Nova-ST: Nano-Patterned Ultra-Dense platform for spatial transcriptomics

Suresh Poovathingal[1,2,3,*,#], Kristofer Davie[1,2,*], Roel Vandepoel[1,2,4], Nicholas Poulvellarie[2], Annelien Verfaillie[5], Nikky Corthout[6] and Stein Aerts[1,2,4,#]

[1] VIB Center for AI & Computational Biology (VIB.AI), Leuven, Belgium.
[2] VIB-KU Leuven Center for Brain & Disease Research, Leuven, Belgium.
[3] Aligning Science Across Parkinson's (ASAP) Collaborative Research Network, Chevy Chase, United States.
[4] Department of Human Genetics, KU Leuven, Leuven, Belgium.
[5] Genomics Core, KU Leuven, Leuven, Belgium
[6] Bio Imaging Core, VIB-KU Leuven Center for Brain & Disease Research, Leuven, Belgium.
[*] Co-first authors
[#] Correspondence to suresh.poovathingal@kuleuven.be and stein.aerts@kuleuven.be

Abstract:

Spatial transcriptomics workflows using barcoded capture arrays are commonly used for resolving gene expression in tissues. However, existing techniques are either limited by capture array density or are cost prohibitive for large scale atlasing. We present Nova-ST, a dense nano-patterned spatial transcriptomics technique derived from randomly barcoded Illumina sequencing flow cells. Nova-ST enables customized, low cost, flexible, and high-resolution spatial profiling of large tissue sections. Benchmarking on mouse brain sections demonstrates significantly higher sensitivity compared to existing methods, at reduced cost.

Characterizing and modeling complex tissues requires an understanding of the spatial cellular organization and of the interactions between cells, in the context of normal and pathological states[1,2]. Multiplexed *in situ* imaging-based assays measure RNA expression at sub-cellular resolution, but their implementation is non-trivial and requires pre-selected gene sets for measurements[3]. On the other hand, spatial barcoding methods enable unbiased whole transcriptome analysis and are routinely used for spatial transcriptomics profiling[4]. Currently however, most of these techniques have a low resolution. In addition, commercial spatial assays have a high cost and therefore offer limited financial feasibility for large scale tissue atlasing efforts.

Recently, several methods based on nano-patterned arrays have emerged that can potentially provide whole transcriptome capture at subcellular resolution [5–7]. These also allow fine tuning of the size of the capture array to achieve near single cell resolution. Stereo-seq from STOmics uses randomly barcoded DNA nanoballs captured on a nanopatterned array[5]. Alternatively, Seq-Scope, repurposes the Illumina MiSeq flow cell to perform spatial RNA sequencing[7]. In the latter method, spatial barcoding is achieved using local bridge amplification of DNA libraries containing random spatial barcodes[7]. Here, the functionalized area of the flow cell where the tissue is overlayed for spatial sequencing is limited (< 2mm of imaging area). In addition, MiSeq flow cells have surface functionalization in contiguous circular patterns, further decreasing the effective spatial footprint (Fig. 1a). Thus, for most practical purposes related to large scale, high throughput spatial sequencing of tissue required for cell-atlasing efforts, Seq-Scope offers limited possibilities.

To develop a cost effective and easy to use spatial transcriptomics platform, we adapted the Seq-Scope workflow to perform spatial sequencing on large tissue sections using high density Illumina patterned flow cells. After assessing different flow cells, we chose NovaSeq 6000 chips for high-density spatial profiling of large tissue sections and developed an experimental and computational workflow on this platform called Nova-ST. Guidance for performing the Nova-ST experimental workflow will become available at protocols.io and https://nova-st.aertslab.org/, while the computational pipeline is available at https://github.com/aertslab/Nova-ST

NovaSeq flow cells are patterned with oligonucleotide-functionalized nano-wells arranged in a hexagonal pattern (Fig. 1b). Compared to the MiSeq, this layout has benefits, including higher spatial density and lower cluster crosstalk. We imaged a NovaSeq flow cell with electron

microscope, which revealed nano-wells of ~300 nm diameter and a well-to-well pitch of ~625 nm (Fig. 1b), this provides a resolution of ~350 spots for RNA capture per 100 μm^2 area. Similar to Seq-Scope[7], our Nova-ST workflow starts with sequencing of HDMI oligonucleotides. HDMI molecules contain sequencing adapters, a spatial identifier sequence and a DraI restriction enzyme cleavable RNA capture site (Fig. 1c). The spatial coordinates on the flow cells are created during the first sequencing pass with HDMI molecules (Fig. 1c, Supplementary Fig. 1, details in Methods section). This resulted in 78.57% of the nanowells receiving a valid HDMI spatial barcode. The resulting data from the HDMI sequencing is processed, whereby HDMIs are extracted alongside their intra-tile coordinates and stored in binary files for downstream usage (further details in Methods). The preliminary HDMI sequencing quality, including the base composition is assessed prior to downstream steps (Supplementary Fig. 2)

Next, the flow cell is prepared for spatial transcriptomics capture (Fig. 1d). The flow cell consists of three sandwiched layers: an upper thin glass layer, a gasket layer providing clearance for the fluidic flow and a lower thick glass layer. Both the thin and thick glass layers have functionalized HDMI surfaces. (Fig. 1d). We first separated the thin and thick glass layers by gently prying, and then scribed the separated glass layers with a portable CNC diamond tipped scribing tool (Fig. 1d, details in Methods), along a rectangular grid of 1cm x 1cm (Fig. 1d). Successful disassembly would yield 48 thick and 48 thin Nova-ST chips (Fig. 1d). Next, the Nova-ST chips are prepared for RNA capture after a series of enzymatic and chemical treatments steps (details in Methods, Fig. 1d, Supplementary Fig. 1). Each Nova-ST chip provides an area equivalent to 10 mm x 8 mm for spatial profiling (Fig. 1d).

To test Nova-ST, we performed spatial transcriptomics experiments with 10 μm fresh-frozen mouse brain coronal sections (Fig. 1d). We performed five biological replicate experiments. One replicate was chosen for deep sequencing (~1.2 billion reads). The Nova-ST analysis pipeline was run on these samples (see Methods, Supplementary Fig. 3). Briefly, tissue covered tiles are identified using a subset of sequencing reads and a chip-based barcode whitelist is created. All sequencing reads are then demultiplexed by their spatial (HDMI) barcode (compared and corrected with the whitelist), while simultaneously being mapped to the genome and quantified using STARsolo[8], whereby coordinates are corrected to put them into a global chip-based space. The HDMI footprint of the tissue indicates high spatial resolution (Fig. 1e and the inset showing the zoomed in area). Next, data is binned into contiguous chunks of 25x25 μm^2, 50x50 μm^2 and

100x100 μm^2, and pre-processing was performed (see Methods). The Nova-ST pipeline results in .loom files that can be visualized in the SCope viewer[9]. Fig.1f-h shows the spatial expression of several marker genes, each localized in the correct brain region, such as Mef2 in the cortex; Hpc and Camk2b in the hippocampus; and Ccn2 in Layer 6b.

