

## Supplementary Materials for

### **Visualization and analysis of gene expression in tissue sections by spatial transcriptomics**

Patrik L. Ståhl, Fredrik Salmén, Sanja Vickovic, Anna Lundmark, José Fernández Navarro, Jens Magnusson, Stefania Giacomello, Michaela Asp, Jakub O. Westholm, Mikael Huss, Annelie Mollbrink, Sten Linnarsson, Simone Codeluppi, Åke Borg, Fredrik Pontén, Paul Igor Costea, Pelin Sahlén, Jan Mulder, Olaf Bergmann, Joakim Lundeberg,\* Jonas Frisén

\*Corresponding author. Email: joakim.lundeberg@scilifelab.se

Published 1 July 2016, *Science* **353**, 78 (2016)  
DOI: 10.1126/science.aaf2403

#### This PDF file includes

Materials and Methods  
Figs. S1 to S11  
Tables S1 and S2  
References

## **Supplementary materials**

### **Visualization and analysis of gene expression in tissue sections by spatial transcriptomics**

#### **Materials and methods**

##### ***Outline of the Spatial Transcriptomics workflow***

The Spatial Transcriptomics method follows the following workflow: array printing, tissue sectioning, tissue fixation, tissue staining, imaging, pre-permeabilization, permeabilization, cDNA synthesis, tissue removal, probe release, library preparation, sequencing, data processing, data visualization and analysis. When the method refers to separate wells, these wells are created on the glass slide by using a rubber mask (ArrayIT hybridization cassette) to separate the surface into reaction compartments, suitable for the size of the printed arrays, where on-chip reactions can take place. Typically, from tissue types that the protocol has been optimized for, 75% of tissue sections put onto the array will produce high quality data. This is an approximate estimate based on reproducibility between samples of the same type, given the RNA (RIN > 8) and tissue morphology is of good quality. We specify “high quality data” as sample data with an average number of genes per feature above a certain threshold value, as well as correct tissue placement on the spatial array (tissue inside the area, few folded areas and missing pieces etc.). The approximate values differ between different types of samples. For mouse olfactory bulb we require at least 3000 genes per feature at average and at least 90% of the bulb to be intact and

placed within the spatial array. For breast cancer samples we require at least 3000 genes per feature at average in the tumor regions and that all of the tumor regions are placed within the spatial array.

Optimization of the method for a particular tissue requires adjustment of the parameters for permeabilization, usually different times or concentrations (decreased with smaller tissues and increased with larger), as well as adjustment of the section thickness.

#### ***Generation of surfaces with poly-T reverse transcription primers***

To investigate the feasibility of generating cDNA from mRNA in tissue sections with maintained spatial resolution we immobilized poly-20TVN capture oligonucleotides (IDT) on the surface of Codelink Activated microscope glass slides (#DN01-0025, Surmodics), according to the manufacturer's instructions (Surmodics). The oligonucleotide immobilized on the surface was:

Surface reverse transcription oligonucleotide:

[AmC6]UUUUUGACTCGTAATACGACTCACTATAAGGACACGACGCTTTCCGATCTNNNNNNNTTTTTTTTTTTTTTVN

#### ***Generation of spatially barcoded arrays for Spatial Transcriptomics***

For comparing the efficiency of cDNA generation on an array with cDNA generation in solution, and for generation of Spatial Transcriptomics data, we arrayed 1007 capture oligonucleotides with 18-mer unique barcodes, a 9-mer semi-randomized UMI and a poly-20TVN capture region (IDT) onto the surface of Codelink Activated Slides according to the manufacturer's (Surmodics) instructions. The oligonucleotides were immobilized in 100 $\mu$ m features with 200 $\mu$ m center-to-

center distance between the spots. According to the slide manufacturer approximately 200 million oligonucleotides were immobilized in each spot. An area in the top left corner of the array was left blank for the purpose of orientation. Six 6200x6600 $\mu$ m arrays were prepared per glass slide. The oligonucleotides immobilized on the spatially barcoded arrays were:

Surface spatially barcoded oligonucleotides:

[AmC6]UUUUUGACTCGTAATACGACTCACTATAAGGACACGACGCTTCCGATCT[18  
MER\_SPATIALBARCODE]WSNNWSNNVTTTTTTTTTTTTTVN

### ***Collection and preparation of olfactory bulbs***

For spatial transcriptomics analysis, adult C57BL/6 mice (>2 months old) were euthanized and olfactory bulbs were immediately isolated and snap-frozen in Isopentane (#M32631, Sigma-Aldrich). Tissue was embedded in cold OCT (#4532, Sakura) before sectioning. The olfactory bulb was sectioned on a cryostat at a thickness of 10 $\mu$ m. Sections were mounted onto spatially barcoded arrays or to surfaces with poly-T reverse transcription primers with one section per well. For smFISH, sections were mounted on cover glasses and for LCM, sections were mounted onto MMI Tool membrane (MMI AG, Glattburg, Switzerland).

For in situ hybridization, euthanized mice were transcardially perfused with PBS (#SH30378.02, GE Healthcare Life Sciences) followed by 4% formaldehyde (#F1635, Sigma-Aldrich). Brains were post-fixed in 4% formaldehyde over-night. Vibratome-sections (25 $\mu$ m) were cut and mounted on glass slides (#J1800AMNZ, Fisher Scientific).

### ***Collection and preparation of breast cancer biopsy***

The biopsy was collected and snap frozen. Tissue was embedded in cold OCT before sectioning.

The biopsy was sectioned on a cryostat at a thickness of 16 $\mu$ m. Every forth section was mounted onto spatially barcoded arrays with one section per well.

### ***Staining and imaging***

Sections were fixed for 10 minutes in neutral formaldehyde solution, diluted 1:9 from 36.5 - 38.0% stock solution (#F8775, Sigma-Aldrich) in 1xPBS (#09-9400, Medicago). Directly following fixation, the formaldehyde was washed away with 1xPBS followed by addition of Propan-2-ol (#A461-1, Fisher Scientific), sections for smFISH were stored at -80°C until processed. For Spatial Transcriptomics experiments and LCM experiments, sections were stained with Mayer's Hematoxylin (#S3309, Dako) and Bluing buffer (#CS702, Dako) followed by Eosin (#HT110216, Sigma-Aldrich) in Tris-base (0.45M Tris, 0.5M Acetic acid, pH 6.0). Sections were rinsed, and for Spatial Transcriptomics experiments, dried sections were mounted with 85% Glycerol (#104094, Merck Millipore) and covered with a coverslip (#BB024060A1, Menzel-Gläser).

Bright field imaging was carried out using the Metafer Slide Scanning Platform (Metasystems). Raw images were stitched together with the VSlide software (Metasystems). Images were visualized and extracted as .jpeg.

### ***Permeabilization and reverse transcription***

For each well, corresponding to each sub-array with a section, 70 $\mu$ l of 1x Exonuclease I Reaction Buffer (#B0293S, NEB) with 0.19 $\mu$ g/ $\mu$ l BSA (#B9000S, NEB) was added and incubated at 37°C

for 30 minutes. Each well was washed with 100 $\mu$ l 0.1x SSC, diluted in deionized water from stock solution (#S6639, Sigma-Aldrich). Next, 70 $\mu$ l of 0.1% pepsin (#P7000-25G, Sigma-Aldrich) dissolved in 0.1M HCl (#318965-1000ML, Sigma-Aldrich) was added to each well and incubated at 37°C for 10 minutes. Each well was washed as previously described and 70 $\mu$ l of reverse transcription mix was added to each well and incubated overnight (ON). For spatially barcoded arrays belonging to samples intended for sequencing, the reverse transcription mix contained 1x First Strand Buffer (#18080-044, Invitrogen), 5mM DTT (#18080-044, Invitrogen), 500 $\mu$ M of each dNTP (#R0192, Fisher Scientific), 0.19 $\mu$ g/ $\mu$ l BSA, 50ng/ $\mu$ l Actinomycin D (#A1410-2MG, Sigma-Aldrich), 1% DMSO (#472301-500ML, Sigma-Aldrich), 20U/ $\mu$ l Superscript III (#18080-04, Invitrogen) and 2U/ $\mu$ l RNaseOUT (#10777-019, Invitrogen). For spatially barcoded arrays and surfaces with poly-T reverse transcription primer intended for generation of fluorescent cDNA footprints, the reverse transcription mix contained the same reagents as above except for 500 $\mu$ M of each dATP/dGTP/dTTP, 12.5 $\mu$ M of dCTP , 25 $\mu$ M of Cyanine 3-dCTP (#NEL576001EA, PerkinElmer). For experiments optimizing the conditions for permeabilization the permeabilization was carried out for 0, 2, 10 and 30 minutes. For experiments optimizing the conditions for cDNA synthesis the cDNA synthesis was carried out for 1 hour, 3 hours, 6 hours and overnight.

To degrade and remove mouse tissue, 70 $\mu$ l Proteinase K (#19131, Qiagen) and PKD Buffer (pH 7.5, according to the manufacturer, #1034963, Qiagen), at a ratio of 1:7, were added to each well and incubated at 56°C for 1 hour. After the incubation, slides washed with 2x SSC with 0.1% SDS at 50°C (#71736-100ML, Sigma-Aldrich) followed by 0.2x SSC and 0.1x SSC at room temperature (RT) and finally spin-dried. For the breast cancer biopsy, the tissue was incubated in

1:100  $\beta$ -Mercaptoethanol (#444203, CALBIOCHEM) in RLT buffer (#79216, Qiagen) at 56°C for 1 hour before the Proteinase K step.

#### ***cDNA library preparation for sequencing***

For each well, 70 $\mu$ l of release mix containing 1.1x Second Strand Buffer (#10812-014, Invitrogen), 8.75 $\mu$ M of each dNTP, 0.20 $\mu$ g/ $\mu$ l BSA and 0.1U/ $\mu$ l USER Enzyme (#M5505, NEB) was added and incubated at 37°C for 1-2 hour and 65 $\mu$ l from each well were collected. 5 $\mu$ l of second strand mix containing 2.7x First Strand Buffer, 3.7U/ $\mu$ l DNA polymerase I (#18010-017, Invitrogen) and 0.18U/ $\mu$ l RNaseH (#18021-014, Invitrogen) was added and samples were incubated at 16°C for 2 hours, 5 $\mu$ l T4 DNA polymerase /#M0203S, NEB) was added and samples were incubated at 16°C for another 20 minutes. 25 $\mu$ l of deionized water or 80mM EDTA (#15575-038, Invitrogen) was added and purified using Agencourt RNAClean XP beads (#A63987, Beckman Coulter) according to the manufacturer's protocol eluted into deionized water. 5.6 $\mu$ l of the samples were mixed with 10.4 $\mu$ l In Vitro Transcription mix with a final content of 1x T7 Reaction Buffer (#AM1333, Ambion), 7.5mM of each NTP (#AM1333, Ambion), 1x T7 Enzyme Mix (#AM1333, Ambion) and 1U/ $\mu$ l SUPERaseIN (#AM2694, Ambion). Samples were incubated at 37°C for 14 hours.

