2P-NucTag: on-demand phototagging for molecular analysis of functionally identified cortical neurons

Authors: Jingcheng Shi[1,2,3,*], Boaz Nutkovich[4,5,*], Dahlia Kushinsky[4,5,*], Bovey Y. Rao[1,2,3,*], Stephanie A. Herrlinger[1,2,*], Emmanouil Tsivourakis[4,5], Tiberiu S. Mihaila[1,2], Margaret E. Conde Paredes[1,2,3,6], Katayun Cohen-Kashi Malina[4,5], Cliodhna K. O'Toole[1,2], , Hyun Choong Yong[1,2], Brynn M. Sanner[1,2], Angel Xie[1,2], Erdem Varol[6], Attila Losonczy[1,2,#, ‡], and Ivo Spiegel[4,#,‡]

Affiliations:
[1] Department of Neuroscience, Columbia University, New York, NY, United States
[2] Mortimer B. Zuckerman Mind Brain Behavior Institute, Columbia University, New York, NY, United States
[3] Doctoral Program in Neurobiology and Behavior, Columbia University, New York, NY, United States
[4] Department of Brain Sciences, Weizmann Institute of Science, Rehovot, Israel
[5] Department of Molecular Neuroscience, Weizmann Institute of Science, Rehovot, Israel
[6] Tandon School of Engineering, New York University, New York, NY, United States

[*] Co-first authors
[#] Co-senior authors
‡ Corresponding authors. Emails: al2856@columbia.edu, ivo.spiegel@weizmann.ac.il

Summary:
Neural circuits are characterized by genetically and functionally diverse cell types. A mechanistic understanding of circuit function is predicated on linking the genetic and physiological properties of individual neurons. However, it remains highly challenging to map the transcriptional properties to functionally heterogeneous neuronal subtypes in mammalian cortical circuits *in vivo*. Here, we introduce a high-throughput two-photon nuclear phototagging (2P-NucTag) approach optimized for on-demand and indelible labeling of single neurons via a photoactivatable red fluorescent protein following *in vivo* functional characterization in behaving mice. We demonstrate the utility of this function-forward pipeline by selectively labeling and transcriptionally profiling previously inaccessible 'place' and 'silent' cells in the mouse hippocampus. Our results reveal unexpected differences in gene expression between these hippocampal pyramidal neurons with distinct spatial coding properties. Thus, 2P-NucTag opens a new way to uncover the molecular principles that govern the functional organization of neural circuits.

One-Sentence Summary: A high-throughput on-demand phototagging approach for direct mapping of molecular and functional properties of cortical neurons in behaving animals.

Main text:

Information processing in neural circuits requires precise interactions between molecularly and functionally diverse populations of neurons. Since gene expression ultimately dictates neuronal connectivity and function, a fundamental goal of neuroscience has been to characterize gene expression profiles of functionally defined neurons and to measure changes in gene expression associated with distinct functional states of neurons[1]. High-throughput transcriptomic approaches such as single-cell/-nucleus RNA-sequencing (sc/snRNA-seq) and spatial transcriptomics have greatly accelerated the identification of gene programs in molecularly distinct types of neurons at single-cell resolution[2–8]. However, the complete biological picture from paired *in vivo* and *ex vivo* functional and anatomical characterization studies of neuronal subtypes[9–13] remains missing, requiring the generation and validation of new subtype-specific molecular tools[14–16]. Recently, correlated *in vivo* Ca^{2+} imaging with *post hoc* spatial transcriptomics has been used to relate gene expression with *in vivo* function, but this approach is limited to spatially sparse GABAergic interneurons[17–20]. Therefore, a method to identify genes that are differentially expressed in densely packed but functionally distinct glutamatergic pyramidal neurons (PNs), could significantly accelerate our understanding of how gene expression determines circuit function and behavior.

The inability to tag single functionally identified cortical PNs *in vivo* in behaving animals presents a significant challenge, as large-scale neural recordings have shown that PNs are highly heterogeneous in their physiological, anatomical, and response properties, and are spatially intermixed within neocortical and hippocampal circuits[16,21–27]. For example, PNs with distinct spatial coding properties are distributed throughout the dense cell body layer of the hippocampus[28–35]. However, the origin of this functional diversity in feature selectivity is largely unknown, and it remains unclear if gene expression differences are associated with discrete and transient functional cell states[35–40]. Thus, there is a critical unmet need for function-forward approaches that directly map *in vivo* physiological and transcriptional profiles in cortical circuits in behaving animals.

Previous attempts using Ca^{2+} and light-dependent labeling of transiently active neurons[41] were limited by their spatial resolution, deficiencies in targeting neurons with high baseline intracellular Ca^{2+} levels, and the inability to label neurons that decrease their activity in response to behavioral state or sensory stimuli. Similarly, immediate early gene-dependent labeling approaches[42–45] lack the temporal and spatial resolution to faithfully report the precise activity patterns and response properties of single neurons. Finally, previous attempts to tag cortical neurons with photoactivatable fluorescent proteins[46] with single-cell precision have been deployed with limited success[47].

Here we introduce a robust *in vivo* pipeline (2P-NucTag), based on a photoactivatable red fluorescent protein (PAmCherry) and a genetically encoded green Ca^{2+} indicator (GCaMP7f), that optimizes a previously described framework[48] mainly used *ex vivo*. Our approach combines large-scale *in vivo* two-photon (2P) functional imaging of cortical PNs with reliable and selective 2P phototagging of nuclei in a subset of neurons based on their functional properties. Using fluorescence-activated cell sorting (FACS) to isolate phototagged neuronal nuclei *post hoc*, combined with our recently developed Meso-seq approach for transcriptomics in ultra-sparse populations[49], we achieve previously unattainable molecular characterization of *functionally* identified PNs *in vivo* in behaving animals.

In vivo two-photon phototagging with 2P-NucTag

A major challenge in combining *in vivo* functional recording with stable tagging in the same neuron is the co-expression of an activity sensor and a photoactivatable tag with spectrally

separable fluorescent imaging and photoactivation. We overcame this challenge by generating a bicistronic construct on a recombinant adeno-associated viral (rAAV) backbone that co-expresses cytosolic GCaMP7f[50] for 2P Ca^{2+} imaging and a nucleus-targeted photoactivatable red fluorescent protein (H2B-PAmCherry) for 2P phototagging[48] under a promoter that is selective for cortical glutamatergic neurons[51] (2P-NucTag, Fig. 1A). Upon injection of the 2P-NucTag rAAV into the CA1 region of the mouse dorsal hippocampus to label CA1 PNs, we found that GCaMP7f is properly expressed in the perinuclear space of the infected PNs (Fig. 1B). Targeting the nuclei of GCaMP-expressing neurons, we achieved rapid nuclear PAmCherry photoconversion using 810-nm excitation light on a three-dimensional acousto-optical deflector microscope (3D-AOD)[52,53] (Fig. 1B; supplementary movie 1, see methods) and orthogonal GCaMP-Ca^{2+} activity imaging with 940 nm excitation light (fig. S1A,B). Photoconverted PAmCherry red fluorescence was detected at >1000 nm excitation (1040 or 1070 nm, see methods) and was localized to the targeted nuclei (Fig. 1B, top). 2P-NucTag enabled the imprinting of arbitrary tagging patterns into the CA1 pyramidal cell layer with 3D-AOD scanning, showing spatially precise photoconversion of H2B-PAmCherry with single-nuclear and even sub-nuclear resolution (Fig. 1B, middle and bottom). We carried out a detailed characterization of wavelength, laser power, and duration-dependence of PAmCherry photoactivation *in vivo* to identify the optimal parameters for spatially precise photoactivation (Fig. 1C). Based on our results, we opted for 810 nm excitation light with 37-42 mW laser power (measured after the objective), 1.3 ms/pixel dwell time over 70 pixel x 70 pixel regions-of-interest (ROIs, with 0.1 μm/pixel resolution). The 810-nm wavelength is spectrally separated from the GCaMP-based Ca^{2+} imaging wavelength at 940 nm, and these photoactivation wavelengths, laser power, and duration parameters yielded robust increases in PAmCherry fluorescence of targeted nuclei (192-379% increase in PAmCherry red fluorescence visualized using1040 nm excitation: n = 36 cells, 295% ± 8% $\Delta F/F$, mean ± s.e.m.) after single scans while minimizing total scan time and power. We confirmed that phototagged nuclei remained detectable over multiple days after photolabeling (Fig. 1D). We next photoactivated a subset of CA1 PNs in a large (700 x 700 μm) field of view (FOV), similar to the FOV size used for *in vivo* 2P population imaging experiments in CA1 (Fig. 1E, left). We found clear photoactivation of individual target nuclei when visualized *in vivo*. We then confirmed that *in vivo* photolabeling was preserved in *post hoc* histological slices (Fig. 1E, middle), and that phototagged cells can be reliably registered across *in vivo* z-stacks and *post hoc* confocal images (Fig. 1E, right, Fig. 1F, Fig. S1C,D, supplementary movie 2). In addition, we segmented the nuclei from both the *in vivo* and *ex vivo* z-stacks to observe the average axial and lateral fluorescence profiles of phototagged nuclei, demonstrating single-nucleus resolution (Fig 1F, right).

To demonstrate the compatibility of the 2P-NucTag construct with downstream cell sorting and transcriptomic applications, we further prepared CA1 PN samples and subjected them to FACS and Meso-seq (fig. S2). We collected 3 samples of AAV-infected nuclei ('Injected, infected'), 2 samples of nuclei that were not infected, but the mice received AAV injections ('Injected, non-infected'), and 2 negative control samples from mice that were not injected with AAVs ('non-injected'). For each sample, 50-100 nuclei were collected and analyzed in bulk via Meso-seq. Between these groups, we found that they have comparable sequencing statistics in terms of total reads and percent of uniquely mapped reads (fig. S2A); notably, these sequencing statistics are similar to those when we applied Meso-seq to AAV-infected visual cortex samples in previous studies[49,54]. Gene expression levels between groups were highly correlated (fig. S2B&C). Reads of representative genes are also similar between groups except for the inhibitory neuron markers (fig. S2D). Thus, our results demonstrate that the previously established Meso-seq

protocol works well with low numbers of hippocampal PN nuclei as input, and that the 2P-NucTag construct does not cause changes in transcriptional properties in AAV-infected PNs. We also obtained *ex vivo* whole-cell patch-clamp intracellular recordings from phototagged and control CA1 PNs (fig. S3A) in acute hippocampal slices: comparing H2B-TAG (infected and photoactivated), H2B (infected), and control (non-infected) cells revealed no differences in intrinsic properties (fig. S3B-G). Together, these results confirm the utility of 2P-NucTag to analyze *ex vivo* the cellular properties of single CA1 PNs that were functionally characterized and phototagged *in vivo* and demonstrate that the 2P-NucTag construct does not affect the intrinsic physiological properties of these PNs. Thus, 2P-NucTag enables high-throughput, indelible phototagging of neuronal nuclei *in vivo* that can also be identified via our registration pipeline for *post hoc* analyses *ex vivo*. Phototagged nuclei further enable downstream RNA-sequencing and *ex vivo* electrophysiology analyses.

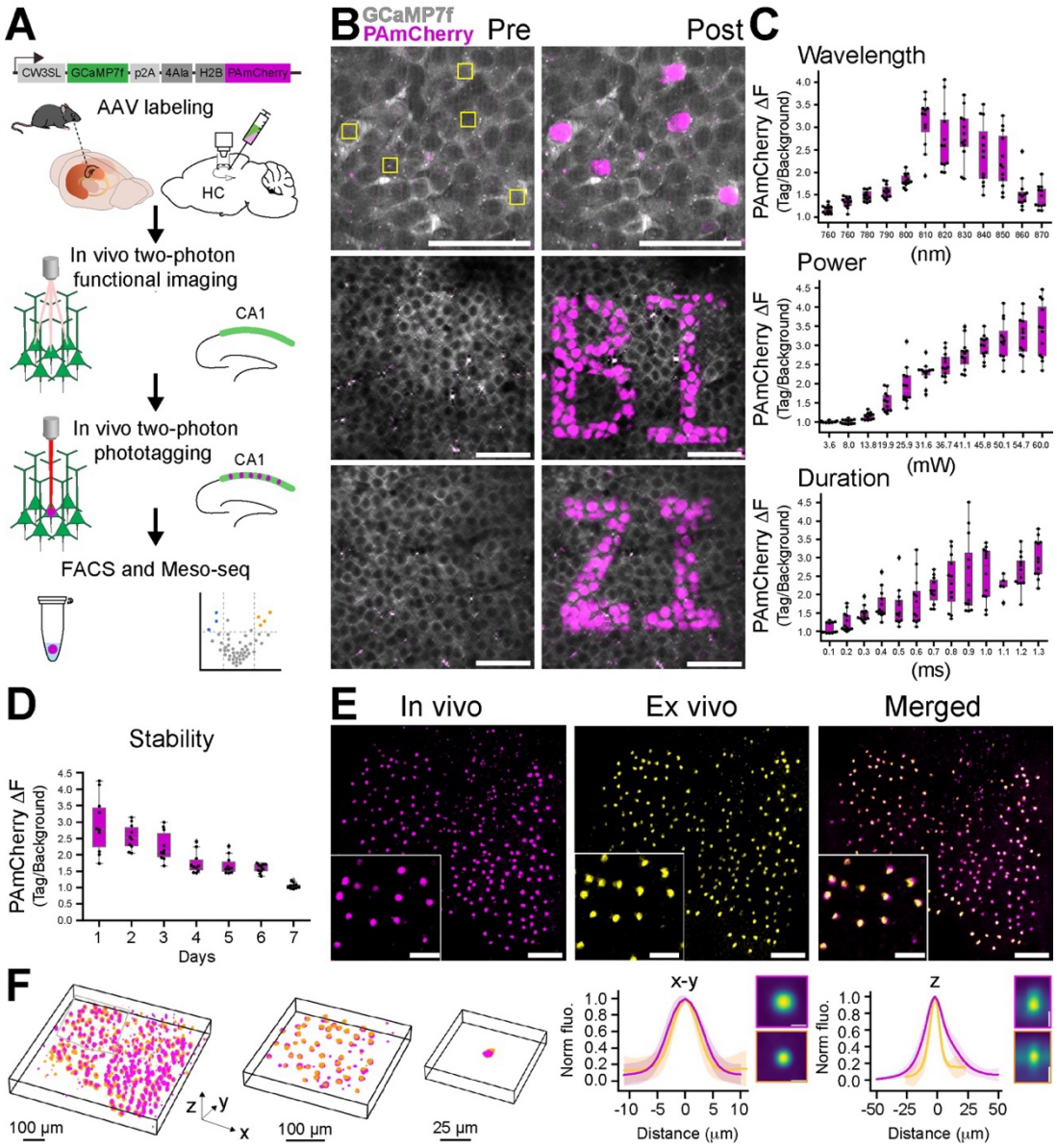

Fig. 1. *In vivo* two-photon phototagging with 2P-NucTag. (**A**) Schematics of the 2P-NucTag pipeline. Top: bicistronic rAAV construct, injection to the hippocampus. Middle: *in vivo* two-photon (2P) GCaMP- Ca^{2+} population imaging followed by 2P PAmCherry photoactivation, fluorescence-activated cell sorting (*FACS*), and mesoscale sequencing (*Meso-seq*). (**B**) Top: representative *in vivo* time-averaged (6 frames average) 2P images of individual cells before (*Pre*) and after (*Post*) *in vivo* two-photon PAmCherry photoactivation in the CA1 pyramidal layer of the mouse dorsal hippocampus. Individual nuclei were photoactivated with 810-nm 2P laser chessboard scanning region-of-interest (ROI, yellow boxes) over target nuclei (70 x 70 pixel for each ROI, 0.1 µm/px, 1.3 ms/px total pixel dwell time, 6,370 ms total scan time per ROI. Laser power was 40 mW measured after the objective) with a 3-dimensional acousto-optical deflector microscope (3D-AOD). Gray: GCaMP7f (940 nm excitation), magenta: PAmCherry, (1040 nm excitation). Scale bar: 50 µm. Middle and bottom: imprints of letters '*BI*' and *'ZI'* following patterned *in vivo* two-photon photoactivation in the hippocampal CA1 pyramidal layer (scale bar, 50 µm). (**C**) Characterization of *in vivo* 2P photoactivation parameters for PAmCherry: duration, wavelength, laser power (measured after the objective) (n = 11-12 cells per condition). Relative change in PAmCherry red fluorescence (ΔF) is based on normalizing the tagged nuclei fluorescence to the fluorescence of neighboring untagged nuclei measured with 1040 nm excitation. **(D)** *In vivo* stability of the PAmCherry fluorescence signal over days after a single photoactivation scan (n = 8 cells). (**E**) Representative time-averaged images from z-stacks of photoactivated nuclei *in vivo* (magenta: PAmCherry, scale bar: 100 µm). Middle: *ex vivo post hoc* confocal z-stack image of the same field of view (FOV, magenta: PAmCherry). Right: registered *in vivo* and *ex vivo* images following non-rigid image transformation (magenta: *in vivo*, yellow: *ex vivo*, see methods). (**F**) Left: 3D overlay of tagged nuclei registered between *in vivo* (magenta) and *ex vivo* (yellow) z-stacks with increasing lateral resolution (as in *E*). Gray box represents the segmented area for subsequent images. Right: normalized lateral (*x-y*, left) and axial (*z*, right) fluorescence profiles (mean ± s.e.m.) of tagged cells *in vivo* (magenta, n = 1 mouse, 200 cells). Yellow: mean ± s.e.m. of *ex vivo* confocal images (as in *E* and *F,* same mouse and nuclei). Inset: (x-y) top: average *in vivo* maximum z–projection, bottom: average *ex vivo* maximum z projection; (z) top: average *in vivo* lateral projection, bottom: average *ex vivo* lateral projection. Scale bar: 10 µm. Boxplots show the 25^{th}, 50^{th} (median), and 75^{th} quartile ranges, with the whiskers extending to 1.5 interquartile ranges below or above the 25^{th} or 75^{th} quartiles, respectively. Outliers are defined as values extending beyond the whisker ranges.