Next, we assessed the reproducibility of Nova-ST by comparing quality metrics across the five replicates (four at lower sequencing depth [76-111 million reads]) (Fig. 2a, Supplementary Fig. 4). The percentage of reads with valid barcodes that are also successfully mapped, range between 81.3% and 85.1% and the percentage of unique UMIs ranges between 36.48% and 78.60% (depending upon sequencing depth). For three different bin sizes (25x25 μm^2 – bin 50, 50x50 μm^2 – bin 100, and 100x100 μm^2 – bin 200), we detected a median of 994, 2821, and 6318 genes with non-zero counts; and 2131, 8268 and 32317 UMIs per bin for the deeply sequenced sample; for the shallowly sequenced samples we detected a median of 251, 878, and 2503 genes, with a median of 374, 1530 and 5920 UMIs (Fig. 2b and Supplementary Fig. 4). Comparing total counts per gene across the filtered datasets at bin 50 shows high reproducibility (Supplementary Fig. 5, Supplementary Table 1).

Next, we assessed RNA diffusion during tissue permeabilization. For this, we localized mitochondrial mRNAs and nuclear RNA (MALAT1) and found these transcripts to be largely non-overlapping, suggesting low levels of RNA diffusion (Fig. 2c). Unbiased clustering of bins identifies individual layers in the cortex as well as subtypes within the hippocampus (Fig. 2d). This resolution allows us to identify clusters of cells that span only 1-2 cell layers, such as layer 6b that forms a separate cluster (Fig. 2d), in agreement with Ccn2 (a layer 6b marker gene) expression shown in Fig. 1h and Fig. 2c.

To test whether the unbiased clustering corresponds to known cell types, we used *cell2location*[10] to map an independent scRNA-seq data set of the mouse cortex and hippocampus to the binned Nova-ST data[11]. This correctly identifies the location of all scRNA-seq cell types within the cortex and hippocampus, and cell types overall correspond with the unbiased clustering (e.g., Layer 2/3, Layer 4/5, and Layer 6 intra-telencephalic (IT) neurons in Fig. 2d).

Finally, we compared the performance (% of usable reads), the sensitivity (UMI and gene counts), and the specificity (correct localization of scRNA-seq cell types by *cell2location*) of Nova-ST with Stereo-seq. To this end, we performed a Stereo-seq experiment on a comparable mouse brain

section, and down-sampled the Stereo-seq data to the same depth as the deeply sequenced Nova-ST run. Nova-ST libraries have similar metrics to Stereo-seq libraries with regards to mapping % and annotated reads (83.51% and 73.04% [Nova-ST] vs 77.9% and 62.55% [Stereo-seq]) (Fig. 2a, e). Nova-ST shows a lower percentage of recovered barcodes, due to non-imaged sections from the NovaSeq, but has superior complexity allowing more genes/UMIs to be detected at lower sequencing depths and greater information obtained in deeper sequencing (25.54% [Nova-ST] versus 9.85% [Stereo-seq] unique UMIs at equal depth) (Fig. 2f). At a bin size of 200, Nova-ST detects a median of 6318 genes compared to 4092 genes with the Stereo-seq platform at the same bin size. Open-ST[12], another spatial sequencing technique developed independently by another group also demonstrates superior sensitivity compared to Stereo-seq. For Nova-ST, we detect a median UMI count of 294 per 100 μm^2 at a sequencing depth of ~15 million reads per mm^2 of total sequenced surface (deeply sequenced sample), which compares to the sensitivity reported by the Open-ST platform[12]. Thus, Nova-ST obtains more genes and UMIs than Stereo-seq at similar bin sizes, allowing us to decrease bin sizes in Nova-ST to achieve increased resolution (Fig. 2b,f). At similar bin sizes (Bin 50), Nova-ST better delineates known spatial clusters within the cortex and hippocampus than Stereo-seq (Fig. 2d,g) as well as obtains better predictions from *cell2location* (Fig. 2d,h). Together, this comparison shows that Nova-ST outperforms the nearest competitive technology on the same tissue, providing greater usable resolution, while allowing for further customization such as cell boundary staining and development for use in profiling other modalities such as open chromatin or protein expression.

References:

1. Vickovic, S. *et al.* High-definition spatial transcriptomics for in situ tissue profiling. *Nat. Methods* **16**, 987–990 (2019).
2. Lein, E., Borm, L. E. & Linnarsson, S. The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing. *Science* **358**, 64–69 (2017).
3. Rao, A., Barkley, D., França, G. S. & Yanai, I. Exploring tissue architecture using spatial transcriptomics. *Nature* **596**, 211–220 (2021).
4. Tian, L., Chen, F. & Macosko, E. Z. The expanding vistas of spatial transcriptomics. *Nat. Biotechnol.* **41**, 773–782 (2023).
5. Chen, A. *et al.* Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell* **185**, 1777-1792.e21 (2022).
6. Fu, X. *et al.* Polony gels enable amplifiable DNA stamping and spatial transcriptomics of chronic pain. *Cell* **185**, 4621-4633.e17 (2022).
7. Cho, C.-S. *et al.* Microscopic examination of spatial transcriptome using Seq-Scope. *Cell* **184**, 3559-3572.e22 (2021).
8. Kaminow, B., Yunusov, D. & Dobin, A. STARsolo: accurate, fast and versatile mapping/quantification of single-cell and single-nucleus RNA-seq data. 2021.05.05.442755 Preprint at https://doi.org/10.1101/2021.05.05.442755 (2021).
9. Davie, K. *et al.* A Single-Cell Transcriptome Atlas of the Aging Drosophila Brain. *Cell* **174**, 982-998.e20 (2018).
10. Kleshchevnikov, V. *et al.* Cell2location maps fine-grained cell types in spatial transcriptomics. *Nat. Biotechnol.* **40**, 661–671 (2022).

11. Yao, Z. *et al.* A taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell* **184**, 3222-3241.e26 (2021).

12. Schott, M. *et al.* Open-ST: High-resolution spatial transcriptomics in 3D. 2023.12.22.572554 Preprint at https://doi.org/10.1101/2023.12.22.572554 (2023).