The samples were purified using Agencourt RNAClean XP beads according to the manufacturer's protocol and eluted into 10 $\mu$ l deionized water. The amount and average fragment length of amplified RNA (aRNA) was determined by using the RNA 6000 Pico Kit (#5067-1513, Agilent) with a 2100 Bioanalyzer (Agilent) according to the manufacturer's protocol. The remaining sample and 2.5 $\mu$ l Ligation adapter (IDT) was added to a final concentration of 0.71 $\mu$ M. Samples were heated at 70°C for 2 min. then placed on ice before 4.5 $\mu$ l ligation mix was added to a final

content of 1x T4 RNA Ligase Reaction Buffer (#B0216L, NEB), 20U/ $\mu$ l T4 RNA Ligase2, truncated (#M0242L, NEB), 4U/ $\mu$ l RNase Inhibitor, Murine (#M0314L, NEB) and 0.5 $\mu$ M Ligation Adapter. Samples were incubated at 25°C for 1 hour. The samples were purified using Agencourt RNAClean XP beads according to the protocol previously described and mixed with 1 $\mu$ l RT-primer (IDT) to a final concentration of 1.7 $\mu$ M and 1 $\mu$ l dNTPs and a final concentration 0.83mM of each dNTP. Samples were heated at 65°C for 5 min and then placed on ice, 8 $\mu$ l reverse transcription mix was added to a final content of 1x First Strand Buffer, 0.05M DTT, 500 $\mu$ M of each dNTP, 1mM RT-primer, 10U/ $\mu$ l Superscript III and 2U/ $\mu$ l RNaseOUT. Samples were incubated at 50°C for 1 hour before placed on ice and the samples were subsequently purified using Agencourt RNAClean XP beads according to the protocol previously described. A total reaction volume of 10 $\mu$ l containing 1xKAPA HiFi HotStart ReadyMix (#KK2601, KAPA Biosystems), 1x EVA green (#31000, Biotium), 0.5 $\mu$ M PCR InPE1.0 (Eurofins), 0.01 $\mu$ M PCR InPE2.0 (Eurofins), 0.5 $\mu$ M PCR Index (Eurofins) and 2 $\mu$ l purified cDNA were amplified by qPCR with the following protocol: 98°C for 3 minutes, followed by cycling at 98°C for 20 seconds, 60°C for 30 seconds and 72°C for 30 seconds. After determining the amount of cycles needed based on the qPCR, the samples were amplified in a total reaction volume of 25 $\mu$ l. The libraries were purified as described previously (18) and samples were eluted in 20 $\mu$ l elution buffer (#19086, Qiagen) and the average fragment length of finished libraries was determined by using the DNA 1000 Kit (#5067-1504, Agilent) with a 2100 Bioanalyzer according to the manufacturer's protocol. The concentration of the finished libraries was determined with Qubit dsDNA HS Assay Kit (#Q32854, Life Technologies) according to the manufacturer's protocol. Finished libraries were diluted to 4nM and sequenced on the Illumina NextSeq platform using paired-end sequencing according to the manufacturer's protocol. Typically, 31 or 51 bases were

sequenced on read one, and 121 or 101 bases were sequenced on read two. The oligonucleotides used during library preparation were:

Ligation adapter:

[rApp]AGATCGGAAGAGCACACGTCTGAACCTCCAGTCAC[ddC]

Second reverse transcription primer:

GTGACTGGAGTTCAGACGTGTGCTCTTCCGA

PCR primer InPE1.0:

AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTCCGATCT

PCR primer InPE2.0:

GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

PCR Index primer:

CAAGCAGAAGACGGCATACGAGATXXXXXXTGACTGGAGTT

### ***Fluorescent imaging and image alignment***

For spatially barcoded arrays belonging to samples intended for sequencing where barcoded cDNA had been released, 70 $\mu$ l of hybridization mix containing 0.96x PBS, 0.2 $\mu$ M Cy3\_anti\_A\_probe (Eurofins) and 0.2 $\mu$ M Cy3\_anti\_Frame\_probe (Eurofins) was added and incubated at RT for 10 minutes. Each well was washed with 100 $\mu$ l 0.1x SSC, then 2x SSC with 0.1% SDS at 50°C for 10 minutes, 0.2x SSC at RT for 1 minute, 0.1x SSC at RT for 1 minute and finally spin-dried. Both spatially barcoded arrays belonging to samples intended for sequencing, and spatially barcoded arrays and surfaces with poly-T reverse transcription primer intended to

show fluorescent cDNA footprints were mounted with SlowFade Gold Antifade Reagent (#S36963, Invitrogen) and covered with a coverslip.

Fluorescent imaging, stitching and image extraction was carried out on the same system and with the same programs as previously described for bright field images. Bright field images and fluorescent images were manually aligned in Adobe Photoshop CS6 (Adobe), by using tissue structures and features visible in both images. The oligonucleotides used for array hybridization were:

Cy3 anti-A probe:

[Cy3]AGATCGGAAGAGCGTCGTGT

Cy3 anti-Frame probe:

[Cy3]GGTACAGAACGCGCATAGCAG

#### ***Estimation of diffusion by fluorescent signal***

Signal intensities were measured in ImageJ for six individual cells (fig. S1E, F). All intensities were then corrected and presented as above background level and the distances between the H&E- and Cy3-signals at 10% of the maximum Cy3-signal intensity measured. An example of the measured distance observations is given in fig. S1G. The average diffusion distance at 10% Cy3-signal was estimated and shown in fig. S1H.

#### ***Extraction and fragmentation of total RNA***

Total RNA was extracted from fresh frozen sections of the mouse main olfactory bulb. Tissue was placed in Lysing Matrix D (#116913050, MP Biomedicals) and was lysed by running the tubes in

a FastPrep (MP Biomedicals). RNA was extracted using the RNeasy Plus Mini Kit (#74134, Qiagen) according to the manufacturer's protocol. Total RNA was fragmented using the NEBNext Magnesium Fragmentation Module (#E6150S, NEB) for 2.5-3 minutes to an average length of 200-300nts according to the manufacturer's protocol. Fragmented RNA was purified with RNeasy MiniElute Cleanup Kit (#74204, Qiagen) according to the fragmentation protocol. The average fragment length of fragmented RNA and RIN values of extracted RNA was determined by using the RNA 6000 Pico Kit with a 2100 Bioanalyzer according to the manufacturer's protocol. The concentration of the finished libraries was determined with Qubit RNA HS Assay Kit (#Q32852, Life technologies) or Qubit RNA BR Assay Kit (#Q10210, Life technologies) according to the manufacturer's protocol.

#### ***Total RNA experiment using spatially barcoded arrays***

Total RNA experiments were carried out in the same way as for experiments using tissue sections on spatially barcoded arrays intended for sequencing with three exceptions; no tissue was mounted onto the array, the glycerol mounting step and bright field imaging step were skipped, and 300ng of fragmented total RNA from mouse olfactory bulb was added to each well in the reverse transcription step.

#### ***In solution cDNA library preparation***

Samples with a total volume of 12 $\mu$ l and a concentration of 1.67mM ST-probe mix (equal mix of six unique barcoded probes), 0.83mM of each dNTPs and 140ng fragmented total RNA from mouse main olfactory bulb were prepared. Samples were heated at 65°C for 5 min. After the heating step, samples were placed on ice before 8 $\mu$ l reverse transcription mix was added to a final

content of 1x First Strand Buffer, 0.05M DTT, 500 $\mu$ M of each dNTP, 1mM ST-probe mix, 10U/ $\mu$ l Superscript II (#18064-014, Invitrogen) 2U/ $\mu$ l RNaseOUT and 1mM TS-probe (Exiqon). Samples were incubated at 42°C for 1 hour followed by 70°C for 15 minutes. Libraries were finished and sequenced in the same way as described after the reverse transcription step under “cDNA library preparation for sequencing”. The oligonucleotides used during in solution cDNA library preparation were:

In solution reverse transcription primers:

[AmC6]UUUUUGACTCGTAATACGACTCACTATAGGGACACGACGCTTCCGATCT[18  
MER\_SPATIALBARCODE]WSNNWSNNVTTTTTTTTTTTTTVN

Template switching probe (TS-probe):

AGACGTGTGCTCTTCCGATCTNNNNNNNNNArGrG+G (+=LNA, r=ribo-base)

### ***ISH in tissue sections***

ISH was performed as described previously (19). Labeled probes were produced with DIG RNA labeling kit (#1175025, Roche). DNA template sequences were amplified from mouse olfactory bulb cDNA using PCR primer (IDT) sequences obtained from Allen Brain Atlas, conjugated to the T3 promoter sequence. The oligonucleotides used for *Rbfox3* template amplification were:

*Rbfox3* forward primer:

CCAGCTGAATATGCCACC

*Rbfox3* reverse primer (including T3 promoter):

ATTAACCCTCACTAAAGGGAGGCCGATGGTGTGATGGTAA.

### ***Single molecule fluorescent in situ hybridization***

The sections were permeabilized with methanol and incubated for 10 minutes at 70°C in Tris-EDTA (pH 8.0.) Slides were washed twice with 2xSSC and the incubated in hybridization buffer (20) containing 250nM fluorescent label probes (LGC Biosearch Technologies) for 4 hours at 38.5°C. Slides were washed four times in 2xSSC, 20% formamide and counterstained with Hoechst. The staining step was followed by a wash in 2xSSC and finally the slides were mounted with Prolong Gold mounting medium (#P36930, ThermoFisher Scientific). Image stacks (0.3μm distance) were acquired using a Nikon Ti-E with motorized stage (Nikon).

The images were analyzed using python with the numpy, scipy.ndimage and scikit-image libraries. Briefly, after background removal using a large kernel gaussian filter, a Laplacian-of-Gaussian was used to enhance the RNA dots. Background objects, significantly larger than the smFISH dots, were removed after image thresholding. The remaining RNA dots were counted in round areas with a diameter of approximately 100μm in GCL, MCL, OPL and GL (3-6 areas per region). The images were stitched, aligned and pseudocolored in Fiji. The count data was compared to Spatial Transcriptomics data in the same regions with the exception that all possibly spots, overlapping only one region in the section, were used.

### ***Laser capture microdissection***

Laser capture microdissection was performed with the Olympus IX81 system (Olympus) and utilizing the mmi Cell Tools (MMI). The microdissected pieces were captured onto 0.2 ml white isolation caps for subsequent cDNA synthesis and library preparation. Round areas with a diameter slightly above 100μm were cut from MCL and GCL. This size was chosen in order to mimic the feature size on Spatial Transcriptomics arrays. After capture, 100ul Lysis buffer from

the RNAqueous-Micro Total kit (#AM1931, Ambion) was added to the tubes followed by an incubation for 30 min at 42 °C. The samples were frozen down at -80 °C until further processing. RNA was extracted using kit according to the manufacturer's protocol. After extraction, the volume was brought down to 2ul using a Speedvac. To each sample, an unique ST-probe (0.1 pmol) was added as well as DTT, 1uM dNTP mix, SuperScript III and RNaseOUT to a final concentration as stated in the standard protocol. Also actinomycin D (dissolved in 10% v/v DMSO) was added, to the same concentration as previously described, and samples were resuspended in 1x First Strand Buffer to a total volume of 10ul. The reverse transcription was performed at 42°C for 60 min followed by a heat inactivation step at 70 °C for 15 minutes. A total RNA experiment on a spatial array was carried out as previously described but with 500ng of fragmented total RNA (per well) from MCF-7 cells. During the release step, the LCM cDNA was spiked in, with a maximum of three samples per well. Libraries for sequencing were carried out as previously described. The oligonucleotides used for LCM cDNA library preparation were:

LCM reverse transcription primers:

[AmC6]UUUUUGACTCGTAATACGACTCACTATAGGGACACGACGCTTCCGATCT[18  
MER\_SPATIALBARCODE]WSNNWSNNVTTTTTTTTTTTTTVN

#### ***Sequence alignment and generation of gene expression database***

In each pair of reads, the first read contained the spatial barcode and the UMI but no gene information. Read two contained the gene information and was first processed using BWA based quality trimming. In addition, stretches of homopolymers exceeding 15 bases were removed together with the sequence following the homopolymer, and all reads shorter than 28 bases were

discarded. Remaining reads were aligned against the mouse genome (GRCm38) for olfactory bulb samples, and against the human genome (GRCh38) for breast cancer samples, using Bowtie2 (21) v2.2.3 with default settings. The number of reads aligning to each gene were counted with htseq-count (22) v0.6.1, using RefSeq gene annotations restricted to validated protein coding and non-coding transcripts (transcript\_ids starting with “NM\_” or “NR\_”) for olfactory bulb samples, and using ensembl gene annotations for breast cancer samples. Mitochondrial genes, ribosomal sequences and the long-non-coding transcript “Malat1” were removed from the analysis due to high abundance and internal priming. Whenever read two in a read pair was aligned to a gene, the spatial barcode from the corresponding read one was assigned to a barcode in the reference file as described previously (23) (“de-multiplexing”). To remove reads that were duplicated during the amplification protocol, a cluster algorithm that sorts the UMIs and iterates over them to create clusters based on minimum hamming distance was used. Finally all information was written to a file to generate a database containing combined gene expression and spatial information.

