

Spatial perturb-seq: single-cell functional genomics within intact tissue architecture

Received: 3 February 2025

Accepted: 3 February 2026

Published online: 21 February 2026

 Check for updates

Kimberle Shen¹, Wan Yi Seow¹, Choong Tat Keng¹, Michelle Gek Liang Lim¹, Daryl Shern Lim¹, Ke Guo¹, Amine Meliani¹, Muhammad Irfan Bin Hajis¹, Bolun Wang¹, Shyam Prabhakar^{1,2,3}, Kok Hao Chen¹ & Wei Leong Chew^{1,4}✉

We develop Spatial Perturb-Seq, an *in vivo* CRISPR technology that interrogates multiple genes within single cells of intact tissues, compatible with both sequencing-based and probe-based spatial technologies. We apply Spatial Perturb-Seq to knock out risk genes for neurodegenerative diseases in the mouse brain, uncovering cell autonomous and cell-cell microenvironmental effects within the spatially intact tissue. Spatial Perturb-Seq functionally screens multiple genes *in situ* and *in vivo*, bypasses cell processing steps that skew cell type representation, identifies intracellular and intercellular effects of knockouts, and identifies candidate genes underlying dysregulated neuronal intercellular communication pathways.

Perturb-seq, which combines pooled CRISPR screens with single-cell RNA-sequencing (scRNA-seq), has dramatically increased the scale of functional genomics studies^{1,2}. Several studies have shown the feasibility of Perturb-seq *in vivo*^{3–5}, in systems where cells maintain their native gene expression patterns. However, spatial organization and cell-cell interactions are not preserved. Perturb-map⁶, relying on protein barcodes detected by imaging methods, has been used to identify genetic factors influencing the spatial organization of solid tumours. Similarly, optical pooled screens allow perturbations to be linked to image-based phenotypes but require extensive technical expertise and specialized microscopic setups^{7–10}. These prior approaches uncouple the CRISPR perturbation detection platforms from transcriptomic sequencing used for cell phenotype identification. In contrast, a spatial functional genomics technology that uses scalable nucleotide barcodes would allow the endogenous cell phenotype, exogenous perturbations, and spatial localizations to be concurrently identified with a single sequencing technology. Here, we present Spatial Perturb-Seq, an *in vivo* Perturb-seq technology that interrogates genetic functions directly in single cells in their native tissue. Spatial Perturb-Seq simultaneously measures whole transcriptomes at single-cell resolution (and therefore cell type), CRISPR barcodes (linked to gene perturbation), spatial coordinates, and cell-cell interactions

(microenvironment) - all through a single Stereo-seq¹¹ and/or Xenium¹² run (Fig. 1a schematic).

Results

We simultaneously interrogated 18 genes within the same animals, using a pooled barcoded CRISPR-gRNA AAV library (Supplementary Data 1) delivered intracranially into the hippocampus of Cas9-expressing mice¹³ at low multiplicity of infection (MOI). Intentional sparse editing ensures perturbed cells are surrounded by non-edited cells, allowing isolation of cell autonomous and non-cell autonomous (i.e. microenvironmental) effects. Most of the immediate neighbours of perturbed cells are barcode-negative, as designed (Supplementary Fig. 1). We used AAV over the more commonly used lentivirus, to avoid possible genotoxic integration¹⁴ and template switching events that uncouple barcodes from their respective gRNAs¹⁵.

To reduce perturbation failure due to non-functional guides, and because gene editing efficiency is increased using multiple guides compared to single guides, we used constructs containing three gRNAs to target each gene. We confirmed the presence of the full-length 3x U6-gRNA constructs before and after AAV transduction (73.7% full-length after one round of PCR) and also observed recombination between U6-gRNA cassettes in a minority of constructs by

¹Genome Institute of Singapore (GIS), Agency for Science, Technology and Research (A*STAR), 60 Biopolis Street, Genome, Singapore, Republic of Singapore. ²Population and Global Health, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Republic of Singapore. ³Cancer Science Institute of Singapore, National University of Singapore, Singapore, Republic of Singapore. ⁴Synthetic Biology Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Republic of Singapore. ✉e-mail: chewwl@a-star.edu.sg

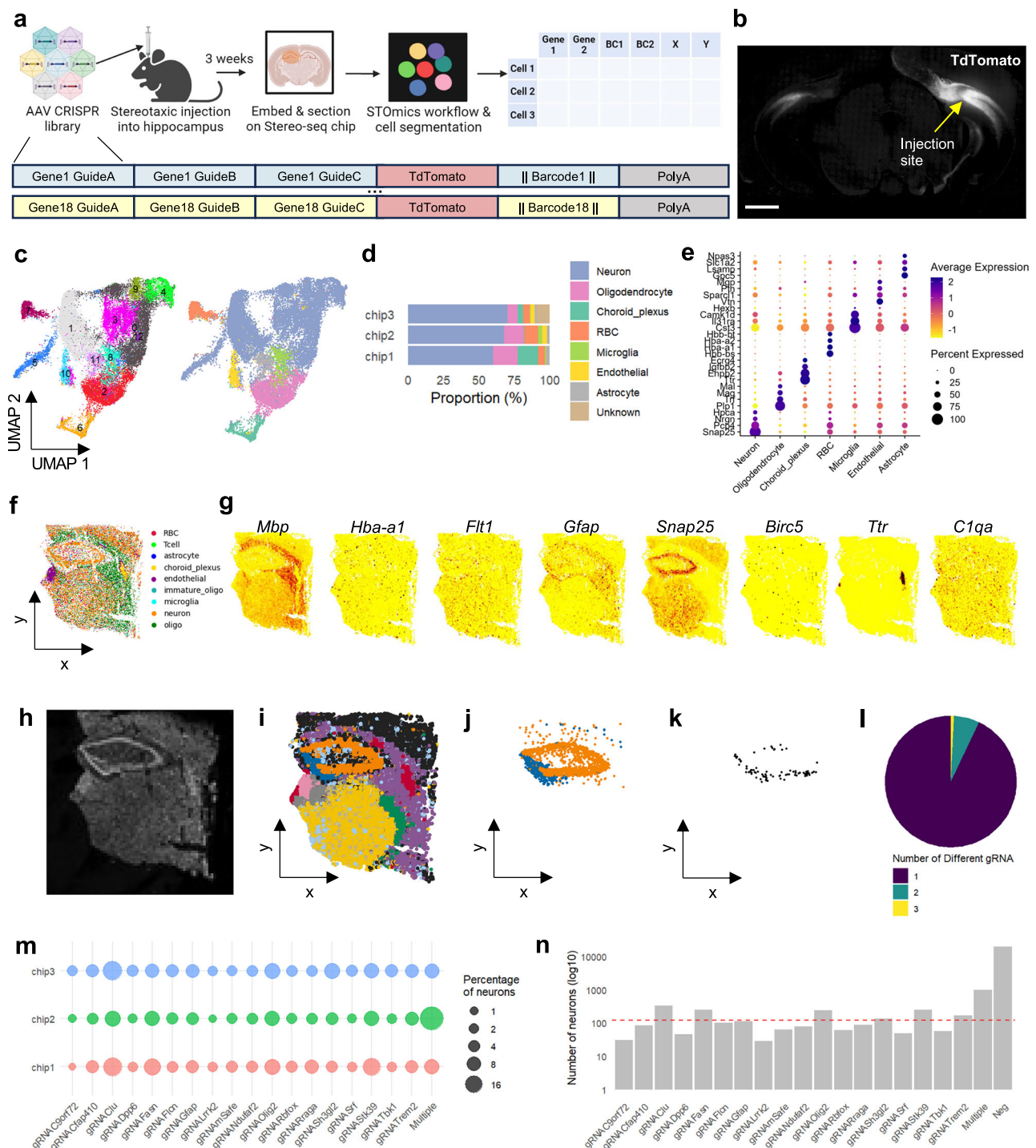


Fig. 1 | In vivo spatial perturb-seq of the adult mouse brain by intracranial injection of a barcoded AAV library. **a** Schematic of the in vivo spatial Perturb-seq pipeline, and depiction of components in our AAV-PHP.eB library, showing 3 gRNAs, TdTomato and 1 barcode in each member. Created in BioRender. Shen, K. (2026) <https://BioRender.com/1ox9bo3>. **b** Confocal imaging of a coronal section of a mouse brain following stereotaxic injection into the hippocampus, showing that transduction spread is limited to the hippocampus region. Scale bar: 1000 μ m. $n = 5$ mice. **c** UMAP projection of cells from chip 1 following Stereo-seq, showing recovery of neurons, oligodendrocytes, choroid plexus cells, RBCs, microglia, endothelial cells, and astrocytes. Left: coloured by cluster number. Right: coloured by cell type. **d** Similar proportions of cell types were recovered for each chip, with the majority cell type being neurons as expected. **e** Dot Plot showing expression of cell-type enriched genes for each cell type. **f** Spatial scatter of annotated mouse

brain cells. **g** Spatial gene plot of canonical cell-type markers major cell populations (Microglia: C1qa; Oligodendrocytes: Mbp; Endothelial cells: Flt1; Neurons: Snap25; Astrocytes: Gfap, Choroid Plexus: Ttr; RBC: Hba-a1; Immature oligodendrocytes: Birc5). **h** ssDNA image for cellular localization on the Stereo-seq chip. **i** Spatial scatter following BANKSY clustering with lambda 0.2, showing clusters recovered. **j** Selection of CA and DG neurons in the hippocampus by focusing on BANKSY clusters 3 and 10. **k** Subset of barcode-positive cells from (j). **l** Piechart representation of the number of unique gRNA barcodes among all gRNA barcode positive cells, showing a small proportion of barcode-positive cells with multiple barcodes detected. **m** Dot Plot showing percentage of barcode-positive neurons among hippocampal neurons in (j) for each chip. **n** Total number of cells for each gRNA across all 3 chips. Min: 29, max: 330, mean: 122. f–l Data from chip 1.

PCR gel densitometry analysis (Supplementary Fig. 2a–c). A second round of PCR in preparation for long-read sequencing further shifted the full-length 3×U6-sgRNA cassettes to the shorter 1x and 2x U6-sgRNA, an expected artifact of in vitro recombination from multiple PCR cycles that skewed amplicon profiles toward shorter fragments (Supplementary Fig. 2d, e). Importantly, the consistent proportion of full-length constructs before and after AAV transduction suggests minimal in vivo recombination (Supplementary Fig. 2d). Next, in silico off-target prediction for all 54 guides identified only 6 sites in the mouse genome with <3 mismatches. All sites with <4 mismatches are shown in Supplementary Data 2. We first determined that the designed gRNAs are effective in cultured cells (Supplementary Fig. 3a). Next, to confirm CRISPR-Cas9 activity in Cas9 mice, we transfected gRNAs into BMDMs derived from Cas9 mice and achieved high editing rates (Supplementary Fig. 3b). We also verified in vivo editing in the tissue of injected mice, observing an average of 0.15% indels for each locus genotyped in the single cell lysate (Supplementary Fig. 3c, d), close to the observed mean of 0.13% of sequenced single cells containing each barcode. Mutational profiling showed distinct indels between 1 and 10 bp (Supplementary Fig. 3e), consistent with precise on-target CRISPR/Cas9-mediated editing at the target sites. This suggests that cells delivered with barcoded CRISPR gRNAs are largely edited at the genomic loci sequenced, consistent with editing efficiencies of single gRNAs in Cas9-mice. The consistency between the intentionally low fraction of cells harbouring individual barcodes and the fraction of cells exhibiting targeted genetic perturbations provides an initial indication that barcode detection by sequencing is a reasonable proxy for inferring linked genetic perturbations.

