precision imaging for brain-wide connectome with cytoarchitectonic landmarks at the cellular level. *Nat. Commun.* **7,** 12142 (2016).

30. Kebschull, J. M. *et al.* High-Throughput Mapping of Single-Neuron Projections by

Sequencing of Barcoded RNA. *Neuron* **91,** 975–987 (2016).

31. Wang, Q., Sporns, O. & Burkhalter, A. Network Analysis of Corticocortical Connections Reveals Ventral and Dorsal Processing Streams in Mouse Visual Cortex. *J. Neurosci.* **32,** (2012).

32. Smith, I. T., Townsend, L. B., Huh, R., Zhu, H. & Smith, S. L. Stream-dependent development of higher visual cortical areas. *Nat. Neurosci.* **20,** 200–208 (2017).

33. Murakami, T., Matsui, T. & Ohki, K. Functional Segregation and Development of Mouse Higher Visual Areas. *J. Neurosci.* **37,** 9424–9437 (2017).

34. Pecka, M., Han, Y., Sader, E. & Mrsic-Flogel, T. D. Experience-Dependent Specialization of Receptive Field Surround for Selective Coding of Natural Scenes. *Neuron* **84,** 457–469 (2014).

Acknowledgements

We would like to thank Ashley Juavinett, Longwen Huang, Sonja Hofer and Petr Znamenskiy for comments on the manuscript. Funding sources: National Institutes of Health (5RO1NS073129 to A.M.Z., 5RO1DA036913 to A.M.Z.); Brain Research Foundation (BRF-SIA-2014-03 to A.M.Z.); IARPA (MICrONS to A.M.Z.); Simons Foundation (382793/SIMONS to A.M.Z.); Paul Allen Distinguished Investigator Award (to A.M.Z.); PhD fellowship from the Boehringer Ingelheim Fonds (to J.M.K.); PhD fellowship from the Genentech Foundation (to J.M.K); Natioanl Natural Science Foundation of China (NSFC 31600847 to Y.H.); European Research Council (NeuroV1sion 616509 to T.D.M.-F.), and Swiss National Science Foundation (SNSF 31003A_169802 to T.D.M.-F.).

Author contributions

Y.H. generated the dataset for fluorescence-based axonal tracing. D.C. and Y.H. traced the cells. R.A.A.C. analyzed the serial 2-P imaging data and axonal projection patterns. J.M.K. and F.I. collected the MAPseq dataset. J.M.K. and A.M.Z. performed the analysis of projection patterns. J.M.K., T.D.M-F. and A.M.Z. wrote the paper.

Author information

The authors declare no conflict of interests.

Figures

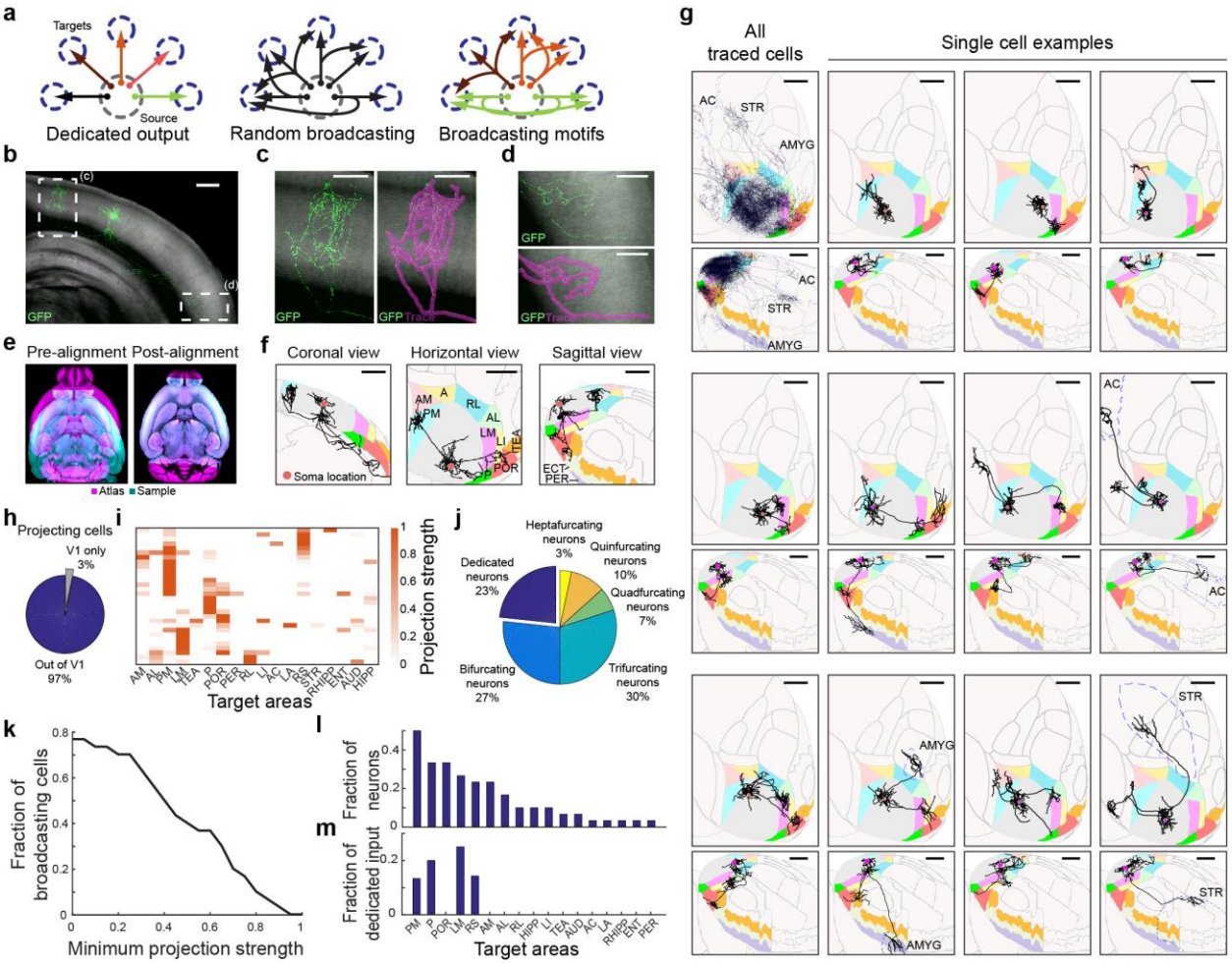


Figure 1: Brain-wide single-cell tracing reveals the diversity of axonal projection patterns of layer 2/3 V1 neurons, with most cells projecting to more than one target area.

(**a**) Three hypothetical modes of inter-areal information transfer from one area to its multiple targets. Neurons (arrows) could each project to a single area (top) or to several areas either randomly (middle) or in predefined projection patterns (bottom). (**b**) Maximum projection of a representative example GFP-filled neuron coronal view acquired by serial-section 2-photon microscopy. Autofluorescence from the red channel is used to show the brain's ultrastructure (gray background).

Scale bar = 600 μm. N = 71. (**c**-**d**) Higher magnification of the medial (**c**) and lateral (**d**) axonal arbor of the example cell. Scale bar = 300 μm. (**e**) Horizontal section through a sample brain (cyan) and Allen reference atlas (ARA; magenta) before (left) and after (right) rigid and non-rigid transformation of the brain to the atlas. (**f**) Coronal, sagittal and horizontal projections of the traced example cell overlaid in ARA space. Target cortical areas are coloured as indicated. Areas: A, anterior; AL: anterolateral; AM: anteromedial; LI: lateroitermediate; LM: lateral; P: posterior; PM: posteromedial; POR: postrhinal; RL: rostrolateral; TEA: temporal association; ECT: ectorhinal; PER: perirhinal. Scale bar = 1 mm. (**g**) Overlay of all traced single neurons (top left) and 11 example cells in Allen Reference Atlas (ARA) space; horizontal view (upper panel) and sagittal view (lower panel). Dashed outlines label non-visual target areas: AC: anterior cingulate cortex; STR, striatum; AMYG: amygdala. Note that these images are for illustration purposes only because a 2D projection cannot faithfully capture the true axonal arborisation pattern in 3D. Scale bar = 1 mm. (**h**) Pie chart illustrating the fraction of traced single neurons that project to at least one target area outside V1, where at least 1 mm of axonal innervation is required for an area to be considered a target. (**i**) Projection pattern of all GFP-filled V1 neurons targeted randomly (upper panel, n=31). The colour-code reflects the projection strengths of each neuron, determined as axon length per target area, normalized to the axon length in the target area receiving the densest innervation. Only brain areas that receive input form at least one neuron, as well as striatum, are shown. Areas: AUD: auditory cortex; ENT: entorhinal; HIPP: hippocampus; LA: lateral amygdala; RHIPP: retrohippocampal region; RS: retrosplenial. (**j**) The number of projection targets for every neuron that projects out of V1. (**k**) The proportion of cells targeting more than one area, when projection targets that receive projections weaker than the indicated projection strength are ignored. For each neuron,

projection strengths are normalized to axon length in the target area receiving the densest innervation. (**l**) The fraction of neurons projecting to each of the 18 target areas of V1. (**m**) The fraction of neurons innervating a single target area ('dedicated' projection neurons) out of all neurons that innervate that area.

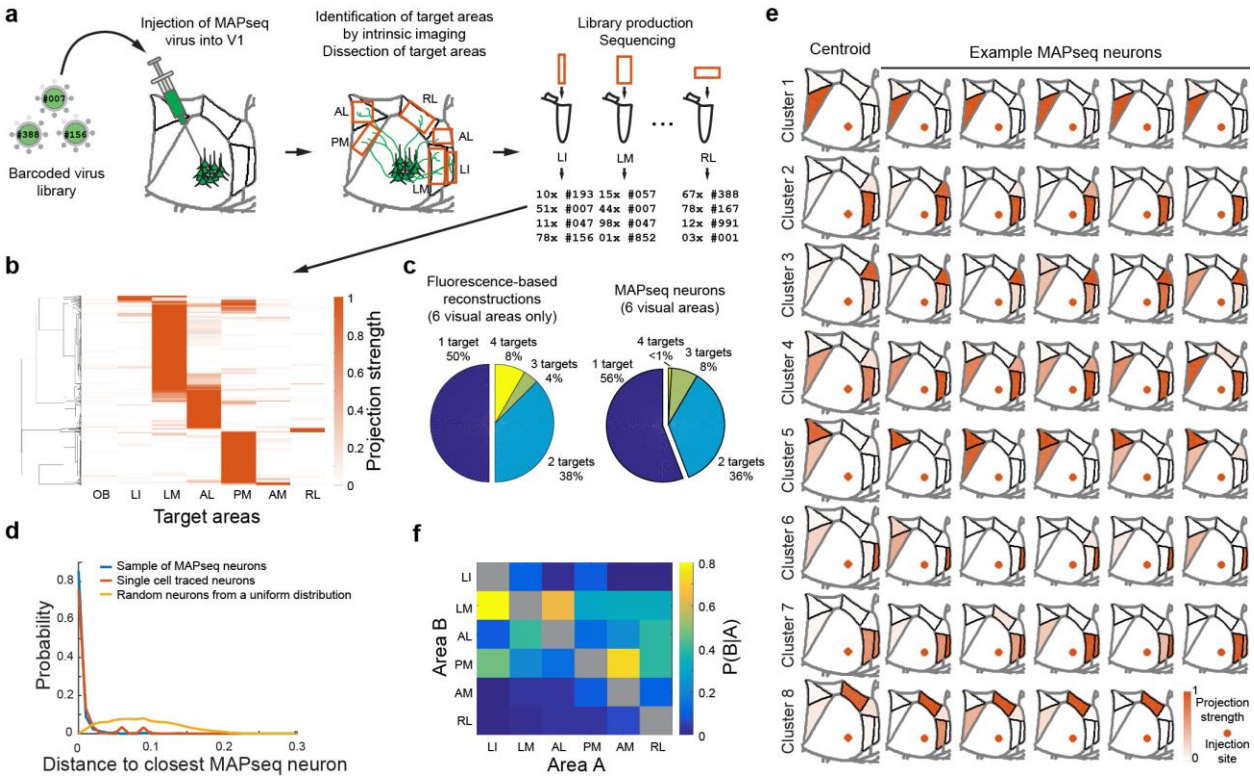


Figure 2: MAPseq projection mapping reveals a diversity of projection motifs. (**a**) Overview of the MAPseq procedure. Six target areas were chosen for analysis: LI, LM, AL, PM, AM and RL. (**b**) Projection strength in the six target areas, as well as the olfactory bulb (OB) as a negative control, of 553 MAPseq-mapped neurons. Projection strengths per neuron are defined as the number of barcode copies per area, normalized to the efficiency of sequencing library generation and to the neuron's maximum projection strength (n=4 mice). (**c**) Number of projection targets of V1 neurons when considering the six target areas only, based on the fluorescence-based axonal reconstructions (left) or the MAPseq data (right). (**d**) Distribution of cosine distances obtained by a bootstrapping procedure (1000 repeats) between MAPseq neurons (blue), fluorescence-based single neuron reconstructions and MAPseq neurons (orange), or random neurons (with projection strengths sampled from a uniform distribution) and MAPseq neurons (yellow). The distance distributions obtained from MAPseq neurons and fluorescence-based single-neuron reconstructions