### ***Comparing libraries and features***

The total expression for each gene in a dataset was calculated by adding the reads from each barcode together. Data from the replicates and the library made in solution was normalized using the median ratio normalization method implemented in DESeq2 (24), followed by adding a pseudocount of 1 before the data was log2-transformed. For each pair-wise comparison, the Pearson correlation was calculated using all genes expressed in any of the two datasets. To compare the number of genes at different read depths in each of three samples (ST, total RNA in solution and total RNA on surface), FASTQ files were down-sampled.

Read coverage over an average “meta-gene” was calculated using the script *geneBody\_coverage2.py* from the RSeQC suite (25).

To display the number of unique genes and transcripts per feature and the coefficient of variation versus the mean expression, 273 features that were covered by at least 50% tissue were selected. All transcripts expressed were used and, except for the amplification duplicate removal (UMI-filtering), no normalization was carried out. To calculate the number of genes and unique reads per feature in selected morphological regions, a subset of features only covered by one region was selected. A total of 37 features from olfactory nerve layer, 41 from glomerular layer, 37 from outer plexiform layer, 13 from mitral cell layer and 60 from granular cell layer were chosen for this purpose. Subsampling was performed as described above.

### ***Spatial gene expression analysis and visualization***

The position of features containing genes of interest was plotted in R using the files containing the spatial gene expression information. Only features with more than 1000 genes were used. Samples were normalized as previously described. To display the enriched spatial expression patterns for selected genes, individual cutoffs were applied based on the total expression of the selected gene in the dataset.

To visualize low horizontal diffusion, ten features from MCL and 30 features from GCL at different distance from MCL were selected. Since only ten features could be found at the closest distance, center of feature at 50-55 $\mu$ m from MCL, this number was also chosen for the other areas. From Allen Brain Atlas, ten genes with a specific expression in MCL, OPL and GL but not in GCL were chosen. Data was normalized as previously described. The expression for each feature at the different distances were visualized as a heat map and directly visualized on top of the tissue.

Also, the sum of unique transcripts for all ten genes was also added together to calculate a value in average per feature for each region. The same procedure was repeated for the LCM samples; however, only five samples from each region were obtained.

### **Differential expression analysis**

For analysis of differential expression between selections based on morphology, ten features from each of three regions were selected. Only features with more than 1000 genes were used and data was normalized as previously described. Features in regions  $\alpha$  and  $\beta$  were located in the granular cell layer, with one region from each hemisphere. Features in region  $\gamma$  were located in the glomerular layer in both hemispheres. Pair-wise comparisons were carried either by taking the mean expression of each gene over all ten features or by comparing single features directly. Differential expression analysis was carried out using DESeq2. Only genes with a log2-fold change  $> 1$  and an adjusted p-value of  $< 0.05$  were considered to be differentially expressed. For each selection, examples of overexpressed genes were marked in the scatterplot together with three housekeeping genes (*Rpn2*, *Pgk1* and *Actb*).

For analysis of differential expression between gene-based selections, data was normalized as previously described. Features were selected based on the presence and absence of three interneuron marker genes, *Camk4*, *Vip* and *Th*. A gene was considered present if the expression was 1 or more normalized counts in a feature and absent if the expression was below 1. To ensure that selections with similar sequencing depth were compared, ten random features from each set were selected. Pair-wise comparison and differential expression analysis were carried out as previously described. The three house keeping genes together with examples of overexpressed genes were marked in the scatterplot.

### ***Dimensionality reduction and hierarchical clustering***

All features from both replicates covered by at least 50% tissue and expressing 1000 genes were chosen for t-SNE, making up 551 features. All genes present in at least one feature, making up a total of 17,201 genes, were used. Data was normalized as previously described. The Rtsne R package was used to perform dimensionality reduction of the expression data. The separated features were plotted in three dimensions and each feature was assigned to a cluster based on Euclidean distance (ward.D2). To determine the optimal number of clusters, the DIndex method from the NbClust R package was used. Each cluster was assigned a color and the features were plotted in the tissue showing the cluster color. To find cluster-specific marker genes, DESeq2 were used, and each cluster was compared to the rest of the features. Genes with a log2-fold change  $> 2$  and an adjusted p-value  $< 0.05$  were collected in a heat map, making up 532 genes. Examples of these genes were plotted in the tissue with the gene expression values colored from dark blue (low) to yellow (medium) to red (high).

Principal component analysis was carried out on features that contained at least 1000 genes and that only overlapped one morphological region. Thus, 186 out of the 273 features located under the tissue were selected, with 37 from olfactory nerve layer, 41 from glomerular layer, 37 from outer plexiform layer, 13 from mitral cell layer and 58 from granular cell layer. Data was normalized as previously described. Principal component analysis (PCA) was carried out using the prcomp() function in R by using all genes present in at least one feature, making up a total of 16,155 genes. To show some examples of genes that contribute most to the separation, the loadings of example genes were plotted over the PCA plots.

### ***Correlation between individual features in adjacent sections***

Replicate 10 and 11 were chosen due to the fact that they were adjacent to each other and that it was possible to find features with the approximate same location. The data was normalized as previously described and Pearson correlations were calculated using all genes present in at least one of the two features that were compared.

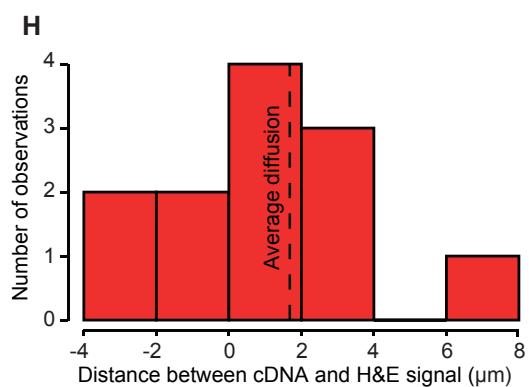
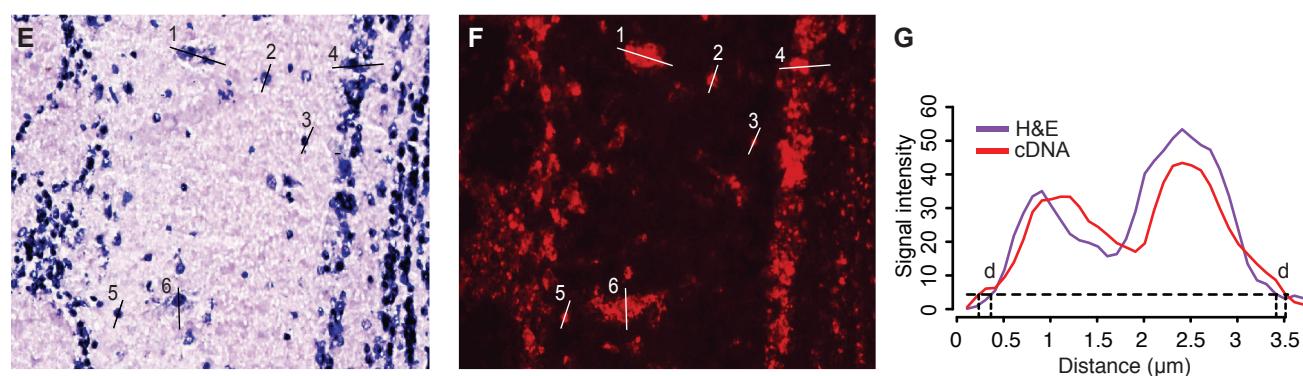
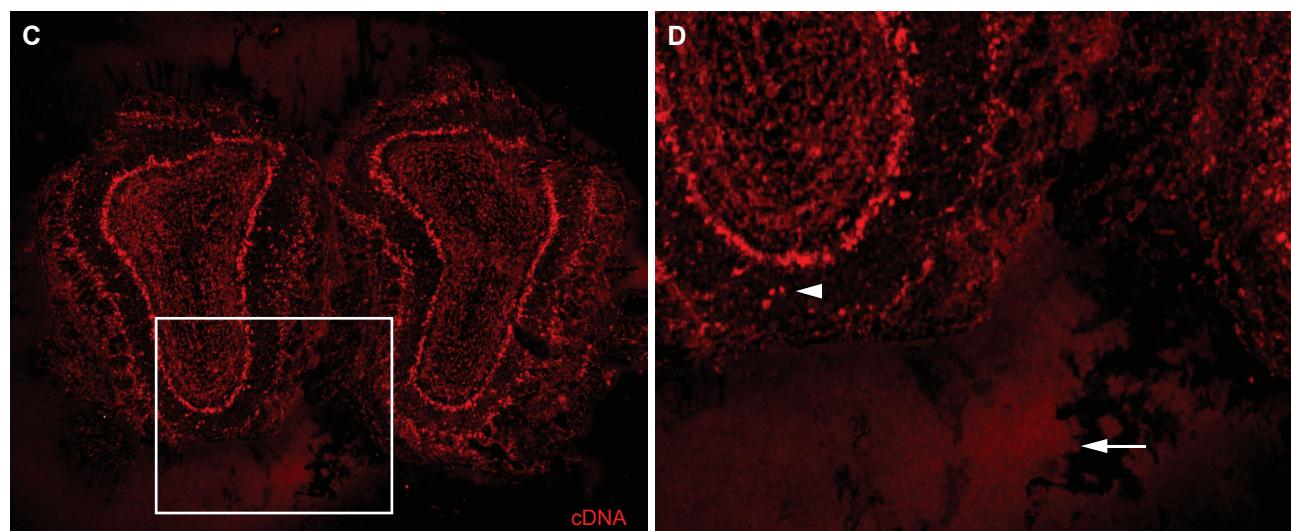
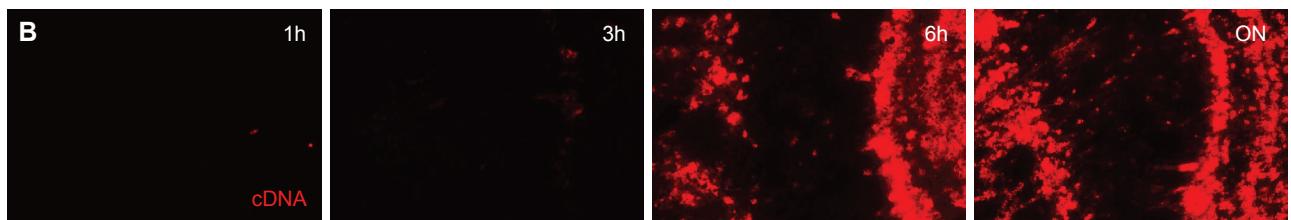
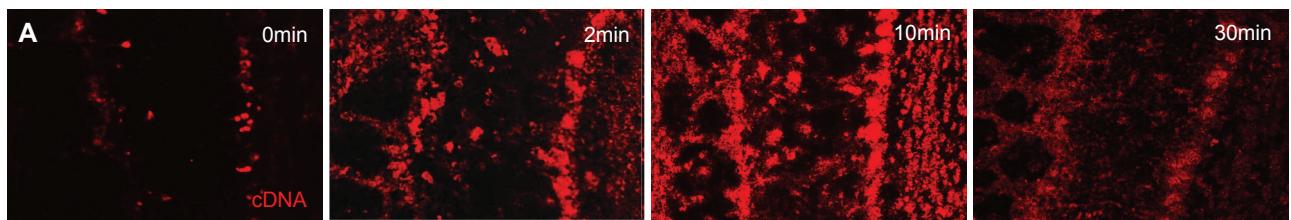
### ***Analysis of breast cancer biopsy***

Areas with invasive ductal cancer, as well as areas of ductal cancer in situ were marked by a pathologist. For analysis of differential expression, features from each of the seven regions were selected (in all four tissue sections, Fig. 4C, D, fig. S11A, B). Only features with more than 1000 genes with none or minimal infiltration of stromal cells were used and data was normalized as previously described. Differential expression analysis was carried out using DESeq2. Only genes with a log<sub>2</sub>-fold change > 2 and an adjusted p-value of < 0.05 (between two regions) were visualized and examples of relevant genes were marked in the heatmaps (Fig. 4E, fig. S11C).

