

# Hybridization-based *in situ* sequencing (HybISS) for spatially resolved transcriptomics in human and mouse brain tissue

Daniel Gyllborg<sup>1</sup>, Christoffer Mattsson Langseth<sup>1</sup>, Xiaoyan Qian<sup>1</sup>, Eunyoung Choi<sup>1</sup>, Sergio Marco Salas<sup>1</sup>, Markus M. Hilscher<sup>1</sup>, Ed S. Lein<sup>2</sup> and Mats Nilsson<sup>1,\*</sup>

<sup>1</sup>Science for Life Laboratory, Department of Biochemistry and Biophysics, Stockholm University, 171 65 Solna, Sweden and <sup>2</sup>Allen Institute for Brain Science, Seattle, WA 98109, USA

Received March 06, 2020; Revised August 19, 2020; Editorial Decision September 10, 2020; Accepted September 11, 2020

## ABSTRACT

Visualization of the transcriptome *in situ* has proven to be a valuable tool in exploring single-cell RNA-sequencing data, providing an additional spatial dimension to investigate multiplexed gene expression, cell types, disease architecture or even data driven discoveries. *In situ* sequencing (ISS) method based on padlock probes and rolling circle amplification has been used to spatially resolve gene transcripts in tissue sections of various origins. Here, we describe the next iteration of ISS, HybISS, hybridization-based *in situ* sequencing. Modifications in probe design allows for a new barcoding system via sequence-by-hybridization chemistry for improved spatial detection of RNA transcripts. Due to the amplification of probes, amplicons can be visualized with standard epifluorescence microscopes for high-throughput efficiency and the new sequencing chemistry removes limitations bound by sequence-by-ligation chemistry of ISS. HybISS design allows for increased flexibility and multiplexing, increased signal-to-noise, all without compromising throughput efficiency of imaging large fields of view. Moreover, the current protocol is demonstrated to work on human brain tissue samples, a source that has proven to be difficult to work with image-based spatial analysis techniques. Overall, HybISS technology works as a targeted amplification detection method for improved spatial transcriptomic visualization, and importantly, with an ease of implementation.

## INTRODUCTION

The field of spatially resolved transcriptomic technologies is emerging as a vital tool to profile gene-expression, continu-

ously pushing current methods to accommodate larger gene panels and larger areas without compromising throughput efficiency. In the era of single-cell transcriptomics, large quantities of data from single-cell RNA-sequencing (scRNA-seq) has revealed the true transcriptomic diversity of cell types within organisms in high detail with such technologies becoming more accessible to laboratories around the world (1–3). The exponential growth of sequencing large numbers of cells within the last decade has allowed for a systematic assignment to defining cell types, cell states or transitions between them more clear (4–6). Capturing only static states of cells has been one drawback to sequencing methods and implementation of computational methods trying to overcome this are currently being used, even with *in situ* hybridization data (7–9). However, one major drawback to scRNA-seq methods that would be difficult to solve computationally is that it requires the cellular dissociation of a tissue to collect cells which could lead to under sampling of vulnerable cell types and more importantly, any spatial context of a cell within a given tissue being lost.

The field of spatial transcriptomics, assigning transcripts a positional location *in situ*, has emerged as a vital tool in the validation of scRNA-seq data and exploring the transcriptomic profiles and cellular architecture across tissues. *In situ* hybridization technologies have come a long way and large consortium efforts are realizing the importance of not only being able to define all the cell types within a human being but also giving them a spatial position in the form of a human cell atlas (10,11). Image-based transcriptomics have shown the possibility to detect a various range of gene numbers within tissue sections to different degrees of sensitivity and accuracy (12–14), some used to explore data driven discoveries of biological studies such as behavior based cellular changes (15,16).

Spatial distribution of RNA transcripts for cell-type mapping can give further understanding to arrangement of complex tissues (16–18). Multiplexed *in situ* hybridization

\*To whom correspondence should be addressed. Tel: +46 730537876; Email: mats.nilsson@scilifelab.se  
Present address: Xiaoyan Qian, CartaNA AB, 171 65 Solna, Sweden.

offers the possibility to explore cellular diversity at sub-cellular resolution in an upscaled approach. As an example, our lab has developed *in situ* sequencing (ISS) to be used to detect RNA isoforms (19), transcriptomic distribution (20), and cell typing across tissue sections (21). The established ISS method based on barcoded padlock probes (PLPs) and amplification through rolling circle amplification (RCA) has shown robust detection of RNA for various applications, however, further improvements are needed to meet the upscaling demands to explore cellular diversity of scRNA-seq data across large tissue areas of various origins.

Here, we demonstrate an improved ISS method based on the principles of PLPs and RCA to overcome the inherent limitations of current ISS to address the requirements to investigate spatial mapping of larger gene panels across entire tissue sections. The second iteration of the ISS method uses a sequence-by-hybridization (SBH) chemistry approach to define the location of mRNA molecules within tissue (HyBISS: hybridization-based *in situ* sequencing). The unique feature of the ISS method is the probe amplification and the integrated barcoding system that can be decoded across sequential rounds of probing, imaging and stripping. Specifically, the HyBISS method takes a new approach to combinatorial barcoding, allowing for more robust detection of molecules and its ability to spatially resolve gene expression *in situ* has been upscaled, supporting the accommodation of larger gene panels, which was one of the main limitations of the sequence-by-ligation (SBL) methods of ISS (20,22). Here, we present HyBISS along with an example application to demonstrate the possibility to explore the architecture of human brain tissue and whole mouse brain coronal sections. This results in data sets that can be used to further improve image-based spatial analysis methods such as transcript-spot calling, cell segmentation, and cell typing. Overall, HyBISS has benchmarked higher than SBL-based ISS, expanding the possibilities of high-throughput image-based *in situ* transcriptomics.

## MATERIALS AND METHODS

### Protocol

A step-by-step protocol can be found online at [protocols.io](https://www.protocols.io) and is updated with any modifications: [dx.doi.org/10.17504/protocols.io.xy4fpyw](https://doi.org/10.17504/protocols.io.xy4fpyw) (Version 1).

### Gene selection

Gene panels were curated both manually and computationally. The panels were based on scRNA-seq data from mouse primary visual cortex (5) and human middle temporal gyrus (23,24). Gene list panels were generated as part of the Chan Zuckerberg Initiative SpaceTx Consortium, Working Groups 2 and 3. From the curated ranked gene list, a subset was selected to develop and perform HyBISS.

### Probe design

Oligonucleotides for PLPs, bridge-probes, detection readout probes, and anchors were ordered from Integrated DNA Technologies (IDT) or IBA Lifesciences. Stocks of 100  $\mu$ M (TE buffer, pH 8.0) for bridge-probes and detection

readout probes were stored at  $-20^{\circ}\text{C}$ . PLPs were ordered as large pools and phosphorylated with T4 Polynucleotide Kinase (New England Biolabs). Sequences of oligonucleotides can be found in Supplementary Table S1. Target sequences for the selected genes were obtained using in-house Python padlock design software package that utilizes ClustalW and BLAST+ ([https://github.com/Moldia/multi\\_padlock\\_design](https://github.com/Moldia/multi_padlock_design)) with the following parameters: arm length, 15; Tm, low 65, high 75; space between targets, 15. After target sequences were obtained, five targets were selected randomly per gene. If fewer targets were found then only those were selected. The backbone of the PLPs include a 20 nucleotide (nt) ID sequence obtained from the Affymetrix GenFlex Tag 16K Array and a 20 nt sequence ‘anchor’ that is common among subsets of PLPs, a total six of these anchor sequences were used, but were not used for decoding and served only as linker sequences in this study. The decoding scheme for then designing bridge-probes was generated via an in-house MATLAB script that specifies number of positions and minimum base difference between two individual barcodes, hamming distance set to two for five cycles of decoding (<https://github.com/Moldia/iss-analysis/tree/master/lib/barcode>).