Data availability:

Raw mouse brain data has been deposited to NCBI's GEO archive:
Nova-ST: GSE256318
Stereo-seq: GSE256319
Accessory data and information on the Nova-ST workflow can be found at:
https://nova-st.aertslab.org/

Code availability:

All code has been deposited on GitHub at: https://github.com/aertslab/Nova-ST

Acknowledgements:

We thank UZ Leuven Genomics Core team for supporting with the sequencing workflow and gifting spent sequencing flow cells for downstream processing and flowcell cutting and processing practice. We would also like to thank Dr. Pieter Baatsen from the Electron Micronscopy platform at VIB KU Leuven Center for Brain and Disease Research for his help with the surface scanning of the NovaSeq chip with Scanning Electron Microscope. This work is funded by the following grants to S.A.: ERC Advanced Grant (101054387_Genome2Cells); Michael J. Fox Foundation for Parkinson's Research (Michael J. Fox Foundation) (ASAP-000430).

Author contributions

S.P., K.D. & S.A. conceived and developed the Nova-ST workflow; S.P. & K.D. performed the experiments and data analysis with help from R.V., N.P., A.V. & N.C.; S.P, K.D. & S.A. wrote the paper with inputs from all the other authors; S.P. & S.A. provided supervision and guidance for the project.

Competing interest

We have no competing interest.

Figures:

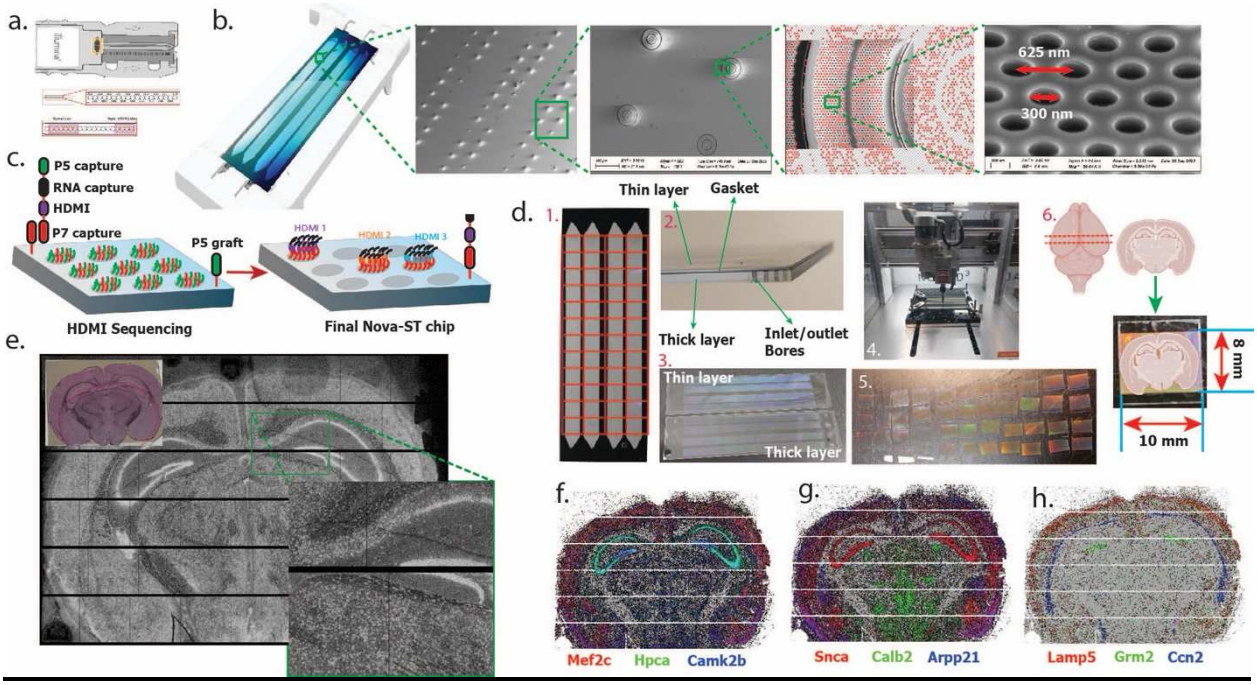

Figure 1: Set up and characterization of Nova-ST platform. a. Details of MiSeq flow cell used in the Seq-Scope platform[7]. Expanded view of the MiSeq's flow channel indicating the functional area for HDMI capture. **b.** Details of NovaSeq 6000 sequencing flow cell. Electron microscopy details of the nanowell pattern on the functional surface of the sequencing flow cell. **c.** Zoomed in illustration of the nano wells showing the HDMI sequencing on the NovaSeq flow cell and the finalized Nova-ST chip after all the enzymatic and chemical processing resulting in single stranded DNA with HDMI sequences and RNA capture domains. **d.** Post processing details after the HDMI sequencing: 1. 1 cm x 1 cm grid along which the flow cells are cut to prepare the Nova-ST chips. 2. Details of the NovaSeq glass chip showing the thin and think glass layers sandwiched by a flow gasket. 3. Separated layers of NovaSeq 6000 flow cell exposing the functional surface of the sequencing flow cell. 4. Scribing of the Nova-ST glass layers using the NOMAD 3 CNC milling machine. 5. Final cut 1cm x 1 cm Nova-ST chips. 6. Illustration of the mouse brain sectioning and positioning on the Nova-ST chip. **e.** Spatial HDMI footprint of the transcripts captured from the mouse brain section in the deeply sequenced (DS) sample. **f-h.** Spatial gene expression mapping of several region specific gene markers visualized with SCope[9].

