

1 **HyDrop enables droplet based single-cell ATAC-seq and single-cell RNA-seq using**
2 **dissolvable hydrogel beads**

3 Florian V. De Rop^{1,2}, Joy N. Ismail^{1,2}, Carmen Bravo González-Blas^{1,2}, Gert J. Hulselmans^{1,2},
4 Christopher C. Flerin^{1,2,4}, Jasper Janssens^{1,2}, Koen Theunis^{1,2,4}, Valerie M. Christiaens^{1,2}, Jasper
5 Wouters^{1,2}, Gabriele Marcassa^{1,3}, Joris de Wit^{1,3}, Suresh Poovathingal^{1,#}, and Stein Aerts^{1,2,#}

6 ¹ VIB-KU Leuven Center for Brain & Disease Research

7 ² Laboratory of Computational Biology, Department of Human Genetics, KU Leuven

8 ³ Laboratory of Synapse Biology, Department of Neurosciences, KU Leuven

9 ⁴ Aligning Science Across Parkinson's (ASAP) Collaborative Research Network, Chevy Chase,
10 MD,= 20815.

11 #Shared last author; correspondence to suresh.poovathingal@kuleuven.be and
12 stein.aerts@kuleuven.be.

Abstract

Single-cell RNA-seq and single-cell ATAC-seq technologies are used extensively to create cell type atlases for a wide range of organisms, tissues, and disease processes. To increase the scale of these atlases, lower the cost, and pave the way for more specialized multi-ome assays, custom droplet microfluidics may provide solutions complementary to commercial setups. We developed HyDrop, a flexible and open-source droplet microfluidic platform encompassing three protocols. The first protocol involves creating dissolvable hydrogel beads with custom oligos that can be released in the droplets. In the second protocol, we demonstrate the use of these beads for HyDrop-ATAC, a low-cost non-commercial scATAC-seq protocol in droplets. After validating HyDrop-ATAC, we applied it to flash-frozen mouse cortex and generated 7,996 high-quality single-cell chromatin accessibility profiles in a single run. In the third protocol, we adapt both the reaction chemistry and the capture sequence of the barcoded hydrogel bead to capture mRNA, and demonstrate a significant improvement in throughput and sensitivity compared to previous open-source droplet-based scRNA-seq assays (Drop-seq and inDrop). Similarly, we applied HyDrop-RNA to flash-frozen mouse cortex and generated 9,508 single-cell transcriptomes closely matching reference single-cell gene expression data. Finally, we leveraged HyDrop-RNA's high capture rate to analyse a small population of FAC-sorted neurons from the *Drosophila* brain, confirming the protocol's applicability to low-input samples and small cells. HyDrop is currently capable of generating single-cell data in high throughput and at a reduced cost compared to commercial methods, and we envision that HyDrop can be further developed to be compatible with novel (multi-) omics protocols.

Introduction

In the past five years, droplet microfluidics have been applied extensively to partition single cells and sequence their nucleic acid content. Two methods that pioneered the field, inDrop (1) and Drop-seq (2), both rely on the same working principle: individual cells are rapidly encapsulated into a nanoliter droplet together with a barcoded bead. Barcoding primers carried by the beads inside the emulsion are then used to index each individual cell's mRNA. This process occurs either inside the droplet, where the cell's mRNA is reverse transcribed using barcoded primers released by the barcoded bead (inDrop) or after emulsion breaking, where the cellular mRNA is anchored onto barcodes carried by resin beads (Drop-seq). The thousands of single-cell transcriptomes can then be processed for next-generation sequencing in a pooled manner. The high throughput and low per-cell reagent consumption associated with this approach have allowed researchers to profile gene expression of tens of thousands of single cells (3–5) at an affordable cost. However, the labour requirements of current open-source droplet microfluidic protocols combined with their limited sensitivity has hindered widespread adoption. Furthermore, academic development of droplet-microfluidic single-cell sequencing technology after inDrop and Drop-seq has been limited. To our knowledge, only one non-commercial droplet-based scATAC-seq protocol has been published so far (6). Despite its elegant conceptual solution to droplet-based combined scATAC and scRNA-seq, the SNARE-seq protocol is still labour-intensive, and the use of resin beads leads to reduced cell capture rates. These resin beads are loaded at dilute concentrations to prevent microfluidic obstruction, and as a result, many droplets are empty, leading to a low cell capture rate (~2%), and reagent waste (7). Instead, deformable hydrogel beads can be stacked and loaded into droplets at a fixed rate without risk of microfluidic failure, thereby increasing cell capture rate to >50% (7,8). Several commercial solutions have emerged since inDrop and Drop-seq (9,10), and their application has been used to generate hundreds of thousands (11,12) - and recently millions (13,14) - of single-cell transcriptomes at a high sensitivity. However, the high cost of these commercial protocols remains prohibitive for many research applications, and their fixed nature limits custom protocol development.

In order to increase the sensitivity and user-friendliness of open-source microfluidic protocols, and to provide a more flexible and open platform, we developed HyDrop, a new hydrogel-based droplet microfluidic method for high-throughput scRNA-seq or scATAC-seq. We adapted inDrop's original isothermal extension bead barcoding protocol (15) to a linear amplification workflow to generate more uniformly barcoded beads. Next, we applied a custom hydrogel bead production process similar to a recently published protocol (16) to generate dissolvable beads, improving barcoded primer release and diffusion. We also optimised Drop-seq's pooled template-switching reverse transcription strategy for application inside the cell/bead emulsion, and optimised the assay's sensitivity by testing several different cDNA library preparation strategies. The combination of these adaptations resulted in a significantly increased sensitivity for HyDrop-RNA compared to inDrop and Drop-seq at no additional cost, and in a more user-friendly workflow. Additionally, the dissolution of the hydrogel beads stabilises the cell/bead emulsion during linear amplification thermocycling. This change allowed us to implement to our knowledge the first open-source single-cell ATAC-seq in droplets using hydrogel beads. We

applied both technologies to mouse cerebral cortex and generated single-cell ATAC-seq and single-cell RNA-seq data that is highly concordant with reference data.

Results

Generation of dissolvable hydrogel beads with barcodes for scATAC-seq and scRNA-seq

We generated barcoded hydrogel beads that can dissolve and release their embedded barcoded oligonucleotide. Polyacrylamide beads incorporating disulfide crosslinkers and short oligonucleotide PCR handles were generated by droplet microfluidics similar to a recently published method (16). A custom droplet microfluidic chip (fig. 3 – fig. supp. 1) was employed to produce beads of approximately 50 μ m diameter. These hydrogel beads were then barcoded using a modified three-round split-pool linear amplification synthesis method (1,15), resulting in 96x96x96 (884,736) barcode possibilities (see supplementary file 1 – “Molecular sequence description of HyDrop bead barcoding” for an in-depth visualisation of the nucleic acid sequences in every step). The terminal sequence used in the final round of barcoding can be varied depending on the assay the beads will be used for (see Methods, fig. 1a). A sequence complementary to the Tn5 transposase adapter was used to capture tagmented chromatin fragments in scATAC-seq and a unique molecular identifiers (UMI)(17) + poly (dT) sequence was used to capture and count poly (A)+ mRNA in scRNA-seq (protocols described further). The barcoded beads were stored in a glycerol-based freezing buffer at -80 °C in order to prevent loss of primers over time (fig. 1 – fig. supp. 1).

We validated the extension of the hydrogel bead primers using fluorescent probes complementary to the beads 3-prime terminal sequence (15) (fig. 1b, fig. 1 – fig. supp. 2) and the sub-barcode purity using fluorescent probes complementary to one of the 96 sub-barcode possibilities (fig. 1c, fig. 1 – fig. supp. 2, 3, table S1, 2). These experiments showed that there was no significant loss of primers or mixing of barcodes throughout the barcoding process, and that the beads are uniform in size and primer content. Additional testing revealed that our modified linear amplification barcoding method produced more uniformly barcoded beads compared to the isothermal amplification protocol described in inDrop (fig. 1 – fig. supp. 2). Furthermore, the disulfide moieties incorporated in both the bead's polymer matrix and oligonucleotide linker are cleaved when exposed to reducing conditions, such as dithiothreitol (DTT). This chemical method of release is more user-friendly compared to the UV-mediated (1,15) primer release as the beads do not have to be shielded from light. In addition to improved primer release compared to non-dissolvable beads (fig. 1 – fig. supp. 4, table S3), dissolvable beads did not disrupt the emulsion during the thermocycling (fig. 1 – fig. supp. 5). We hypothesize that this effect is due to the lower physical stress generated by the dissolved hydrogel as opposed to the solid hydrogel. This added benefit allowed us to implement scATAC-seq (see further) on HyDrop. Finally, by varying the concentration of the acrydite primer during bead synthesis, lower or higher amounts of cleavable barcoded primers could be generated. We found that when the concentration of acrydite primer incorporated in the hydrogel matrix was high (50 μ M, similar to InDrop), excess unreacted barcodes could not be sufficiently filtered out in further downstream steps. These primers were then carried over to subsequent reactions, leading to random barcoding of free fragments after droplet merging, and

subsequently to cell-mixed expression or chromatin accessibility profiles. The bead primer concentration with an optimal balance between sensitivity and library purity was found to be 12 μ M for both scATAC-seq and scRNA-seq (fig. 1 – fig. supp. 6).

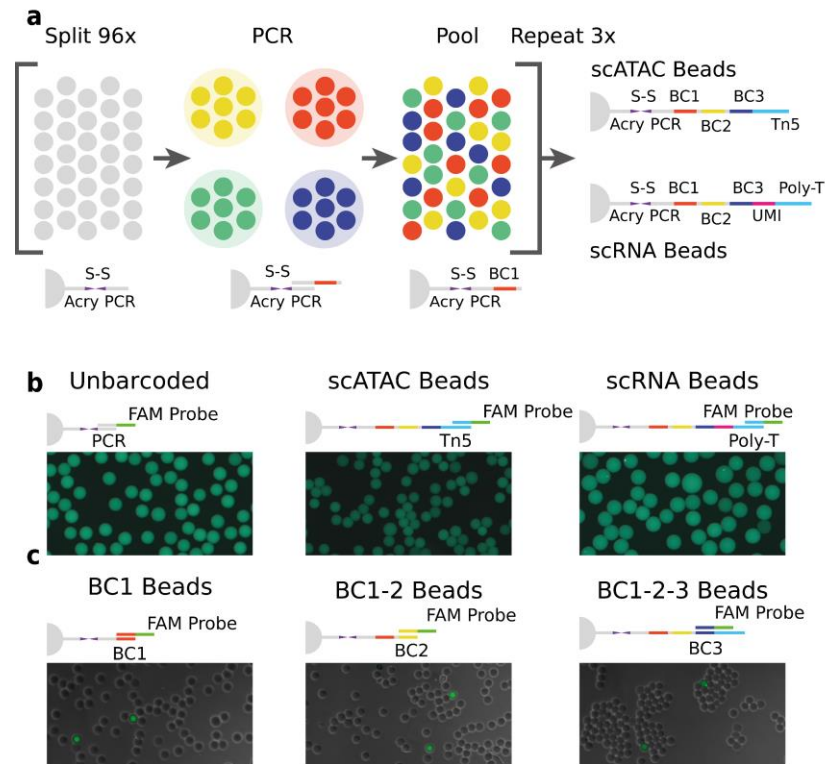


Figure 1. Technical overview of HyDrop barcoded bead production and quality control. a. Split-pool process for barcoding of dissolvable hydrogel beads. Beads are sequentially distributed over 96 wells, sub-barcoded, re-pooled, and distributed three times to generate 96x96x96 (884,736) possible barcode combinations. Different 3'-terminal capture sequences are possible depending on the oligonucleotide sequence appended in the last step. **b.** Semi-quantitative assessment of bead primer incorporation by FISH after every sub-barcoding step shows that bead fluorescence uniformity is retained throughout the barcoding process. **c.** FISH with probes complementary to only one of 96 sub-barcode possibilities shows that approximately 1/96 beads exhibit fluorescence for a selected sub-barcode probe. Fluorescence signal is overlaid with a brightfield image at 50% transparency to indicate positions of non-fluorescent beads (see figure 1 – supplementary figures 2 to 6 for additional quality control experiments and full images).

Implementation and accuracy assessment of HyDrop-ATAC

We implemented a new open-source protocol for single-cell ATAC-seq using HyDrop's dissolvable barcoded hydrogel beads. Nuclei were Tn5 tagmented in bulk and co-encapsulated with HyDrop-ATAC beads. Pitstop, a selective small molecule clathrin inhibitor, was supplemented during nuclei extraction and tagmentation to increase nucleus permeability to Tn5

(18). Inside the droplet, the hydrogel beads dissolve and release their uniquely barcoded primers inside the droplet due to the presence of DTT carried by the nuclei/PCR mix. Subsequent thermocycling of the emulsion denatures the Tn5 protein complex and releases accessible chromatin fragments within the droplet. These fragments were then linearly amplified and cell-indexed by the bead's barcoded primers after which the emulsion was broken and the indexed ATAC fragments were pooled, PCR amplified, and sequenced (fig. 2, see supplementary file 1 – “Molecular sequence description of HyDrop-ATAC” for an in-depth visualisation of the nucleic acid sequences in every step).

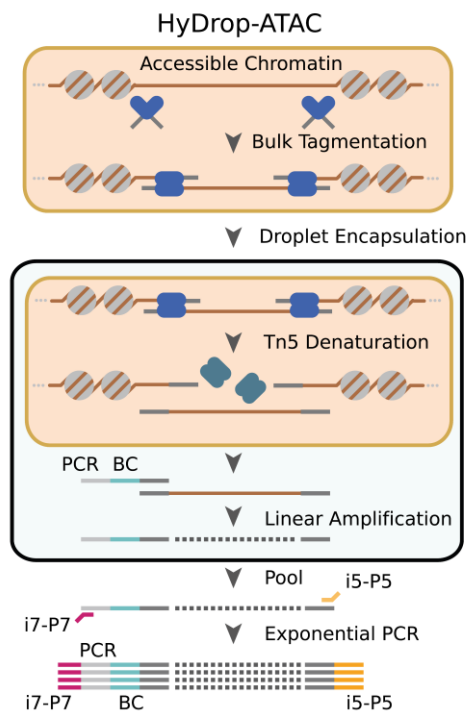


Figure 2. Schematic overview of HyDrop-ATAC. Nuclear membrane is visualised in salmon, water droplet is visualised in blue. Nuclei were Tn5 tagged in bulk and co-encapsulated with HyDrop-ATAC beads, where the hydrogel beads dissolve and release their uniquely barcoded primers. Thermocycling of the emulsion releases accessible chromatin fragments which are then linearly amplified and cell-indexed by the bead's barcoded primers within the droplet. The emulsion is then broken and the indexed ATAC fragments are pooled, PCR amplified, and sequenced

In order to co-encapsulate beads and nuclei with a high capture rate and minimal microfluidic complexity, we developed a custom microfluidic chip (fig. 3 – fig. supp. fig. 1). The chip design features one inlet for beads, one inlet for cells or nuclei and one inlet for the emulsion carrier oil. This configuration is slightly more convenient to operate compared to inDrop's 4 channel setup. Several layers of passive filters near the inlet ports mitigate dust and debris buildup during droplet generation to prevent obstruction of the channels. Beads and nuclei were loaded via a tip reservoir to reduce non-linear flow behaviour and the potential accumulation of cells/nuclei and hydrogel beads associated with narrow tubes (19,20) (fig. 3a, b). Due to the stability of all

flows and the deformable nature of the hydrogel beads, > 90% occupancy of hydrogel beads in droplets could be achieved (8) (fig. 3c), resulting in a final cell recovery of ~65%.

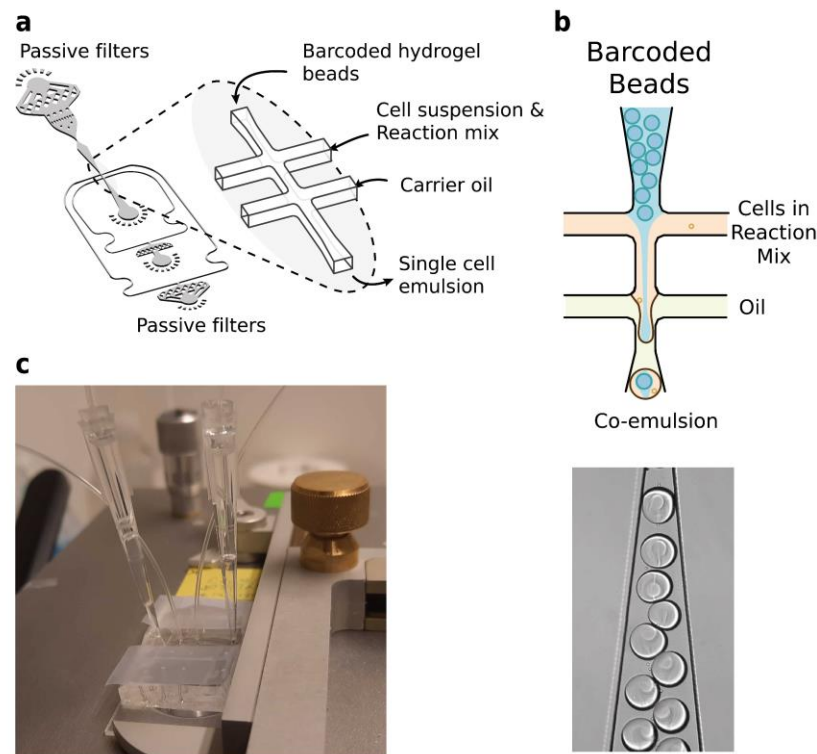


Figure 3. Microfluidics setup employed in HyDrop emulsion generation. **a.** Microfluidic chip setup on the Onyx platform. Cells and beads are loaded into pipette tips and plugged into a HyDrop Chip. Flow of oil and aqueous phases is achieved by Onyx displacement syringe pumps. **b.** HyDrop chip design has three inlets: one each for carrier oil, barcoded hydrogel beads and cell/reaction mix (see figure 3 - figure supplement 1 for full design). Passive filters at each inlet prevent dust and debris from entering the droplet generating junction. **c.** Diagram and snapshot of cell/bead encapsulation into microdroplets.