### Tissue collection and preparation

All animal procedures were approved by the Institutional Animal Care and Use Committee at the Allen Institute for Brain Science (Protocol No. 1511). Mouse tissue was obtained from the Allen Institute for Brain Science under the SpaceTx consortium. Fresh whole mouse brain tissue cryopreserved in optimal cutting temperature (OCT) media was stored at  $-80^{\circ}\text{C}$  until sectioning. Tissue was sectioned with a cryostat (CryoStar™ NX70) at 10  $\mu$ m and collected on SuperFrost Plus microscope slides.

Human postmortem tissue collection and processing was performed as previously described in Hodge *et al.* (24), in accordance with the provisions of the United States Uniform Anatomical Gift Act of 2006 described in the California Health and Safety Code section 7150 (effective 1 January 2008) and other applicable state and federal laws and regulations. Human tissue was received from the Allen Institute for Brain Science under the SpaceTx consortium. 10  $\mu$ m thick sections were received already mounted on adhesive microscope slides on dry ice and stored at  $-80^{\circ}\text{C}$  until used. Tissue from  $-80^{\circ}\text{C}$  was first allowed to reach room temperature (RT) ( $\sim 5$  min) before 3% (v/v) formaldehyde fixation (5 min for mouse, 30 min for human tissue).

### HyBISS

We refer to the accompanying [www.protocols.io](https://www.protocols.io) link for an in-depth protocol on the method, reagents used and their concentrations. *Reverse transcription*: After fixation, sections were permeabilized with 0.1 M HCl for 5 min and washed with PBS. SecureSeal™ Hybridization Chambers (Grace Bio-Labs) were applied around tissue sections and mRNA was reverse transcribed priming with random decamers, RNase inhibitor and reverse transcriptase (BLIRT) overnight at  $37^{\circ}\text{C}$ . *PLP hybridization and ligation*: Tissue sections were fixed for 40 min post reverse transcription

and subsequently washed with PBS. Phosphorylated PLPs were hybridized at a final concentration of 10 nM/PLP, and ligated in the same reaction with Tth Ligase (BLIRT) and RNaseH. This was performed at 37°C for 30 min and then moved to 45°C for 1.5 h. *Rolling circle amplification*: Sections were washed with PBS and RCA was performed with phi29 polymerase (BLIRT) and Exonuclease I (Thermo Fisher Scientific) overnight at 30°C. *Autofluorescence quenching*: Human sections were treated with TrueBlack Lipofuscin Autofluorescence Quencher (TLAQ) (Biotium) per manufactures instructions, for 45 s and immediately washed with PBS. SecureSeal chambers were then removed. *Bridge-probe hybridization*: Bridge-probes (10 nM) were hybridized at RT for 1 h in hybridization buffer (2× SSC, 20% formamide). *Readout detection probe hybridization*: This was followed by hybridization of readout detection probes (100 nM) and DAPI (Biotium) in hybridization buffer for 2 h at RT. Sections were washed with PBS and mounted with SlowFade Gold Antifade Mountant (Thermo Fisher Scientific).

### Imaging

Imaging was performed using a standard epifluorescence microscope (Zeiss Axio Imager.Z2) connected to external LED source (Lumencor® SPECTRA X light engine). Light engine was setup with filter paddles (395/25, 438/29, 470/24, 555/28, 635/22, 730/40). Images were obtained with a sCMOS camera (2048 × 2048, 16-bit, ORCA-Flash4.0 LT Plus, Hamamatsu), automatic multi-slide stage (PILine, M-686K011), and Zeiss Plan-Apochromat objectives 20x (0.8 NA, air, 420650-9901), 40× (1.4 NA, oil, 420762-9900). Filter cubes for wavelength separation included quad band Chroma 89402 (DAPI, Cy3, Cy5), quad band Chroma 89403 (Atto425, TexasRed, AlexaFluor750), and single band Zeiss 38HE (AlexaFluor488). In SBL-based ISS, one of the base libraries was either conjugated to TexasRed or Atto425 fluorophores that requires substitution of the 555/28 filter paddle with a compatible 575 one to image TexasRed. Regions were outlined and saved in order to perform repetitive cycle imaging.

### Stripping

After imaging, sections were prepared for sequential cycles. Readout detection probes and bridge-probes were stripped off RCPs. After five washes with 2× SSC, sections were incubated in stripping solution (65% formamide, 2XSSC) for 30 min at 30°C. This was followed by 5 washes with 2× SSC. Now the next cycle of bridge-probes could be hybridized as previously.

### Image analysis

Imaging data was analyzed with in-house custom software that handles image processing and gene calling. All code is written in MATLAB and is freely available at <https://github.com/Moldia>. Many steps follow previous publications (20,21,25,26).

Multispectral imaging for channels occurs for multiple cycles. Each image consists of multiple tiles that contain a

10% overlap to allow for stitching alignment between fields of view. A z-stack focal depth of 10 μm with a spacing of 0.5 μm deep allows for coverage of entire tissue. After imaging, Zeiss ZEN software was used to maximum intensity project each field of view to obtain a flattened two-dimensional image which were then stitched together. After exporting images in .tif format, images were aligned between cycles and then were split into multiple smaller images, henceforth referred to as tiles. The tiled images were subsequently top-hat filtered, the RCPs were segmented and the intensity of each RCP was measured, for each channel. If an anchor stain was not used, a pseudo-anchor was created by making a composition of the four readout detection probe channels into one merged image for each cycle. The above steps, after the tiling, were done in CellProfiler 2.2.0. The intensity measurements were saved in a comma separated values and subsequently used to decode the RCPs in MATLAB (repository: <https://github.com/Moldia/Tools>). These intensity measurements are used to calculate a quality score, defined as follows: the maximum intensity in the set of channels is divided by the sum of the all of the channels. This means that we can have a quality score of 0.25–1. Similar intensities in all channels would result in a quality score on the lower end of this spectrum and at the upper end, the opposite, i.e. more dissimilar intensities between the maximum intensity channel and the other channels.

### SBH versus SBL analysis

Each method was performed in parallel on two sequential mouse brain coronal sections using four PLPs per target gene *Actb*, *Pgk1*, *Gapdh* and *Polr2a*, allowing for the possibility to decode in one cycle. The protocol for the steps to decode SBL chemistry was as previously published (21) and at protocols.io: [dx.doi.org/10.17504/protocols.io.bb2gqibw](https://doi.org/10.17504/protocols.io.bb2gqibw). Three regions of interest (ROI) were selected from entire imaged sections to cover isocortex (ROI1), hippocampal formation (ROI2) and thalamus (ROI3) (Supplementary Figure S2B). The size of each ROI was 6000 × 6000 pixels for Exp\_1/Exp\_2, and 3900 × 3900 pixels for Exp\_3/Exp\_4. The Exp\_# on the same tissue refer to different imaging cycles. These ROIs were then exported as .tiff files. *Intensity measurements over RCPs*: For each channel image, the RCPs intensity was measured over the RCP, 10 pixels in each direction from the middle, which generated a total of 21 intensity measurements. *Average intensity over the RCPs*: The average intensity and the standard deviation over the RCPs were calculated. This was then visualized using line plots. *Maximum intensity*: The top 300 intensities were then extracted from the middle of each RCP that was located, generating 300 intensities values from each channel. The maximum intensities were then visualized with violin plots. *Signal-to-noise ratio (SNR)*: For each of the RCPs found, the noise was defined as the mean intensity of the four outer positions, i.e. pixel number -10, -9, 9 and 10. The intensity value for each position was then divided by the noise. The mean and the standard deviation for these values were then calculated and plotted. *Control experiment*: In the control experiment (Supplementary Figure S2G), an anchor probe was hybridized to both sections and the



average intensity, SNR and the SNR in the center of RCP was compared for round 1 and round 5, with SBL and SBH chemistries being performed between rounds. *Spot finding at different signal intensity thresholds*: starfish, a Python library for processing images, was used to resolve spots at different signal intensity thresholds. Images ( $2000 \times 2000$  pixels) from SBH and SBL experiments, were processed using a custom starfish pipeline. Spots were separated using BlobDetector, which is a multi-dimensional gaussian spot detector that can be tailored to detect spots of different sizes and intensities. To assess the number of spots found in different signal intensity thresholds, the BlobDetector was tuned to different thresholds. The number of spots for each threshold was recorded for SBH and SBL and subsequently plotted as a barplot.