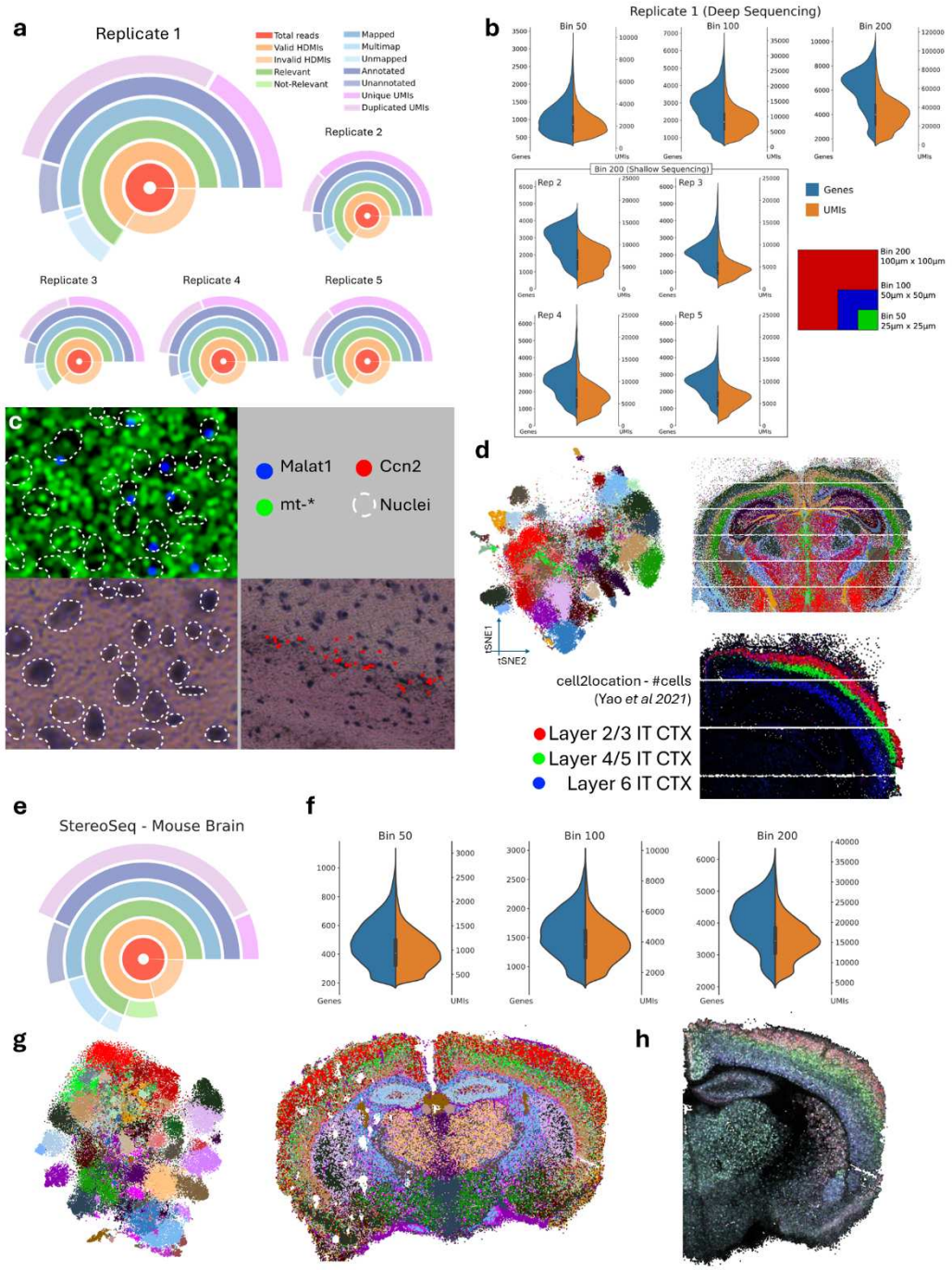

Figure 2: Performance and benchmarking of Nova-ST platform using mouse brain section a. Circle plots showing various metrics from the Nova-ST libraries, including mapping percentage, number of annotated reads and sequencing saturation. **b.** Violin plots of both number of genes detected and number of UMIs detected at bin 50, 100 and 200 for deep sequencing (DS) and bin 200 for four replicates of shallow sequenced libraries. **c.** Two areas of tissues showing i) Malat1 expression localized to cell nuclei, distinct from mitochondrial reads and ii) Ccn2 expression limited to cortical layer 6b. **d.** t-SNE and spatial visualizations for Nova-ST DS showing Leiden clustering. Cell2location predictions for different layers in the cortex are displayed in the cutout.

e. Circle plot showing various metrics from the Stereo-seq library. **f.** Violin plots of both number of genes detected and number of UMIs detected at bin 50, 100 and 200 for the Stereo-seq sample **g.** t-SNE and spatial visualizations showing Leiden clustering of the Stereo-seq sample. **h.** Cell2location predictions for cortex layers (key shown in **d.**) in the Stereo-seq sample.

Methods:

1. Mouse tissue preparation for cryo-sectioning

All animal experiments were performed according to the KU Leuven ethical guidelines and approved by the KU Leuven Ethical Committee for Animal Experimentation (approved protocol no. ECD P007/2021). Six to eight weeks old mice (C57BL/6J) used in the study were kept on a 14 h light, 10 h dark–light cycle from 7:00 to 21:00. Brains were dissected and immediately snap-frozen in isopentane for 10 min. Afterwards, brains were embedded in Tissue-Tek OCT cryo embedding compound. Coronal cryosections (10 µm) were performed at CT=14°C, OT=11°C. The used area of mouse cryosections is in the somatosensory areas adjacent to the posterior parietal association areas. For each brain tissue, 5-10 OCT scrolls of 70 um section thickness were collected into DNA lo-bind 2 ml eppendorf tubes. Ice cold PBS was used for washing the tissue to remove the OCT matrix. Total RNA was extracted from the washed tissue sections using the innuPREP mini RNA kit (Analytik Jena; Cat. No. AJ 845-KS-2040250). The RNA quality was assessed using RNA Nano kit (Agilent). Only tissue with RIN > 7 was used for the spatial analysis.

2. Stereo-seq optimization of tissue permeabilization

Tissue optimization was performed using the Stereo-seq Permeabilization kit (Cat. No. 111KP118) and Stereo-seq chip set P (Cat. No. 110CP118) according to the manufacturer's protocol (Stereo-seq permeabilization set user manual, Ver A1). Briefly, 4 permeabilization chips were removed from the storage buffer and washed with nuclease free water and dried at 37°C. Next, 4 consecutive 10 µm tissue sections were prepared from the tissue cryo-block and placed on the permeabilization chip, the tissue layer was thawed to attach it to the surface of the chip. After drying the tissue on a 37°C hot plate, the chip was then dipped into 100% methanol at -20°C and incubated for 30 mins to fix the tissue. Post fixation, the tissue permeabilization test was performed on these chips by permeabilizing the tissue with PR enzyme prepared in 0.01N HCl (pH 2.0), at 4 different time points ranging from 6 mins to 30 mins. After the permeabilization, the chips were rinsed with 0.1X SSC buffer before reverse transcription. Reverse transcription was carried out at 42°C for 1 hour in dark. Tissue removal was performed at 55°C for 1 hour using the TR enzyme to clear the tissue before imaging. Fluorescence imaging was performed in the TRITC channel with 10X objective, following the imaging guidelines provided by the manufacturer (Guidebook for Image QC &

microscope assessment and imaging, Ver A5). The optimal permeabilization time was assessed based on the strongest fluorescence signal with the lowest signal diffusion (crispness of the RNA footprint). Based on our assessment, we found the most optimal permeabilization time for the mouse brain to be 12 mins.