To assess the purity of scATAC-seq libraries generated by HyDrop-ATAC, we performed two mixed-species experiments. First, we generated single-nuclei ATAC-seq libraries from a 50:50 mixture of human breast cancer (MCF-7) and a mouse melanoma cell line generated previously (21). We developed a custom pre-processing and mapping pipeline for HyDrop-ATAC data (21). After filtering the cell barcodes for a minimum TSS enrichment score of 7 and unique fragment count of 1,000, we recovered 1,353 cells from an input of 2,000, and a median of 2,705 unique fragments per cell. 88% of all barcode reads were paired with the whitelist when 1 mismatch was allowed along the entire length of the 30 bp barcode. An additional 1.16% and 1.13% of HyDrop barcodes could be assigned to the whitelist when 2 and 3 mismatches were allowed respectively. The remaining 9% of barcodes sequences could not be identified and are most likely a result of frame-shift due to incorporation errors in the synthetic oligonucleotide production process. Notably, only 39.7% of barcode reads could be matched with the whitelist in

a public inDrop dataset (3), when 1 mismatch was allowed. When pre-filtered cell barcode sequences were used as a whitelist for a public mouse retina Drop-seq dataset (2), 49.5% of reads could be matched. We identified 98.4% of cells as either human or mouse at a minimum purity of 95% fragments mapping to either species (fig. 4a). Next, we generated libraries from a mixture of MCF-7/PC-3/Mouse cortex (45:45:10) to evaluate whether two human cell types can be distinguished. A spike-in of 10% mouse cells was used as an internal control. We recovered 2,602 human cells, 466 mouse cells, and 93 species doublets after filtering for 95% species purity. Clustering human cells (together with the MCF-7 cells from the first species mixing experiment to evaluate batch effects) recovered two distinct populations, each exhibiting specific ATAC-seq peaks near MCF-7 or PC-3 marker genes (fig. 4 – fig. supp. 1a). Aggregated reads per cluster showed typical ATAC-seq profiles concordant with public bulk ATAC-seq data (22) (fig. 4b, fig. 4 – fig. supp. 1b).

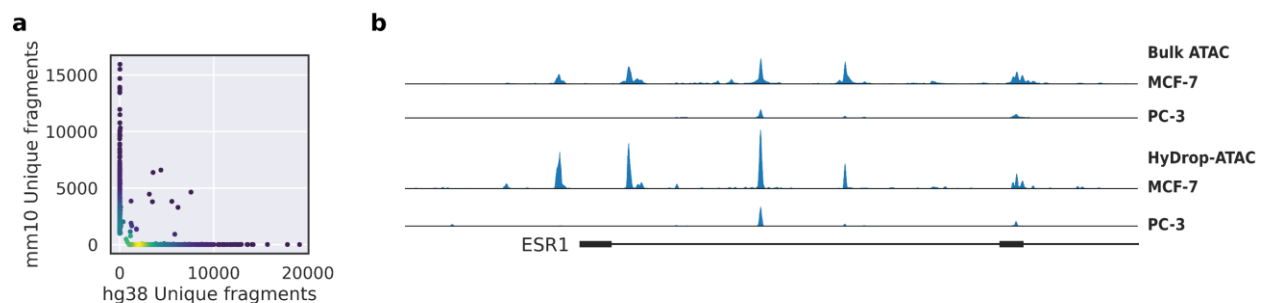


Figure 4. Validation of HyDrop-ATAC on mixed-species cell lines - a. Scatterplot of the number of unique fragments detected in a 50:50 mixture of human MCF-7 and mouse melanoma cells coloured by local density estimation. **b.** RPGC-normalized aggregate genome tracks comparing HyDrop-ATAC and bulk ATAC-seq profiles of human MCF-7 and PC-3 cell lines around the Estrogen receptor 1 (ESR1) locus, scaled to maximum of all four samples. Aggregate enrichment profile of reads around transcription start site (TSS), see fig. 4 - fig. supp. 1 for clustering of these cells and count correlation with public data. Supplementary source data files available for figure a. **Application of HyDrop-ATAC on flash-frozen mouse cortex recapitulates cerebral cellular heterogeneity and cell-type specific accessibility profiles**

To evaluate the performance of HyDrop-ATAC on primary tissue, we generated single cell libraries from snap-frozen, dissected adult mouse brain cortex. Libraries were sequenced to approximately 75% duplication rate. After filtering for a minimum of 1,000 unique nuclear fragments, a TSS enrichment score of 5, and removing 506 cells (6%) detected as doublets by Scrublet (23), we recovered a total of 7,996 single nuclei. Cells passing the filters had a median of 4,148 fragments per cell, a median TSS enrichment score of 13, and a median of 53% of fragments in peaks, reflecting high-quality cells and low levels of background signal (fig. 5a-d). Even though the number of unique fragments per cell (~4K) is lower than that of commercial methods (e.g., 17-20K per cell for 10x Genomics, see Methods), HyDrop-ATAC achieves comparable results in terms of TSS enrichment and FRIP scores. We used cisTopic (24) to

reduce the dimensionality of the dataset and the Leiden algorithm (25) to cluster cells (fig. 5e). Cell annotation using the aggregated ATAC signal around several neocortex markers (26,27) recovered 19 distinct cell types, similar to previously published scATAC-seq mouse cortex data (28,29) (fig. 5 – fig. supp. 1). For example, we identified oligodendrocyte precursors and mature oligodendrocytes, marked by exclusive accessibility nearby *Sox10* and *Pdgfra2*, respectively. Within ganglionic eminence-derived interneurons, we were able to further distinguish medial ganglionic eminence-derived subtypes with specific ATAC-seq signal near either *Vip* or *Lamp5*, and caudal ganglionic eminence-derived subtypes with accessibility near either *Sst* or *Pvalb*. Finally, HyDrop-ATAC data revealed distinct cell-type specific differentially accessible regions (fig. 6a, b).

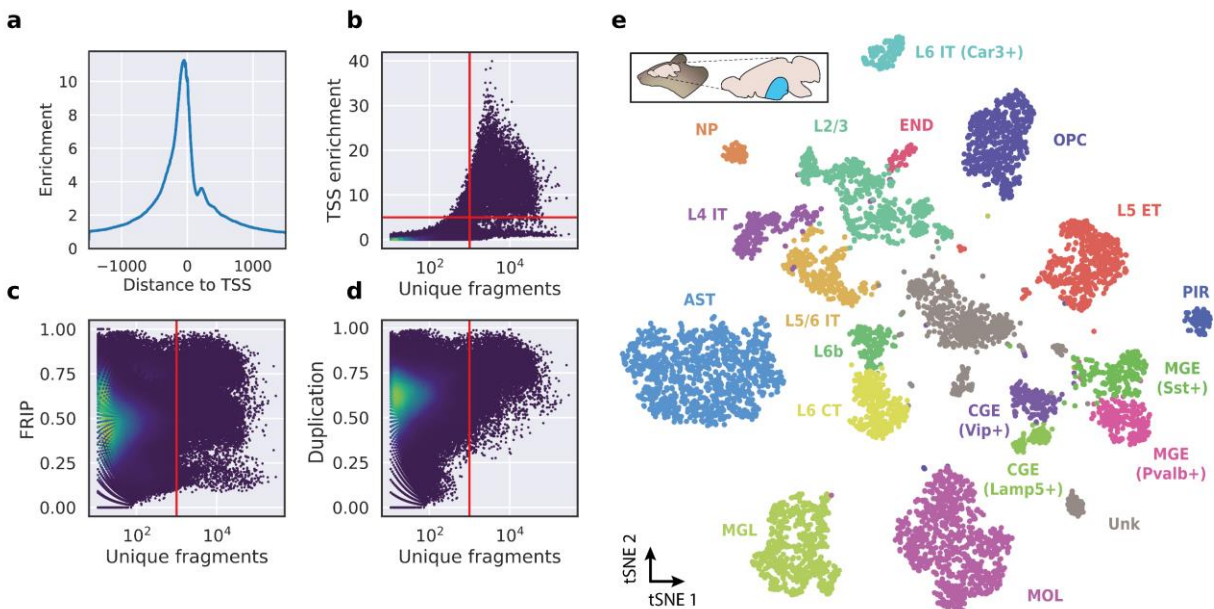


Figure 5. Application of HyDrop-ATAC on flash-frozen mouse cortex recapitulates cellular heterogeneity. Aggregate enrichment of ATAC fragments near transcription start-sites (TSS) (a), TSS enrichment per barcode (b), fraction of reads in peaks (FRIP) per barcode (c) and duplication rate per barcode (d) in mouse cortex HyDrop-ATAC data. A minimum TSS enrichment of 5 and a unique number of fragments of 1000 are used as cut-off values to separate cells from background (red lines). Cells are coloured by local density estimation. e. UMAP projection of 7,996 mouse cortex nuclei annotated with cell type inferred by accessibility near marker genes. Abbreviations: microglia (MGL), mature oligodendrocytes (MOL), oligodendrocyte precursors (OPC), astrocytes (AST), endothelial cells (END), piriform cortex neurons (PIR), caudal and medial ganglionic eminence derived neurons (CGE, MGE), layers 2-6 intratelencephalic (IT), L5 extratelencephalic (ET), L5/6 near projecting excitatory neurons (NP), L6 corticoencephalic (CT), and deep L6 excitatory neurons (L6b). See fig. 5 – fig. supp. 1 for cluster marker gene activities. Supplementary source data files available for figures b, c and d.

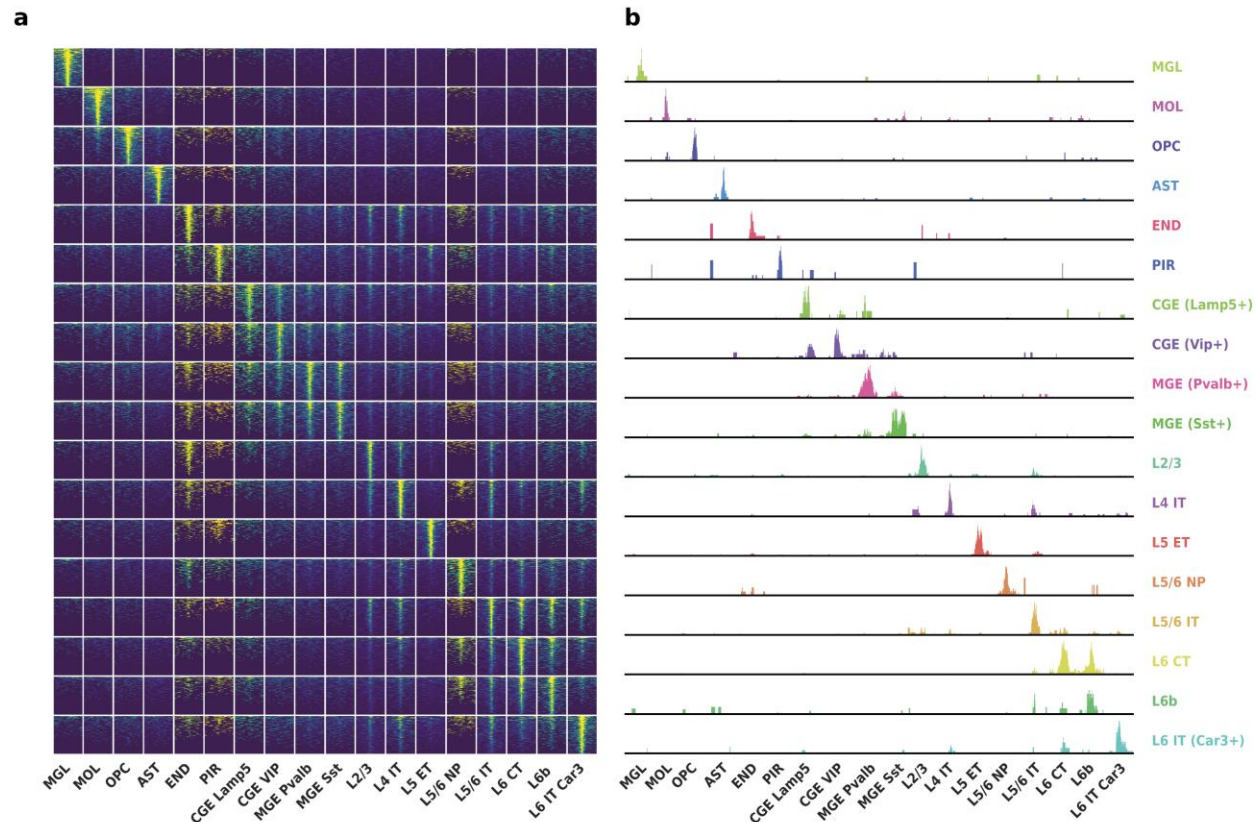


Figure 6. Differentially accessible regions between cell types recovered by HyDrop-ATAC on mouse cerebral cortex. **a.** Aggregate accessibility of top 1000 differentially accessible regions per cluster. **b.** Row-scaled, CPM-normalized aggregate genome track covering the top 1 differentially accessible region (DAR) for each cluster.

Implementation of HyDrop-RNA as a hybrid method between inDrop and Drop-seq

We next implemented a new scRNA-seq assay using barcoded bead primers carrying a 3-prime poly (dT) sequence. Single cells or nuclei were resuspended in a reverse transcriptase mix and co-encapsulated into microdroplets with 3-prime poly (dT) HyDrop beads. The same microfluidic chip design was used for both HyDrop-RNA and HyDrop-ATAC. Cells are lysed inside the droplets upon contact with the lysis buffer in which the barcoded beads were suspended. Simultaneously, barcoded primers were released from the hydrogel bead after exposure to DTT present in the reverse transcriptase mix. Reverse transcription inside the emulsion generates thousands of barcoded single-cell cDNA libraries in parallel. The emulsion was then broken and the single-cell transcriptome libraries were processed further in a pooled manner (fig. 7, see supplementary file 3 - "Molecular sequence description of HyDrop-RNA" for an in-depth visualisation of the nucleic acid sequences in every step), similarly to the InDrop protocol (15). Although both assays are based on hydrogel beads, HyDrop-RNA differs significantly from InDrop. HyDrop-RNA employs a template switching oligo (TSO) reverse transcription technique (similar to Drop-seq), rather

than an *in vitro* transcription/random hexamer priming workflow. This change simplifies and speeds up the protocol significantly with no reduction in sensitivity.

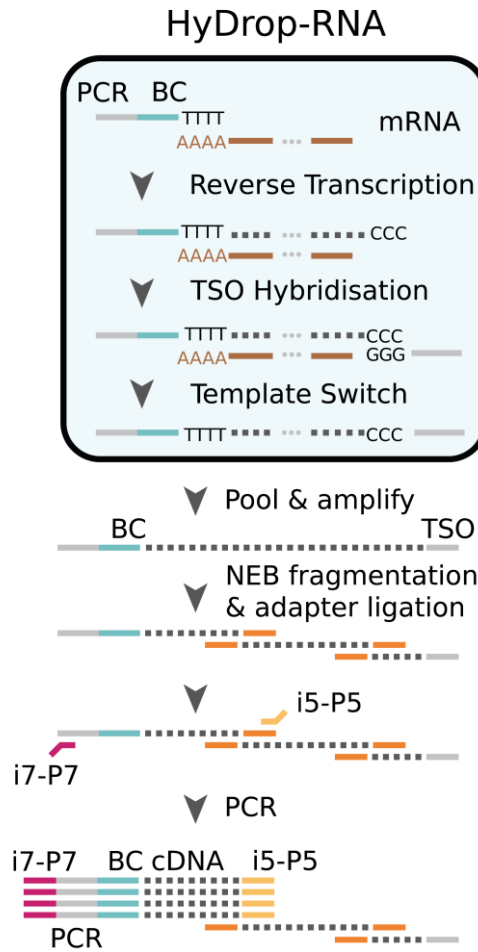


Figure 7. Schematic overview of HyDrop-ATAC. Single cells or nuclei are resuspended in a reverse transcriptase mix and co-encapsulated into microdroplets with 3-prime poly (dT) HyDrop beads. Cells are lysed inside the droplets and barcoded primers are released from the hydrogel bead. Reverse transcription inside the emulsion generates thousands of barcoded single-cell cDNA libraries in parallel which are processed further in a pooled manner after breaking the emulsion.

To improve the assay's sensitivity, we interrogated the impact of several alterations to the protocol's reaction chemistry by testing them on a 50:50 human-mouse (human melanoma, mouse melanoma) mixture. We first investigated the use of pooled exonuclease I treatment after reverse transcription to remove unused barcode primers. We reasoned that these unused barcoded primers could potentially prime transcripts during the subsequent bulk IS-PCR, leading to a loss of purity of transcripts associated with a given barcode. As evident from the increase in pure cells detected, we found that exonuclease I treatment indeed improved assay

purity (Fig. supp. 1). We then tested the implementation of a locked nucleic acid (LNA) in the TSO, as it has been shown to increase assay sensitivity due to increased stability of the TSO-cDNA complex (30). We also investigated the addition of GTP/PEG to the in-droplet reverse transcription reaction. Both the addition of PEG as a molecular crowding agent and GTP as a stabilizer of the TSO complex has been shown to improve assay sensitivity in SMART-seq3 (31). Finally, we wondered whether a second strand synthesis library construction approach could outperform the TSO/IS-PCR approach. In order to test this, we compared both alkaline hydrolysis and enzymatic treatment (RNAse H) to remove the RNA strand from the first strand product, and evaluated the performance of both Bst 2.0 DNA polymerase and Klenow (exo-) (32,33) fragment for second strand synthesis. We found that the classical TSO and IS-PCR protocol supplemented with GTP/PEG performed best, yielding a median of 2,110 UMIs and 1,325 genes per cell with a species-purity of 90.1% (fig. 8a, fig. 8 – fig. supp. 1). Accordingly, the GTP/PEG method was used in all further HyDrop-RNA experiments. Applying this protocol to a 50:50 human-mouse (MCF-7, mouse melanoma) mixture recovered 1,235 human and 846 mouse cells with 169 species doublets at a cutoff of 95% species purity, with a median of 1,439 UMIs and 904 genes per cell (fig. 8b).

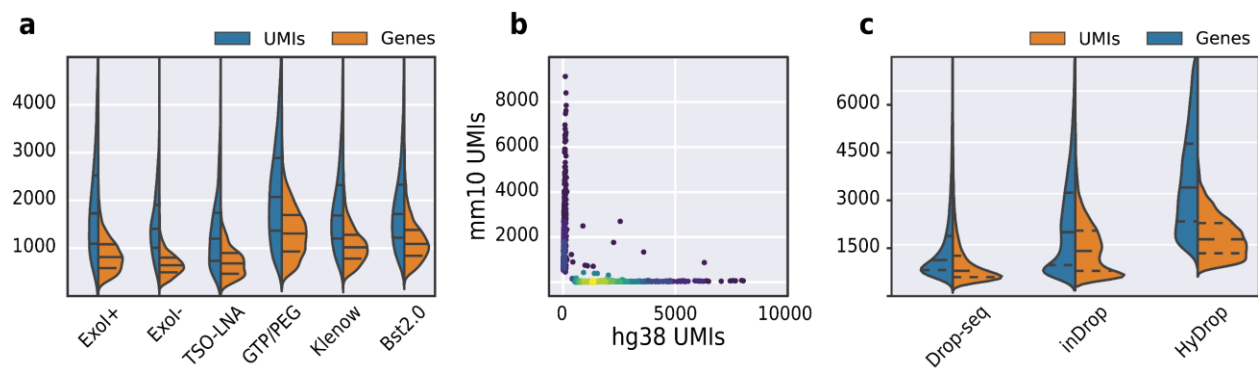


Figure 8. Validation and benchmarking of HyDrop-RNA on species-mixed cell line samples and mouse cerebral cortex. **a.** Comparison of UMI and gene count of HyDrop-RNA with and without Exo I treatment post-droplet merging, with the use of a locked nucleic acid (LNA) template switching oligo (TSO) and with GTP/PEG, BST2.0 and Klenow fragment library preparation. Inner lines represent Q1, median and Q3. See figure 8 – supplement 1 for species purity plots of these experiments. **b.** Scatterplot of human and mouse UMIs detected in a 50:50 mixture of human MCF-7 and mouse melanoma cells coloured by local density estimation. **c.** Comparison of UMI and gene count of public inDrop mouse cortex data, public Drop-seq mouse retina data, and HyDrop-RNA mouse cortex data. Inner lines represent Q1, median and Q3. See figure 8 – figure supplement 2 for additional quality comparison including commercial methods. Supplementary source data files available for figures a, b and c.

