HT-smFISH: a cost effective and flexible workflow for high-throughput single molecule RNA imaging

Adham Safieddine, Emeline Coleno, Frederic Lionneton, Abdel-Meneem Traboulsi, Soha Salloum, Charles-Henri Lecellier, Thierry Gostan, Virginie Georget, Cédric Hassen-Khodja, Arthur Imbert, et al.

HAL Id: hal-03834928

https://hal.sorbonne-universite.fr/hal-03834928

Submitted on 31 Oct 2022

HT-smFISH: a cost effective and flexible workflow for high-throughput single molecule RNA imaging

Adham Safieddine[1,2,*], Emeline Coleno[1,3], Frederic Lionneton[4], Abdel-Meneem Traboulsi[1], Soha Salloum[1,3] , Charles-Henri Lecellier[1,5], Thierry Gostan[1], Virginie Georget[4], Cédric Hassen-Khodja[4], Arthur Imbert[6,7], Florian Mueller[8], Thomas Walter[6,7], Marion Peter[1,9], Edouard Bertrand[1,3,10,*]

1: Institut de Génétique Moléculaire de Montpellier, University of Montpellier, CNRS, Montpellier, France

2: Sorbonne Université, CNRS, Institut de Biologie Paris-Seine, Laboratoire de Biologie du Développement, F-75005, Paris, France

3 : Institut de Génétique Humaine, University of Montpellier, CNRS, Montpellier, France

4 : BioCampus Montpellier, Montpellier, Cedex, France

5 : LIRMM, University of Montpellier, CNRS, Montpellier, France

6 : MINES ParisTech, PSL-Research University, CBIO-Centre for Computational Biology, Fontainebleau, France

7 : Institut Curie, Paris, Cedex, France

8: Imaging and Modeling Unit, Institut Pasteur, UMR 3691 CNRS, C3BI USR 3756 IP CNRS, Paris, France

9 : Institut de Recherche en Cancérologie de Montpellier U1194, University of Montpellier, INSERM, ICM, CNRS, Montpellier, France

10: Equipe Labélisée Ligue Nationale Contre le Cancer, University of Montpellier, CNRS, Montpellier, France

*: To whom correspondence should be addressed: adham.safieddine@sorbonne-universite.fr
and edouard.bertrand@igh.cnrs.fr

EDITORIAL SUMMARY: In HT-smFISH, multiple RNA probes are generated by parallel *in vitro* transcription from a large pool of unlabeled oligonucleotides. This reduces costs per targeted RNA compared to many smFISH methods, is easily scalable, and flexible in design.

PROPOSED TWEET: HT-smFISH is a cost effective and flexible method for high-throughput single molecule RNA probe synthesis and imaging. @AdhamSafieddine @IGMM_Montpel @IBPS_Paris

PROPOSED TEASER: High-throughput single molecule RNA imaging

Abstract (<250 words)

The ability to visualize RNA in its native subcellular environment with single molecule FISH (smFISH) has reshaped our understanding of gene expression and cellular functions. A major hindrance of smFISH is the difficulty to perform systematic experiments in medium or high throughput formats, principally because of the high cost of generating the individual fluorescent probe sets. Here we present High-Throughput smFISH (HT-smFISH), a simple and cost-efficient method for imaging hundreds to thousands of single endogenous RNA molecules in 96-well plates. HT-smFISH uses RNA probes transcribed *in vitro* from a large pool of unlabeled oligonucleotides. This allows generating individual probes for many RNAs species replacing commercial DNA probe sets. HT-smFISH thus reduces costs per targeted RNA compared to many smFISH methods, is easily scalable, and flexible in design. We provide a protocol that combines oligo pool design, probe set generation, optimized hybridization conditions, and guidelines for image acquisition. The pipeline requires knowledge of standard molecular biology tools, cell culture and fluorescent microscopy. It is achievable in ~20 days. In brief, HT-smFISH is tailored for medium to high-throughput screens that image RNAs at single molecule sensitivity.

Introduction

Proper compartmentalization of biological molecules is essential for all forms of life. The sub-cellular localization of RNA has received increasing attention owing to its contribution to various aspects of cellular function[1,2]. For example, mRNA localization can fine tune gene expression by mediating local translation allowing rapid on-site production of needed proteins in response to signaling events, stress, or cell cycle changes[3–6]. Local protein synthesis has been described in several models such as developing *Drosophila* embryos[7], neurons[8], yeast[9], various oocytes[10,11], and even commonly used human cell lines[6,12–14]. RNA localization also mediates translation-independent functions. This is not only true for long non-coding RNAs whose functions directly depend on their localization[15,16]; but also for mRNAs that serve a structural role[17], or contribute to phase separation mechanisms[18]. Much like proteins, it is becoming clear that the function of an RNA molecule is intimately linked to its localization.

Principles of single molecule FISH

Much of our understanding of RNA localization and trafficking stems from imaging-based methods. Indeed, such approaches often preserve cells and organelles, thus revealing transcripts in their endogenous molecular habitats. A breakthrough in fluorescent *in situ* hybridization (FISH) techniques took place around 20 years ago with the invention of single molecule FISH (smFISH)[19]. Here, multiple fluorescently labeled probes (oligonucleotides) are designed to hybridize with the same RNA target. Background comes from unbound single probes while individual transcripts are much brighter due to cumulative hybridizations of many probes on the same target. This breaks the single molecule barrier and allows resolving individual transcripts that appear as diffraction-limited spots using widefield illumination. The signal to noise ratio (SNR) is, in theory, proportional to the number of fluorophore-labeled probes bound to the RNA. In practice, 10-50 probes often allow adequate and specific detection with even more probes producing a better signal. This

creates an incentive to increase the number of probes targeting a transcript, which is often hindered by costly synthesis procedures.

Single molecule FISH variants

Initial smFISH methods uses oligonucleotides of 50-60 bases labeled with 4-5 fluorophores[19]. Challenging targets and demanding specimens have inspired numerous smFISH variants[20]. What many of these variants have in common is that they involve indirect labeling schemes that use unlabeled primary probes hybridizing with the RNA target itself, fused to overhangs called Readout sequences that in turn bind fluorescently labeled oligonucleotide detectors. Such designs provide two advantages: (i) they lower the cost of probe synthesis since the primary probe sets are themselves unlabeled and the used readout sequence is the same across different targets (as is the case in single molecule inexpensive FISH, smiFISH[21], and this protocol); (ii) they can be used to drastically increase the SNR, whereby Readout sequences serve as platforms to recruit amplifier probes than in turn bind many detectors through a tree-like architecture (as is the case in branched DNA FISH[22] and FISH-STICs[23] for example). Such approaches are particularly well adapted for optically demanding samples such as tissues and embryos with increased autofluorescence, or short targets such as specific exons or even microRNAs. However, they may also carry disadvantages including complex hybridization protocols, false positives caused by probe aggregation, lower detection rates due to a smaller number of primary probes, and the large size of probes rendering certain subcellular regions inaccessible.