### Crosstalk measurements

Crosstalk between channels was assessed by cytofluorograms. For each pixel position, the intensity of one channel was plotted on the x-axis and the intensity of another channel on the y-axis. This means that the pure signals will populate the paraxial regions and the brightest spots populate the ends of the axes. Mixed pixels are in the diagonal area of the plot.

## RESULTS

### Amplification of padlock probes and sequence-by-hybridization for the detection of RNA transcripts

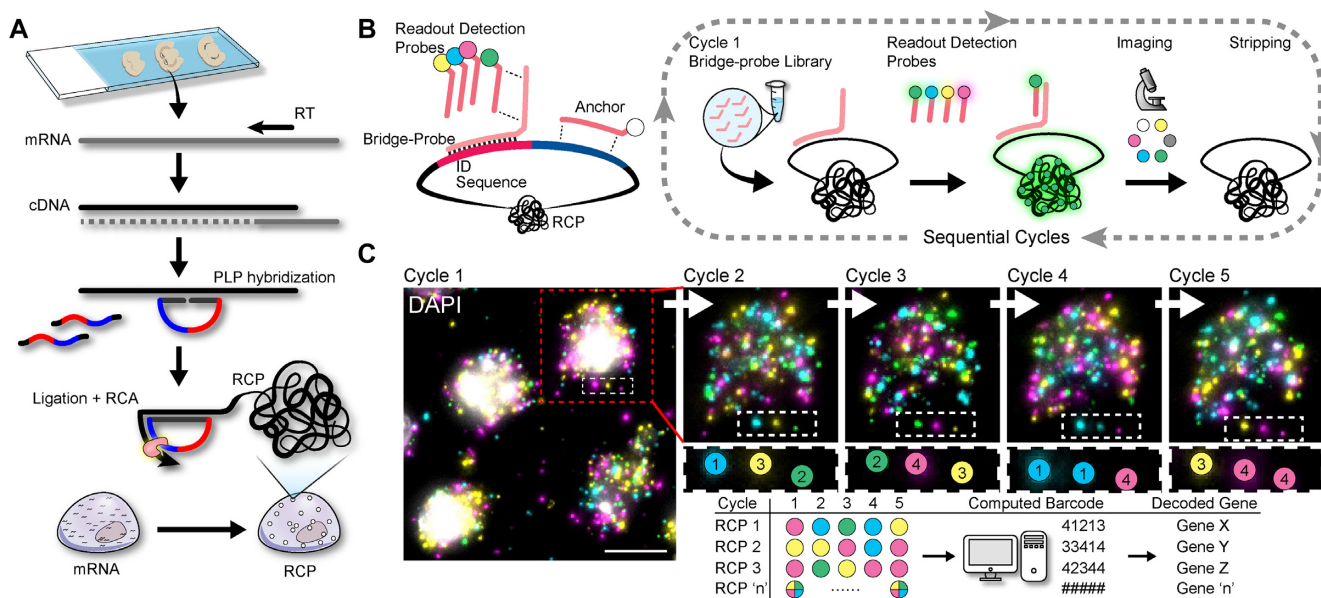
Similar to previous established ISS methods, PLPs are designed to target specific cDNA sequences from reversed transcribed mRNA *in situ* (20) (Figure 1A). An in-house padlock design pipeline produces multiple gene unique target sequences based on user input parameters (see Materials and Methods). The customizable backbone of the PLPs contains two parts: a unique ID sequence that is paired to each gene of interest targeted and a general 'anchor' sequence shared by all PLPs. Here, we design PLPs to target a 30 nt gene unique cDNA sequence split into two 15 nt arms, and a backbone with a 20 nt ID sequence and a 20 nt anchor sequence, resulting in a final 70 nt long PLP (Supplementary Figure S1A). The 20 nt ID sequence is a pre-determined sequence unique for the transcript target (e.g. five PLPs targeting five regions of one gene transcript have the same ID sequence). The backbone and ligated target arms of the PLPs are amplified through RCA and results in a rolling circle product (RCP), a sub-micron amplicon detectable with probes and a conventional epifluorescence microscope. The RCP amplicon contains multiple, repetitive ID sequence targets, allowing for the specificity of complimentary hybridization of readout oligonucleotides termed bridge-probes. Bridge-probes are 39 nt long, 17 base pair (bp) binding complementary to the amplified ID sequence, a 2 nt non-complimentary linker, and 20 nt that consists of one of four sequences binding complementary to a 20 bp readout detection probe conjugated to one of four assigned fluorophores (Figure 1B, Supplementary Figure S1B). Each ID sequence is associated to a combinatorial coding scheme (barcode, e.g. 43221) that is decoded sequentially with a cy-

cle dependent bridge-probe library (Supplementary Figure S1C).

Ligation of juxtaposed 15 nt PLP target-end arms allows for high specificity and RCA of the targeted transcripts in position. Sequential rounds of bridge-probe hybridization, readout detection probing, imaging and stripping allows for a highly multiplexed assay without the round limitation of previous SBL-based ISS methods (Figure 1B). Currently, as presented here, HybISS is set up with four readout detection probes per cycle. This allows for a target panel of genes ( $G$ ) that can be measured in a combinatorial manner with four fluorophores ( $F$ ) per cycle ( $c$ ), and thus  $G=F^c$ , without any cycle limitation and currently showing excellent tissue integrity and RCP maintenance over ten cycles (Supplementary Figure S1D, E), roughly calculated to be a loss of 0.28% RCPs per round. Additionally, unlike SBL-based ISS method which is limited to four fluorophores (one per barcode base: A, T, C or G), the number of fluorophores possible in HybISS is limited only by a microscope's ability to distinguish fluorophores, allowing for increased combinatorial capabilities with proper dye and filter selection. Furthermore, subset bridge-probe panel selection allows for the flexibility to do successive pools of panels on the same tissue sample in the case of overcrowding probes and does not compromise experimental design for probe selection as individual gene readouts can be excluded by simply removing corresponding bridge-probe and done in later cycles (Supplementary Figure S1F). This has not been possible in previous ISS methods in that every PLP would be detected every round and problematic genes such as high expressers would have to be run separately. Decoding is performed through sequential combinatorial decoding scheme similar to previous ISS methods, where each barcode has a hamming distance of at least two, allowing for correction and probabilistic assignment of barcodes with better accuracy (Figure 1C) (20,21). Moreover, with bridge-probes, this now also allows the possibility for further error correction by including 'zeroes' in the decoding scheme where bridge-probes are simply excluded for a subset of genes in different rounds of probing. All these new features integrated into HybISS increases its flexibility and adaptability to suit the design of experiments.