are statistically indistinguishable (Kolmogorov-Smirnov one-sided two sample test; $p=0.94$; $\alpha=0.05$), whereas the distributions obtained from both MAPseq neurons or fluorescence-based reconstructed neurons are statistically different form the distribution obtained using random neurons (Kolmogorov-Smirnov two sample test; $p<10^{-3}$; $\alpha=0.05$). (**e**) Centroids and example cells for eight clusters obtained by k-means clustering of all MAPseq cells using a cosine distance metric. Target areas are coloured to indicate the projection strength of the plotted neuron. Projections strengths are normalize as in (**b**). (**f**) The probability of projecting to one area (Area A) given that the same neuron is projecting to another area (Area B) based on the MAPseq dataset.

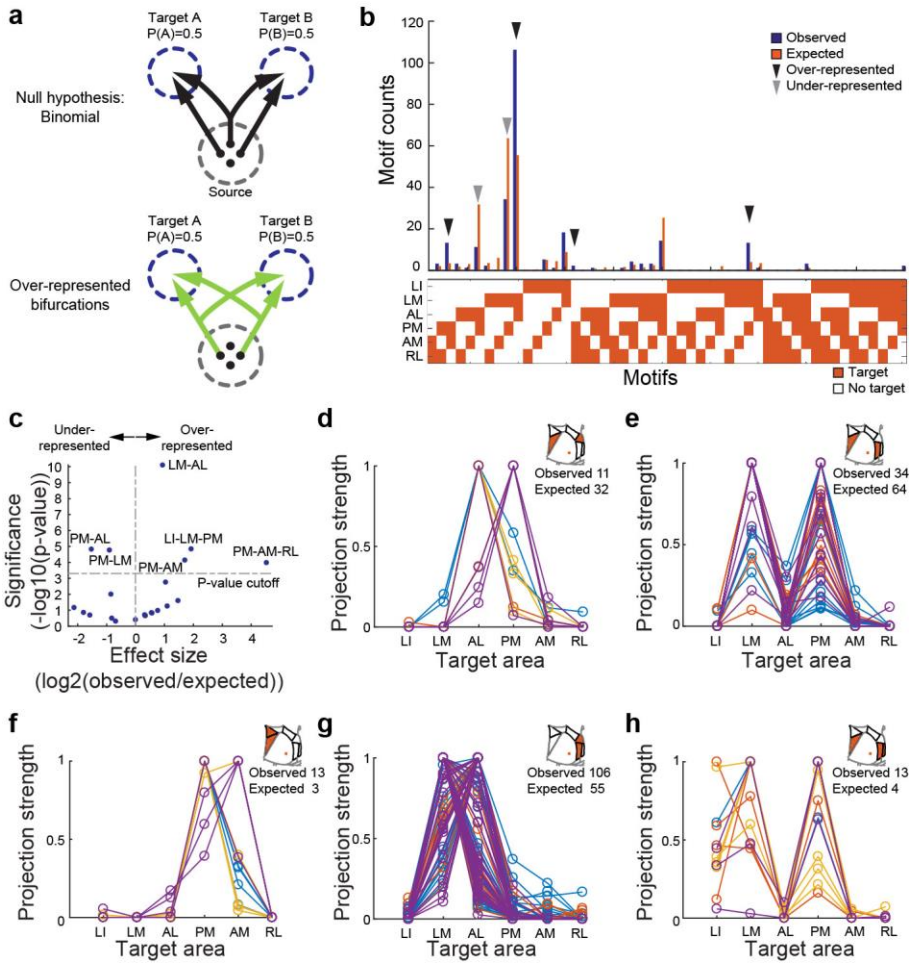


Figure 3: Over- and under-represented projection motifs of neurons in primary visual cortex. (**a**) The null hypothesis of independent projections to two target areas (left) and an example deviation (over-represented bifurcation) from the null hypothesis (right). (**b**) The observed and expected abundance of all possible bi-, tri- and quadfrucation motifs in the MAPseq dataset. Significantly over- or under-represented motifs, based on a binomial test with Bonferoni correction (see Methods), are indicated by black and grey arrowheads, respectively. N=553 neurons from 4 animals. (**c**) Statistical significance of over- and under-represented broadcasting motifs and associated effect sizes, based on a binomial test with Bonferoni correction (see Methods). N=553 neurons from 4 animals. (**d**-**h**) The projection strengths of the individual neurons (one per line) giving rise to the six under-represented (**d**,**e**) or over-represented (**f**-**h**) projection motifs. For each neuron,

390 the projections strength in each target area is normalized to the neuron's maximum projection
391 strength. Lines of the same color represent neurons mapped in the same brain (n=4 mice).

Methods

The anatomical single-cell tracing experiments were conducted at The Biozentrum, University of Basel, Switzerland. We licensed and performed all experimental procedures in accordance with Basel Canton animal welfare guidelines using both male and female adult (>8 weeks of age) C57BL/6 mice. Detailed protocols and all software are available at: *http://mouse.vision/han2017*

Fluorescence-based single neuron tracing

Two-photon guided single-cell electroporation. We performed surgery as described previously[34]. Briefly, we anesthetized animals with a mixture of fentanyl (0.05 mg kg^{-1}), midazolam (5 mg kg^{-1}) and medetomidine (0.5 mg kg^{-1}), and maintained stable anaesthesia by isoflurane (0.5% in O_2). We performed all electroporations on a custom linear scanning 2-photon microscope, equipped to image both a green and a red channel and running ScanImage 5.1[35]. For electroporation we used a patch pipette (12-16 MΩ) filled with plasmid DNA (pCAG-eGFP (Addgene accession 11150) or pAAV-EF1a-eGFP-WPRE (generous gift from Botond Roska; sequence file can be found in the Supplemental Materials, 100 ng/µl) and AlexaFluor 488 (50 µM) in intracellular solution, and delivered electroporation pulses (100 Hz, -14 V, 0.5 ms for 1 s) with an Axoporator 800A (Molecular Probes) when pushed against a target cell. We verified successful electroporation by dye filling of the cell body, and then sealed the skull with a chronic window using 1.5% agarose in HEPES-buffered artificial cerebrospinal fluid and a cover slip. We finally confirmed plasmid expression two days after electroporation by visualization of GFP epifluorescence through the chronic imaging window. Three to 10 days after electroporation, we transcardially perfused anesthetized mice with 10 ml 0.9% NaCl followed by 50 ml 4% paraformaldehyde in 0.1 M phosphate buffer (pH 7.4). We removed the brains from the skull and post-fixed them in 4%

paraformaldehyde overnight at 4 °C. We then stored the fixed brains in PBS at 4 °C until imaging with serial-section 2-photon tomography.

Serial-section 2-photon tomography. We embedded the fixed brains in 5% oxidised agarose (derived from Sigma Type I agarose) and covalently cross-linked the brain to the agarose by incubation in an excess of 0.5–1% sodium borohydrate ($NaBH_4$, Sigma) in 0.05 M sodium borate buffer overnight at 4°C. We then imaged embedded brains using a TissueVision 2-photon scanning microscope[23,36], which cut physical sections of the entire brain every 50 μm coronally, and acquired optical sections every 10 μm in two channels (green channel: 500-560 nm; red channel: 560-650 nm) using 940 nm excitation laser light (Mai Tai eHP, Spectraphysics). Each imaged section is formed from overlapping 800x800 μm "tiles". We imaged with a resolution of 1 μm in x and y and measured an axial point spread function of ~5 μm FWHM using ScanImage 5.1.

Image processing and cell tracing. We stitched raw image tiles using a custom MATLAB-based software, *StitchIt*. *StitchIt* applies illumination correction based on the average tiles for each channel and optical plane, and subsequently stitches the illumination-corrected tiles from the entire brain. We then navigated through the stitched brain space using *MaSIV* (https://github.com/alexanderbrown/masiv), a MATLAB-based viewer for very large 3-D images, and traced axons using a custom, manual neurite-tracer extension for *MaSIV*. The tracer was not blinded, as no comparison across experimental conditions was to be performed. No power calculations were performed.

To assign each voxel of the imaged brains to a brain area, we segmented each brain using areas defined by the Allen Reference Atlas (ARA, Common Coordinate Framework v3; © 2015 Allen Institute for Brain Science. Allen Brain Atlas API. Available from: *brain-map.org/api/index.html*), after smoothing with a single pass of an SD=0.5 voxel Gaussian kernel using the Nifty "seg-maths"

436 tool as described previously[37]. Briefly, we downsampled one imaging channel to a voxel size of
437 25 µm and converted it to MHD format using *StitchIt*. We then registered the volume to the ARA
438 average template brain using Elastix[38] by applying rigid affine transformation followed by non-
439 rigid deformation with parameters as described previously[39,40]. We examined registration quality
440 using a custom Python/PyQt5 application, *Lasagna,* which overlays the Allen template brain and
441 the registered sample brain and is extendable to allow the overlay of traced cells, or the overlay of
442 ARA area borders onto a down-sampled brain. In order to transform the traced cells into ARA
443 space (sample to ARA) we calculated the inverse transform to the one calculated by Elastix (ARA
444 to sample) and applied this to the traced points.

445 **Analysis of traced neurons.** To avoid potential incomplete filling of neurons from biasing the
446 results of our analyses, we excluded cells with non-arborizing primary branches in the ipsilateral
447 hemisphere from the analysis. Out of a total of 71 traced cells, we excluded 28 cells that exhibited
448 abrupt, non-callosal terminations, as well as 5 cells that were backlabeled from the striatum, thus
449 restricting our analysis to ipsilateral projection patterns of 31 cells in V1 and 7 in other higher
450 visual areas. Moreover, axonal branches terminating contralaterally or after entering the corpus
451 callosum were considered as callosal terminations and were included in the analysis (compare
452 ref[6]). We calculated the first order projection statistics only using the ARA-registered cells that
453 satisfied these criteria. To reduce any artifacts associated with ARA registration or individual brain
454 variability in boundaries between brain areas, we excluded any axon within 50 µm from any brain
455 area boundary from the analysis. We then calculated the projection strength of each neuron to each
456 area as the total length of axon of that neuron in an area. To determine the number of projection
457 targets for every cell, we used a minimum projection strength of 1 mm axon length per target area.