3. Stereo-seq spatial transcriptomics analysis

The spatial transcriptomics analysis was performed using the Stereo-seq Transcriptomics kit (Cat. No. 111ST114) according to the manufacturer's protocol (Stereo-seq Transcriptomics set user manual, Ver A2). Briefly, as with the permeabilization analysis, the T-chip was removed from the storage buffer and washed with nuclease free water and dried at 37°C. Next, a 10 µm tissue section from a desired region of interest was prepared from the tissue cryo-block and placed on the T-chip and thawed the tissue layer to attach to the surface of the chip. After drying the tissue on 37°C hot plate, the chip was then dipped into 100% methanol at -20°C and incubated for 30 mins to fix the tissue. The fixed tissue was then stained using the Qbit ssDNA reagent (Thermo Cat. No. Q10212). Fluorescence imaging of the single stranded DNA staining was performed in the FITC channel with a 10X objective, following the imaging guidelines provided by the manufacturer (Guidebook for Image QC & microscope assessment and imaging, Ver A5). Prior to permeabilization, the ssDNA-stained image was also subjected to QC analysis using the imageQC software as per manufacturer's recommendations. As with the permeabilization protocol, the tissue permeabilization was carried out with PR enzyme prepared in 0.01N HCl (pH 2.0) at 37°C. The optimal permeabilization time estimated from the tissue permeabilization analysis was used for the transcriptomics analysis. After washing the chip, reverse transcription mix was added to the chip and incubated at 42°C for at least 3 hrs. Tissue removal from the stereo seq chip was achieved by incubating the chip in the TR buffer at 55°c for 10 minutes. cDNA release and collection was performed by incubating the chip in cDNA release mix overnight at 55°C and the released cDNA was purified with Ampure XP beads (Beckman Coulter; Cat. No. A63882) using the manufacturer's recommendation. After quality assessment using a bioanalyzer (Agilent), sequencing library preparation was performed using transposase assisted tagmentation reaction. Indexed PCR and library purification was performed to prepare the final sequencing library as per manufacturer's recommendations. Final Stereo-seq libraries were sequenced on MGI/BGI sequencing platforms and were sequenced at the MGI Latvia sequencing facility.

4. HDMI sequencing

The HDMI generation in this work was done using the Illumina NovaSeq 6000 S4 kit; 35 cycles (PN: 20044417). In accordance with the original publication[7], HDMI32DraI-32 ultramer (IDT technologies – standard desalted purification) was used for first level sequencing on the NovaSeq S4 flow cell to generate the HDMI array. The HDMI32DraI-32 ultramer (details of all oligonucleotide sequences used in this work is provided in Supplementary Table 2) was diluted to 1 uM concentration and the actual concentration of the oligonucleotide was titrated using qPCR to estimate the final concentration to be loaded for sequencing. Briefly, we used the Kapa Library Quantification kit (Roche, KK4824) to quantify the Oligonucleotides. Based on the concentration estimated from qPCR, libraries were denatured and loaded at a final concentration of 300 pM on the NovaSeq 6000 following the manufacturer's instructions. Custom read primer Read1-DraI was also ordered from IDT technologies with PAGE purification[7]. The read primer was diluted to 0.3µM with HT1 buffer and loaded into the custom read primer 1 position in the NovaSeq reagent cartridge. The sequencing configuration used for reading the HDMI barcodes was 37(R1)-0(I1)-0(I2)-0(R2). At the end of the 32nd cycle, the instrument was manually aborted without initiating a post run wash. The S4 flow cell was then retrieved for immediate downstream postprocessing, it can also be stored safely at 4°C for at least 2 weeks. Users not having direct access to NovaSeq 6000 instrument can instruct sequencing facility to perform the HDMI sequencing step and transport the sequenced flow cell at 4°C. Prior to shipping, the inlet and outlet ports of the flow cell should be sealed using a PDMS biopsy plugs (see the details below) to ensure the liquids in the flow channels does not dry out.

5. Post sequencing processing

Sealing of the inlet and outlet for the flow channels was achieved by plugging them with 1.5-2mm Polydimethylsiloxane (PDMS) cylinders. To prepare these cylinders, the monomer and catalyst of SYLGARD™ 184 Silicone Elastomer Kit (Dow chemicals) was prepared in a 10:1 weight ratio. The components were mixed thoroughly and vacuum degassed. To polymerize, the mix was poured into a 3 cm petri dish and incubated at 80°C for 2 hours to complete the polymerization process. The PDMS slab was then extracted from the petri dish and wrapped into aluminum foil. 2 mm PDMS cylinders were prepared from this slab using a 2 mm biopsy punch (World Precision Instruments; Cat No. 504531).

The HDMI flow cells were then subjected to downstream enzymatic treatment. Firstly, the flow channels were cleaned with 200 µl of nuclease free water, NFW (Thermo Fisher; Cat No. 10977035). This was repeated for a total of three times. During each wash step, after filling the channels with reagents, a vacuum source (general vacuum pump, e.g. VWR, Cat. No. SART16694-1-60-06) was used to completely remove the reagents from the channel. Aspiration was continued until the channels became completely dry. To expose the RNA capture handle, the double stranded DNA was cut using restriction endonuclease DraI (NEB Inc. Cat. No. R0129L). The flow channels were first cleaned with 200 µl 1X rCutSmart™ Buffer. Then all flow channels were filled with 200 µl DraI reaction mix (1X rCutSmart™ Buffer, 2U/ul DraI Enzyme). Excess liquid overflowing from the outlet ports was aspirated using the vacuum pump. After making sure there were no air pockets trapped in the flow channels, the 2mm PDMS cylinder blocks were forced into the inlet and outlet ports to seal them using thin tipped forceps. The flow cell assembly was then placed into a humidification chamber (Nunc Square BioAssay Dishes; Thermo Fisher – Cat. No. 240835). For humidification, the flow cell was placed along with wet paper tissue and the petri dish was sealed using Parafilm. The flow cell was then incubated overnight at 37°C for the completion of the endonuclease reaction.

The flow channels were washed three time with 200 µl of NFW. After the vacuum aspiration of water from the flow channels, the channels were filled with 200 µl of 1X Exonuclease buffer (NEB Inc. M0293L). After the removal of the exonuclease buffer, the flow channels were filled with 200 µl of Exonuclease reaction mix (1X Exonuclease reaction buffer, 1U/µl of Exonuclease enzyme and 0.14 U/µl of Quick Calf Intestinal Phosphatase (NEB Inc. M0525L)). After aspirating the excess reaction mix from the inlet/outlet of the sequencing flow cells, the ports were sealed with fresh 2 mm PDMS cylinders. The reaction mix was then incubated at 37°C for 45 mins in the same humification chamber. After the completion of the exonuclease reaction, the HDMI flow cell assembly was retrieved, and the channels were washed three times with 200 µl of NFW. After each wash the liquid was completely withdrawn from the flow channels using a vacuum pump.