### HybISS compared to SBL-based ISS using reference gene panel for benchmarking

In order to demonstrate some of the improvements of HybISS technology, we compare it to SBL-based ISS by designing sets of PLPs to bind identical target sequences of a panel of mouse brain reference genes (*Actb/Gapdh/Pgk1/Polr2a*) and PLP backbones to support the different SBL and SBH sequencing chemistries, resulting in the same length PLPs (Supplementary Figure S2A, Supplementary Table S1). SBL and SBH experiments were run in parallel on sequential mouse brain coronal sections showing even distribution of reference genes across tissue sections (Figure 2A, B), although SBL shows an increase in overall background intensities (Figure 2C). To further assess intensity differences, we measured intensity of RCPs in three comparable ROIs (Supplementary Figure S2B). First, maximum signal intensities were measured for the top 300



**Figure 1.** HybISS method overview, hybridization-based *in situ* sequencing. (A) Overview of ISS, first reverse transcribing mRNA transcripts to cDNA. Gene specific PLPs target cDNA with juxtaposed ends next to each other that allow for ligation of PLP. Only transcripts that are ligated are enzymatically amplified by RCA. (B) Schematic overview of HybISS. Every cycle consists of hybridizing bridge-probes to RCPs and reading them out with fluorophore conjugated readout detection probes. For sequential cycles, bridge-probes are then stripped off to allow for rehybridizing next round of bridge-probes. (C) Example images of 5 cycles of a single cell. Sequential cycles with different bridge-probe libraries allows for the decoding of target transcripts within a cell. Scale bar: 10  $\mu$ m.

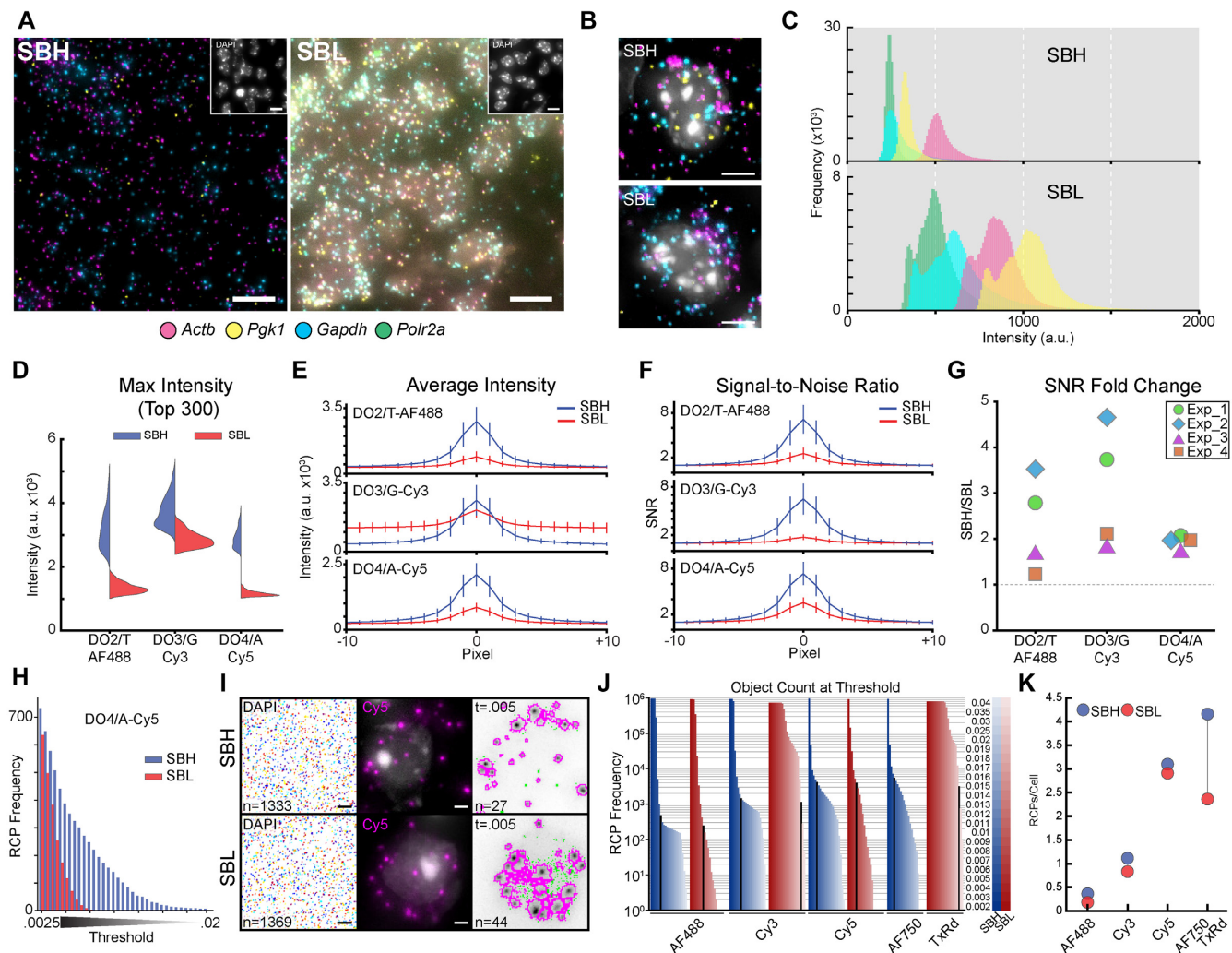
RCPs per channel, showing a significant increase for SBH-based HybISS across comparative channels (Figure 2D). Additionally, intensity was measured across a 21-pixel line centered at random RCPs within the ROIs to measure relative intensity to background. Not only was there an increase in average intensity values, background noise was also reduced in some channels of HybISS and importantly resulting in an increased SNR in all channels measured (Figure 2E, F). This indicates a more robust detection method, distinguishing RCPs with more ease due to better SNR. Comparable results were found in other tissues and cycles (Supplementary Figure S2C–E), resulting in an increase of SNR across several experiments (Figure 2G). This was controlled for by hybridizing only anchors to the respective RCPs and, as expected, no differences seen in average RCP intensity values, suggesting no effect on PLP design or in initial steps from hybridizing PLPs to RCA between the two iterations of ISS (Supplementary Figure S2F, G). However, after several rounds, we begin to see increases in background noise in SBL chemistry leading to higher SNR in SBH-based chemistry, most likely due to the harsher stripping conditions and additional enzymatic steps in SBL-based chemistry.

RCP localization and determination involves multiple steps of image processing but is highly dependent on pixel intensities being over a set threshold level. If a decoding strategy, such as *starfish* (27) is applied to an example SBH or SBL image, it is clear that at every threshold, SBH resolves more spots compared to SBL (Figure 2H). With a simple thresholding strategy to localize objects using CellProfiler, it becomes clear that SBH can distinguish objects more consistently over a range of thresholds where SBL chemistry produces much greater variation (Figure 2I, J).

After manual selection of optimal threshold for each condition and channel, we see an increase in detection efficiency for RCPs per cell (Figure 2K). Overall, HybISS benchmarks higher than SBL-based ISS in all parameters tested that confirms the significant upgrade in technology resulting in more robust and reliable signals.