458 **MAPseq**

MAPseq sample processing. To define the V1 injection site and target higher visual areas LI, LM, AL, PM, AM and RL, we used optical imaging of intrinsic signals as previously described[13,41]. Briefly, we first implanted a customized head plate and then thinned the skull to increase its transparency. After 2-3 days of recovery, we sedated the mice (chlorprothixene, 0.7 mg/kg) and lightly anesthetized them with isoflurane (0.5-1.5% in O_2), delivered via a nose cone. We illuminated visual cortex with 700 nm light split from an LED source into 2 light guides, performing imaging with a tandem lens macroscope focused 250-500 µm below the cortical surface and a bandpass filter centered at 700 nm with 10 nm bandwidth (67905; Edmund optics). We acquired images at 6.25 Hz with a 12- bit CCD camera (1300QF; VDS Vosskühler), frame grabber (PCI-1422; National Instruments) and custom software written in LabVIEW (National Instruments). We visually stimulated the contralateral eye of mice with a monitor placed at a distance of 21 cm and presented 25-35° patches of 100% contrast square wave gratings with a temporal frequency of 4 Hz and a spatial frequency of 0.02 cycles per degree for 2 s followed by 5 s of grey screen (mean luminance of 46 cd/m^2). To establish a coarse retinotopy of the targeted area, we alternated the position of the patches: we used two different elevations (approx. 0 and 20°) and two different azimuths (approx. 60 and 90°); at each position we acquired at least 17 trials. We obtained intrinsic signal maps by averaging the responses during the stimulation time using ImageJ (National Institute of Mental Health, NIH) and mapping the location of the estimated spots of activation onto a previously acquired blood vessel picture.

We then pressure injected (Picospritzer III, Parker) 100 nl of 1x10^{10} GC/ml barcoded MAPseq Sindbis virus[30] with a diversity of >8x10^6 different barcode sequences unilaterally at a depth of 100-200 µm from the brain surface into V1 of a total of four 8-10 week old C57BL/6 females. In

addition, we labeled the six higher visual areas by placing a DiI-coated micropipette into retinotopically matched positions according to intrinsic signal maps. For this, we allowed 2-5 µl of a 2.5 mg/ml DiI (Invitrogen D3911) in ethanol solution to dry on the outside of a pulled micropipette tip until some DiI crystals were visible. Mice were sacrificed 44-48 hours post-injection by decapitation, and their brain immediately extracted and flash frozen on dry ice.

We cut 180 µm thick coronal sections using a cryostat at -10ºC blade and sample holder temperature, and melted each slice onto a clean microscope slide before rapidly freezing it on dry ice again. We then dissected each target area and the injection site using cold scalpels while keeping the brain sections frozen on a metal block cooled to approximately -20ºC in a freezing 2.25M $CaCl_2$ bath[42]. During dissection, we identified each dissected area using a fluorescent dissection microscope to visualize viral GFP expression and DiI stabs labeling each target area (Extended Data Fig. 7). Throughout the procedure, we took care to avoid sample cross-contamination by never reusing tools or blades applied to different areas and changing gloves between samples. To measure noise introduced by contamination, we collected samples of the olfactory bulb from each brain, which served as a negative control.

We then processed the dissected samples for sequencing largely as previously described[30], but pooling all samples after first strand cDNA synthesis. Briefly, we extracted total RNA from each sample using Trizol reagent (Thermo Fisher) according to the manufacturer's instructions. We mixed the sample RNA with spike-in RNA (obtained by *in vitro* transcription of a double stranded ultramer with sequence 5'-GTC ATG ATC ATA ATA CGA CTC ACT ATA GGG GAC GAG CTG TAC AAG TAA ACG CGT AAT GAT ACG GCG ACC ACC GAG ATC TAC ACT CTT TCC CTA CAC GAC GCT CTT CCG ATC TNN NNN NNN NNN NNN NNN NNN NNN NAT

CAG TCA TCG GAG CGG CCG CTA CCT AAT TGC CGT CGT GAG GTA CGA CCA CCG CTA GCT GTA CA-3' (IDT)[30]) and reverse transcribed the RNA mixture using gene specific primer 5'-CTT GGC ACC CGA GAA TTC CAN NNN NNN NNN NNX XXX XXX XTG TAC AGC TAG CGG TGG TCG-3', where X_8 is one of >300 trueseq like sample specific identifiers and N_{12} is the unique molecular identifier, and SuperscriptIV reverse transcriptase (Thermo Fisher) according to the manufacturer's instructions. We then pooled all first strand cDNAs, purified them using SPRI beads (Beckman Coulter) and produced double stranded cDNA as previously described[43]. We then treated the samples using ExonucleaseI (NEB) and performed two rounds of nested PCR using primers 5'-CTC GGC ATG GAC GAG CTG TA-3' and 5'-CAA GCA GAA GAC GGC ATA CGA GAT CGT GAT GTG ACT GGA GTT CCT TGG CAC CC GAG AAT TCC A-3' for the first PCR and primers 5'-AAT GAT ACG GCG ACC ACC GA-3' and 5'- CAA GCA GAA GAC GGC ATA CGA-3' for the second PCR using Accuprime Pfx polymerase (Thermo Fisher). Finally, we gel extracted the resulting PCR amplicons using Qiagen MinElute Gel extraction kit according to the manufacturer's instructions and sequenced the library on a Illumina NextSeq500 high-output run at paired-end 36 using the SBS3T sequencing primer for paired-end 1 and the Illumina small RNA sequencing primer 2 for paired-end 2.

MAPseq data analysis. Based on the sequencing results, we constructed a barcode matrix M of (number of barcodes) x (number of dissected areas) with entry $M_{i,j}$ representing the absolute counts of barcode i in area j as previously described[30]. We de-multiplexed the sequencing results, extracted the absolute counts of each barcode in each sample based on the UMI sequence and error corrected the barcode sequences, before matching barcode sequences to the virus library and constructing matrix M by matching barcode sequences across areas. We then filtered the barcode matrix for 'high-confidence' cell bodies inside the dissected area of V1 by requiring a minimum of

10 counts in at least one target area, an at least 10-fold difference between the cell body location in V1 and the most abundant target area in data normalized to the efficiency of library production as measured by the amount of recovered spike-in RNA counts, and an absolute minimum barcode count of 300 in V1. We then normalized the raw barcode counts in each area by the relative spike-in RNA recovery to the olfactory bulb sample, merged the results from all four processed brains into a single barcode matrix and used this matrix for all further analysis.

To determine whether a particular neuron projected to any given target area, we chose a conservative threshold of at least 5 barcode counts, based on the highest level of barcode expression in the olfactory bulb negative control sample.

Calculation of statistical significance of projection motifs. To calculate the statistical significance of broadcasting projection motifs, we compared against the simplest model in which we assumed that each neuron projected to each area independently. To generate predictions of this model, we first estimated the probability of projecting to each area, assuming independent projections. We define the probability $P(A_i)$ that a given neuron projects to the i^{th} area A_i as

$$P(A_i) = \frac{N_{A_i}}{N_{total}},$$

where N_{A_i} is the number of neurons in the sample that project to area A_i, $i = 1..k$ for k analyzed target areas, and N_{total} is the total number of neurons in the sample.

In our MAPseq experiments, we do not have direct access to N_{total}, since for technical reasons we only include neurons that have at least one projection among the dissected targets. Since in principle some neurons might project to none of the areas dissected (see Fig. 3**a**), failure to include

these would lead to an underestimate of N_{total}. However, assuming independence of projections we can infer N_{total} from the available measurements.

To estimate N_{total}, we first observe that

$$\mathrm{P}(\text{project to at least one area}) + \mathrm{P}(\text{project to no area}) = 1$$

$$\Leftrightarrow \frac{N_{obs}}{N_{total}} + \prod_{j=1}^{k}\left(1 - \frac{N_{A_j}}{N_{total}}\right) = 1$$

where N_{obs} is the total number of neurons observed to project to at least one area. For $k=6$ areas, we can expand this expression to

$$\left(N_{obs} - \sum_{j=1}^{6} N_{A_j}\right) N_{total}^5 + \sum_{i=1}^{6}\sum_{\substack{j=1 \\ j\neq i}}^{6} N_{A_i} N_{A_j} N_{total}^4 - \sum_{i=1}^{6}\sum_{\substack{j=1 \\ j\neq i}}^{6}\sum_{\substack{k=1 \\ k\neq j}}^{6} N_{A_i} N_{A_j} N_{A_k} N_{total}^3$$

$$+ \sum_{i=1}^{6}\sum_{\substack{j=1 \\ j\neq i}}^{6}\sum_{\substack{k=1 \\ k\neq j}}^{6}\sum_{\substack{l=1 \\ l\neq k}}^{6} N_{A_i} N_{A_j} N_{A_k} N_{A_l} N_{total}^2$$

$$- \sum_{i=1}^{6}\sum_{\substack{j=1 \\ j\neq i}}^{6}\sum_{\substack{k=1 \\ k\neq j}}^{6}\sum_{\substack{l=1 \\ l\neq k}}^{6}\sum_{\substack{m=1 \\ m\neq l}}^{6} N_{A_i} N_{A_j} N_{A_k} N_{A_l} N_{A_m} N_{total} + \prod_{i=1}^{6} N_{A_i} = 0.$$

Noting that this is a quintic equation in N_{total}, we can use a root finder to solve for N_{total} numerically, and use the result to calculate $P(A_i)$.

Using the derived N_{total} and $P(A_i)$, we can calculate the p-value for every possible broadcasting motif by calculating the value of the binomial cumulative distribution function, for a total of N_{total} tries, the empirical number of observed counts (successes), and P(motif) assuming independent

projections. We calculated the p-value of all possible bi-, tri- and quadfurcations, and determined significantly over- or under-represented broadcasting motifs at a significance threshold of α=0.05 after Bonferoni correction.

Data availability

All sequencing data are publicly accessible on the Sequence Read Archive under accessions SRR5274845 (ZL097 for mouse 4 and mouse 5) and SRR5274844 (ZL102 for mouse 6 and mouse 7). All single cell tracing results are accessible on http://mouse.vision/han2017 and will be uploaded to http://neuromorpho.org.

Code availability

All software are available at: *http://mouse.vision/han2017*

Additional References

35. Pologruto, T. A., Sabatini, B. L. & Svoboda, K. ScanImage: flexible software for operating laser scanning microscopes. *Biomed. Eng. Online* **2,** 13 (2003).

36. Mayerich, D., Abbott, L. & McCormick, B. Knife-edge scanning microscopy for imaging and reconstruction of three-dimensional anatomical structures of the mouse brain. *J. Microsc.* **231,** 134–43 (2008).

37. Niedworok, C. J. *et al.* aMAP is a validated pipeline for registration and segmentation of high-resolution mouse brain data. *Nat. Commun.* **7,** 711879 (2016).

38. Klein, S., Staring, M., Murphy, K., Viergever, M. A. & Pluim, J. elastix: A Toolbox for Intensity-Based Medical Image Registration. *IEEE Trans. Med. Imaging* **29,** 196–205 (2009).

582 39. Kim, Y. *et al.* Whole-Brain Mapping of Neuronal Activity in the Learned Helplessness
583 Model of Depression. *Front. Neural Circuits* **10,** 3 (2016).