We first evaluated the methodology with conventional Perturb-seq via single-cell RNA-seq (scRNA-seq) (Supplementary Fig. 4). Tropism of the AAV-PHP.eB serotype used here was indeed highest in neurons¹⁶, with high tropism also observed for astrocytes and oligodendrocytes, but not for microglia and endothelial cells (Supplementary Fig. 4a–e). Focussing first on oligodendrocytes, we recovered known oligodendrocyte subtypes: MOL1, MOL2, MOL5/6, and DA1fn¹⁷ (Supplementary Fig. 5a, b). The proportion of barcode-positive MOL1 oligodendrocytes was significantly higher than barcode-positive MOL5/6 oligodendrocytes (Supplementary Fig. 5c, d), suggesting that injection injury and AAV transduction could steer oligodendrocytes to a less mature, MOL1 state, consistent with reports showing MOL1 enrichment at injury sites¹⁸. It is also possible that AAV-PHP.eB has stronger tropism for MOL1 oligodendrocytes. Oligodendrocytes are responsible for making and maintaining the myelin sheath, a process that involves lipid metabolism, which we examined through the expression of two sets of genes in oligodendrocytes: myelin genes (*Mobp*, *Mog*, *Opalin*, *Plp1*, *Mbp*, *Cnp*, *Mag*, *Mal*) and lipid metabolism genes (*Srebfl1*, *Hmgcr*, *Scd1*, *Scd2*, *Acaca*, *Dgat1*, *Cpt1a*, *Elovl6*). Olig2-KO does not significantly reduce the expression of canonical mature myelin genes (Supplementary Fig. 5e), consistent with its role in oligodendrocyte specification but not maintenance¹⁹. *Rraga* and *Flcn* perturbations cause slight downregulation of myelination genes in mice (Supplementary Fig. 5e), suggesting that the *Rraga*-*Flcn*-Tfeb pathway could be a promising target for promoting remyelination. This is consistent with *Flcn* and *Rraga* being required for myelination in zebrafish²⁰, even though their function in mammalian myelination has not been recognized. In addition, *Fasn*, a key regulator of fatty acid metabolism, is required for the expression of lipid metabolism genes in oligodendrocytes (Supplementary Fig. 5f).

With scRNA-seq, 11% of oligodendrocytes were positive for barcodes, with an average of 26 oligodendrocytes per perturbation (Supplementary Fig. 6a, b). To assess the penetrance of transcriptomic changes induced by each perturbation, we performed Mixscale analysis²¹ (Supplementary Fig. 6c, d). Mixscale enables analysis of knockout penetrance, helping to alleviate challenges in Perturb-seq experiments where it is unfeasible to genotype all target genes.

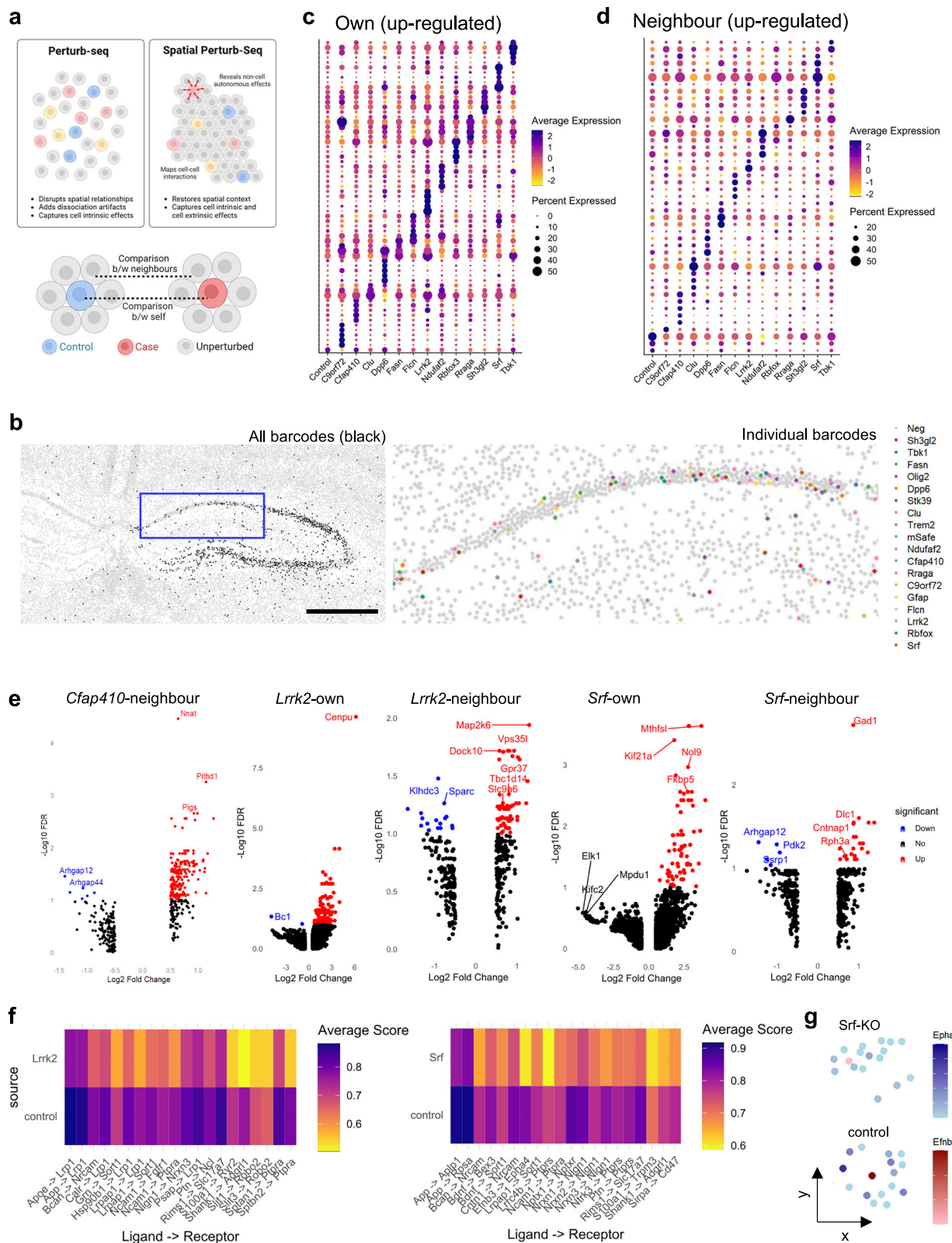
Mixscale results showed that most gene perturbations were associated with transcriptomic changes in oligodendrocytes, except for *Srf*, *Lrrk2*, *Ndufaf2*, and *Cfap410* in oligodendrocytes, perhaps due to low baseline expression of three of these genes (*Srf*, *Lrrk2*, and *Cfap410*) in oligodendrocytes (Supplementary Fig. 6e).

Neurons were underrepresented in our scRNA-seq data, as they are known to be fragile and depleted during the cell dissociation process used for scRNA-seq²². Hence, we performed additional single-nuclei RNA-seq (snRNA-seq), where we recovered majority neurons (Supplementary Fig. 4f–h). However, for the snRNA-seq run, without a separate dial-out amplification of the barcodes during library preparation, Perturb-seq with snRNA-seq showed insufficient barcode detection compared to scRNA-seq (Supplementary Fig. 4i, j), likely due to lower mature mRNA content in the nuclei. These data point to the considerations in maintaining accurate cell type representation arising from the cell dissociation and processing methodologies employed, and the necessity of additional amplification steps for snRNA-seq.

Having assessed the methodology with conventional Perturb-seq, we next introduced spatial transcriptomic readout through Spatial Perturb-Seq. Applying in vivo Spatial Perturb-Seq with Stereo-seq (Fig. 1a, b), we obtained spatial sequencing data for 229,775 cells over 3 experiments and 4 mice (Supplementary Fig. 7a, b and Supplementary Data 3). Libraries were prepared without barcode-specific PCR-enrichment. The dataset is then processed with DeepCell²³ for cell segmentation (Supplementary Fig. 7c), where an average of 462 and 773 expressed genes were detected per cell and per neuron, respectively (Supplementary Data 3). We recovered neurons, red blood cells, oligodendrocytes, microglia, astrocytes, choroid plexus and endothelial cells (Fig. 1c–g) with spatial patterns consistent with gene expression modules (Supplementary Fig. 7d, e). The consistency in manual annotation and spatial hotspot analyses of gene modules provides confidence in the data quality and cell segmentation. We employed BANKSY²⁴ to select regions of interest, by adjusting the lambda parameter to vary the cells' embedding. At low lambda values, BANKSY functions in single cell-typing mode, whereas at high lambda values, it identifies spatial domains (Supplementary Fig. 8a). We used lambda 0.2 to achieve the goal of spatially informed cell-typing to identify hippocampal neurons (Fig. 1h–k). BANKSY's neighbour-calling framework identified cell-intrinsic (own) and microenvironment (neighbour) DEGs (Supplementary Fig. 8b). All 18 barcodes are represented in the sequencing data (Supplementary Fig. 9a, b), with most barcode-positive neurons containing only 1 unique barcode (Fig. 1l–n). Cells with multiple barcodes (12.5% of barcode-positive cells) were excluded from further analysis. Among all cells profiled by Stereo-seq, 2.1% are barcode-positive (Supplementary Data 3), closely aligning with the 2.4% observed in the scRNA-seq dataset, suggesting comparable barcode detection sensitivity.

Spatial Perturb-Seq enables the interrogation of both cell-autonomous and non-cell autonomous microenvironmental effects of perturbations within native intact tissues (Fig. 2a), operating by the detection of perturbation-specific barcodes (Fig. 2b). Aggregating perturbed cells across the entire hippocampus, the top 5 significantly upregulated genes for each perturbation were distinct among the 18 perturbations (Fig. 2c, d), highlighting the specificity of transcriptomic changes induced by each genetic knockout. To compare the effects of the different perturbations, we quantified the number of DEGs ($p < 0.05$, $lfc > 0.5$) for each perturbation compared to control, and their average effect size (Supplementary Fig. 8b). As expected, perturbations cause a stronger cell autonomous effect, compared to their influence on the microenvironment (defined as 15 closest neighbouring cells). *Cfap410* perturbation caused the highest number of DEGs in its non-perturbed cellular neighbourhood, consistent with its role as a player in cilia function and synaptic plasticity in neurons²⁵.