Multiplexed smFISH strategies

Recent years have seen the rapid development of smFISH techniques aimed at simultaneously resolving hundreds to thousands of RNA species, adding the spatial dimension to transcriptomic studies (recently chosen as "method of the year"[24,25]). Such approaches often rely on the use of

multiple readout sequences on a single mRNA, creating a powerful system able to encode many species with a relatively small number of readouts (up to 2^n-1 encoded RNA species, with n being the number of readouts). These readout sequences are revealed by sequential rounds of precise detection, imaging, and bleaching (therefore requiring dedicated microfluidics to perform rounds of hybridization), as well as specific algorithms to decode and interpret the image. Implementations of multiplexed smFISH labeling include: (i) multiplexed error robust FISH (merFISH[26]); (ii) sequential FISH (seqFISH[27], seqFISH+[28], and other variants[29,30]); (iii) and split probe FISH (splitFISH[31]). These technologies can detect up to 10,000 transcripts in cells and cleared tissues. Sophisticated multiplexed smFISH approaches are readily competing with single cell sequencing methods in the gene expression field thanks to the single molecule sensitivity and detailed spatial information they provide[32–35].

In-situ sequencing strategies

Other approaches than can image many RNA species rely on _in-situ_ sequencing. Several such techniques exist such as fluorescent in situ sequencing (FISSEQ[36]), targeted in-situ sequencing[37], barcode _in-situ_ targeted sequencing (BaristaSeq[38]), and barcoded oligonucleotides ligated on RNA targets and amplified and sequenced _in-situ_ (BOLORAMIS[39]). These methods use barcoded padlock or padlock-like probes, which bind a cDNA reverse transcribed from the target RNA _in situ_, or directly the RNA itself. These probes are then circularized, amplified, and sequenced over several rounds to allow barcode and target identification. _In-situ_ sequencing has also been recently combined with expansion microscopy in Expansion Sequencing (Exseq[40]), extending its usability in tissues and offering higher spatial resolutions. Finally, spatially-resolved transcript amplicon readout mapping (STARmap[41]) improves on target specificity by using a modified, two-part padlock design. Amplification occurs only if both primers bind the same target RNA, drastically reducing non-specific amplification. Overall, amplification-based _in-situ_ sequencing

methods are adapted for imaging and multiplexing in tissue samples due to the amplified signal they provide.

Development of HT-smFISH and target users

SmFISH is a powerful tool for investigating RNA localization, gene expression, and cell-to-cell heterogeneity at the single molecule level[20]. To this day, scaling up this technique remains a challenge for both specialized and non-specialized labs mainly due to the incremental costs of probe sets and reagents. Traditional smFISH protocols use fluorescently labeled oligonucleotides, and we previously developed a popular low-throughput smFISH variant that used non-fluorescent primary probes (smiFISH[21]), thereby simplifying and lowering the cost of the probes. Here, we developed a convenient and cost-effective High-Throughput version of smiFISH, HT-smFISH.

HT-smFISH aims to image hundreds to thousands of RNAs at the single molecule level. It is simple and can by performed by many molecular biology labs in weeks to months. This is achievable through designing a single oligo pool containing a mixture of probe sets against the RNAs of interest present in low copies. From this mix, gene-specific probe sets are differentially amplified in 96-well plates with one well containing one probe set to be used for the *in-situ*. This in-house amplification of the oligo pool thus lowers probe synthesis costs per RNA imaged (provided that around 150-250 RNAs at least are targeted): ~15-20 $/€ per RNA compared to ~500 $/€ in conventional smFISH and ~150 $/€ in smiFISH[21]. HT-smFISH includes a simple probe design algorithm and a non-toxic, formamide-free hybridization buffer. The *in-situ* hybridization also works in 96-well plates, making the technique particularly suited for medium to large smFISH screens. In addition, data can be interpreted by direct visual examination without complicated image analysis algorithms, although quantitative measurements can be done with dedicated tools such as FISH-quant[42], locFISH[43], FISH-quant v2[44], RS-FISH[45], StarFISH[46], deepBlink[47], and

DypFISH[48]. HT-smFISH has flexible experimental design where individual probes can be easily re-used many times and labeled with different fluorophores. The protocol is also compatible with other co-labeling procedures such as immunofluorescence, and is applicable to various cell lines or tissue sections, making it broadly useful for many labs.

Overview of the Procedure

The first stage of HT-smFISH involves designing and amplifying an oligo pool against the target RNAs of interest. This oligo pool contains a mixture of all the probes present in low copies. Individual probe sets recognizing one RNA target are specifically amplified using two rounds of PCR and in vitro transcription (IVT) in 96-well plates. This produces primary RNA probe sets at drastically reduced costs compared to commercially purchased ones. The next stage encompasses two parts: (i) primary probe set labeling and the *in-situ* hybridization itself. During labeling, the unlabeled probe sets produced from the first stage are annealed to fluorescent oligos forming duplexes. These duplexes are used in the *in-situ* hybridization step where the presented workflow focuses on cells grown in 96-well plates with one well dedicated to image one RNA target. The provided hybridization conditions, however, are readily applicable in other experimental designs and formats, tissue sections, and are compatible with secondary labeling such as immunofluorescence. Lastly, the final stage of HT-smFISH involves image acquisition, data management, and RNA image analysis.

Limitations of HT-smFISH

In HT-smFISH, we recommend using 50-100 probes per RNA target to acquire reliable and consistent signals, with optimal sensitivity for hundreds of RNA targets. Considering that the average length of a hybridization sequence is around 28 nts, this naturally limits the usability of the method to RNAs of ~1400 nts to accommodate sufficient probes. Fewer probes can be used,

down to a minimum of 20 for shorter targets, and probes can also be overlapped and tiled to increase probe density. In such cases, signal to noise ratios and the percentage of probe sets giving interpretable signal tend to decrease (around 68% success rate based on a test of ~150 probe sets). Transcripts shorter than 500 bases remain difficult to detect by HT-smFISH in a systematic fashion.

Moreover, while HT-smFISH can be multiplexed, this is limited to imaging only up to four RNAs simultaneously by using distinct fluorescent dyes in which the pipeline does not use combinatorial or sequential labeling strategies. Also, HT-smFISH involves two rounds of PCR and one IVT reaction that must be quality controlled and are essential for the performance of the final in-situ detection. During hybridization, RNA probe sets are used and this requires working in a strict RNase-free environment, avoiding repeated probe set freeze-thaw cycles through aliquoting, and proper storage (probes are stable for years at -80°C). Finally, the cost benefits of HT-smFISH are best achieved with experiments involving 150-250 total RNA targets, in which an entire oligo pool of 10,000 oligos is utilized. If desirable, costs can further be reduced through in-house production of enzymes such as T7 RNA polymerase but this requires some knowledge of protein purification.