584 40. Renier, N. *et al.* Mapping of Brain Activity by Automated Volume Analysis of Immediate
585 Early Genes. *Cell* **165,** 1789–1802 (2016).

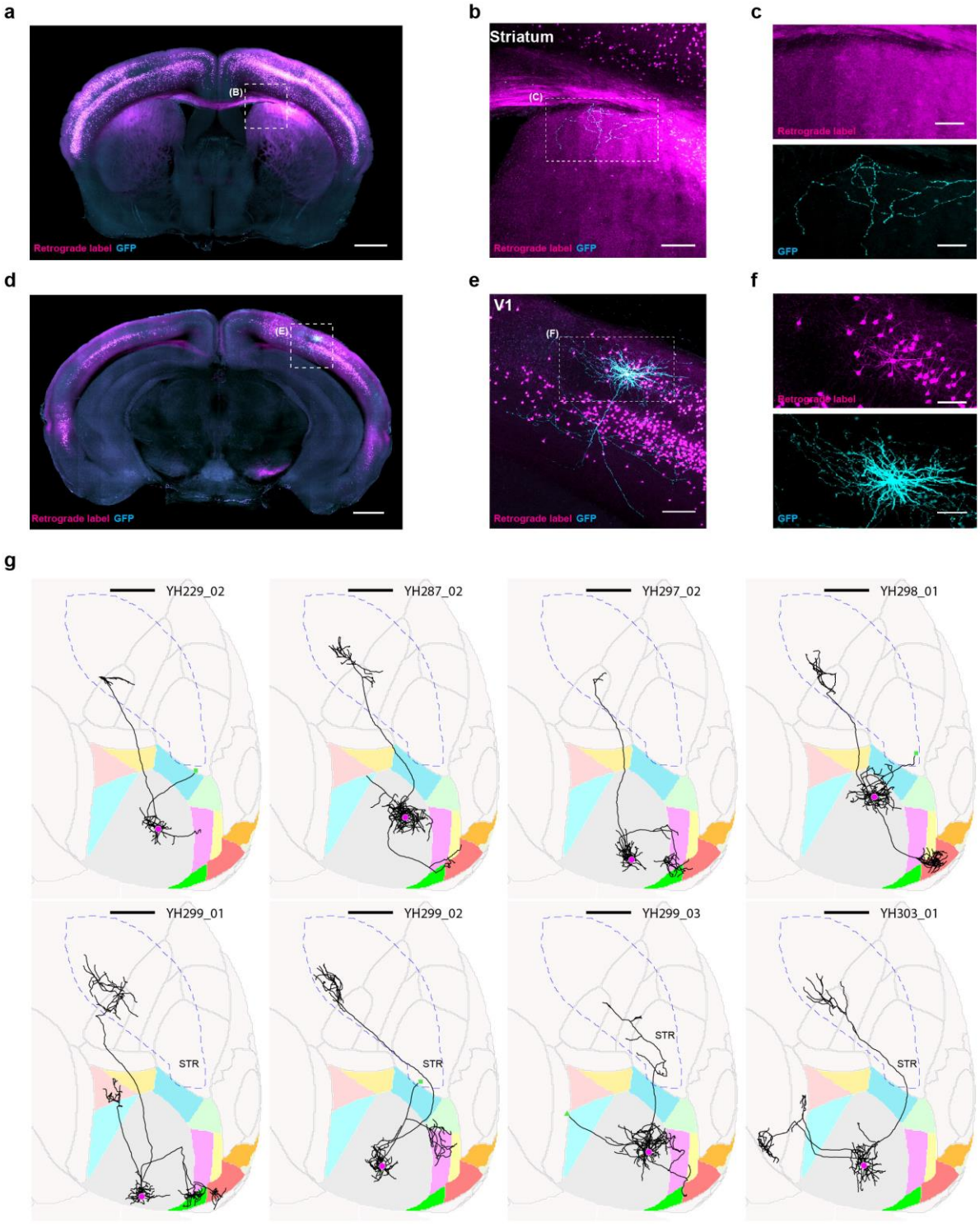
586 41. Roth, M. M. *et al.* Thalamic nuclei convey diverse contextual information to layer 1 of
587 visual cortex. *Nat. Neurosci.* **19,** 299–307 (2015).

588 42. Bryan, W. P. & Byrne, R. H. A calcium chloride solution, dry-ice, low temperature bath. *J.*
589 *Chem. Educ.* **47,** 361 (1970).

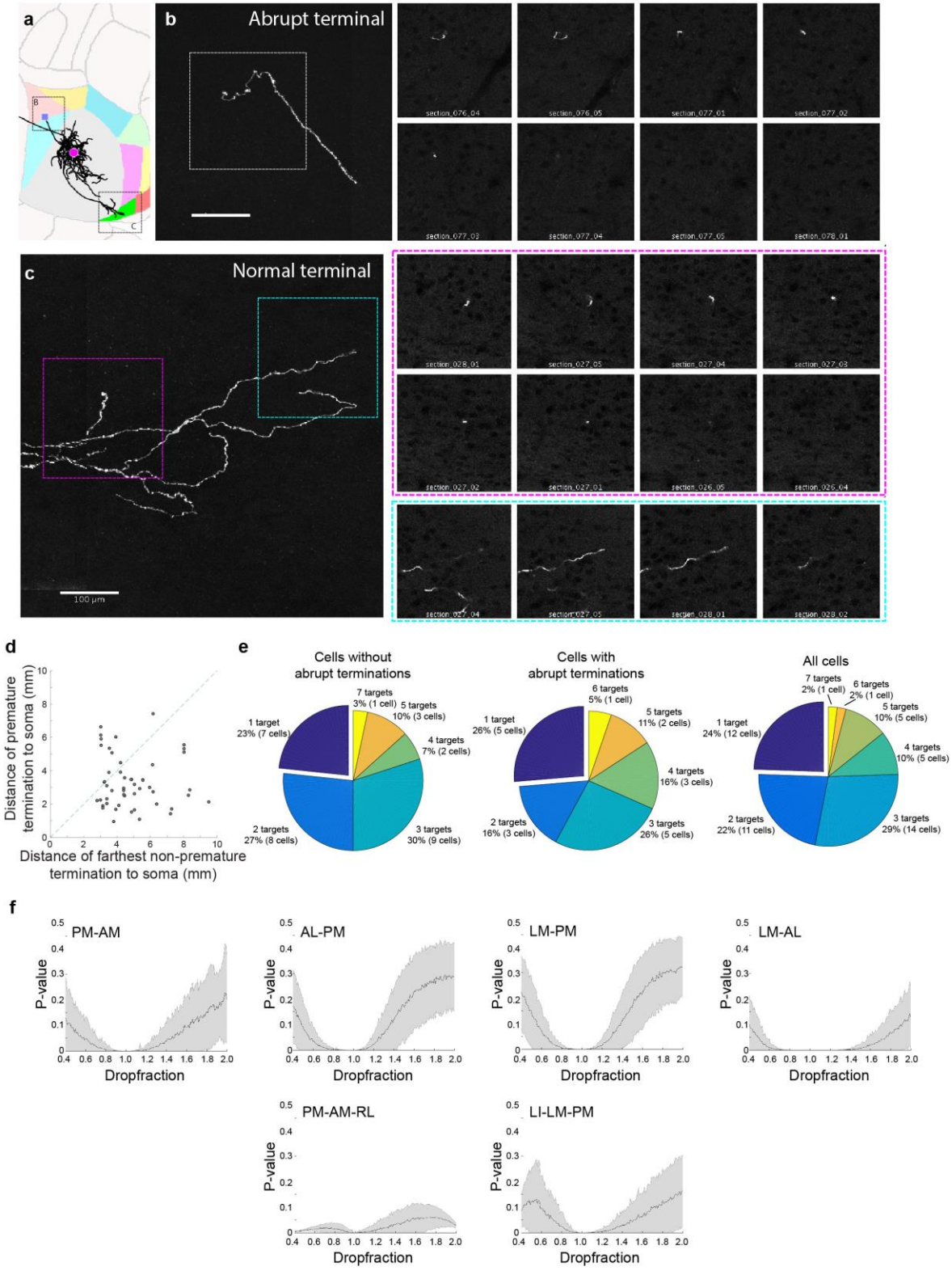
590 43. Morris, J., Singh, J. M. & Eberwine, J. H. Transcriptome Analysis of Single Cells. *JoVE*
591 e2634--e2634 (2011). doi:10.3791/2634