Knock out of *Lrrk2* ($n = 29$ cells) led to 213 cell-autonomous DEGs observed ($p < 0.05$, $lfc > 0.5$), the highest number amongst all



perturbed genes (Fig. 2e, Supplementary Fig. 8b). *Lrrk2* is a key player in neuronal signalling and Parkinson's disease pathology^{26,27}. *Bcl1*, a lncRNA found in dendrites that regulates translation of specific mRNAs in synapses²⁸, is downregulated upon *Lrrk2* KO. Among the micro-environmental changes induced by *Lrrk2* KO: downregulation of *Sparc*; upregulation of *Vps35l* (*Vps35* is known to functionally interact with *Lrrk2*²⁹), *Dock10* (involved in dendritic spine formation³⁰), and *Gpr37*

(known to be associated with PD³¹). The microenvironmental effects stemming from *Lrrk2* KO can potentially underlie *Lrrk2*-mediated pathology in neuronal signalling and Parkinson's disease.

Another striking finding is that from *Srf* ($n=50$ cells), a transcription factor regulating genes involved in neuronal growth and synaptic plasticity³². Among the genes dysregulated upon *Srf*-KO, most are target genes identified through ChIP-seq data from the ENCODE

Fig. 2 | Analysis of cell-autonomous and non-cell autonomous effects of perturbations and cell-cell communication in hippocampal neurons. **a** Schematic depicting Spatial Perturb-Seq and the two different comparisons that can be made – cell autonomous and non-cell autonomous, i.e. (1) transcriptomes of perturbed cells (own). (2) transcriptomes of the wildtype neighbours of perturbed cells (neighbours). Created in BioRender. Shen, K. (2026) <https://BioRender.com/x0zc3s8>. **b** FOV of a subsection of Stereo-seq data, highlighting all barcode-positive cells (left) and distinct barcodes shown in different colours (right, subset of left panel boxed in blue), with cells negative for barcodes in grey. Data taken from chip 2. **c**, **d** Dot Plot showing top DE genes for each perturbation, compared to all other perturbations, for own transcriptome (**c**), and neighbour cells' transcriptome (**d**).

e Volcano plots showing DE genes ($\text{lfc} > 0.5$, $\text{fdr} < 0.05$) for Cfp410-KO neighbours (left-most), Lrrk-KO and Srf-KO own and neighbour cells. Only $\text{lfc} > 0.5$ genes are shown. mSafe-KO was used as the control group. **f** Heatmap visualizing ligand-receptor expression for top 20 differentially expressed ligand-receptor pairs, comparing communication scores between Lrrk2-KO or Srf-KO neurons and their wildtype neighbours vs communication between control neurons and their wildtype neighbours. **g** Spatial plot of each source (ligand-expressing) cell in red and their 15 wildtype neighbour (receptor-expressing) cell in blue, showing expression of *Epha4* ligand in perturbed cells and *Efnb3* receptor in neighbour cells. Top: Srf-KO. Bottom: control. Scale bar represents 1 mm.

Transcription Factors Target dataset³³, though some did not reach the fdr threshold of < 0.05 . In the Srf-KO microenvironment, *Arhgap12* (involved in cytoskeletal and actin dynamics³⁴) and *Ssrp1* (co-activator of Srf³⁵) were downregulated, *Gad1* (involved in the synthesis of neurotransmitter GABA³⁶) was upregulated.

We delved deeper into cell-cell communication, as it is essential for neuronal function. To ensure accuracy of cell-cell communication analysis, we studied cells by keeping to the same neuronal and spatial niches, as ligand-receptor communication occurs among neighbouring cells in proximity. We used the LIANA receptor-ligand framework³⁷ to interrogate the dataset and prioritize biologically important ligand-receptor pairs (Fig. 2f). We observed a 19% reduction in the SCA score for Lrp1 signalling in the cellular neighbours of Lrrk2-KO neurons, compared to control neurons. Lrp1 mediates important neuronal functions, including α -synuclein uptake³⁸, suggesting a possible mechanistic link between Lrrk2 activity and PD pathogenesis. The data showed a reduction in Nlgn1 signalling in Srf-KO neighbours, highlighting the importance of *Srf* in synaptic function. To visualize these cell communication networks, individual cells showing Efnb3-Epha4 signalling in Srf-KO and control-KO microenvironments is shown on a spatial plot (Fig. 2g). Source cells and their 15 closest neighbours (target cells), are shown according to their XY coordinates, with expression levels of the ligand in red (source cells) and expression levels of the receptor in blue (target cells). These analyses highlight the ability of Spatial Perturb-Seq to interrogate genetic determinants of intercellular cell communication signalling pathways within the native tissue.

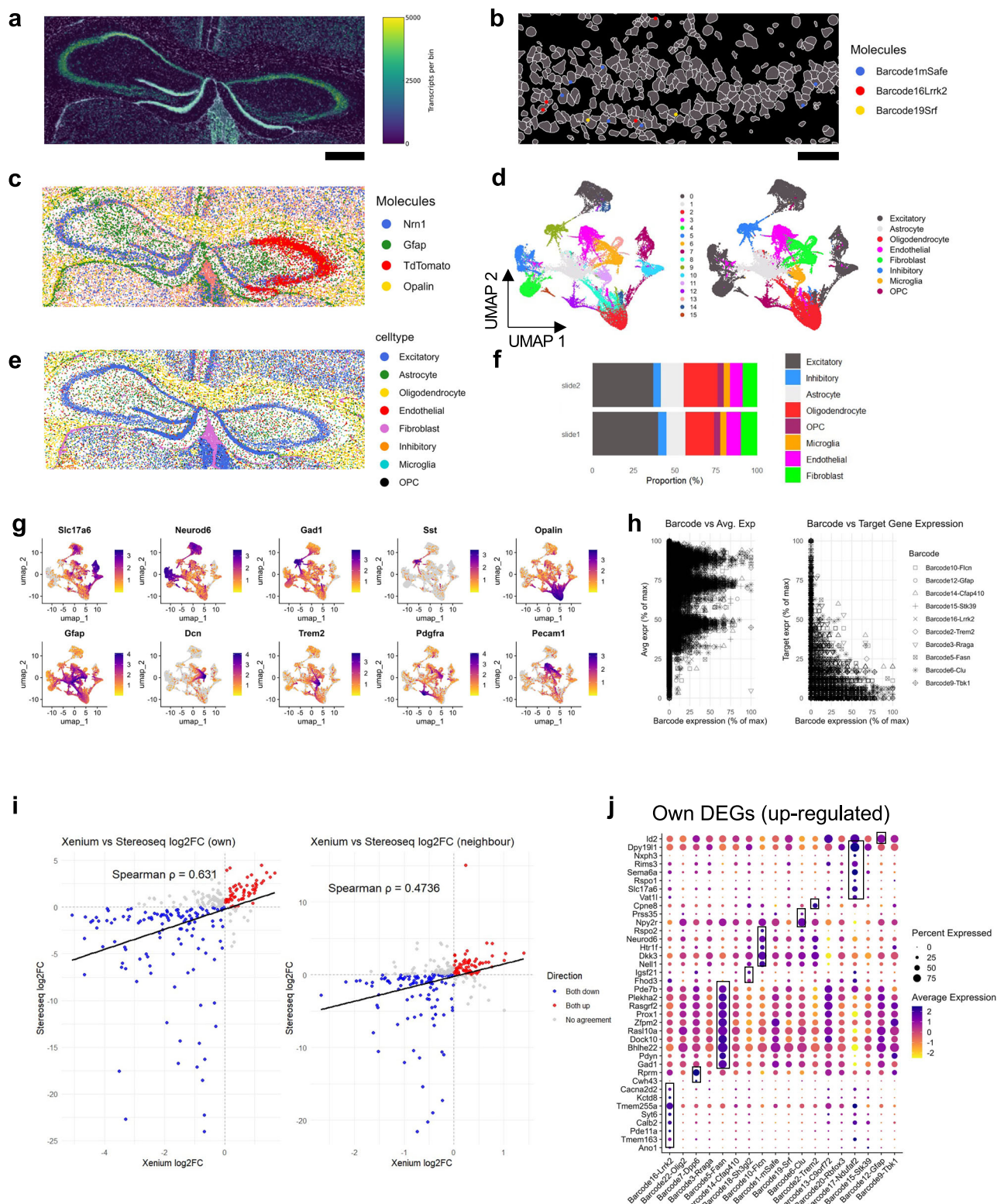
To test the compatibility of Spatial Perturb-Seq in an orthogonal, probe-based platform, we turned to the Xenium¹² platform with a 247-gene probe panel and custom probes against our 450 bp barcodes (Supplementary Data 4, 5). Transcript density maps showed expected cell densities across the tissue (Fig. 3a). Cell segmentation boundaries demarcate cells with or without individual molecules of detected barcodes (Fig. 3b). TdTomato transcript localized to the expected region of the neuronal cell bodies in the injected side of the hippocampus (Fig. 3c). All 18 barcodes were robustly detected using the Xenium platform (Supplementary Fig. 10a, b), demonstrating their dual functionality for sequencing-based and probe-based spatial platforms. Among recovered cell types, neurons were the most abundant (Fig. 3d–f), with cell-type identities annotated via canonical marker expression (Fig. 3g). Excitatory neurons showed the highest transduction with AAV-PHP.eB (Supplementary Fig. 10c–f). Importantly, we again found a strong linkage between detected barcodes and target gene knockout, evident by data showing that target gene expression in neurons decreases with increasing expression of corresponding gRNA-barcode, while the average global transcript levels are not affected (Fig. 3h, Supplementary Fig. 11).

We next investigated the concordance of spatial transcriptomic signatures between Xenium and Stereo-seq (Supplementary Fig. 12a). For both platforms, average $\log_2$ fold-change was calculated for all Xenium-targeted genes (excluding barcodes), comparing barcode-positive vs. barcode-negative cells. Fold-changes from Xenium were

compared to those from Stereo-seq across all DEGs. A Spearman correlation of 0.63 was observed for DEGs in barcode-positive cells themselves (“own DEGs”) and 0.47 for DEGs in neighboring cells (“neighbour DEGs”) (Fig. 3i). This suggests that perturbation effects observed in barcode-positive cells are moderately preserved across both platforms. We then looked at DEGs for each barcode compared to the rest of the barcodes. Among upregulated genes, only 9 out of 18 perturbations led to DEGs with the 247-gene Xenium probe panel used (Fig. 3j), consistent with the reduced gene representation compared to the transcriptome-wide, unbiased representation from Stereo-seq. In the context of Spatial Perturb-Seq, whole-transcriptome profiling (i.e. Stereo-seq) might be more suited for unbiased discovery. In cells with Lrrk2-KO, transcripts for *Cacna2d2*, *Kctd8*, *Tmem255a*, *Syt6*, *Calb2*, *Pde11a*, *Tmem163*, and *Ano1* were upregulated, consistent with the known function of Lrrk2-KO in disrupting calcium signalling in neurons³⁹ and leading to dysregulation of genes involved in calcium homeostasis. Consequently, genes such as *Cacna2d2* (a subunit of voltage-dependent calcium channels), *Calb2* (calbindin 2, a calcium-binding protein), and *Ano1* (a calcium-activated chloride channel) may exhibit altered expression in Lrrk2-KO cells.