### HybISS performed on whole mouse brain coronal section for purpose of gene transcript detection

To demonstrate the potential and validate HybISS, we explored multiplexing in mouse brain sections. The benchmarking gene panel used was curated by the SpaceTx Consortium (part of the Chan Zuckerberg Initiative and Human Cell Atlas project) (10,27), out of which 119 genes were selected for HybISS PLP design (Supplementary Table S1). This probe panel was designed to represent expression of genes to distinguish various cell types within the primary visual cortex (5) but was applied here to a whole coronal adult mouse brain section ( $\sim 60$  mm<sup>2</sup>, 10  $\mu$ m thick). Being able to infer data over large areas of the central nervous system, or any tissue, is desired but current spatial methods are limited in their throughput possibilities, imaging only small regions of interest with data storage and time becoming limiting factors. Here, we were able to explore tissue architecture in the entire coronal section in an efficient manner (each cycle, 40 $\times$  imaging, <5 h,  $\sim 700$ GB data), requiring minimal additional input for increased data output. We show the robustness to detect hundreds of target genes and map them in a whole section (Figure 3A, B). HybISS was able to resolve the list of genes after combinatorial decoding from five cycles (Figure 3B, C), with many genes showing discrete pat-



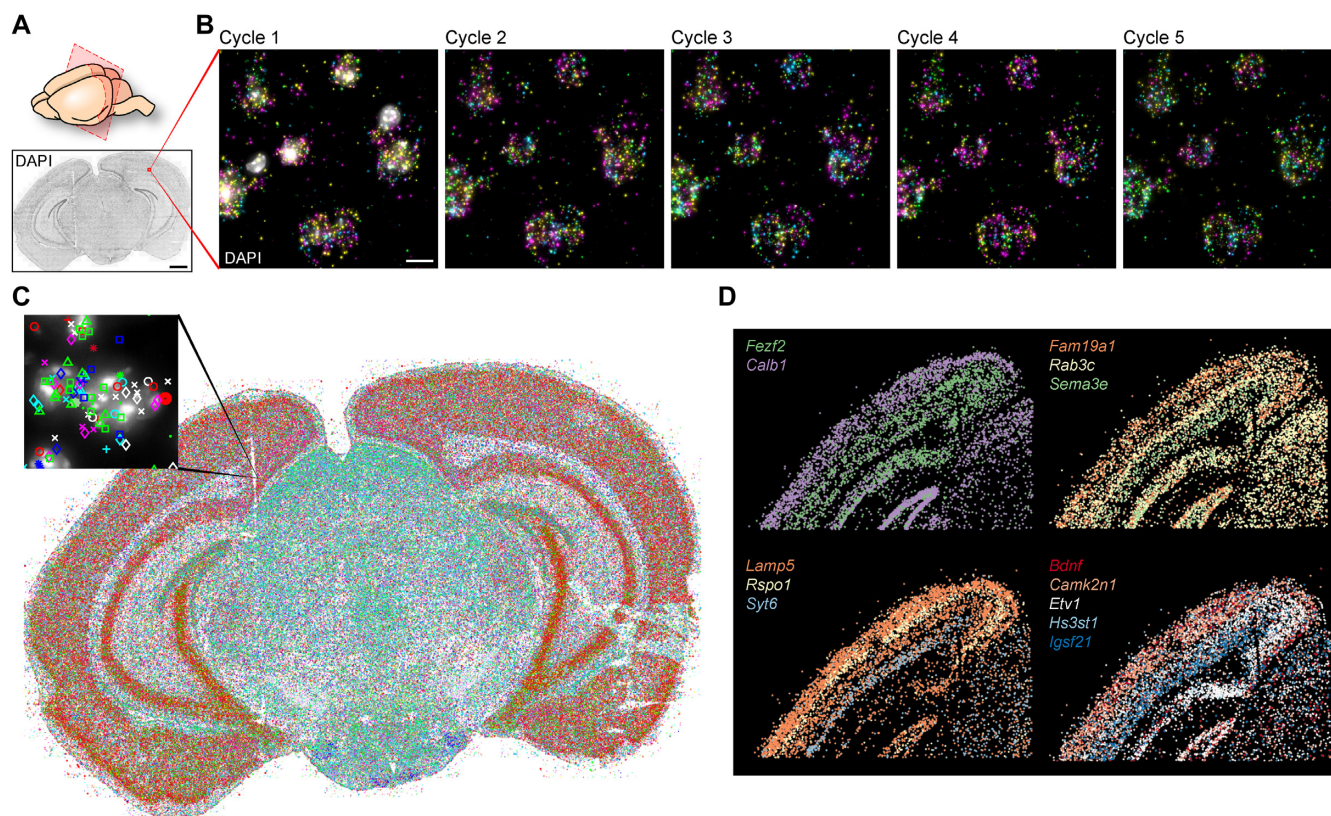
**Figure 2.** Comparison of HyBISS vs SBL-based ISS. (A) Representative images on the distribution of reference genes in sequential mouse coronal sections targeted by SBH- or SBL-based chemistries. Each channel has been scaled to the same intensity for comparison. Inset displays nuclear DAPI. Scale bar: 10  $\mu\text{m}$ . (B) Magnification of a single nuclei marked by DAPI with RCPs, intensities have been rescaled for clarity. Scale bar: 5  $\mu\text{m}$  (C) Binned pixel intensity from images in panel (A) for each channel. Green = Cy3, blue = Cy5, red = Atto425, yellow = AF488. (D) Max intensity of top 300 RCPs from three ROIs in each channel measured, comparing SBL- and SBH-based chemistries. (E) Average intensities of measured RCPs in ROIs. Pixel intensity measured across a 21-pixel line bisecting RCPs (blue = SBH, red = SBL). (F) Calculated SNR from intensity values in panel (E), using outer 2-pixels at each end as the measurement for noise. (blue = SBH, red = SBL). (G) Ratio of SBH/SBL SNR values in several replicates across the three channels measured. (H) RCP detection with various thresholds in a  $2000 \times 2000$  px ROI in the Cy5 channel for SBH and SBL using *starfish* (35) Blob Detector, indicating more RCPs found in SBH at each threshold. (I) Left: ROI of  $5000 \times 5000$  px from SBH and SBL experiments with segmented nuclei. Middle: Example raw image of a single cell with Cy5 channel. Right: Example threshold detection (0.005) from CellProfiler and number of objects (magenta outline) detected. Scale bar: left 100  $\mu\text{m}$ , middle 2  $\mu\text{m}$  (J) CellProfiler Object counts for a range of thresholds in all channels from  $5000 \times 5000$  px images in SBH and SBL chemistry. Black bar indicates optimal manual threshold and used for calculation in panel (K) (blue = SBH, red = SBL). (K) RCP count per cell from analyzed image used in (J). SBH DAPI count = 1333, SBL DAPI count = 1369.

terns, including laminar structures of cortical tissue (Figure 3D). Various genes have very distinct spatial distribution throughout the mouse brain and can be used to validate the method by comparing to other reference atlases such as the Allen Mouse Brain Atlas (2019) (28) (Supplementary Figure S3A). Here, we confirm the expression distribution of individual genes not only in the visual cortex but also regions like the hippocampus and thalamus. Although mouse tissue provides a valuable data resource, to understand relevant biological processes in humans, the study of human samples is quintessential.

### Multiplexing with HyBISS on human brain tissue

Data obtained from human samples would be more relevant and insightful but is limited due to restricted availability of good quality samples and difficulty in adapting and scaling experimental methods. Human tissue regions are much larger than model organisms and thus investigating comparable regions to mice in humans today still requires significant progress in multiple fields. Well preserved human samples to be used for various single-cell methods are rare and are known to be difficult to work with compared to mouse counterparts where fresh tissue and proven proto-





**Figure 3.** HybISS on mouse coronal brain section. (A) Whole mouse coronal section used for HybISS,  $\sim 60 \text{ mm}^2$ . Scale bar: 1 mm. (B) Representative images of HybISS using PLPs to map 119 genes over 5 cycles in section. Only first image includes counterstain for nuclei with DAPI. Scale bar:  $10 \mu\text{m}$ . (C) MATLAB output of the decoding of 119 genes across entire section, each color/symbol marking a single transcript detected. Inset shows zoomed-in representative image of gene marker plot mapped on DAPI image. (D) Selection of a subset of genes shown in (C) that have a distinct spatial laminar distribution within the neocortex.