Phototagging of functionally identified PNs in the hippocampus *in vivo*

To demonstrate the utility of 2P-NucTag for on-demand labeling of *functionally* defined PNs, we deployed 2P-NucTag to label hippocampal PNs located in the dorsal CA1, a region with well-established spatial coding heterogeneity[28–35]. Only a subset of hippocampal PNs ('place cells') exhibit reliable spatial tuning for a location ('place field') during exploration[55]. Place cell identity is thought to be randomly allocated within a seemingly homogenous population of PNs[56], and is highly dynamic, with the majority of cells changing their spatial tuning properties over the timescale of days[35–40]. The spatial intermingling of place cells with active-non-place and silent PNs[57–60] without apparent topographical organization in the densely packed CA1 pyramidal layer[35,60] allows us to test the utility of 2P-NucTag in selectively labeling PNs occupying these distinct functional states.

To selectively label CA1 PNs with distinct spatial coding properties, we trained mice in a spatial navigation task for water rewards in a linear virtual reality environment[61–63] and performed *in vivo* 2P GCaMP- Ca^{2+} imaging of PNs (Fig. 2A,B). We reliably detected GCaMP-Ca^{2+} signals from individual PNs (Fig. 2C, fig. S1B). We analyzed the basic characteristics of GCaMP7f in the bicistronic 2P-NucTag construct and confirmed that the basic properties of the indicator are similar to those for a previously published single-construct version of GCaMP7f[50] (fig. S1B). We classified all PNs in the imaging FOV as place cells, active non-place cells, or silent cells (Fig. 2D, see methods). Following the functional identification of PNs, we generated spatial masks of place cell locations in the FOV and used these masks to guide 2P phototagging of all identified place cells in the imaging FOV with 3D-AOD (n = 5 mice, Fig. 2E, F). In a separate set of mice (n = 4 mice), we tagged a subset of CA1 PNs that exhibited no detectable activity during the

imaging session ('silent cells' see methods, fig. S4, Table 1), whereby we photoactivated a similar number of silent cells as place cells in each mouse despite their higher abundance in our recordings. We could reliably register photoactivated nuclei to the spatial masks of functional profiles that we generated for cells of interest (Fig. 2F, fig. S4D,E, fig. S5, see methods). We quantified the number of place cell nuclei that we successfully photoactivated and compared it to the number of spatial masks we generated for each animal. Across all animals, we were able to phototag 93.3% ± 4.2% (n = 5 mice) of all targeted place cell masks and 53.6 ± 6.6% of all targeted silent cells (Fig. 2F).

To assess the accuracy of 2P photoactivation in the densely packed CA1 pyramidal layer, we quantified the number of tagged nuclei for each targeted cell and found that photoactivation is well restricted to the target cells' nuclei with limited off-target labeled nuclei, which had a smaller increase in mCherry fluorescence (Fig. 2G, fig. S4F). Furthermore, 2P GCaMP-Ca^{2+} imaging at 940 nm over the course of the imaging session (18-40 min, see methods) resulted in a minimal increase of mCherry red fluorescence (37.4% ± 16.2% ΔF/F, n = 5 mice, Fig. 2G, fig. S4F, 1070 nm excitation). Targeted place cells demonstrated higher mCherry fluorescence (487.7% ± 67% ΔF/F, n = 5 mice) compared to the background and off-target fluorescence (210.4% ± 32.1% ΔF/F, n = 5 mice). Mouse behavior and the quality of GCaMP recordings were consistent between the two groups of mice (silent and place, fig. S4G-I). In sum, 2P-NucTag offers on-demand *in vivo* labeling of functionally defined neurons with high efficacy and accuracy.

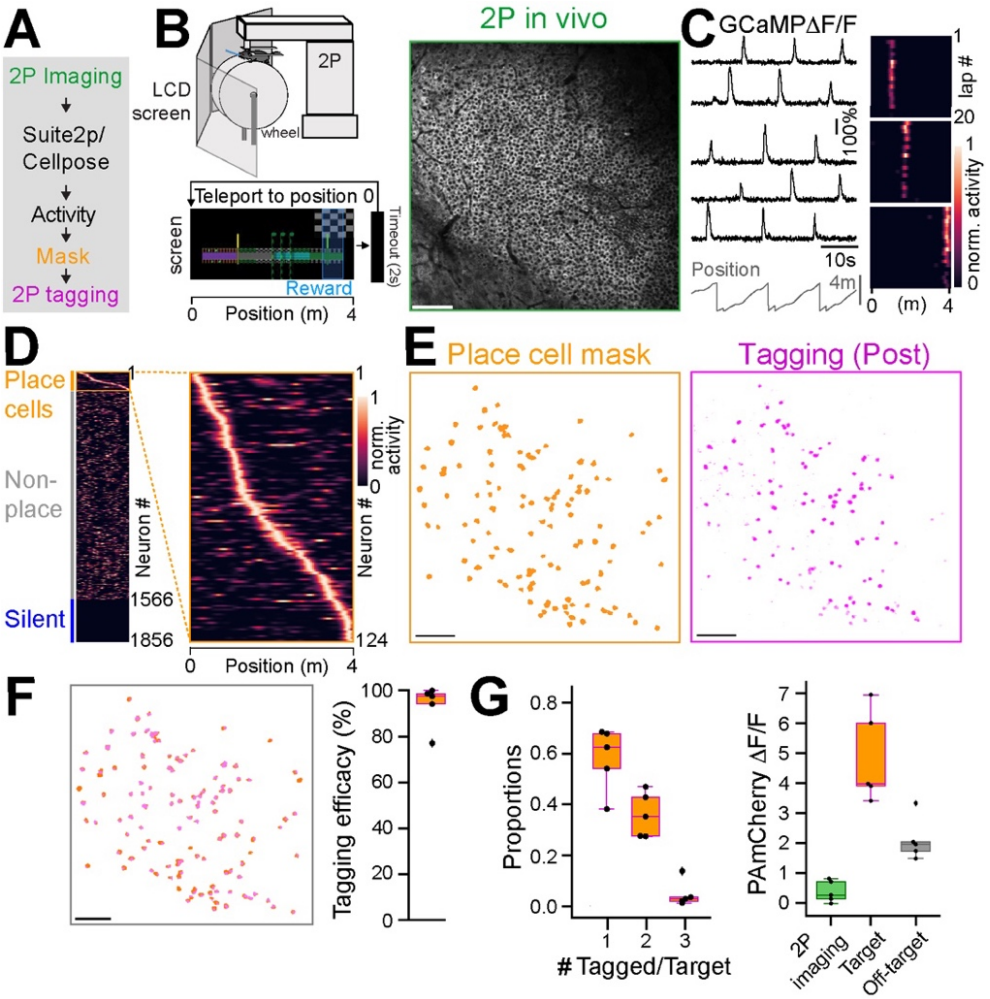

Fig. 2. Selective phototagging of place cells in the hippocampus with 2P-NucTag. (**A**) Pipeline for two-photon (*2P*) phototagging of functionally identified hippocampal neurons during spatial navigation. (**B**) Left: schematics of 2P imaging setup in virtual reality (VR). Head-fixed mice are trained to run for a water reward in a 4-m long linear VR corridor projected onto LCD screens surrounding the animal. At the end of the corridor, mice are teleported back to the start position after a 2-second delay. Right: example 2P field of view (FOV) of GcaMP in the CA1 pyramidal layer. Scale bar: 100 μm.(**C**) Left: Traces of relative GcaMP-Ca^{2+} fluorescence changes (ΔF/F) from five example CA1 place cells during VR spatial navigation. Right: heatmaps of normalized ΔF/F activity from three example place cells over 20 laps during VR navigation. (**D**) Left: heatmap of all CA1PNs detected with Suite2p/Cellpose in the FOV shown in *B*. Identified place cells are marked with an orange box. Right: Zoomed-in heatmap of place cell tuning curves. (**E**) Left: spatial mask (orange) of identified place cells from *D* in the FOV. Right: PAmCherry fluorescence (magenta) of tagged nuclei after 2P phototagging. Scale bar: 100 μm. (**F**) Left: overlay of spatial masks of identified CA1PNs and tagged nuclei for the FOVs in *E*. Scale bar: 100 μm. Right: tagging efficacy, defined as the fraction of successfully tagged place cell nuclei (93.3% ± 4.2%, mean ± s.e.m., n = 5 mice). (**G**) Left: Proportion of single, double, and triple-tagged nuclei following phototagging of a single place cell. Right: relative change in PAmCherry red fluorescence (1070 nm excitation) for non-tagged cells in the FOV after 2P imaging (green), after 2P phototagging of targeted place cell nuclei (orange) and off-target nuclei (gray, n = 5 mice). Boxplots show the 25th, 50th (median), and 75th quartile ranges, with the whiskers extending to 1.5 interquartile ranges below or above the 25th or 75th quartiles, respectively. Outliers are defined as values extending beyond the whisker ranges.

Transcriptional profiling of functionally identified CA1 PNs

The ability to tag single cells *in vivo* enables a powerful new form of hypothesis generation and testing by which, for the first time, functionally defined cells with known behavioral relevance can be isolated and characterized. To demonstrate this, we sought to interrogate transcriptomic signatures of the functional 'place' and 'silent' cell states of CA1 PN.

Following functional imaging and in vivo phototagging, brain tissue containing dorsal CA1 was collected, and nuclei were dissociated and stained with CoreLite 488-conjugated NeuN antibody and DAPI to identify neuronal nuclei (see methods). In vivo photoactivated nuclei were identified by bright mCherry fluorescence (Fig. 3A,B) and were separated out from the non-photolabeled neuronal nuclei by fluorescent-activated cell sorting (FACS), enabling an estimated 18-67% recovery of all photoactivated nuclei identified during in vivo imaging (Fig. 3C). To determine whether place cells and silent cells differed in their gene expression programs, RNA-seq was performed on both populations of sorted nuclei by Meso-seq, an approach that enables reliable identification of differentially expressed genes in ultra-low amounts for FACS-isolated neuronal nuclei (i.e., tens of sorted nuclei, Fig. 3D)[49]. After isolating the phototagged nuclei, libraries were generated with the Meso-seq protocol and were sequenced at a depth of 40-60 million reads per library (fig. S4A). Reads were aligned to the mm39 mouse genome assembly with STAR, counted with HTSeq, and gene expression patterns were compared between nuclei isolated from mice in which place cells were tagged (n = 5) and mice in which silent cells were tagged (n = 4) via PyDESeq2.

In both populations, canonical CA1 PN marker genes (*Map2, Actb, Dlg4, Neurod6*) were highly and non-differentially expressed, supporting the cellular precision of our tagging approach (Fig. 3E,F). However, 219 genes were identified as differentially expressed in a significant manner between silent and place cells (Fig. 3E, see Supplementary Table). In some cases, specific genes were reliably identified in one group and absent from the other. For example, no counts were measured for an inward rectifying potassium channel (*Kcnj12*) in place cell samples, but the expression was present in all silent cell samples. Conversely, mRNA from two transcription factors of the zinc finger protein family (*Zfp84, Zfp977*) and a protocadherin gene (*Pcdha8*) were identified in every place cell sample and in none of the silent cell samples (Fig. 3G). To ensure that the observed differences did not arise from other sources of variation between our place and

silent cell samples, we performed several additional analyses, investigating potential contributions from anatomical positioning, transcriptional responses to neural activity, cellular health, and sex. Genes for which expression has been shown to vary along the dorsoventral and proximodistal axes of CA1 were not differentially expressed in our place and silent-cell nuclei (fig. S6B)[29,64], suggesting that any observed differences are unlikely to have originated from differences in anatomical positioning during cell tagging. Immediate early gene levels (*Fos, Arc, Egr1, Npas4*) did not differ (Fig. 3F)[65–67], suggesting that the ≥25 hours between the last behavior session and tissue collection was sufficient time to eliminate the impact of immediate transcriptional responses to neural activity. Apoptotic gene counts also did not differ between the two groups (Fig. S6), suggesting that there was no significant difference in cell health between the silent and place cells[68]. Finally, to test for gross molecular differences between the cells in an unbiased fashion, we performed gene set enrichment analysis on fifty hallmark gene sets from the mouse molecular signatures database and found no significant differences between silent and place cell samples (Fig. 3H). Since the sex of mice used for both groups was not well balanced in this study (Table 1), it is possible that gene expression differences originated from the mouse's sex rather than the functional identity of the tagged cells. To explore this possibility, we compared gene expression in randomly tagged (function-blind) CA1 PN nuclei from both male (n = 2) and female (n = 2) mice. Although we found that known sex-specific genes[69] were differentially expressed, none of the genes identified as enriched in either silent or place cells had significant sex associations (Fig. S7C,D). To assess if our differentially expressed genes-of-interest signified upregulation or downregulation in place cells from the mean or, conversely, downregulation or upregulation in silent cells, we compared gene expression in place cells and silent cells to randomly tagged cells separately. We found that *Kcnj12* expression was enriched in silent cells compared to random cells. *Zfp84* and *Zfp977* were significantly downregulated in silent cells compared to random cells, while *Pcdha8* expression was significantly enriched in place cells compared to random cells (Fig. S7A,B).

We were particularly interested in the enriched expression of *Kcnj12* in silent cells: this gene encodes for an inward rectifying potassium channel[70–72], its expression pattern suggests that there could be a difference in the intrinsic electrophysiological properties of silent and non-silent cells. Thus, to validate the expression of *Kcnj12*, we first performed RNAScope Fluorescent in situ hybridization (FISH) in brain sections with tagged silent CA1 PNs (fig. S8). Using a probeset for detecting *Kcnj12* (fig. S8A,B), we then quantified the number of *Kcnj12* puncta in tagged silent cells and a number-matched subset of randomly selected non-tagged cells (n = 170 cells each, 2 mice). This analyses confirmed that tagged silent cells indeed express more *Kcnj12* transcripts than non-tagged cells (fig. S8C,D). To test whether silent CA1 PNs differ in their intrinsic electrophysiological properties, we next performed targeted whole-cell patch-clamp recordings from tagged silent and nontagged cells *ex vivo* in acute hippocampal slices (fig. S9). These analyses revealed that silent cells had a hyperpolarized resting membrane potential as compared to random non-tagged cells.

In summary, our results demonstrate that, for the first time, we have the ability to apply differential gene expression analyses on populations of cells from the same brain region that differ only by their functional identity. Moreover, our in-depth sequencing approach can indeed identify differences in gene expression between functionally defined cells tagged *in vivo* with 2P-NucTag, and these functionally-molecularly distinct neurons can be further analyzed at high-resolution with downstream molecular, cellular and electrophysiological approaches.