### ***Comparison of Spatial Transcriptomics data and ISH data***

Allen Brain Atlas ISH gene expression data for sagittal sections were downloaded through the API using a modified version of the example script provided at [http://api.brain-map.org/examples/doc/structures/download\\_data.py.html](http://api.brain-map.org/examples/doc/structures/download_data.py.html). To compare the number of genes that could be detected with Spatial Transcriptomics versus Allen Brain Atlas, genes with expression energy of at least 1 (default) in Allen Brain Atlas, and 10 unique transcripts (slightly above one detected transcript per million mapped) from Spatial Transcriptomics RNA-seq data were used. Genes that were only detected with one of the methods, as well as the genes that could be detected

with both methods are presented as a Venn diagram. The expression and number of genes that are specifically detected with only one of the method are showed in scatterplots. The biological functions of the 1070 ABA-specific genes were determined by using Panther and DAVID.



**Fig. S1. Spatially localized cDNA synthesis from a tissue section on an oligo(dT)-array**

**(A)** A time series showing the fluorescent cDNA footprint corresponding to varying pepsin permeabilization incubation times. The signal increases from 0 minutes, through 2 minutes to 10 minutes, and has started to blur out at 30 minutes.

**(B)** A times series showing the fluorescent cDNA footprint corresponding to varying incubation times for cDNA synthesis. The amount of cDNA increases from 1 hour, through 3 hours to 6 hours, and remains at this intensity over night (ON).

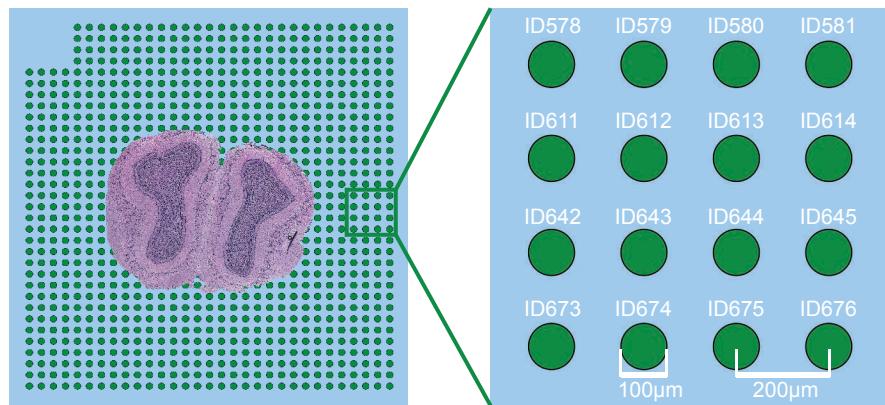
**(C)** Non-specific fluorescent cDNA signal can be observed around the edges of the tissue. This signal can be linked to diffusion of RNA from the edges of the section. In a magnified image it can be observed that this does not affect the spatial restriction of the fluorescent cDNA signal under the tissue.

**(D)** Magnification of box in **(C)**. Arrowhead indicates cDNA under the tissue with preserved spatial restriction. Arrow indicates cDNA synthesized from RNA diffused from the edges of the tissue.

**(E, F)** Measuring areas for six cells, marked in H&E- (black lines) and Cy3-channels (white lines) respectively. The measuring areas include both cells and connective tissue in order to quantify the diffusion distance.

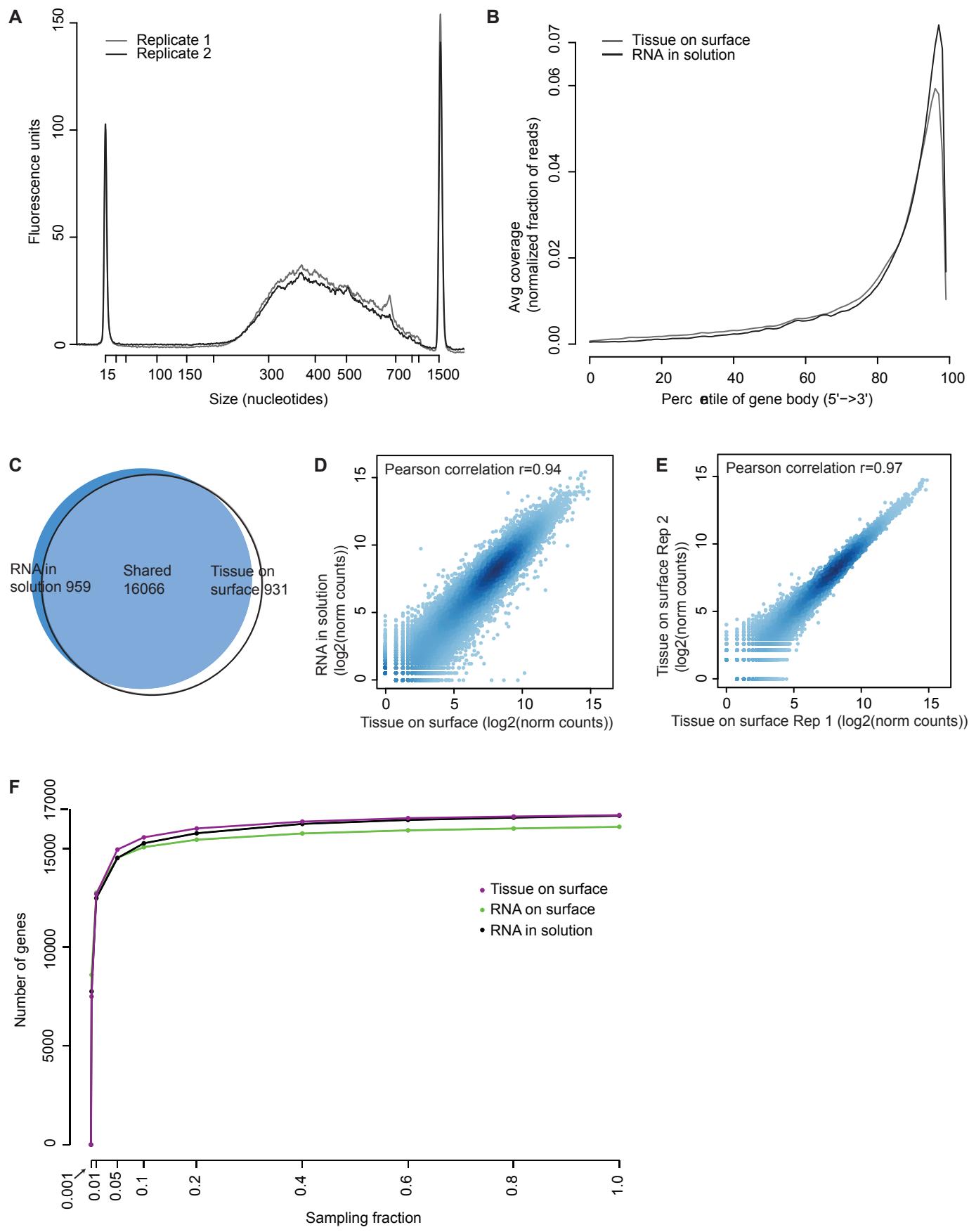
**(G)** As an example, signal intensities for cell 2 in both channels with depicted distance ( $d$ ) at 10% of Cy3-signal intensity, is shown.

**(H)** Histogram showing the observed distances, where Cy3-signal can be detected outside the cells.



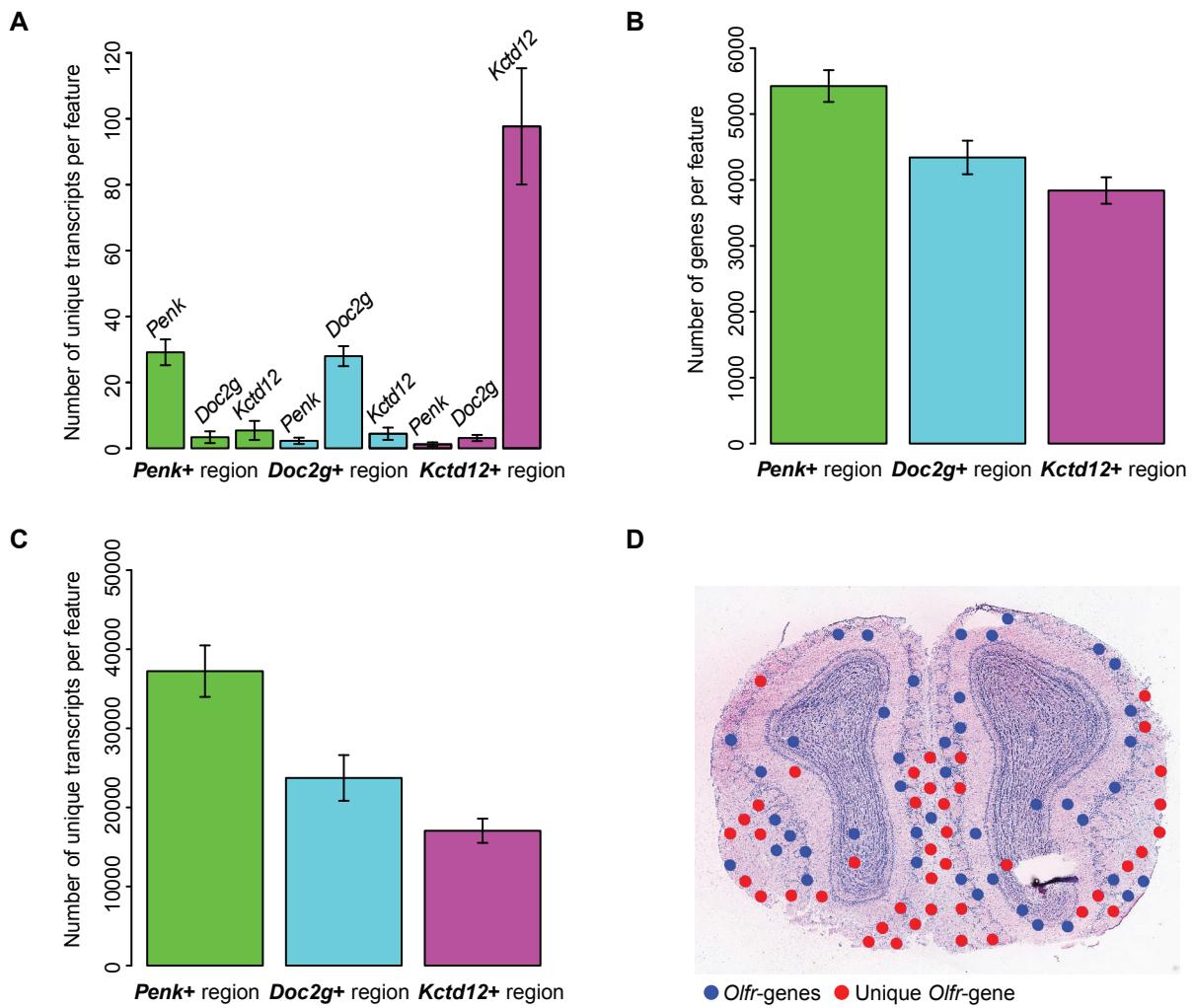
**Fig. S2. Spatially barcoded array**

Layout of barcoded capture probes on the spatially barcoded array overlaid with a tissue section. The diameter of each spot is 100µm and the spot center to center distance is 200µm.