To further assess the concordance of transcriptomic signatures between Xenium and Stereo-seq, we next performed Spearman correlation analysis for all nine perturbations where DEGs were detected in Xenium, comparing the expression profiles on a per-perturbation basis rather than aggregating across barcode-positive and -negative cells (Supplementary Fig. 12b). We again observed similar directionality within the DEGs, including in *Srf*- and *Lrrk2*-perturbed cells, with Spearman coefficients ranging from a modest 0.1 to 0.5, which aligns with expected biological and technical variability from the two different platforms, especially when working with fewer cell and transcript numbers (Supplementary Fig. 12c, d). The similar directionality among these DEGs, despite different fold change magnitude, suggests different dynamic ranges between the probe-based and sequencing-based platforms. As further validation, we analyzed the Allen Brain Cell Atlas as another orthogonal dataset, specifically focusing on validating the robustness and reproducibility of Lrrk2-mediated gene expression changes in both “own” and “neighbour” cells (Supplementary Fig. 12e). *Bcl1* expression decreases with lower *Lrrk2* expression (Spearman = 0.853), consistent with our findings. Among “neighbour” DEGs, *Map2k6* and *Dock10* shows a modest inverse correlation (Spearman = -0.217 and -0.110 respectively), and *Sparc* expression is positively correlated (Spearman = 0.562). Hence, the Lrrk2-mediated DEG in *Bcl1* (own) and in *Map2k6*, *Dock10*, *Sparc* (neighbours) are validated with three orthogonal datasets of Stereo-seq, Xenium, and Allen Brain Cell Atlas MERFISH.

The hippocampal neuronal niche is not uniform but is instead segmented into anatomical sub-regions with cellular and microdomain-specific differences. Such regional differences could affect the interpretation of results derived from aggregating perturbed cells across different anatomical regions, a practice currently common in some dissociated and spatially resolved Perturb-Seq methodologies. Perturbed neurons within the dentate gyrus (DG) and CA1 sub-regions



revealed a substantial overlap in both upregulated and downregulated genes, indicating shared transcriptional responses in these two sub-regions. However, region-specific transcriptomic differences were also observed (Supplementary Fig. 13). We attempted to evaluate if the difference in transcriptomic response stems from the neurons residing within distinct anatomical regions, however, further analysis revealed that such direct comparison was not feasible because the DG and CA1 regions contain neurons of different intrinsic cellular subtypes. Spatial Perturb-Seq may eventually discriminate the perturbation responses

for cells of the same intrinsic cellular subtype residing within different anatomical regions (which non-spatial approaches cannot resolve), especially when sequencing scales to capture sufficient cells of the same intrinsic cell types.

Discussion

Spatial Perturb-Seq enables single-cell, whole-transcriptome profiling of cell autonomous and non-cell autonomous effects of genetic perturbations in intact tissue. This ability to segment a genetic

Fig. 3 | Spatial perturb-seq with probe-based xenium. **a** Transcript density map showing transcripts per bin (bin size: 20 μ m). **b** Magnified area showing cell segmentation boundaries and individual barcode molecules. Three barcodes are plotted: Barcode1-mSafe, Barcode16-Lrrk2, Barcode19-Srf. **c** Spatial plot showing transcript molecules in expected positions: *Nrn1* (neurons) in blue and *Gfap* (astrocytes) in green in the hippocampus, *Opalin* (oligodendrocytes) in yellow in the corpus callosum, *TdTomato* in the neuronal region of the injected hippocampus. **d** UMAP projection colored by cluster number on the left and annotated cell type on the right, showing recovery of the major cell types in the brain: excitatory neurons, inhibitory neurons, astrocytes, oligodendrocytes, OPCs, microglia, endothelial cells, and fibroblasts. Cell types were annotated based on canonical marker gene expression. **e** Spatial plot of the different cell types, showing expected positions in the brain. **f** Similar proportions of cell types were recovered for each slide, with the majority cell type being neurons as expected. **g** UMAP representation

showing the expression of canonical cell-type markers (Excitatory neurons: *Slc17a6*, *Neurod6*; Inhibitory neurons: *Sst*, *Gad1*; Microglia: *Trem2*; Oligodendrocytes: *Opalin*; OPCs: *Pdgfra*; Astrocytes: *Gfap*; Endothelial cells: *Pecam1*; Fibroblasts: *Dcn*). **h** Each dot is a cell, showing average expression of all transcripts (left) and target (right) counts against barcode counts. X- and Y-axis are normalized, where 100% represents counts from the highest expressing cell. See Supplementary Fig. 11. **i** Concordance of spatial transcriptomic signatures between Xenium and Stereo-seq. For both platforms, average log₂ fold-change was calculated for all Xenium-targeted genes (excluding barcodes), comparing barcode-positive vs. barcode-negative cells. Fold-changes from Xenium were compared to those from Stereo-seq across all DEGs and shown via a scatter plot. **j** Dot Plot showing DE genes for each perturbation, compared to all other perturbations. Not all perturbations had detectable DEGs with the gene panel used. Each box highlights DEGs specific to the perturbation indicated by the column.

perturbation's autonomous and non-cell autonomous effects reveals insights into cell-cell communication that would otherwise be lost via conventional dissociated single cell sequencing. Several recent technologies represent important steps toward enabling spatially resolved and multiplexed CRISPR screens, each with distinct strengths and limitations (Supplementary Data 6). Perturb-FISH⁴⁰ offers high spatial resolution but due to its reliance on targeted probes, it lacks whole-transcriptome coverage, which limits its ability for unbiased screens, consistent with our observations with the 247-gene Xenium probe panel. Perturb-map⁶ relies on the antibody-based Pro-Code⁴¹ system, limiting scalability and requiring imaging-based readouts and complex instrumentation. Perturb-DBiT⁴², which is microfluidics-based, requires more extensive technical optimization and does not offer single-cell resolution. Perturb-Multimodal⁴³ offers high plex and resolution, as well as multimodal measurements of transcriptomic, proteomic and morphological readouts, but is technically complex, relying on RCA-MERFISH. These methods underscore growing momentum in the field and the critical need for functional genomics with spatial and microenvironmental contexts. Importantly, these highlight progress towards compatibility of the methodology with a broader array of user-friendly detection platforms.

To assess compatibility of our system with an alternative probe-based detection system, we performed Xenium and showed robust detection of our barcodes and a tight linkage in decreased expression of target genes with increasing gRNA-barcode counts (Supplementary Fig. 11). We also performed RNA FISH against our nucleotide barcodes and found high specificity FISH spots ($R^2 = 0.78$) (Supplementary Fig. 14), suggesting feasible compatibility with other probe-based spatial transcriptomics platforms such as FISH (with individual probe technical optimizations) or the more readily available MERFISH⁴⁴ and CosMx⁴⁵. Further introduction of cell-type-specific barcode expression or sequential injections of different barcodes could endow capabilities for spatial lineage tracing⁴⁶ and spatial molecular timing^{47,48}.

The primary bottlenecks of Spatial Perturb-Seq are the cost of sequencing, as both perturbed cells and their unperturbed neighbours must be sequenced, and the naturally limited cell numbers in a tissue niche. However, the steady decline in sequencing costs and concurrent increase in scale will enable larger sample sizes to be sequenced to increase statistical power. Power analysis done with powsimR⁴⁹ to assess the statistical robustness of our platform showed FDR < 10% and TPR > 60–80% for 125–625 cells per perturbation (Supplementary Fig. 15), demonstrating the suitability of multiplexed experiments using this platform. As is the case with most spatial technologies, the accuracy of cell segmentation underpins the accuracy of Spatial Perturb-Seq, though significant advancements are being made in the field. Spatial Perturb-Seq unlocks a strategy to functionally interrogate genes at single-cell resolution within intact tissues, enabling the profiling of perturbation effects in both the target cells and their wildtype microenvironment.

Methods

Experimental methods

Ethics statement. Mice experiments were performed following protocols approved by A*STAR IACUC and in compliance with The National Advisory Committee for Laboratory Animal Research (NACLAR) Guidelines.

Mice

All animal work was performed under the guidelines of the Institutional Animal Care and Use Committee (IACUC protocol #211646). Mice were housed on a standard light cycle (12 h light/12 h dark) under pathogen-free conditions, with ambient temperature and humidity maintained within institutional guidelines. Female Rosa26-CAG-Cas9 C57BL/6J mice (Jackson Laboratory JAX 024858) aged 8–16 weeks were used.

Stereotaxic injections

Animals were anesthetized with ketamine/xylazine (injected intraperitoneally) and placed on a thermostatically controlled heating pad. Animals were checked for sedation. Hair was removed from the incision site on the scalp with a shaver and the area was cleaned with an alcohol wipe. Corneas were protected with an ophthalmic lubricating ointment. Animals were then positioned in a stereotaxic head holder with the skull firmly barred. An incision on the scalp was made and the scalp reflected over the brain region of interest. For all experiments, the left hippocampus was targeted with the following stereotaxic coordinates relative to the bregma: AP −1.7, ML +1.6, DV −1.8. A handheld dental drill was used to drill a <1 mm diameter hole in calcified bone in the AP, ML region of interest. The cannula was inserted into the region of interest and the 5×10^8 viral particles of AAVs were injected at a rate of 2–5 nL/s in a maximal volume of 1.5 μ L, so as not to cause tissue injury. Following injection, the cannula was withdrawn after 5 minutes. The incision on the scalp was closed with Vetbond Tissue Adhesive. Atipamezole was administered intraperitoneally for anaesthesia reversal and Buprenorphine was administered subcutaneously for pain relief. Following this, animals were placed on a thermostatically controlled heating pad for recovery and monitoring. For 3 days post-surgery, Buprenorphine was administered for pain relief, and animals were monitored for signs of distress and wound inflammation. Mice were kept for 3 weeks under standard conditions before being sacrificed by CO₂ asphyxiation for tissue extraction and processing.

gRNA library design

We designed three guides each for 18 loci – 17 genes and a safe harbour locus (*mSafe*) as a control (54 guides in total). All guides and their predicted off-target binding are included in Supplementary Data 1 and 2. The three guides for the same locus are linked to the same barcode on one vector, to facilitate detection of the genetic knockout

via a single unique barcode. Each vector contains TdTomato, with 18 vectors in total. We focused our library on target genes associated with GWAS. From the NHGRI-EBI Catalog of human genome-wide association studies⁵⁰, we chose GWAS hits associated with Alzheimer's Disease (https://www.ebi.ac.uk/gwas/efotraits/MONDO_0004975) (CLU, LRRK2, NDUFAF2, RBFOX1, TREM2), Amyotrophic Lateral Sclerosis (https://www.ebi.ac.uk/gwas/efotraits/MONDO_0004976) (C9orf72, CFAP410, DPP6, TBK1), and Parkinson's disease (SH3GL2, STK39) (https://www.ebi.ac.uk/gwas/efotraits/MONDO_0005180). These genes have mouse orthologs and are expressed in the mouse brain. We also included 6 positive control genes (Fasn, Flcn, Gfap, Olig2, Rrag, Srf) and a safe harbour site. With these 18 genes/loci of interest, we designed three gRNAs each using the online tool [CHOPCHOP](#), with *Mus musculus* (mm10/GRCm38) as the genome and with default parameters. Best-scoring guides (based on their on-target and off-target scoring system) within the first few exons were selected.

