A spatial atlas of inhibitory cell types in mouse hippocampus

Xiaoyan Qian*[1], Kenneth D. Harris*†[2], Thomas Hauling[1,2], Dimitris Nicoloutsopoulos[2], Ana B. Muñoz-Manchado[3], Nathan Skene[2,3], Jens Hjerling-Leffler[3], Mats Nilsson†[1]

1. Science for Life Laboratory, Department of Biochemistry and Biophysics, Stockholm University, Sweden
2. Institute of Neurology, University College London, UK
3. Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden

*Equal contribution.
† Correspondence: kenneth.harris@ucl.ac.uk; mats.nilsson@scilifelab.se

Understanding the function of a tissue requires knowing the spatial organization of its constituent cell types. In the cerebral cortex, single-cell RNA sequencing (scRNA-seq) has revealed the genome-wide expression patterns that define its many, closely related cell types, but cannot reveal their spatial arrangement. Here we introduce *probabilistic cell typing by in situ sequencing* (pciSeq), an approach that leverages prior scRNA-seq classification to identify cell types using multiplexed *in situ* RNA detection. We applied this method to map the inhibitory neurons of hippocampal area CA1, a cell system critical for memory function, for which ground truth is available from extensive prior work identifying the laminar organization of subtly differing cell types. Our method confidently identified 16 interneuron classes, in a spatial arrangement closely matching ground truth. This method will allow identifying the spatial organization of fine cell types across the brain and other tissues.

The cortex contains many cell types, which differ in their spatial organization, morphology, connectivity, physiology, and gene expression. Although the diversity of cortical cells was known to classical neuroanatomists, the true complexity of cortical cells has only become clear since the development of transcriptomics[1–6]. Many cell types previously thought to be homogeneous in fact contain multiple fine subclasses, which share expression of the great majority of their genes, but are distinguished by relatively small sets of genes. Hippocampal area CA1 contains at least 20 subtypes of molecularly-distinct inhibitory interneurons, which are arranged in a stereotyped spatial organization[7–9], and recordings have shown that even closely related cell types have different *in vivo* functions[10]. Analysis of CA1 inhibitory classes by scRNA-seq yields clusters strikingly consistent with these classically-defined types[5]. Mapping the spatial organization of CA1 interneurons is thus not only important to understand the brain's memory circuits, but also provides a powerful way to validate spatial cell type mapping approaches, using the spatio-molecular ground truth provided by this system.

Here we provide a spatial map of CA1 interneuron types, using a new approach to *in situ* cell typing based on *in situ* RNA expression profiling. While several approaches to multiplexed *in situ* RNA detection and cell type classification have been proposed[11,12] (e.g. *in situ* sequencing[13], FISSEQ[14], SNAIL[15], MERFISH[16], CorrFISH[17], osmFISH[18] and seqFISH[19]), none have yet shown the ability to distinguish fine cell types. The current method achieves this using three design criteria suggested by scRNA-seq analysis. First, for typing of closely related cells such as CA1 inhibitory neurons, of the order 100 genes suffices to classify cells accurately[5]. Second, false positive reads present much more of a problem for cell typing than low detection efficiency. Indeed, because RNA expression levels follow a negative binomial distribution[20] (whose standard deviation scales with its mean), detecting just a few copies of a genuinely-expressed gene can confidently identify cell types, provided one can be sure these are not misdetections. Third, using the cell types revealed by previous large-scale scRNA-seq analysis to define types avoids needing to relearn cell classes from noisier *in situ* data. We show that this combination allows cell typing of closely-related neuronal classes, verified by the ground truth available from CA1's laminar architecture.

Our cell typing method consists of three steps (**Supplementary Figure S1**). First, we select a set of marker genes sufficient for identifying cell types, using previous scRNA-seq data. Second, we apply *in situ* sequencing to detect expression of these genes at cellular resolution in tissue sections. Third, gene reads are assigned to cells, and cells to types using a probabilistic model derived from scRNA-seq clusters.