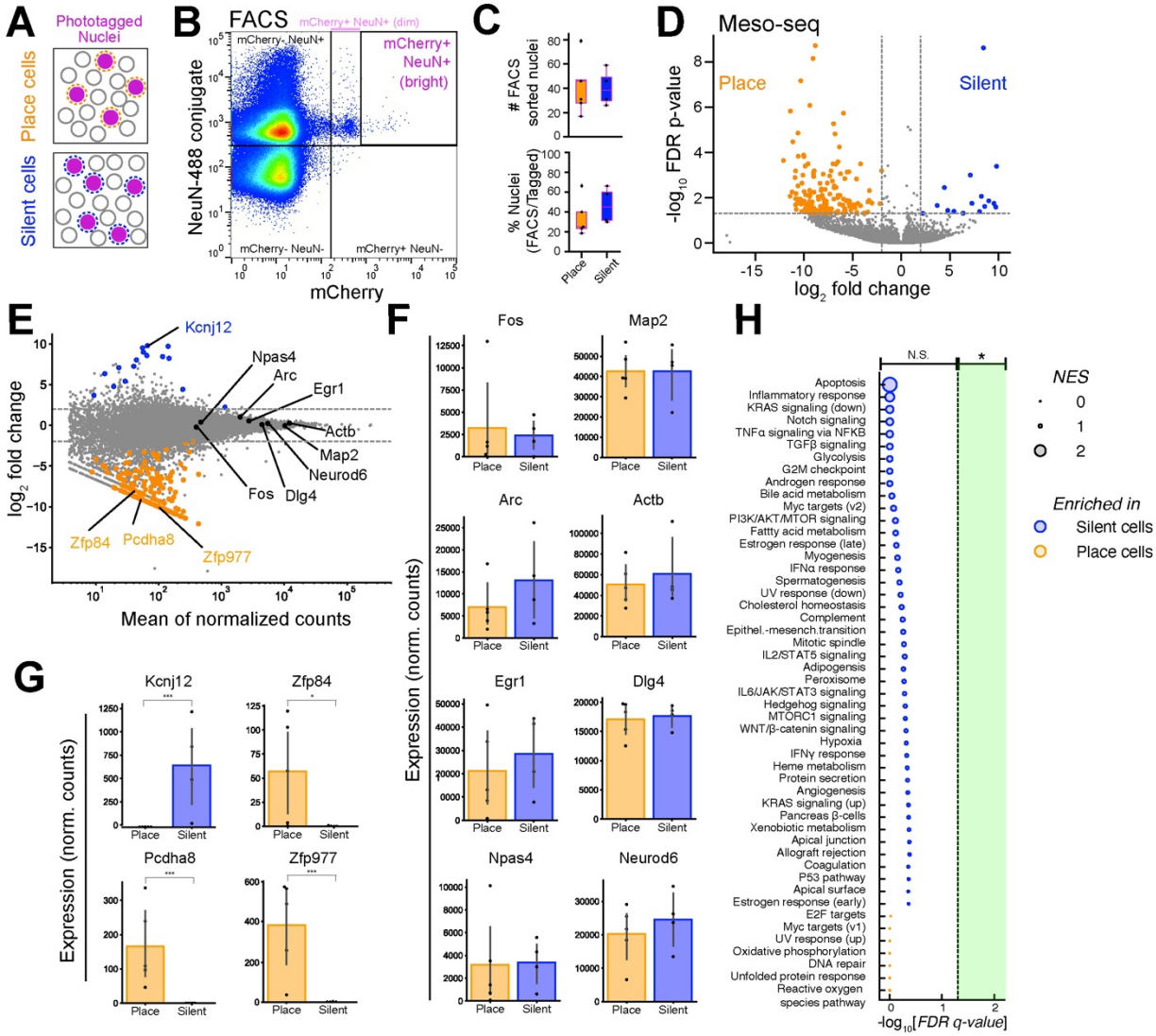

Fig. 3. *Post hoc* transcriptional profiling of phototagged place and silent cells. (**A**) Schematics of *in vivo* photoactivated nuclei. 'Place' cell sample and 'Silent' cell samples from different mice were collected for FACS and Meso-seq. (**B**) Representative FACS graph. Gating for mCherry was set after the first 5000 events of DAPI+ nuclei to the border of the 'dim' mCherry+ population to separate out the sparse and high-intensity mCherry+ population. Bright mCherry+ NeuN+ populations were collected as the photoactivated nuclei. (**C**) Top: number of FACS sorted nuclei from 'place' and 'silent' samples (n = 9, 17-79 sorted nuclei, 40.3 ± 6.43, mean ± s.e.m.). Bottom: Proportion of FACS sorted nuclei compared to the number of *in vivo* photoactivated nuclei (n = 9, 18.45% to 66.39% FACS recovery, 39.94 ± 6.30%, mean ± s.e.m.). (**D**) Volcano plot of Meso-seq differential expressed gene (DEG) analysis for 'place' and 'silent' cells (significantly different genes are shown in orange and blue. Orange: enriched in place cells; blue: enriched in silent cells). (**E**) Meso-seq MA plot depicting DeSeq2 normalized gene counts versus log2 fold change of silent/place samples. Genes that are significantly different are labeled in orange and blue (same as above). Genes shown in panels *F* & *G* are highlighted and labeled in *E*. (**F**) Bar graph showing the normalized counts for genes that are not differentially expressed (FDR adjusted p-value. *<0.05, **<0.001**,** ***<0.001, PyDeSeq2. Otherwise, comparisons are not significant). (**G**) Bar graph showing the normalized counts for differentially expressed genes (FDR adjusted p-value. *<0.05, **<0.001**,**<0.001, PyDeSeq2. Otherwise, comparisons are not significant). (**H**) Gene ontology analysis performed on all differentially expressed genes. Vertical line: FDR-adjusted p value of 0.05. NES = normalized enrichment score. Boxplots show the 25th, 50th (median), and 75th quartile ranges, with the

whiskers extending to 1.5 interquartile ranges below or above the 25th or 75th quartiles, respectively. Outliers are defined as values extending beyond the whisker ranges.

Discussion

Comprehensive molecular-functional characterization of cortical circuits remains elusive due to the limited toolkit for correlated *in vivo* functional recording techniques and *post hoc* molecular analyses. Here, we introduced 2P-NucTag, a robust function-forward approach combining *in vivo* imaging, on-demand single-nucleus tagging, and *post hoc* transcriptomics. This pipeline seamlessly integrates 2P Ca^{2+} imaging in behaving mice, 2P phototagging of functionally identified neuronal nuclei, FACS-isolation of the tagged nuclei, and subsequent analyses of isolated nuclei with a mesoscale protocol for in-depth transcriptomics in ultra-sparse neuronal populations.

We demonstrated the utility of 2P-NucTag by selectively labeling and analyzing transcriptional profiles of two functionally orthogonal subpopulations of hippocampal PNs, 'place cells' and 'silent cells'. Unexpectedly, our data revealed a number of genes that are putatively differentially expressed between these distinctly transient physiological cell states[35–40]. The low counts for some of these genes indicate that further replicates may be necessary to confirm their differential expression. Nevertheless, several of the identified genes generate new hypotheses about the origin of these two functional cell identities, which had canonically been considered otherwise indistinguishable. Specifically, the marked enrichment of an inward rectifying potassium channel (*Kcnj12*) in silent cells implicates intrinsic excitability as a putative cellular and molecular mechanism that might influence the activity and spatial coding properties of PNs in the hippocampus.[73] This result and model align with previous studies showing that that intrinsic excitability is a major factor governing cell recruitment to a memory trace[74–77]. Future experiments will be needed to test the function of specific genes that we have identified with our tagging technique, which could reveal a non-random allocation of functional activity onto CA1 PNs during learning and experience. In particular, future application of 2P-NucTag to selectively tag PNs with highly distinct and shared coding properties, including place field propensity[28,36,78,79], recruitment to 'replay' and 'preplay' events[80–83] and participation in functionally interconnected subnetworks[84] may uncover genetic and developmental[85–89] backbones of functional heterogeneity among PNs.

Our results showed low levels of unintended PAmCherry photoactivation during functional imaging as well as limited off-target labeling. Although the activation level of these background and off-target neurons is lower and the actual targeted cells' photoactivation can be distinguished through *in vivo* imaging, FACS, and *ex vivo* confocal imaging, this presents a complicating factor when setting gating levels for FACS. Therefore, in this study, we sorted only the brightest population of cells with high levels of mCherry fluorescence (Fig. 3B). Future optimization of the phototagging vector and the 2P tagging parameters could help refine photoactivation specificity. Nevertheless, our proof-of-principle implementation of 2P-NucTag to functionally heterogeneous hippocampal PNs that are spatially intermixed in the densely packed pyramidal cell layer demonstrates that 2P-NucTag should be readily applicable to neocortical tissue with lower cell density.

Beyond this first application to the mouse hippocampus, 2P-NucTag will have broad appeal for correlated structural, molecular, and functional analysis of neural circuits. Our approach offers the ability to bridge molecular and functional architecture in cortical circuits and facilitates the identification of candidate genes that determine the distinct circuit functions of neurons. Our pipeline should be potent for uncovering molecular signatures of phenomenologically described

neuronal subsets with distinct feature selectivity, task-related activity, and longitudinal stability that are spatially intermixed in cortical circuits[21,23,25,90–94].

Our approach can also be extended in multiple directions. Firstly, as 2P-NucTag is compatible with *post hoc* histological analysis of imaged tissue, it can be combined with spatially resolved transcriptomics[17,18,95,96] following automated registration between *in vivo* and *post hoc* images[97,98]. Secondly, *ex vivo* electrophysiological recordings from tagged cells would allow for detailed *ex vivo* physiological, molecular, and anatomical characterization of *in vivo* tagged cells[9–12]. Thirdly, in principle, our pipeline is compatible with sc/snRNA-seq, given the single-cell resolution achieved through targeted photoactivation. Fourthly, further optimization of phototagging with multi-color fluorophores may enable simultaneous labeling of multiple functionally defined cell types in the same animal, while cytosolic tags[46] can promote anatomical labeling of subcellular axonal or dendritic compartments for downstream structural and connectivity analyses. Lastly, the on-demand nature of photoactivation and its days-long stability following single photoactivation makes our pipeline a valuable tool for precise analysis of transcriptional trajectories following cellular plasticity events during, for example, behavioral learning[52,67,99,100].

In summary, our novel phototagging approach aids in expanding our knowledge of the fundamental relationship between the molecular and functional architecture of mammalian cortical circuits and allows for the identification of candidate genes that determine the distinct circuit functions of seemingly homogeneous populations of neurons. Beyond the proof-of-principle implementation to hippocampal circuits, our approach provides a general framework and roadmap for linking genes to cells across neural circuits and model organisms.

Materials and Methods

Animals

All animal care and experiment procedures were in accordance with the guidelines of the National Institute of Health. Animal protocols were approved by the Columbia University Institutional Animal Care and Use Committee and the Weizmann Institute of Science Institutional Animal Care and Use Committee. Mice were group-housed under normal lighting conditions in a 12-hour light/dark cycle. *Ad libitum* water was provided until the beginning of training for the spatial navigation task.

Plasmids and Viral Constructs

pAAV-CW3SL-GCaMP7f-4Ala-H2B-PAmCherry was generated by standard cloning techniques. GCaMP7f was PCR-amplified from Addgene plasmid #104492 with a 3' Primer that contained sequences encoding four Alanine residues and the P2A sequence (both in frame with the coding sequence of GCaMP7f). H2B-PAmCherry was amplified from Addgene plasmid #133419. The PCR products were then subcloned by Gibson-assembly into Addgene plasmid #61463 after removing EGFP from this plasmid by restriction with ClaI and EcoRI. The sequence of the cloned plasmid was validated by Sanger sequencing and GCaMP7f + 4xAla, P2A and H2B-PAmCherry were all found to be in frame. The plasmid was packaged into AAVDJ at a viral titer of 7.26E+15 essentially as described[49].

Surgery

All procedures were performed with mice under anesthesia using isoflurane (4% induction, 1.5% maintenance in 95% oxygen). Mice's body temperature was maintained using a heating pad both during and after the procedure. Surgeries were performed on a stereotaxic instrument (Kopf Instruments). Before incision, mice were given subcutaneous meloxicam, as well as bupivacaine at the incision site. Doses were calculated based on the animal's weight. An incision above the skull was made to expose bregma and lambda for vertical alignment. Skull surfaces were cleaned and scored to improve dental cement adhesion. For viral injection, a glass capillary loaded with rAAV is attached to a Nanoject device (Drummond Scientific).

For all experimental mice, viruses were injected unilaterally in the left dorsal CA1 at 4 depths using the coordinates: -2.2 AP, -1.75 ML, and -1.2, -1.1, -1.0, -0.9 DV (relative to Bregma). At each depth, 75nl of AAVDJ-CW3SL-GCaMP7f-4Ala-H2B-PAmCherry was injected. After injection, surgical sites were closed with sutures. Three days after injection, the skull was exposed and a 3mm craniotomy was made centered at the same coordinate of the injection site. Dura was removed, and the cortex was slowly aspirated with continuous irrigation of cold 1X PBS until the fiber tract above the hippocampus was visible. A 3-mm imaging cannula fitted with a 3mm glass coverslip was implanted over the craniotomy site. Cannulas were secured by Vetbond. A custom titanium headpost for head-fixation was secured first with C&B Metabond (Parkell) and then dental acrylic. At the end of each procedure, the mice received a 1.0 ml saline injection subcutaneously and recovered in their home cage with heating applied. Mice were monitored for 3 days after the procedure.

Behavior paradigm

Mice were first water-deprived and habituated to handling and head fixation at least 7 days after implant surgery. They were then exposed to a 4-m long linear virtual reality (VR) corridor[61–63] that

stayed consistent in the training and recording. At the end of the environment, an inter-trial interval of 2 seconds of blank screen was included before the start of the next lap. For the next 10-14 days, mice were trained to run through the virtual environment and lick for a 5% sucrose reward. The rewards were first randomly distributed across the environment, and the number of rewards was slowly reduced from 30 at the beginning of the training to 2 when the mouse was deemed ready for recording. The final reward location was fixed toward the end of the VR environment. Mice were trained to run at least 30-60 laps in the environment. During behavioral imaging, mice were imagined during a single VR session (range: 18-40 min, 26 ± 3 min, n = 9 mice).

In vivo two-photon imaging and data processing

2P functional imaging was conducted using an 8-kHz resonant scanner (Bruker) and a 16x near-infrared (NIR) water immersion objective (Nikon, 0.8 NA, 3.0-mm working distance). For population imaging, a field of view of 700 μm x 700 μm was acquired at 30 Hz, 512 x 512 pixels using a 940-nm laser (Chameleon Ultra II, Coherent, 45-91 mW after the objective). Red (PAmCherry) and green (GCaMP7f) channels were separated by an emission cube set (green, HQ525/70 m-2p; red, HQ607/45 m-2p; 575dcxr, Chroma Technology), and fluorescence signals were collected with GaAsP photomultiplier tube modules (7422P-40, Hamamatsu). Following the acquisition of two-photon imaging data, Ca^{2+} imaging data was structured and aligned with behavior data using the SIMA analysis package[101]. CA1 ROIs were detected using the Suite2p (v0.14.2) package[102]. To allow detection of all potential ROIs regardless of their activities during the recording, Suite2p was run with Cellpose ('anatomical_only')[103] for the ROI detection step. The pre-trained cyto2 model included in the published Cellpose package was used for ROI detection. When capturing two-photon z-stack images of photoactivated PAmCherry nuclei, a fixed wavelength 1070-nm laser (Fidelity-2W, Coherent) was used for excitation.

In vivo two-photon phototagging

Photoactivation was conducted using a three-dimensional random-access acousto-optical (3D-AOD) microscope (3D Atlas, Femtonics)[19,53]. Mice were head-fixed and anesthetized with isoflurane to minimize motion and increase the spatial precision of phototagging. The same 16x NIR water immersion objective was used to find the same field of view as in the functional recordings. Photoactivation was performed at 810-nm (Chameleon Ultra II, Coherent). Two-photon images of the mCherry red fluorescence were taken before and after photoactivation using a 1040-nm excitation laser (Alcor 1040-5W, Spark Lasers). Red (mCherry) and green (GCaMP7f) channels were separated by an emission cube set (green, HQ520/60 m-2p; red, HQ650/160 m-2p; 565dcxr, Chroma Technology), and fluorescence signals were collected with GaAsP photomultiplier tube modules (7422P-40, Hamamatsu). Two-photon images of the mCherry red fluorescence were taken before and after photoactivation using a 1040-nm excitation laser (Alcor 1040-5W, Spark Lasers). For photoactivation, a 7 x 7 μm, 0.1 μm /pixel scanning pattern was placed on the cell to be photoactivated. Each pixel was activated for a total dwell time of 1.3 ms with a laser power of 40 mW. This gave the total scanning time of each cell at 6,370 ms. Following photoactivation, a z-stack was taken for each mouse to assess the photoactivation efficacy.