592

Extended Data Figures

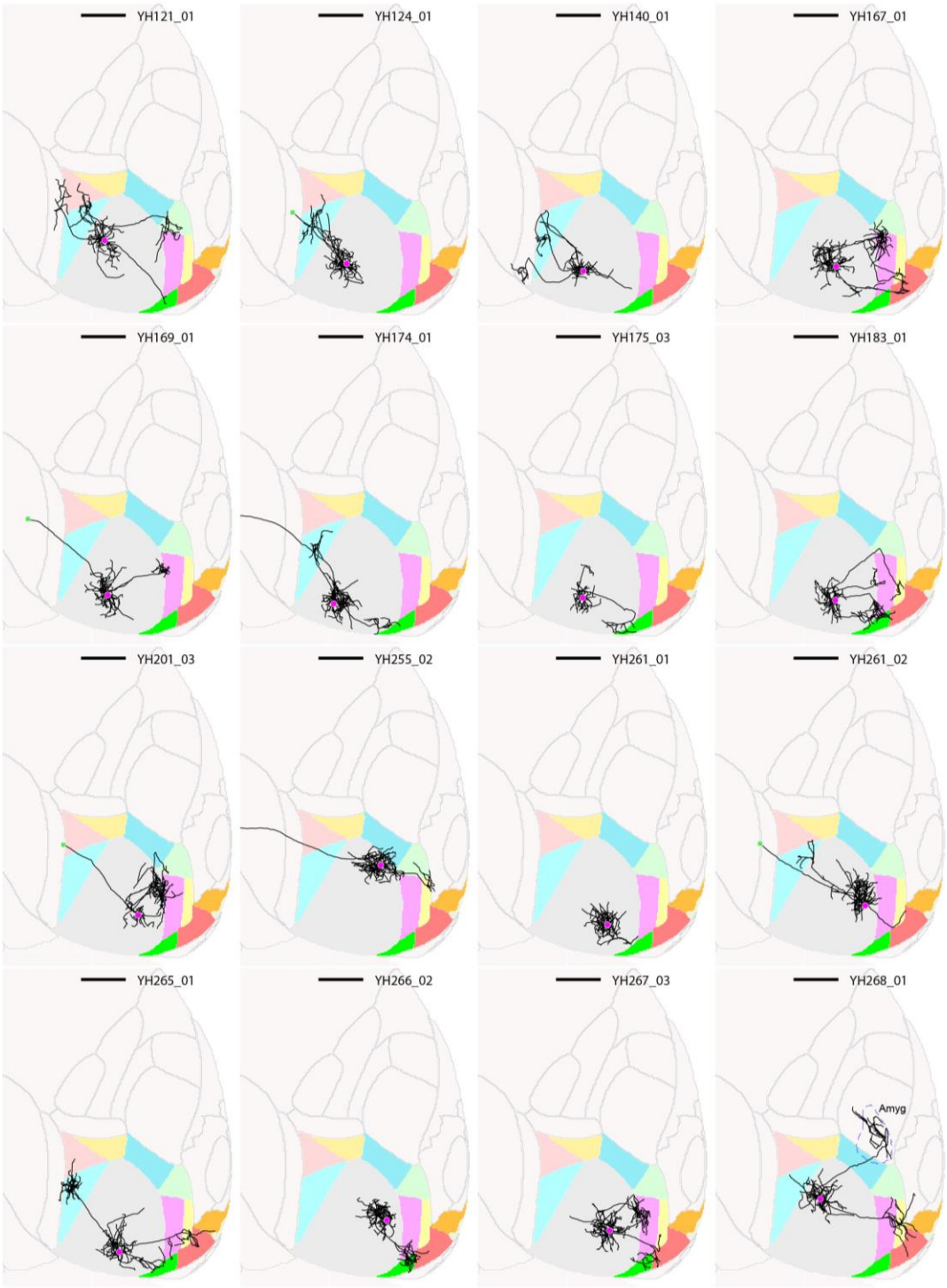


Extended Data Figure 1: Single-neuron tracing protocol efficiently fills axons projecting to the ipsilateral striatum. We retrogradely labeled striatum projecting cells by stereotactically injecting cholera toxin subunit B conjugated with AlexaFluro594 or PRV-cre into the visual striatum of wild type mice or tdTomato reporter mice (Ai14, JAX), respectively (magenta). With visual guidance of two-photon microscopy, we electroporated single retrogradely labeled cells in V1 with a GFP expressing plasmid (cyan). (**a**) Coronal, maximum intensity projections of visual striatum. Scale bar = 1 mm. (**b**) Higher magnification view of the visual stratum. Scale bar = 0.2 mm. (**c**) Single channel images of the same axonal arbor as in (**b**). (**d**) Coronal maximum intensity projection containing V1. Scale bar = 1 mm. (**e**) Higher magnification view of V1. Scale bar = 0.2 mm. (**f**) Single channel images of V1. Scale bar = 0.2 mm. (**g**) Horizontal ARA-space projections of eight retrogradely labeled and electroportated cells. Cell ID numbers are indicated at the top right of each thumbnail. Scale bar = 1 mm. Note that one additional cell was retrogradely labeled and electroporated, which revealed its axonal projection to the striatum, but it is not shown because the brain was too distorted to allow accurate atlas registration.


a
b
Abrupt terminal
c
Normal terminal
100 µm
d
Distance of premature termination to soma (mm)
Distance of farthest non-premature termination to soma (mm)
e
Cells without abrupt terminations
7 targets 3% (1 cell)
5 targets 10% (3 cells)
4 targets 7% (2 cells)
3 targets 30% (9 cells)
2 targets 27% (8 cells)
1 target 23% (7 cells)
Cells with abrupt terminations
6 targets 5% (1 cell)
5 targets 11% (2 cells)
4 targets 16% (3 cells)
3 targets 26% (5 cells)
2 targets 16% (3 cells)
1 target 26% (5 cells)
All cells
7 targets 2% (1 cell)
6 targets 2% (1 cell)
5 targets 10% (5 cells)
4 targets 10% (5 cells)
3 targets 29% (14 cells)
2 targets 22% (11 cells)
1 target 24% (12 cells)
f
PM-AM
AL-PM
LM-PM
LM-AL
PM-AM-RL
LI-LM-PM
P-value
Dropfraction

609

Extended Data Figure 2: Some axonal branches terminate abruptly without arborizing, while other branches of the same neuron arborise extensively within different target areas and appear to be completely filled. (**a**) Horizontal view of a representative cell in ARA space. The abrupt termination is labeled with a purple square. N=28 abruptly terminating cells. (**b**) The abrupt termination of the example cell shown as a maximum z-projection (left) and in the individual z-sections (right). (**c**) Two normal terminations of the same cell, shown as a maximum z-projection (left) and in two color-coded series of z-sections (right). (**d**) Distance of abrupt termination from cell body vs. distance of farthest regular termination of the same cell. Dashed line indicates the unity line. (**e**) Pie charts illustrating the distribution of target numbers of all projecting neurons without abrupt terminations (as shown in the main figures; *left*), of projecting cells with abrupt terminations (*centre*) and of all projecting cells (no abrupt terminations + abrupt terminations; *right*). (**f**) To test the effect of false negatives on our analyses, we simulated the random loss or gain of projections from the MAPseq dataset, while maintaining overall area projection probabilities. N=553 neurons; 400 repeats. P-values based on a binomial test for all six projection motifs determined as significantly over- or underrepresented in our dataset are plotted after removing (dropfraction < 1) or adding (dropfraction >1) connections. Mean (black line) and s.d. (shaded area) are indicated.

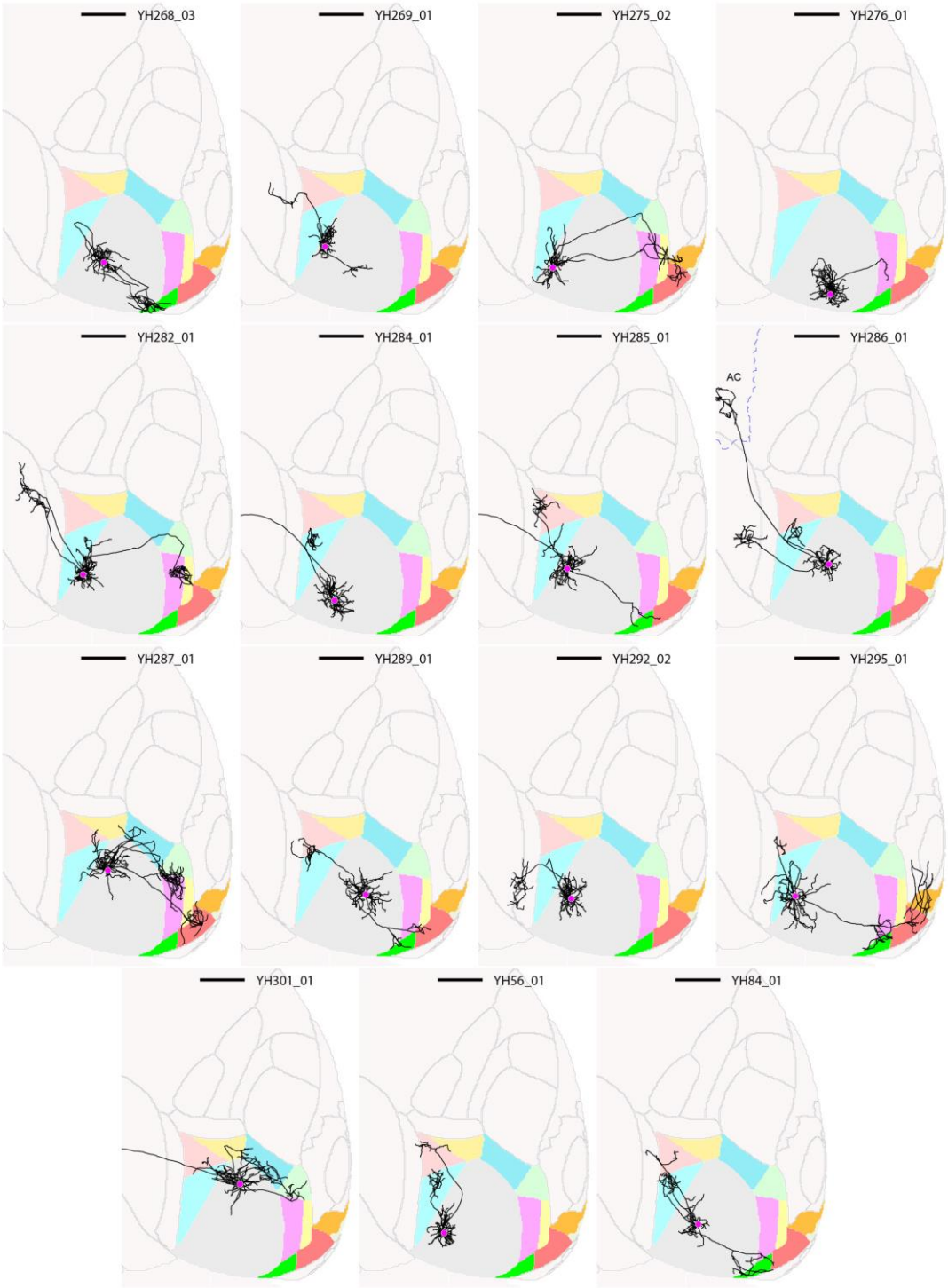


627

628 **Extended Data Figure 3: Thumbnails of traced layer 2/3 V1 neurons, part 1.** Horizontal views

629 of the ARA space are shown, and cell ID numbers are indicated at the top right of each thumbnail.