To select a gene panel, we developed an algorithm that searches for a subset of genes that can together identify scRNA-seq cells to their original clusters, after downsampling expression levels to match the lower efficiency of *in situ* data (see Methods). The gene panel was selected using a database of inhibitory neurons from hippocampus[5] (**Supplementary Figure S2**) as well as isocortex[2], and the results were manually curated prior to final gene selection. The algorithm identified familiar interneuron markers such as *Sst, Vip, Npy,* and *Cck*, but also many genes identified only by transcriptomic analysis (e.g. *Cxcl14*, *Ntng1*, *Id2*). The final gene sets were chosen by manual curation of these lists, excluding genes that were likely to be strongly expressed in all cell types even if at different levels, and favoring genes which have been used in classical immunohistochemistry (**Supplementary Table S1, Supplementary Figure S3**). A further set of three genes were excluded after initial experiments, as their expression was widespread in neuropil and did not help identify cell types (*Slc1a2*, *Vim*, *Map2*). The final panel contained 99 genes.

To generate RNA expression profiles, we modified the barcode-targeted *in situ* sequencing method described by Ke *et al*[13] (**Supplementary Figure S4**). A library of padlock probes was generated to match the selected genes. For each gene, we selected a set of target sequences along the exon sequence, such that the two arms of each padlock probe together match a 40-basepair sequence along on the mRNA. In addition to the recognition arms, each padlock probe contained a 4-basepair barcode, an "anchor sequence" allowing all amplicons to be labelled simultaneously, and a 20-basepair hybridization sequence allowing for additional readouts. For more weakly
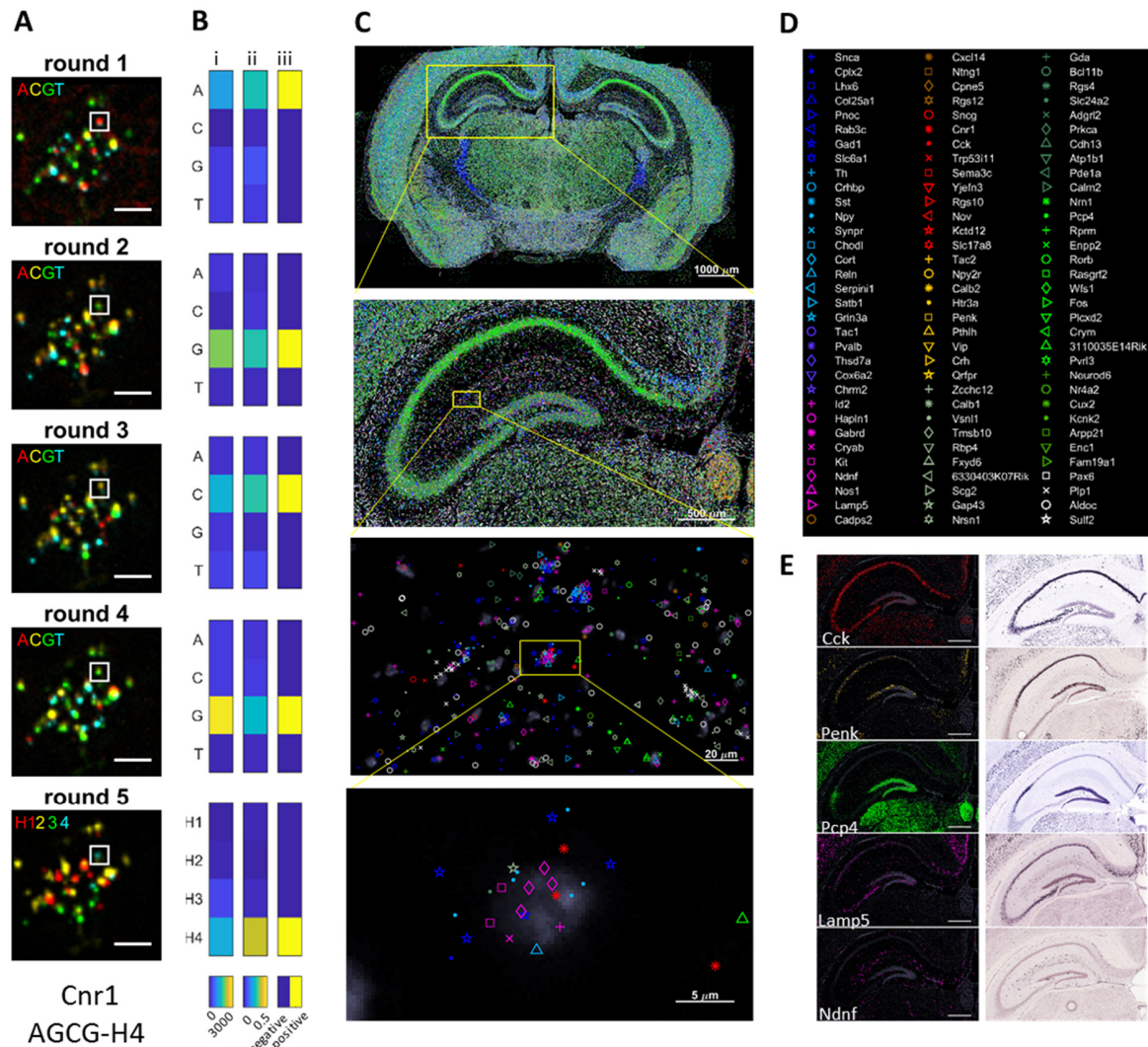
2