Cloning of gRNAs plasmid pool

The plasmid pAAV-CAG-TdTomato (codon diversified) was ordered from Addgene (#59462). The plasmid backbone (2 µg) was digested with HindIII/SalI (NEB) for 1 h at 37 °C followed by an inactivation step for 20 min at 80 °C. DNA fragments consisting of both the overlapping sequences with the digested pAAV-CAG-TdTomato backbone and the 450 bp barcodes were ordered from IDT. PCR reactions were used to amplify the DNA fragments. The Gibson assembly reaction was set as follows: 100 ng of digested plasmid backbone, 10 ng of amplified DNA fragment, 10 µL NEBuilder HiFi DNA Assembly Master Mix (NEB, E2621L) and H₂O up to 20 µL total reaction. The reaction was incubated for 1 h at 50 °C. Sul of the Gibson reaction was used for transformation using Stbl3 cells (Thermo Fisher, Cat#C737303). Sequencing reactions (Biobasic) were used to confirm the correct insertion of the barcodes. Following barcode cloning, three gRNAs were designed for each locus of interest to be inserted into each bar-coded plasmid. Golden gate cloning strategy was employed to arrange the three gRNAs in tandem, each driven by a U6 promoter. The gRNAs were then inserted into the NdeI-digested pAAV-CAG-TdTomato using Gibson assembly.

AAV production and purification

AAVs were generated in-house⁵¹. Briefly, AAVs were packaged via a triple transfection of 293AAV cell line (AAV-100, Cell Biolabs, San Diego, CA, USA). The cells were plated in a HYPERFlask 'M' (Corning) in growth media, which consisted of DMEM + glutaMax + pyruvate + 10% FBS (Thermo Fisher), that was supplemented with 1X MEM non-essential amino acids (Gibco). Transfection was done when the cells were between 70% and 90% confluent. Media was replaced with fresh pre-warmed growth media before transfection. For each HYPERFlask 'M', 200 µg of pHelper (Cell Biolabs), 100 µg of pRepCap encoding capsid proteins for serotype DJ or 2J, and 100 µg of pZac-CAS1-GFP or pZac-CMV-CasRx gRNAs were mixed in 5 mL of DMEM, and 2 mg of PEI "MAX" (Polysciences) (40 kDa, 1 mg/mL in H₂O, pH 7.1) was added for PEI: DNA mass ratio of 5:1. The mixture was incubated for 15 min before being transferred to the cell media drop-wise. The day after transfection, the media were changed to DMEM + glutamax + pyruvate + 2% FBS. 48–72 h after transfection, the cells were harvested by scraping or dissociation with 1X PBS (pH 7.2) + 5 mM EDTA and then pelleted at 1500 g for 12 min. Cell pellets were resuspended in 1–5 mL of lysis buffer (Tris HCl pH 7.5 + 2 mM MgCl + 150 mM NaCl), before freeze-thawing thrice using a dry-ice-ethanol bath and a 37 °C water bath. Cell debris was clarified via centrifuging at 4000 g for 5 min, and the supernatant was collected. To remove the unpackaged nucleic acids, the collected supernatant was treated with 50 U/mL of Benzonase (Sigma–Aldrich) and 1 U/mL of RNase cocktail (Invitrogen) for 30 min at 37 °C. Next, the lysate was loaded on top of a discontinuous density gradient that consisted of 6 ml each of 15%, 25%, 40%, and 60%

Optiprep (Sigma–Aldrich) in a 29.9 mL Optiseal polypropylene tube (Beckman–Coulter). The tubes were then ultra-centrifuged at 330,000 g, for 1.5 h at 18 °C, using a Type 70 Ti rotor. The 40% fraction was extracted and dialyzed with 1X PBS (pH 7.2) supplemented with 35 mM NaCl, using Amicon Ultra-15 (100 kDa MWCO) (Millipore). qPCR was then carried out with the ITR-sequence-specific primers and probes, using the ATCC reference standard material 8 (ATCC) as a standard, to determine the titres of the purified AAV vector stocks.

Guide vector recombination assay

HEK293 cells were seeded in 48-well plates and transduced with recombinant adeno-associated viruses (AAVs) for 48 h at multiplicity of infection (MOI) of 10000. The transduction conditions included transduction with either AAV-PhP.eB-TdTomato-gRNAs or AAV-PhP.eB-GFP and a non-transduced control. At 48 h post-transduction, the cells were washed three times with phosphate-buffered saline (PBS), followed by cell lysis using 50 µL of QuickExtract™ DNA Extraction Solution (Lucigen) per well. The lysates were incubated according to the manufacturer's instructions, and 1 µL of the resulting lysate was used as template DNA for downstream PCR analysis. PCR was performed using a primer pair designed to detect recombination events at the target locus. The primers were synthesized based on a calculated melting temperature (T_m) of 55 °C: Forward primer: 5'-CACTTGGCAGTA-CATCAAGTGT-3', Reverse primer: 5'-GCCATTTACCGTCATTGACGT-3'. Each 25 µL PCR reaction contained 1X high-fidelity PCR buffer, 0.4 µM of each primer, 0.5 U of high-fidelity Q5 DNA polymerase (NEB), and 1 µL of lysate template. The thermal cycling conditions were as follows: initial denaturation at 98 °C for 1 min; 35 cycles of denaturation at 95 °C for 15 s, annealing at 54 °C for 15 s, and extension at 70 °C for 1 min; followed by a final extension at 70 °C for 5 min. PCR products were analyzed on a 1% agarose gel stained with Gel Red. Electrophoresis was performed at 120 V for approximately 40 min. Bands were visualized using a gel documentation system, and densitometry measurements were done in ImageJ.

Tissue dissection, dissociation and FACS

After mice were sacrificed by CO₂ asphyxiation, intracardial perfusion with ice-cold PBS was carried out to remove blood cells. Brains were dissected and either embedded in OCT for sectioning or dissociated into single cells. Dissociation of brain tissue was done with the Neural Dissociation Kit (Miltenyi Biotec) following the manufacturer's instructions. Briefly, the brain tissue of interest was minced with a razor blade, followed by enzymatic digestion at 37 °C and triturated. The cell suspension was then strained through a 70 µm cell strainer and centrifuged at 300 g for 10 min. The cell pellets were then re-suspended in Hibernate A with 10 µM Calcein-AM (Invitrogen) and 1 µg/mL Propidium Iodide (PI) (Stemcell), for FACS sorting. Live, single cells were selected by positive selection of Calcein-AM and negative selection of PI. Immediately after FACS sorting, cells were used for the generation of 10×3' chromium libraries and sequenced on Novaseq (Illumina).

Coverslip functionalization and sample preparation for FISH

Coverslip functionalization was carried out before use for tissue sectioning. Coverslips (Warner Instruments, cat. no. 64–1500) were cleaned in 1 M KOH for 1 h before being rinsed thrice in Mili Q water. The coverslips were rinsed with 100% methanol before being functionalised in an amino-silane solution (3% vol/vol (3-aminopropyl) triethoxysilane (Merck cat no. 440140), 5% vol/vol acetic acid (Sigma, cat. no. 537020) for 2 min at room temperature. Following this, the coverslips were rinsed thrice with Mili Q water before being dried overnight at 47 °C in an oven. The mouse brain sample was frozen in optimal cutting temperature compound and sectioned into 7 µm sections onto functionalized coverslips using a cryostat. The sections were fixed using 4% vol/vol paraformaldehyde in 1× PBS for 15 min, then rinsed with 1× PBS and stored at –80 °C.

FISH

The encoding probes were diluted in a 10% hybridization solution that was composed of 10% deionized formamide (Ambion™ Cat: AM9342) (vol/vol), 1 mg/ml yeast tRNA (Life Technologies, cat. no. 15401-011) and 10% dextran sulfate (Sigma, cat. no. D8906) (wt/vol) in 2× SSC. A final concentration of 10–50 nmol per probe was used. Mouse brain sections were permeabilized with 70% ethanol overnight at 4 °C. After permeabilization, samples were rinsed twice with 2× SSC. Next, encoding probe staining was performed and samples were stained for 16 h at 37 °C. After encoding probe staining, samples were washed in 10% formamide wash buffer (10% deionized formamide in 2× SSC) at 37 °C for 15 min, twice. The samples were rinsed thrice with 2× SSC, before DAPI (Sigma, cat. no. D9564) staining was carried out.

Microscopy for FISH

Imaging was carried out on a home-build microscope⁵². A Nikon CFI Plan Apo Lambda 60 × 1.4-n.a. oil-immersion objective for imaging with an exposure time of 1 s. The microscope was built around a Nikon Ti2-E body with a Marzhauser SCANplus IM 130 mm × 85 mm motorized X–Y stage, and the Nikon Perfect Focus System maintained focus during sample imaging. We used a Coherent Obis 405 100-mW laser for the DAPI channel, and MPB Communications fiber lasers were used as illumination for tdTomato (592 nm, 2RU-VFL-P-500-592-B1R (500 mW)), Cy5 (647 nm, 2RU-VFL-P-1000-647-B1R (1000 mW)) and IRDye 800CW (750 nm, 2RU-VFL-P-500-750-B1R (500 mW)). Samples were mounted onto the microscope stage via the use of a flow chamber (Biopetech, cat. no. FCS2). This allowed buffer exchange to be carried out via the use of a computer-controlled fluidics system. The sample was stained at room temperature for 15 min with readout probe solution via buffer exchange, prior to imaging. The readout probe solution consists of 10 nM of each fluorescently labelled readout probe, 10% deionized formamide (vol/vol) and 10% dextran sulfate (wt/vol) in 4× SSC. After hybridization, 2× SSC was flowed in before a rinse with 10% formamide wash buffer. 2× SSC was flowed again before imaging buffer. The imaging buffer contains 2× SSC, 10% glucose, 50 mM Tris-HCl pH 8, 2 mM Trolox (Sigma, cat. no. 238813), 40 µg/mL catalase (Sigma, cat. no. C30), and 0.5 mg/mL glucose oxidase (Sigma, cat. no. G2133). After imaging, samples were washed with 55% formamide wash buffer containing 0.1% TritonX-100 to remove the fluorescent signals.