Comparison of HT-smFISH with other smFISH variants

Traditional smFISH methods use costly fluorescent oligonucleotides as probes and several variants were developed that use unlabeled probes and decrease the cost while increasing the signal-to-noise ratio (smiFISH[21], bDNA-FISH[22], FISH-STICs[23], padlock-FISH[49,50], HCR[51]). However, all these variants use chemically synthesized probes at scales sufficient for direct usage in smFISH experiments. Therefore, none of them work in high-throughput, at least at an affordable price[22]. HT-smFISH has the unique ability to generate individual smFISH probe sets in a high-throughput manner, which can be used in a simple protocol to perform smFISH in 96-well plates.

As stated above, HT-FISH has limited multiplexing abilities: the number of RNA species that can be simultaneously detected using HT-smFISH is the number of available fluorescently distinct dyes. This is in contrast to other techniques that combine barcoding with sequential hybridization (MERFISH[26] , seqFISH[27–29], and splitFISH[31]), which can simultaneously detect up to thousands of RNAs. However, these sequential FISH techniques involve complex barcoding schemes and a large number of precisely calibrated hybridization rounds that require dedicated microfluidic equipment. Moreover, custom algorithms are needed to align acquisitions, decode signals, and produce the individual image corresponding to each RNA species. In addition, clearing and embedding techniques are essential to increase the SNR, which can damage intracellular structures. Such multiplexed smFISH techniques also have limited flexibility: once a probe set for a selected group of RNA species has been designed, it must always be used entirely to label all the RNAs to permit decoding. As a consequence, probes for individual RNA species or a subset of them cannot be directly used. Finally, sequential FISH techniques cannot decode RNA identities when several RNA co-localize, precluding their use for abundant RNAs and in cases where different RNA species accumulate at the same place (P-bodies for instance). Thus, while the multiplexed smFISH techniques can be implemented in specialized labs and are powerful for correlative studies or in cases where biological samples are precious (biopsies for instance), they are poorly adapted for use in non-specialized labs, in cases where flexibility is needed, or where the RNAs of interest co-localize.

Similarly, sequencing-based methods often rely on an *in situ* reverse transcription (RT) step, which is limited by random priming, and variable priming efficiency *in-situ* (FISSEQ[36] and BaristaSeq[38]). Some RT-free methods exist (such as BOLORAMIS[39]), in which padlocks directly hybridize to *in-situ* RNAs. This however lowers the technique's sensitivity as it detects only 11-35% of RNAs compared to traditional smFISH[39]. Another common limitation of *in-situ* sequencing is the miss-hybridization of single padlocks that are amplified creating an important source of

background. While certainly powerful, sequencing based methods tend to involve several precise enzymatic amplifications steps (each carrying bias and raising costs) and require specialized chemistry and sample preparation (such as clearing, expansion[40], and 3D hydrogel DNA chips[41]) to increase resolution.

HT-smFISH therefore offers a simple alternative for performing medium to large-scale smFISH screens, usable in many molecular biology labs, and where probes for individual RNAs can almost indefinitely be re-used once generated, whether in high or low throughput format.

Experimental design

The HT-smFISH pipeline is presented in Figure 1 and consists of:

Oligo pool design

HT-smFISH probes start with an oligo pool, a complex solution containing 10,000 to 100,000 DNA oligonucleotides, each present in low copies and requiring amplification to generate the gene-specific probe sets. Each DNA oligonucleotide contains three elements: (i) the hybridization sequence that recognizes the RNA of interest (Hyb); (ii) two sequences flanking the hybridization sequence, called Readout X and Y, and common to all the oligonucleotides of the oligo pool; (iii) barcode sequences at both extremities of each oligo (BC 1 and BC 2), which are unique to the probe set targeting a particular RNA (Figure 2). Generally, 50-100 individual probes recognize every target transcript. Typical HT-smFISH oligo pools contain either 12,000 or 92,000 oligos, for 100-200 or 1000-2000 RNA targets respectively, and the design of the oligo pool is generated by Oligostan-HT, a custom script that optimizes the change in Gibbs free energy (ΔG) of hybridizing probes and removes repeated sequences. Oligostan-HT uses a multifasta file containing all the target sequences as an entry point (see Box 1), and it generates 50 to 100 oligo probes per RNA. These probes are filtered for repeated sequences and provide optimal

11

hybridization properties (see Box 2). The output file contains all the oligonucleotide sequences and displays both the hybridization sequence alone, and the full oligonucleotide to be ordered (BC 1-Readout Y-Hyb-Readout X-BC 2). To reduce the costs and simplify handling, the same set of barcodes are always used when designing different oligo pools.

Probe set generation

To generate RNA probe sets starting from an oligo pool, we use two PCR amplifications and an in vitro transcription reaction, all performed in 96-well plates (the scale is adaptable, Figure 3). Briefly, the 1st PCR uses gene-specific barcode primers to amplify a probe set of 50-100 oligos targeting an RNA of interest, with each well containing the primer pair to amplify the probe set for a different target RNA. For instance, the primers for the PCR in well 1 correspond to BC 1 and the reverse complement of BC 2. Those for well 2 correspond to BC 3 and the reverse complement of BC 4 and so on. Barcode primers are ordered and maintained in separate 96-well plates to simplify handling (see Supplementary Table 1 for sequences). The Oligostan-HT output file indicates the plate and well numbers that correspond to the correct barcode primer pair for each RNA target (Supplementary Table 2). The 2nd PCR uses readout primers to remove gene-specific barcodes and add a T7 promoter to all probe sets. Finally, the PCR products are used with T7 RNA polymerase to *in vitro* transcribe primary single stranded RNA probe sets, with each well containing one unlabeled probe set recognizing one RNA target. This reaction is carried out using a high yield enzyme and produces large quantities of single stranded RNA. The final RNA probe set integrities and quantities are measured using an appropriate capillary electrophoresis system. Another option is reverse transcribing the

RNA if users prefer working with DNA probe sets. However, we found that this has significant costs and further delivers low probe yields.

Probe set labeling

In this step, Readout sequences X and Y of the primary RNA probes are pre-hybridized to fluorescently labeled locked nucleic acids (LNAs) in 96-well plates, using a single heating/cooling step. Each LNA is labeled with two TYE 563 molecules, marking each probe-LNA duplex with four fluorophores in total (Figure 4A). This configuration produces reliable signals and allows a flexible experimental design by using LNAs labeled with different dyes. When performing single color RNA labeling experiments, we recommend the TYE 563 fluorophore as it provides the highest SNR. For multi-color labeling, we suggest using the fluorophores ATTO 488 (green channel), TYE 665 (far-red channel), and TYE 705 (near infrared channel, appropriate filter required). Since the LNA and Readout sequences are identical regardless of the fluorophore, pre-hybridization should first be done in different wells. Individually labeled duplexes (in different colors) are then combined in the *in-situ* hybridization step.