Figure 1. Molecular maps. **A)** Pseudocolor images showing barcode sequencing readout for a region corresponding to one cell. Top to bottom, base-specific fluorophores in the four cycles of sequencing by ligation, and for the fifth cycle of barcode specific hybridization. The white square shows a single RCP of barcode AGCG-H4. Scale bars: 5 µm. **B)** Gene-calling for this RCP. Left: pseudocolor representation of raw fluorescence intensities; Middle, intensity after crosstalk compensation; Right, best fit barcode (AGCG-H4, encoding the gene *Cnr1*). **C)** Distribution of 99 genes at different zoom levels. From top to bottom: a complete coronal mouse brain section; left hippocampus; the border of stratum radiatum and stratum lacunosum moleculare; finally, zoom-in to reads for the cell whose raw fluorescence is shown in panel (A). **D)** Code symbols for the 99 marker genes. **E)** Comparison of the distribution of five markers in the hippocampus as determined by pciSeq (left column) with the distribution shown in the Allen Mouse Brain Atlas (right column). Scale bars: 500 µm.

expressed genes, we designed probes matching multiple target sequences along the mRNA length, which aided their detection without compromising detection of others (**Supplementary Figure S5**). In total we designed 755 probes for 99 genes, but used only 161 barcodes out of 1024 ($=4^5$) possible combinations to allow error correction (for probe sequence and barcodes see **Supplementary Table S2)**.

3

To apply the method to a tissue section, mRNA is enzymatically converted to cDNA *in situ* and then degraded. The padlock probe library is applied, and a ligation enzyme selectively circularizes only probes perfectly matching their cDNA target sequences. The circularized probes are rolling-circle amplified (RCA), generating sub-micron sized DNA molecules (rolling-circle products: RCPs), each carrying hundreds of copies of the probe's barcode. The RCP barcodes are identified with a standard fluorescence microscope (20x objective) in five rounds of multi-color imaging (**Figure 1A**). Finally, RCPs for two genes which express so strongly their signal would swamp others *(Sst* and *Npy)* are detected separately in a 6th round by hybridizing fluorescent probes to their target recognition sequences. Data is analyzed using a custom pipeline, including point-cloud registration to deal with chromatic aberration in the images, and compensation for optical or chemical crosstalk between bases in the sequencing readout (**Figure 1B; Supplementary Figure S6** and Methods). These improved chemical and analytic methods achieved a density of reads sufficient for fine cell type assignment.

Our first experiments were performed using probes against a subset of 84 genes on four coronal sections of mouse brain (10 µm fresh frozen). After verifying that detected expression patterns match *in situ* hybridization data from the Allen Mouse Brain Atlas[21], we continued with two further experiments using the full 99-gene panel, on two and eight coronal sections, respectively. All fourteen sections were from one P25 male CD1 mouse and covered different parts of the dorsal hippocampus (**Supplementary Figure S7**). Each section contained roughly 120,000 cells and in total 15,424,317 reads passed quality control (**Supplementary Table S3)**. To display reads corresponding to 99 genes on a single map, we displayed each read with symbols whose colors grouped genes often expressed by similar cell types, and glyph distinguished genes within these color groups (**Figure 1, C and D**).