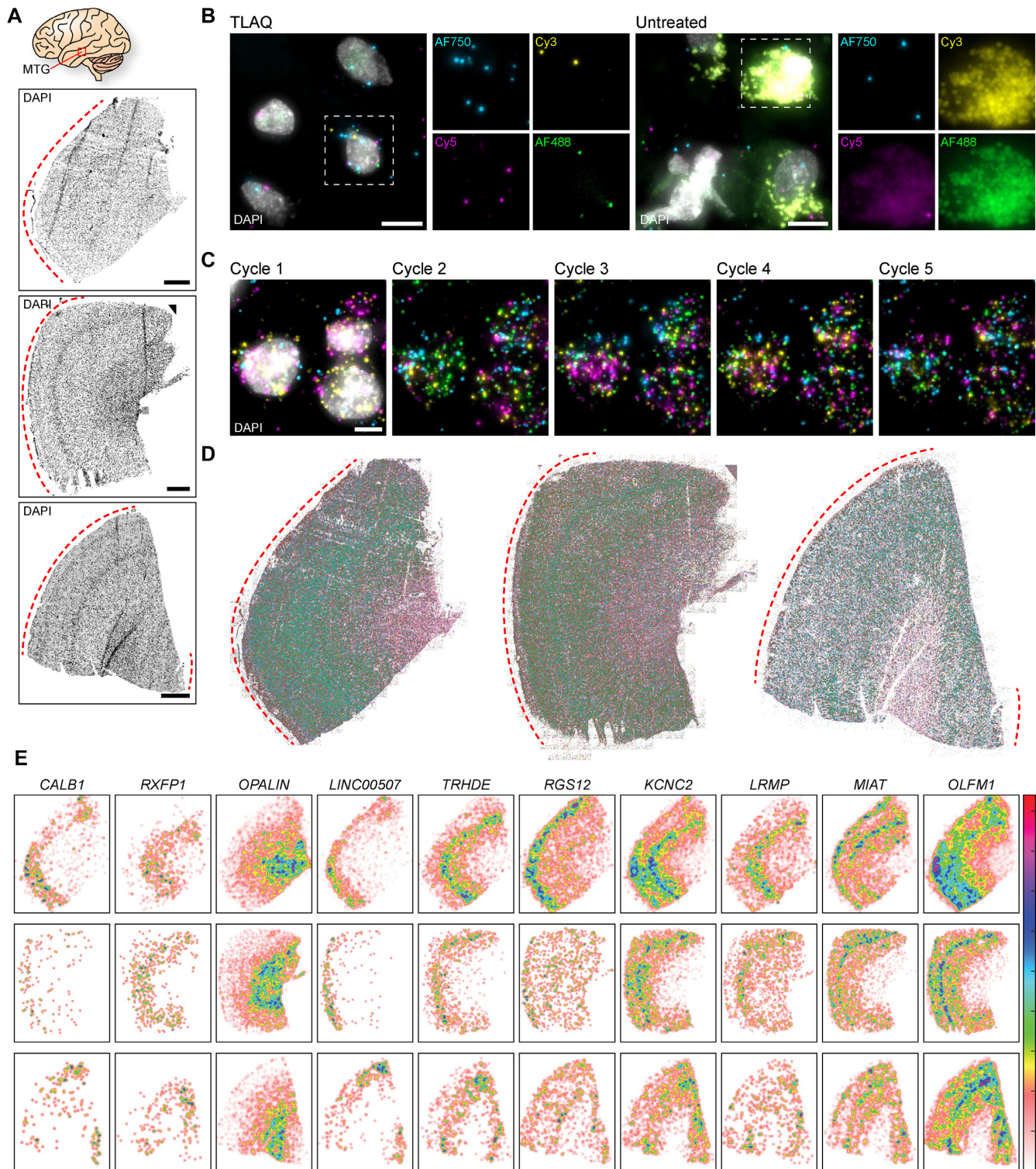
cols are more abundant and therefore being able to extract all information possible from a single tissue would provide a wealth of information, and is critical when studying damaged tissue such as the case for neurodegenerative disorders. Furthermore, determining exact anatomical positioning in human samples is more difficult due to their increased area and lack of good reference controls as compared to mouse samples.

Here, HybISS was performed on three sections of human brain tissue from the middle temporal gyrus (24) (whole sections,  $\sim 29 \text{ mm}^2$ ,  $45 \text{ mm}^2$ ,  $25 \text{ mm}^2$ ,  $10 \mu\text{m}$  thick) (Figure 4A). One inherent problem with human brain samples is lipid-containing residues of lipofuscin that causes strong autofluorescence in multiple imaging channels and is found in many tissues of aging humans. Many spatial methods and imaging techniques are sensitive to this autofluorescence and overcoming this with any quenching methods does not completely solve the problem and computational clearing strategies are not perfect and require variable user input. Here, we implement HybISS together with simple autofluorescence quenching (TrueBlack Lipofuscin Autofluorescence Quencher, TLAQ) that overcomes this problem in human brain sections and does not require any additional advanced microscopic tools or computational clearing methods (Figure 4B). Testing with a human reference gene panel (*ACTB/CYCI/ACTG1/NDUFB4*), in al-

most all imaged channels without treatment where lipofuscin was present, RCPs could not be distinguished from the background of lipofuscin (Supplementary Figure S4A, B). This was further corroborated with investigating crosstalk between channels with and without TLAQ treatment, indicating the effect of lipofuscin on background noise and true signal determination (Supplementary Figure S4C, D). TLAQ allows for clear separation of signals in imaged channels from the background and it only adds five minutes to the protocol and a one-time treatment maintains well over imaging cycles.

With only this additional treatment, we applied HybISS and were able to resolve a similar SpaceTx Consortium curated gene list from human middle temporal gyrus single-nucleus RNA-sequencing data from which a subset of 120 were targeted and decoded (Figure 4C, D). We are able to show comparable results to mouse sections without any additional difficulties, indicating a robust method to be used for spatial gene-expression profiling investigations. We further examined the spatial distribution of various gene transcripts with kernel density estimation plots across the three sections, many showing distinct patterns, including laminar distribution of cortical tissue (Figure 4E). Collectively, this data shows a proof of concept that it is possible to do multiplexed *in situ* hybridization studies in human tissue in a high throughput and robust manner.





**Figure 4.** HyBISS on human middle temporal gyrus brain sections. **(A)** DAPI nuclei stain of human tissue sections from middle temporal gyrus. Dashed red line demarcates the outer pial surface of tissue section. Area approx. 29 mm<sup>2</sup>, 45 mm<sup>2</sup>, 25 mm<sup>2</sup> top to bottom panel. Scale bar: 1 mm. **(B)** Representative images of the effects of lipofuscin in human brain tissue that can be treated with TLAQ, and HyBISS amplification overcomes any residual background noise. Scale bar: 10 μm. **(C)** Magnified field of view from section in panel (a) of several cells across 5 cycles of HyBISS. First cycle includes DAPI to show nuclei location. Scale bar: 5 μm. **(D)** Spatial distribution of decoded HyBISS transcripts of 120 gene panel across the three tissue sections. Left, 1 649 212 spots; middle, 1 936 227 spots; right, 602 681 spots. **(E)** Kernel density estimation plots for a subset of individual gene transcripts that show distinct spatial distribution, including laminar anatomy of cortical tissue.