HyDrop-RNA on flash-frozen mouse cortex recovers cerebral cell types

We then used HyDrop-RNA to generate 9,508 single nuclei transcriptomes from snap-frozen mouse cortex in a single experiment. At a sequencing saturation of approximately 60% duplicates, we achieve a median of 3,389 UMIs and 1,658 genes per cell before, and 3,404

UMIs and 1,662 genes per cell after doublet filtering (fig. 8c), compared to the median of 1,920 UMI and 1,321 genes detected by inDrop snRNA-seq on mouse auditory cortex neurons (3) and the median of 1,071 UMI and 763 genes detected by Drop-seq on mouse retina neurons (2). Both datasets also exhibited a lower read alignment efficiency compared to HyDrop, with Drop-seq mapping 52%/21%, and inDrop mapping 48%/28% of reads to the mouse genome/transcriptome, while HyDrop reaches 88%/55%. HyDrop's "sequencing efficiency" is higher than both inDrop and Drop-seq's: 7.44% of HyDrop's sequenced reads end up as mapped transcripts with a unique molecular identifier, whereas this metric is 3.24% for the inDrop dataset and 4.46% for the Drop-seq dataset (fig. 8 – figure supplement 2). 10x Chromium v2 gene expression reference data reports a median number of genes of 775-2,679 and a median number of UMIs of 1,127-6,570 on E18 and adult mouse brain nuclei, and a sequencing efficiency of 2.54%-20.9% (see methods). Comparing the top per-cluster differentially expressed genes with markers from the Allen Brain Atlas SMART-seq data (26) revealed 30 distinct populations corresponding to previously identified cell types (fig. 9a-d, fig. 9 – fig. supp. 1). In addition to the major neuronal and glial populations previously detected in our HyDrop-ATAC experiment, we detect a small population of vascular leptomeningeal cells (VLMC) and layer 2 intratelencephalic neurons from the medial entorhinal area (L2 IT ENTm). We also detect both D1 and D2 medium spiny neurons (MSN) as a result of residual striatal tissue and layer 3 *Scnn1a*+ neurons from the retrosplenial and anterior cingulate area (L3 RSP ACA), concordant with Atlas SMART-seq data (26).

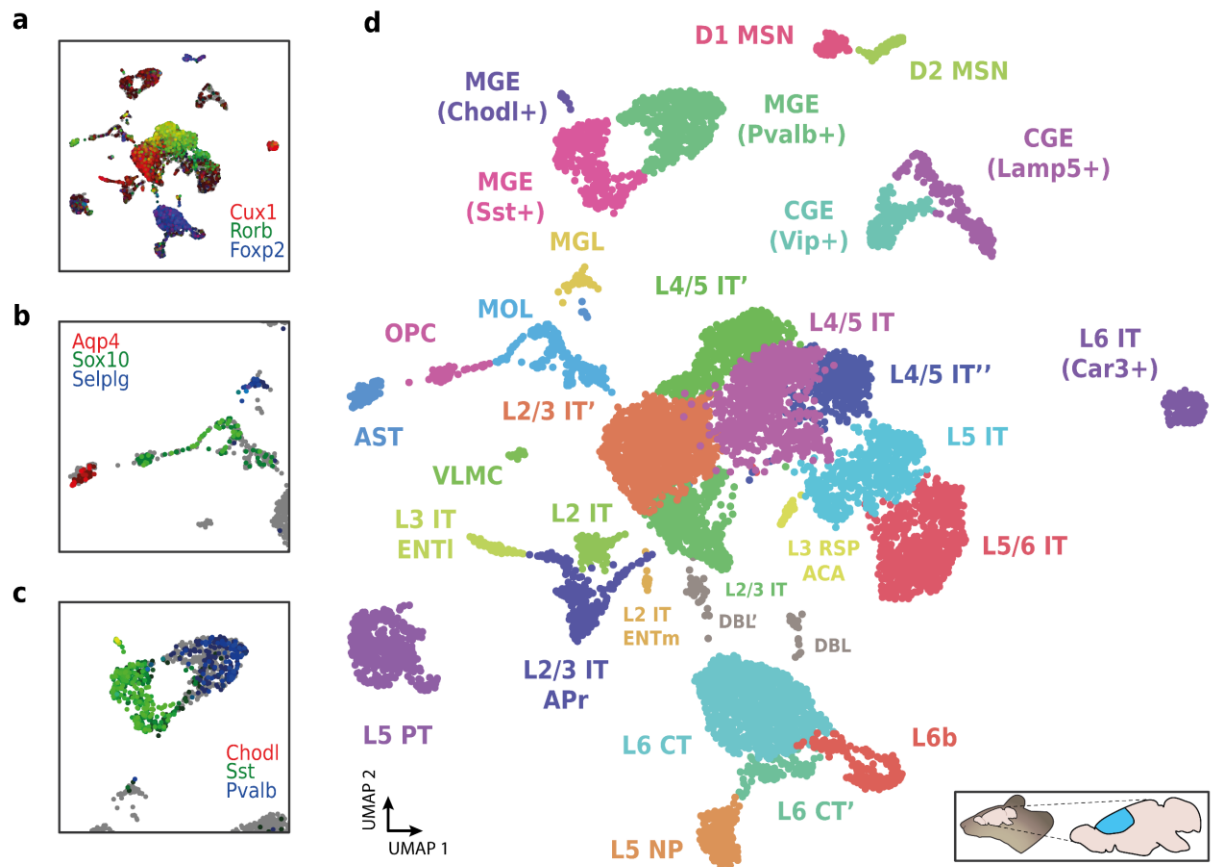


Figure 9. Application of HyDrop-RNA on flash-frozen mouse cerebral cortex. Mouse cortex UMAP colored by log-scaled UMI counts of Cux1, Rorb, Foxp2 **(a)**, Aqp4, Sox10, Selp1g **(b)**, Chodl, Sst and Pvalb **(c)**. Colors are scaled to minimum and maximum values. **d.** UMAP projection of 9507 mouse cortex nuclei annotated with cell type inferred by marker gene expression. Abbreviations: microglia (MGL), mature oligodendrocytes (MOL), oligodendrocyte precursors (OPC), astrocytes (AST), endothelial cells (END), piriform cortex neurons (PIR), caudal and medial ganglionic eminence derived neurons (CGE, MGE), layers 2-6 intratelencephalic (IT), pyramidal tract (PT), near projecting excitatory neurons (NP) and corticoencephalic (CT) neurons, layer 2 intratelencephalic medial entorhinal area neurons (L2 IT ENTm), L2/3 intratelencephalic area prostriata neurons (L2/3 IT APr), layer 3 intratelencephalic entorhinal neurons (L3 IT ENTl), layer retrosplenial and anterior cingulate area neurons (L3 RSP ACA), deep L6 excitatory neurons (L6b), D1 and D2 medium spiny neurons (MSN), and vascular leptomeningeal cells (VLMC). See figure 9 - supplementary figure 1 for expression of top differentially accessible genes within these clusters.

HyDrop-RNA on FAC-sorted *Drosophila* neurons confirms high capture rate even on small input samples

To assess HyDrop-RNA's performance on low cell input samples, we also performed the protocol on approximately 1500 FAC-sorted neurons from the *Drosophila* brain. We dissected brains in which mCherry expression was driven in specific cell populations by a Gal4 driver line (R74G01-Gal4) and used mCherry-positive cells as input for HyDrop-RNA (fig. 10a). Of the 1,500 cells obtained after FACS sorting, we recovered 973 fly brain cells with a median of 1,307 UMIs and 640 genes (fig. 10b). In comparison, in-house Drop-seq performed on the entire fly brain recovered a median of 579 UMIs and 289 genes per cell at a deeper sequencing saturation (11). Annotation of the 973 single-cell transcriptomes obtained by HyDrop confirmed the presence of all three Kenyon cell subtypes alongside T1 and Tm1 neurons, as expected from our stainings and previous reports (34). Surprisingly, we also detected a small population of Mi1 neurons (fig. 10c-e, fig. 10 – fig. supp. 1).

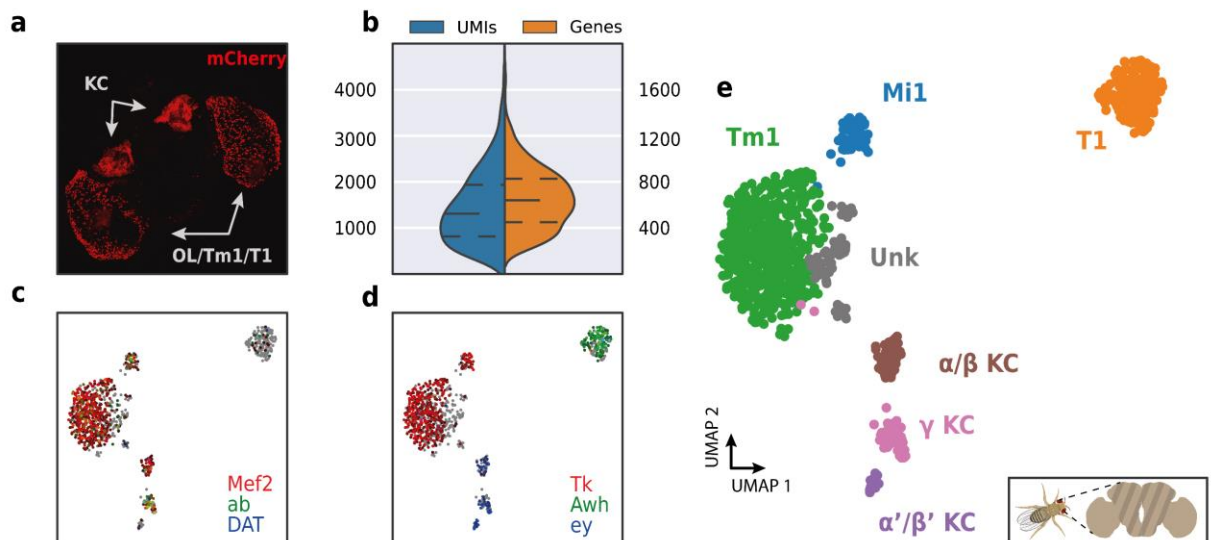


Figure 10. Application of HyDrop-RNA on FAC-sorted *Drosophila* neurons. **a.** Confocal maximum intensity projection of R74G01-Gal4>UAS-mCherry brain. **b.** Violin plot of UMIs and genes detected in nuclei derived from FAC-sorted fly neurons. Inner lines represent Q1, median and Q3. Fly neuron UMAP colored by log-scaled UMI counts of Mef2, ab, DAT (**c**) and Tk, Awh, ey (**d**). Colors are scaled to minimum and maximum values. **e.** UMAP of 973 FAC-sorted *Drosophila* neurons annotated with cell types inferred by marker gene expression (KC, Kenyon cells; Tm1, transmedullary neuron; Mi1, medullary intrinsic neuron). See figure 10 - supplementary figure 1 for expression of marker genes within these clusters. Supplementary source data file available for figure b.

Discussion

We developed HyDrop, an open-source platform for single-cell RNA- or single-cell ATAC-seq using hydrogel beads. We applied HyDrop to generate thousands of mouse, human, and *Drosophila* single-cell gene expression and chromatin accessibility profiles to demonstrate the protocol's applicability to a variety of different biological samples. HyDrop-ATAC and HyDrop-RNA experiments on mouse and fly tissues recapitulated the cellular heterogeneity of these complex samples and strongly agreed with reference datasets (11,26,28,29). We found HyDrop-RNA to outperform its open-source predecessors (1,2,6) both in terms of sensitivity and user-friendliness. Moreover, at a per-cell library cost of < \$0.03 it does so at a significantly lower cost compared to commercial droplet-microfluidic alternatives (9,35,36).

Optimization and modification of HyDrop's reaction chemistry and bead composition may lead to further improvements in sensitivity. For example, we consider that benchmarking additional bead barcoding strategies, such as direct on-bead DNA synthesis or barcode ligation (37) instead of linear amplification may further improve overall barcode quality. Ligation strategies would also allow the incorporation of oligonucleotide modifications in the barcode's capture sequence, a possible necessity for adaptation of HyDrop to novel assays. For example, this change would permit HyDrop's extension to protocols such as s3-ATAC (38), which leverages the use of uracil bases and uracil-intolerant polymerases to double the amount of usable fragments generated by Tn5. The current 96³ design could also be changed to 384² (like inDrop) or even 768². This two-step barcoding approach would result in shorter barcode sequences, allowing for cost-saving in sequencing and increased flexibility when pooling HyDrop libraries together with other library types such as 10X Genomics 3' WTA libraries. However, rescaling the barcodes in this manner would also necessitate the acquisition of a larger pool of sub-barcodes - from three 96 well plates to 8 and 16 plates respectively - and reduce the total barcode complexity from 884,736 to 147,456 and 589,824. Finally, one of the main challenges of implementing custom microfluidics systems is the considerable price of accurate stepper motor syringe pumps (~\$2000 per independent channel) and the need for trained personnel. However, several simplified alternatives have recently emerged (39,40), and we believe their application with HyDrop could further reduce the start-up costs and complexity of single-cell experiments.

We envision that HyDrop's reduction in both cost and labour will accelerate the scaling of large-scale atlasing efforts and bring the benefits of single-cell sequencing to smaller projects. Additionally, HyDrop's dissolvable bead synthesis and linear amplification barcoding toolkit

408 could potentially be exploited to produce more complex beads incorporating multiple capture
409 sequences, leading to implementation of single-cell (multi-) omics assays such as the capture of
410 accessible chromatin, (m)RNA, and surface epitopes (41) or intracellular proteins (42) from
411 single cells.

Materials & methods

Microfluidic droplet generator manufacturing

Microfluidic droplet generators were produced using standard SU-8 lithography and PDMS lithography according to well established protocols (43). Briefly, the design for droplet generators were made in AutoCAD R2014 and the designs are printed onto a chrome mask using a laser writer. The SU-8 lithography is performed on a 4 inch silicon wafer using SU 8-2050 (Microchem) negative photoresist using UV aligner (EVG-620). As per manufacturers' recommendation, spin coating of the wafer with SU8 was performed at 500 rpm (ramp 100 rpm/s) - 10 s and 2000 rpm (ramp: 300 rpm/s) - 30 s, to achieve a feature height of 70-80 μ m). For preparing the PDMS chip, a mixture of PDMS monomer and crosslinker (Dow Corning SYLGARD 184) was prepared at a ratio of 10:1 and mixed thoroughly. The mixture was degassed in vacuum for 45 minutes and poured onto an SU-8 master template and baked at 80 °C for 4 hours. Inlet ports were cut using a 1 mm biopsy needle after which the chips were exposed to high-voltage plasma for 30 s and bonded onto a glass slide. 5 μ L of 2% Trichloro (1H,1H,2H,2H-perfluorooctyl)silane in HFE was injected into each channel, incubated for 10 mins at room temperature and excess oil was removed by applying pressurized air. Chips were finally baked at 100 °C for 2 hours (more detailed methods for photolithography and PDMS lithography in <https://dx.doi.org/10.17504/protocols.io.bvpin5ke>).

Barcoded hydrogel bead manufacturing & storage

Dissolvable hydrogel beads are synthesized similar to a previously published protocol¹⁴ and barcoded according to a modified inDrop protocol (15). For synthesizing 2-3 mL batch of beads, 2 mL of Bead Monomer Mix (6% acrylamide, 0.55% bisacryloylcystoyleamine, 10% TBSET (10 mM Tris-HCl pH 8, 137 mM NaCl, 2.7 mM KCl, 10 mM EDTA, 0.1% Triton X-100), 12 μ M acrydite primer, 0.6% ammonium persulfate) was encapsulated into 50 μ m diameter droplets in HFE-7500 Novac oil with EA-008 surfactant (RAN Biotech). 1 mL aliquots of the resulting emulsion was layered with 400 μ L of mineral oil and incubated at 65 °C for 14 hours. Excess mineral oil and the emulsion oil was removed and 2-3 washes with 1 mL of droplet breaking solution (20% PFO in HFE) was performed. Beads were pelleted at 5000 xg, 4 °C for 30 seconds and washed twice in 1 mL of 1% SPAN-80 in hexane. Beads were sequentially washed in TBSET until all hexane phase was removed.

Beads were washed twice in Bead Wash Buffer (10 mM Tris-HCl pH 8, 0.1% Tween-20), twice in PCR Wash Buffer (10 mM Tris-HCl pH8, 50 mM KCl, 1.5 mM MgCl₂, 0.1% Tween-20). The subsequent liquid handling in the 96-well plate is performed using Hamilton microlab STAR robot. 22.5 μ L of beads were distributed to a 96 well plate. 2.5 μ L of 100 μ M sub-barcode primer and 25 μ L of KAPA HiFi Hotstart master mix (Roche) was added to each well and the plate was thermocycled (95 °C 3 min., 5 cycles of 98 °C 20 s, 38 °C 4 min., 72 °C 2 min., 1 cycle of 98 °C 1 min., 38 °C 10 min., 72 °C 4 min., followed by a final hold on 4 °C) with intermittent vortexing during every annealing step. 50 μ L of STOP-25 (10 mM Tris-HCl pH 8, 25 mM EDTA, 0.1% Tween-20, 100 mM KCl) was added to each well to deactivate the polymerase and its contents were pooled. Remaining beads in wells were washed out with 100 μ L of STOP-25 and the

beads were rotated at room temperature for 30 minutes. Beads were then washed with STOP-10 (10 mM Tris-HCl pH 8, 10 mM EDTA, 0.1% Tween-20, 100 mM KCl) and rotated for 10 minutes in Denaturation Solution (150 mM NaOH, 85 mM BRIJ-35). Beads were washed twice in Denaturation Solution and twice more in Neutralisation Solution (100 mM Tris-HCl pH 8, 10 mM EDTA, 0.1% Tween-20, 100 mM NaCl). The sub-barcoding step was repeated twice more for a total of 3 sub-barcodes.