Expression patterns were consistent with expectation at multiple levels of detail. At the whole-brain level (**Figure 1C**, top), differences between regions could be seen, for example with the inhibitory thalamic reticular nucleus dominated by inhibitory-associated genes (blue) and the CA1 pyramidal layer dominated by pyramidal-associated genes (green). Zooming in to the hippocampus (**Figure 1C**, 2nd row) revealed differences between cell layers, for example with stronger expression of genes associated with *Sst* neurons (cyan) in stratum oriens, neurogliaform-associated genes (pink) in stratum lacunosum-moleculare, and non-neuronal genes (white) in the white matter. Zooming further to single neurons (bottom two rows) showed genes that appeared to be grouped together in combinations as expected from scRNA-seq. Expression patterns of all genes present in the Allen Mouse Brain Atlas[21] matched at a corresponding coronal level (examples in **Figure 1E**). Read densities were consistent between experiments, even with different gene panels, further supporting the reliability of the technique (r = 0.93; **Supplementary Figure S8A**).
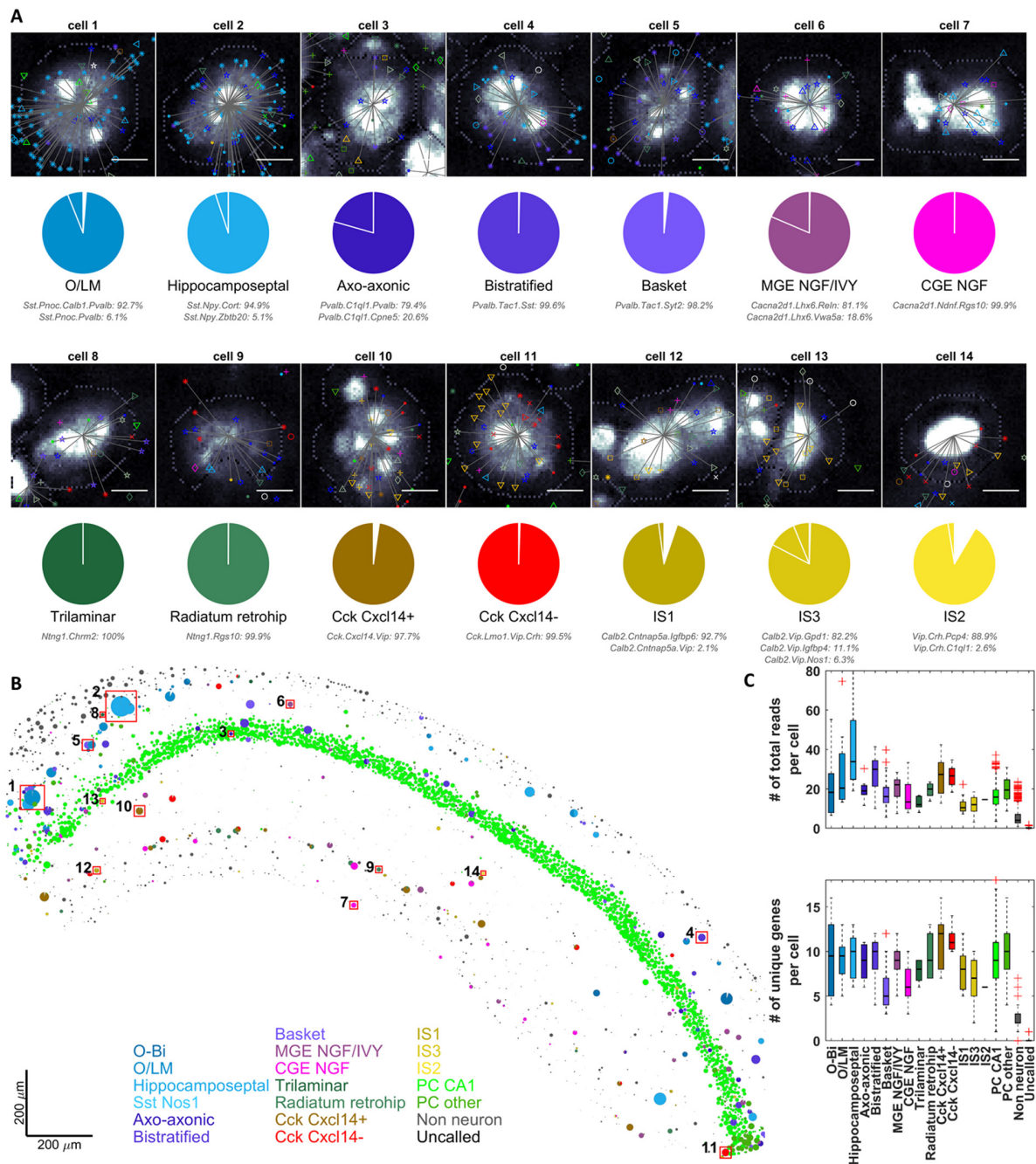
4