## DISCUSSION

The HybISS method is a further advancement of the ISS technology in that it strives to meet the requirements and demands in the spatial transcriptomics field without compromising the throughput and area limitation factors that many other spatial techniques are running into. We show the ease of application to human brain tissue with a large gene panel selection to be able to distinguish gene transcripts across entire tissue sections, including whole mouse brain coronal sections. The simplicity and flexibility of HybISS make it a versatile image-based spatial transcriptomic method that can be easily implemented and adapted for a wide variety of scientific questions. Customizing gene lists will be vital as it is a targeted approach and curating custom gene panels from scRNA-seq data, such as probabilistic cell typing by *in situ* sequencing (pciSeq) (21), to create cell-type atlases will provide a useful tool in the HybISS method.

Due to the amplification method of RCA, an epifluorescence microscope using widefield objectives (20×/40×) can be used for rapid scanning over large regions, essentially limited to what can fit on a slide. Amplification-based strategies also aid in overcoming background noise found in many tissue sources. HybISS shows an increase in RCP intensity compared to previous SBL-based methods that allows for robust spot calling, decoding and processing of data. It has been stated that efficiency of SBL chemistry is 20–30% (29) and is probably due to multiple step requirement of anchor binding and competitive correct base library hybridization with degenerate sequences and ligation, all having various efficiencies, while SBH chemistry is a non-competitive direct hybridization step requiring no enzymes or degenerate sequences. Additionally, base libraries are 9 nt with degenerate bases that could make them sticky, increasing the general background noise during detection. Moreover, the harsher stripping conditions of SBL could be affecting tissue integrity as well. Although efficiency of PLP ligation and RCA should not be different, we observe effects to efficiently detect RCPs more robustly as a clear advantage. Cleaner and more reliable data aids in downstream analysis such as creating gene networks (30), cell typing (21) and molecular maps (31).

The full potential and implementation of HybISS has not yet been reached. Fine tuning of PLP and target sequence design can increase efficiency, henceforth improve confidence in analysis such as cell typing, although not necessary as even data with lower detection efficiency can yield comprehensive cell maps (21). Larger gene panels (several fold increase) will be feasible with little to no adjustment to the method presented here. Additionally, as with previous ISS methods, HybISS and PLPs maintain the ability to detect single nucleotide variations due to their high specificity and sensitivity to be able to distinguish known point mutations (32) or editing sites (19) in tissues for example. SBL-based ISS has already shown its power in other non-neuronal tissue, in the form of cancer diagnostics (20,26) and tuberculosis granuloma (25) among others. Therefore, we see no reason for HybISS to improve multiplexed transcript detection in other fields and applications.

We see the great potential of implementing HybISS in spatial cell atlas projects. SBL-based ISS has been demon-

strated to assign cell types in mouse hippocampus (21) and neural crest development (18). ISS has also been applied and analyzed together with scRNA-seq and data from untargeted transcriptome-wide Spatial Transcriptomics (33) to develop a spatiotemporal atlas of the developing human heart (34). As demonstrated here on human tissue, the capacity is in place to start creating a comprehensive spatial reference maps for the Human Cell Atlas (10) specifically in brain tissue.

Moreover, we provide an image-based spatial transcriptomic data set that can be used to develop computational spatial methods for such analysis as cell segmentation and cell type assignment, as implemented in *starfish* for example (27,35). We find the unique insight into the human cortical tissue to be of interest to all and dissemination of such information important. Applying further advancements that have been incorporated, at least in part, to other spatial techniques to HybISS are currently being pursued to try to push the capabilities of the technology. This could be in the form of microfluidic devices to automate production (36), reducing time, or investigating how these cells are actually arranged in a three-dimensional space and not simply sequential sections.

Implementing tools to further our knowledge of the human body will give us a better understanding of it when in the diseased state. Just like the sequencing of the human genome changed the perspective of human life, a global effort has put us in an era where we are about to define all the cell types within the body. How these cells are organized, connected, and function will give insight to what happens when things go wrong. Here, we present HybISS, the next generation of ISS as an *in situ* image-based spatial technology that is reliable, robust, and scalable to meet the emerging demands of the single-cell field.

## DATA AVAILABILITY

All code is available online at <https://github.com/Moldia>. The raw figures (several terabytes) are available from the corresponding author upon reasonable request. Analysis files for sections in Figure 3 and 4 are available at Figshare (doi: 10.17045/sthlmuni.12821840) including gene spot coordinates and corresponding DAPI images.

## SUPPLEMENTARY DATA

[Supplementary Data](#) are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank all members of the SpaceTx Consortium for their work contribution that made this possible. We thank Mats Nilsson lab members for their insight and comments.

*Author contributions:* D.G. designed and performed the experiments, analyzed data and wrote the manuscript. C.M.L. performed experiments, analyzed data. E.C. performed experiments. X.Q., S.M.S. and M.M.H. analyzed data. E.S.L. and SpaceTx Consortium provided tissue and gene list curation. M.N. conceived and supervised the study. All authors read the manuscript and provided feedback.

## FUNDING

Chan Zuckerberg Initiative, an advised fund of Silicon Valley Community Foundation for funding of the SpaceTx consortium and the Human Cell Atlas; Human Developmental Cell Atlas grants from the Erling-Persson Family Foundation and the Knut and Alice Wallenberg Foundation; Swedish Research Council [2019-01238]; EASI-Genomics a European Union's Horizon 2020 research and innovation [824110]; Swedish Brain Foundation (Hjärnfonden) [PS2018-0012]. Funding for open access charge: Swedish Research Council [2019-01238].

*Conflict of interest statement.* X.Q., M.M.H. and M.N. hold shares in the company CartaNA AB that commercializes *in situ* sequencing reagents.