Hydrogel beads were sequentially filtered using a 70 µm strainer (Falcon). For both the HyDrop-ATAC and RNA beads were pelleted at 300 xg, 4 °C and resuspended in 5 mL of Bead Freezing Buffer (150 mM NaCl, 125 mM Tris-HCl pH 7, 10 mM MgCl₂, 4% Tween-20, 0.75% Triton X-100, 30% glycerol, 0.3% BSA). Beads were pelleted at 300 xg, 4 °C and resuspended in 5 mL of Bead Freezing Buffer a second time and incubated at 4 °C for at least 3 hours. Beads were pelleted at 1000 xg, 4 °C and the pellet was aliquoted into 35 µL aliquots and stored at -80 °C for long term storage (further method details in [dx.doi.org/10.17504/protocols.io.bxsapnae](https://doi.org/10.17504/protocols.io.bxsapnae))

Hydrogel bead fluorescence in-situ hybridisation quality control

Bead QC was performed as described previously (15). Briefly, 10 µL of hydrogel beads were resuspended in 1 mL of hybridization buffer (5 mM Tris-HCl pH 8.0, 5 mM EDTA, 0.05% Tween-20, 1 M KCl) and centrifuged for 1 min. at 1000 xg. The wash was repeated once more, and 960 µL of the supernatant was removed. 4 µL of 200 µM specific FAM probe was added depending on which part of the barcode needed testing (see supplementary oligonucleotide table). The beads were incubated at room temperature in the dark for 30 min. Beads were washed thrice in QC buffer and visualised under a Zeiss Axioplan 2 microscope and a Zeiss Colibri 7 light source (300 ms exposure time, 80% light source intensity). Occasional punctate intensities (diameter ~1 µm) were observed in quality control performed on 10x beads and are thought to be dye precipitates or crystals.

Fluorescence images were analysed using opencv2 (44). Briefly, foreground and background were determined using Otsu's method. Circles were detected in the foreground using Hough's algorithm. The mean grayscale intensity of the background was subtracted from the grayscale intensity at the centre of each detected circle. The centre intensity method was favored over a volumetric intensity measurement over the entire bead as slight differences in focus between images can affect intensity near the edge of the bead. For the 1/96 sub-barcode experiment, both fluorescence and brightfield images were taken of the same beads incubated with one out of 96 possible sub-barcodes. Total number of beads was counted manually from the brightfield image, while positive beads were counted using the image analysis tool from the fluorescence images. A mask of the positive beads from the fluorescence image was overlaid onto the brightfield image to mark positive beads for visualisation. Image analysis tool code and raw data of all used images available at https://github.com/aertslab/hydrop_data_analysis.

Cell culture and cell dissociation

MCF-7 cells (NCI-DTP Cat# MCF7, RRID:CVCL_0031) were cultured in RPMI1640 (ThermoFisher 11875093) medium supplemented with 10% FBS (ThermoFisher 10270-106),

1% penicillin/streptomycin (Life Technologies 15140122), and 10 ug/mL insulin (Sigma Aldrich I9278) and passaged twice per week. PC-3 cells (NCI-DTP Cat# PC-3, RRID:CVCL_0035) were cultured in RPMI1640 medium supplemented with 10% FBS and 1% penicillin/streptomycin and passaged twice per week. Mouse melanoma cells were cultured in DMEM (ThermoFisher 13345364) supplemented with 10% FBS and 1% penicillin/streptomycin and passaged once per week. MM087 melanoma cells were cultured in F-10 Nutrient mix supplemented with 10% FBS and 1% penicillin/streptomycin and passaged once per week. All cell lines tested negative for mycoplasma prior to use. The identity of PC-3, MCF-7 and mouse melanoma cell lines were correctly identified since their aggregate scATAC-seq data correlated highly with public data (see figure 4b, figure 4 – supplement 1b for PC-3 and MCF-7, mouse melanoma data not shown). The identity of our MM087 was confirmed by SNP comparison with the MM087 genome for our manuscript “Robust gene expression programs underlie recurrent cell states and phenotype switching in melanoma” (Wouters et al., <https://doi.org/10.1038/s41556-020-0547-3>).

Cells were washed in PBS and dissociated into single cell suspensions by adding 1.5 mL of 0.05% Trypsin (Life Technologies 25300054) and waiting for 5 to 7 minutes. The single-cell suspension was centrifuged at 500 rcf for 5 min at 4 °C and the resulting pellet was resuspended in PBS. This PBS wash was repeated once more and the single-cell suspension was processed further.

Fly rearing and cell dissociation

GMR74G01-Gal4 (BDSC Cat# 39868, RRID:BDSC_39868) and UAS-mCherry (BL#38425) flies were obtained from Bloomington Drosophila Stock Center. The resulting cross strain was raised on standard cornmeal-agar medium at 25 °C at a 12h light/dark cycle. 50 adult brains were dissected in DPBS and transferred to a tube containing 100 µL of cold DPBS solution. Samples were centrifuged at 500 rcf for 1 min and the supernatant was replaced by 50 µL of dispase (3 mg/mL, Sigma-Aldrich, D4818, 2 mg) and 75 µL of collagenase I (100 mg/mL, Invitrogen, 17100-017). Brains were dissociated in a Thermoshaker (Grant Bio PCMT) at 500 rpm for 2 h at 25 °C, with pipette mixing every 15 min. Cells were subsequently washed with 1000 µL of cold DPBS solution and centrifuged at 500 rcf for 5 min at 4 °C and resuspended in 250 µL of DPBS with 0.04% BSA. Cell suspensions were passed through a 10 µm pluriStrainer (ImTec Diagnostics, 435001050). Cells were sorted based on viability and mCherry positivity using the Sony MA900 cell sorter. Sorted cells were collected into Eppendorf tubes pre-coated with 1% BSA.

Cell line nuclei extraction

A pellet of 1 million dissociated cells or fewer was incubated on ice in 200 µL of ATAC Lysis Buffer (1% BSA, 10 mM Tris-HCl pH 7.5, 10 mM NaCl, 0.1% Tween-20, 0.1% NP-40, 3 mM MgCl₂, 70 µM Pitstop in DMSO, 0.01% Digitonin) for 5 to 7 minutes. 1 mL of ATAC Nuclei Wash Buffer (1% BSA, 10 mM Tris-HCl pH 7.5, 0.1% Tween-20, 10 mM NaCl, 3 mM MgCl₂) was added and the nuclei were pelleted at 500 xg, 4 °C for 5 minutes. The resulting pellet was resuspended in 100 µL of ice-cold PBS and filtered with a 40 µm strainer (Flowmi).

Mouse cortex dissection

All animal experiments were conducted according to the KU Leuven ethical guidelines and approved by the KU Leuven Ethical Committee for Animal Experimentation (approved protocol numbers ECD P183/2017). Mice were maintained in a specific pathogen-free facility under standard housing conditions with continuous access to food and water. Mice used in the study were 57 days old and were maintained on 14 h light, 10 h dark light cycle from 7 to 21 hours. In this study, cortical brain tissue from female P57 BL/6Jax was used. Animals were anesthetized with isoflurane, and decapitated. Cortices were collected, divided in four equal quadrants along the dorso-ventral and anterior-posterior axis, and immediately snap-frozen in liquid nitrogen. For HyDrop-ATAC, a ventral/posterior quadrant from the left hemisphere was used. For HyDrop-RNA, a dorsal/anterior quadrant was used from the left hemisphere of a second mouse.

Snap-frozen mouse cortex nuclei extraction

For the preparation of nuclei for scRNA-seq, we used a modified protocol from the recently published single nuclei preparation toolbox³⁷ to extract nuclei from snap-frozen mouse cortex samples. Briefly, a ~1 cm³ frozen piece of mouse cortex tissue was transferred to 0.5 mL of ice-cold homogenisation buffer (Salt-tris solution - 146 mM NaCl, 10 mM Tris 7.5, 1 mM CaCl₂, 21 mM MgCl₂, 250 mM Sucrose, 0.03% Tween-20, 0.01% BSA, 25 mM KCl, 1 mM 2-Mercaptoethanol, 1X cOmplete protease inhibitor, 0.5U/ul of RNase In Plus (Promega)) in a Dounce homogenizer mortar and thawed for 2 minutes. The tissue was homogenised with 10 strokes of pestle A and 5 strokes of pestle B until a homogeneous nuclei suspension was achieved. The resulting homogenate was filtered through a 70 µm cell strainer (Corning). The homogenizer and the filter is rinsed with an additional 500 µL of homogenization buffer. The tissue material was pelleted at 500xg and the supernatant was discarded. The tissue pellet was resuspended in a homogenization buffer without Tween-20. An addition 1.65 ml of homogenization buffer was topped up and mixed with 2.65 ml of Gradient Medium (75 mM sucrose, 1 mM CaCl₂, 50% Optiprep, 5 mM MgCl₂, 10 mM Tris 7.5, 1 mM 2-Mercaptoethanol 1X cOmplete protease inhibitor, 0.5U/ul of RNase In Plus (Promega)). 4 mL of 29% Iodexanol cushion was prepared with a Diluent medium (250 mM Sucrose, 150 mM KCl, 30 mM MgCl₂, 60 mM Tris 8) and was loaded into an ultracentrifuge tube. 5.3 mL of sample in homogenization buffer + gradient medium was gently layered on top of the 29% Iodexanol cushion. Sample was centrifuged at 7700 xg, 4 °C for 30 minutes and the supernatant was gently removed without disturbing the nuclei pellet. Nuclei were resuspended in 100 µL of Nuclei buffer (1% BSA in PBS + 1U/ul of RNase Inhibitor).

For the preparation of nuclei for ATAC seq, we used a slightly modified protocol to extract nuclei from snap-frozen mouse cortex samples. Briefly, a ~1 cm³ frozen piece of mouse cortex tissue was transferred to 1 mL of ice-cold homogenisation buffer (320 mM Sucrose, 10 mM NaCl, 3 mM Mg (OAc), 10 mM Tris 7.5, 0.1 mM EDTA, 0.1% IGEPAL-CA360, 1X cOmplete protease inhibitor and 1 mM DTT) in a Dounce homogenizer mortar and thawed for 2 minutes. The tissue was homogenised with 10 strokes of pestle A and 5 strokes of pestle B until a homogeneous nuclei suspension was achieved. The resulting homogenate was filtered through a 70 µm cell strainer (Corning). 2.65 mL of ice-cold gradient medium was added to 2.65 mL of homogenate and mixed well. 4 mL of 29% Iodexanol cushion (129.2 mM Sucrose, 77.5 mM KCl, 15.5 mM MgCl, 31 mM Tris-HCl pH 7.5, 29% Iodexanol) was loaded into ultracentrifuge tube. 5.3 mL of

572 sample in homogenization buffer + gradient medium was gently layered on top of the 29%
573 Iodoxanol cushion. Sample was centrifuged at 7700 xg, 4 °C for 30 minutes and the supernatant
574 was gently removed without disturbing the nuclei pellet. Nuclei were resuspended in 100 µL of
575 Nuclei buffer (1% BSA in PBS).

576 **HyDrop-ATAC library preparation**

577 50 000 nuclei were resuspended in 50 µL of ATAC Reaction Mix (10% DMF, 10% Tris-HCl pH
578 7.4, 5 mM MgCl₂, 5 ng/µL Tn5, 70 µM Pitstop in DMSO, 0.1% Tween-20, 0.01% Digitonin) and
579 incubated at 37 °C for 1 hour. 100 µL of 0.1% BSA in PBS was added and the nuclei were
580 pelleted at 500 xg, 4 °C for 5 minutes and resuspended in 40 µL of 0.1% BSA in PBS.

581 Tagmented nuclei were added to 100 µL of PCR mix (1.3X Phusion HF Buffer, 15% Optiprep,
582 1.3 mM dNTPs, 39 mM DTT, 0.065 U/µL Phusion HF Polymerase, 0.065 U/µL Deep Vent
583 Polymerase, 0.013 U/µL ET SSB). PCR mix was co-encapsulated with 35 µL of freshly thawed
584 HyDrop-ATAC beads in HFE-7500 Novac oil with EA-008 surfactant (RAN Biotech) on the Onyx
585 microfluidics platform (Droplet Genomics). The resulting emulsion was collected in aliquots of 50
586 µL total volume and thermocycled according to the linear amplification program (72 °C 15 min.,
587 98 °C 3 min., 13 amplification cycles of [98 °C 10 s, 63 °C 30 s, 72 °C 1 min.], followed by a final
588 hold on 4 °C). 125 µL of recovery agent (20% PFO in HFE), 55 µL of GITC Buffer (5 M GITC, 25
589 mM EDTA, 50 mM Tris-HCl pH 7.4) and 5 µL of 1 M DTT was added to each separate aliquot of
590 50 µL thermocycled emulsion and incubated on ice for 5 minutes. 5 µL of Dynabeads was
591 added to the aqueous phase and incubated for 10 minutes. Dynabeads were pelleted on a Nd
592 magnet and washed twice with 80% EtOH. Elution was performed in 50 µL of EB-DTT-Tween
593 (10 mM DTT, 0.1% Tween-20 in EB (10 mM Tris-HCl, pH 8.5)). A 1x Ampure bead purification
594 was performed according to manufacturer's recommendations. Elution was performed in 30 µL
595 of EB-DTT (10 mM DTT in EB). Eluted library was further amplified in 100 µL of PCR mix (1x
596 KAPA HiFi, 1 µM index i7 primer, 1 µM index i5 primer). Final library was purified in a 0.4x-1.2x
597 double-sided Ampure purification and eluted in 25 µL of EB-DTT (10 mM DTT in EB). User-
598 friendly protocol available on [dx.doi.org/10.17504/protocols.io.bxsbpnan](https://doi.org/10.17504/protocols.io.bxsbpnan). A detailed cost
599 analysis of HyDrop-ATAC can be found in supplementary sheet 2.

600 **HyDrop-RNA single cell library preparation**

601 For a recovery of 2000 cells, 3795 cells were resuspended in 85 µL of RT mix (1x Maxima RT
602 Buffer, 0.9 mM dNTPs, 25 mM DTT, 1.3 mM GTP, 15 % Optiprep, 1.3 U/µL RNase inhibitor, 15
603 U/µL Maxima hRT, 12.5 µM TSO, 4.4% PEG-8000). RT mix was co-encapsulated with 35 µL of
604 freshly thawed HyDrop-RNA beads in RAN oil on the Onyx microfluidics platform. The resulting
605 emulsion was collected in aliquots of 50 µL total volume and thermocycled according to the RT
606 program (42 °C for 90 min., 11 cycles of [50 °C for 2 min., 42 °C for 2 min.], 85 °C for 5 min.,
607 followed by a final hold on 4 °C). 125 µL of recovery agent (20% PFO in HFE), 55 µL of GITC
608 Buffer (5 M GITC, 25 mM EDTA, 50 mM Tris-HCl pH 7.4) and 5 µL of 1 M DTT was added to
609 each separate aliquot of 50 µL thermocycled emulsion and incubated on ice for 5 minutes. 99
610 µL of Ampure XP beads was added to the aqueous phase and incubated for 10 minutes.
611 Ampure beads were pelleted on a Nd magnet and washed twice with 80% EtOH. Elution was

performed in 30 µL of EB-DTT-Tween (10 mM DTT, 0.1% Tween-20 in EB). Exonuclease treatment was performed by adding 4 µL of 10x NEBuffer 3.1, 4 µL of Exo I, and 2 µL of dH₂O to 30 µL of eluted library. The Exo I reaction mix was incubated at 37 °C for 5 min., 80 °C for 1 min. for heat inactivation. followed by a final hold at 4 °C. 2 µL of 1 M DTT was added and a 0.8x Ampure XP purification was performed according to manufacturer's recommendations. cDNA was eluted in 40.5 µL of EB-DTT (10 mM DTT in EB) and added to ISPCR mix (40 µL library, 50 µL 2x KAPA HiFi, 10 µL 10 µM TSO-P primer). PCR cycling was performed according to the ISPCR program (95 °C for 3 min., 13 cycles of [98 °C for 20 s, 63 °C for 20 s, 72 °C for 3 min.], 72 °C for 5 min. followed by a final hold at 4 °C. 2 µL of 1 M DTT was added and a 0.6x Ampure XP purification was performed according to manufacturer's recommendations. cDNA was eluted in 28.5 µL of EB-DTT. Final sequencing library was prepared according to the following customised NEB Ultra II FS protocol (NEB E7805S). 80 ng of amplified cDNA was fragmented in Ultra II fragmentation mix (26 µL of amplified cDNA, 7 µL of NEBNext Ultra II FS Reaction Buffer, 2 µL of NEBNext Ultra II FS Enzyme Mix) on the following thermocycling program: 37 °C for 10 min., 65 °C for 30 min. and a final hold at 4 °C. 15 µL of EB was added and a 0.8x Ampure purification was performed according to manufacturer's recommendation and eluted in 35 µL. Fragmented library was adapter-ligated in NEBNext Ultra II adapter ligation mix (35 µL of fragmented library, 30 µL of NEBNext Ultra II Ligation Master Mix, 1 µL of NEBNext Ligation Enhancer, 2.5 µL of NEBNext Adapter for Illumina) at 20 °C for 15 min., with 4 °C final hold. 28.5 µL of EB was added and a 0.8X Ampure purification was performed according to manufacturer's recommendation and eluted in 30 µL. Eluted library was amplified in PCR master mix (50 µL 2x KAPA HiFi, 10 µL 10 µM Hy-i7 primer, 10 µL 10 µM Hy-i5 primer, 30 µL eluted library) in the following thermocycling program: 95 °C for 3 min., 13 cycles of [98 °C for 20 s, 64 °C for 30 s, 72 °C for 30 s], 72 °C for 5 min. and a final hold at 4 °C. Sequencing-ready library was purified using a 0.8x Ampure purification and eluted in 30 µL of EB. User-friendly protocol available on <https://dx.doi.org/10.17504/protocols.io.bva5n2g6>.

HyDrop-RNA optimisation trials

We performed 6 trials on a 50:50 mixture of human melanoma (MM087) and mouse melanoma (MMel). Trials were performed as described in the general HyDrop-RNA protocol, but with the following changes. All trials, except for the GTP/PEG trial, were performed using the following RT reaction mix (1.6x Maxima h-RT buffer, 1.6 mM dNTPs, 47 mM DTT, 15% Optiprep, 1.6 U/µL RNase Inhibitor, 15.7 U/µL Maxima hRT, 12.5 µM TSO). For the Exo- condition, the Exonuclease I treatment was skipped. For all other conditions the Exonuclease I treatment was performed as described above. For the TSO-LNA trial, a locked nucleic acid TSO was used instead of the regular TSO. For the GTP/PEG trial, all steps were performed as described in the main protocol.