For phototagging of 'place cells' and 'silent cells', photoactivation experiments were conducted a day after functional recording sessions. The imaging field of view of 700 μm x 700 μm was matched between the 3D-AOD microscope and the time-averaged GCaMP image from functional recording. After confirming the same field of view as functional imaging, the viewport was

zoomed in to a dimension of 250 x 250 μm for more effective identification of targeted cells according to the generated spatial masks. Following photoactivation of all cells, a z-stack image was taken for each mouse to confirm the tagging accuracy. Mice were given *ad libitum* water after functional imaging for at least 12 hours before photoactivation. During the session, mice were monitored every 10 min for breathing rate and reflexes. Heating and eye ointment were applied. Following the photoactivation, mice were returned to the home cage to recover with a heating pad.

Tissue dissociation and preparation of nuclei for FACS

For all animals used in this manuscript, tissue collection was performed at the same time of the day (5 pm) and at least one hour after the photoactivation. Mice were euthanized using CO_2. The headpost and metal cannulas were removed. Dorsal CA1 of the hippocampus was collected by first using a 3-mm biopsy punch to cut a circular section of tissue the same size as the craniotomy. A microspatula was used to remove the shallow section of tissue that contained dorsal CA1. The tissue was placed in a 1.5 ml RNase-free Eppendorf tube and snap-frozen in liquid nitrogen. Tissues were stored at -80 ℃ until the start of the nuclei isolation.

Nuclei were prepared for FACS sorting essentially as previously described (*46*). In short, to isolate the nuclei, each collected tissue was transferred to a dounce tissue homogenizer (DWK Life Sciences) with 1ml of homogenization buffer (10 mM Tris Buffer, 250mM Sucrose, 25 mM KCl, 5mM $MgCl_2$, 0.1mM DTT, 0.1% Triton X-100, 1X Protease Inhibitor Cocktail, and 40U/μl RNAsin Plus RNase Inhibitor in nuclease-free water). Loose and tight pestles were used to break apart the tissue 10 times each. Following the homogenization, 1 ml of homogenization buffer was added to each dounce, and the homogenate was pipetted up and down to further break apart tissue before being passed through a 30 μm cell strainer (Miltenyi Biotec) and collected in a 15 ml conical tube.

The homogenate was centrifuged at 4 ℃, 700g for 8 min. The supernatants were then removed from the visible cell pellet. The cell pellet was resuspended with 800 μl of blocking buffer (1X PBS, 1% BSA, and 40 U/μl RNasin Plus RNase Inhibitor in nuclease-free water). The resuspension was incubated on ice for 15 min and transferred to a 1.5 ml tube. 2 μl of CoraLite Plus 488-conjugated NeuN Monoclonal antibody (Proteintech) was added to the resuspension, and incubated on an orbital rotator at 4 ℃ for 30 min. After incubation, nuclei were centrifuged at 4 ℃, 700 g for 8 min. The supernatant was removed, and the cell pellet was resuspended with 1000 μl of blocking buffer. DAPI was added to the suspension at a final concentration of 0.001 mg/ml, and the samples were passed through a 40- μm Flowmi Cell Strainer (Bel-Art). All samples were kept on ice until the start of FACS.

Fluorescence-activated cell sorting (FACS)

The sorting was performed at the Zuckerman Institute Flow Cytometry Core using a MoFlo Astrios Cell Sorter (Beckman Coulter). Event rates were kept between 5000-10000 events per second. The cell sorter uses a linear array of lasers ordered as 640 nm, 488 nm, 561 nm, 532 nm, 405 nm and 355 nm from top to bottom. For experiments described in this manuscript, 488 nm, 561 nm and 405 nm lasers were used to detect the fluorescence of NeuN, mCherry and DAPI respectively. A gating control sample was used to set the gates for DAPI, NeuN and mCherry. Dissociated nuclei were passed through the cell sorter to collect those with high mCherry signals. Bright and Dim mCherry gates were determined after the first 5000 events of DAPI-positive nuclei. Gating was

set to collect only the sparse and bright population that showed a high mCherry signal, and these events were collected as the photoactivated nuclei ('Bright' mCherry). Nuclei were collected into SMART-Seq CDS sorting buffer that contains 1X lysis buffer, SMART-Seq Oligo-dT and RNase inhibitor. All sorted samples were kept on dry ice as recommended by the SMART-Seq protocol until the start of first-strand synthesis. To increase the likelihood of collecting nuclei in this ultra-sparse population, aborted events for 'Bright' mCherry and all events for other positive mCherry were collected into the blocking buffer (200 ml). This suspension was passed through the sorter again with the same fluorescence gating following the completion of the first sorting to capture the bright nuclei.

Meso-seq

We followed the previously published Meso-seq protocol[49] with minor modifications. Sorted nuclei were collected in a lysis buffer following the SMART-Seq mRNA LP (with UMIs) protocol. Reverse transcription for cDNA was followed by cDNA amplification using 17-18 PCR cycles. Purified cDNA was prepared for sequencing the library using the SMART-seq Library Preparation Kit. Libraries were amplified using 14 PCR cycles. The concentration of the final library was determined using a Qubit3.0 Fluorometer (Invitrogen), and the average DNA fragment size was determined using a Bioanalyzer (Agilent). Sequencing was performed on a NextSeq 2000 sequencer with P2-100 reagents (Illumina). Libraries were diluted and pooled according to the recommendation of the sequencing kit.

***Ex vivo* electrophysiology in acute hippocampus slices**

Mice were transcardially perfused with ice-cold sucrose dissection media (26 mM NaHCO3, 1.25 mM NaH2PO4, 2.5 mM KCl, 10 mM MgSO4, 11 mM glucose, 0.5 mM CaCl2, 234 mM sucrose; 340 mOsm). Brains were then dissected and sliced, while being kept in ice-cold sucrose dissection media, into coronal sections (300 µm thick) containing the hippocampal CA1 using a Leica VT1200S vibratome. Slices were incubated in high osmotic concentrated artificial cerebrospinal fluid (aCSF) (28.08 mM NaHCO3, 1.35 mM NaH2PO4, 132.84 mM NaCl, 3.24 mM KCl, 1.08 mM MgCl2, 11.88 mM glucose, 2.16 mM CaCl2; 320 mOsm) at 32°C for 30 minutes immediately after slicing. Then, slices were incubated in normal osmotic concentrated artificial cerebrospinal fluid (26 mM NaHCO3, 1.25 mM NaH2PO4, 123 mM NaCl, 3 mM KCl, 1 mM MgCl2, 11 mM glucose, 2 mM CaCl2; 300 mOsm) at 32°C for 30 minutes and subsequently at room temperature. All solutions were saturated with 95%-O2/5%-CO2, and slices were used within 6 hours of preparation. Whole-cell patch-clamp recordings were performed in aCSF at 32°C from neurons in the visual cortex. Recording pipettes were pulled from borosilicate glass capillary tubing with filaments (OD 1.50 mm, ID 0.86 mm, length 10 cm) using a P-700 micropipette puller (Sutter Instruments) and yielded tips of 3–5 MΩ resistance. Recordings were sampled at 20 kHz and filtered at 3 kHz. Data were acquired via Clampex10 using a Multiclamp 700B amplifier and digitized with an Axon Digidata 1550B data acquisition board (Axon Instruments). Tagged neurons were identified based on mCherry nuclear fluorescence using a pE-300 white MB LED light (CoolLED) with GYR (525-660nm) spectrum, combined with an Olympus Cy5 Filter Cube Set (ex. 604-644nm; em. 672-712nm).

To ensure that the recorded cells were indeed phototagged neurons, Alexa 594 Hydrazide (10uM) was added to the internal solution to allow co-localization of the fluorescence of the tagged neuron to the one that was patched using confocal imaging in PFA(4%)-fixed slices.

For measuring the intrinsic properties, the following internal solution was used: 135 mM k-gluconate, 4mM KCl, 10mM HEPES, 10mM Pcreatine, 4mM Mg-ATP, 4mM GTP-Na and 2mM Na2-ATP. Intrinsic properties were calculated by giving 1.2s long current steps (20pA).

Tissue collection and processing for *in situ* hybridization

Mice were anesthetized with isoflurane and transcardially perfused with 20 mL of ice-cold 0.01M phosphate base saline (PBS, Sigma) followed by 20 mL ice-cold 4% paraformaldehyde (PFA, Electron Microscopy Sciences) in PBS. Brains were post-fixed in 4% PFA for 24 hours and then saturated with a 10%, 20%, and 30% sucrose solution sequentially over 48 hours until they sunk to the bottom of each successive solution. 30% Sucrose-saturated brains were then embedded in OCT (Optimal Cutting Temperature Compound, Sakura, cat#4583), frozen, stored overnight at -80 ℃, and sliced transversely at 20 µm thickness with a cryostat. Sections were stored at -80 ℃ on slides and used for RNAScope *in situ* hybridization.

RNAScope Fluorescent In Situ Hybridization (FISH)

20 µm fixed frozen sections of the frozen tissue block were taken and the RNAScope™ Multiplex Fluorescent Reagent Kit v2 – User Manual was followed (cat#: 323100). The Kcnj12 targeting probe was designed and generated by Advanced Cell Diagnostics Inc (Kcnj12: cat#: 525171-C3, Entrez Gene ID: 16515, GenBank Accession #: NM_010603.6). Slides were sequentially dehydrated using ethanol solutions of increasing concentrations (50%, 70% and 100%) for 5 min at RT. 5-8 drops of H_2O_2 were added to each sample, followed by a 10 min incubation at RT. After rinsing with distilled water, antigen retrieval was performed for 5 min at 99°C. To digest sections, RNAscope Protease III was applied to the sections for 30 min at 40°C. The probes were hybridized for 2 h at 40°C and amplified with AMP1 (30 min), AMP2 (30 min), AMP3 (15 min); each was incubated at 40°C. The probe was fluorescently tagged with 1:2000 TSA Vivid Fluorophore 650 (PN 323273). Slides were counterstained with DAPI for a nuclear stain to identify viable cells and mounted in ProLong Gold Antifade Mountant. 20 µm sections were imaged in 3 µm z-steps using an inverted confocal microscope with 20x oil objective (A1 HD25, Nikon Instruments Inc.).

Data analysis

Quantification and Statistical Analysis

All statistical details for comparisons are described in the text. No statistical methods were used to determine sample sizes. Boxplots show the 25th, 50th (median), and 75th quartile ranges with the whiskers extending to 1.5 interquartile ranges below or above the 25th or 75th quartiles, respectively. Outliers are defined as values extending beyond the whisker ranges. For comparisons between two populations with non-normal distributions, the Mann-Whitney U test was used. For comparisons between gene expression datasets, the Wald test followed by multiple corrections via the Benjamini and Hochberg method was used as described in PyDESeq2[104].

Event detection

Fluorescence GCaMP traces were deconvolved using OASIS for fast nonnegative deconvolution[105]. As in ref.[106], these putative spike events were filtered at 3 median absolute deviations (MAD) above the raw trace, using a predetermined signal decay constant of 400 ms. The binarized signal was used to qualify whether a neuron was active at the respective frame. In

our analysis, we do not claim to uncover true spiking events in these neurons but use deconvolution for denoising and diminishing Ca^{2+} autocorrelation.

Spatial tuning curves

The virtual environment was divided into 100 evenly spaced bins (4 cm), which were then utilized to bin a histogram of each cell's neuronal activity. Neuronal activity was filtered to include activity from when the animal was running above 3 cm/s and to exclude activity during the 2-sec teleportation at the end of the 4-m track. The spatial tuning curves were normalized for the animal's occupancy and then smoothed with a Gaussian kernel (σ = 12 cm) to obtain a smoothed activity estimate.

Place/silent cell detection

Place fields were detected by identifying locations in the virtual environment where a neuron was more active than expected by chance. We circularly shifted each neuron's deconvolved spike trace and recomputed the smoothed, trial-averaged spatial tuning curve of the shifted trace to generate a shuffled null tuning curve per cell. We repeated this procedure 1000 times in order to calculate the 95th percentile of null tuning values at every spatial bin to generate a threshold for a $p<0.05$ significance curve. Spatial tuning curves that surpassed the null threshold were marked as candidate place fields, and the place field width was calculated as the total bins where the tuning curve exceeded the shuffled null tuning curve. To restrict our analysis to neurons with specific firing fields, we additionally required that place fields have a width greater than 8 cm and less than one-third of the virtual environment (1.3 m). To ensure that the place field activity was stable, we also required that all place cells had activity for at least 20 laps.

If the binarized signal trace for a cell did not have any detected events via OASIS deconvolution, the cell was classified as silent.

PAmCherry fluorescence quantification

(Fig. 1C,D)

Red PAmCherry fluorescence was calculated as the tagged nuclei fluorescence versus the background fluorescence of nearby untagged nuclei for the image taken post-tagging. This was intentional to control for periodic two-photon imaging at 940 nm that may cause an increase in fluorescence for every cell in the FOV.

Fluorescence intensity distribution analysis

(Fig. 1F, right)

Red PAmCherry tagged nuclei were segmented using Cellpose with manual curation performed within the Cellpose GUI for the max axial projections of the *ex vivo* z-stack (excitation: 568 nm nm) and *in vivo* z-stack (1070 nm) to generate masks. The masks were used to segment the 3-D volumes of each cell and the respective fluorescence profiles were normalized and aligned based on the peak value. The average and standard error were computed based on the aligned fluorescence profiles of the cells from the respective *in vivo* and *ex vivo* volumes.

Two-photon background excitation fluorescence quantification

[Fig. 2G(2P imaging), fig. S2F (2P imaging)]

Changes in PAmCherry red fluorescence due to two-photon fluorescence excitation at 940 nm from functional recordings were examined. By taking the average fluorescence of the pre-imaging red channel image and post-imaging red channel image detected at 1070 nm, ΔF/F was computed using the change in fluorescence between the average post-imaging red image and pre-imaging red image divided by the pre-imaging red image average [(post – pre) / pre].

Tagged and Off-target Fluorescence Change Quantification

[Fig. 1G (target, off-target), fig. S2F (target, off-target)]

Tagged nuclei were segmented using Cellpose with manual curation performed within the Cellpose GUI for the max axial projection of the *in vivo* -stacks from each mouse to generate masks. These masks were used to segment a 3D volume for each nuclei that was targeted across all mice. The number of off-target nuclei was manually quantified laterally and axially per targeted nuclei.

Masks for off-target nuclei were hand-drawn on the max axial projection to exclude the targeted nuclei and were at most 2 nuclei bodies away. Experimental background fluorescence masks were drawn on the surrounding areas with successfully tagged cells and excluded all targeted and off-target nuclei.

The fluorescence values for all tagged and off-target nuclei were percentile-filtered to exclude the lower 10% of fluorescence values to account for vignetting effects and mask inhomogeneities over the max axial projection. The percentile filtered fluorescence values for tagged and off-target nuclei were averaged, and the ΔF/F was computed by taking the difference between the average tagged or off-target fluorescence value and the average background fluorescence and then normalizing by the average background fluorescence [tagged: (tagged – background)/background; off-target: (off-target – background)/ background]. Note this background value is distinct from the one described in the two-photon background excitation fluorescence quantification.

GCaMP quantifications

[fig. S1AB, fig. S2I]

Frequency: The total number of deconvolved events by OASIS was normalized by the total duration of the recording.

For each cell, the average transient was segmented within a 15-second time window and computed by averaging along aligned deconvolved spike times. If multiple detected events were within 50 frames, the events were treated as a single transient. A cell was only used if there were at least 3 detected transients within the total trace to exclude cells without obvious GCaMP-Ca^{2+} dynamics. For each average transient, we computed the median value based on 5 seconds pre peak transient and the range 5-10 seconds after the peak transient (given that the average transient took significantly less than 5 seconds to resolve) to act as the baseline.

Amplitude: The difference between the max value of the average transient and the baseline value was computed

Half-Rise Time: The time between the half-max value prior to the max and the max value of each transient was computed. We excluded cells where the average half-max value was not observed prior to the transient and performed 99^{th} percentile filtering to remove extreme outliers. Due to

sampling rate limitations, we cannot comment on the true half-rise time, so these are approximations.
Half-Decay Time: We computed the time between the max and the half-max value following the transient peak. To remove extreme outliers, we performed 99th percentile filtering.