**Fig. S3. Comparison of cDNA libraries generated in solution or on the surface of a spatially barcoded array**

- (A)** Bioanalyzer DNA 1000 (Agilent) trace, showing average size of libraries generated from cDNA synthesized from tissue on a spatial array.
- (B)** Read coverage over an average “meta gene” shows a distinct peak at the 3'-end. Read coverage from an RNA in solution library and a tissue on array library are overlaid.
- (C)** Venn diagram of genes detected in extracted total RNA in solution and a tissue section processed on the surface of an array.
- (D)** Scatterplot of gene expression for an in-solution and an array sample. The samples show an even representation of highly and lowly expressed genes.
- (E)** Scatterplot of gene expression between two replicates processed on an array.
- (F)** Diversity of transcriptome data from extracted total RNA from tissue processed in solution, tissue section processed on spatially barcoded array and extracted total RNA from tissue processed on spatially barcoded array.



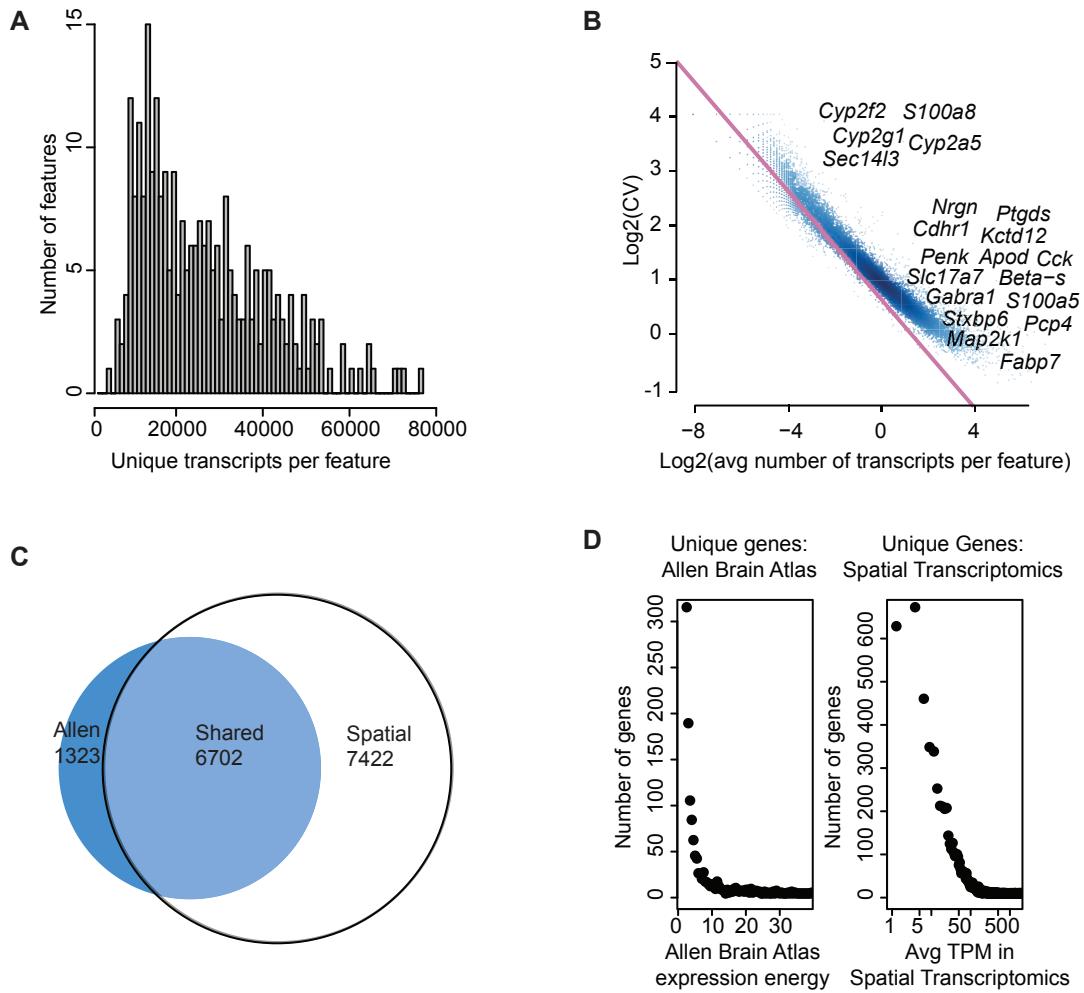
**Fig. S4. Expression and distribution of *Penk*, *Doc2g*, *Kctd12* and *Olfr*-genes**

**(A)** Average expression and distribution of *Penk*, *Doc2g* and *Kctd12* in regions that are enriched for the same threeee genes as shown in **Fig. 2B**. Error bars represent standard deviation.

**(B)** Average number of genes in each of the three regions enriched for *Penk*, *Doc2g* or *Kctd12* as shown in **Fig. 2B**. Error bars represent standard deviation.

**(C)** Average number of unique transcripts in each of the three regions enriched for *Penk*, *Doc2g* or *Kctd12* as shown in **Fig. 2B**. Error bars represent standard deviation.

**(D)** Distribution of features with detected lowly expressed olfactory receptor transcripts (*Olfr*). Blue denotes a feature where *Olfr*-genes are detected. Red denotes a feature where a unique *Olfr*-gene to this tissue sample is detected. In total 77 unique *Olfr* transcript types were detected across 57 different features.



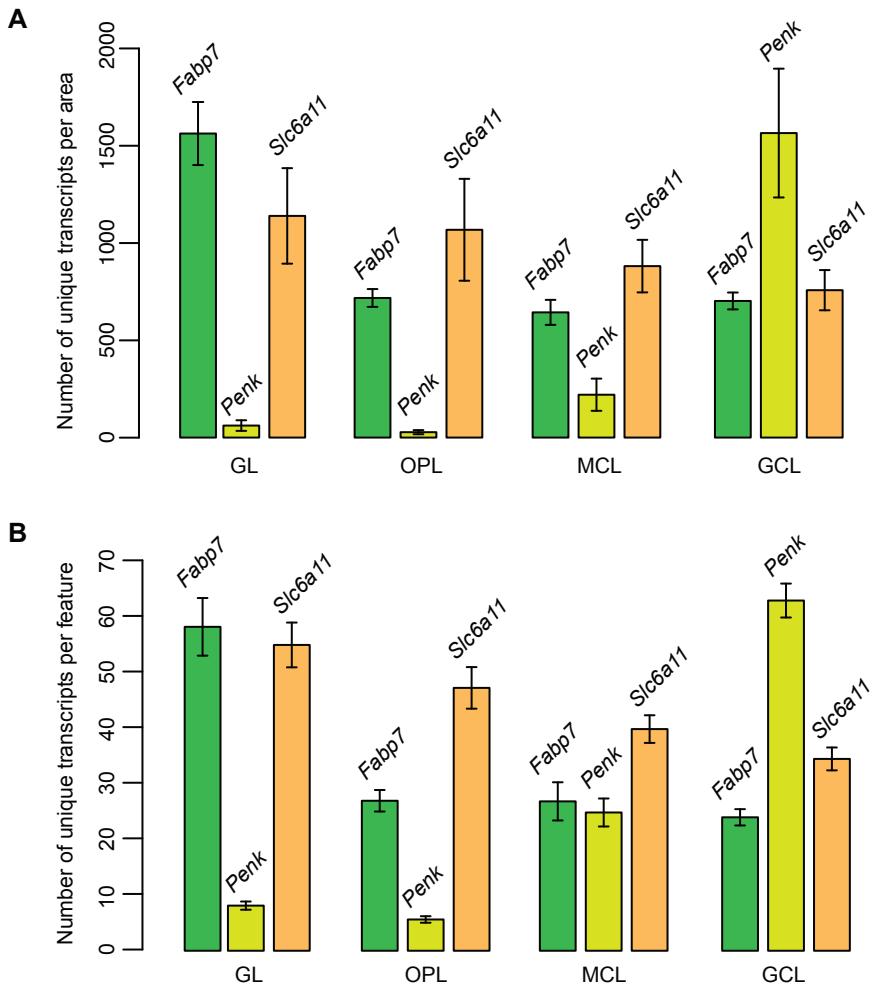
**Fig. S5. Capture of gene expression data by Spatial Transcriptomics compared to ISH database**

**(A)** Distribution of unique transcripts captured per feature (UMI-filtered) under the tissue during an experiment on the Spatially barcoded array, in which 273 features were covered by tissue.

**(B)** CV over average expression per feature for UMI-filtered transcripts. Red line shows theoretical Poisson distribution. Example genes with divergent variation are labeled.

**(C)** Comparison of genes detected in Allen Brain Atlas (ABA) ISH data and Spatial Transcriptomics RNA-seq data. Using the default expression cutoff (for differential expression analysis) of 1, for the ABA dataset and a cutoff of 10 unique transcripts per gene (corresponding to approximately 1 transcript per million mapped) for Spatial Transcriptomics RNA-seq data, the Spatial Transcriptomics RNA-seq data detects almost twice as many genes. Of the 1323 genes detected in mouse olfactory bulb in the ABA but not with Spatial Transcriptomics, 253 could not be identified in RefSeq, used for annotation in Spatial Transcriptomics, despite searching for gene aliases via information in the Allen API, or mapping ABA probe sequences (also retrieved from the Allen API) to RefSeq mRNA annotations. Of the remaining 1070 ABA-specific genes that could be identified, 630 had no evidence of expression (no observed reads) in Spatial Transcriptomics, and 440 had expression below the defined threshold. In terms of biological function, transcription factor genes (especially homeobox genes) and immune genes were overrepresented in this set of 1070 ABA-specific genes.

**(D)** Number of genes that are detected with one of the method but not with the other, against the expression values. Most of the genes that are detected using only one of the technologies are weakly detected with the other technology; ~90% of the genes that are not detected with Spatial Transcriptomics have an expression energy below 10 in ABA, while ~85% of the genes not detected in ABA have a mean TPM below the overall mean TPM in Spatial Transcriptomics.



**Fig. S6. Count data for smFISH and Spatial Transcriptomics**

(A) The average number of unique transcripts, for *Penk*, *Fabp7* and *Slc6a11*, detected in GL, OPL, MCL and GCL with smFISH.