In situ hybridization

In high-throughput usage, hybridization is routinely done in 96-well plates, either with a plastic bottom for lower costs, or a glass bottom for higher quality imaging. The provided hybridization conditions are also suitable for other experimental formats. In the hybridization mixture, HT-smFISH replaces formamide by urea, which is non-toxic, and also uses a simpler hybridization buffer. This makes it more compatible with other in situ labeling methods such as immunolabeling or GFP fluorescence (which is not destroyed during the procedure, Figure 4B). Examples of HT-smFISH performed in cell lines are

presented in Figure 5 and highlight the high quality of the signals as well as the cell-to-cell heterogeneity in subcellular mRNA localization frequently observed. Probes generated using HT-smFISH are also usable for detecting RNA in primary cells and even frozen tissue sections using identical hybridization conditions as typical cell lines (Figure 6). HT-smFISH has been succesfuly applied to HeLa, HEK293T, HCT116, and RPE1 cells; as well as mice-derived primary neurons and colon tissue sections.

Whenever applicable, we recommend four general controls (see Figure 7 for example control images):

1. A negative control consisting of the hybridization mixture with neither primary probes nor fluorescent LNAs included. This control is useful for assessing the auto-fluorescence of cells.

2. A negative control consisting of the hybridization mixture with the fluorescent LNAs included, but not the primary probe set. This is important for ensuring that fluorescent LNAs do not produce a single molecule-like signal on their own (for example through aggregation). This control is useful for assessing potential false positive signal.

3. A positive control consisting of using a probe set that is known to produce a high quality signal. This control allows for detecting potential issues with other components (non-oligonucleotide elements) of the hybridization mixture. A transcript produced from a housekeeping gene can be chosen in general. However, depending on the biological question, RNAs with specific features can be used (we routinely use KIF1C localizing to cell extensions, and DYNC1H1 localizing in translation factories for instance[12,52]). Such RNAs will serve as a baseline when deriving and comparing quantitative features of RNA localization patterns. They are also very useful to assess image quality.

4.   In case a secondary labeling is used, include all negative controls for that procedure. If immunofluorescence is performed on top of the HT-smFISH for instance, including a well without any antibody and another with only the secondary antibody is ideal. Such controls help evaluate the impact of the additional staining procedure on the quality of the FISH signal and ensure immunolabeling specificity.

Imaging

This can be done by a variety of instruments, typically including widefield or confocal microscopes (we routinely use an Opera Phenix automated spinning disk). Imaging 96-well plates can be automated using simple macros. Typically, we acquire 3D image stacks spanning the entire cellular thickness. Maximum intensity projection (MIP) is then used to generate single 2D images (Figure 4C). MIPs however, can lead to underestimation of highly expressed RNAs and RNAs concentrated in a specific subcellular compartment. In such cases, we recommend examining full stacks, comparing with negative controls (RNAs not enriched in any compartment for instance), and using dedicated quantification methods such as cluster decomposition[44]. For visualization and analysis, we construct mosaics that span several fields of views (4x4 fields tiled per well for example). Typical HT-smFISH images of cell lines are presented in Figure 5, and highlight the high quality of the signals as well as the cell-to-cell heterogeneity in subcellular mRNA localization frequently observed. Images of RNA detected in mouse colon tissue sections and primary neurons are also presented in Figure 6.

Image analysis

Image analysis usually consists of three steps. First, mRNAs molecules are detected in the smFISH channel. When RNA aggregations are present in the image (such as P-bodies

or transcription sites), these can be decomposed in individual RNA molecules using a Gaussian mixture model[44]. Second, nuclei are segmented using the DAPI channel while cells can be segmented using either: (i) the smFISH channel (using cytoplasmic background resulting from the RNA labeling, may be difficult depending on the RNA stain and the sample to be segmented), (ii) or a dedicated channel containing a CellMask[TM] labeling (a cytoplasm-nucleus or cell membrane dye available in several colors, recommended for robust segmentation). The best segmentation results are often obtained by using recent deep learning methods[53]. The detected RNA molecules are then assigned to a cell. Third, an analysis at the single cell level is performed, where several spatial or morphological features relative to the cells and the mRNAs molecules are computed. These features include the proportion of mRNAs in specific subcellular regions, count of RNA foci, average mRNAs distance to cell membrane or nucleus, mRNAs dispersion index, etc… These features are then used for descriptive statistics, hypothesis testing, and visualization or as input for a machine learning classifier trained to recognize mRNAs localization patterns[43,48,44].

Materials

Biological Materials

- HeLa Kyoto (RRID:CVCL_1922,ATCC), HEK293T (RRID:CVCL_0045, ATCC) or any other cells of interest, growing in culture.

CAUTION Regularly test cell lines for mycoplasma contaminations and authenticity.

Reagents

CRITICAL All commercial reagents and buffers that are not mentioned in the Reagent Setup section should be stored according to the manufacturer's instructions, unless otherwise specifically noted.

Probe set generation

- Single stranded DNA oligo pools from Genescript or other preferred supplier. Details concerning target selection, design, ordering, storage and use are detailed in Box 1, 2 and 3.
- dNTP mix, 10 mM each (Thermo Fisher Scientific, cat. no. R0193).
- Phusion Hot Start II DNA Polymerase (Thermo Fisher Scientific, cat. no. F549L).
- Forward and reverse barcode primers for amplifying probe sets via PCR 1. These can be ordered from Integrated DNA Technologies (IDT) in 96-well plates at 100 µM or any other preferred vendor. Guidelines concerning barcode design are presented in Box 2 and 3 and a supplementary file containing the sequences to be ordered is provided (Supplementary Table 1).
- GoTaq® G2 Hot Start Taq Polymerase (Promega, cat. no. M7405).

- Forward and reverse primers (that bind Readout X and Y) for adding a T7 promoter via PCR 2. These can be ordered from Integrated DNA Technologies (IDT) at 100 µM or any other preferred vendor. See Table 1 for primer sequences.

- UltraPure™ DNase/RNase-Free Distilled Water (Thermo Fisher Scientific, cat. no. 10977035) or water of comparable purity.

- TE 1X, pH 8 (Sigma, cat. no. 93283-500ML).

- NucleoSpin Gel and PCR Clean-up kit (MACHEREY-NAGEL, cat. no. 740658.1).

- DNA 1K Reagent Kit (PerkinElmer, cat. no. CLS760673).

- UltraPure™ 1M Tris-HCl, pH 8.0 (Thermo Fisher Scientific, cat. no. 15568025).

- DL-Dithiothreitol (DTT) 1 M solution, molecular biology grade (Sigma, cat. no. 43816-10ML).