Image denoising

To correct for vertical scanning artifacts, we utilized combined wavelet and Fourier filters described in ref.[107] [github: https://github.com/DHI-GRAS/rmstripes]. Symlet 20 wavelets were used with varying levels of decomposition (2-5) for discrete wavelet transform to perform vertical striping correction in static images.

***In vivo* and *ex vivo* image registration**

The *in vivo* and *ex vivo* images were transformed into 3D volumetric images for registration. The *in vivo* sequential images were concatenated across the z direction, i.e., depth, stacking the 2D images into a volumetric representation using MATLAB. The *ex vivo* images composed of 2D slices in each section were concatenated into 3D volumetric data using a stitching algorithm developed as a precursor for automatic *ex vivo* and *in vivo* registration[97] [github: https://github.com/ShuonanChen/multimodal_image_registration]. The discontinuity between *ex vivo* sections results in an unknown spatial correlation between them, requiring registration between sections. Common cells between sections (i.e., sections one and two) were used as reference markers for registration. The common cells were manually selected using a Napari GUI in Python and were utilized as reference markers to inform the scaling and affine transformations to be applied The scaling and affine transformations were run automatically, transforming the second section to align with the first section. In the scaling transformation, the relative distances between cells in a section were compared to the cells in the first section, inducing an enlargement or shrinkage of the second section to match the first. In the affine transformation, the second section was geometrically transformed to align correctly with the first section. The transformations were obtained and applied to each slice in each section, and then each slice was concatenated together across the z direction to form a volumetric image.

The *in vivo* and *ex vivo* 3D images were then adjusted to have a uniform pixel size in all dimensions (1 μm in x,y and z), ensuring matching FOVs, and equivalent resolution across both images. Time-averaged representations of *in vivo* and *ex vivo* volumetric stacks were attained by employing maximum intensity projection (MIP) representations in FIJI, compressing the stacks into 2D images[108]. The *in vivo* and *ex vivo* registration was carried out using a non-rigid registration algorithm for the (i) 3D volumetric stacks and (ii) MIP (2D) images. All cells common to both *in vivo* and *ex vivo* images were manually selected as the centroid of each cell using a Napari GUI in Python. The common cells were utilized as features to inform the scaling, affine transformation, and deformation transformation, which were applied to the *ex vivo* image. The scaling and affine transformations were run automatically. In the scaling transformation, the distances between the cells in the *ex vivo* image and matching cells in *in vivo* induce enlargement or shrinking of the *ex vivo* image to match the *in vivo* image. In the affine transformation, the *ex vivo* image was geometrically transformed to align with the *in vivo* image using the matching cells. The images in the GUI are updated to reflect the changes induced by the scaling and affine transformation. The deformation transformation uses a vector field, smoothed with Gaussian filtering, to move cells and deform the image, ensuring features in the transformed *ex vivo* image align with the *in vivo*

image. The deformation transformation was iteratively employed, beginning with a Gaussian kernel size of 100, reducing to a kernel size of less than 10, with the user inspecting the alignment and making manual adjustments to the cell centroid position in the GUI. The completed transformed *ex vivo* image was then overlaid with the *in vivo* image.

RNA-sequencing data analysis

Sequencing data from the Illumina Sequencer was first post-processed through the Illumina DRAGEN secondary analysis pipeline to de-multiplex based on a unique index for each sample. RNA-seq reads were aligned to the mouse genome (mm39) using STAR[109]. Unique reads were counted using HTSeq[110]. HTSeq generated reads were then analyzed for differential expression using PyDESeq2[104]. Following HTSeq counts, any genes with expression in less than 3 samples were discarded. FDR adjusted p values were used to determine significantly different genes. Both 'place' and 'silent' cell samples were analyzed against randomly tagged, function-blind, and sex-matched samples generated from mouse CA1 tissue in the same way as described in this methods section above. The same comparisons were made to identify differentially expressed genes for 'place' versus 'random', and 'silent' versus 'random'. Within this 'random' dataset, male-female samples were compared to identify differentially expressed genes influenced by sex. All DEGs from these comparisons were cross-referenced to find common hits.

Gene set enrichment analysis

Gene set enrichment analysis was performed on HTSeq-generated counts using the gseapy package[111] in Python 3.11. Enrichment of the fifty hallmark pathways from the molecular signatures database for *Mus musculus* (version 2023.2) was compared in place and silent cells. Comparisons were done with t-test and 1000 permutations.

RNAscope FISH signal quantification and statistical analysis

Cell detection was performed using Cellpose, which identified individual cells based on DAPI nuclear staining. This segmentation was manually curated. The detection and quantification of RNAscope probe signals was performed using QuPath's Subcellular Detection tool (version 0.5.1). The detection threshold was set to ensure accurate identification of signal dots. To compare mean RNA expression levels between tagged and non-tagged cells, a Mann-Whitney U Test was used. For this comparison, a random subset of non-tagged cells was selected to match the number of tagged cells.

Acknowledgments

We thank Ira Schieren and Max Wallach (Columbia, Zuckerman Institute, ZI) and Dr. Efrat Hagai (Flow Cytometry Unit, Weizmann Institute of Science) for help with FACS, Columbia University Herbert Irvine Comprehensive Cancer Center Molecular Pathology Shared Resources for help with the Bioanalyzer, the Bendesky lab in ZI and Dr. Hadas Keren-Shaul, Revital Ronen at the Weizmann, Nancy and Stephen Grand Israel National Center (G-INCPM) for help with RNA-sequencing, George Zakka (Losonczy Lab) for technical support, Erica Rodriguez (Salzman lab) in ZI for help with tissue processing, the Lomvardas lab in ZI for sequencing advice and help, the Gogos lab in ZI for molecular bench space, the Polleux lab in ZI for help with cell counter and confocal microscopy. We thank Franck Polleux, Joseph Gogos, Steven A. Siegelbaum, and Darcy Peterka (Columbia, ZI) for feedback on the manuscript.

Funding

SAH is supported by the Burroughs Wellcome Fund. DK is supported by a fellowship from the Israel Ministry of Absorption (IMOAb) and by the Horowitz Foundation. AL is supported by NIMHR01MH124047, NIMHR01MH124867, NINDSR01NS121106, NINDSU01NS115530, NINDSR01NS133381, NINDSR01NS131728, NIARF1AG080818. IS is supported by an ISF personal grant (2354/19) and a BSF US-Israel binational grant (2021281) and he is the incumbent of the Friends and Linda and Richard Price Career Development Chair and a scholar in the Zuckerman STEM leadership program.

Author contributions:

Conceptualization: JS, BN, DK, AL, IS
Methodology: JS, BN, DK, BR, SAH, ET, KCK, MECP, HCY, BMS, AX
Investigation: JS, BR
Software: BR, JS, TSM, CKO, EV
Formal analysis: JS, BR, TSM, ET, AL
Validation: JS, BN, DK, KCK, BR
Visualization: JS, BR, TSM, SAH, CKO, ET, MECP, KC-KM
Data curation: JS, BR, ET, TSM
Funding acquisition: AL, IS
Resources: AL, IS
Project administration: AL, IS
Supervision: AL, IS
Writing – original draft: JS, BN, TSM, AL, IS
Writing – review & editing: all authors

Competing interests: Authors declare that they have no competing interests.

Data and material availability All data, code, and materials used to generate figures and perform statistical tests are available at Zenodo or NWB upon publication. RNA-seq data will be made freely available on NCBI GEO upon acceptance of our manuscript. Accession numbers for publicly available reagents are included in supplementary materials and methods.

References

1. Zeng, H. What is a cell type and how to define it? *Cell* **185**, 2739–2755 (2022).
2. Macosko, E. Z. *et al.* Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell* **161**, 1202–1214 (2015).
3. Tasic, B. *et al.* Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat Neurosci* **19**, 335–346 (2016).
4. Zeisel, A. *et al.* Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347**, 1138–1142 (2015).
5. Harris, K. D. *et al.* Classes and continua of hippocampal CA1 inhibitory neurons revealed by single-cell transcriptomics. *PLoS Biol* **16**, e2006387 (2018).
6. Piwecka, M., Rajewsky, N. & Rybak-Wolf, A. Single-cell and spatial transcriptomics: deciphering brain complexity in health and disease. *Nat Rev Neurol* **19**, 346–362 (2023).
7. Zhang, M. *et al.* Molecularly defined and spatially resolved cell atlas of the whole mouse brain. *Nature* **624**, 343–354 (2023).
8. Moses, L. & Pachter, L. Museum of spatial transcriptomics. *Nat Methods* **19**, 534–546 (2022).
9. Cadwell, C. R. *et al.* Multimodal profiling of single-cell morphology, electrophysiology, and gene expression using Patch-seq. *Nat Protoc* **12**, 2531–2553 (2017).
10. Cossell, L. *et al.* Functional organization of excitatory synaptic strength in primary visual cortex. *Nature* **518**, 399–403 (2015).
11. Ko, H. *et al.* Functional specificity of local synaptic connections in neocortical networks. *Nature* **473**, 87–91 (2011).
12. Scala, F. *et al.* Phenotypic variation of transcriptomic cell types in mouse motor cortex. *Nature* **598**, 144–150 (2021).

13. Bock, D. D. *et al.* Network anatomy and in vivo physiology of visual cortical neurons. *Nature* **471**, 177–182 (2011).

14. Taniguchi, H. *et al.* A Resource of Cre Driver Lines for Genetic Targeting of GABAergic Neurons in Cerebral Cortex. *Neuron* **71**, 995–1013 (2011).

15. Vormstein-Schneider, D. *et al.* Viral manipulation of functionally distinct interneurons in mice, non-human primates and humans. *Nat Neurosci* **23**, 1629–1636 (2020).

16. O'Toole, S. M., Oyibo, H. K. & Keller, G. B. Molecularly targetable cell types in mouse visual cortex have distinguishable prediction error responses. *Neuron* **111**, 2918-2928.e8 (2023).

17. Bugeon, S. *et al.* A transcriptomic axis predicts state modulation of cortical interneurons. *Nature* **607**, 330–338 (2022).

18. Condylis, C. *et al.* Dense functional and molecular readout of a circuit hub in sensory cortex. *Science* **375**, eabl5981 (2022).

19. Geiller, T. *et al.* Large-Scale 3D Two-Photon Imaging of Molecularly Identified CA1 Interneuron Dynamics in Behaving Mice. *Neuron* **108**, 968-983.e9 (2020).

20. Vancura, B., Geiller, T., Grosmark, A., Zhao, V. & Losonczy, A. Inhibitory control of sharp-wave ripple duration during learning in hippocampal recurrent networks. *Nat Neurosci* 1–10 (2023) doi:10.1038/s41593-023-01306-7.

21. Cohen-Kashi Malina, K. *et al.* NDNF interneurons in layer 1 gain-modulate whole cortical columns according to an animal's behavioral state. *Neuron* **109**, 2150-2164.e5 (2021).

22. Hansel, D. & Vreeswijk, C. van. The Mechanism of Orientation Selectivity in Primary Visual Cortex without a Functional Map. *J. Neurosci.* **32**, 4049–4064 (2012).

838 23. Hirokawa, J., Vaughan, A., Masset, P., Ott, T. & Kepecs, A. Frontal cortex neuron types
839 categorically encode single decision variables. *Nature* **576**, 446–451 (2019).
840 24. Mancinelli, S. & Lodato, S. Decoding neuronal diversity in the developing cerebral cortex:
841 from single cells to functional networks. *Current Opinion in Neurobiology* **53**, 146–155
842 (2018).
843 25. Ohki, K. & Reid, R. C. Specificity and randomness in the visual cortex. *Current Opinion in*
844 *Neurobiology* **17**, 401–407 (2007).
845 26. Peng, H. *et al.* Morphological diversity of single neurons in molecularly defined cell types.
846 *Nature* **598**, 174–181 (2021).
847 27. Treves, A. & Rolls, E. T. Computational analysis of the role of the hippocampus in memory.
848 *Hippocampus* **4**, 374–391 (1994).
849 28. Buzsáki, G. & Mizuseki, K. The log-dynamic brain: how skewed distributions affect network
850 operations. *Nat Rev Neurosci* **15**, 264–278 (2014).
851 29. Cembrowski, M. S. *et al.* Spatial Gene-Expression Gradients Underlie Prominent
852 Heterogeneity of CA1 Pyramidal Neurons. *Neuron* **89**, 351–368 (2016).
853 30. Mallory, C. S. & Giocomo, L. M. Heterogeneity in hippocampal place coding. *Current*
854 *Opinion in Neurobiology* **49**, 158–167 (2018).
855 31. Pettit, N. L., Yap, E.-L., Greenberg, M. E. & Harvey, C. D. Fos ensembles encode and shape
856 stable spatial maps in the hippocampus. *Nature* **609**, 327–334 (2022).
857 32. Soltesz, I. & Losonczy, A. CA1 pyramidal cell diversity enabling parallel information
858 processing in the hippocampus. *Nat Neurosci* **21**, 484–493 (2018).
859 33. Sosa, M., Gillespie, A. K. & Frank, L. M. Neural Activity Patterns Underlying Spatial
860 Coding in the Hippocampus. in *Behavioral Neuroscience of Learning and Memory* (eds.

Clark, R. E. & Martin, S. J.) 43–100 (Springer International Publishing, Cham, 2018). doi:10.1007/7854_2016_462.

34. Tanaka, K. Z. *et al.* The hippocampal engram maps experience but not place. *Science* **361**, 392–397 (2018).

35. Ziv, Y. *et al.* Long-term dynamics of CA1 hippocampal place codes. *Nat Neurosci* **16**, 264–266 (2013).

36. Lee, J. S., Briguglio, J. J., Cohen, J. D., Romani, S. & Lee, A. K. The Statistical Structure of the Hippocampal Code for Space as a Function of Time, Context, and Value. *Cell* **183**, 620-635.e22 (2020).

37. Kentros, C. G., Agnihotri, N. T., Streater, S., Hawkins, R. D. & Kandel, E. R. Increased Attention to Spatial Context Increases Both Place Field Stability and Spatial Memory. *Neuron* **42**, 283–295 (2004).

38. Grosmark, A. D., Sparks, F. T., Davis, M. J. & Losonczy, A. Reactivation predicts the consolidation of unbiased long-term cognitive maps. *Nat Neurosci* **24**, 1574–1585 (2021).

39. Keinath, A. T., Mosser, C.-A. & Brandon, M. P. The representation of context in mouse hippocampus is preserved despite neural drift. *Nat Commun* **13**, 2415 (2022).

40. Colgin, L. L., Moser, E. I. & Moser, M.-B. Understanding memory through hippocampal remapping. *Trends in Neurosciences* **31**, 469–477 (2008).

41. Moeyaert, B. *et al.* Improved methods for marking active neuron populations. *Nat Commun* **9**, 4440 (2018).

42. DeNardo, L. A. *et al.* Temporal evolution of cortical ensembles promoting remote memory retrieval. *Nat Neurosci* **22**, 460–469 (2019).

883 43. Guenthner, C. J., Miyamichi, K., Yang, H. H., Heller, H. C. & Luo, L. Permanent Genetic
884 Access to Transiently Active Neurons via TRAP: Targeted Recombination in Active
885 Populations. *Neuron* **78**, 773–784 (2013).
886 44. Lee, D., Hyun, J. H., Jung, K., Hannan, P. & Kwon, H.-B. A calcium- and light-gated switch
887 to induce gene expression in activated neurons. *Nat Biotechnol* **35**, 858–863 (2017).
888 45. Sørensen, A. T. *et al.* A robust activity marking system for exploring active neuronal
889 ensembles. *eLife* **5**, e13918 (2016).
890 46. Datta, S. R. *et al.* The Drosophila pheromone cVA activates a sexually dimorphic neural
891 circuit. *Nature* **452**, 473–477 (2008).
892 47. Lien, A. D. & Scanziani, M. In vivo Labeling of Constellations of Functionally Identified
893 Neurons for Targeted in vitro Recordings. *Front Neural Circuits* **5**, 16 (2011).
894 48. Lee, D., Kume, M. & Holy, T. E. Sensory coding mechanisms revealed by optical tagging of
895 physiologically defined neuronal types. *Science* **366**, 1384–1389 (2019).
896 49. Apelblat, D., Roethler, O., Bitan, L., Keren-Shaul, H. & Spiegel, I. Meso-seq for in-depth
897 transcriptomics in ultra-low amounts of FACS-purified neuronal nuclei. *Cell Reports*
898 *Methods* **2**, 100259 (2022).
899 50. Dana, H. *et al.* High-performance calcium sensors for imaging activity in neuronal
900 populations and microcompartments. *Nat Methods* **16**, 649–657 (2019).
901 51. Scheyltjens, I. *et al.* Evaluation of the expression pattern of rAAV2/1, 2/5, 2/7, 2/8, and 2/9
902 serotypes with different promoters in the mouse visual cortex. *Journal of Comparative*
903 *Neurology* **523**, 2019–2042 (2015).