10X library preparation and sequencing

Libraries from dissociated single cells were generated using Chromium Next GEM Single Cell 3' Reagent Kits v3.2 (Dual Index) (10X Genomics). Cell numbers were counted by FACS, and 20,000 cells per mouse were used to generate each library. Libraries were generated according to the manufacturer's instructions, with 12 cycles of cDNA amplifications, 25 ng of cDNA input and 14 cycles of sample index PCR. Barcode sequences were not amplified. Libraries generated had a fragment length of 400–450 bp and were sequenced on an Illumina Nova-seq (PE150).

iSeq library preparation and sequencing

To confirm the presence of indels in brain tissue after AAV transduction, 50 000 TdTomato+ cells from 4 brains were lysed with Quick-Extract (Lucigen) according to the manufacturer's instructions. Uninjected brain tissue was used as WT control. Two rounds of PCR were used. For PCR1, genotyping primers were designed using Primer3 (<https://primer3.ut.ee/>) to amplify a 150–300 bp region around the Cas9 cut site with 34 PCR cycles. The following sequences were added to the forward and reverse genotyping primers, respectively, for the second round of amplification: CTTTCCCTACACGACGCTTCCGA TCTNNNNNN, GGAGTTCAGACGTGTGCTCTTCCGATCT. NNNNN is used to introduce diversity at the start of the iSeq read to improve read quality. For PCR2 used to index samples, 6 PCR cycles were used.

For both PCR1 and PCR2, Q5 High Fidelity Master Mix (NEB) was used according to the manufacturer's recommended PCR temperatures and parameters. Indexed PCR products from PCR2 were pooled, run on a gel, and gel-purified using Wizard SV Gel and PCR Clean-Up System (Promega). DNA concentration was measured with Qubit HS dsDNA kit (Vazyme). The library was spiked with 2% PhiX library (Illumina) and sequenced on an iSeq 100 system.

Stereo-seq chip preparation

OCT blocks were stored at –80 °C and equilibrated at –20 °C for 2 h prior to sectioning. OCT blocks were cryosectioned at a thickness of 10 µm using a CM1950 cryostat. Following satisfactory QC results (RIN = 8.5–8.6), mouse brain coronal sections were collected on the Stereo-seq chip surface. Tissue sections were adhered to a Stereo-seq chip surface and incubated at 37 °C for 5 min. The tissues were then fixed in methanol and incubated at 20 °C for 30 min. The tissue sections were then permeabilized at 37 °C for 10 min and then washed with 0.1× SSC buffer. The RNA released from permeabilized tissues was captured using DNB probes and reverse-transcribed at 42 °C for 3 h. After in situ reverse transcription, tissues were removed with the tissue removal buffer. The chips were then incubated with 400 µL cDNA release buffer overnight at 55 °C, and then washed once with 400 µL of 0.1× SSC buffer. The released cDNA was then collected and purified using 0.8× Ampure XP Beads. The cDNA was then amplified using the following PCR conditions 95 °C for 5 min, 15 cycles of 98 °C for 20 s, 58 °C for 20 s, then 72 °C for 3 min, and a final incubation at 72 °C for 5 min. PCR products were purified using 0.6× Ampure XP Beads and the concentrations of cDNA were quantified using a Qubit™ dsDNA Assay Kit. Three Stereo-seq chips were processed from brains collected across four mice. The layout and metrics for each chip are detailed in Supplementary Fig. 7 and Supplementary Data 3.

Stereo-seq library preparation and sequencing

STOMICS libraries were constructed with the STOMICS Gene Expression kit. The cDNA was first checked for the presence of barcodes by PCR. Briefly, 20 ng of cDNA were fragmented at 55 °C for 10 min. The fragmented products were then amplified using the following PCR conditions one cycle at 95 °C for 5 min, 13 cycles at 98 °C for 20 s, 58 °C for 20 s and 72 °C for 30 s, and lastly one cycle at 72 °C for 5 min. The PCR products were purified using Ampure XP Beads for DNB generation and were finally sequenced (100 bp PE) using T7 sequencer.

Preparation of slides for Xenium V1 assay

The Xenium v1 assay workflow was carried out according to the manufacturer's instructions (10x Genomics, User Guide CG000749). Briefly, fresh frozen tissue sections (10 µm) placed on Xenium slides were fixed and permeabilized with minor modifications to the manufacturer's protocol (10x Genomics, User Guide, CG000581). Before immersing sections in 1% SDS, sections were photobleached in 1× PBS for 1 h at 4 °C. Following fixation and permeabilization, probe hybridization with a combination of both the pre-designed mouse brain panel (247 genes) and an add-on custom panel of barcode sequences was carried out. Probes that were bound to the target RNA were ligated at both ends to generate a circular DNA probe. Rolling circle amplification of the circularized DNA probe then generated multiple copies of the gene-specific barcode for each RNA target. This was followed by morphology-based cell segmentation staining (boundary stain used: ATP1A1 + CD45 + E-Cadherin) before continuing with autofluorescence quenching and nuclei staining. Slides were then loaded onto the Xenium Analyzer for high throughput and automated in situ imaging and analysis. Fluorescence-labelled oligos bound to the amplified DNA probes and tissues underwent multiple rounds of fluorescent probe hybridization, imaging, and probe removal to generate a codeword specific for each barcode. Each codeword was then converted into a gene identity. Two Xenium slides were prepared from adjacent

sections of a single OCT block with brains from two mice (run metrics detailed in Supplementary Data 5). In total, Xenium assays were performed on eight brain sections across 2 mice.

Long-read sequencing for assessing integrity of U6-sgRNA cassette

Amplicon libraries from 3x U6-sgRNA constructs were prepared for long-read sequencing using the Oxford Nanopore Technologies Ligation Sequencing Kit V14 (SQK-LSK114) following the manufacturer's standard protocol. Briefly, the library preparation workflow included DNA end repair and preparation, adapter ligation, and cleanup steps. The prepared sequencing library was loaded onto a MinION Flow Cell R10.4.1 (FLO-MIN114). Sequencing was performed on a MinION Mk1C device for 48 hours of continuous data acquisition. Real-time basecalling and data streaming were conducted using the MinKNOW software.

Computational methods

Barcode design. One of our goals is to design nucleotide barcodes long enough to be probed against that are tolerated in cells. We wrote a script that can generate a set of barcodes with these features: having an appropriate length—long enough for probing against, but short enough to avoid inducing deleterious effects in the cell (450 nt); not containing stop codons, to avoid nonsense mediated decay; having a high edit distance between each barcode; having a high edit distance between different segments of the same barcode, to allow design of multiple probes against the same barcode (with each probe being 15–50 nt long); being distinct from the human and mouse transcriptome, to avoid confusion with endogenous genes. With these criteria, we generated a set of 22 barcodes and used 18 for our library, with one barcode per gene (Fig. 1a schematic).

FISH encoding probe design

Fourteen encoding probes were designed against the 450 bp barcode 22 using the Stellaris Probe Designer software (LGC Biosearch Technologies). Parameters used were: Masking level = 5 (mouse), Oligo length = 19–20 nt, Min. spacing length = 2nt. Each encoding probe sequence was flanked on both sides by the readout probe sequence 5'-TCTGTTTGACGCGCT-3' with a spacer nucleotide (A) in between the readout and the encoding probe region. The concatenated sequences and the readout probe (5'IRD800CWN/AGCGCGTCAAACAGA) were purchased from Integrated DNA Technologies (IDT). Encoding probes used were: 5'-TCTGTTTGACGCGCT A ATACGTAAGGCAAGTCCGAC A TCTGTTTGACGCGCT-3'

5'-TCTGTTTGACGCGCT A TGCTTTCTGTACGTGGGTAG A TCTGTTTGACGCGCT-3'

5'-TCTGTTTGACGCGCT A ATAATGGCACGAAGCCAGAC A TCTGTTTGACGCGCT-3'

5'-TCTGTTTGACGCGCT A AGGAAAGGCGTGAGGGTACG A TCTGTTTGACGCGCT-3'

5'-TCTGTTTGACGCGCT A CTCCACATTTGATTGGTTTG A TCTGTTTGACGCGCT-3'

5'-TCTGTTTGACGCGCT A CAGCATGACGGGCAGTGAAC A TCTGTTTGACGCGCT-3'

5'-TCTGTTTGACGCGCT A GTTCGACGTACCAAGCAATT A TCTGTTTGACGCGCT-3'

5'-TCTGTTTGACGCGCT A CGTGGGAACCGTATTGGAAT A TCTGTTTGACGCGCT-3'

5'-TCTGTTTGACGCGCT A ATGCCGCGTCTTGCCTAC A TCTGTTTGACGCGCT-3'

5'-TCTGTTTGACGCGCT A CATGATCGCCACAATCGATG A TCTGTTTGACGCGCT-3'

5'-TCTGTTTGACGCGCT A AAATTTCCGATTGCAGCGTT A TCTGTTTGACGCGCT-3'

5'-TCTGTTTGACGCGCT A CCGCTGACTGTAATAGGTAC A TCTGTTTGACGCGCT-3'

5'-TCTGTTTGACGCGCT A GGTTGAGATCGGAAGCAACG A TCTGTTTGACGCGCT-3'

5'-TCTGTTTGACGCGCT A ACTGACTCGGAGTGATGTTT A TCTGTTTGACGCGCT-3'

10X sequencing data processing

Sequencing results from 10X libraries were processed and demultiplexed with Cell Ranger 7.1.0 pipeline version (10X Genomics). An edited mouse reference genome (mm10-2020) with the 450 bp barcodes added was used for alignment and generating the UMI count matrices³³. Briefly, we used Cell Ranger's mkref function to construct a custom index for read alignment, with the addition of the barcode sequences and names to the "genes.gtf" and "genome.fa" files in the mm10 mouse reference genome. This allows for barcode counts to be included in count matrices. Sequencing saturation of all reads and of barcode reads was calculated for all generated libraries and were >0.6 and >0.5, respectively. Filtered count matrices were analyzed using the Seurat package (version 4.3.0.1) in R (version 4.3.1). Briefly, QC was performed to keep only cells with percentage mitochondrial reads below 15%, and nFeature_RNA between 200 and 7000, to exclude empty droplets, multiplets and dying cells. Counts were then normalized and scaled with the SCTransform() function. Following this, Principal Component Analysis (PCA) was run on the top 2000 most variable features, using the RunPCA() function. Clusters were identified with the FindNeighbors() function by generating a K-nearest neighbour graph with 10–16 dimensions, and clustered using the Louvain algorithm with a resolution of 0.3–0.4, using the FindClusters() function. The cells were then represented by a 2-dimensional Uniform Manifold Approximation and Projection (UMAP) graph. Lastly, coarse cell-type annotation was performed using a combination of analyzing the most highly expressed marker genes of each Seurat cluster, and expression of known canonical marker genes for each brain cell type (neurons, oligodendrocytes, astrocytes, endothelial cells, microglia). As each cluster was represented by all batches, no batch correction or integration methods were applied. For the analysis of oligodendrocyte perturbation phenotypes, oligodendrocytes expressing barcodes linked to *mSafe*, *Dpp6*, *Lrrk2*, *Gfap*, *Cfap410*, *Rbfox3* are grouped as "control", as these genes are not expected to impact oligodendrocyte function.

Assignment of barcode/gRNA status for 10X and Stereo-seq sequencing data

For assignment of barcode-positive and barcode-negative status, negative status was assigned to cells with no expression of any barcodes, and all other cells were positive. For assignment of gRNA status, cells were classified by the expression of barcodes as follows: No barcodes expressed = Neg; more than 1 unique barcode expressed = Multiple; exactly 1 unique barcode expressed = the gene corresponding to that barcode.