- Spermidine, molecular biology grade (Sigma, cat. no. 85558-1G).

- Magnesium chloride solution (1 M $MgCl_2$ in H_2O), molecular biology grade (Sigma, cat. no. 63069-100ML).

- Pyrophosphatase powder (Sigma, cat. no. I1643-500UN).

CRITICAL Store at -20°C for several years.

- HiScribe™ T7 RNA Polymerase kit (New England Biolabs, cat. no. E2040S). Alternatively, we use homemade T7 RNA polymerase prepared according to Rio D. et al[54]. In such a case, a ribonucleotide solution mix, 25mM each is required (such as New England Biolabs, cat. no. N0466L).

CRITICAL homemade T7 RNA pol can be stored for at least 2 years at -20°C without freezing. If using a commercial enzyme, follow the manufacturer's instructions.

- RNA Assay Reagent Kit (PerkinElmer, cat. no. CLS960010).

Cell culture

- DMEM, high glucose, GlutaMAX Supplement (Thermo Fisher Scientific, cat. no. 31966047).

- Fetal bovine serum (FBS) (eurobio SCIENTIFIC, cat. no. CVFSVF0601).

  CRITICAL: Aliquot and store FBS at -20 °C. Avoid freeze-thaw cycles.

- Penicillin-streptomycin, 10 000 U/ml) (Thermo Fisher Scientific, cat. no. 15140122)

- PBS, pH 7.4 (Thermo Fisher Scientific, cat. no. 10010023).

- 10x Trypsin solution (Sigma, cat. no. 59427C-100ML).

- Trypan blue stain (Thermo Fisher Scientific, cat. no. 15250061).

- Absolute ethanol (VWR, cat. no. MFCD00003568).

- 32% Paraformaldehyde solution (PFA, Electron Microscopy Sciences, cat. no. 15714S).

  CAUTION Handle PFA-containing solutions under a chemical hood. PFA is moderately and severely toxic by skin contact and inhalation respectively. PFA is a probable carcinogen.

In situ Hybridization

- Urea (Sigma, cat. no. U5378-5KG).

- Dextran sulfate sodium salt from *Leuconostoc spp.* (Sigma, cat. no. D8906-100G).

- tRNA from *Escherichia coli* (Sigma, cat. no. R1753-2KU).

- UltraPure™ DNase/RNase-Free Distilled Water (Thermo Fisher Scientific, cat. no. 10977035) or comparable water.

- Around 200 μg of locked nucleic acids (LNA) X and Y labeled with TYE563 fluorophore (Qiagen, custom synthesis, sequences given in Table 1).

- Triton™ X-100 for molecular biology (Sigma, cat. no. T8787-50ML).

- 20x Saline Sodium Citrate (SSC) buffer (Sigma, cat. no. S6639-1L).

- DAPI (4',6-diamidino-2-phénylindole, dichlorhydrate) (Thermo Fisher Scientific, cat. no. D1306).

- CellMask™ Deep Red stain (Thermo Fisher Scientific, cat. no. H32721) or any another desired cytoplasm-nucleus staining reagent.

- P-phenylenediamine (Sigma, cat. no. P6001-50G).

- 10x RNase-free PBS, pH 7.4 (Thermo Fisher Scientific, cat. no. AM9624).

- Sodium carbonate-bicarbonate buffer, pH 9.6 (Sigma, cat. no. C3041-50CAP).

- Glycerol for molecular biology (Merck, cat. no. 1040950250).

Equipment

- Electronic multichannel micropipette, ideally with 8 channels (e.g. Integra, cat. no. 4624).

- Multichannel aspiration system, ideally with an 8 needle head (e.g. Integra VACUSAFE, cat. no 158 310 complemented with an 8-channel adapter).

- Filtered sterile pipette tips (e.g. TipOne, VWR).

- Serological pipettes, assorted (Corning, or equivalent).

- Nuclease-free micro centrifuge tubes, 1.5, 2, or 5 ml depending on scale (e.g. Eppendorf, cat. no. 0030 125.150).

- 1 L autoclaved beakers and bottles.

- 0.2μm filtration system for volumes of around 1 L (Sigma, cat. no. SEGPU1145).

- Falcon centrifuge tubes, polypropylene, 15 ml (Dominique Dutscher, cat. no. 352096).

- Falcon centrifuge tubes, polypropylene, 50 ml (Dominique Dutscher, cat. no. 352070).

- 96-well PCR plate (Thermo Fisher Scientific, cat. no. AB-0600).

- PCR plate seal (Dominique Dutscher, cat. no. 4Ti-0500).

- PCR Thermocycler adapted to 96 well plates (e.g. Thermo Fisher Scientific, cat. no. 4375786).

- Eppendorf Twin.tec® PCR colorless plates, 96 wells (Sigma, cat. no. EP0030128648-25EA).

- Falcon TC-treated cell culture plate, 10 cm (Greiner Bio One, cat. no. 664160).

- Countess cell-counting chamber slides (Thermo Fisher Scientific, cat. no. C10314).

- Countess II cell counter (Fisher scientific, cat. no. AMQAX1000), or any other system for cell quantification.

- Light microscope (e.g. Thermo Fisher Scientific, cat. no.AMF5000).

- pH meter.

- Benchtop vortexer (e.g. Dominique Dutscher, cat. no. 079008).

- Tabletop centrifuge with an adapter fitting 96-well plates (Eppendorf, cat. no. 5810000010 or an equivalent system).

- Cell incubator at 37 °C, with regulated humidity and CO_2 (Thermo Fisher Scientific, cat. no. 51030284 or comparable system).

- Bench top water bath with adjustable temperature (Sigma, cat. no. Z615498).

- 96-well glass bottom (Dominique Dutscher, cat. no. 655892), or plastic bottom (Perkin Elmer, cat. no. 6055302) plates.

- NanoDrop or an equivalent nucleic acid measurement system capable of measuring nucleic acids coupled to fluorescent molecules.

- Opera Phenix™ High Content Screening System (PerkinElmer), or other fluorescent microscopes.

- Speed vac lyophilization system with a 96-well plate adapter (e.g. Eppendorf concentrator plus, cat. no. 5305000304).

- PerkinElmer LabChip GX capillary electrophoresis system or any equivalent system capable of quantifying DNA and RNA samples from 96-well plates.

- DNA 1K / 12K / Hi Sensitivity Assay LabChip (PerkinElmer, part no. 760517).

- DNA 5K / RNA / Charge Variant Assay LabChip (PerkinElmer, part no. 760435).

- Tecan Freedom EVO 200 liquid handler system (optional, can automate certain steps of the pipeline).

Software

- Harmony High-Content Imaging and Analysis Software (PerkinElmer), or any equivalent image acquisition software.

- ImageJ (https://imagej.nih.gov/ij/).