630 Scale bar = 1 mm.

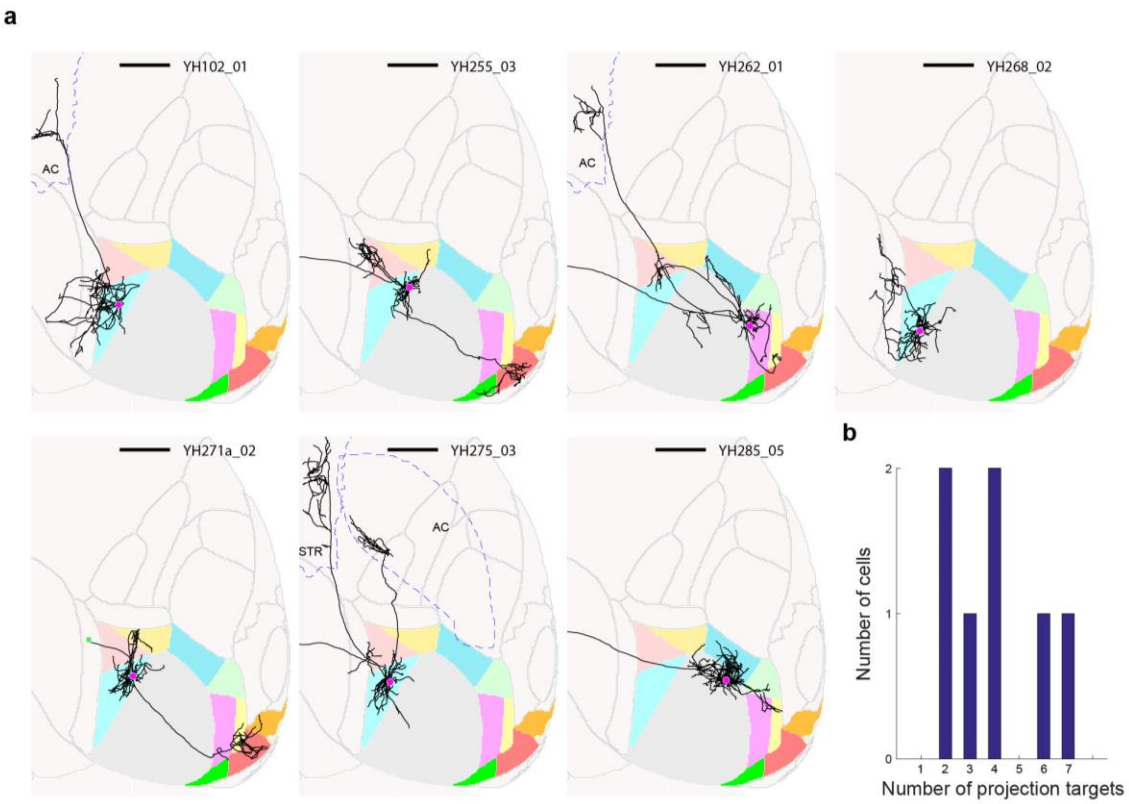


631

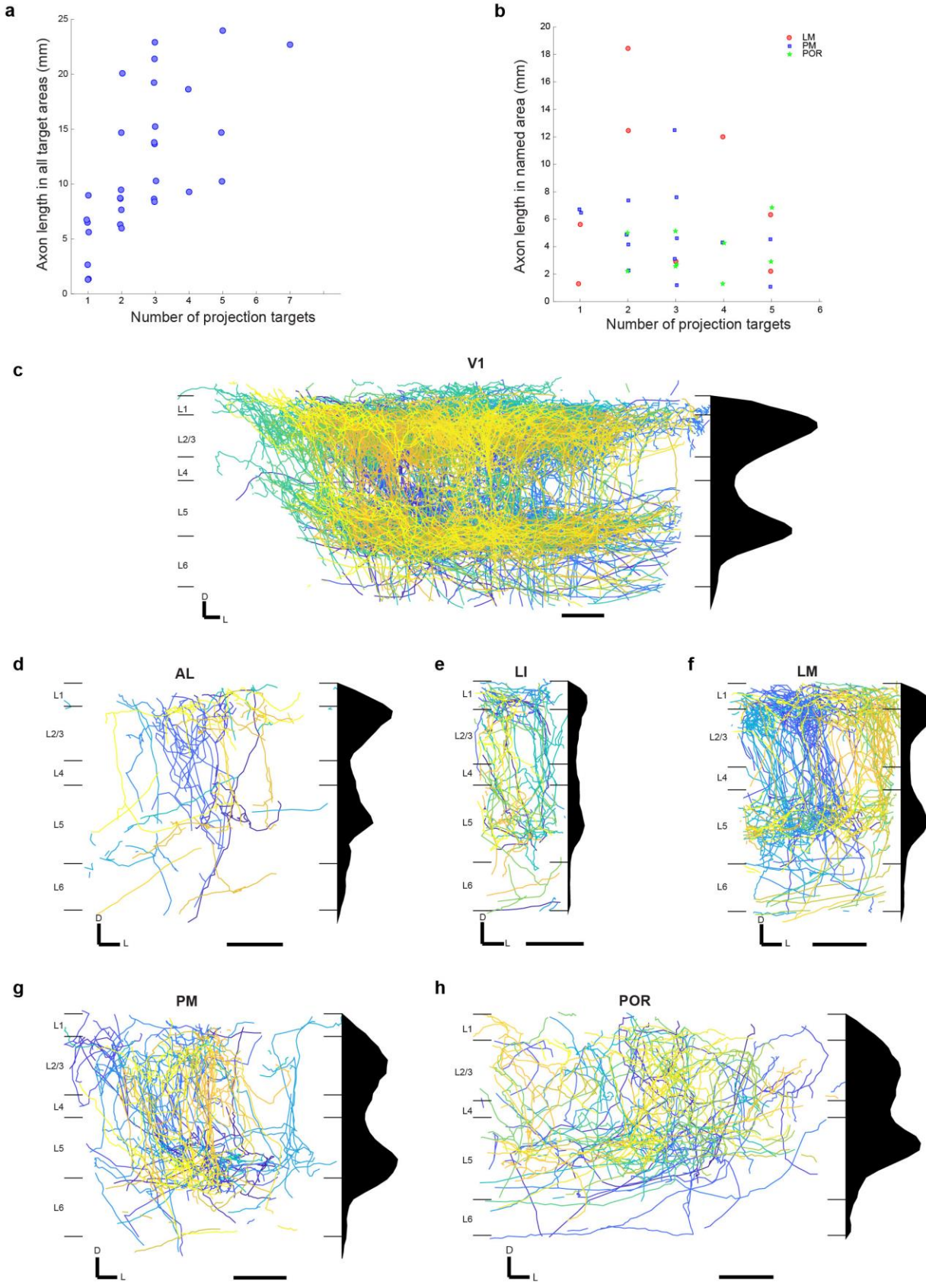
632 **Extended Data Figure 4: Thumbnails of traced layer 2/3 V1 neurons, part 2.** Horizontal views

633 of the ARA space are shown, and cell ID numbers are indicated at the top right of each thumbnail.

634 Scale bar = 1 mm.

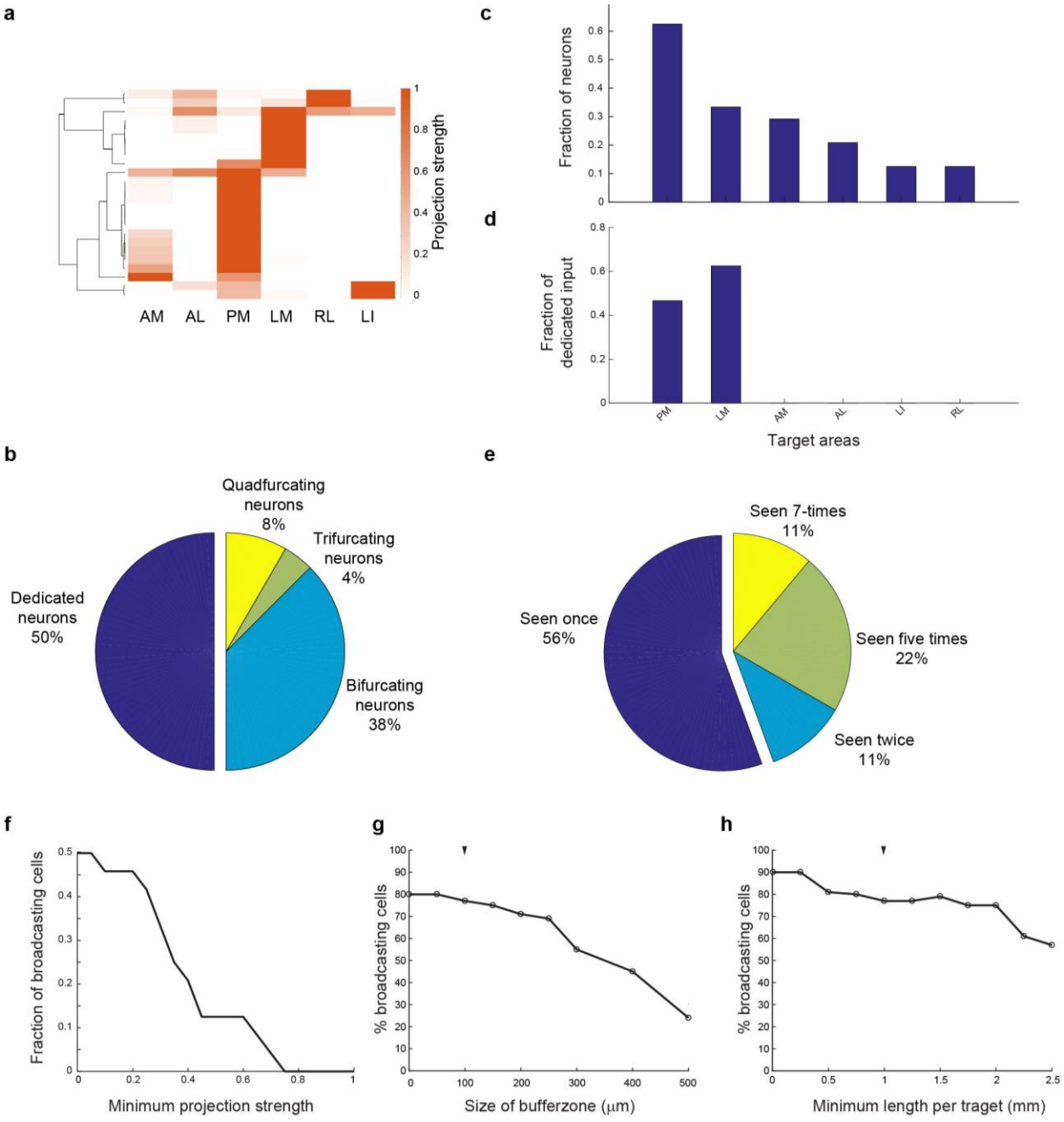


Extended Data Figure 5: Individual neurons in higher visual areas project to more than one target area. (**a**) Thumbnails of all traced neurons with cell bodies not in V1. Brain area identity is color-coded as in Figure 1. Cell identity is indicated at the top right of each thumbnail. Scale bar = 1 mm. (**b**) Histogram of the number of target areas per cell.


a
Axon length in all target areas (mm)
Number of projection targets
b
Axon length in named area (mm)
Number of projection targets
LM
PM
POR
c
V1
L1
L2/3
L4
L5
L6
D
L
d
AL
e
LI
f
LM
g
PM
h
POR

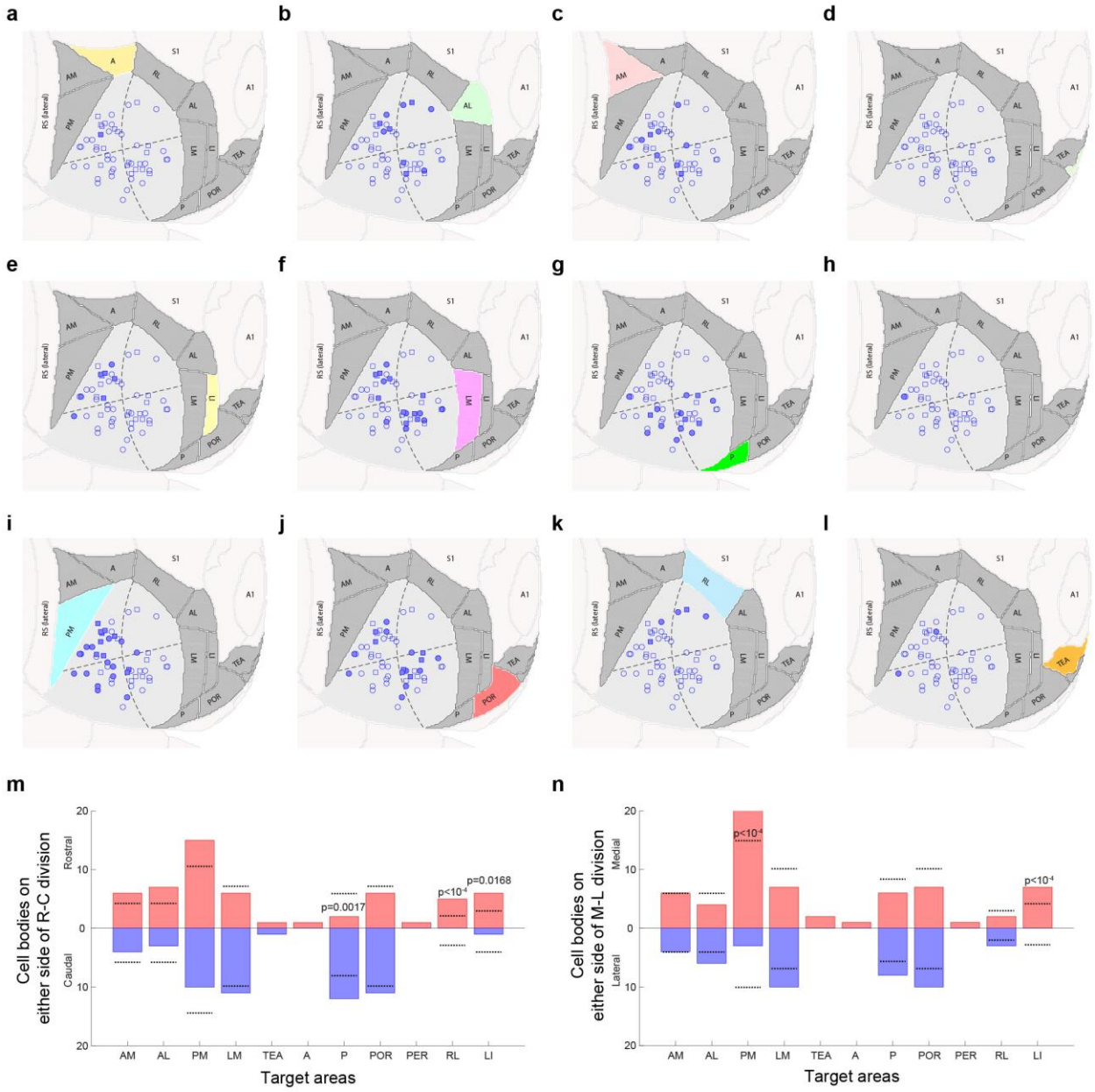
640

Extended Data Figure 6: Density of axonal innervation by area and layer of V1 layer 2/3 projection neurons. (**a**) Total axon length plotted as a function of the number of targets innervated by every V1 projection neuron. (**b**) Axon length in area LM, PM or POR plotted as a function of the total number of targets innervated by each neuron projecting to the respective area. (**c**-**h**) The axons of V1 neurons in target areas most densely innervate layers 2/3 and 5, with some density in layer 1, but less in layers 4 and 6, often recapitulating the laminar axonal profile within V1. Coronal views of each area are shown in ARA space (*left*) and axonal arbors of each neuron innervating the area are color coded. Scale bar = 200 μm. A histogram of the laminar innervation is shown (*right*). Note that cells with abrupt terminations outside the shown area were included in this analysis. Areas depicted are (**c**) V1, (**d**) AL, (**e**) LI, (**f**) LM, (**g**) PM, (**h**) POR. White matter axons are not shown.