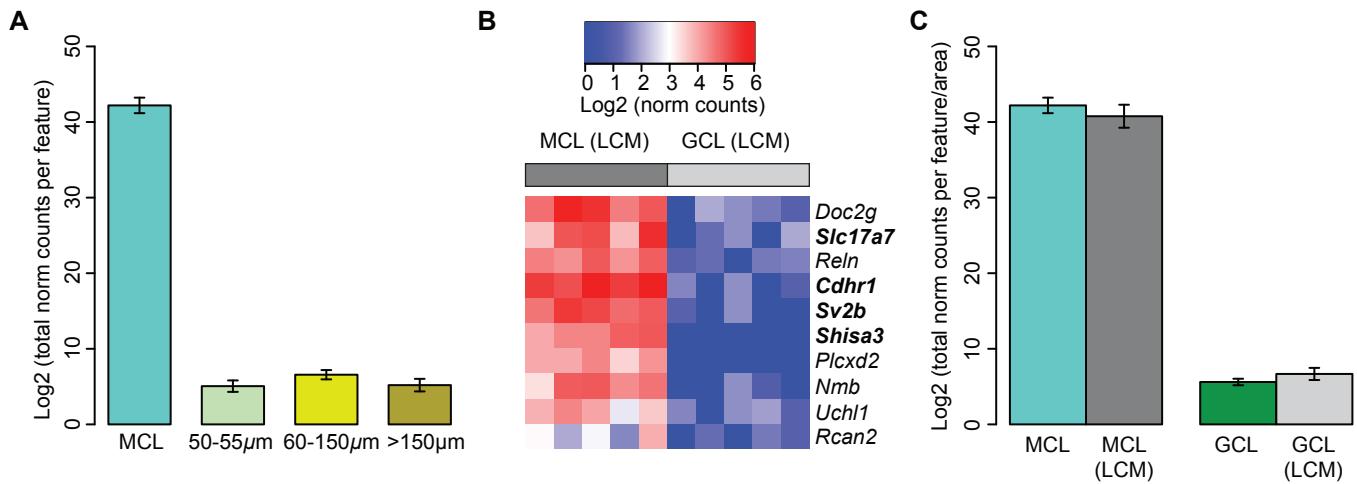
Error bars represent standard deviation.

(B) The average number of unique transcripts, for *Penk*, *Fabp7* and *Slc6a11*, detected in GL, OPL, MCL and GCL with Spatial

Transcriptomics. The expression patterns show similarity to

smFISH (A) with a sensitivity of  $6.9 \pm 1.5\%$ . Error bars represent

standard deviation.

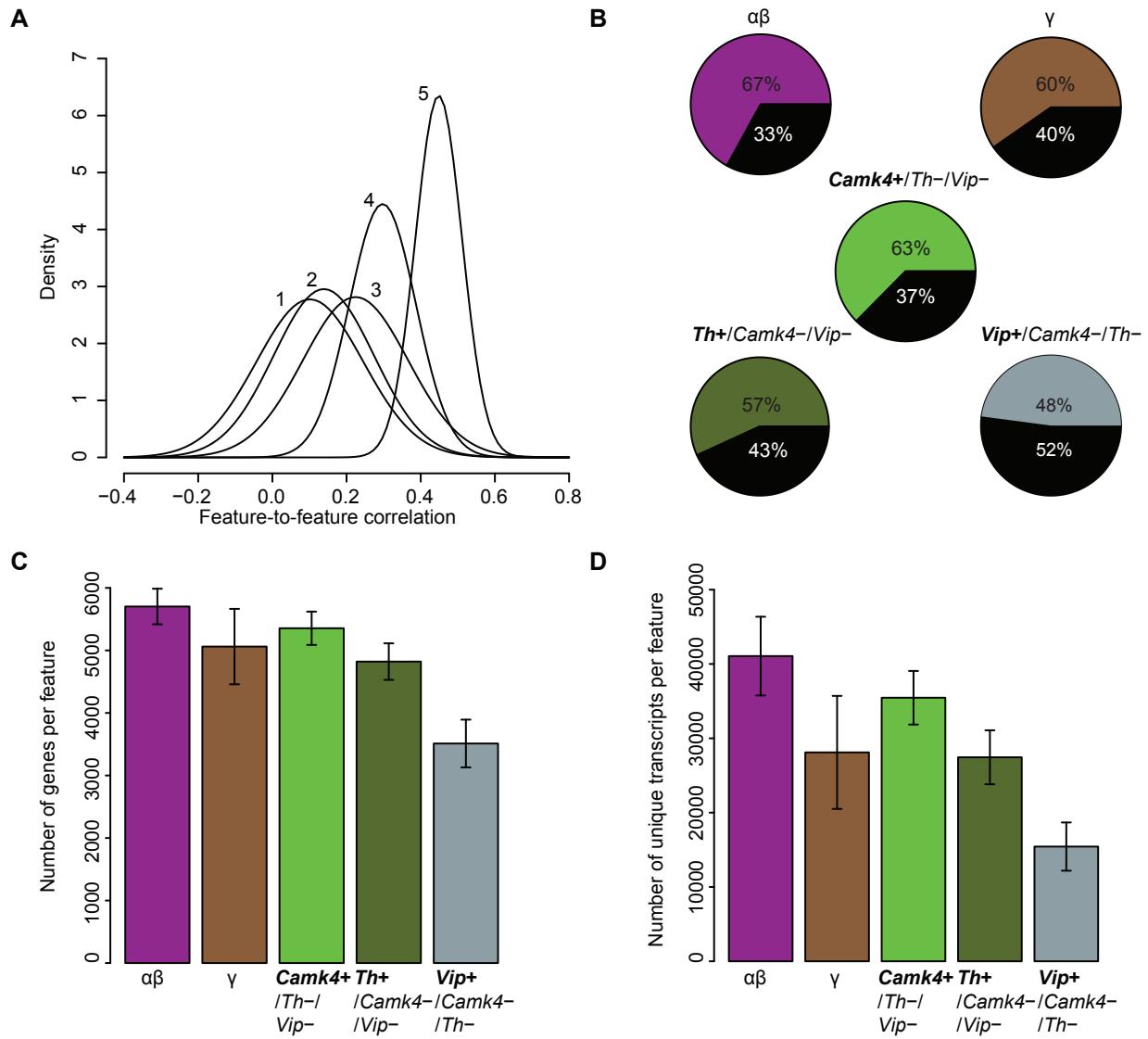


**Fig. S7. Expression and distribution of layer-specific marker-genes**

**(A)** The average sum of unique transcripts, for all ten genes (Fig. 2E), per feature in MCL and in GCL at various distances from MCL. Data shows very little expression in GCL compared to MCL with no significant difference in expression in relation to distance from MCL. Error bars represent standard deviation.

**(B)** LCM data was obtained from round areas with a diameter of  $\sim 100\mu\text{m}$ . Heatmap, from these experiments, showing expression of genes enriched in MCL. The genes are highly expressed in MCL, but show non-or low expression in GCL with similar distribution as detected with Spatial Transcriptomics (Fig. 2E).

**(C)** The average sum of unique transcripts, for all ten layer-specific marker-genes, per feature or area in MCL and GCL, detected with Spatial Transcriptomics (Fig. 2E) or LCM (A). The expression profiles are very similar using both methods with very few transcripts detected in GCL as compared to MCL. Error bars represent standard deviation.



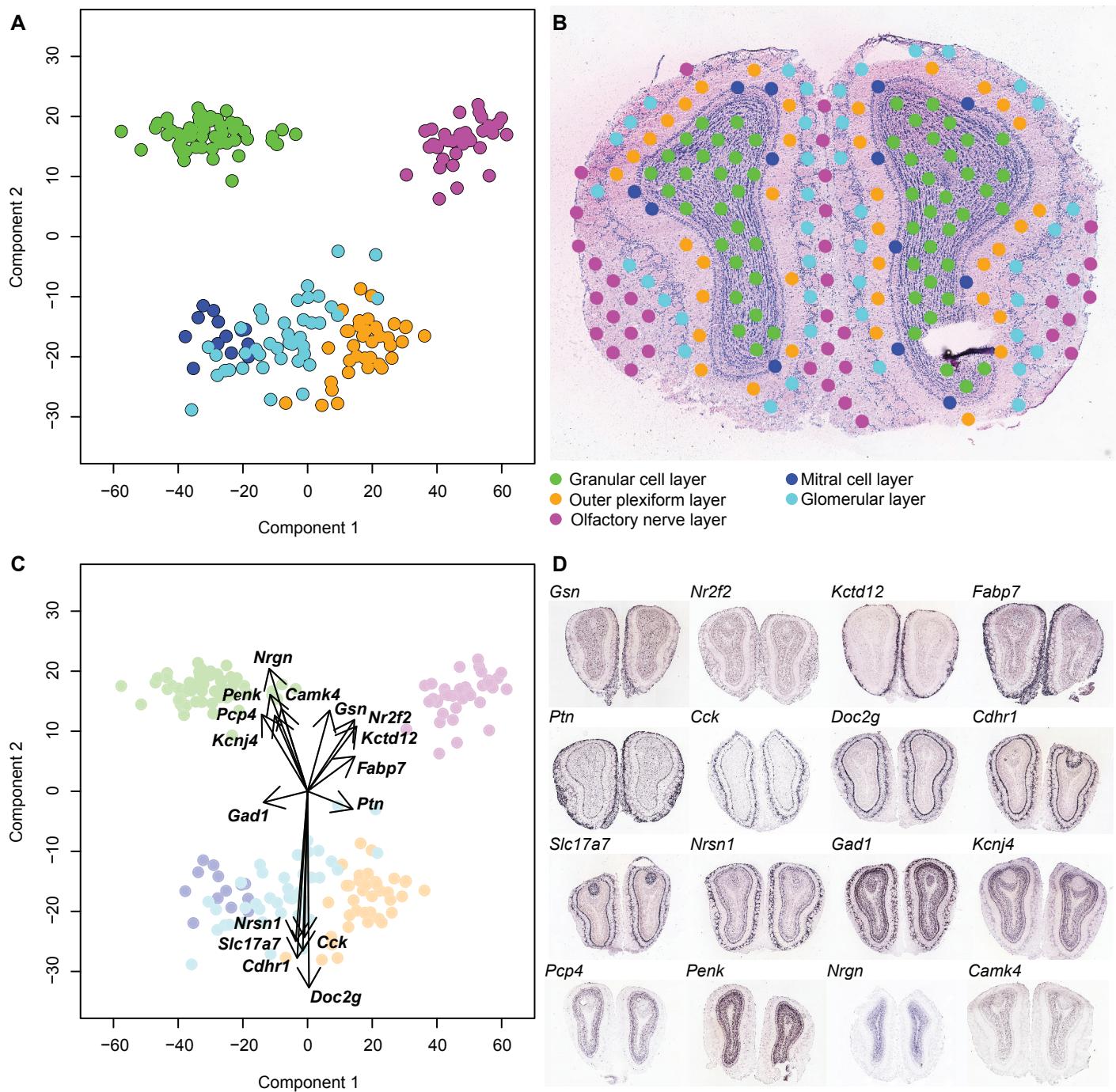
**Fig. S8. Feature-to-feature correlation and gene detection for different regions**

(A) Normal distribution of Pearson correlations between individual features within the same morphological region or between features within different morphological or marker-based regions. Correlations; 1 (*Vip+/Camk4-/Th-* to *Th+/Camk4-/Vip-*):  $r=0.10 \pm 0.14$ , 2 (*Camk4+/Vip-/Th-* to *Vip+/Camk4-/Th-*):  $r=0.14 \pm 0.14$ , 3 (*Camk4+/Vip-/Th-* to *Th+/Camk4-/Vip-*):  $r=0.22 \pm 0.14$ , 4 ( $\alpha$  to  $\gamma$ ):  $r=0.30 \pm 0.06$ , 5 ( $\alpha$  to  $\beta$ ):  $r=0.45 \pm 0.09$ .