Figure 2. Cell type map of CA1 from an example experiment (experiment 4-3 right hemisphere). **A)** Reads are assigned to cells, and cells to classes using a probability model based on scRNA-seq data. Top row: distribution and assignment of reads for fourteen example cells. Colored symbols indicate reads (color code as in **Figure 1D**). Grayscale background image indicates DAPI stain with watershed segmentation as dotted line. Straight lines join reads to the cell for which are assigned highest probability. Scale bars: 5 μm. Bottom row: pie charts showing probability distribution of each class for the same example cells. Colors indicate broad cell types; segments show probabilities for individual scRNA-seq clusters (named underneath). **B)** Spatial map of cell types across CA1. Cells are represented by pie charts, with size proportional to the number of reads assigned to the cell. Numbers identify the example cells in A. **C)** Box-and-whisker representation of total read count per cell of each type (top) and average number of unique genes per cell of each type (bottom).

5

We focused our cell typing analysis on hippocampal area CA1 (**Supplementary Figure S9**), where we can use the known laminar locations of several identified cell types to validate the method. We used the approach to identify the types of 27,338 cells typed as CA1 neurons, from 28 hippocampi. Data files for all experiments are available at https://figshare.com/s/88a0fc8157aca0c6f0e8, and an online viewer showing reads and probabilistic cell type assignments is at http://insitu.cortexlab.net**.**

A fundamental challenge for *in situ* cell typing is assigning genes to cells. mRNA molecules can be found throughout a cell's cytoplasm, and boundaries between cells are difficult to obtain in 2D imaging. In the current approach, we counterstained all sections with DAPI to reveal cellular nuclei. Applying standard watershed segmentation to these cells yielded boundaries that contained many, but not all the genes that manual inspection suggested belonged to them (**Figure 2A**). To solve this problem, we developed an algorithm which leverages scRNA-seq data to assign genes to cells more accurately than would be possible by spatial location alone (**Figure 2A,** straight lines). The model uses scRNA-seq clusters to define an expected read density for each gene in each cell type, and assigns a prior probability of assigning reads to cells based the DAPI image. It employs a variational Bayesian algorithm to compute a posterior probability distribution for the cell assigned to each read, and the type assigned to each cell. Note that the algorithm does not take into account a cell's laminar location, allowing this to be used for validation.

To represent the results on a spatial map, we displayed each cell's class assignments by a pie-chart, of size proportional to total gene count, and with the angle of each color-coded slice indicating the probability of assignment to that class (**Figure 2B;** see also **Supplementary Figure S10;** for all cell type maps, see **Supplementary Appendix;** online viewer at http://insitu.cortexlab.net). The algorithm was able to confidently identify 16 inhibitory classes: 3 types of interneuron-selective cell; 2 types of Cck cell; 2 types of neurogliaform cell; 2 types of GABAergic projection cell; 3 types of parvalbumin cell and 4 types of somatostatin cell (**Supplementary Tables S4 and S5**). Further subdivisions of these types, identified by scRNA-seq clusters, were called with varying levels of confidence**.** Although our panel was primarily aimed at distinguishing inhibitory neurons, we also obtained confident distinction of two types of pyramidal cell, as well as non-neuronal cells. The probabilistic algorithm allows diagnostics which show which genes provided evidence for calling as one type over another (**Supplementary Figure S11**). The average number of gene reads per cell was over 20 for most targeted cell types, and the number of unique genes detected per cell was in the range 5 to 10 (**Figure 2C**).

The algorithm's cell type assignments conformed closely to known combinatorial patterns of gene expression in fine CA1 inhibitory subtypes. For example, the identification of *Sst* cells as O/LM or hippocamposeptal correlated with further expression of *Reln* or *Npy*[22,23] (**Figure 2A**; cells 1,2); identification of *Pvalb* cells as axo-axonic, basket or bistratified correlated with further expression of *Pthlh, Satb1/Tac1*, or *Sst/Npy* [6,22,24] (Cells 3-5)*;* identification of neurogliaform cells as CGE-derived or MGE-derived/Ivy correlated with further expression of *Ndnf/Kit/Cxcl14* or *Lhx6/Nos1* [2,5,25,26]*;* identification of long-range projection neurons as trilaminar or radiatum-

6

retrohippocampal correlated with expression of *Chrm2* or *Ndnf/Reln* [23,27] (Cells 8,9)*; Cck* cells were identified as two subtypes correlated with expression of *Cxcl14,* with both subtypes expressing *Cnr1* and further subdivided by *Vip* expression [5,28,29] (Cells 10-11); finally, inhibitory-selective cells were divided into three classes correlated with the combinatorial expression of *Calb2* and *Vip* [30,31] (Cells 12-14). Across all experiments, the patterns of both classical and novel interneuron markers were consistent with scRNA-seq results, as well as the known biology of CA1 interneurons (**Supplementary Figure S12**). Moreover, the cell type composition was consistent between the left and right hemispheres (**Supplementary Figure S8B**).