- Oligostan-HT, available in Docker Hub:

  (https://hub.docker.com/r/oligostan/oligostan_ht_rna)

- FISH-quant v2 (https://github.com/fish-quant)[42,44], RS-FISH (https://github.com/PreibischLab/RS-FISH)[45] StarFISH (https://github.com/spacetx/starfish)[46], deepBlink (https://github.com/BBQuercus/deepBlink)[47], or DypFISH (https://github.com/cbib/Dypfish)[48] for the analysis of smFISH images.

Reagent Setup

Preparation of homemade 10x *in vitro* transcription buffer (optional)

Start by preparing a 1M spermidine solution by dissolving 10 mg of spermidine powder in 68.8 µl RNase-free water. Prepare single-use aliquots (20 µl for instance) and store them at -20°C for several months in airtight Eppendorf tubes.

Prepare the 10x IVT buffer by mixing the components indicated in the table below. Alternatively, use a commercial "high yield" T7 buffer. Store at -20 °C for several years, follow manufacturer's instructions if using a commercial buffer.

Component	Amount (µl)	Final concentration in 10x buffer (mM)
1 M Tris-HCl pH 8	400	400
1 M DTT	100	100
1 M spermidine	20	20
1 M MgCl₂	200	200
Ultra-pure H₂O	280	-
Total	1000	-

Resuspension of LNA X and Y

Resuspend LNA X and Y tubes containing 200 µg of material in RNAse-free water to obtain a concentration of 200 ng/ µl.

Aliquot LNA X and Y and stored at -20 °C for several year (we recommend small aliquots of around 5 µl depending on the required scale), protected from direct sources of light.

CRITICAL We recommend measuring the exact concentration using an appropriate NanoDrop. Actual concentration may slightly vary due to synthesis variations. This variation could impact signal-to-noise ratios. If probes do not work well one should optimize amount of primary probes as well as the excess of LNA X and Y (by default 20%).

Resuspension of tRNA

Dissolve the 99.5 mg (2000 units) of tRNA in 4970 µl of nuclease-free water to obtain a solution at 20 mg/µl. Aliquot 1 ml of this solution in nuclease-free micro centrifuge tubes.

Store at -20°C for medium term use (around one year), or -80 °C for long term use.

Preparation of a 70% (wt/vol) by weight dextran sulfate solution

Weigh 28 g of dextran sulfate sodium salt in a 50 ml falcon tube. To this mass, add 30 ml of nuclease-free water. Homogenize and incubate at 65°C in a water bath. Vortex from time to time to dissolve the salt. Complete dissolution can take several hours. Adjust the final volume to 40 ml.

Stable at room temperature (25-28°C) and can be stored for around 2 years.

Preparation of 1 L of 10 M urea

Weigh 600 g of urea in an autoclaved beaker. Transfer this amount into an autoclaved or nuclease-free 1 L bottle. Add 500ml of nuclease-free water and homogenize. Incubate in a water bath at 50°C. Keep adding water until the urea completely dissolves reaching a volume of 1 L. Filter the solution through 0.2 µm pores.

Store at room temperature for a maximum period of one month (the fresher, the better). Plan ahead for large screens.

CRITICAL Urea at 10 M tends to precipitate out of solution and forms crystals visible by naked eye around temperatures lower than 20-25 °C. Before using urea in the hybridization mixture, make sure it is completely solubilized by incubating the solution in a water bath around 30-37 °C for 5-10 mins.

Preparation of DNA and cytoplasm staining solution for one 96-well plate

Resuspend 10 mg of DAPI in 2 ml nuclease-free deionized water to obtain a 5mg/ml stock solution. Likewise, prepare a stock CellMask solution at 10 mg/ml according to the manufacturer's instructions. To make a working solution for staining the nucleus and cytoplasm after the in-situ hybridization, add 1.4 µl of DAPI at 5 mg/ml and 1.4 µl of the stock far-red CellMask (or any other desired color) at 10 mg/ml to 7ml of PBS.

Store the stock DAPI and CellMask solutions either at 4°C for short-term storage (around 6 months), or -20 °C for long-term storage (up to 2 years) in the dark. Prepare working dilutions fresh, up to a few days before use and store them in the darl at 4°C.

Preparation of homemade mounting medium (optional)

Mix 50 mg of P-phenylenediamine with 5 ml of 10x PBS pH7.5 and homogenize well in a 50 ml falcon tube. Adjust the pH to 8 by adding a pH 9.6 sodium carbonate-bicarbonate buffer. Add around 45 ml of glycerol to reach a volume of 50 ml. Homogenize well by pipetting and vortexing. Aliquot appropriate volumes (10 ml for example) and store at -80°C for several years. Store working aliquots at -20°C for usage up to a year.

Procedure

Target selection and oligo pool design TIMING 4 d

1. Design and order an oligo pool containing probe sets against the target RNAs of interest and its associated barcode primers. Details and guidelines concerning target selection, barcode and probe set design, and oligo pool characteristics are found in Boxes 1, 2, and 3.

PAUSEPOINT The oligo pool and barcode primer plate can be stored at -20°C for several years.

Probe set generation TIMING 5-8 d

CRITICAL We recommend using nuclease-free barrier tips when handling all reagents in this section since the quality of the PCRs and in vitro transcription will directly affect the probe set performance.

CRITICAL Keep the stock tubes of the various enzymes used in this section at -20°C and use cold tube racks while performing the experiment.

2. Before starting PCR 1, prepare the following components:

   - For each probe set to be amplified, adjust the concentration of the forward and reverse primers which will anneal to the unique barcode sequences to 5 µM in nuclease free water (plan this ahead of starting the PCR, see Figure 7 for sequence example, and Supplementary Table 1 and 2 for an example of an oligo pool with barcodes).

   - Dilute the oligo pool 100 times in nuclease free water for a final volume of at least 265 µl per one PCR plate.

   PAUSEPOINT Diluted oligo pools and barcode primers can be stored at -20°C for several years.

3. In a nuclease-free tube, prepare a common mixture for PCR 1 according to the following (volumes are given for one well and one 96-well plate):

Component	Volume for one well (µl)	Volume for one 96-well plate (96 wells + 10 wells as pipetting margins)	Final concentration in a 50 µl reaction
Nuclease-free water	31	3286	-
5xPhusion hot start II DNA polymerase buffer	10	1060	1x
10 mM dNTPs	1	106	200 µM
Oligo pool diluted 100 times	2.5	265	-
Phusion hot start II DNA polymerase at 2 U/µl	0.5	53	1 U per reaction
Total	45	4770	-

4. Dispense 45 µl of this mixture per well in a 96-well PCR plate.

5. Add 2.5 µl of each barcode primer at 5 µM as detailed in the Oligostan-HT output file (Supplementary Table 2, also see Box 3) for a final concentration of 250 nM in 50 µl per well, and mix while pipetting.