6. **Flow cell disassembly for Nova-ST chip preparation**

The flow cell assembly was next placed into an oven at 50°C for 20 mins to dry the flow channels. The orientation of the top and bottom glass layers was identified with respect to the inlet and outlet ports of the sequencing flow cells. This is required for identifying the spatial location of the Nova-

ST chips during the spatial transcriptomics analysis. The glass chip was then removed from the plastic housing by manually pulling out the plastic brackets that clamp down the inlet and outlet ports, releasing the glass chip. Then, using a fine scalpel, the gasket layer between the thin and thick glass surfaces (Fig. 1d) was scored gently to separate the thin and thick glass layer. This scoring must be done carefully without damaging the functional surface of the HDMI array. After gently prying and separating the glass layers, paper masking tape was glued to the back of the glass layers (3M, Cat No. 3M 201E 48MM), excess tape was trimmed off. In this work we have used the NOMAD 3 CNC milling machine from Carbide3D to score the glass surface into a 1cm x 1cm cutting grid (Fig. 1d). This CNC milling machine comes with a 130W spindle and has a working area of 200x200mm and 76mm in height. Less powerful machines can be used for the purpose described above. A diamond drag bit with a 90-degree tip, from the CNC milling machine manufacturer was used. The tip angle keeps the scoring as narrow as possible and ensures better penetration into the glass compared to a 120-degree tip. The bit is also equipped with a spring inside to adjust the force. In this case, it was adjusted so that little pressure is applied to the glass, while the tool length measurement probe of the machine can still detect the bit.

The scoring pattern was created using the machine suppliers dedicated software, Carbide Create. The glass plates dimensions were defined in the software as well as the desired scoring pattern. The scoring depths was then adapted according to the glass thickness and the direction. The thicker glass plate (1.2mm thick) was scored with a 0.6mm depth in the width direction (shortest side) and 0.2mm depth in the length direction. For the thin section (thickness 0.3mm), depths of 0.4mm and 0.1mm were used. It is important to know that the depth of cut defined in the software is not the actual depth. The actual depth differs due to the spring that retracts at the glass contact. This explains the higher depth of cut defined compared to the glass thickness. Each score was performed with a single pass of the tool.

Once the pattern was defined, the machine code (Gcode) was sent to the CNC machine via Carbide Motion, another software supplied by the manufacturer and was used to control machine movements. The plates were clamped to the table for scribing. After scoring the glass layers, the cutting was performed by using the glass running pliers (SPEEDWOX). To reduce the damage caused by the pliers on the functional surface, rubber tips were used. Before using the pliers, the pliers and the rubber tips were wiped with RNA & DNAZap followed by cleaning with 100% ethanol to ensure they were free of contamination. For cutting and preparing the Nova-ST chips,

the pliers were used to cut the vertical score lines, by applying gentle pressure in the middle of the glass layers along the score line. After breaking the chips along the vertical score line, the masking tape was cut using a scalpel or razor. This was followed by breaking the glass chips along the horizontal score lines to produce the 1cm x 1cm Nova-ST chips. The Nova-ST chips were then pried out of the masking tape using forceps and the chips placed into 24 well plates. The location of the chip was recorded on the wells. This was repeated across the whole flow cell to produce 96 1cm x 1cm Nova-ST chips from both the thin and thick layers of the NovaSeq chip. In our experience breaking and preparation of the Nova-ST chips from the thick glass layer is non-trivial and sometimes getting perfect cut along the score lines is difficult. It's highly recommended to practice the cutting on trial flow cells before attempting on the HDMI flow cells. The Nova-ST chips in the 24 well plates were then subjected to following steps to remove the second strand and to store them for long term. The Nova-ST chips were washed 3X times with 0.1N NaOH. For each wash the chips were incubated in the caustic solution for 5 mins, to ensure efficient denaturation of the second strand. Each of the Nova-ST chips were then washed 3X times with 1 ml nuclease free water followed by 2 times wash with 1 ml of 1X TE buffer (IDTE solution; IDT, Cat. No. 11-05-01-09). After the last wash, the chips were stored in IDT 1X TE buffer for long term storage at 4°C.

7. RNA quality assessment

Prior to the tissue optimization and spatial transcriptome analysis on the Nova-ST chips, every tissue analyzed in this work was subject to RNA quality assessment. In brief, 5-10 serial sections of 50-70 um thickness were cryo-sectioned from a region farther away from the region of interest. These serial sectioned OCT scrolls were put into a 2 ml lo-bind tube (Eppendorf Cat. No. 0030108078) and stored at -80°C. The tissue scrolls were washed with 1 ml of ice-cold PBS at 4°C. Total RNA extraction from the spun-out tissue was performed using innuPREP mini RNA kit (Analytik Jen; Cat. No. AJ 845-KS-2040250). Manufacturer recommendations were followed to extract total RNA. Elution was performed in 30 µl of NFW. The quality of the total RNA was assessed using Pico RNA kit (Agilent)

8. Optimization of tissue permeabilization for Nova-ST workflow

Optimal tissue permeabilization for mouse brain samples analyzed in this work was estimated using the 10X Visium Spatial Optimization kit (PN 1000192). Briefly, serial sections of the tissue

were sectioned from the OCT embedded tissue matrix and placed on the capture spots of Visium Spatial Tissue Optimization slide (10X Genomics, PN: 3000394). To estimate the optimal permeabilization time, pepsin (Sigma Aldrich; Cat. No. P7000) at a concentration of 1 mg/ml in 0.1N HCl (Fisher Scientific Cat. No. AA35644K2) was used with different incubation times (5, 10, 15, 20, 25, 30, 35 mins). The rest of the protocol was followed as per the manufacturer's recommendations (10X Genomics; Visium Spatial Gene Expression Reagent Kits – Tissue Optimization; CG000238 Rev E) to determine the most optimal time for tissue permeabilization. Imaging was performed using a Nikon NiE upright microscope equipped with a 10x Plan Apo lambda 0.45 air lens and a black and white sCMOS camera Prime BSI (Teledyne Photometrics). The setup was controlled by NIS-Elements (5.42.04, Nikon Instruments Europe B.V.). TRITC was excited with 550nm (CoolLED pE-800) and collected with a 577-630nm emission filter. A large tilescan was acquired to cover the entire tissue and chip area using 10% overlap and a focus surface. Like Stereo-seq, the optimal permeabilization time was assessed based on the strongest fluorescence signal with the lowest signal diffusion (crispness of the RNA footprint). Based on our assessment, we found the permeabilization time of 27 minutes optimal for mouse brain sections.

9. Nova-ST workflow: Tissue preparation, permeabilization & first strand synthesis

As with Stereo-seq experiments, the OCT embedded tissues were sectioned to a thickness of 10µm, placed on the capture area of the Nova-ST chip and melted. If the samples were not immediately processed for transcriptome capture, the Nova-ST chip was re-frozen on cryoblock and stored in -80°C and in our experience the quality of the tissue is not impacted with 2-3 weeks of storage of the tissue section at -80°C.