As the pciSeq algorithm did not use cell's laminar location to determine its class assignment, we were able to validate both the *in situ* cell typing method, and the scRNA-seq classification it relies on, by verifying that cell classes it identifies are found in appropriate layers. The layers in which cell types were identified were consistent with known ground truth (**Figure 3**). Cells identified as O/LM, hippocamposeptal, or O-Bi showed a strong preference for their known locations in *stratum oriens* (*so*), while
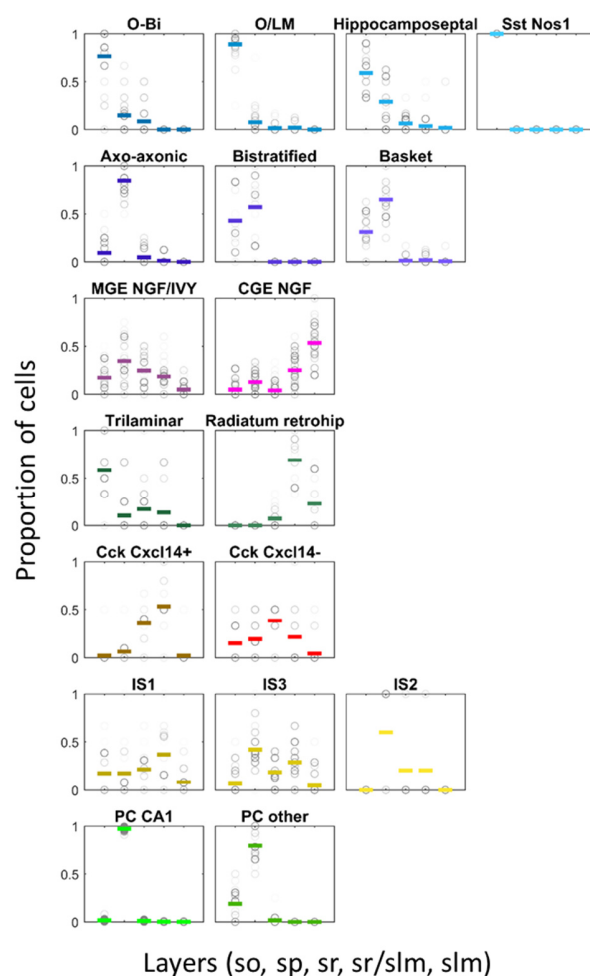

Figure 3. Fraction of each cell class found each CA1 layer. Circles indicate means of a single experiment with gray level representing number of cells of that class in the experiment; colored lines denote grand mean over all 28 hippocampi. In each plot, the 5 x-axis positions represent layers: stratum oriens (so), stratum pyramidale (sp), stratum radiatum (sr), border or strata radiatum and lacunosum-moleculare (sr/slm); stratum lacunosum-moleculare (slm).

neurons identified as bistratified also could be found in *stratum pyramidale* (*sp*)[32,33] (Sst/Nos1 cells were too rare to reliably localized; **Supplementary Figure S13**). Cells identified as basket cells were found in *sp* and less often *so*, while rarer cells identified as axo-axonic were found in the pyramidal layer nearly exclusively[34]. Amongst cells identified with neurogliaform cells, those identified as having developmental origin in the medial ganglionic eminence (including Ivy cells) were found throughout all layers, while those identified as having origins in caudal ganglionic eminence were found nearly exclusively in *stratum lacunosum-moleculare* (*slm*), consistent with previous reports[25,26]. Of the two transcriptomic classes identified with long-range projecting GABAergic