Gene editing analysis

Following preparation and sequencing of iSeq libraries as described above, quantification of the percentage of indels was calculated with CRISPresso2 (v2.0.20) with default parameters (default_min_aln_score 60, quantification_window_center -3, exclude_bp_from_left 15, exclude_bp_from_right 15, quantification_window_size 1, plot_window_size 20, min_bp_quality_or_N 0). % Indel for each guide is calculated as % Modified (edited) - % Modified (WT). For all loci, there was at least 20k aligned reads.

Stereo-seq data processing

For processing of fastq files, the SAW pipeline version V7.0.0 and Image Studio version 3.0.0 were used. Genome and gtf files used for alignment were generated by combining the mouse reference GRCm39 and the 450bp barcodes. CID (Coordinate Identity)

sequences were mapped to the coordinates on the Stereo-seq chip, allowing 1 mismatch. Reads are filtered based on Q30 phred quality score, indicating less than 1 in 1000 chance of base calling error. Following alignment to the reference, a gene count file was generated quantifying deduplicated, annotated reads. This gene count matrix is then aligned to the image (image registration). Cell segmentation was conducted using nuclear-stained tissue images and gene expression data, with the DeepCell algorithm. Following the generation of a Stereo object from the cellbin GEF file, a h5ad file and Seurat object was generated. For QC, cells with percent.mito > 20, and number of counts between <100 or > 6000 were discarded. After QC, cells were processed through the standard Seurat pipeline, as described under the “10X sequencing data and processing” methods section. Neurons expressing barcodes linked to *mSafe*, *Gfap*, *Stk39*, *Trem2*, and *Olig2* were grouped as control, as they are lowly expressed in neurons and/or are not expected to exert a great phenotyping change.

Selection of hippocampal neurons with BANKSY

BANKSY was run on each chip separately, after data scaling, and before PCA using npcs = 30. We used lambda 0.2 to achieve spatially-informed cell type embeddings. Higher lambda values resulted in spatial domain segmentation with separate sections on the same Stereo-seq chip being clustered separately. The following parameters were used: spatial_mode = “knn_r”, ndim = 2, k_geom = 15. Clusters corresponding to hippocampal CA1, 3 and dentate gyrus (DG) neurons, the hippocampal neuronal niche, were kept for analysis.

Differential gene expression (DGE) and analysis

DGE analysis was performed with Seurat’s FindMarkers function with the following parameters were: min.pct = 0.1, logfc.threshold = 0.5, test.use = “wilcox”. For reporting of the number of DEGs, the same number of cells for each perturbation was used, to ensure comparability. For calculation of neighbour DEGs, 15 neighbours were identified with BANKSY (spatial_mode = “knn_r”) for each perturbed cell in the group, and barcode-positive neighbours were excluded. The wildtype neighbours of the 2 groups (case and control), were then compared. Genes with $p < 0.05$ and $lfc > 0.5$ were counted as DEGs. For volcano plots, p values were adjusted with the p.adjust function to obtain fdr values. Genes with $fdr < 0.05$ and $lfc > 0.5$ were marked as significant.

Cell-cell communication (CCC) analysis

We used the LIANA (Ligand-Receptor Analysis) tool, a computational pipeline for prioritizing ligand-receptor interactions based on databases like CellPhoneDB, connectomeDB2020, and CellChat, to elucidate cell-cell communication between neurons. As cells within the same microenvironment are more likely to be interacting, we focussed on the communication between perturbed cells and their immediate neighbours. For each of these 3 perturbed groups, Srf-KO, Lrrk2-KO, and control, the perturbed cells and their wildtype/unperturbed neighbours were identified. 15 neighbours were identified for each perturbed cell with BANKSY (spatial_mode = “knn_r”), and barcode-positive neighbours were excluded. For each pair (Srf-KO + neighbours, Lrrk2-KO + neighbours, control-KO + neighbours), we applied LIANA with statistical methods “sca” and “natmi” to calculate and prioritize top interactions, as these are most likely to be biologically relevant. We modelled the perturbed cells as the source cells and their wildtype neighbours as the target cells. We then identified common ligand-receptor pairs from the prioritized list between Lrrk2-KO and control, and Srf-KO and control, and show the top 20 ligand-receptor pairs with biggest differences in sca.LRscore.

Processing of xenium spatial transcriptomics data

Xenium spatial transcriptomics data were processed following the Seurat v5 Spatial Vignette (https://satijalab.org/seurat/articles/seurat5_spatial_vignette_2). Xenium output files were first converted to Seurat

objects with the LoadXenium() function with the following parameters: mols.qv.threshold = 20, cell.centroids = TRUE, molecule.coordinates = TRUE, segmentations = “cell”, flip.xy = FALSE. Gene expression data were normalized and scaled using default parameters with the SCTransform() function, followed by dimensionality reduction with the RunPCA() and RunUMAP() with dims = 1:10. Clustering was performed with the FindNeighbors() with dims = 1:10, and FindClusters() functions. Images were produced with the ImageDimPlot() function. Cell types were annotated based on the expression of canonical marker genes. Cells with barcode count >1 were classified as barcode-positive.

Allen Brain Cell Atlas data and image preprocessing

Images of hippocampal neurons from the Allen Brain Cell Atlas (<https://portal.brain-map.org/atlas-and-data/bkp/abc-atlas>) coloured by gene expression, were segmented into equal grid squares in ImageJ (10 by 10 pixels). Mean fluorescence intensity per gene per grid was exported as a CSV file. Spearman’s ρ was computed for (i) within-grid (“own-cell”) and (ii) adjacent-grid (“neighbor-cell”) gene pairs (radius of 20 pixels; 12 cells), focusing on *Lrrk2* expression correlations. Spearman’s rank correlation was chosen because it is non-parametric and robust to non-linear relationships. Correlations were interpreted as measures of local co-variation rather than causality. “Own-cell” correlations represent transcriptional co-expression within the same spatial microdomain, whereas “neighbor-cell” correlations capture potential intercellular coordination or shared microenvironmental influences.

Calculation of R2 values

The Cell Profiler software was used to calculate the coefficient of determination (R^2 value) between TdTomato protein intensity and the number of FISH spots of TdTomato RNA, Barcode, and Actb control, from confocal images of TdTomato protein and FISH staining. A cropped image was first analyzed using IdentifyPrimaryObjects (for analysis of TdTomato protein fluorescent intensity or FISH spots), then gridded using the DefineGrid module (200 grid squares). Next, the IdentifyObjectsInGrid module was done, followed by the RelateObjects module to relate parent objects (grid squares) to child objects (TdTomato protein fluorescent intensity or FISH spots). The Pearson R^2 coefficient was then computed (one point per grid area) with a fixed intercept at the origin, to assess the correlation between TdTomato protein fluorescent intensity and the number of FISH spots.

Power analysis with powsimR

First, we modelled the noise in the dataset, as this affects the sensitivity of the platform. To do this, the mean-dispersion relationship was estimated using a negative binomial model, using count data from neurons from chip2 randomly down-sampled to 10000 cells, and normalized with scan. These distribution statistics were then used to set up simulations, using the following parameters: nsims = 25, p.DE = 0.1, pLFC = 2, LibSize = “equal”. Lastly, marginal and conditional FDR and TPR are evaluated in the evaluateDE function with the following parameters: alpha.type = “adjusted”, MTC = “BH”, alpha.nominal = 0.1, stratify.by = “dispersion”, filter.by = “dispersion”, strata.filtered = 1, target.by = “lfc”, delta = 0. As the average number of cells we obtained per perturbation was 122, marginal and conditional FDR and TPR were shown with the following number of cells: 5 vs 5, 25 vs 25, 125 vs 125, 625 vs 625.

Statistics and reproducibility

Statistical analyses were done using GraphPad Prism 10.6.1 for Windows unless otherwise stated. Data are represented as mean, with error bars representing Standard Error of the Mean (SEM), with specific statistical analyses and sample sizes indicated in the respective figure legends. Unless otherwise stated, unpaired two-tailed Student’s t test in GraphPad was used to compare two groups (control and perturbed).

A p-value of <0.05 was considered significant. Statistical tests in Supplementary Figs. 5e, 12c, 12d, 13b were done using ggsignif R package (version 0.6.4) using the geom_signif() function with default parameters, and with white geom_boxplot() showing the median, first and third quartiles in the violin plots. All mice used were female adults aged 8–16 weeks, and no animals were excluded from the study. Blinding was not performed and was not required, as all mice served as biological replicates without separate experimental groups of mice. Negative controls (mSafe) were internal to each mouse because perturbations were multiplexed within the same brain. Power calculations were performed with the powsimR R package (version 1.2.5), enabling reproducible estimation of expected statistical power.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Raw and processed sequencing data have been deposited at NCBI's Gene Expression Omnibus (GEO) with accession numbers [GSE274447](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE274447) (Stereo-seq) and [GSE274058](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE274058) (scRNA-seq). Source data are provided with this paper.

Code availability

Custom code is available at Github (<https://github.com/kimberle9/spatialperturbseq>) and archived on Zenodo⁵⁴ (<https://doi.org/10.5281/zenodo.17959756>). The repository is released under the MIT License, an Open Source Initiative–approved license. There are no restrictions on access or reuse. Analysis of the processed data in this study was done using open-source R packages Seurat (<https://github.com/satijalab/seurat>), BANKSY, powsimR (<https://github.com/bvieth/powsimR>), scCustomize (<https://github.com/samuel-marsh/scCustomize>), and open-source Python packages Stereopy (<https://stereopy.readthedocs.io/en/latest/>) and SAW.