For the Klenow fragment second strand synthesis trial, the purified first strand product was treated with 1 µL of E. coli RNase H (NEB M0297S). The mixture was incubated at 37 °C for 30 minutes after which the enzyme was inactivated using 10 mM EDTA. The single stranded product was purified using 1.2x Ampure XP bead purification (BD sciences) and eluted in 25 µL of EB buffer. dN-SMRT primer was added to the single strand product to a final concentration of

2.5 μ M and the mixture was denatured by incubation at 95 °C for 5 minutes. The sample was then allowed to cool to room temperature and incorporated in the Klenow enzyme mix (1x Maxima h-RT buffer, 1 mM dNTP, 1U/ μ L of Klenow Exo-; NEB M0212L) was added to the single strand library. The Klenow enzyme mix was incubated at 37 °C for 60 min. The second strand reaction was stopped by heating the product at 85 °C for 5 min. The sample was purified using 1X Ampure XP and eluted in 40 μ L of EB buffer. The purified second strand product was amplified with ISPCR primers as described above.

For the BST 2.0 polymerase second strand synthesis trial, the purified first strand product was treated with 1 μ L of E. coli RNase H (NEB M0297S). The mixture was incubated at 37 °C for 30 minutes after which the enzyme was inactivated using 10 mM EDTA. The single stranded product was purified using 1.2X Ampure XP bead purification (BD sciences) and eluted in 25 μ L of EB buffer. dN-SMRT primer was added to the single strand product to a final concentration of 2.5 μ M and the mixture was denatured by incubation at 95 °C for 5 minutes. The sample was then allowed to cool to room temperature and incorporated in the Bst 2.0 enzyme mix (1X Isothermal amplification buffer, 1 mM dNTP, 1U/ μ L of Bst 2.0 DNA polymerase; NEB M0537L) was added to the denatured library and the mixture was incubated at 55 °C for 10 mins and 60 °C for 45 minutes. The second strand reaction was stopped by heating the product at 85 °C for 5 minutes. The sample was purified using 1X Ampure XP and eluted in 40 μ L of EB buffer. The purified second strand product was amplified with ISPCR primers as described above.

Sequencing

HyDrop-ATAC libraries were sequenced on Illumina NextSeq500 or NextSeq2000 systems using 50 cycles for read 1 (ATAC paired-end mate 1), 52 cycles for index 1 (barcode), 10 cycles for index 2 (sample index) and 50 cycles for read 2 (ATAC paired-end mate 2).

HyDrop-RNA libraries were sequenced on Illumina NextSeq2000 systems using 50 cycles for read 1 (3' cDNA), 10 cycles for index 1 (sample index, custom i7 read primer), 10 cycles for index 2 (sample index) and 58 cycles for read 2 (barcode + UMI, custom read 2 primer).

HyDrop-ATAC data processing

Barcode reads were trimmed to exclude the intersub-barcode linear amplification adapters using a mawk script. Next, the VSN scATAC-seq pre-processing pipeline (45) was used to map the reads to the reference genome and generate a fragments file for downstream analysis. Here, barcode reads were compared to a whitelist (of 884,736 valid barcodes), and corrected, allowing for a maximum 1 bp mismatch. Uncorrected and corrected barcodes were appended to the fastq sequence identifier of the paired-end ATAC-seq reads. Reads were mapped to the reference genome using bwa-mem with default settings, and the barcode information was added as tags to each read in the bam file. Duplicate-marking was performed using samtools markdup. In the final step of the pipeline, fragments files were generated using Sinto (<https://github.com/timoast/sinto>). For mixed-species data, cells were filtered for a minimum of 1000 unique fragments and a minimum TSS enrichment of 7. For mouse cortex data, higher level analysis such as clustering and differential accessibility were performed using cisTopic

(24). In brief, cells were filtered for a minimum of 1000 unique fragments and a minimum TSS enrichment of 5. Fragments overlapping mouse candidate cis-regulatory regions (46) were counted, and the resulting matrix was filtered for potential cell doublets using a Scrublet (23) threshold of 0.35. Cells were Leiden-clustered based on the cell-topic probability matrix generated by an initial cisTopic LDA incorporating 51 topics, at a resolution of 0.9 with 10 neighbours. A consensus peak set was generated from per-cluster peaks and used to recount fragments. Cells were filtered using the same filtering parameters and a new model with 50 topics was trained. Cells were again Leiden-clustered based on the cell-topic probability matrix generated by the second LDA, at a resolution of 0.9 with 10 neighbours. Region accessibility was imputed based on binarised topic-region and cell-topic distributions. Gene activity was imputed based on Gini index-weighted imputed accessibility in a 10 kb up/downstream decaying window around each gene including promoters. Leiden clusters were annotated based on imputed gene accessibility around marker genes (26,27). Differentially accessible regions were called using one-versus-all Wilcoxon rank-sum tests for each cell type, with an adjusted p-value of 0.05 and log2FC of 1.5. RPGC-normalized aggregate genome coverage bigwigs were generated from BAM files using DeepTools (47). Per-cluster genome coverage tracks were generated using pyBigWig.

HyDrop-RNA data processing

Barcode reads were trimmed to exclude the intersub-barcode linear amplification adapters using a mawk script. Reads were then mapped and cell-demultiplexed using STARsolo (48) in CB_UMI_Complex mode. The resulting STARsolo-filtered count matrices were further analysed using Scanpy (49). In short, cells were filtered on expression of a maximum of 4000 genes, and a maximum of 1% UMIs from mitochondrial genes. Genes were filtered on expression in a minimum of 3 cells. Potential cell doublets were filtered out using a Scrublet (23) threshold of 0.25. The filtered expression matrix was scaled to total counts and log-normalized. Total counts and mitochondrial reads were regressed out and UMAP embedding was performed after PCA. Cells were annotated and fine-tuned based on differential gene expression of marker genes sourced from either the Davie et al. Drosophila brain atlas (11) or the Allen Brain RNA-seq Database (26). Raw inDrop (SRR10545068 to SRR10545079) and Drop-seq (SRR1853178 to SRR1853184) sequencing data was downloaded from SRA. Data was sampled to the HyDrop-RNA mouse cortex sample sequencing depth (52738 reads per cell) and mapped to mouse reference genome using STARsolo, allowing 1 mismatch in the cell barcodes (like in our HyDrop-RNA and HyDrop-ATAC analyses). Full analysis process is documented on the Hydrop GitHub repository (https://github.com/aertslab/hydrop_data_analysis/tree/main/HyDrop-RNA_publicdata_comparison). Public reference 10x single-cell ATAC-seq data was sourced from <https://support.10xgenomics.com/single-cell-atac/datasets> (“Flash frozen cortex, hippocampus, and ventricular zone from embryonic mouse brain (E18)”, “Fresh cortex from adult mouse brain (P50)”). Public reference 10x single-cell gene expression data was sourced from <https://support.10xgenomics.com/single-cell-gene-expression/datasets> (“1k Brain Nuclei from an E18 Mouse”, “2k Brain Nuclei from an Adult Mouse (>8 weeks)”). Public PC-3 and MCF-7 ATAC-seq data was sourced from ENCODE (ENCFF772EFK: doi:10.17989/ENCSR422SUG, ENCFF024FNF: doi:10.17989/ENCSR499ASS).

Data was visualised using a combination of Seaborn (50) and Matplotlib (51). A vector image representing mouse head and cortex was sourced from SciDraw (52).

Ethical approval

All animal experiments were conducted according to the KU Leuven ethical guidelines and approved by the KU Leuven Ethical Committee for Animal Experimentation (approved protocol numbers ECD P037/2016, P014/2017, and P062/2017). All use of cell lines was approved by the KU Leuven Ethical Committee for Research under project number S63316.

Data availability

The datasets generated during and/or analysed during the current study are available on the Gene Expression Omnibus (GEO) with the primary accession number GSE175684 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE175684>), and on SCoPe (https://scope.aertslab.org/#/HyDrop/*/welcome). Step-by-step user protocols are available on Protocols.io (<https://www.protocols.io/workspaces/aertslab>). Data and image analysis tutorials for HyDrop are available on GitHub (https://github.com/aertslab/hydrop_data_analysis).

Funding

This work was supported by an ERC Consolidator Grant to S.A. (no. 724226_cis- CONTROL), by the KU Leuven (grant no. C14/18/092 to S.A.), FWO (grant no. G0B5619N to S.A.), by the joint efforts of The Michael J. Fox Foundation for Parkinson's Research (MJFF) and the Aligning Science Across Parkinson's (ASAP) initiative, by the Foundation Against Cancer (grant no. 2016-070 to S.A.), a postdoctoral research fellowship from Stichting tegen Kanker (Foundation against Cancer), the Belgian Cancer Society (to J.W.), a PhD fellowship from the FWO to F.D., and a technology development grant from VIB Tech Watch. MJFF administers the grant ASAP-000430 on behalf of ASAP and itself. Computing was performed at the Vlaams Supercomputer Center and high-throughput sequencing at the Genomics Core Leuven. For the purpose of open access, the author has applied a CC-BY public copyright license to the Author Accepted Manuscript (AAM) version arising from this submission."

Acknowledgements

We thank Andrew Adey for his great advice on Tn5 and scATAC-seq. We also thank Sebastián Najle, Céline Vallot, and Arnau Sebé-Pedrós for many discussions on droplet microfluidics, Frederik Ceyssens and the KU Leuven Nanocenter for their support with microfabrication, Ghanem Ghanem for his kind donation of the MM087 melanoma lines, Jean-Christophe Marine for his kind donation of the mouse melanoma lines, and Koen De Wispelaere for his assistance during his master internship.

Competing interests

772 The authors declare no competing interests or commercial affiliations.

773

References

1. Klein AM, Mazutis L, Akartuna I, Tallapragada N, Veres A, Li V, et al. Droplet Barcoding for Single-Cell Transcriptomics Applied to Embryonic Stem Cells. *Cell*. 2015 May 21;161(5):1187–201.
2. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*. 2015 May 21;161(5):1202–14.
3. Kalish BT, Barkat TR, Diel EE, Zhang EJ, Greenberg ME, Hensch TK. Single-nucleus RNA sequencing of mouse auditory cortex reveals critical period triggers and brakes. *Proc Natl Acad Sci*. 2020 May 26;117(21):11744–52.
4. Yap E-L, Pettit NL, Davis CP, Nagy MA, Harmin DA, Golden E, et al. Bidirectional perisomatic inhibitory plasticity of a Fos neuronal network. *Nature*. 2021 Feb;590(7844):115–21.
5. Karaiskos N, Wahle P, Alles J, Boltengagen A, Ayoub S, Kipar C, et al. The *Drosophila* embryo at single-cell transcriptome resolution. *Science*. 2017 Oct 13;358(6360):194–9.
6. Chen S, Lake BB, Zhang K. High-throughput sequencing of the transcriptome and chromatin accessibility in the same cell. *Nat Biotechnol*. 2019 Dec;37(12):1452–7.
7. Zhang X, Li T, Liu F, Chen Y, Yao J, Li Z, et al. Comparative Analysis of Droplet-Based Ultra-High-Throughput Single-Cell RNA-Seq Systems. *Mol Cell*. 2019 Jan 3;73(1):130–142.e5.
8. Abate AR, Chen C-H, Agresti JJ, Weitz DA. Beating Poisson encapsulation statistics using close-packed ordering. *Lab Chip*. 2009 Sep 21;9(18):2628–31.
9. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, et al. Massively parallel digital transcriptional profiling of single cells. *Nat Commun*. 2017 Jan 16;8(1):14049.
10. Sarkar A, Mei A, Paquola ACM, Stern S, Bardy C, Klug JR, et al. Efficient Generation of CA3 Neurons from Human Pluripotent Stem Cells Enables Modeling of Hippocampal Connectivity In Vitro. *Cell Stem Cell*. 2018 May 3;22(5):684–697.e9.
11. Davie K, Janssens J, Koldere D, De Waegeneer M, Pech U, Kreft Ł, et al. A Single-Cell Transcriptome Atlas of the Aging *Drosophila* Brain. *Cell*. 2018 Aug 9;174(4):982–998.e20.
12. Svensson V, da Veiga Beltrame E, Pachter L. A curated database reveals trends in single-cell transcriptomics. *Database* [Internet]. 2020 Jan 1 [cited 2021 May 19];2020(baaa073). Available from: <https://doi.org/10.1093/database/baaa073>
13. Datlinger P, Rendeiro AF, Boenke T, Senekowitsch M, Krausgruber T, Barreca D, et al. Ultra-high-throughput single-cell RNA sequencing and perturbation screening with combinatorial fluidic indexing. *Nat Methods*. 2021 May 31;1–8.
14. Ren X, Wen W, Fan X, Hou W, Su B, Cai P, et al. COVID-19 immune features revealed by a large-scale single-cell transcriptome atlas. *Cell*. 2021 Apr 1;184(7):1895–1913.e19.

- 811 15. Zilionis R, Nainys J, Veres A, Savova V, Zemmour D, Klein AM, et al. Single-cell barcoding
812 and sequencing using droplet microfluidics. *Nat Protoc.* 2017 Jan;12(1):44–73.
- 813 16. Wang Y, Cao T, Ko J, Shen Y, Zong W, Sheng K, et al. Dissolvable Polyacrylamide Beads
814 for High-Throughput Droplet DNA Barcoding. *Adv Sci.* 2020;7(8):1903463.
- 815 17. Kivioja T, Vähärautio A, Karlsson K, Bonke M, Enge M, Linnarsson S, et al. Counting
816 absolute numbers of molecules using unique molecular identifiers. *Nat Methods.* 2011 Nov
817 20;9(1):72–4.
- 818 18. Mulqueen RM, DeRosa BA, Thornton CA, Sayar Z, Torkenczy KA, Fields AJ, et al.
819 Improved single-cell ATAC-seq reveals chromatin dynamics of in vitro corticogenesis.
820 *bioRxiv.* 2019 May 15;637256.
- 821 19. Sinha N, Subedi N, Wimmers F, Soennichsen M, Tel J. A Pipette-Tip Based Method for
822 Seeding Cells to Droplet Microfluidic Platforms. *JoVE J Vis Exp.* 2019 Feb
823 11;(144):e57848.
- 824 20. Hur SC, Henderson-MacLennan NK, McCabe ERB, Carlo DD. Deformability-based cell
825 classification and enrichment using inertial microfluidics. *Lab Chip.* 2011 Mar 7;11(5):912–
826 20.
- 827 21. Dankort D, Curley DP, Carlidge RA, Nelson B, Karnezis AN, Damsky Jr WE, et al. Braf
828 V600E cooperates with Pten loss to induce metastatic melanoma. *Nat Genet.* 2009
829 May;41(5):544–52.
- 830 22. ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human
831 genome. *Nature.* 2012 Sep 6;489(7414):57–74.
- 832 23. Wolock SL, Lopez R, Klein AM. Scrublet: Computational Identification of Cell Doublets in
833 Single-Cell Transcriptomic Data. *Cell Syst.* 2019 Apr 24;8(4):281-291.e9.
- 834 24. Bravo González-Blas C, Minnoye L, Papasokrati D, Aibar S, Hulselmans G, Christiaens V,
835 et al. cisTopic: cis-regulatory topic modeling on single-cell ATAC-seq data. *Nat Methods.*
836 2019 May;16(5):397–400.
- 837 25. Traag V, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected
838 communities. *Sci Rep.* 2019 Dec;9(1):5233.
- 839 26. Yao Z, van Velthoven CTJ, Nguyen TN, Goldy J, Seden-Cortes AE, Baftizadeh F, et al. A
840 taxonomy of transcriptomic cell types across the isocortex and hippocampal formation. *Cell*
841 [Internet]. 2021 May 17 [cited 2021 May 19]; Available from:
842 <https://www.sciencedirect.com/science/article/pii/S0092867421005018>
- 843 27. Zeisel A, Hochgerner H, Lönnerberg P, Johnsson A, Memic F, Zwan J van der, et al.
844 Molecular Architecture of the Mouse Nervous System. *Cell.* 2018 Aug 9;174(4):999-
845 1014.e22.
- 846 28. Preissl S, Fang R, Huang H, Zhao Y, Raviram R, Gorkin DU, et al. Single-nucleus analysis
847 of accessible chromatin in developing mouse forebrain reveals cell-type-specific
848 transcriptional regulation. *Nat Neurosci.* 2018 Mar;21(3):432–9.

- 849 29. Li YE, Preissl S, Hou X, Zhang Z, Zhang K, Fang R, et al. An Atlas of Gene Regulatory
850 Elements in Adult Mouse Cerebrum. *bioRxiv*. 2020 May 17;2020.05.10.087585.
- 851 30. Picelli S, Faridani OR, Björklund ÅK, Winberg G, Sagasser S, Sandberg R. Full-length
852 RNA-seq from single cells using Smart-seq2. *Nat Protoc*. 2014 Jan;9(1):171–81.
- 853 31. Hagemann-Jensen M, Ziegenhain C, Chen P, Ramsköld D, Hendriks G-J, Larsson AJM, et
854 al. Single-cell RNA counting at allele and isoform resolution using Smart-seq3. *Nat*
855 *Biotechnol*. 2020 Jun;38(6):708–14.
- 856 32. Hughes TK, Wadsworth MH, Gierahn TM, Do T, Weiss D, Andrade PR, et al. Second-
857 Strand Synthesis-Based Massively Parallel scRNA-Seq Reveals Cellular States and
858 Molecular Features of Human Inflammatory Skin Pathologies. *Immunity*. 2020 Oct
859 13;53(4):878-894.e7.
- 860 33. Stickels RR, Murray E, Kumar P, Li J, Marshall JL, Di Bella DJ, et al. Highly sensitive
861 spatial transcriptomics at near-cellular resolution with Slide-seqV2. *Nat Biotechnol*. 2021
862 Mar;39(3):313–9.
- 863 34. Konstantinides N, Kapuralin K, Fadil C, Barboza L, Satija R, Desplan C. Phenotypic
864 Convergence: Distinct Transcription Factors Regulate Common Terminal Features. *Cell*.
865 2018 Jul 26;174(3):622-635.e13.
- 866 35. Satpathy AT, Granja JM, Yost KE, Qi Y, Meschi F, McDermott GP, et al. Massively parallel
867 single-cell chromatin landscapes of human immune cell development and intratumoral T
868 cell exhaustion. *Nat Biotechnol*. 2019 Aug;37(8):925–36.
- 869 36. Lareau CA, Duarte FM, Chew JG, Kartha VK, Burkett ZD, Kohlway AS, et al. Droplet-
870 based combinatorial indexing for massive-scale single-cell chromatin accessibility. *Nat*
871 *Biotechnol*. 2019 Aug;37(8):916–24.
- 872 37. Delley CL, Abate AR. Modular barcode beads for microfluidic single cell genomics.
873 *bioRxiv*. 2020 Sep 11;2020.09.10.292326.
- 874 38. Mulqueen RM, Pokholok D, O'Connell BL, Thornton CA, Zhang F, O'Roak BJ, et al. High-
875 content single-cell combinatorial indexing. *bioRxiv*. 2021 Jan 12;2021.01.11.425995.
- 876 39. Langer K, Bremond N, Boitard L, Baudry J, Bibette J. Micropipette-powered droplet based
877 microfluidics. *Biomicrofluidics*. 2018 Jul;12(4):044106.
- 878 40. Kim B, Hahn YK, You D, Oh S, Choi S. A smart multi-pipette for hand-held operation of
879 microfluidic devices. *Analyst*. 2016 Oct 3;141(20):5753–8.
- 880 41. Stoeckius M, Hafemeister C, Stephenson W, Houck-Loomis B, Chattopadhyay PK,
881 Swerdlow H, et al. Simultaneous epitope and transcriptome measurement in single cells.
882 *Nat Methods*. 2017 Sep;14(9):865–8.
- 883 42. Rivello F, Buijtenen E van, Matuła K, Buggenum JAGL van, Vink P, Eenennaam H van, et
884 al. Single-cell intracellular epitope and transcript detection revealing signal transduction
885 dynamics. *bioRxiv*. 2020 Dec 2;2020.12.02.408120.