neurons, those identified as trilaminar cells were located primarily in *so*[23,35,36]; while those identified as radiatum retrohippocampal were found at the border of *stratum radiatum* (*sr*) and *slm* [23,27,37,38]. Cck interneurons were divided into two primary classes, with the *Cxcl14*-positive class located primarily in *sr*, close to the *slm* border, and the *Cxcl14*-negative class in all layers, as previously predicted[5]. Amongst interneuron-selective subtypes, cells identified as IS1 were found in all layers as expected[31], while IS3 cells were located primarily in *sp* and *sr,* but very rare in *slm*[28] (IS2 cells were too rare for reliable quantification of their laminar distribution). Reassuringly, cells identified as CA1 pyramidal were found exclusively in *sp*, while cells identified as non-CA1 excitatory could also be occasionally found in *so*, with the latter reflecting primarily subicular pyramidal cells, which can also be found extending into *so* at the CA1/subiculum boundary. Thus, our method identified 18 types of CA1 neuron with the expected laminar distributions, both confirming the predictions of our cell type identifications, and the *in situ* sequencing method itself.

To conclude, we have shown it is possible to identify closely related cell classes *in situ* by levering information from scRNA-seq to design a probe panel and cell identification algorithm for *in situ* sequencing. A key to this method's success is its very low false-positive gene detection rate. Indeed, while the 16 inhibitory classes identified by this algorithm closely match known biology, a recent SNAIL-based method[15] with higher total read density but also possibly higher erroneous reads identified just four or five classes, which were dominated by single genes and do not match combinatorial expression patterns suggested by scRNA-seq and classical analyses. We found that sixteen classes of CA1 inhibitory cell could be identified with just 99 genes chosen using scRNA-seq data. Large-scale scRNA-seq projects are now underway for the whole body, and the data required to design panels and apply this method to all tissues will soon be available. The pciSeq approach requires only low-magnification imaging (in contrast to FISH-based methods), and so may be applied high throughput, raising the possibility of body-wide spatial cell type maps in the near future.

References

1. Zeisel, A. *et al.* Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science* **347,** 1138–1142 (2015).

2. Tasic, B. *et al.* Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat. Neurosci.* **19,** 335–346 (2016).

3. Tasic, B. *et al.* Shared and distinct transcriptomic cell types across neocortical areas. *bioRxiv* 229542 (2017). doi:10.1101/229542

4. Zeisel, A. *et al.* Molecular Architecture of the Mouse Nervous System. *Cell* **174,** 999-1014.e22 (2018).

5. Harris, K. D. *et al.* Classes and continua of hippocampal CA1 inhibitory neurons revealed by single-cell transcriptomics. *PLoS Biol.* **16,** e2006387 (2018).

6. Paul, A. *et al.* Transcriptional Architecture of Synaptic Communication Delineates GABAergic Neuron Identity. *Cell* **171,** 522-539.e20 (2017).

7. Freund, T. F. & Buzsaki, G. Interneurons of the hippocampus. *Hippocampus* **6,** 347–470 (1996).

8. Pelkey, K. A. *et al.* Hippocampal GABAergic Inhibitory Interneurons. *Physiol. Rev.* **97,** 1619–1747 (2017).

9. Somogyi, P. Hippocampus: intrinsic organization. in *Handbook of Brain Microcircuits* (eds. Shepherd, G. M. & Grillner, S.) (Oxford University Press, 2010).

10. Klausberger, T. & Somogyi, P. Neuronal diversity and temporal dynamics: the unity of hippocampal circuit operations. *Science* **321,** 53–7 (2008).

11. Strell, C. *et al.* Placing RNA in context and space - methods for spatially resolved transcriptomics. *FEBS J.* (2018). doi:10.1111/febs.14435

12. 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).

13. Ke, R. *et al.* In situ sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods* **10,** 857–860 (2013).

14. Lee, J. H. *et al.* Highly multiplexed subcellular RNA sequencing in situ. *Science* **343,** 1360–3 (2014).

15. Wang, X. *et al.* Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* **361,** (2018).

16. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348,** aaa6090 (2015).

17. Coskun, A. F. & Cai, L. Dense transcript profiling in single cells by image correlation decoding. *Nat. Methods* **13,** 657–660 (2016).

18. Codeluppi, S. *et al.* Spatial organization of the somatosensory cortex revealed by cyclic smFISH. *bioRxiv* 276097 (2018). doi:10.1101/276097

19. Shah, S., Lubeck, E., Zhou, W. & Cai, L. In Situ Transcription Profiling of Single Cells Reveals Spatial Organization of Cells in the Mouse Hippocampus. *Neuron* **92,** 342–357 (2016).