52. Geiller, T., Priestley, J. B. & Losonczy, A. A local circuit-basis for spatial navigation and memory processes in hippocampal area CA1. *Current Opinion in Neurobiology* **79**, 102701 (2023).

53. Katona, G. *et al.* Fast two-photon in vivo imaging with three-dimensional random-access scanning in large tissue volumes. *Nat Methods* **9**, 201–208 (2012).

54. Roethler, O. *et al.* Single genomic enhancers drive experience-dependent GABAergic plasticity to maintain sensory processing in the adult cortex. *Neuron* **111**, 2693-2708.e8 (2023).

55. O'Keefe, J. & Dostrovsky, J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Research* **34**, 171–175 (1971).

56. Witharana, W. k. l. *et al.* Nonuniform allocation of hippocampal neurons to place fields across all hippocampal subfields. *Hippocampus* **26**, 1328–1344 (2016).

57. Rich, P. D., Liaw, H.-P. & Lee, A. K. Large environments reveal the statistical structure governing hippocampal representations. *Science* **345**, 814–817 (2014).

58. Thompson, L. T. & Best, P. J. Place cells and silent cells in the hippocampus of freely-behaving rats. *J. Neurosci.* **9**, 2382–2390 (1989).

59. Wilson, M. A. & McNaughton, B. L. Dynamics of the Hippocampal Ensemble Code for Space. *Science* **261**, 1055–1058 (1993).

60. Zaremba, J. D. *et al.* Impaired hippocampal place cell dynamics in a mouse model of the 22q11.2 deletion. *Nat Neurosci* **20**, 1612–1623 (2017).

61. Bowler, J. C. & Losonczy, A. Direct cortical inputs to hippocampal area CA1 transmit complementary signals for goal-directed navigation. *Neuron* **0**, (2023).

926 62. Bowler, J. C. *et al.* behaviorMate: An Intranet of Things Approach for
927 Adaptable Control of Behavioral and Navigation-Based Experiments. *bioRxiv*
928 2023.12.04.569989 (2024) doi:10.1101/2023.12.04.569989.
929 63. Priestley, J. B., Bowler, J. C., Rolotti, S. V., Fusi, S. & Losonczy, A. Signatures of rapid
930 plasticity in hippocampal CA1 representations during novel experiences. *Neuron*
931 S0896627322002628 (2022) doi:10.1016/j.neuron.2022.03.026.
932 64. Cembrowski, M. S., Wang, L., Sugino, K., Shields, B. C. & Spruston, N. Hipposeq: a
933 comprehensive RNA-seq database of gene expression in hippocampal principal neurons.
934 *eLife* **5**, e14997 (2016).
935 65. Bloodgood, B. L., Sharma, N., Browne, H. A., Trepman, A. Z. & Greenberg, M. E. The
936 activity-dependent transcription factor NPAS4 regulates domain-specific inhibition. *Nature*
937 **503**, 121–125 (2013).
938 66. Sheng, M. & Greenberg, M. E. The regulation and function of c-*fos* and other immediate
939 early genes in the nervous system. *Neuron* **4**, 477–485 (1990).
940 67. Spiegel, I. *et al.* Npas4 Regulates Excitatory-Inhibitory Balance within Neural Circuits
941 through Cell-Type-Specific Gene Programs. *Cell* **157**, 1216–1229 (2014).
942 68. Apoptosis: A Review of Programmed Cell Death - Susan Elmore, 2007.
943 https://journals.sagepub.com/doi/10.1080/01926230701320337.
944 69. Xu, J., Burgoyne, P. S. & Arnold, A. P. Sex differences in sex chromosome gene expression
945 in mouse brain. *Human Molecular Genetics* **11**, 1409–1419 (2002).
946 70. Nichols, C. G. & Lopatin, A. N. INWARD RECTIFIER POTASSIUM CHANNELS. *Annual*
947 *Review of Physiology* **59**, 171–191 (1997).

71. Reimann, F. & Ashcroft, F. M. Inwardly rectifying potassium channels. *Current Opinion in Cell Biology* **11**, 503–508 (1999).

72. Shcherbatyy, V. *et al.* A Digital Atlas of Ion Channel Expression Patterns in the Two-Week-Old Rat Brain. *Neuroinform* **13**, 111–125 (2015).

73. Epsztein, J., Brecht, M. & Lee, A. K. Intracellular Determinants of Hippocampal CA1 Place and Silent Cell Activity in a Novel Environment. *Neuron* **70**, 109–120 (2011).

74. Santoni, G. *et al.* Chromatin plasticity predetermines neuronal eligibility for memory trace formation. *Science* **385**, eadg9982 (2024).

75. Yiu, A. P. *et al.* Neurons Are Recruited to a Memory Trace Based on Relative Neuronal Excitability Immediately before Training. *Neuron* **83**, 722–735 (2014).

76. Josselyn, S. A. & Frankland, P. W. Memory Allocation: Mechanisms and Function. *Annual Review of Neuroscience* **41**, 389–413 (2018).

77. Sano, Y. *et al.* CREB Regulates Memory Allocation in the Insular Cortex. *Current Biology* **24**, 2833–2837 (2014).

78. Mizuseki, K. & Buzsáki, G. Preconfigured, Skewed Distribution of Firing Rates in the Hippocampus and Entorhinal Cortex. *Cell Reports* **4**, 1010–1021 (2013).

79. Alme, C. B. *et al.* Place cells in the hippocampus: Eleven maps for eleven rooms. *Proceedings of the National Academy of Sciences* **111**, 18428–18435 (2014).

80. Dragoi, G. & Tonegawa, S. Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature* **469**, 397–401 (2011).

81. Dragoi, G. & Tonegawa, S. Selection of preconfigured cell assemblies for representation of novel spatial experiences. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**, 20120522 (2014).

971 82. Grosmark, A. D. & Buzsáki, G. Diversity in neural firing dynamics supports both rigid and
972 learned hippocampal sequences. *Science* **351**, 1440–1443 (2016).
973 83. Huszár, R., Zhang, Y., Blockus, H. & Buzsáki, G. Preconfigured dynamics in the
974 hippocampus are guided by embryonic birthdate and rate of neurogenesis. *Nat Neurosci* **25**,
975 1201–1212 (2022).
976 84. Geiller, T. *et al.* Local circuit amplification of spatial selectivity in the hippocampus. *Nature*
977 **601**, 105–109 (2022).
978 85. Cossart, R. & Khazipov, R. How development sculpts hippocampal circuits and function.
979 *Physiological Reviews* **102**, 343–378 (2022).
980 86. Donato, F., Jacobsen, R. I., Moser, M.-B. & Moser, E. I. Stellate cells drive maturation of the
981 entorhinal-hippocampal circuit. *Science* **355**, eaai8178 (2017).
982 87. Deguchi, Y., Donato, F., Galimberti, I., Cabuy, E. & Caroni, P. Temporally matched
983 subpopulations of selectively interconnected principal neurons in the hippocampus. *Nat*
984 *Neurosci* **14**, 495–504 (2011).
985 88. Kveim, V. A. *et al.* Divergent recruitment of developmentally defined neuronal ensembles
986 supports memory dynamics. *Science* **385**, eadk0997 (2024).
987 89. Marissal, T. *et al.* Pioneer glutamatergic cells develop into a morpho-functionally distinct
988 population in the juvenile CA3 hippocampus. *Nat Commun* **3**, 1316 (2012).
989 90. Solstad, T., Boccara, C. N., Kropff, E., Moser, M.-B. & Moser, E. I. Representation of
990 Geometric Borders in the Entorhinal Cortex. *Science* **322**, 1865–1868 (2008).
991 91. Taube, J. S., Muller, R. U. & Ranck, J. B. Head-direction cells recorded from the
992 postsubiculum in freely moving rats. I. Description and quantitative analysis. *J. Neurosci.* **10**,
993 420–435 (1990).

994 92. Eichenbaum, H. Time cells in the hippocampus: a new dimension for mapping memories.
995 *Nat Rev Neurosci* **15**, 732–744 (2014).
996 93. Driscoll, L. N., Duncker, L. & Harvey, C. D. Representational drift: Emerging theories for
997 continual learning and experimental future directions. *Current Opinion in Neurobiology* **76**,
998 102609 (2022).
999 94. Josselyn, S. A. & Tonegawa, S. Memory engrams: Recalling the past and imagining the
1000 future. *Science* **367**, eaaw4325 (2020).
1001 95. Xu, S. *et al.* Behavioral state coding by molecularly defined paraventricular hypothalamic
1002 cell type ensembles. *Science* **370**, eabb2494 (2020).
1003 96. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved,
1004 highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).
1005 97. Chen, S. *et al.* Multimodal Microscopy Image Alignment Using Spatial and Shape
1006 Information and a Branch-and-Bound Algorithm. in *ICASSP 2023 - 2023 IEEE International*
1007 *Conference on Acoustics, Speech and Signal Processing (ICASSP)* 1–5 (2023).
1008 doi:10.1109/ICASSP49357.2023.10096185.
1009 98. Rao, B. Y. *et al.* Non-parametric Vignetting Correction for Sparse Spatial Transcriptomics
1010 Images. in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*
1011 (eds. de Bruijne, M. et al.) 466–475 (Springer International Publishing, Cham, 2021).
1012 doi:10.1007/978-3-030-87237-3_45.
1013 99. Bittner, K. C., Milstein, A. D., Grienberger, C., Romani, S. & Magee, J. C. Behavioral time
1014 scale synaptic plasticity underlies CA1 place fields. *Science* **357**, 1033–1036 (2017).
1015 100. Magee, J. C. & Grienberger, C. Synaptic Plasticity Forms and Functions. *Annual Review of*
1016 *Neuroscience* **43**, 95–117 (2020).

101. Kaifosh, P., Zaremba, J. D., Danielson, N. B. & Losonczy, A. SIMA: Python software for analysis of dynamic fluorescence imaging data. *Front. Neuroinform.* **8**, (2014).

102. Pachitariu, M. *et al.* Suite2p: beyond 10,000 neurons with standard two-photon microscopy. *bioRxiv* 061507 (2017) doi:10.1101/061507.

103. Stringer, C., Wang, T., Michaelos, M. & Pachitariu, M. Cellpose: a generalist algorithm for cellular segmentation. *Nat Methods* **18**, 100–106 (2021).

104. Muzellec, B., Teleńczuk, M., Cabeli, V. & Andreux, M. PyDESeq2: a python package for bulk RNA-seq differential expression analysis. *Bioinformatics* **39**, btad547 (2023).

105. Friedrich, J., Zhou, P. & Paninski, L. Fast online deconvolution of calcium imaging data. *PLOS Computational Biology* **13**, e1005423 (2017).

106. O'Hare, J. K. *et al.* Compartment-specific tuning of dendritic feature selectivity by intracellular Ca2+ release. *Science* **375**, eabm1670 (2022).

107. Münch, B., Trtik, P., Marone, F. & Stampanoni, M. Stripe and ring artifact removal with combined wavelet — Fourier filtering. *Opt. Express, OE* **17**, 8567–8591 (2009).

108. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nature Methods* **9**, 676–682 (2012).

109. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

110. Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).

111. Fang, Z., Liu, X. & Peltz, G. GSEApy: a comprehensive package for performing gene set enrichment analysis in Python. *Bioinformatics* **39**, btac757 (2023).

Supplementary Figures

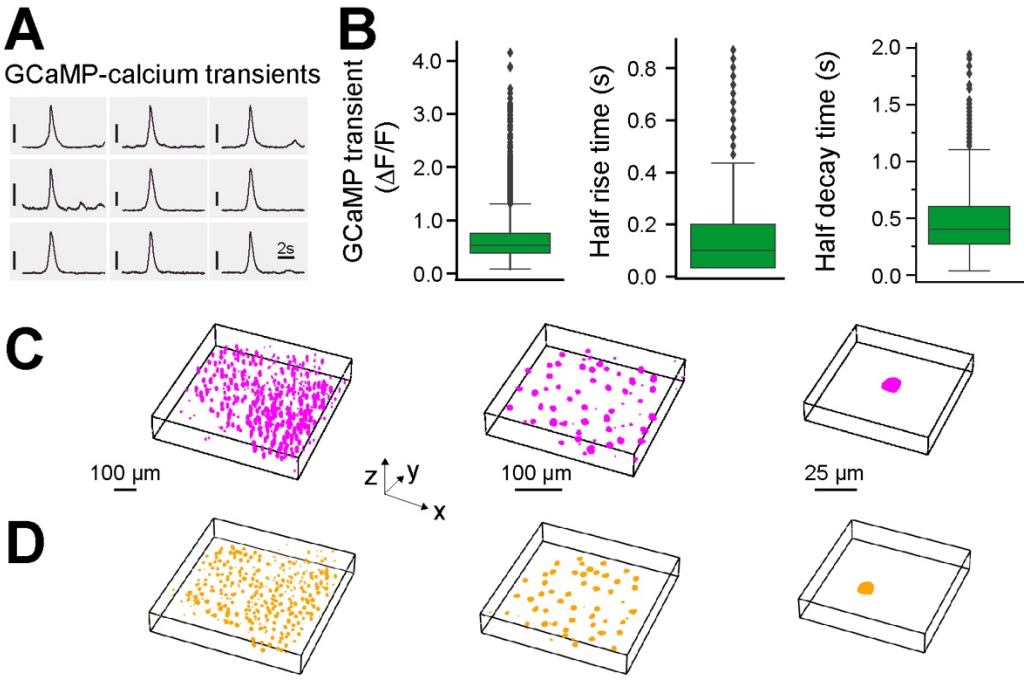

Fig. S1. Additional data on 2P-NucTag. (**A**) Representative average GCaMP-Ca^{2+} transients from nine CA1 PNs, vertical scale bar (50% ΔF/F). (**B**) Left: GCaMP amplitude (54.6% ± 0.3% ΔF/F, n = 8190 cells in 9 mice). Middle: GCaMP half rise time (0.15s ± 0.001s, n = 7797 cells in 9 mice), Right: GCaMP half decay time (0.44s ± 0.003s, n = 8282 cells in 9 mice, 940 nm excitation, See methods for cell exclusion criterion). Boxplots show the 25th, 50th (median), and 75th quartile ranges, with the whiskers extending to 1.5 interquartile ranges below or above the 25th or 75th quartiles, respectively. Outliers are defined as values extending beyond the whisker ranges. (**C**) Left: *in vivo* 3D visualization of the entire field of view (FOV). Middle: subset of *in vivo* 3D visualization. Right: representative cell from *in vivo* 3D visualization. (**D**) Left: *ex vivo* confocal 3D visualization of entire FOV. Middle: subset of *ex vivo* 3D visualization. Right: representative cell from *ex vivo* 3D visualization. *C* and *D* are corresponding to *Figure 1F*.