(B) The average percentage (colored parts) of genes detected in both of two individual features from the same morphological or marker-based regions. The black parts show the average percentage of genes that are not detected in both features.

(C) Average number of genes in each of the five regions as shown in Fig. 3A, D. Error bars represent standard deviation.

(D) Average number of unique transcripts in each of the five regions as shown in Fig. 3A, D. Error bars represent standard deviation.



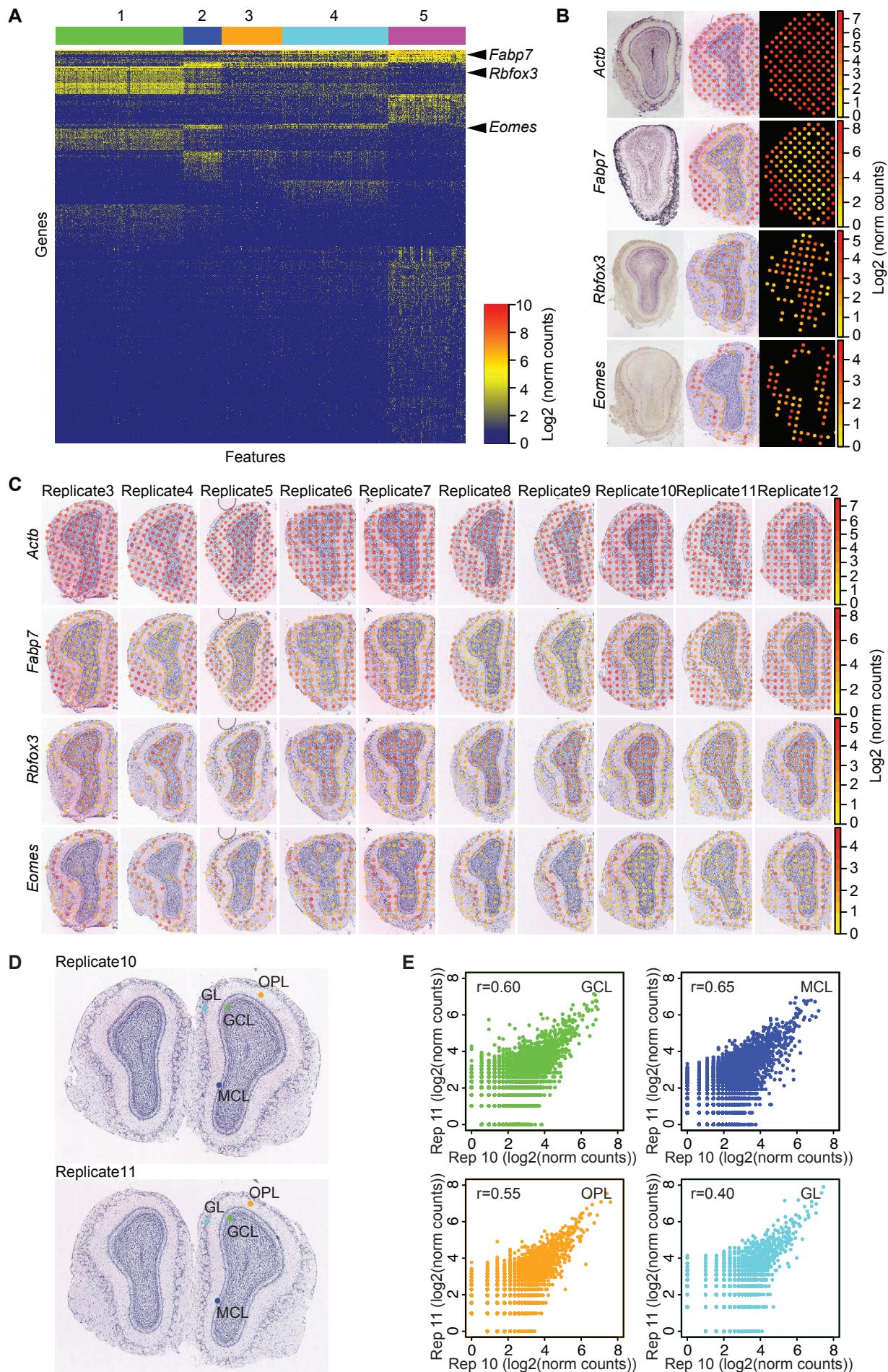
**Fig. S9. Principal component analysis of tissue domains**

**(A)** A principal component analysis (PCA) of gene expression profiles from 186 features taken from five regions under the tissue. Each dot is colored accordingly to which region it was selected from.

**(B)** The position of each feature that was selected for PCA.

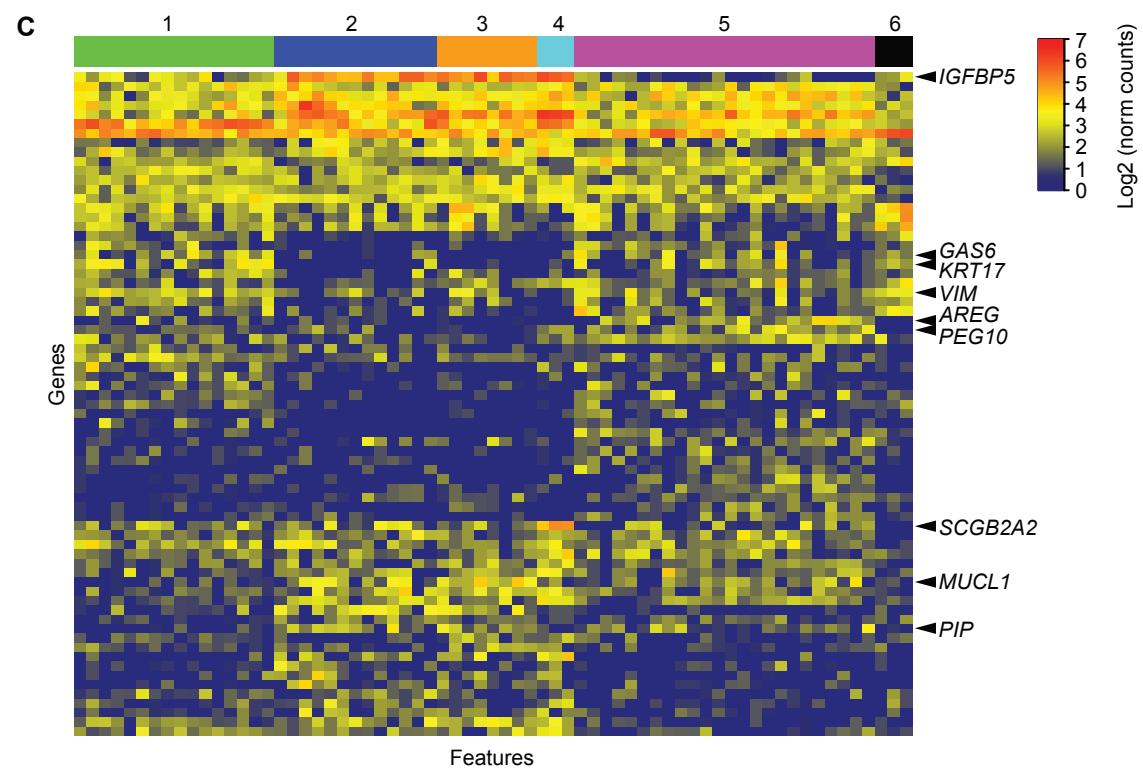
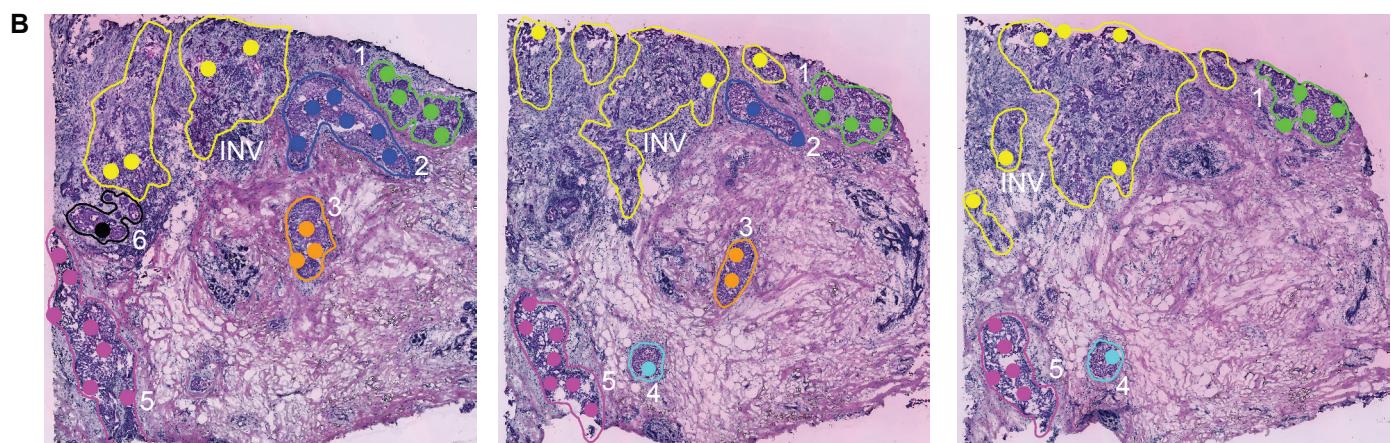
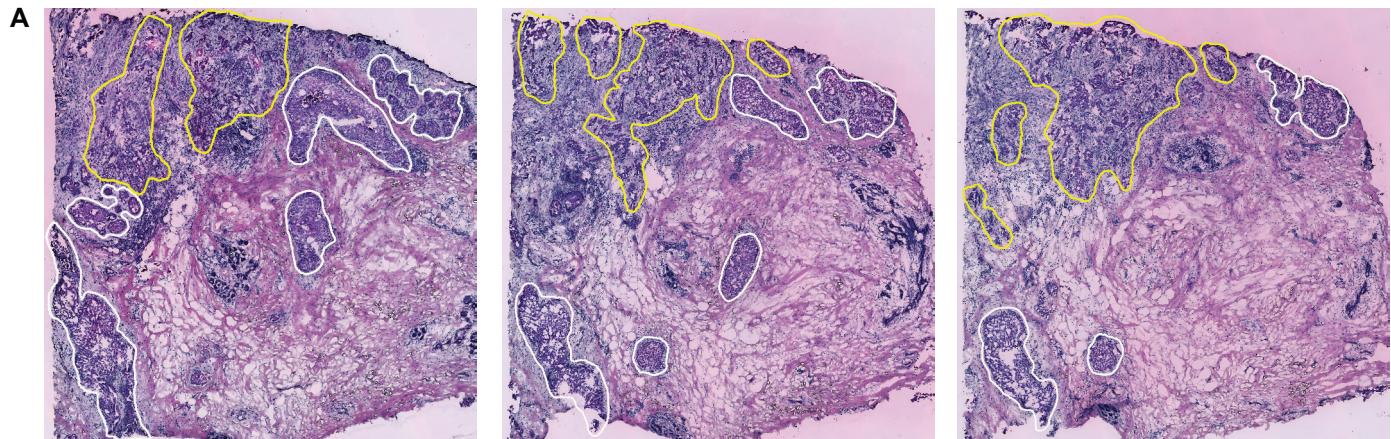
**(C)** PCA bi-plot showing some of the genes (loadings) that contribute the most to the separation.