CRITICAL STEP Take care not to cross-contaminate wells by primers. Non-specific amplification can potentially generate misidentified probe sets.

6. Carry out the PCR using a thermocycler according to this program:

Number of cycle	Denaturation	Annealing	Extension
1	98°C, 30 sec		
2-24	98°C, 30 sec	59°C, 30 sec	72°C, 30 sec
25			72°C, 5 min

Pre-heat the machine lid to 98 °C to avoid liquid evaporation and changes in component concentration.

CRITICAL STEP To achieve sufficient probe set amplification, 24 cycles of PCR is a good starting point. However, it is important to determine the exact number of cycles needed for each oligo pool as we found out this can vary from one oligo pool to another. Since we tested the same barcodes across different oligo pools, this is likely related to variations in the oligo pool synthesis process. With each new oligo pool, we advise to first generate ~20 probe sets to determine the optimal number of PCR 1 rounds before amplifying all probe sets. Too few cycles lead to invisible PCR 1 and PCR 2 bands and low RNA yields, while too many cycles generate high molecular weight smears in the PCR reactions. Once established, PCR parameters are reliable and replicable within the same oligo pool.

CRITICAL STEP We found the annealing temperature of 59 °C to work sufficiently well with our generated barcodes. Consider adjusting this temperature when using different barcodes.

PAUSE POINT The PCR product can be stored at -20°C for several months and before being used in PCR 2. The plate can also be stored at -80°C for long-term storage. This is particularly useful if the downstream RNA probe set needs to be extensively and repeatedly used at distant moments in time, or if it is accidentally damaged. In such cases, it is possible to generate RNA probe sets again, starting from this step.

7. Dilute part of the PCR reaction product 10 times in 1x Tris-EDTA (pH 8.0) and run a capillary electrophoresis as a quality control step. The exact dilution and loaded amount will vary depending on the analysis system but we routinely dilute 5 µl of this PCR reaction in 1x Tris-EDTA for a total volume of 50 µl. Load the diluted PCR product in a LabChip GX system (equipped with the DNA 1K/12K/Hi sensitivity LabChip and prepared with the DNA 1K Reagent Kit). See Figure 8A for an example of a capillary electrophoresis.

CRITICAL STEP It is important to quality control at least several wells to make sure the correct fragment was sufficiently amplified. Although the exact quantity will differ between amplified probe sets, an order of about 2 ng/µl (after diluting 10 times) is often suitable.

TROUBLESHOOTING

8. Before starting PCR 2, prepare the following components (see Figure 7 and Table 1 for primer sequences and expected amplification product):

- Adjust the concentration of the forward primer corresponding to a readout Y sequence with a T7 promoter at its 5' end (see sequence in Table 1) to 10 µM in at least 530 µl nuclease free water per 96-well plate.

- Adjust the concentration of the reverse primer corresponding to the reverse complimentary sequence of readout X (see sequence in Table 1) to 10µM in at least 530 µl nuclease free water per 96-well plate.

PAUSEPOINT Diluted primers for PCR 2 can be stored at -20°C for several years.

9. In a nuclease-free tube, prepare a common mixture for PCR 2 according to the following (volumes are given for one well and one 96-well plate):

Component	Volume for one well (µl)	Volume for one 96-well plate (96 wells + 10 wells for pipetting margins)	Final concentration in a 100 µl reaction

Nuclease-free water	60.5	6413	-
5x Colorless GoTaq® Flexi Buffer	20	2120	1x
25 mM MgCl$_2$	6	636	1.5 mM
10 mM dNTPs	2	212	200 µM
Forward readout Y primer at 10 µM	5	530	500 nM
Reverse readout X primer at 10 µM	5	530	500 nM
GoTaq® G2 Hot Start Taq Polymerase at 5 U/ µl	0.5	53	2.5 U per reaction
Total	99	10494	-

10. Mix well by vortexing or pipetting.

11. Dispense 99 µl of this mixture per well in a 96-well PCR plate.

12. Add 1 µl of the undiluted PCR 1 product plate per well. Take care not to cross-contaminate wells with PCR 1 products.

13. Carry out the PCR using a thermocycler according to this program:

Number of cycle	Denaturation	Annealing	Extension
1	98°C, 2 mins		
2-14	98°C, 30 sec	66°C, 30 sec	72°C, 30 sec
15			72°C, 5 min

Pre-heat the machine lid to 98 °C to avoid liquid evaporation and changes in component concentration.

PAUSE POINT The PCR product can be stored at -20°C for several months before being column purified.

14. Column purify the PCR 2 amplification product using a NucleoSpin Gel and PCR Clean-up kit according to the manufacturer's instructions, and elute in 140 µl. This step can be performed with a robotized pipetting system (for example we routinely use a Tecan Freedom EVO 200 liquid handler).

PAUSE POINT The purified product can be stored at -20°C for several months before being used in subsequent steps.

15. Dilute part of the PCR reaction product 25 times in 1x Tris-EDTA (pH 8.0) and run a capillary electrophoresis as a quality control step. The exact dilution and loaded amount will vary depending on the system but we routinely dilute 2 µl of this PCR reaction in 1x Tris-EDTA for a total volume of 50 µl, and load in a LabChip GX system (equipped with the DNA 1K/12K/Hi sensitivity LabChip and prepared with the DNA 1K Reagent Kit). See Figure 8B for an example of a capillary electrophoresis.

16. Lyophilize all remaining PCR 2 product using a speed vac (3 hours at 40 °C with no centrifugation).

17. Add 8 µl of nuclease-free water (example: UltraPure™ water) to each well for resuspension.

18. Use a PCR film to cover the 96-well plate well. Make sure the film covers all wells tightly. This is important to avoid liquid transfer across wells during subsequent vortexing.

19. Incubate for 15 minutes at 37 °C, while vortexing briefly and gently from time to time.

20. Spin down (or centrifuge briefly) the plate to make sure the droplets attached to the film are collected in the wells.

PAUSE POINT This plate containing purified PCR 2 amplification products in 8 µl can be stored at -20°C for several months before being used in subsequent in vitro transcription reactions.

21. Before starting the vitro transcription, prepare a 1 U/µl Pyrophosphatase solution. Dissolve the needed amount of stock pyrophosphatase powder for a specific experiment (at around 500 U/mg of powder) in nuclease-free water. Prepare at least 32 µl of solution at 1 U/µl for every full 96-well plate of probe sets to be transcribed. Prepare fresh and do not freeze-thaw the enzyme. This step may be skipped if a high-yield commercial T7 RNA polymerase is used.