Standard Hematoxylin and Eosin (H&E) staining procedure was used to stain the tissue. Briefly, the Nova-ST chip with frozen tissue section was taken from -80°C storage and immediately melted on a 37°C block for 1 min. The tissue was then fixed in methanol at -20°C for 30 mins. Post fixation, the tissue was dehydrated by adding 150µl of isopropyl alcohol (IPA) and incubating for 1 min. After removal of the IPA, the Nova-ST chip was air dried for 3 mins (or until the chip is completely dried). 200 µl Mayer's haematoxylin (Agilent, Cat. No. S3309) was added to the chip and incubated for 7 mins. Using a forceps the chip was washed in excess NFW. 150µl of bluing buffer (Agilent, Cat. No. CS702) was added to Nova-ST chip and incubated for 2 mins. The chip

was again washed with excess NFW. The tissue was then treated with 200 µl of Eosin-Y buffered solution (10% v/v of Eosin-Y (Sigma, Cat. No. HT110216) in 0.45 M Tris Acetic acid solution pH 6.0) and incubated for 1 min. The chip was dried at 37°C for 5 mins (or until the water was completely evaporated) prior to imaging. Brightfield imaging was performed using a Nikon NiE upright microscope equipped with a 10x Plan Apo lambda 0.45 air lens and a color camera DFK 33UX264 (The Imaging Source, LLC). The setup was controlled by NIS-Elements (5.42.04, Nikon Instruments Europe B.V.). A large tile scan was acquired to cover the entire tissue and chip area using 10% overlap and a focus surface. Post imaging, the sample was immediately processed for the spatial transcriptomics workflow.

Pepsin reagent prepared in 0.01N HCl (pH 2.0) (1 mg/ml) was preheated in a 37°C oven. After H&E staining and imaging, using forceps, the Nova-ST chip was placed into a 3 cm petridish. 300µl of prewarmed pepsin was added to the H&E stained tissue and permeabilization was performed at 37°C for the optimal permeabilization time estimated in the previous step. After the permeabilization step, the pepsin solution was blotted off from the Nova-ST chip and the chip was transferred into a 24 well plate. The permeabilization reaction was stopped by sequentially washing the chip with 0.1X SSC (20X SSC; Thermo Fisher; Cat. No: 15557044), followed by 300 µl of 1X RT wash buffer (1X Maxima h- Reverse Transcriptase buffer; Cat. No EP0753, 1U/µl Lucingen NxGen RNase Inhibitor; Cat. No. 30281-2). Finally, 300 µl of First Strand mix was added to the well (1X Maxima h- Reverse Transcriptase buffer, 1U/µl Lucingen NxGen RNase Inhibitor, 4% Ficoll PM-400; Sigma Aldrich Cat. No. F4375-10G, 1 mM dNTP; Thermo Fisher Cat. No. R1121, 10U/µl Maxima RTase). The wells were covered with multiple layers of square patches of Parafilm. The 24 well plate was then sealed and put into oven at 42°C for overnight incubation for first strand synthesis.

10. **Nova-ST workflow: Exonuclease treatment**

In this step, exonuclease treatment was performed on the Nova-ST chip to remove single stranded HDMI capture tags without the first strand product to avoid the undesired secondary downstream reactions. After the first strand reaction, the FSS mix was removed from and the Nova-ST chip and was washed with 300 µl of 0.1X SSC. The chip was washed with 300 µl of 1X Exonuclease I buffer before adding 300 µl Exonuclease reaction mix (1X Exonuclease reaction buffer, 1U/µl of Exonuclease enzyme). The reaction was incubated at 37°C for 45 mins.

11. **Nova-ST workflow: Tissue clearance**

After incubation, the tissue on the surface of the Nova-ST chips was cleared. In this step, the exonuclease reaction mix was removed from the well and 300 µl of Tissue clearance reagent was added to the well containing the Nova-ST chip (100 mM of Tris pH 8.0; Thermo Fisher Cat. No AM9856, 200 mM of NaCl; Thermo Fisher Cat. No. AM9760G, 2% SDS; Thermo Fisher Cat. No. 24730020, 5 mM EDTA; Thermo Fisher Cat. No. 15575020, 16U/µl Proteinase K; NEB Inc Cat. No. P8107S). The reaction was incubated at 37°C for 45 mins to complete the tissue removal reaction.

12. **Nova-ST workflow: Second strand synthesis**

Before the subsequent processing, clearance of the tissue was ensured from the surface of the Nova-ST chip. Then the Nova-ST chip was washed three times with 3 ml of NFW. The chip was them washed three times with 0.1N NaOH. During each wash, the chip was incubated in 0.1N NaOH for 5 mins to remove the mRNA strand. The Nova-ST chip was then washed three times with 0.1M Tris-HCl (pH 7.5) (Thermo Fisher; Cat. No. 15567027) followed by 3X times wash with NFW. Using forceps, the chip was transferred to another well in the 24 well plate. Before transferring excess water was blotted off from the bottom of the chip surface. The chip was then subjected to second strand synthesis reaction. 300µl of Second strand synthesis reaction mix (1X NEB Buffer 2, 10 mM RPE randomer, 1 mM dNTP, 0.5 U/µl Klenow Fragment (3'→5' exo-); NEB Inc Cat. No. M0212L) was added to the well. The chip was incubated in the second strand reaction mix for 2 hours at 37°C.

13. **Nova-ST workflow: Second strand product extraction, cleanup and random primer extension PCR**

The Nova-ST flow cell was washed three times with 3 ml of NFW. Using forceps, the chip was transferred from the well to a 3 cm petri dish. Before transferring excess water was blotted off from the bottom of the chip surface. 90µl of 0.1 M NaOH was added to the surface of the Nova-ST chip, it was then incubated on the chip for 5 mins. After the incubation, the liquid was harvested into a DNA lo-bind 1.5 ml eppendorf tube. This process was repeated two additional times. After the final collection, the volume of the RPE collect was estimated and 0.28 times the volume of Tris 7.0 (Thermo Fisher Cat. No. AM9851) was added to the RPE collection to neutralize the reaction. After two minutes of incubation, the RPE product was purified using 1.8X Ampure XP

beads as per manufacturer's recommendation. The magnetic bead elution was performed with 44 µl of EB buffer (Qiagen Cat. No. 19086). 42 µl of elute was taken for the RPE PCR. PCR was performed on purified RPE product by adding the PCR mix (1X KAPA HiFi master mix; Roche Cat. No. KK2602, 1 µM RPE forward primer, 1 µM RPE reverse primer). The following RPE PCR program was used for the product amplification: 95°C – 3 mins, 14 cycles of {95°C – 30s; 60°C – 1 min; 72°C – 1 min}, final extension of 72°C – 5mins. The PCR product was then purified with 0.8X Ampure XP beads with elution in 40 µl of EB buffer.