- 886 43. Xia Y, Whitesides GM. *Soft Lithography*. *Annu Rev Mater Sci*. 1998;28(1):153–84.
- 887 44. Bradski G. *The OpenCV Library*. Dr Dobbs J Softw Tools. 2015 Aug 15;
- 888 45. Maxime De Waegeneer, Chris Campbell Flerin, Kris Davie, Gert Hulselmans. *vib-*
889 *singlecell-nf/vsn-pipelines: v0.25.0* [Internet]. Zenodo; 2021 [cited 2021 May 19]. Available
890 from: <https://zenodo.org/record/4468513>
- 891 46. Moore JE, Purcaro MJ, Pratt HE, Epstein CB, Shores N, Adrian J, et al. Expanded
892 encyclopaedias of DNA elements in the human and mouse genomes. *Nature*. 2020
893 Jul;583(7818):699–710.
- 894 47. Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, et al. *deepTools2: a*
895 *next generation web server for deep-sequencing data analysis*. *Nucleic Acids Res*. 2016
896 Jul 8;44(W1):W160-165.
- 897 48. Kaminow B, Yunusov D, Dobin A. *STARsolo: accurate, fast and versatile*
898 *mapping/quantification of single-cell and single-nucleus RNA-seq data*. *bioRxiv*. 2021 May
899 5;2021.05.05.442755.
- 900 49. Wolf FA, Angerer P, Theis FJ. *SCANPY: large-scale single-cell gene expression data*
901 *analysis*. *Genome Biol*. 2018 Feb 6;19(1):15.
- 902 50. Waskom ML. *seaborn: statistical data visualization*. *J Open Source Softw*. 2021 Apr
903 6;6(60):3021.
- 904 51. Hunter JD. *Matplotlib: A 2D Graphics Environment*. *Comput Sci Eng*. 2007 May;9(3):90–5.
- 905 52. Kennedy A. *mouse brain silhouette* [Internet]. 2020 [cited 2021 May 19]. Available from:
906 <https://zenodo.org/record/3925919>

907

908 **Droplet based single-cell ATAC-seq and single-cell RNA-seq with dissolvable hydrogel**
909 **beads using HyDrop**

910 Florian V. De Rop^{1,2,4}, Joy N. Ismail^{1,2}, Carmen Bravo González-Blas^{1,2}, Gert J. Hulselmans^{1,2},
911 Christopher C. Flerin^{1,2,4}, Jasper Janssens^{1,2}, Koen Theunis^{1,2,4}, Valerie M. Christiaens^{1,2}, Jasper
912 Wouters^{1,2}, Gabriele Marcassa^{1,3}, Joris de Wit^{1,3}, Suresh Poovathingal^{1,4,#}, and Stein Aerts^{1,2,4,#}

913 ¹ VIB-KU Leuven Center for Brain & Disease Research

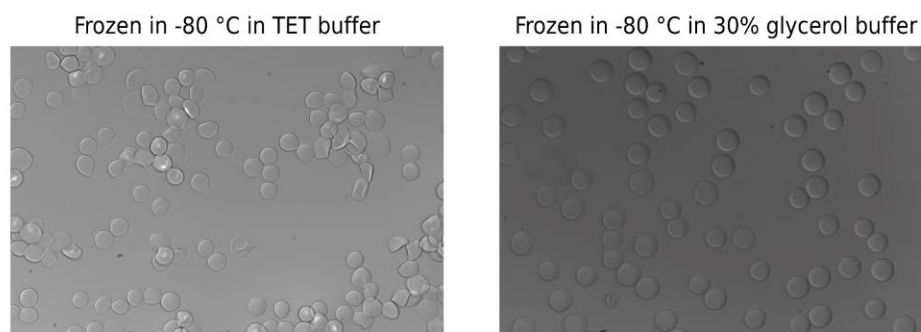
914 ² Laboratory of Computational Biology, Department of Human Genetics, KU Leuven

915 ³ Laboratory of Synapse Biology, Department of Neurosciences, KU Leuven

916 ⁴ Aligning Science Across Parkinson's (ASAP) Collaborative Research Network, Chevy Chase,
917 MD,= 20815.

918 #Shared last author; correspondence to suresh.poovathingal@kuleuven.be and
919 stein.aerts@kuleuven.be.

920



921
922 **Figure 1 – figure supplement 1. Hydrogel bead integrity after being frozen and thawed.**
923 Beads were frozen overnight in TET buffer or 30% glycerol buffer and thawed at room
924 temperature. Rough edges in the TET buffer sample indicate freezing/thawing damage.
925

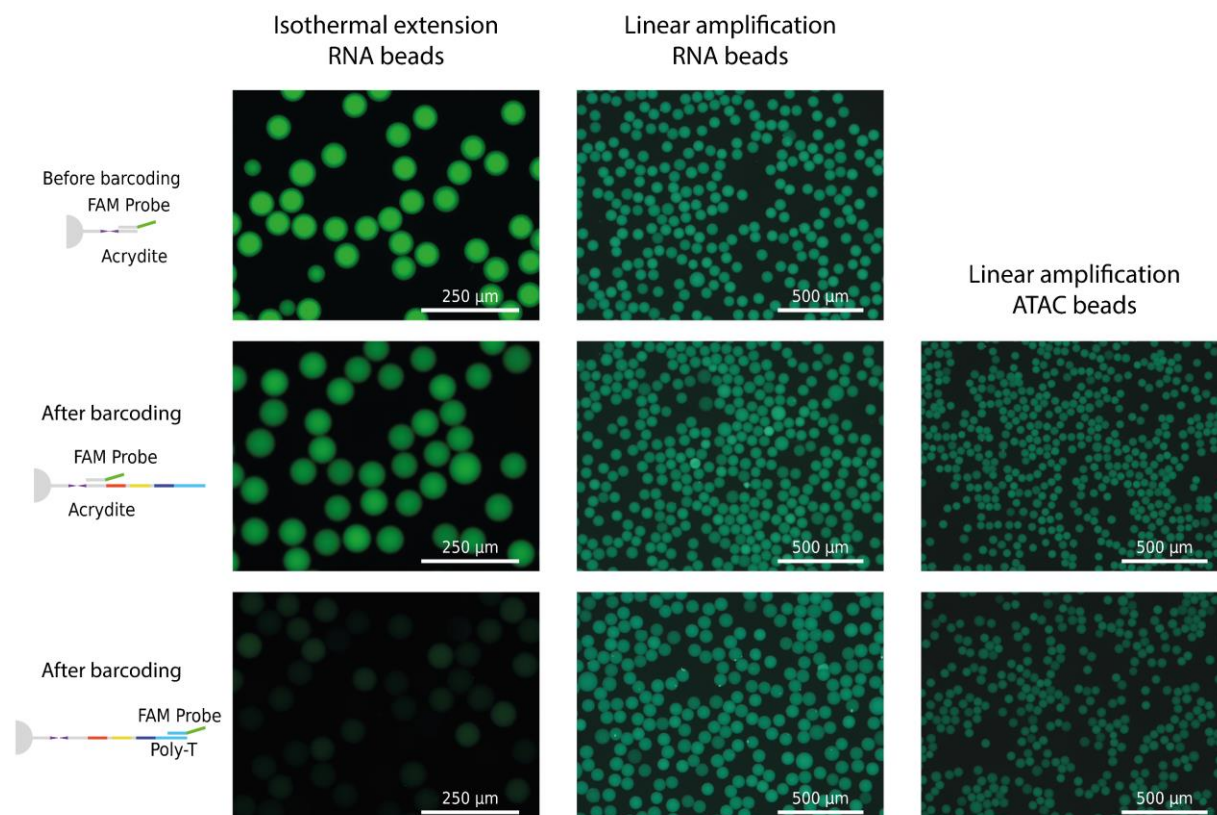


Figure 1 – figure supplement 2. Fluorescence signal of freshly barcoded beads. Beads were incubated with probes complementary to acrydite primer or capture site (poly-T or Tn5 adapter site) before and after barcoding.

Method	Acrydite (uM)	Type	Stage	Probe	n	dmean (μm)	imean
Isothermal ext.	50	RNA	Pre BC	Acrydite	34	57.63 ± 1.68	102.03 ± 1.1
			Post BC	Acrydite	43	61.44 ± 2.21	76.18 ± 14.3
			Post BC	PolyA	40	59.6 ± 2.84	16.51 ± 5.85
Linear amp.	12	RNA	Pre BC	Acrydite	271	63.16 ± 2.4	60.42 ± 4.91
			Pre BC	Acrydite	357	62.53 ± 2.59	52.3 ± 6.64
			Post BC	PolyA	243	64.87 ± 2.73	51.95 ± 6.71
		ATAC	Post BC	Acrydite	517	45.91 ± 1.3	44.82 ± 5.3
			Pre BC	ATAC	351	45.96 ± 1.09	36.33 ± 5.9

Table S1. Count, mean diameter and mean intensity of linear amplification and isothermal extension beads. Beads were incubated in fluorescent probes complementary to basal acrydite primer or poly-A or ATAC capture site. Absolute size of 1 standard deviation is added in mean diameter and mean intensity columns.

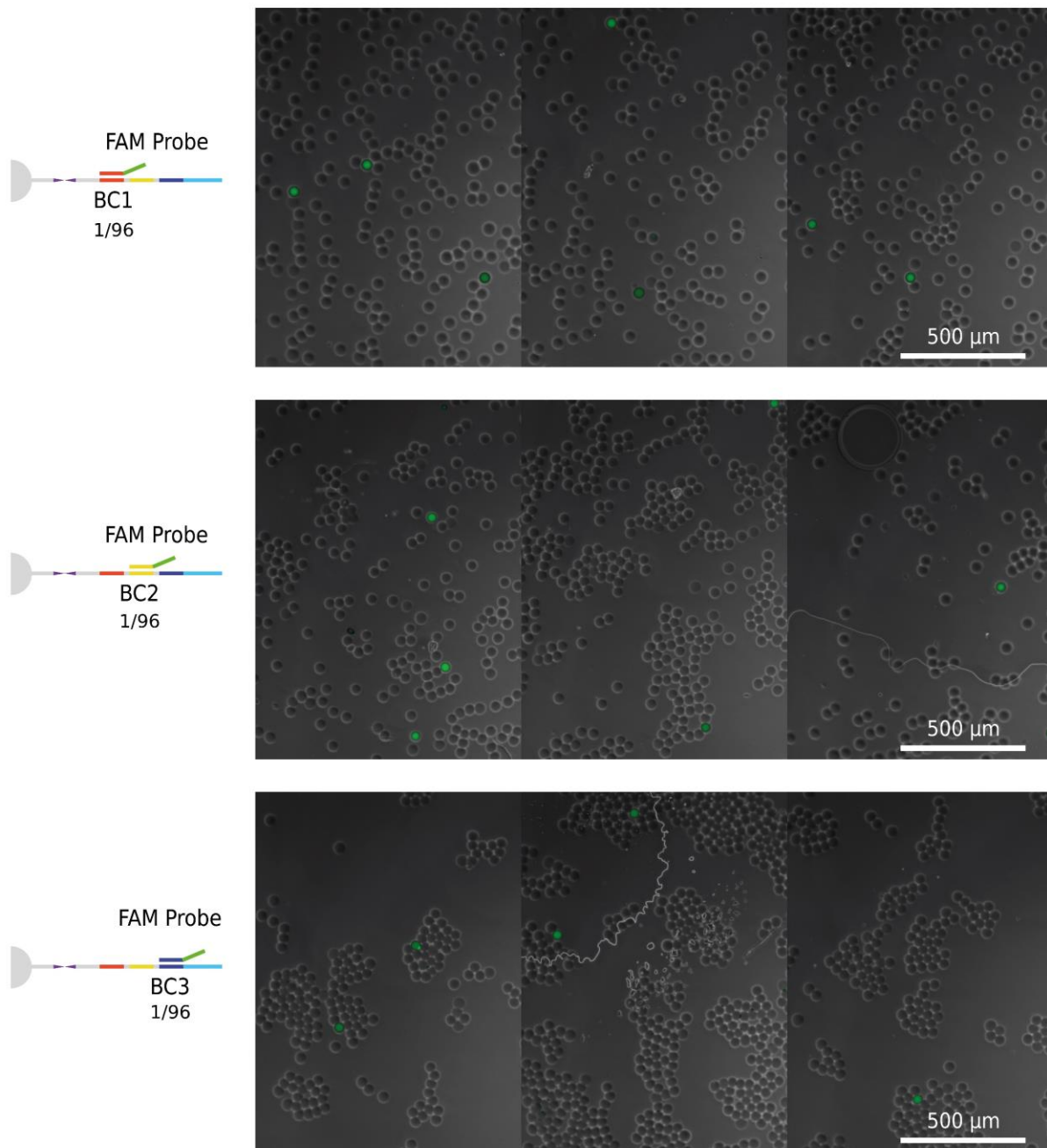


Figure 1 – figure supplement 3. Brightfield images of barcoded beads incubated with FAM oligonucleotide probe complementary to one out of 96 possible sub-barcodes. Segmented fluorescence masks of positive beads overlaid over brightfield image of same beads to visualise positives.

943

	n	Positive	Rate (%)	1 in ...
Round 1	570	7	1.23%	81.4
Round 2	425	7	1.65%	60.7
Round 3	669	5	0.75%	133.8
Total	1664	19	1.14%	87.6

944

945

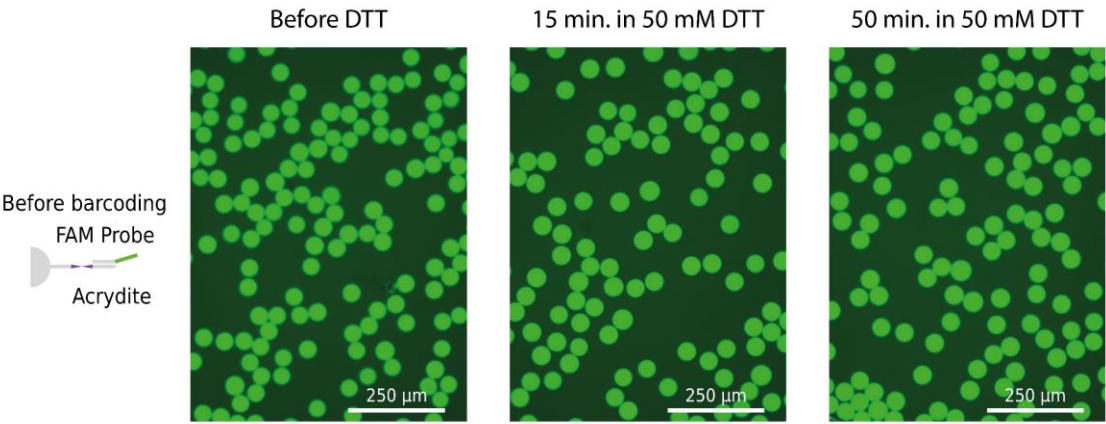
946

947

948

949

Table S2. Table supplementary to figure 1 – figure supplement 3. Total counts and fluorescence-positive counts of beads incubated with FAM oligonucleotide probe complementary to one out of 96 barcode possibilities.



950

951

952

953

954

955

Figure 1 – figure supplement 4. Fluorescence signal of non-dissolvable 50 uM HyDrop-RNA beads. Unbarcoded beads were incubated with probes complementary to the acrydite primer. Images taken immediately after 15 and 50 minute incubation in 50 mM of DTT, and a short wash using QC buffer.

Time in DTT	n	dmean (µm)	imean
0 min.	108	47.77 ± 1.69	82.2 ± 2.63
15 min.	96	48.14 ± 1.48	81.98 ± 2.32
50 min.	112	46.57 ± 1.81	80.24 ± 3.22

956

957

958

959

Table S3. Count, mean diameter and mean background-adjusted intensity of beads incubated in 50 mM of DTT for 15 and 50 minutes as well as negative control. Absolute size of 1 standard deviation is added in mean diameter and mean intensity columns.