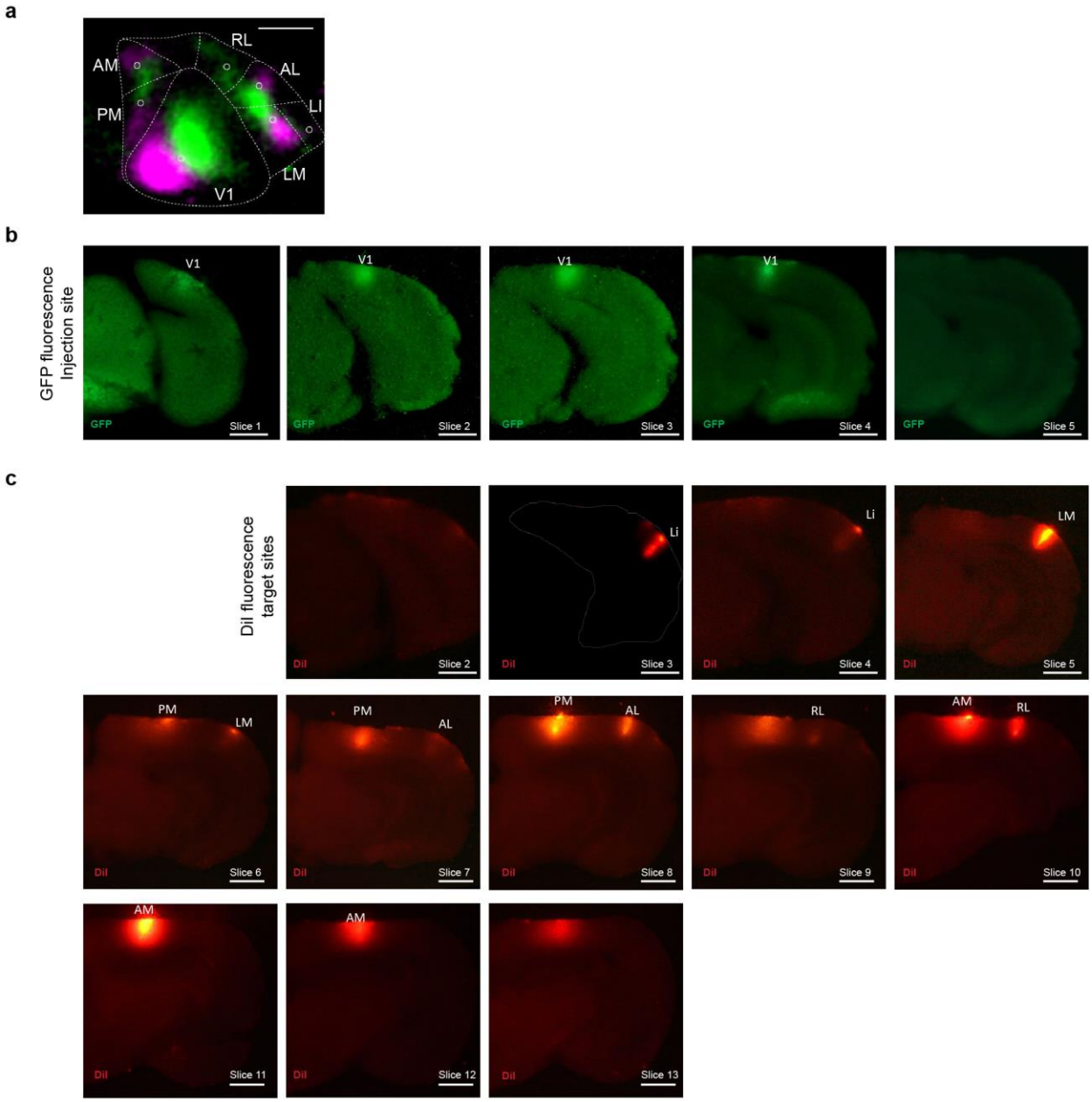
Extended Data Figure 7: Conclusions from fluorescence-based single neuron tracing data hold true if analysis is restricted to subset of target areas. (**a**) The projection patterns of reconstructed GFP-filled neurons when only the six target areas LI, LM, AL, PM, AM, and RL are considered. Projection strengths are normalized to the maximum projection of each neuron, and only neurons projecting to at least one target area are shown. (**b**) Pie chart showing the distribution

658 of target area numbers per projecting neuron. (**c**) Bar graph illustrating the fraction of all cells
659 projecting to each target area. (**d**) The fraction dedicated input per area. (**e**) The number of times
660 each binarized projection motif is observed. (**f**) The fraction of broadcasting cells as a function of
661 the minimum projection strength (relative to the primary target) that each area needs to receive to
662 be considered a target. (**g**) The fraction of broadcasting cells as a function of increasing buffer
663 zones between areas within which axons are ignored, assuming a minimum projection of 1 mm of
664 axon per target area. (**h**) The fraction of broadcasting cells as a function of the minimal amount of
665 axon per area for it to be considered a target, assuming buffer zones of 100 μm width.



Extended Data Figure 8: Location of cell bodies in V1 as a function of their projection targets. (**a**-**l**) Horizontal views of ARA space are shown. The location of all traced V1 neurons are indicated as circles (cells with no abrupt terminations) or squares (cells with abrupt terminations). In every plot the cells projecting to the highlighted higher visual area are colored in solid blue. Target areas considered are (**a**) A, (**b**) AL, (**c**) AM, (**d**) ECT, (**e**) LI, (**f**) LM, (**g**) P, (**h**) PER, (**i**) PM, (**j**) POR, (**k**) RL, (**l**) TEA. (**m**-**n**) Quantification of cell body location in the rostro-caudal (**m**)

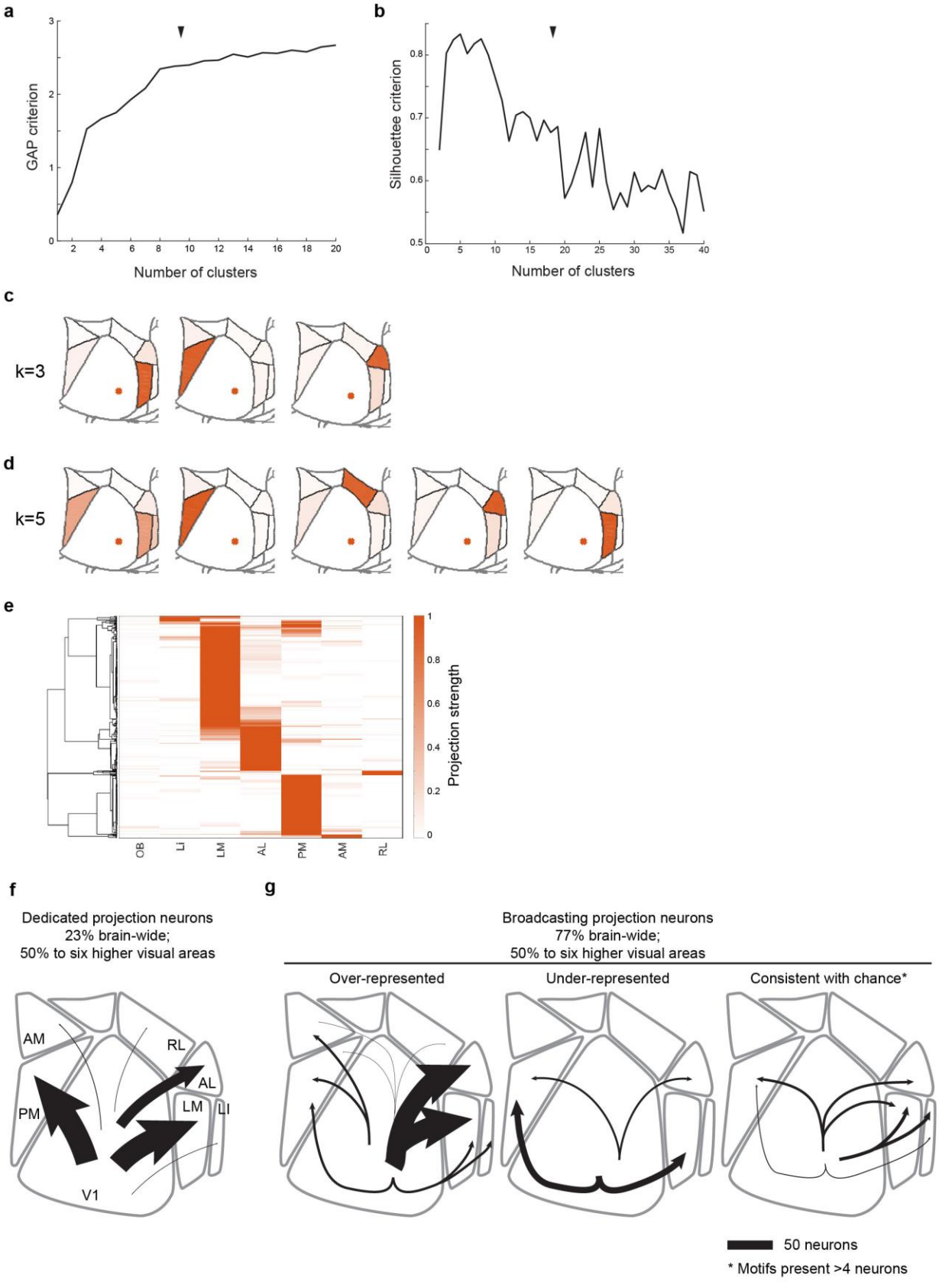
673 and medio-lateral (**n**) direction. Dotted lines indicate expected number of cells based on a boot-
674 strapping procedure, where we randomly selected neurons from the available positions to project
675 to each area and repeated the process 10,000 times. P-values were derived from the bootstrapping
676 probability distribution and are indicated for projection targets significantly deviating from this
677 expectation (α=0.05). P-values below 10^{-4} are not exact and are therefore indicated as a range..