14. **Nova-ST workflow: Index PCR and sequencing**

The RPE amplified library was quantified using the Qubit dsDNA Quantification kit (Thermo Fisher Cat. No. Q32851) and the size distribution of the RPE amplified library was estimated using High Sensitivity DNA kit (Agilent). Based on the measurements 10 nM RPE library dilution was performed using NFW. Final indexing PCR was performed by adding the PCR mix (1X KAPA HiFi master mix, 1 µM WTA forward primer, 1 µM WTA reverse primer and 2nM of RPE library). The following WTA PCR program was used for the product amplification: 95°C – 3 mins, 14 cycles of {95°C – 30s; 60°C – 30 s; 72°C – 30 s}, final extension of 72°C – 5mins. Two rounds of purification were performed on the amplified PCR product with 0.8X Ampure XP beads and the final elution in 60 µl of EB buffer. Sequencing of the Final Nova-ST libraries was performed on NextSeq 2000 sequencing platform. The concentration of the libraries were normalized to 2 nM using RSB buffer with Tween-20 (Illumina; Cat. No. 20512944). The 2 nM library was further diluted to 800 pM before loading to the instrument. The libraries were sequenced with the following sequencing specification: R1 = 34 bps; I1 = 8 bps; I2 = 8bps; R2 = 91 bps.

15. **Data Analysis – HDMIs**

Raw sequencing data was loaded into Illumina's Sequence Analysis Viewer to first, broadly check that the base composition matches the expected sequence (Supplementary Fig. 2) and secondly, to identify any tiles which were not sequenced – these were noted for exclusion in later steps. Raw data was next demultiplexed using Illumina's bcl2fastq, creating one set of fastq files per tile of the flow cell using the following command: `bcl2fastq -R ${RUN_FOLDER} -o Demultiplexed/${TILE_NO} -r 1 -p 1 -w 1 --tiles s_<${TILE_NO}> --use-bases-mask=y32n* --minimum-trimmed-read-length=32 --write-fastq-reverse-complement.` For each read in each fastq file, the barcode sequence was first checked

against the expected degenerate sequence and then the tile number, X coordinate, Y coordinate and read sequence, were recorded and saved to disk. A small subset (10,000) of valid reads was saved to a separate file for later use.

As the X and Y coordinates obtained during the previous step were local coordinates and specific to the tile that each read came from, to properly reconstruct the spatial location of every read from the chip, these coordinates were put within a global context, we use the fiducial markers (concentric circles) present within each tile to align them to each other. This step can either be performed on a per-chip basis, or across all tiles from the HDMI generation, but was performed per chip for the data in this study. To do this, for each tile processed, a numpy array (max_x, max_y) was created and the value at each coordinate where a valid barcode is found was set to 1. These matrices were trimmed and reshaped to bin the data in 25x25 bins, simplifying processing. Matrices were normalized to a maximum of 255 and then converted to greyscale images using OpenCV. These images were inverted, denoised (`cv.fastNlMeansDenoising, h=100`) and thresholded (min 128, max 255) to extract and image of just the fiducial markers. Next, Hough circle detection was used at 2 different radii (1. `minRadius=40,maxRadius=80`, 2. `minRadius=15,maxRadius=30`) to detect the coordinates of the inner and outer fiducial circles within each tile and the centroid of each set was calculated when all (8) circles were identified successfully. Where all 8 fiducials were not identified, centroid coordinates were interpolated using the adjacent tiles. Distances between fiducial circles were measured in the H&E and electron microscopy images (in both rows/swaths and columns) and these distances were used to calibrate the scale of the coordinates extracted from the fastq files to nm. HDMI coordinates were then corrected per tile to place the fiducial circles the correct distance from each other in both directions, beginning with the top left tile of a Nova-ST chip. Swaths 1, 3 and 5 begin at the same position, swaths 2, 4 and 6 are offset by a single tile, and this was accounted for in this correction.

16. Data Pre-processing – Spatial Libraries

Barcodes from the first 1 million reads from read 1 of the RNA sequencing were extracted and compared with the subset taken in HDMI Basic Processing to identify the tiles in the Nova-ST chip used which have reads (i.e. were under the tissue section). All barcodes from the HDMI data of the tiles covered by the section were then extracted into a whitelist for STARsolo (2.7.10b), these were trimmed to 31bp to allow STARsolo to perform a Hamming error-correction. For each

read of the spatial library, the UMIs present in read 2 (the first 9 base pairs) were extracted and appended to the corresponding read from read 1 and STARsolo was run with the following parameters: `--soloType CB_UMI_Simple --soloCBwhitelist ${BARCODE_WHITELIST_FILE} --soloCBstart 1 --soloCBlen 31 --soloUMIstart 32 --soloUMIlen 8 --soloBarcodeMate 0 --soloBarcodeReadLength 0 --soloFeatures Gene GeneFull --soloCBmatchWLtype 1MM --soloUMIdedup 1MM_All --soloCellFilter None --outSAMtype BAM SortedByCoordinate --outSAMattributes NH HI AS nM CR CY UR UY CB UB sS --readFilesIn ${READ_2} ${READ_1_PLUS_UMI}`

Matrices from STARsolo and the corrected coordinates were combined into a GEM file, a format used by BGI's Stereo-seq pipeline, to enable the loading and binning functions of Stereopy, as well as consistent analysis of the two data types.

17. Data Analysis – Spatial Libraries

Stereo-seq matrices were loaded into Python using Stereopy (v0.6) at three different bin sizes, 50, 100 and 200, for Nova-ST, the following sizes were used 728, 1456, 2912, where each size corresponds to the same dimensions as the Stereo-seq bins. Following data loading, samples were further analyzed using Scanpy, in brief: Standard quality control metrics were calculated and bins with too few genes detected were removed (Bin 50/728: 75 Genes, Bin 100/1456: 250 Genes, Bin 200/2912: 500 Genes), bins were not filtered by mitochondrial percentage. Bins were normalized to a total of 10,000 counts and log transformed before highly variable genes were identified. A regression for total counts and mitochondrial percentage was applied and counts were scaled to unit variance with a mean of 0, counts above 10 were clipped to 10. A principal component analysis was performed, followed by neighborhood detection, UMAP generation and leiden clustering at several resolutions.

For cell2location, the single cell mouse cortex and hippocampus data from Yao et al 2021 was loaded from the provided h5 files into an Anndata object and associated metadata was added. This object was then subsampled without replacement to contain half of the original cells to simplify training the cell2location model. A cell2location RegressionModel instance was created using the single cell data with external_donor_name_label as the batch key and subclass_label at the labels key, and the model was trained with 500 max epochs. The cell abundance estimations were

exported using the following parameters: num_samples=1000, batch_size=2500. Next, both the spatial and single cell data were subset to only include genes present in both datasets and a Cell2location model was set up using the spatial data (at bin50), the exported reference cell states and N_cells_per_location=2. This model was then trained using the following parameters: max_epochs=30000, batch_size=None, train_size=1) and the final estimated cell abundances was exported with num_samples=1000 and batch_size equal to $1/10^{th}$ of the dataset size. Anndata objects were converted into loom files, including the spatial coordinates and cell2location prediction scores for visualization in Scope.