**(D)** ISH images of the genes detected as driving the separation in the PCA. ISH images are from the Allen Institute.



**Fig. S10. Expression patterns of cluster-specific marker-genes**

- (A)** Heatmap showing expression patterns of 532 cluster-specific marker-genes from the tissue sections shown in **Fig 4B**.
- (B)** Spatially resolved expression profiles of cluster-specific marker-genes, *Fabp7*, *Rbfox3* and *Eomes* identified in **(A)**, as well as house keeper *Actb* (the two right panels) and ISH-images from the Allen Brain Atlas as reference (the left panel). The spatial expression patterns show high similarity between Spatial Transcriptomics and ISH.
- (C)** Spatially resolved expression profiles in ten tissue sections (replicates) taken from five different animals. The spatial expression patterns show high similarity between replicates.
- (D)** The position of four selected features, from different morphological regions (GCL, MCL, OPL and GL), in each of two adjacent tissue sections (Replicate 10 and 11) that are located at the approximate same positions in each of the sections.
- (E)** Scatterplot with Pearson correlations between features shown in **(D)**. The correlations range from 0.40-0.65 between individual features located at the approximate same positions in each of the two adjacent sections.



**Fig. S11. Adjacent sections from breast cancer biopsy**

- (A) Three more histological sections from the same breast cancer biopsy as shown in **Fig. 4C**, with invasive ductal cancer (INV) and six separate areas of ductal cancer in situ (1-6).
- (B) Features, in the three sections shown in (A), selected for Spatial Transcriptomics analysis. Only features in the INV area without or with minimal stromal infiltration were selected.
- (C) Heatmap of differentially expressed genes between the six ductal cancer in situ areas shown in (B) and in **Fig. 4D**. The revealed expression patterns show a remarkable difference in genes expression profiles between areas of cancer in situ.

**Table S1**

	mRNA and lncRNA	rRNA	mtRNA	Pseudogenes, repeats, artifacts	Unique transcripts after duplicate removal
On surface	23%	23%	24%	30%	9.6M (of 400M reads)
In solution	34%	6%	17%	43%	18 M (of 290M reads)

**Table S1.** Statistics from sequencing a library generated through cDNA synthesis from a tissue section on the surface of a spatially barcoded array and a library generated through cDNA synthesis of extracted RNA in solution.

**Table S2**

	Average genes	Average transcripts
Granular cell layer	5390 +/- 1260	37400 +/- 14100
Mitral cell layer	6070 +/- 858	45400 +/- 16400
Outer plexiform layer	3640 +/- 648	17900 +/- 5850
Glomerular layer	5680 +/- 1030	38000 +/- 15800
Olfactory nerve layer	3390 +/- 763	14200 +/- 5310

**Table S2.** Average number of genes and unique transcripts captured from different cell layers of the adult mouse olfactory bulb through cDNA synthesis from a tissue section on the surface of a spatially barcoded array.

## References and Notes

1. Z. Wang, M. Gerstein, M. Snyder, RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63 (2009). [Medline doi:10.1038/nrg2484](#)
2. N. Crosetto, M. Bienko, A. van Oudenaarden, Spatially resolved transcriptomics and beyond. *Nat. Rev. Genet.* **16**, 57–66 (2015). [Medline doi:10.1038/nrg3832](#)
3. R. Satija, J. A. Farrell, D. Gennert, A. F. Schier, A. Regev, Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **33**, 495–502 (2015). [Medline doi:10.1038/nbt.3192](#)
4. P. A. Combs, M. B. Eisen, Sequencing mRNA from cryo-sliced *Drosophila* embryos to determine genome-wide spatial patterns of gene expression. *PLOS ONE* **8**, e71820 (2013). [Medline doi:10.1371/journal.pone.0071820](#)
5. K. Achim, J. B. Pettit, L. R. Saraiva, D. Gavriouchkina, T. Larsson, D. Arendt, J. C. Marioni, High-throughput spatial mapping of single-cell RNA-seq data to tissue of origin. *Nat. Biotechnol.* **33**, 503–509 (2015). [Medline doi:10.1038/nbt.3209](#)
6. Materials and methods are available as supplementary materials on *Science* Online.
7. T. Hashimshony, F. Wagner, N. Sher, I. Yanai, CEL-Seq: Single-cell RNA-Seq by multiplexed linear amplification. *Cell Reports* **2**, 666–673 (2012). [Medline doi:10.1016/j.celrep.2012.08.003](#)
8. R. N. Van Gelder, M. E. von Zastrow, A. Yool, W. C. Dement, J. D. Barchas, J. H. Eberwine, Amplified RNA synthesized from limited quantities of heterogeneous cDNA. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 1663–1667 (1990). [Medline doi:10.1073/pnas.87.5.1663](#)
9. R. Vassar, S. K. Chao, R. Sitcheran, J. M. Nuñez, L. B. Vosshall, R. Axel, Topographic organization of sensory projections to the olfactory bulb. *Cell* **79**, 981–991 (1994). [Medline doi:10.1016/0092-8674\(94\)90029-9](#)
10. K. D. Pruitt, G. R. Brown, S. M. Hiatt, F. Thibaud-Nissen, A. Astashyn, O. Ermolaeva, C. M. Farrell, J. Hart, M. J. Landrum, K. M. McGarvey, M. R. Murphy, N. A. O’Leary, S. Pujar, B. Rajput, S. H. Rangwala, L. D. Riddick, A. Shkeda, H. Sun, P. Tamez, R. E. Tully, C. Wallin, D. Webb, J. Weber, W. Wu, M. DiCuccio, P. Kitts, D. R. Maglott, T. D. Murphy, J. M. Ostell, RefSeq: An update on mammalian reference sequences. *Nucleic Acids Res.* **42** (D1), D756–D763 (2014). [Medline doi:10.1093/nar/gkt1114](#)
11. S. Zechel, P. Zajac, P. Lönnberg, C. F. Ibáñez, S. Linnarsson, Topographical transcriptome mapping of the mouse medial ganglionic eminence by spatially resolved RNA-seq. *Genome Biol.* **15**, 486 (2014). [Medline](#)
12. D. Grün, A. van Oudenaarden, Design and analysis of single-cell sequencing experiments. *Cell* **163**, 799–810 (2015). [Medline doi:10.1016/j.cell.2015.10.039](#)
13. E. S. Lein, M. J. Hawrylycz, N. Ao, M. Ayres, A. Bensinger, A. Bernard, A. F. Boe, M. S. Boguski, K. S. Brockway, E. J. Byrnes, L. Chen, L. Chen, T.-M. Chen, M. C. Chin, J. Chong, B. E. Crook, A. Czaplinska, C. N. Dang, S. Datta, N. R. Dee, A. L. Desaki, T. Desta, E. Diep, T. A. Dolbeare, M. J. Donelan, H.-W. Dong, J. G. Dougherty, B. J. Duncan, A. J. Ebbert, G. Eichele, L. K. Estin, C. Faber, B. A. Facer, R. Fields, S. R.

Fischer, T. P. Fliss, C. Frenzley, S. N. Gates, K. J. Glattfelder, K. R. Halverson, M. R. Hart, J. G. Hohmann, M. P. Howell, D. P. Jeung, R. A. Johnson, P. T. Karr, R. Kawal, J. M. Kidney, R. H. Knapik, C. L. Kuan, J. H. Lake, A. R. Laramee, K. D. Larsen, C. Lau, T. A. Lemon, A. J. Liang, Y. Liu, L. T. Luong, J. Michaels, J. J. Morgan, R. J. Morgan, M. T. Mortrud, N. F. Mosqueda, L. L. Ng, R. Ng, G. J. Orta, C. C. Overly, T. H. Pak, S. E. Parry, S. D. Pathak, O. C. Pearson, R. B. Puchalski, Z. L. Riley, H. R. Rockett, S. A. Rowland, J. J. Royall, M. J. Ruiz, N. R. Sarno, K. Schaffnit, N. V. Shapovalova, T. Sivisay, C. R. Slaughterbeck, S. C. Smith, K. A. Smith, B. I. Smith, A. J. Sodt, N. N. Stewart, K. R. Stumpf, S. M. Sunkin, M. Sutram, A. Tam, C. D. Teemer, C. Thaller, C. L. Thompson, L. R. Varnam, A. Visel, R. M. Whitlock, P. E. Wohnoutka, C. K. Wolkey, V. Y. Wong, M. Wood, M. B. Yaylaoglu, R. C. Young, B. L. Youngstrom, X. F. Yuan, B. Zhang, T. A. Zwingman, A. R. Jones, Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445**, 168–176 (2007). [Medline doi:10.1038/nature05453](#)

14. A. Mahfouz, M. van de Giessen, L. van der Maaten, S. Huisman, M. Reinders, M. J. Hawrylycz, B. P. Lelieveldt, Visualizing the spatial gene expression organization in the brain through non-linear similarity embeddings. *Methods* **73**, 79–89 (2015). [Medline](#)
15. L. J. P. van der Maaten, G. E. Hinton, Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
16. M. Kittaneh, A. J. Montero, S. Glück, Molecular profiling for breast cancer: a comprehensive review. *Biomarkers Cancer* **5**, 61–70 (2013).
17. C. Gjerdrum, C. Tiron, T. Høiby, I. Stefansson, H. Haugen, T. Sandal, K. Collett, S. Li, E. McCormack, B. T. Gjertsen, D. R. Micklem, L. A. Akslen, C. Glackin, J. B. Lorens, Axl is an essential epithelial-to-mesenchymal transition-induced regulator of breast cancer metastasis and patient survival. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 1124–1129 (2010). [Medline doi:10.1073/pnas.0909333107](#)
18. S. Lundin, H. Stranneheim, E. Pettersson, D. Klevebring, J. Lundeberg, Increased throughput by parallelization of library preparation for massive sequencing. *PLOS ONE* **5**, e10029 (2010). [Medline](#)
19. N. Schaeren-Wiemers, A. Gerfin-Moser, A single protocol to detect transcripts of various types and expression levels in neural tissue and cultured cells: In situ hybridization using digoxigenin-labelled cRNA probes. *Histochemistry* **100**, 431–440 (1993). [Medline doi:10.1007/BF00267823](#)
20. A. Lyubimova, S. Itzkovitz, J. P. Junker, Z. P. Fan, X. Wu, A. van Oudenaarden, Single-molecule mRNA detection and counting in mammalian tissue. *Nat. Protoc.* **8**, 1743–1758 (2013). [Medline doi:10.1038/nprot.2013.109](#)
21. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012). [Medline doi:10.1038/nmeth.1923](#)
22. S. Anders, P. T. Pyl, W. Huber, HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015). [Medline doi:10.1093/bioinformatics/btu638](#)

23. P. I. Costea, J. Lundeberg, P. Akan, TagGD: Fast and accurate software for DNA Tag generation and demultiplexing. *PLOS ONE* **8**, e57521 (2013). [Medline](#)  
[doi:10.1371/journal.pone.0057521](https://doi.org/10.1371/journal.pone.0057521)
24. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014). [Medline](#) [doi:10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8)
25. L. Wang, S. Wang, W. Li, RSeQC: Quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184–2185 (2012). [Medline](#) [doi:10.1093/bioinformatics/bts356](https://doi.org/10.1093/bioinformatics/bts356)