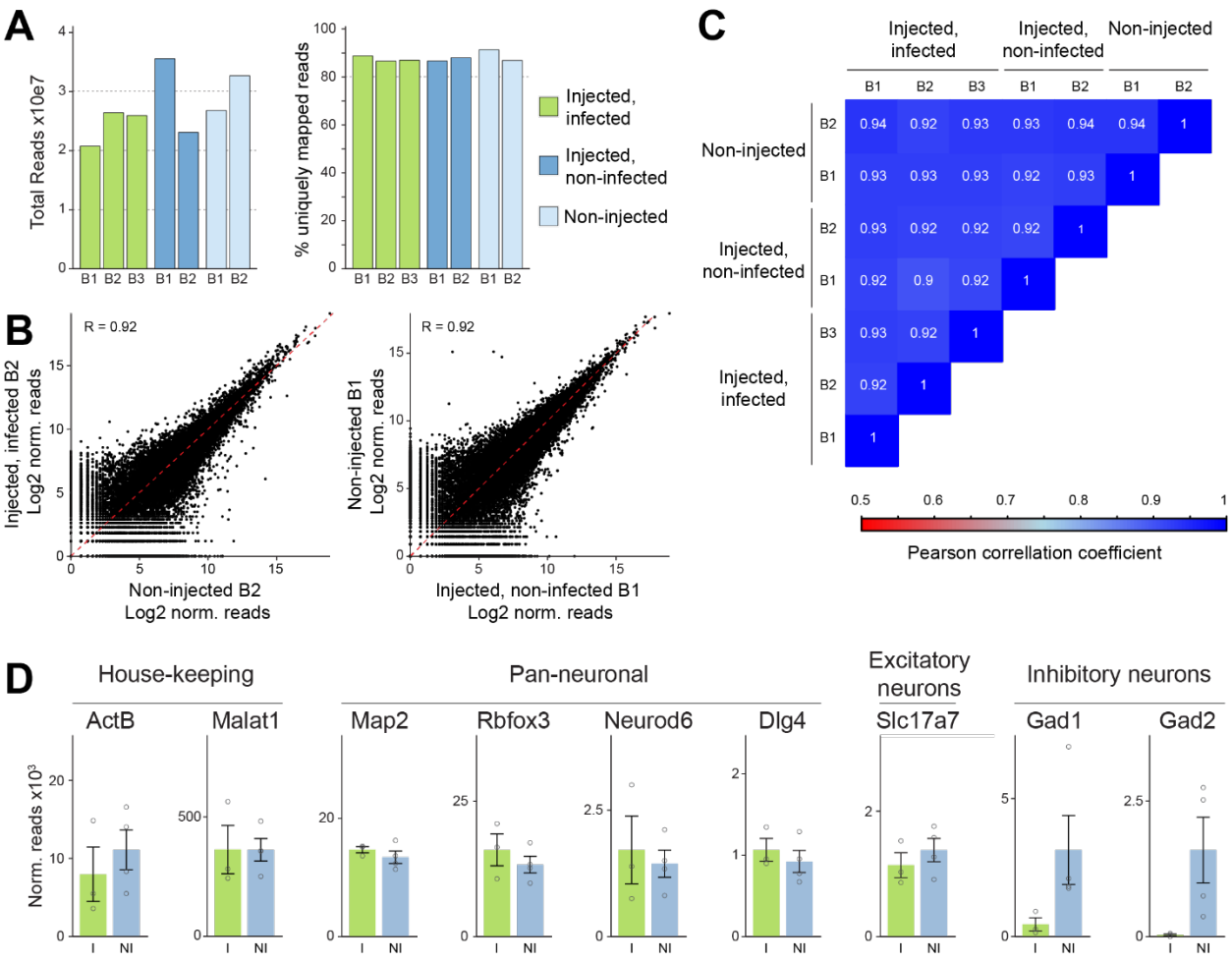

Fig. S2. Meso-seq on hippocampal CA1 pyramidal neurons with 2P-NucTag construct. (**A**) Total reads and percentage of uniquely mapped reads from all libraries (B = biological replicate). (**B,C**) Pearson-correlations of all pair-wise comparisons of the expression levels of all expressed genes in the 7 libraries. (**B**) Two example comparisons. (**C**) Correlation coefficients of all comparisons (B = biological replicate). (**D**) Expression values of example genes in non-infected neurons (NI = pool of "Non-injected" and "Injected, non-infected") and infected neurons (I = "Injected, infected") (error-bars = SEM).

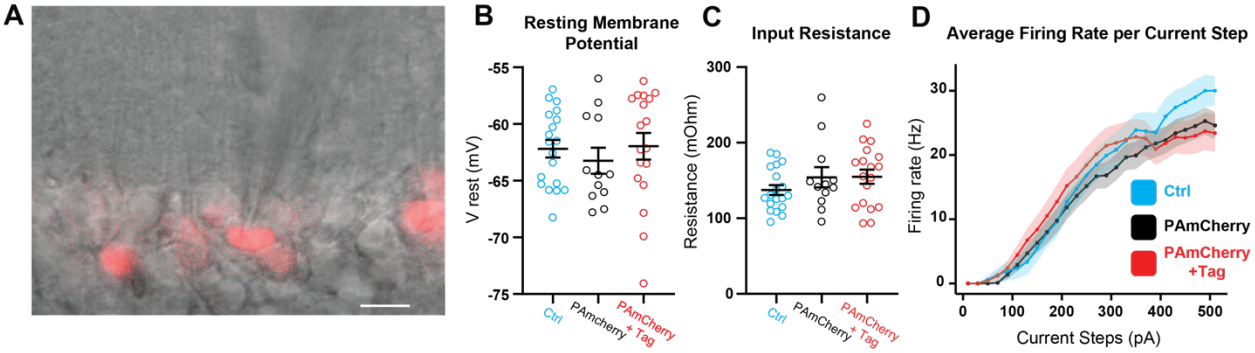

Fig. S3. The 2P-NucTag construct does not alter the intrinsic properties of the CA1 pyramidal neurons. (**A**) Representative image of CA1 pyramidal neurons with (red fluorescence) and without photoactivation during whole cell patch clamp recording. Scale bar: 10 µm. (**B,C**) Quantified data of Resting membrane potential and Input resistance in CA1 excitatory neurons. Blue - control neurons (*Ctrl*), recorded in CA1 of the contralateral non-infected hemisphere. Black- are the neurons recorded from the infected hemisphere but have not been phototagged (*PAmCherry*). Red (*PAmCherry+Tag*) are the phototagged neurons (Ctrl n = 19 cells from 7 mice; PAmCherry n = 12 cells from 8 mice; PAmCherry+Tag n = 18 cells from 7 mice; Statistics Kruskal-Wallis test with post-hoc Dunnett's T3 multiple comparisons test: no statistically significant differences were observed). (**D**) Average firing rate per current step (F-I curve) in each condition. (error bars represent SEM in all data panels).

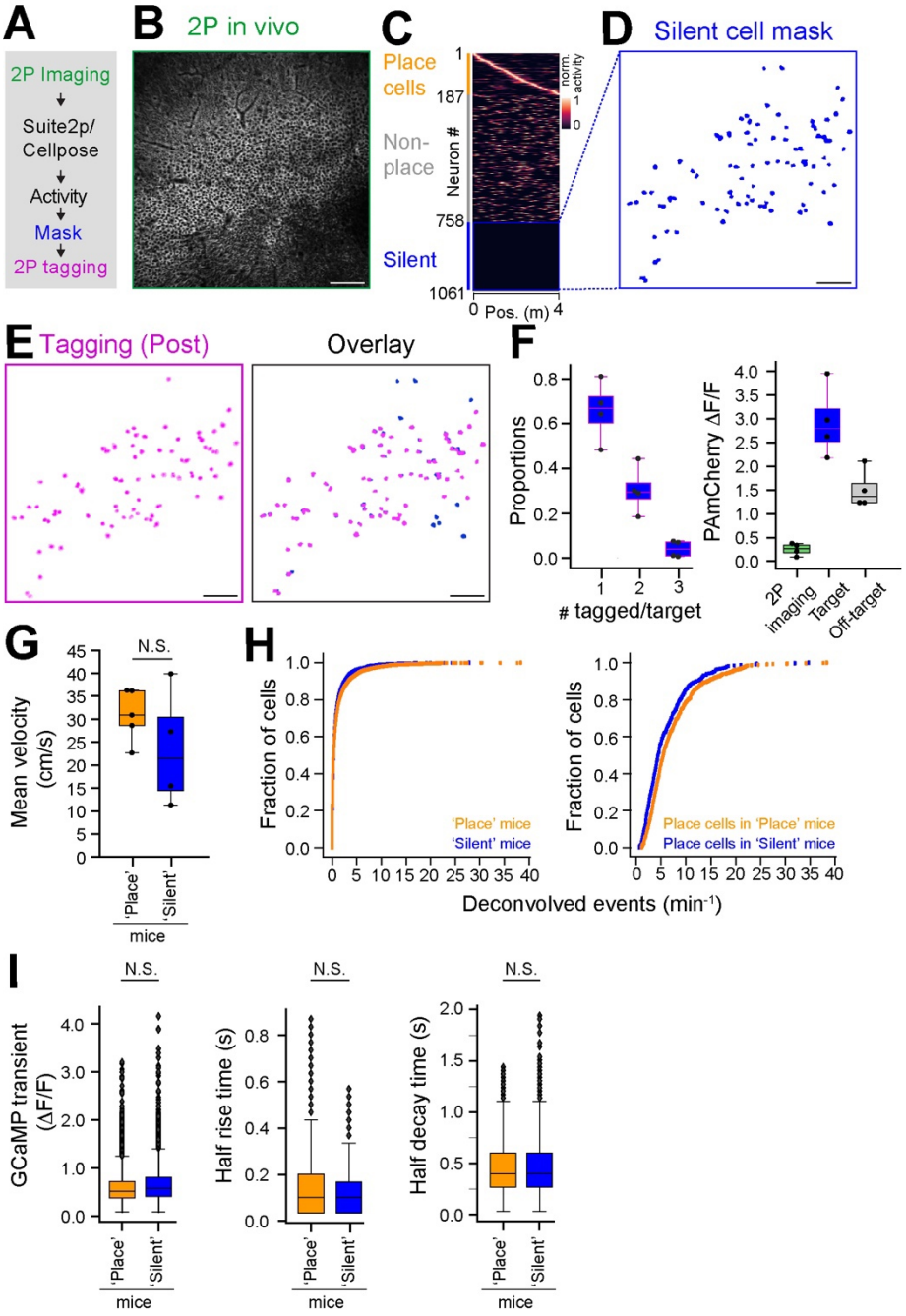

Fig. S4. *In vivo* 2P-NucTag of silent cells. (**A**) Pipeline for two-photon (*2P*) phototagging of functionally identified 'silent' neurons in CA1 during VR spatial navigation (as in *Fig. 1B*). (**B**) Example 2P imaging field of view (FOV) of GCaMP in the CA1 pyramidal layer from a mouse with "silent" cells targeted. Scale bar: 100 µm. (**C**) CA1PNs detected with Suite2p/Cellpose in the FOV shown in *B*. (**D**) Spatial mask (blue) of identified 'silent' cells from the FOV in *B*. Scale bar: 100 µm. (**E**) Left: PAmCherry fluorescence (magenta) of tagged nuclei after 2P phototagging. Right: overlay of spatial masks of identified CA1PNs and tagged nuclei for the FOVs in *D* and *E*. Note that we only tagged a subset of silent cells present in the FOV. Scale bar: 100 µm. (**F**) Left: Proportion of single, double, and triple-tagged nuclei following phototagging of a single silent cell. Right: relative PAmCherry fluorescence change for non-tagged cells in the FOV after 2P imaging (green), after 2P phototagging of targeted silent cell nuclei (blue) and off-target nuclei (gray, n = 4 mice). (**G**) Average velocity of the mice during virtual reality navigation task (Mann-Whitney U test, p-value = 0.286). (**H**) Left: deconvolved events per minute from all cells across all mice from 2P GCaMP-

Ca^{2+}imaging (averaged across mice, n = 5 'Place' mice, n = 4 'Silent' mice, Mann-Whitney U Test, p-value = 0.28). Right: deconvolved events per minute from place cells across all mice (averaged across mice, place n = 5, silent n = 4, Mann-Whitney U Test, p-value = 0.90). (**I**) Left: GCaMP transient amplitude of all cells between groups (averaged across mice, n = 5 'Place' mice, n = 4 'Silent' mice, Mann-Whitney U Test, p-value = 0.14), Middle: GCaMP half rise time of all cells between groups (averaged across mice, n = 5 'Place' mice, n = 4 'Silent' mice, Mann-Whitney U Test, p-value = 0.14), Right: GCaMP half decay time of all cells between groups (averaged across mice, n = 5 'Place' mice, n = 4 'Silent' mice, Mann-Whitney U Test, p-value = 0.81). Boxplots show the 25^{th}, 50^{th} (median), and 75^{th} quartile ranges, with the whiskers extending to 1.5 interquartile ranges below or above the 25^{th} or 75^{th} quartiles, respectively. Outliers are defined as values extending beyond the whisker ranges.

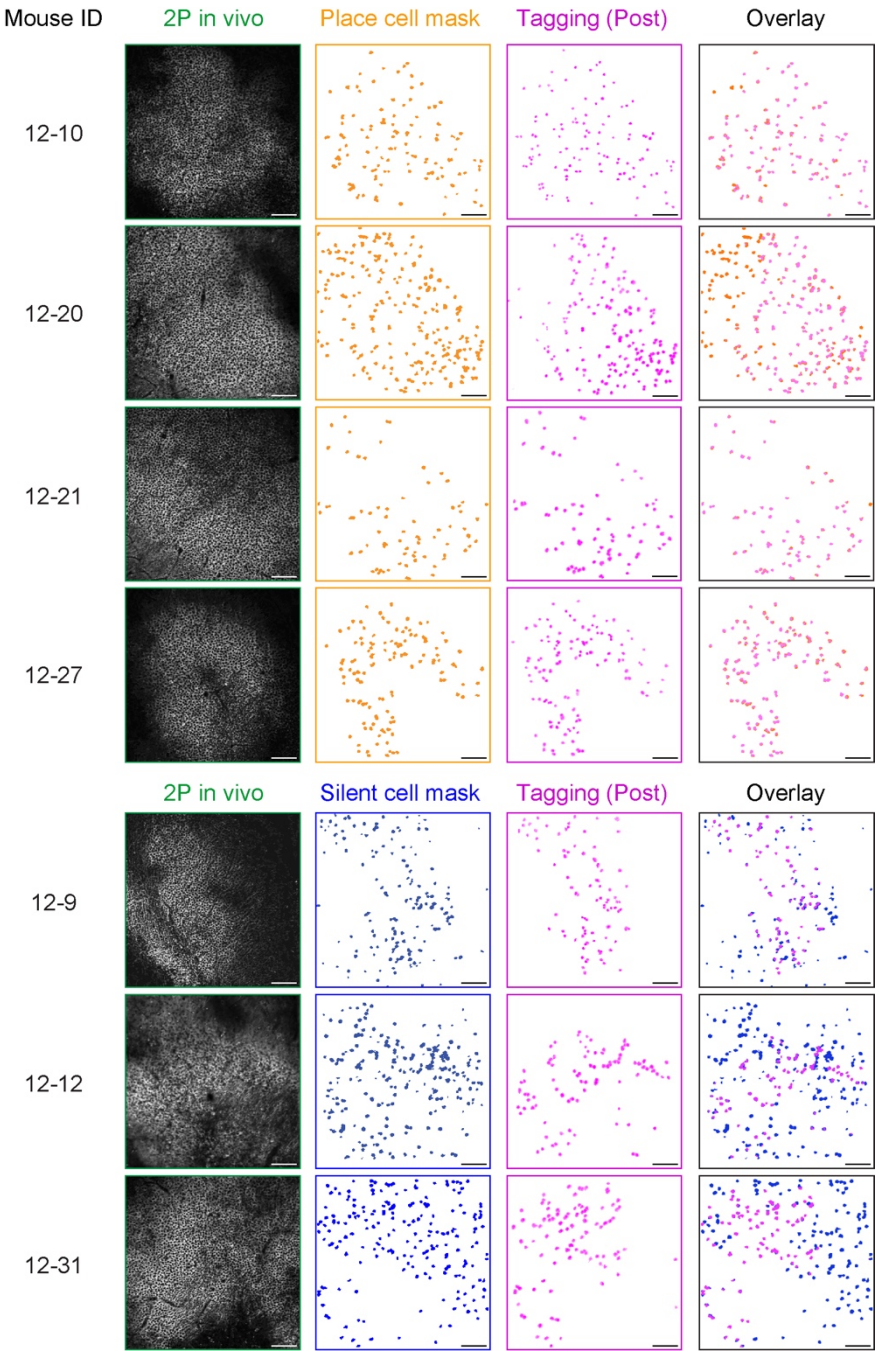

Fig. S5. Additional data on phototagging. From left to right by column: animal ID (animals in *Figure 2* and *figure S2* are not shown here), *in vivo* two-photon (*2P*) imaging fields of view (FOVs), functionally defined masks (orange for place cells and blue for silent cells), PAmCherry fluorescence (magenta) of tagged nuclei after 2P phototagging and overlay of spatial masks from identified CA1PNs and tagged nuclei for the respective FOV in the same row. Note that we only tagged a subset of silent cells present in the FOV in 'Silent' mice, in order to approximate the number of phototagged place cells in 'Place' mice. Scale bar: 100 μm.