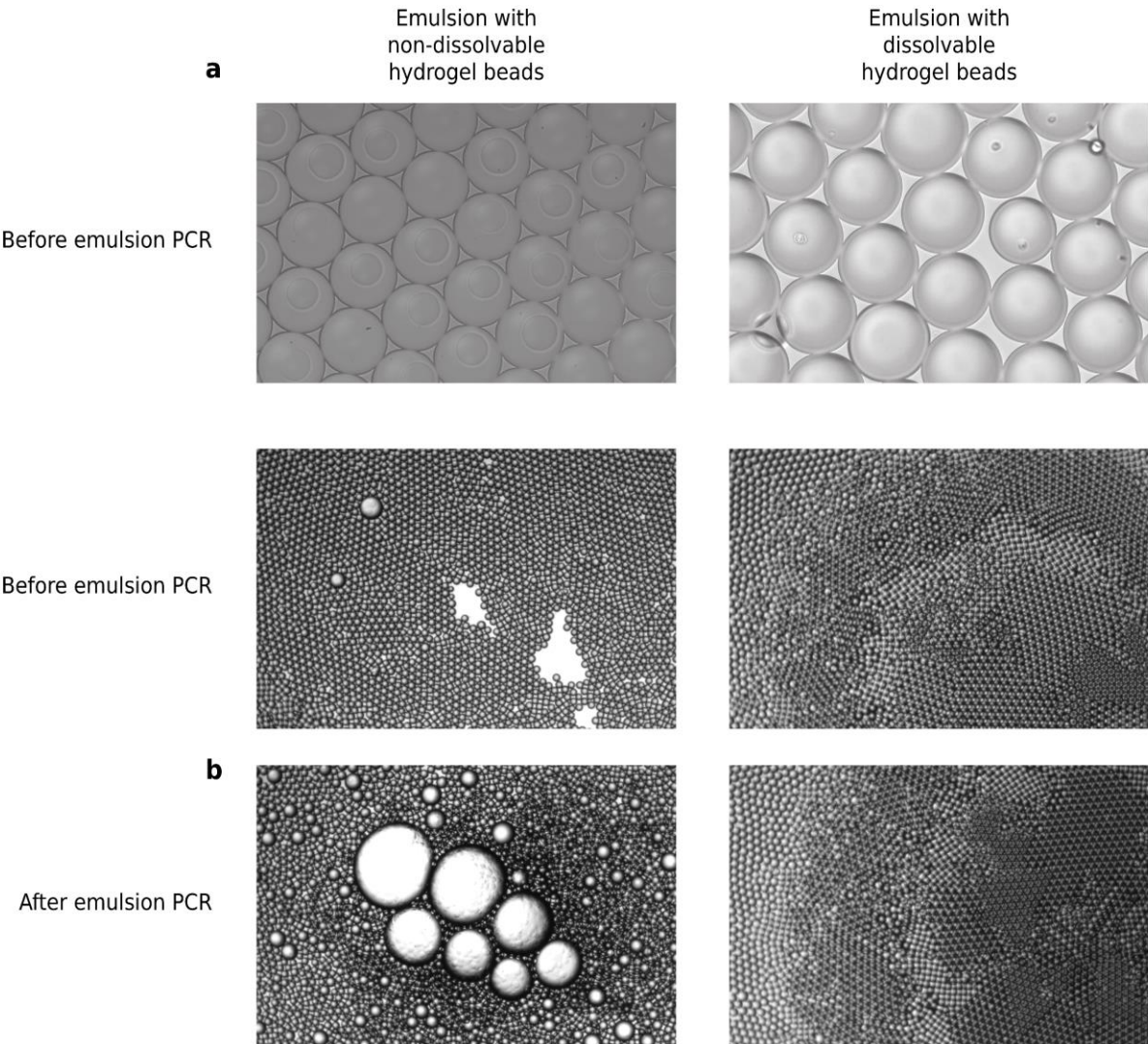


Figure 1 – figure supplement 5. a. HyDrop emulsions made with non-dissolvable and dissolvable beads. Picture taken within 5 minutes after the emulsion was made. **b. HyDrop-ATAC emulsions made with non-dissolvable and dissolvable beads before and after thermocycling.** Picture taken within 5 minutes after the thermocycling was completed.

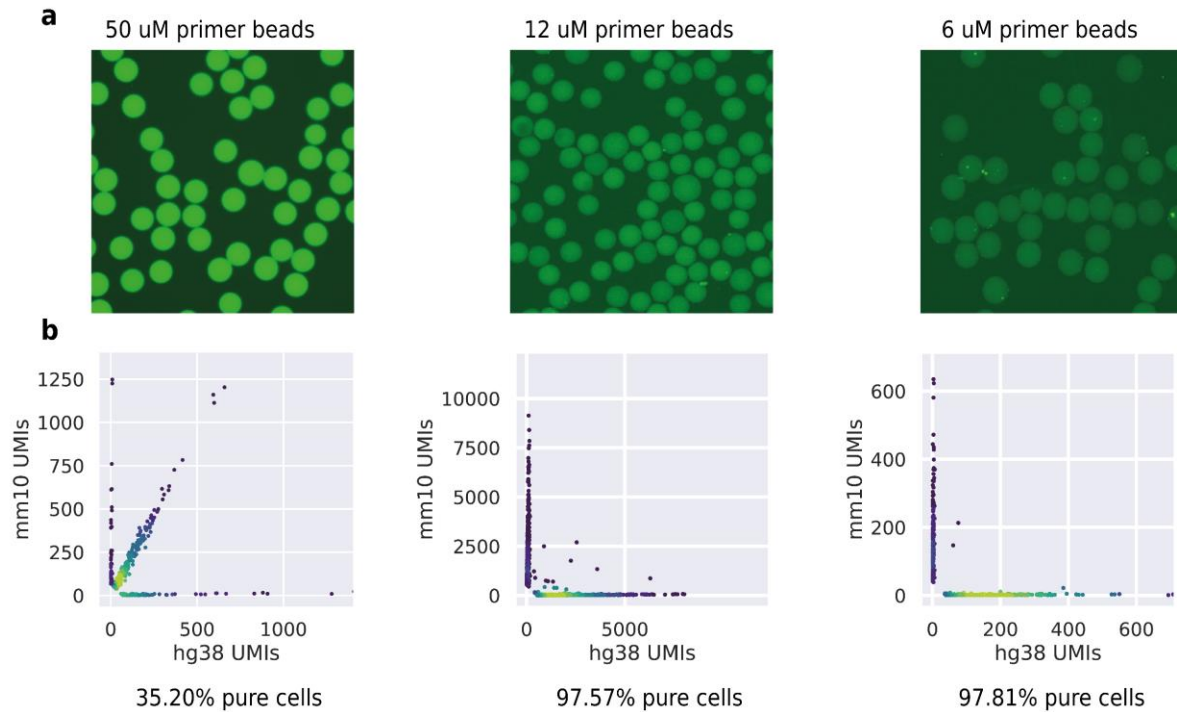


Figure 1 – figure supplement 6. a. Fluorescence signal of 50, 12 and 6 μ M primer concentration beads. Unbarcoded beads were incubated with probes complementary to acrydite primer. **b. Purity plots for HyDrop-RNA species-mixing experiments performed using 50, 12 and 6 uM beads.** Cells were assigned to a species if >95% of their UMIs belonged to transcripts mapping to that species' genome.

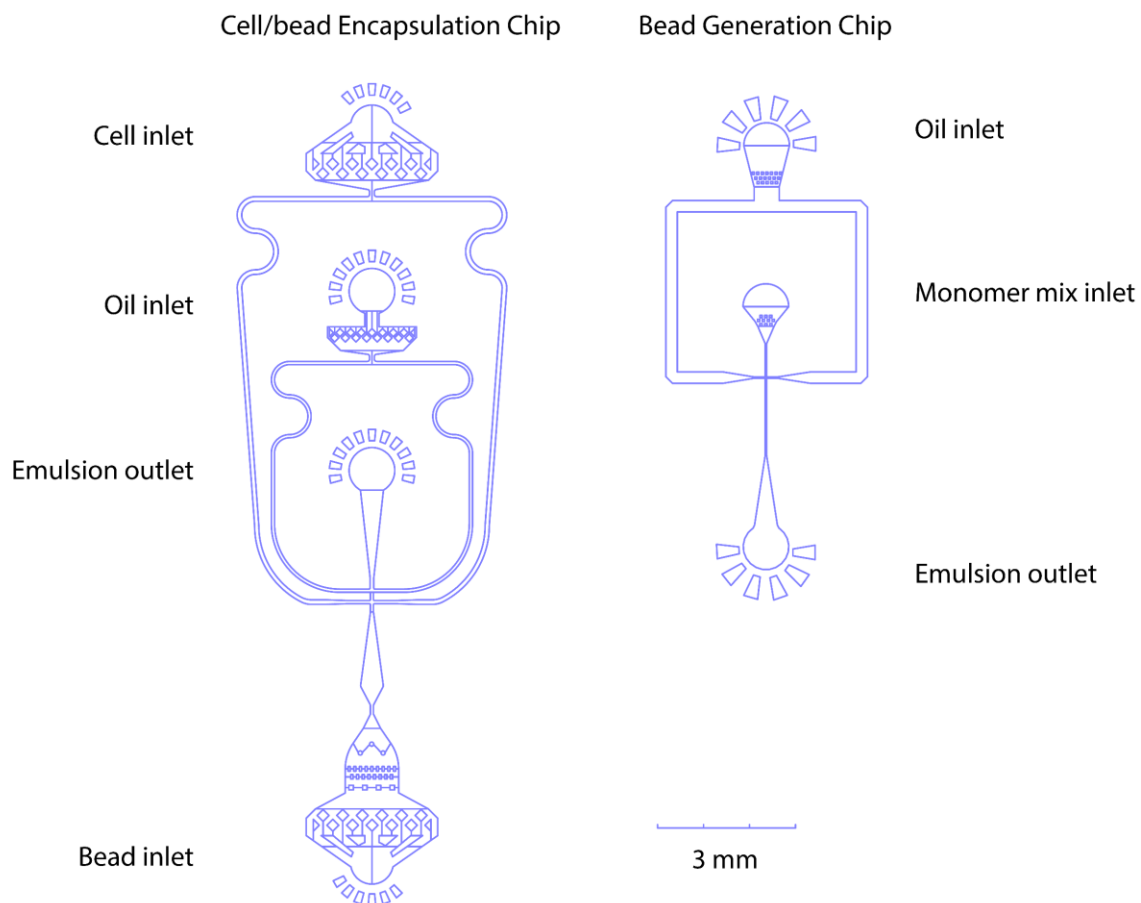


Figure 3 - figure supplement 1. Microfluidic chip designs for cell/bead encapsulation and hydrogel bead generation - Oil, monomer mix, cell suspensions and beads can be pumped in through their respective inlets. At the flow focusing point, aqueous phases (cells/beads/monomer mix) are emulsified by collision with the oil phase at the flow-focusing point. Emulsion is collected at the outlet channel.

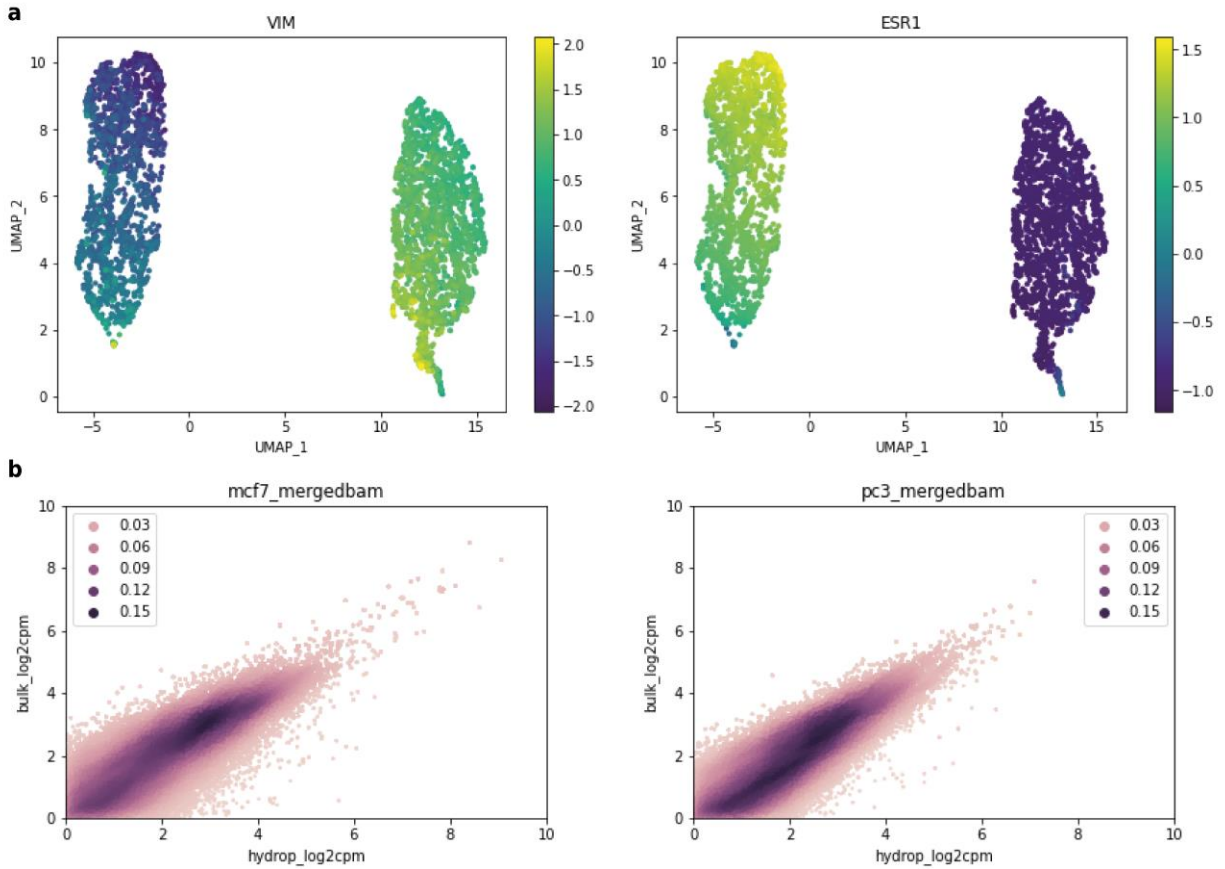


Figure 4 – figure supplement 1. a. MCF-7/PC-3 HyDrop-ATAC UMAP colored by imputed gene activity for Vimentin (VIM) and Estrogen Receptor 1 (ESR1) – VIM is a marker gene for PC-3 cells, ESR1 is a marker gene for MCF-7 cells. b. Correlation of counts in peak regions of public bulk ATAC-seq and HyDrop-ATAC data – Regions were derived from public data, and counts within these regions were correlated for both the HyDrop and bulk data. Color scale encodes each count's local Gaussian kernel density estimation score.

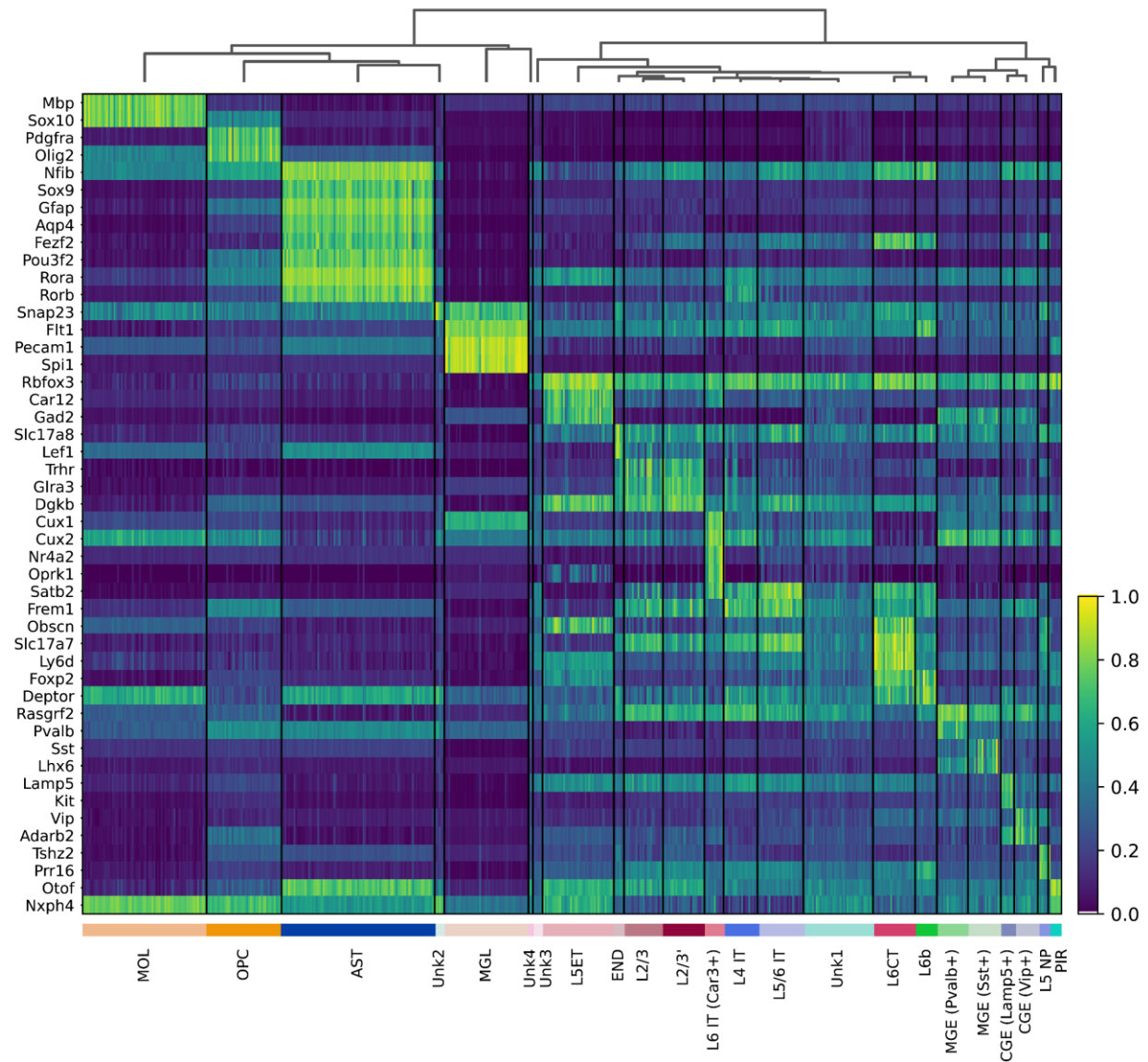


Figure 5 – figure supplement 1. Heatmap of mouse cortex HyDrop-ATAC gene activity –
 Gene activity was imputed by normalised accessibility within a 10 kb window around the gene.