## REFERENCES

- Poulin, J.F., Tasic, B., Hjerling-Lefler, J., Trimarchi, J.M. and Awatramani, R. (2016) Disentangling neural cell diversity using single-cell transcriptomics. *Nat. Neurosci.*, **19**, 1131–1141.
- Zeng, H. and Sanes, J.R. (2017) Neuronal cell-type classification: challenges, opportunities and the path forward. *Nat. Rev. Neurosci.*, **18**, 530–546.
- Ding, J., Adiconis, X., Simmons, S.K., Kowalczyk, M.S., Hession, C.C., Marjanovic, N.D., Hughes, T.K., Wadsworth, M.H., Burks, T., Nguyen, L.T. *et al.* (2020) Systematic comparison of single-cell and single-nucleus RNA-sequencing methods. *Nat. Biotechnol.*, **38**, 737–746.
- Zeisel, A., Munoz-Manchado, A.B., Codeluppi, S., Lonnerberg, P., La Manno, G., Jureus, A., Marques, S., Munguba, H., He, L., Betscholtz, C. *et al.* (2015) Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. *Science*, **347**, 1138–1142.
- Tasic, B., Yao, Z., Grayback, L.T., Smith, K.A., Nguyen, T.N., Bertagnolli, D., Goldy, J., Garren, E., Economo, M.N., Viswanathan, S. *et al.* (2018) Shared and distinct transcriptomic cell types across neocortical areas. *Nature*, **563**, 72–78.
- Zeisel, A., Hochgerner, H., Lönnerberg, P., Johnson, A., Memic, F., van der Zwan, J., Häring, M., Braun, E., Borm, L.E., La Manno, G. *et al.* (2018) Molecular architecture of the mouse nervous system. *Cell*, **174**, 999–1014.
- La Manno, G., Soldatov, R., Zeisel, A., Braun, E., Hochgerner, H., Petukhov, V., Lidschreiber, K., Kastrioti, M.E., Lönnerberg, P., Furlan, A. *et al.* (2018) RNA velocity of single cells. *Nature*, **560**, 494–498.
- Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D.M., Hill, A.J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F.J. *et al.* (2019) The single-cell transcriptional landscape of mammalian organogenesis. *Nature*, **566**, 496–502.
- Xia, C., Fan, J., Emanuel, G., Hao, J. and Zhuang, X. (2019) Spatial transcriptome profiling by MERFISH reveals subcellular RNA compartmentalization and cell cycle-dependent gene expression. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 19490–19499.
- Regev, A., Teichmann, S.A., Lander, E.S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M. *et al.* (2017) The human cell atlas. *Elife*, **6**, e27041.
- Mereu, E., Lafzi, A., Moutinho, C., Ziegenhain, C., MacCarthy, D.J., Alvarez-Varela, A., Battle, E., Sagar, Grün, D., Lau, J.K. *et al.* (2020) Benchmarking single-cell RNA-sequencing protocols for cell atlas projects. *Nat. Biotechnol.*, **38**, 747–755.
- Lein, E., Borm, L.E. and Linnarsson, S. (2017) The promise of spatial transcriptomics for neuroscience in the era of molecular cell typing. *Science*, **358**, 64–69.
- Strell, C., Hilscher, M.M., Laxman, N., Svedlund, J., Wu, C., Yokota, C. and Nilsson, M. (2019) Placing RNA in context and space - methods for spatially resolved transcriptomics. *FEBS J.*, **286**, 1468–1481.
- Asp, M., Bergenstråhle, J. and Lundeberg, J. (2020) Spatially resolved transcriptomes-next generation tools for tissue exploration. *Bioessays*, **1900221**, e1900221.
- Eng, C.-H.L., Lawson, M., Zhu, Q., Dries, R., Koulena, N., Takei, Y., Yun, J., Cronin, C., Karp, C., Yuan, G.-C. *et al.* (2019) Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nature*, **568**, 235–239.
- Moffitt, J.R., Bambach-Mukku, D., Eichhorn, S.W., Vaughn, E., Shekhar, K., Perez, J.D., Rubinstein, N.D., Hao, J., Regev, A., Dulac, C. *et al.* (2018) Molecular, spatial, and functional single-cell profiling of the hypothalamic preoptic region. *Science*, **362**, eaau5324.
- Codeluppi, S., Borm, L.E., Zeisel, A., La Manno, G., van Lunteren, J.A., Svensson, C.I. and Linnarsson, S. (2018) Spatial organization of the somatosensory cortex revealed by osmFISH. *Nat. Methods*, **15**, 932–935.
- Soldatov, R., Kaucaka, M., Kastrioti, M.E., Petersen, J., Chontorotzea, T., Englmaier, L., Akkuratova, N., Yang, Y., Häring, M., Dyachuk, V. *et al.* (2019) Spatiotemporal structure of cell fate decisions in murine neural crest. *Science*, **364**, eaas9536.
- Lundin, E., Wu, C., Widmark, A., Behm, M., Hjerling-Lefler, J., Daniel, C., Öhman, M. and Nilsson, M. (2020) Spatiotemporal mapping of RNA editing in the developing mouse brain using *in situ* sequencing reveals regional and cell-type-specific regulation. *BMC Biol.*, **18**, 6.
- Ke, R., Mignardi, M., Pacureanu, A., Svedlund, J., Botling, J., Wählby, C. and Nilsson, M. (2013) *In situ* sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods*, **10**, 857–860.
- Qian, X., Harris, K.D., Hauling, T., Nicoloutsopoulos, D., Muñoz-Manchado, A.B., Skene, N., Hjerling-Lefler, J. and Nilsson, M. (2020) Probabilistic cell typing enables fine mapping of closely related cell types *in situ*. *Nat. Methods*, **17**, 101–106.
- Ke, R., Mignardi, M., Hauling, T. and Nilsson, M. (2016) Fourth generation of next-generation sequencing technologies: promise and consequences. *Hum. Mutat.*, **37**, 1363–1367.
- Hawrylycz, M.J., Lein, E.S., Guillozet-Bongaarts, A.L., Shen, E.H., Ng, L., Miller, J.A., van de Lagemaat, L.N., Smith, K.A., Ebbert, A., Riley, Z.L. *et al.* (2012) An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, **489**, 391–399.
- Hodge, R.D., Bakken, T.E., Miller, J.A., Smith, K.A., Barkan, E.R., Grayback, L.T., Close, J.L., Long, B., Johansen, N., Penn, O. *et al.* (2019) Conserved cell types with divergent features in human versus mouse cortex. *Nature*, **573**, 61–68.
- Carow, B., Hauling, T., Qian, X., Kramnik, I., Nilsson, M. and Rottenberg, M.E. (2019) Spatial and temporal localization of immune transcripts defines hallmarks and diversity in the tuberculosis granuloma. *Nat. Commun.*, **10**, 1823.
- Svedlund, J., Strell, C., Qian, X., Zilkens, K.J.C., Tobin, N.P., Bergh, J., Sieuwerts, A.M. and Nilsson, M. (2019) Generation of *in situ* sequencing based OncoMaps to spatially resolve gene expression profiles of diagnostic and prognostic markers in breast cancer. *EBioMedicine*, **48**, 212–223.
- Perkel, J.M. (2019) Starfish enterprise: finding RNA patterns in single cells. *Nature*, **572**, 549–551.
- Lein, E.S., Hawrylycz, M.J., Ao, N., Ayres, M., Bensinger, A., Bernard, A., Boe, A.F., Boguski, M.S., Brockway, K.S., Byrnes, E.J. *et al.* (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature*, **445**, 168–176.
- Drmanac, R., Sparks, A.B., Callow, M.J., Halpern, A.L., Burns, N.L., Kerami, B.G., Carnevali, P., Nazarenko, I., Nilsen, G.B., Yeung, G. *et al.* (2010) Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays. *Science*, **327**, 78–81.
- Salamon, J., Qian, X., Nilsson, M. and Lynn, D.J. (2018) Network visualization and analysis of spatially aware gene expression data with InsituNet. *Cell Syst.*, **6**, 626–630.
- Partel, G., Hilscher, M.M., Milli, G., Solorzano, L., Klemm, A.H., Nilsson, M. and Wählby, C. (2019) Identification of spatial compartments in tissue from *in situ* sequencing data. bioRxiv doi: <https://doi.org/10.1101/765842>, 18 September 2019, preprint: not peer reviewed.
- Larsson, C., Grundberg, I., Söderberg, O. and Nilsson, M. (2010) *In situ* detection and genotyping of individual mRNA molecules. *Nat. Methods*, **7**, 395–397.
- Ståhl, P.L., Salmén, F., Vickovic, S., Lundmark, A., Navarro, J.F., Magnusson, J., Giacomello, S., Asp, M., Westholm, J.O., Huss, M. *et al.* (2016) Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*, **353**, 78–82.



34. Asp,M., Giacomello,S., Larsson,L., Wu,C., Fürth,D., Qian,X., Wärdell,E., Custodio,J., Reimegård,J., Salmén,F. *et al.* (2019) A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart. *Cell*, **179**, 1647–1660.
35. Axelrod,S., Carr,A.J., Freeman,J., Ganguli,D., Long,B., Tung,T. and Others (2018) Starfish: Open Source Image Based Transcriptomics and Proteomics Tools. <http://github.com/spacetx/starfish>.
36. Maiño,N., Hauling,T., Cappi,G., Madaboosi,N., Dupouy,D.G. and Nilsson,M. (2019) A microfluidic platform towards automated multiplexed in situ sequencing. *Sci. Rep.*, **9**, 3542.