References

- Dixit, A. et al. Perturb-seq: dissecting molecular circuits with scalable single-cell RNA profiling of pooled genetic screens. *Cell* **167**, 1853–1866.e17 (2016).
- Adamson, B. et al. A multiplexed single-cell CRISPR screening platform enables systematic dissection of the unfolded protein response. *Cell* **167**, 1867–1882.e21 (2016).
- Santinha, A. J. et al. Transcriptional linkage analysis with in vivo AAV-Perturb-seq. *Nature* **622**, 367–375 (2023).
- Zheng, X. et al. Massively parallel in vivo perturb-seq reveals cell-type-specific transcriptional networks in cortical development. *Cell* <https://doi.org/10.1016/j.cell.2024.04.050> (2024).
- Jin, X. et al. In vivo Perturb-Seq reveals neuronal and glial abnormalities associated with autism risk genes. *Science (New York, N.Y.)* **370**, eaaz6063 (2020).
- Dhainaut, M. et al. Spatial CRISPR genomics identifies regulators of the tumor microenvironment. *Cell* **185**, 1223–1239.e20 (2022).
- Gu, J. et al. Mapping multimodal phenotypes to perturbations in cells and tissue with CRISPRmap. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-024-02386-x> (2024).
- Kudo T. et al. Multiplexed, image-based pooled screens in primary cells and tissues with PerturbView. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-024-02391-0> (2024).
- Feldman, D. et al. Optical pooled screens in human cells. *Cell* **179**, 787–799.e17 (2019).
- Wang, C., Lu, T., Emanuel, G., Babcock, H. P. & Zhuang, X. Imaging-based pooled CRISPR screening reveals regulators of lncRNA localization. *Proc. Natl. Acad. Sci. USA* **116**, 10842–10851 (2019).
- Chen, A. et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* **185**, 1777–1792.e21 (2022).
- Marco Salas, S. et al. Optimizing xenium in situ data utility by quality assessment and best-practice analysis workflows. *Nat. Methods* **22**, 813–823 (2025).
- Platt, R. J. et al. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell* **159**, 440–455 (2014).
- Montini, E. et al. The genotoxic potential of retroviral vectors is strongly modulated by vector design and integration site selection in a mouse model of HSC gene therapy. *J. Clin. Invest.* **119**, 964–975 (2009).
- Hill, A. J. et al. On the design of CRISPR-based single-cell molecular screens. *Nat. Methods* **15**, 271–274 (2018).
- Chan, K. Y. et al. Engineered AAVs for efficient noninvasive gene delivery to the central and peripheral nervous systems. *Nat. Neurosci.* **20**, 1172–1179 (2017).
- Pandey, S. et al. Disease-associated oligodendrocyte responses across neurodegenerative diseases. *Cell Rep.* **40**, 111189 (2022).
- Floriddia, E. M. et al. Distinct oligodendrocyte populations have spatial preference and different responses to spinal cord injury. *Nat. Commun.* **11**, 5860 (2020).
- Wang, J. et al. Olig2 ablation in immature oligodendrocytes does not enhance CNS myelination and remyelination. *J. Neurosci.* **42**, 8542–8555 (2022).
- Meireles, A. M. et al. The lysosomal transcription factor TFEB represses myelination downstream of the rag-regulator complex. *Dev. Cell* **47**, 319–330.e5 (2018).
- Jiang, L. et al. Systematic reconstruction of molecular pathway signatures using scalable single-cell perturbation screens. *Nat. Cell Biol.* **27**, 505–517 (2025).
- Hu, P. et al. Dissecting cell-type composition and activity-dependent transcriptional state in mammalian brains by massively parallel single-nucleus RNA-seq. *Mol. cell* **68**, 1006–1015.e7 (2017).
- Bannon, D. et al. DeepCell Kiosk: scaling deep learning-enabled cellular image analysis with Kubernetes. *Nat. Methods* **18**, 43–45 (2021).
- Singhal, V. et al. BANKSY unifies cell typing and tissue domain segmentation for scalable spatial omics data analysis. *Nat. Genet.* **56**, 431–441 (2024).
- Yang, S. et al. Pathogenicity and functional analysis of CFAP410 mutations causing cone-rod dystrophy with macular staphyloma. *Front. Med.* **10**, 1216427 (2023).
- Tolosa, E., Vila, M., Klein, C. & Rascol, O. LRRK2 in Parkinson disease: challenges of clinical trials. *Nat. Rev. Neurol.* **16**, 97–107 (2020).
- Usmani, A., Shavarebi, F. & Hiniker, A. The cell biology of LRRK2 in Parkinson's disease. *Mol. Cell. Biol.* **41**, e00660-20 (2021).
- Briz, V. et al. The non-coding RNA BC1 regulates experience-dependent structural plasticity and learning. *Nat. Commun.* **8**, 293 (2017).
- Bu, M. et al. Inhibition of LRRK2 kinase activity rescues deficits in striatal dopamine physiology in VPS35 p.D620N knock-in mice. *NPJ Parkinson's. Dis.* **9**, 167 (2023).
- Jaudon, F. et al. The RhoGEF DOCK10 is essential for dendritic spine morphogenesis. *Mol. Biol. Cell* **26**, 2112–2127 (2015).
- Morató, X. et al. The Parkinson's disease-associated GPR37 receptor interacts with striatal adenosine A2A receptor controlling its cell surface expression and function in vivo. *Sci. Rep.* **7**, 9452 (2017).
- Kuzniewska, B. et al. SRF is required for sprouting and maintenance of neuronal connections in the hippocampus. *J. Neurosci.* **36**, 5699–5710 (2016).
- Ma'ayan Lab. (n.d.). SRF gene set—ENCODE Transcription Factor Targets. Harmonizome. Retrieved 2024 from <https://maayanlab>.

cloud/Harmonizome/gene_set/SRF/ENCODE+Transcription+Factor+Targets

34. Ba, W. et al. ARHGAP12 functions as a developmental brake on excitatory synapse function. *Cell Rep.* **14**, 1355–1368 (2016).
35. Zeng, S. X., Dai, M. S., Keller, D. M. & Lu, H. SSRP1 functions as a co-activator of the transcriptional activator p63. *EMBO J.* **21**, 5487–5497 (2002).
36. Mitchell, A. C., Jiang, Y., Peter, C. & Akbarian, S. Transcriptional regulation of GAD1 GABA synthesis gene in the prefrontal cortex of subjects with schizophrenia. *Schizophrenia Res.* **167**, 28–34 (2015).
37. Dimitrov, D. et al. Comparison of methods and resources for cell-cell communication inference from single-cell RNA-Seq data. *Nat. Commun.* **13**, 3224 (2022).
38. Chen, K. et al. LRP1 is a neuronal receptor for α -synuclein uptake and spread. *Mol. neurodegeneration* **17**, 57 (2022).
39. Ludtmann, M. H. R. et al. LRRK2 deficiency induced mitochondrial Ca^{2+} efflux inhibition can be rescued by $\text{Na}^{+}/\text{Ca}^{2+}/\text{Li}^{+}$ exchanger upregulation. *Cell Death Dis.* **10**, 265 (2019).
40. Binan, L. et al. Simultaneous CRISPR screening and spatial transcriptomics reveal intracellular, intercellular, and functional transcriptional circuits. *Cell* **188**, 2141–2158.e18 (2025).
41. Wroblewska, A. et al. Protein barcodes enable high-dimensional single-cell CRISPR screens. *Cell* **175**, 1141–1155.e16 (2018).
42. Baysoy, A. et al. Spatially resolved in vivo CRISPR screen sequencing via perturb-DBiT. *bioRxiv* <https://doi.org/10.1101/2024.11.18.624106> (2024).
43. Saunders, R. A. et al. Perturb-Multimodal: A platform for pooled genetic screens with imaging and sequencing in intact mammalian tissue. *Cell* **188**, 4790–4809.e22 (2025).
44. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science (New York, N.Y.)*, **348**, aaa6090 (2015).
45. He, S. et al. High-plex imaging of RNA and proteins at subcellular resolution in fixed tissue by spatial molecular imaging. *Nat. Biotechnol.* **40**, 1794–1806 (2022).
46. Bramlett, C. et al. Clonal tracking using embedded viral barcoding and high-throughput sequencing. *Nat. Protoc.* **15**, 1436–1458 (2020).
47. Kong, W. et al. CellTagging: combinatorial indexing to simultaneously map lineage and identity at single-cell resolution. *Nat. Protoc.* **15**, 750–772 (2020).
48. Jindal, K. et al. Single-cell lineage capture across genomic modalities with CellTag-multi reveals fate-specific gene regulatory changes. *Nat. Biotechnol.* **42**, 946–959 (2024).
49. Vieth, B., Ziegenhain, C., Parekh, S., Enard, W. & Hellmann, I. pow-simR: power analysis for bulk and single cell RNA-seq experiments. *Bioinformatics* **33**, 3486–3488 (2017).
50. MacArthur, J. et al. The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res.* **45**, D896–D901 (2017).
51. Chew, W. L. et al. A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat. Methods* **13**, 868–874 (2016).
52. Goh, J. J. L. et al. Highly specific multiplexed RNA imaging in tissues with split-FISH. *Nat. Methods* **17**, 689–693 (2020).
53. Keng, C. T. et al. Multiplex viral tropism assay in complex cell populations with single-cell resolution. *Gene Ther.* **29**, 555–565 (2022).
54. Shen, K. spatialperturbseq (Version 1.0.0) [Computer Software]. Zenodo <https://doi.org/10.5281/zenodo.17959756> (2025).

Acknowledgements

The authors thank Nikita Gupta, Vipul Singhal, Nigel Chou, Grace Yeo, Timothy Stuart, Adaikalavan Ramasamy and Chia Minghao for assistance and insights on bioinformatics analysis; Torsten Wustefeld for provision of Cas9 mice; Fu Yu and Christine Chiam for use of stereotaxic apparatus; Bing Shao Chia and Joshua James for assistance with library cloning; Maurice Lee and Norbert Ha for assistance with initial FISH experiments; Liew Jun Xian for cloud platform setup; Caleigh Tan for reviewing the manuscript; A*STAR's Immunology Network (SigN) Flow Cytometry platform for help with FACS experiments. A*STAR's SigN Flow Cytometry platform is supported by SigN, A*STAR, and the National Research Foundation (NRF), Immunomonitoring Service Platform (Ref: ISP: NRF2017_SISFP09) grant. This work is supported by A*STAR Core Funding, A*STAR Central Research Fund UIBR SC18/21-1089UI, National Medical Research Council (NMRC) Open Fund—Individual Research Grant (OF-IRG) OFIRG24jul-0096, and Open Fund—Young Individual Research Grant (OF-YIRG) OFYIRG23jul-0050. This project is supported by the Singapore Ministry of Health's National Medical Research Council through the Programme for Research in Epidemic Preparedness and Response (PREPARE), under its PREPARE Vaccines & Therapeutics Co-op Open Grant (PREPARE-OC-VT-2024-008). K.H.C. is supported by the National Medical Research Council of Singapore grant OFIRG20nov-0056 and the National Research Foundation grant NRF-CRP25-2020-0001.

Author contributions

K.S., C.T.K. and W.L.C. conceived and conceptualized the technology framework. K.S., C.T.K., W.Y.S., K.H.C. and W.L.C. designed the experiments. K.S. did stereotaxic injections, tissue collection and processing. W.Y.S. performed the FISH experiments. M.G.L.L. processed slides for Xenium. S.P. provided intellectual contributions to Xenium experiments. D.L.S., C.T.K. and K.S. performed library cloning. D.L.S. and A.M. produced the AAVs. B.W. and C.T.K. prepared libraries for long-read sequencing of the 3xU6-sgRNA cassettes. M.I.B.H. and K.S. wrote the script for the generation of the barcodes. K.G. bred and maintained Cas9 mice. K.S. performed bioinformatics analysis. K.S., W.Y.S. and W.L.C. wrote the manuscript with contributions from the rest of the authors.

Competing interests

The authors (K.S., C.T.K., M.I.B.H., W.L.C., W.Y.S.) are listed as inventors on a patent application related to all of this work (Patent Cooperation Treaty Patent Application No. PCT/SG2025/050446). Applicant: A*STAR. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-026-69677-6>.

Correspondence and requests for materials should be addressed to Wei Leong Chew.

Peer review information *Nature Communications* thanks Qi Liu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026