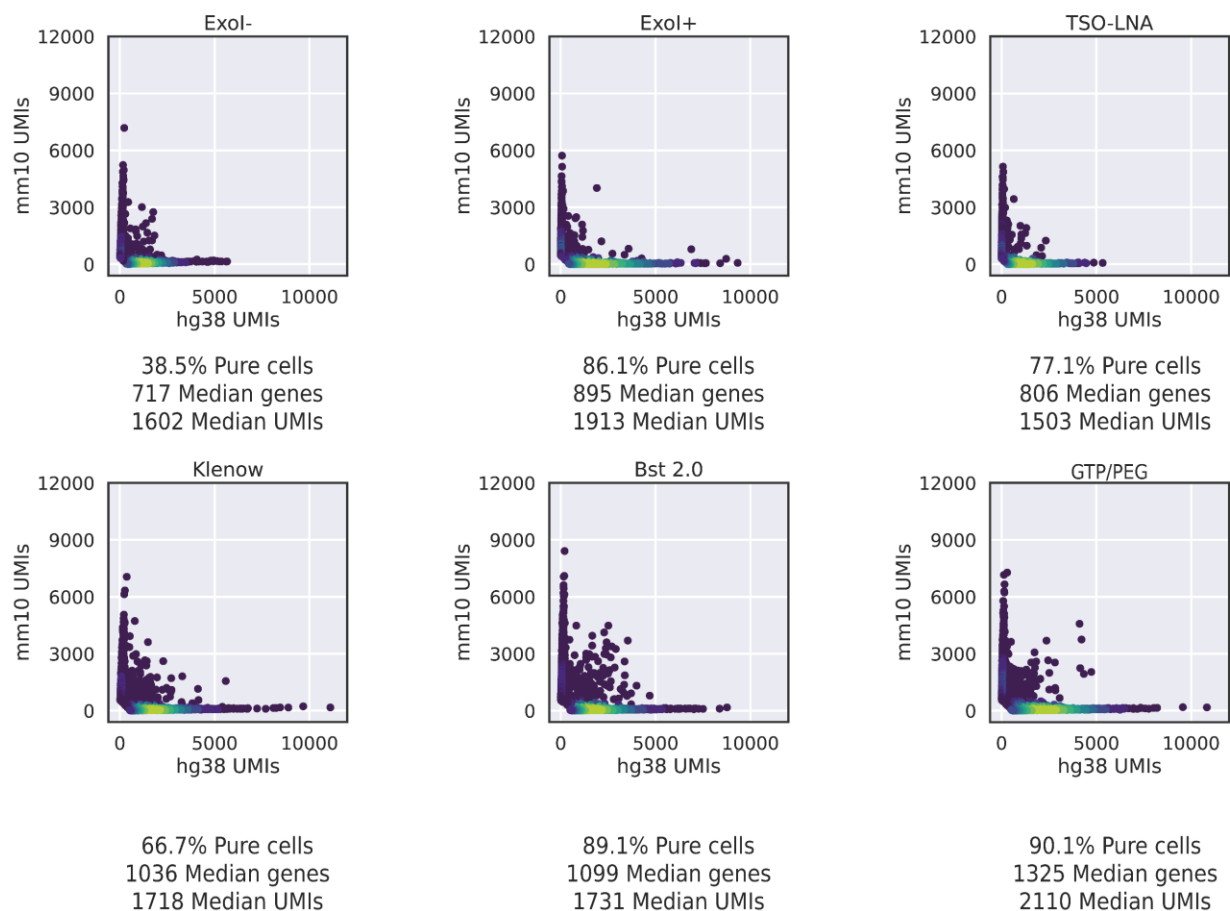


Figure 8 – figure supplement 1. HyDrop-RNA species-mixing purity plots for several different protocol versions. Protocol versions include non-Exo I treated, Exo I treated, TSO-LNA, Klenow fragment, Bst2.0, and GTP/PEG HyDrop-RNA libraries.

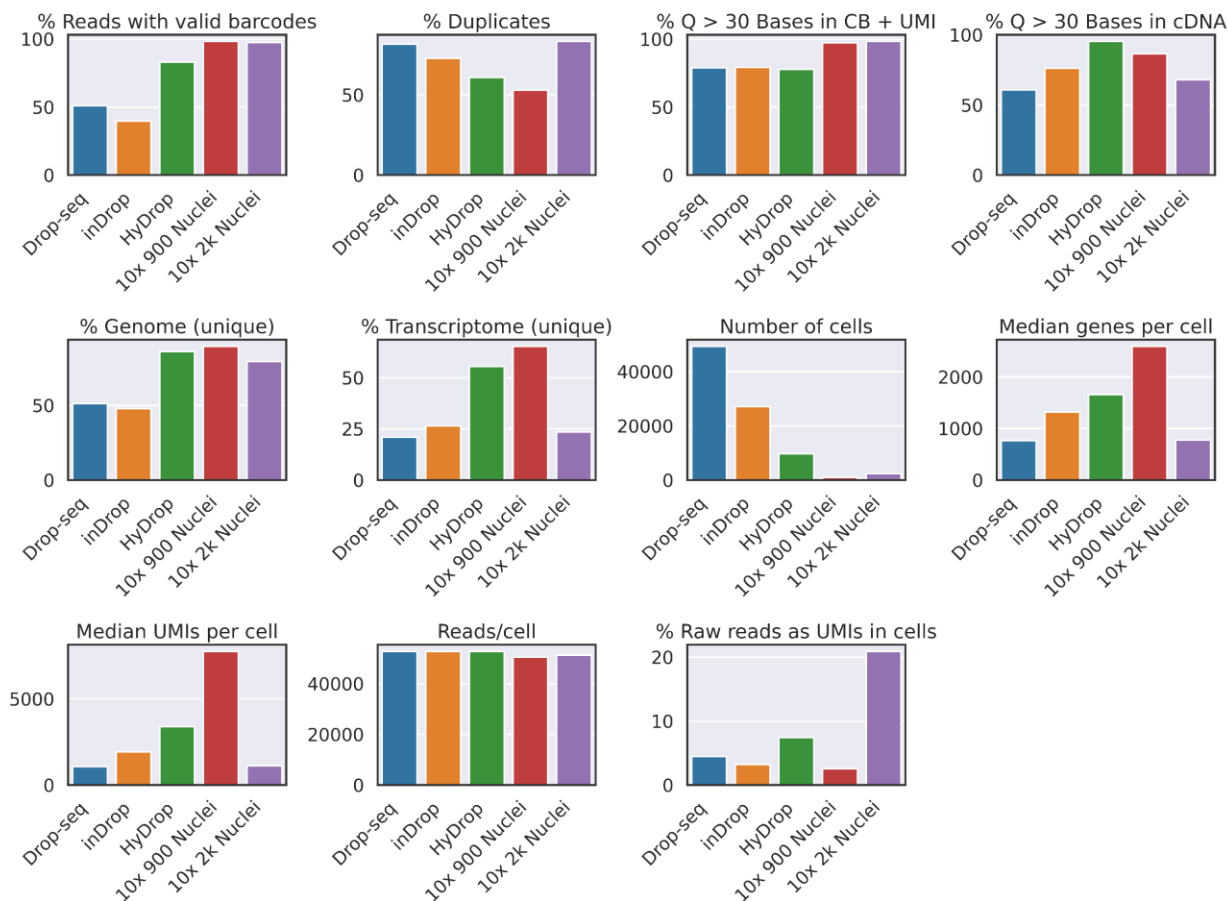
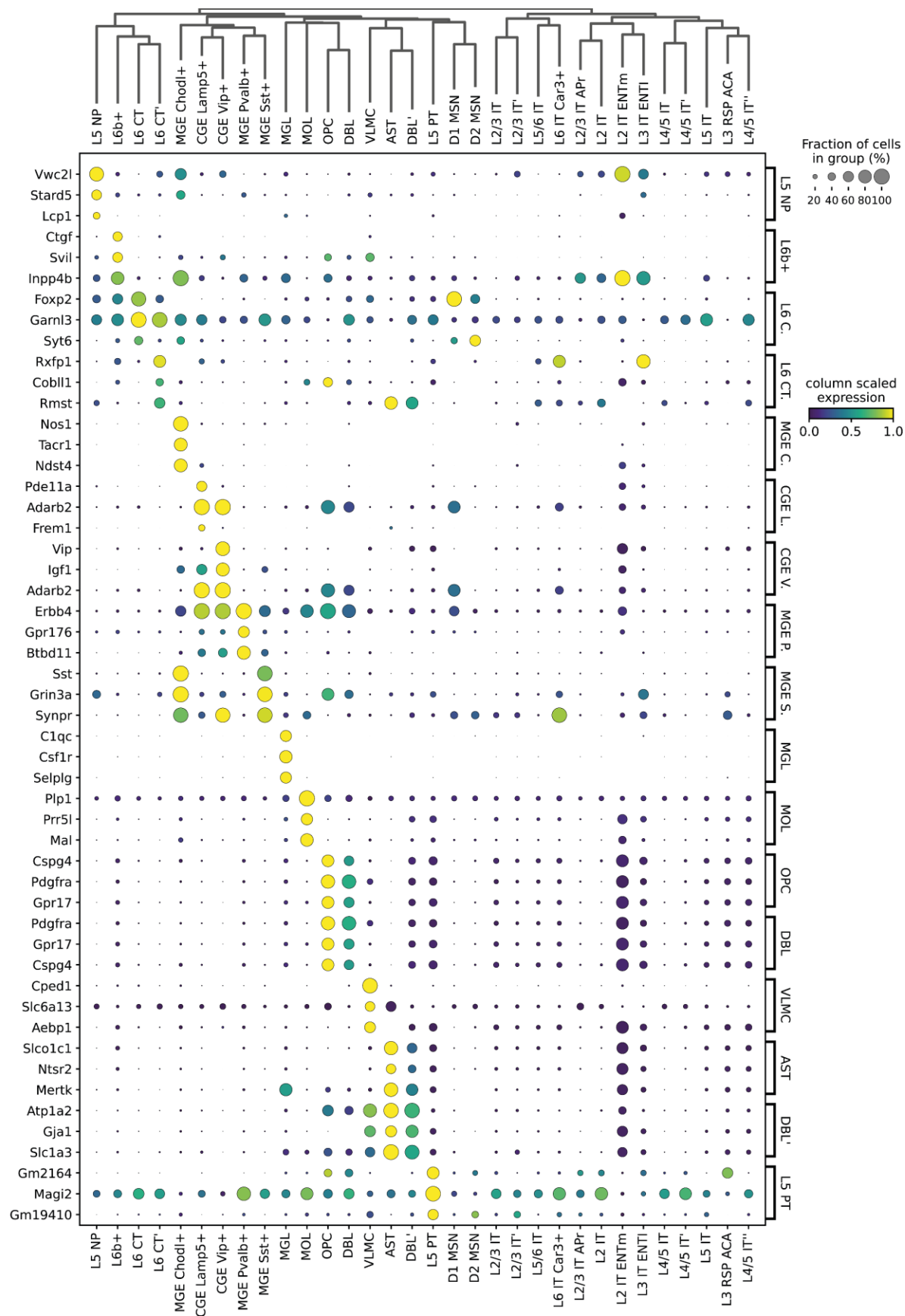


Figure 8 – figure supplement 2. Collection of quality control metrics for inDrop, Drop-seq, HyDrop-RNA and 10x datasets on mouse brain cells. “% Reads with valid barcodes” is calculated after allowing 1 mismatch to the whitelist. “% Duplicates” is calculated on mapped transcripts using UMIs. “% Genome” and “% Transcriptome” are the percentages of all raw reads mapping to the reference genome and transcriptome respectively. “% Raw reads as UMIs in cells” is the sum of UMIs in filtered cells divided by the total number of sequenced reads. Supplementary source data files available.



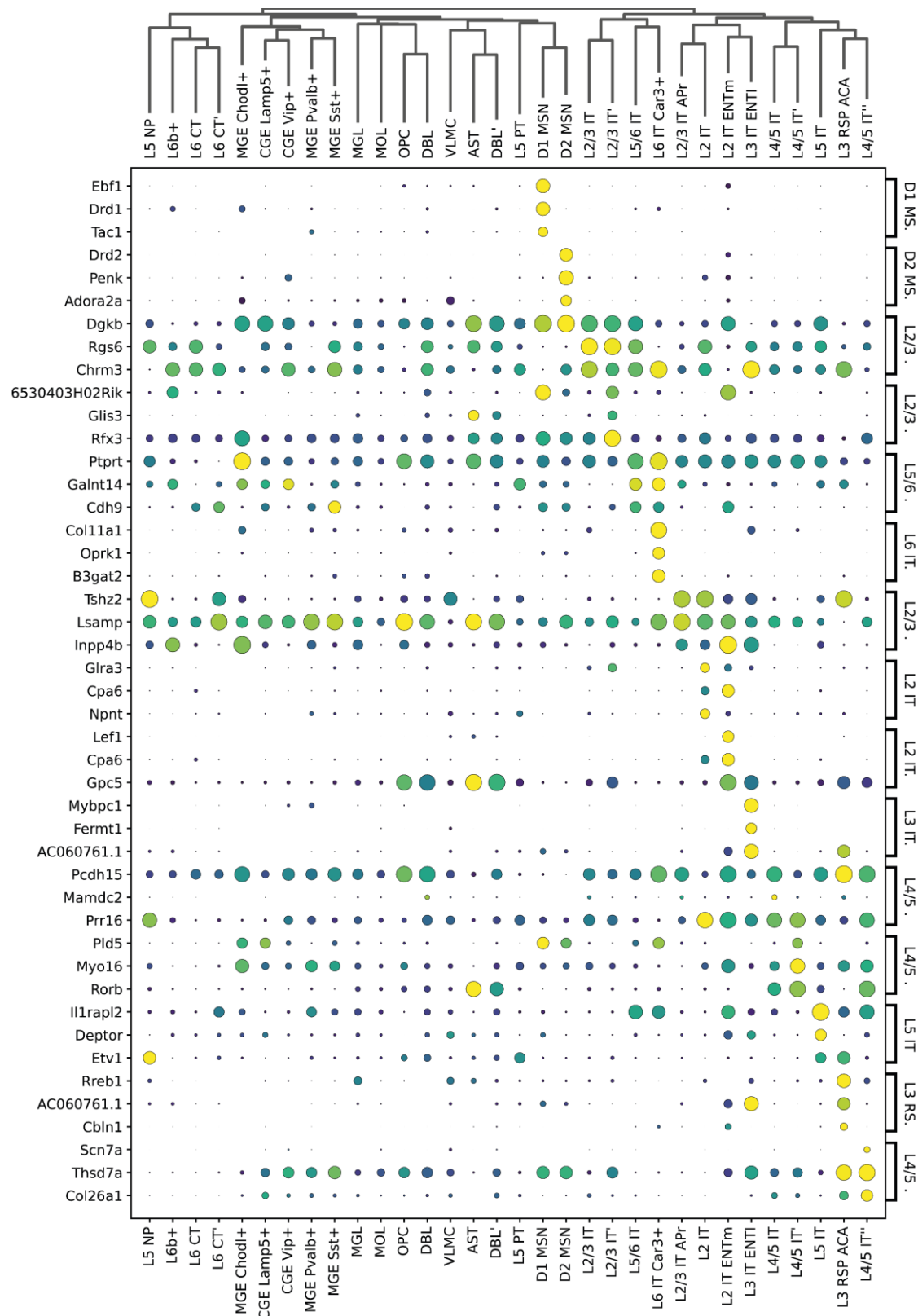
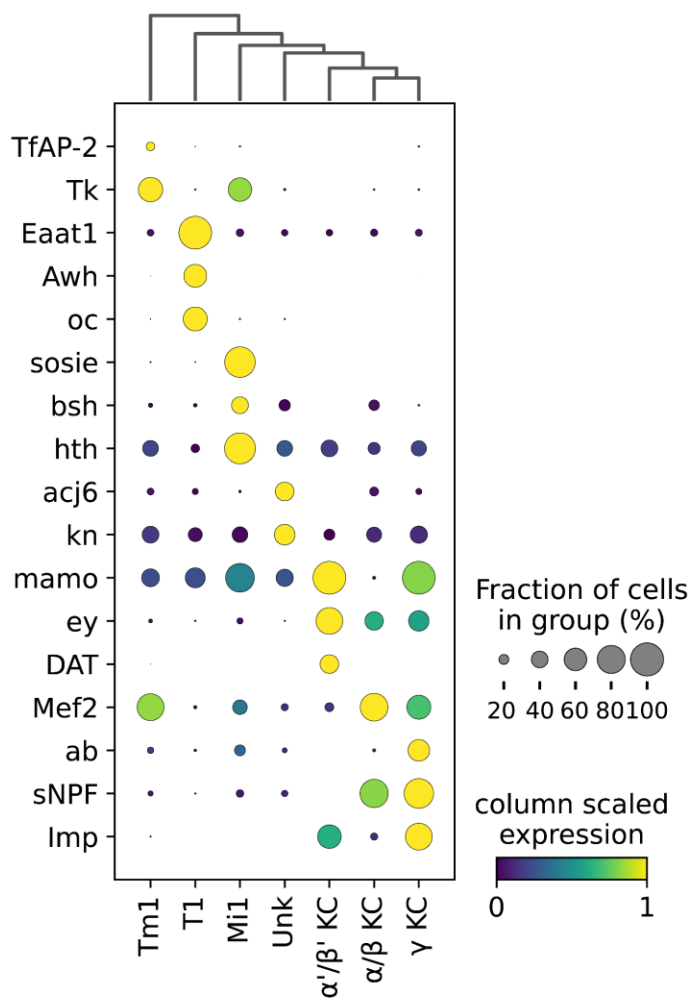


Figure 9 – figure supplement 1. Expression of HyDrop-RNA mouse cortex top 3 differentially expressed genes from each cluster. Dot sizes denote fraction of cluster in

1016 which the gene (row) was expressed. Dot colour encodes the column min/max-scaled
 1017 expression of the gene (row).



1018
 1019
 1020 **Figure 10 – figure supplement 1. Expression of marker genes of HyDrop-RNA on FAC-**
 1021 **sorted fly brain cells.** Dot sizes denote fraction of cluster in which the gene (row) was
 1022 expressed. Dot colour encodes the column min/max-scaled expression of the gene (row).
 1023

1024 **List of figure source data**

1025 Figure 4a source data file – Counts of fragments mapping to either mouse or human reference
1026 genome for each cell.

1027 Figure 5b, c, d source data file – Metadata for each cell, including number of unique fragments,
1028 TSS enrichment, FRIP and duplication rate.

1029 Figure 8a source data file – Number of genes and UMIs detected for each cell.

1030 Figure 8b source data file – Counts of transcripts mapping to either mouse or human reference
1031 genome for each cell, separated by method.

1032 Figure 8c source data file – Number of genes and UMIs detected for each cell, separated by
1033 method.

1034 Figure 8 – supplementary figure 2 – Mapping and quality control statistics for inDrop, HyDrop-
1035 RNA, Drop-seq and 10x scRNA datasets

1036 Figure 10b source data file – Number of genes and UMIs detected for each cell.

1037

1038 **List of supplementary files**

1039 "Supplementary File 1 – "Molecular sequence description of HyDrop bead barcoding": visual
1040 description of progression of nucleotide sequences involved in every step of the barcoding
1041 process

1042 "Supplementary File 2 – "Molecular sequence description of HyDrop-ATAC": visual description
1043 of progression of nucleotide sequences involved in every step of HyDrop-ATAC

1044 "Supplementary File 3 – "Molecular sequence description of HyDrop-RNA": visual description of
1045 progression of nucleotide sequences involved in every step of HyDrop-RNA

1046 "Supplementary File 4 - "Reagents and oligonucleotide list": list of all reagents and
1047 oligonucleotides used in HyDrop

1048 "Supplementary File 5 - "Price calculation for HyDrop bead barcoding and HyDrop-ATAC":
1049 detailed price calculation for HyDrop bead barcoding and HyDrop-ATAC