20. Robinson, M. D. & Smyth, G. K. Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics* **9,** 321–332 (2008).

21. Lein, E. S. *et al.* Genome-wide atlas of gene expression in the adult mouse brain. *Nature* **445,** 168–76 (2007).

22. Katona, L. *et al.* Sleep and movement differentiates actions of two types of somatostatin-expressing GABAergic interneuron in rat hippocampus. *Neuron* **82,** 872–86 (2014).

23. Jinno, S. *et al.* Neuronal diversity in GABAergic long-range projections from the hippocampus. *J Neurosci* **27,** 8790–804 (2007).

24. Viney, T. J. *et al.* Network state-dependent inhibition of identified hippocampal CA3 axo-axonic cells in vivo. *Nat. Neurosci.* **16,** 1802–1811 (2013).

25.   Fuentealba, P. *et al.* Ivy cells: a population of nitric-oxide-producing, slow-spiking GABAergic neurons and their involvement in hippocampal network activity. *Neuron* **57,** 917–29 (2008).

26.   Tricoire, L. *et al.* Common origins of hippocampal Ivy and nitric oxide synthase expressing neurogliaform cells. *J Neurosci* **30,** 2165–76 (2010).

27.   Yamawaki, N. *et al.* Long-range inhibitory intersection of a retrosplenial thalamocortical circuit by apical tuft-targeting CA1 neurons. *bioRxiv* 427179 (2018). doi:10.1101/427179

28.   Acsady, L., Gorcs, T. J. & Freund, T. F. Different populations of vasoactive intestinal polypeptide-immunoreactive interneurons are specialized to control pyramidal cells or interneurons in the hippocampus. *Neuroscience* **73,** 317–34 (1996).

29.   Katona, I. *et al.* Presynaptically located CB1 cannabinoid receptors regulate GABA release from axon terminals of specific hippocampal interneurons. *J Neurosci* **19,** 4544–58 (1999).

30.   Blasco-Ibanez, J. M., Martinez-Guijarro, F. J. & Freund, T. F. Enkephalin-containing interneurons are specialized to innervate other interneurons in the hippocampal CA1 region of the rat and guinea-pig. *Eur J Neurosci* **10,** 1784–95 (1998).

31.   Gulyás, A. I., Hájos, N. & Freund, T. F. Interneurons containing calretinin are specialized to control other interneurons in the rat hippocampus. *J. Neurosci. Off. J. Soc. Neurosci.* **16,** 3397–3411 (1996).

32.   Klausberger, T. *et al.* Spike timing of dendrite-targeting bistratified cells during hippocampal network oscillations *in vivo*. *Nat. Neurosci.* **7,** 41–47 (2004).

33.   Losonczy, A., Zhang, L., Shigemoto, R., Somogyi, P. & Nusser, Z. Cell type dependence and variability in the short-term plasticity of EPSCs in identified mouse hippocampal interneurones. *J. Physiol.* **542,** 193–210 (2002).

34.   Buhl, E. H. *et al.* Physiological properties of anatomically identified axo-axonic cells in the rat hippocampus. *J Neurophysiol* **71,** 1289–307 (1994).

35.   Ferraguti, F. *et al.* Metabotropic glutamate receptor 8-expressing nerve terminals target subsets of GABAergic neurons in the hippocampus. *J Neurosci* **25,** 10520–36 (2005).

36.   Sik, A., Penttonen, M., Ylinen, A. & Buzsaki, G. Hippocampal CA1 interneurons: an in vivo intracellular labeling study. *J Neurosci* **15,** 6651–65 (1995).

37.   Miyashita, T. & Rockland, K. S. GABAergic projections from the hippocampus to the retrosplenial cortex in the rat. *Eur J Neurosci* **26,** 1193–204 (2007).

38.   Jinno, S. Structural organization of long-range GABAergic projection system of the hippocampus. *Front Neuroanat* **3,** 13 (2009).

Acknowledgements

Author contributions

XQ wrote DNA probe design software, performed experiments, analyzed data, designed in situ sequencing protocol, prepared figures, wrote manuscript. KDH designed and wrote analysis software, wrote manuscript. TH designed in situ sequencing protocol. DN

12

designed and wrote online web viewer. AMM prepared tissue samples. NS contributed to gene panel selection. JHL supervised tissue sample collection. MN designed in situ sequencing protocol, supervised experiments, wrote manuscript.

Competing interests
XQ, TH, MN hold shares in Cartana AB.