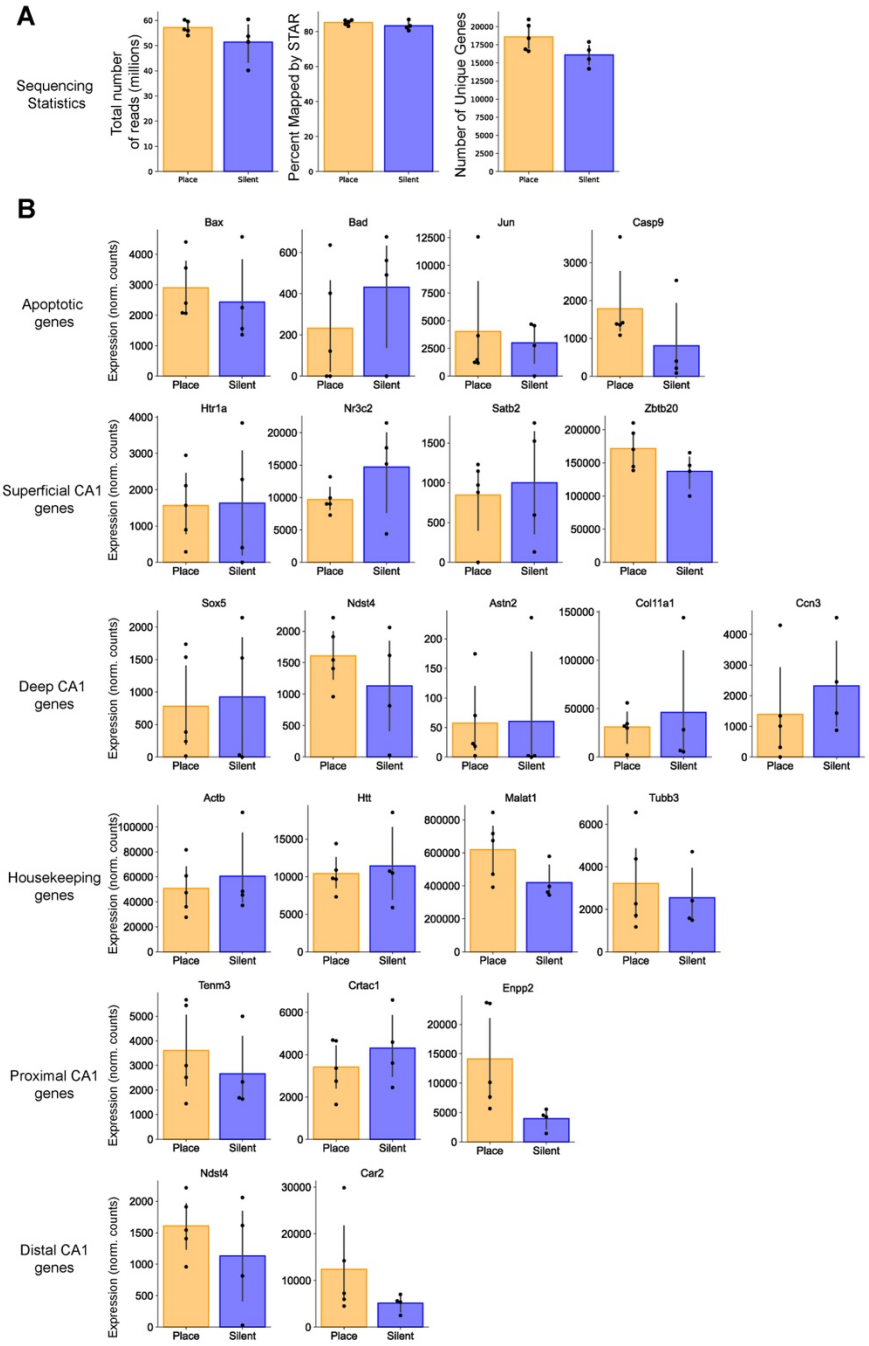

Fig. S6. Additional data on transcriptomics analysis of place and silent cells. (**A**) Sequencing statistics for all 'place' and 'silent' cell samples. Total number of reads - 40 to 60 million reads, 54.64 ± 2.09, n = 9. Percent mapped by STAR - 80.58 to 86.92%. 84.33 ± 0.73. Number of unique genes - 14176 to 20990, 17486 ± 714. (**B**) Normalized counts for groups of genes plotted for 'place' versus 'silent'. Here we show that gene expression of apoptotic genes, superficial CA1 genes, deep CA1 genes, housekeeping genes, proximal CA1 genes, and distal CA1 genes are not different between the two groups (FDR adjusted p-value. *<0.05, **<0.001**,** ***<0.001, PyDeSeq2. All comparisons in this figure are not significant).

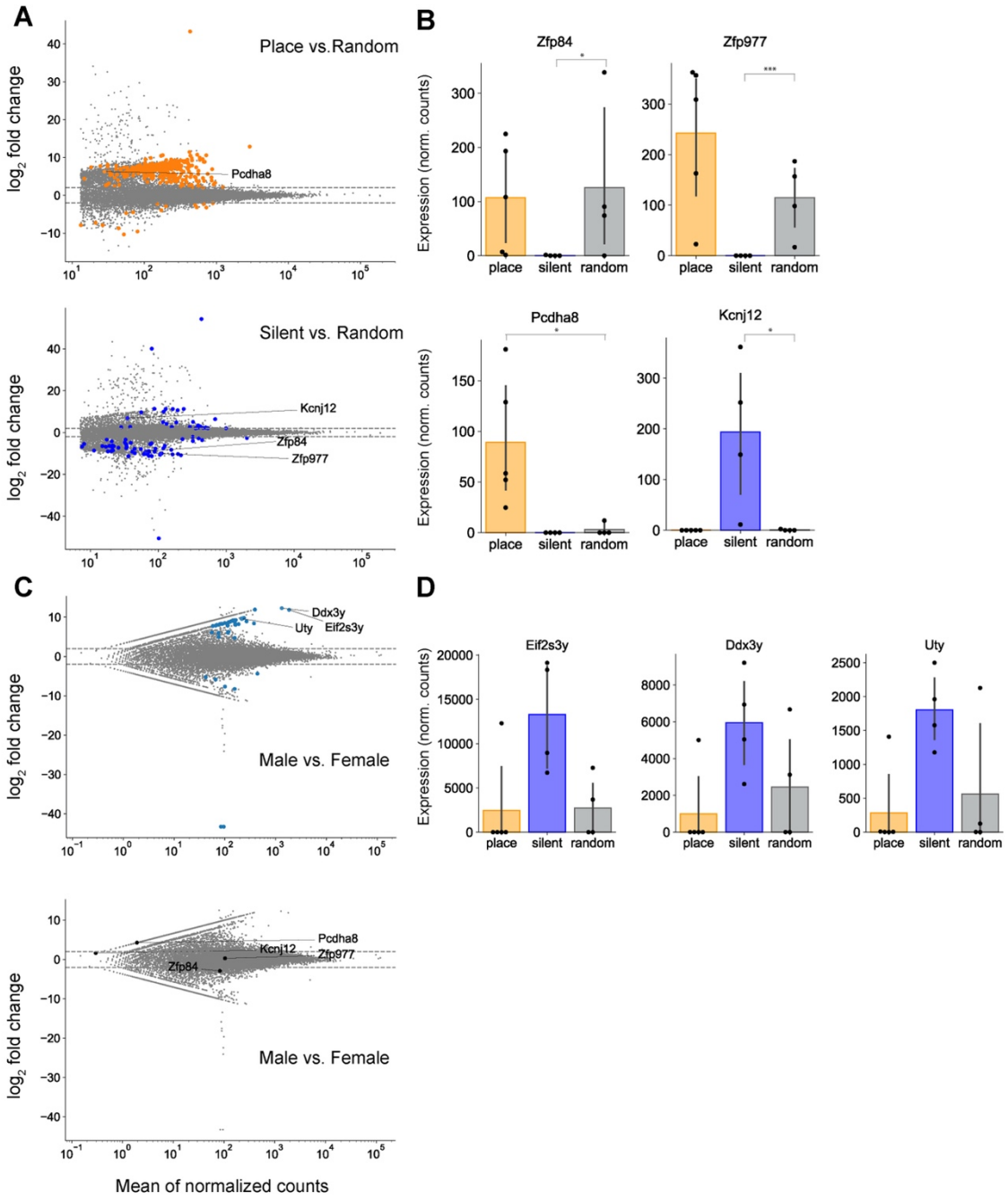

Fig. S7. Additional data on transcriptomics analysis of place and silent cells. (**A**) Top: MA plot of 'place' versus 'random'. Differentially expressed genes (DEGs) are labeled in orange. Bottom: MA plot of 'silent' versus 'random'. DEGs are labeled in blue. Both: DEGs that are common for 'place' versus 'random' were highlighted and labeled. (**B**) Normalized counts for 4 example genes that are significantly differentially expressed across comparisons (FDR adjusted p-value. *<0.05, **<0.001, ***<0.001, PyDeSeq2. Showing here a comparison of 'place' versus 'random' or 'silent' versus 'random'. 'Place' versus 'silent' comparisons were shown in Fig. 3). (**C**) MA plot of male versus female for the 'random' dataset. Top: Y-linked genes that are differentially expressed between sex are highlighted and labeled. Bottom: same four genes in panel A and B are highlighted and labeled. They are not differentially expressed between sex (FDR adjusted p-value. *<0.05, **<0.001, ***<0.001, PyDeSeq2. Otherwise, comparisons are not significant). (**D**) Normalized counts for 3 example DEGs between male and female.

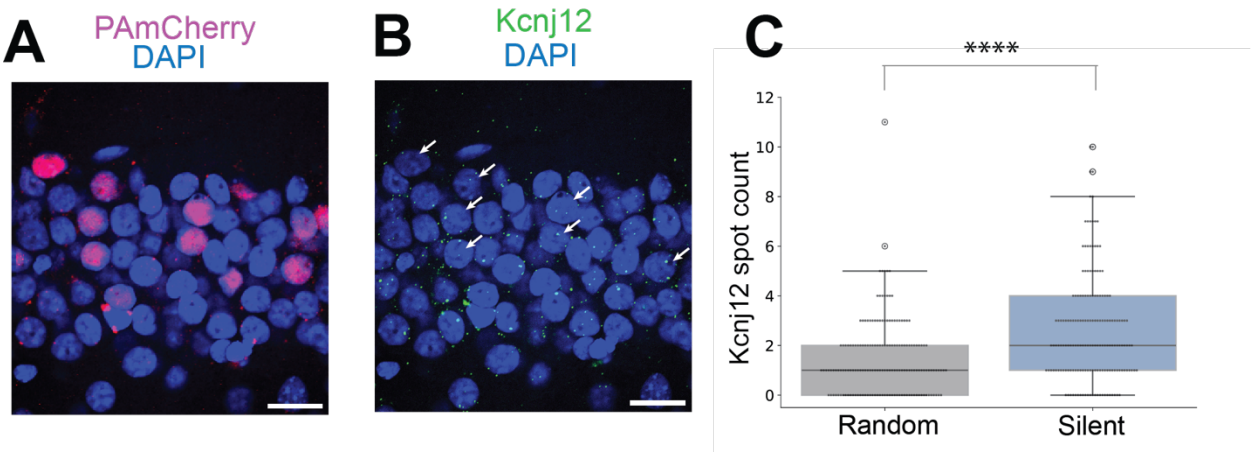

Fig. S8. Spatial distribution of *Kcnj12* transcripts in silent and random cells. (**A**) Confocal horizontal image showing tagged silent cells expressing PAmCherry (magenta with nuclei counterstained by DAPI (blue). Scale bar: 20 µm. (**B**) Confocal image of the same tagged tissues as in (A), hybridized with a probe for *Kcnj12* transcripts (green) using RNAscope Multiplex Assay v2. Nuclei are counterstained with DAPI (blue). Scale bar: 20 µm. For images shown in A&B, red mCherry image was obtained at 20x zoom pre-RNAScope; blue DAPI channel and green Kcnj12 channels were obtained at 60x zoom post RNAScope. (**C**) Box plot comparing the expression levels of *Kcnj12* in tagged cells (n=170, n=2 mice, median=2, IQR=3) and randomly selected non-tagged cells (n=170, n=2 mice, median=1, IQR=2). Mann-Whitney U Test, p-value = 9.56e-08). Boxplots show the 25th, 50th (median), and 75th quartile ranges, with the whiskers extending to 1.5 interquartile ranges below or above the 25th or 75th quartiles, respectively. Outliers are defined as values extending beyond the whisker ranges.

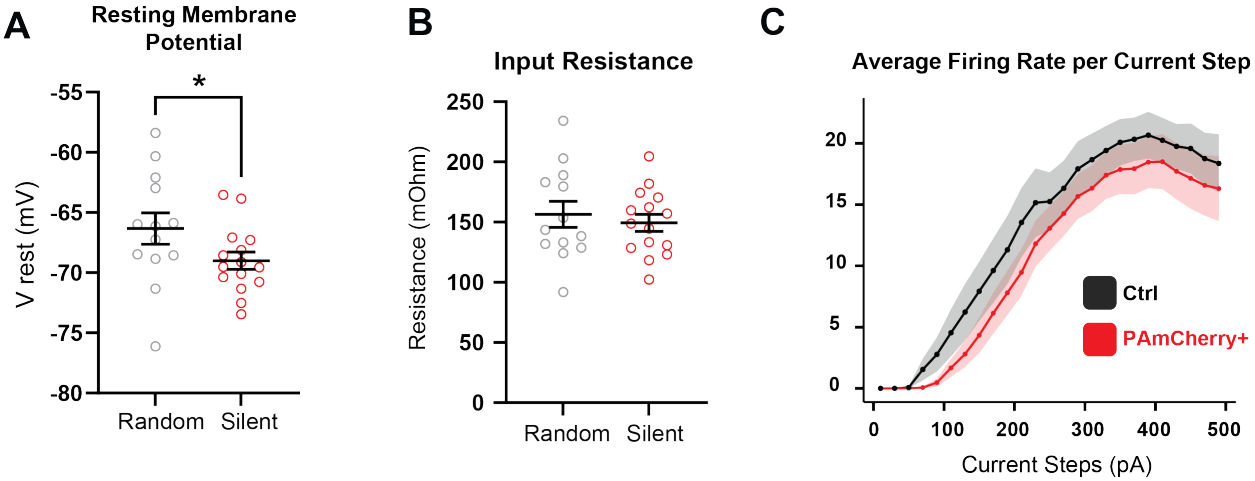

Fig. S9. Silent cells display a decrease in resting membrane potential. (**A,B**) Quantified data of Resting membrane potential and Input resistance in Random (non-silent, *Ctrl*) and silent (PAmCherry+) CA1 PNs (Random n = 19 cells from 4 mice; Silent n = 12 cells from 4 mice). (Mann-Whitney U test. Resting Membrane Potential: p-value=0.0401; Input Resistance: p-value=0.6832; * $p<0.05$). (**C**) Average firing rate per current step in each condition. (Error bars represent SEM in all data panels).

Table 1

Mouse	Sex	Group	Place #	Silent #	Active Non-Place #	Mask #	Tag #
12-10	Male	Place	119	208	1754	119	112
12-20	Female	Place	250	470	1232	218	168
12-21	Female	Place	89	1040	701	74	73
12-27	Female	Place	132	547	548	119	119
12-37	Female	Place	126	290	1440	119	115
12-9	Male	Silent	9	152	655	152	87
12-12	Male	Silent	28	230	287	230	89
12-31	Male	Silent	129	305	1238	200	97
12-34	Male	Silent	187	303	571	113	79

Supplementary Movie 1. Phototagging
Real-time movie of *in vivo* two-photon imaging and phototagging of neurons. Imaging was performed at 1040 nm to visualize change in PAmCherry fluorescence. During phototagging, the 810-nm laser was scanned over target nuclei and the PMT was blanked. A frame average of 64 frames was applied for resolution and clarity. Final video was edited to include scale bar and laser switches. Video time is not representative of actual recording time.

Supplementary Movie 2. Registered *in vivo* and *ex vivo* image stacks
Three-dimensional rendering of registered z-stacks of *in vivo* and *ex vivo* tissue volume of the CA1 pyramidal layer with phototagged nuclei. Corresponding to *Fig. 1F* and Fig *S1C,D*. Magenta: PAmCherry *in vivo* (captured on Bruker 2P microscope, wavelength = 1070 nm). Orange: PAmCherry *ex vivo* (captured on A1 HD25, Nikon Instruments Inc., wavelength = 568 nm)

Supplementary Table 1. List of DEGs between place and silent cells
List of DEGs between place and silent cells showing log 2-fold change, mean of normalized count (baseMean), and FDR adjusted p-value. All genes in the table have an FDR-adjusted p-value less than 0.05 when comparing silent versus place cells. log2FoldChange is computed for silent versus place (positive: enriched in silent cells, negative: enriched in place cells).
File: Supplementary_Table.