678

679 **Extended Data Figure 9: MAPseq dissection strategy.** We identified the to-be-dissected higher
680 visual areas by performing intrinsic imaging of visual cortex in response to stimuli at different
681 positions in the contralateral visual field and mapping the resulting changes in intrinsic signals. (**a**)
682 A representative retinotopic map, with responses to the two 25° visual stimuli pseudocolored in
683 green and magenta (stimulus 1 position: 90° azimuth, 20° elevation; stimulus 2 position: 60° azi-

muth, 20° elevation). Based on this map, we fluorescently labelled retinotopically matched positions in the to-be-dissected cortical areas with a DiI stab (white circles). Putative borders between the higher visual areas are indicated in dashed lines for orientation. Scale bar = 1 mm. N=4 animals. (**b**) The MAPseq virus injection site is discernible in consecutive frozen 180 μm thick coronal sections, using GFP fluorescence. Scale bar = 1 mm. (**c**) DiI injections targeted to matched retinotopic positions in six target areas identified by intrinsic signal imaging. DiI epifluorescence images of each 180 μm thick slice are shown, and dissected areas are labeled. Scale bar = 1 mm.


a
GAP criterion
Number of clusters
b
Silhouettee criterion
Number of clusters
c
k=3
d
k=5
e
Projection strength
OB
Li
LM
AL
PM
AM
RL
f
Dedicated projection neurons
23% brain-wide;
50% to six higher visual areas
AM
RL
AL
PM
LM
LI
V1
g
Broadcasting projection neurons
77% brain-wide;
50% to six higher visual areas
Over-represented
Under-represented
Consistent with chance*
50 neurons
* Motifs present >4 neurons

691

Extended Data Figure 10: Clustering of MAPseq data and data summary. (**a**) GAP and (**b**) Silhouette criteria for k-means clustering of the MAPseq neurons as a function of the number of clusters. Black arrow heads indicate chosen number of clusters (k=8). (**c,d**) Centroids for alternative, near-optimal cluster number choices with (**c**) k=3 and (**d**) k=5. (**e**) Hierarchical clustering results of the MAPseq dataset using a cosine distance metric. Color intensity in (**c**,**d**,**e**) indicated projection strengths. (**f**,**g**) Summary of single-neurons projections from V1. (**f**) Cells targeting single higher visual areas (dedicated projection neurons) comprise the minority of layer 2/3 V1 projection neurons. Among the areas analysed by MAPseq, dedicated projection neurons predominantly innervate cortical areas LM or PM. (**g**) Cells projecting to two or more areas (broadcasting projection neurons) are the dominant mode of information transfer from V1 to higher visual areas. In the six areas analysed by MAPseq, broadcasting neurons innervate combinations of target areas in a non-random manner, including those that are more or less abundant than expected by chance. Line width indicates the absolute abundance of each projection type as observed in the MAPseq dataset.

Supplemental Notes

Fluorescence-based single neuron reconstructions

Supplemental Note 1: Abrupt terminations in GFP filled cells

The mapping of axonal processes by fluorescence-based single neuron reconstructions relies on the complete and roughly homogeneous filling of all axonal processes with GFP protein. If one branch is not completely filled, its full reach will not be detected, and will thus cause a false negative projection.

To minimize such false negatives in our dataset, we took a number of precautions both during the experiment and during data analysis. First, we used a strong promoter to express GFP protein within each electroporated cell, thus maximizing the amount of fluorescent protein expressed in each cell. We then monitored the expression of GFP by epifluorescence imaging through a cranial window, to determine the optimal time of GFP expression on a per cell basis, aiming for high GFP expression but avoiding the cytotoxic effects of too high transgene expression. After fixation and imaging, we assessed the health of every cell by the integrity of the axonal processes – a very sensitive indicator of cell health – and discarded any cell with fragmented axonal processes before tracing. Only 7/71 traced cells exhibit minor blebbing of the axon, such that we can conclude that axonal retraction due to poor health of a cell is not a contaminating factor in our dataset. Finally, we stringently screened for incomplete fills in our imaging datasets in cells that appeared healthy. Incomplete fills are generally accepted to show a gradual fading of fluorescence along the length of a process, as is intuitive for an incomplete fill by a diffusing fluorescent molecule. Such gradual terminations are very rare in our dataset (5 fading terminations detected in total from 52 cells with approximately 100 terminations each), suggesting a low false positive rate. In addition to the gradual terminations, we also observed abrupt terminations, where a process is brightly labeled and simply stops without branching or fading fluorescence either in the white or gray matter (Extended Data Fig. 2**a,b**). Such abrupt terminations are a common occurrence in single cell fills, irrespective of the exact method used for single cell filling, and may thus reflect the true extent of the axon (see for example refs[1,2]). We find that the vast majority of abrupt terminations observed in our dataset are more proximal to their cell body than the farthest filled process of the same cell (Extended Data Fig. 2**d**). These data suggest that abrupt terminations in our data are not the result of a distance dependent failure to fill an axonal branch, and are in agreement with the prevalent interpretation of abrupt terminations as real. Nonetheless, we decided to exclude all cells with abrupt terminations from our analysis as a cautionary measure to produce the best possible dataset to act as a gold standard for a comparison to our MAPseq dataset. Note, that all conclusions presented in this work, in particular the high number of target areas per V1 neuron, are robust to the ex/inclusion of abrupt terminations (Extended Data Fig. 2**e**).

MAPseq

Like any other method, MAPseq is subject to false positives (detection of an extra, artefactual projection) and false negatives (failing to identify a real connection). Please refer to ref[3], for a detailed discussion of the effect of fibers of passage, co-infections, infection of more than one cell with the same barcode sequence, and various other sources of false negatives and false positives. Below, we briefly discuss the most important considerations.

Supplemental Note 2: Unique labeling of cells by viral infection

As described in more detail previously[3], in MAPseq we deliver barcodes to cells by random viral infection. In the simplest scenario we aim to deliver one unique barcode per labeled cell, such that each cell can unequivocally be identified by a single sequence. In practice however, there are two scenarios that deviate from this simple model.

On the one hand, we might infect cells with more than one virus particle and thus label each cell with more than one barcode sequence. Such multiple labeling results in MAPseq overestimating the total number of traced cells, but does not result in a false projection pattern and maintains the relative proportions of projection types. Such multiple labeling will therefore not lead to any false positive results.

On the other hand, degenerate labeling, that is labeling more than one cell with the same barcode sequence, produces artificial projection patterns that result from the merged projection pattern of all the cells labeled with the same barcode. We avoid degenerate labeling in MAPseq by infecting cells with a very large virus library. The fraction of uniquely labeled cells for a given size of virus library with an empirically determined barcode probability distribution can be described by

$$F = 1 - \sum_{i=1}^{N} p_i(1 - (1 - p_i)^{k-1}) ,$$

where p_i is the probability of barcode i=1..N to be chosen from the virus library, k is the number of labeled cells and N is the total number of barcodes in the library[3]. Given the size of the library used in this study (~10^7 distinct barcode sequences) and the number of cells traced per brain (~300), the vast majority of cells (>99 %) will be uniquely labeled. In previous work[3] we validated these theoretical predictions by multiple independent experimental methods.

Supplemental Note 3: Other sources of false positives and negatives.

Beyond errors introduced by degenerate labeling (see above), MAPseq is subject to false positives and negatives from other sources. False negatives are introduced into the dataset when the strength of a real projection falls below the detection floor of MAPseq. Conversely, MAPseq false positives

are introduced when barcodes are detected in areas in which they were not originally present in (e.g. by sample cross-contamination).

Several lines of evidence suggest that MAPseq false negative and positive rates are low. In previous experiments[3], we measured the efficiency of MAPseq to be very similar to that of Lumafluor retrobeads (91.4 ± 6 % (mean ± std error))[3], and therefore concluded that MAPseq false negative rates are comparable to those of other, well established methods. Similarly, we found MAPseq false positive rates to be low (1.4 ± 0.8 % (mean ± std error))[3].

In the present study we improve on these previous estimates by comparing MAPseq data directly to the gold standard of single neuron tracing. To do so, we first used a bootstrapping procedure to measure the minimum pairwise cosine distances between each member of a randomly sampled set of MAPseq neurons and the remaining MAPseq neurons. We then measured the minimum pair-wise cosine distance between the fluorescence-based single neuron reconstructions and the remaining MAPseq neurons. As a negative control, we further measured the minimum pairwise distance between a set of random neurons (with their projection strengths drawn from a uniform distribution) and the remaining MAPseq neurons. We then compared the distribution of minimum distances for the three sets of measurement and found that while the MAPseq-to-MAPseq and fluorescence-based reconstructions-to-MAPseq distributions are statistically indistinguishable (two sample Kolmogorov-Smirnov test, p=0.9439), both the MAPseq-to-MAPseq and fluorescence-based reconstructions-to-MAPseq distributions are statistically different from the random neuron-to-MAPseq distribution (Fig. 2**d**; two sample Kolmogorov-Smirnov test, $p=2.75\text{x}10^{-4}$ and $p=8.76\text{x}10^{-5}$, respectively). Taken together, these results indicate that the statistics of projections inferred by MAPseq are indistinguishable from those obtained by fluorescence-based single neuron reconstructions.

Supplemental Note 4: False negatives cannot explain the observed structure in the projectional dataset.

In order to investigate the effect of the MAPseq false negative rate on our analysis of over- and underrepresented projection motifs, we simulated how our results changed by removing or adding projections to the MAPseq dataset, while keeping the average area projection probabilities fixed. As expected, randomly adding or removing connections did not produce structure in the dataset, consistently increasing p-values (i.e. making them less significant) for all six significantly over- or underrepresented projection motifs reported in Fig. 3 (Extended Data Fig. 2**f**).

Supplemental Tables

Area 1	Area 2	Area 1 projecting cells	Area 2 projecting cells	Area 1&2 projecting cells	% double labeled cells relative to total	% double labeled cells relative to Area 1	% double labeled cells relative to Area 2
Li	LM	43	323	34	10.2	79.1	10.5
Li	AL	43	207	4	1.6	9.3	1.9
Li	PM	43	228	20	8.0	46.5	8.8
Li	AM	43	30	0	0.0	0.0	0.0
Li	RL	43	18	0	0.0	0.0	0.0
LM	AL	323	207	132	33.2	40.9	63.8
LM	PM	323	228	71	14.8	22.0	31.1
LM	AM	323	30	10	2.9	3.1	33.3
LM	RL	323	18	6	1.8	1.9	33.3
AL	PM	207	228	31	7.7	15.0	13.6
AL	AM	207	30	7	3.0	3.4	23.3
AL	RL	207	18	7	3.2	3.4	38.9
PM	AM	228	30	22	9.3	9.6	73.3
PM	RL	228	18	7	2.9	3.1	38.9
AM	RL	30	18	2	4.3	6.7	11.1

Supplemental Table 1: Simulated double retrograde tracing based on MAPseq anterograde tracing data. We determined whether any given MAPseq neuron targeted any one area using the same projection criterion used for the analysis in Fig. 3. For any pair of the six higher visual areas analyzed using MAPseq, we then determined the number of cells that projected to either area in the pair, or to both — effectively simulating double retrograde tracing from the two areas in the pair to V1. We here show the raw counts of cell projecting to each area, and the percentages of cells that project to the indicated pairs of areas, i.e. “double labeled” cells.

References

1. Kita, T. & Kita, H. The subthalamic nucleus is one of multiple innervation sites for long-range corticofugal axons: a single-axon tracing study in the rat. *J. Neurosci.* **32,** 5990–9 (2012).
2. Yamashita, T. *et al.* Membrane Potential Dynamics of Neocortical Projection Neurons Driving Target-Specific Signals. *Neuron* **80,** 1477–1490 (2013).
3. Kebschull, J. M. *et al.* High-Throughput Mapping of Single-Neuron Projections by Sequencing of Barcoded RNA. *Neuron* **91,** 975–987 (2016).