22. In a nuclease-free tube, prepare a common mixture for the in vitro transcription reaction according to the following, volumes are given for one well and one 96-well plate (corresponding to 96 wells + 10 wells as pipetting margins):

Component	Volume for one well (µl)	Volume for one 96-well plate (96 wells + 10 wells for pipetting margins)	Final concentration in a 20 µl reaction
Nuclease-free water	3.3	349.8	-
10x homemade in vitro transcription buffer	2	212	1x
25 mM Nucleotides	6	636	7.5 mM
Homemade T7 RNA polymerase[54]	0.4 (to be determined depending on the homemade batch)	42.4 (to be determined depending on the homemade batch)	to be determined in small-scale reactions

1 U/µl Pyrophosphatase	0.3	31.8	0.3 U per reaction
Total	12	1272	-

CRITICAL STEP The amount of T7 RNA polymerase is essential for synthesizing sufficient RNA. We recommend doing small-scale tests to determine the exact amount and volume of the enzyme to use, for both homemade and commercial enzymes.

23. Add 12 µl of this mixture to each well in the plate containing purified PCR 2 amplification products suspended in 8 µl nuclease-free water for a total in vitro transcription reaction volume of 20 µl. This can be automated using a liquid handling robot.

24. Incubate the plate at 37°C for 4 hrs.

25. Dilute by adding 200 µl 2.8mM EDTA, pH 8 in nuclease-free water. Homogenize the mixture by pipetting while diluting.

PAUSE POINT This plate containing in vitro transcribed RNA probe sets can be stored at -80°C for several years (with 96-hole covers). We highly recommend avoiding freeze-thaw cycles when possible as they will lower the quality of the probe set. If individual probe sets are needed, we recommend aliquoting a fraction for short-term use and storing it at -20 or -80 °C while keeping the main plate at -80°C.

26. Quantify the amount and size of in vitro transcribed RNA using an appropriate capillary electrophoresis based system. We routinely use the same LabChip GX previously used to dose the PCR products (equipped with a DNA 5K / RNA / Charge Variant Assay LabChip prepared using the RNA Assay Reagent Kit). We often dilute the RNA two to three times in 1x Tris-EDTA (pH 8.0) for a total volume of 50 µl (volumes will depend on the exact system used, quantification performed according to the manufacturer's instructions). We routinely obtain a concentration of around 200-1000 ng/µl. A typical

profile is presented in Figure 8C. Make sure the used system and kit have a detection range that falls within the expected RNA concentration. If not, then adjust the sample dilution before measurement.

CRITICAL STEP It is essential to quantify the in vitro transcribed RNA as the two key factors determining FISH signal quality are the primary probe concentration and LNA/probe ratio.

TROUBLESHOOTING

27. Dilute a fraction of all RNA probe sets to 25 ng/µl in nuclease-free water for use in the high-throughput in-situ hybridization step (which requires 50 ng of probe set or 2µl of this working dilution). We routinely use a Tecan Freedom EVO 200 liquid handler system to do this.

PAUSE POINT Plates containing in vitro transcribed RNA (undiluted and adjusted at 25 ng/ µl) can be kept at -20°C for short-term storage (2-3 months), or -80°C for the long-term (recommended for the undiluted RNA probe sets). However, if dilutions are to be stored at -20°C for more than 3 months, make sure the freeze temperature does not fluctuate for best labeling performance. Expect a reduction of labeling quality if the diluted RNA probe sets go through many rounds of freezing-thawing. It is possible to add RNAse inhibitors for heavily used probe sets, although we usually do not use any. The vanadyl ribonucleoside complex (VRC) can be used, but it has a tendency to aggregate and for this reason can stick to the fixed cells, generating some auto fluorescence.

Cell plating, fixation, and permeabilization TIMING 2-3 d

CRITICAL This protocol focuses on performing HT-smFISH on cells grown in 96-well plates. For preparing frozen tissue sections for RNA labeling using HT-smFISH probes and hybridization conditions, refer to Box 4.

28. Supplement DMEM medium with 10% (vol/vol) FBS and 1% (vol/vol) penicillin-streptomycin. This complete medium should be stored at 4 °C and used within 2-3 months.

29. Seed 2x10^4 cells per well in a 96-well plate (see Reagent Setup, cell should be 20-25% confluent at the time of seeding), preferably with a glass bottom for best imaging quality (Dominique Dutscher, cat. no. 655892).

    CRITICAL STEP Some microscopy systems (including the Opera confocal system used in this protocol) cannot image the wells at the edge of the plate. In such cases, avoid seeding these wells.

30. Adjust the final volume per well to 200-300 µl.

31. Allow cells to reach the desired imaging confluency, ideally 60-80%. This depends on the doubling time, but often happens in 2 days.

32. Prepare a 4% PFA (vol/vol) solution by diluting the 32% stock solution in PBS. Prepare the dilution just before fixing the cells.

    CAUTION Handle PFA-containing solutions under a chemical hood. PFA is moderately and severely toxic by skin contact and inhalation respectively. PFA is a probable carcinogen.

    CRITICAL We recommend using a solution from a vial sealed under inert gas and opened for no more than three weeks and stored at 4°C in the dark after opening. Commercial unopened PFA vials contain an inert gas and can be stored at room temperature for several years (at least four, may vary according to reference)

33. Fix cells by replacing the growth medium with a volume of 4% PFA solution required to cover the well (200 µl for example).

34. Incubate cells in the presence of 4% PFA for 20 minutes at room temperature.

35. Aspirate PFA and rinse cells twice with 200 µl PBS.

36. While cells are in PBS, prepare a 70% (vol/vol) ethanol solution in RNase free water.

37. Replace the PBS with 200 µl of 70% ethanol solution.

38. Wrap the plate with wrapping film (example: parafilm) to minimize ethanol evaporation.

39. Store the plate overnight at 4 °C to permeabilize cells.

PAUSE POINT The plate can either be used the next day for *in-situ* hybridization or kept in 70% ethanol at 4 °C for up to 2 months as long as the wells do not dry.

In-situ hybridization and mounting TIMING 2-4 d

CRITICAL This section describes the in-situ hybridization procedure in 96-well plates. As described above (step 29) the outermost wells lining the plate cannot be imaged by certain microscopy systems resulting in 60 hybridizations per plate.

CRITICAL It is important to assign a few wells as negative and positive controls depending on the experimental aim(s). Recommendations for controls are found in the Experimental Design section, and example images in Figure 7.

40. Measure the exact concentration of the resuspended LNA X and Y (theoretically at 200 ng/µl, using a Nanodrop for example) since the exact value may differ due to synthesis variations. Make sure to spin down LNA X and Y tubes after defrosting and mix by pipetting to homogenize the solution.

CRITICAL STEP It is essential to measure the concentration of the LNAs as the two key factors determining FISH signal quality are the primary probe concentration and LNA/probe ratio. If signal quality is poor, try to optimize the amount of probes and LNA/probe ratio on ~10 probe sets.

41. In a nuclease-free PCR tube, prepare the LNA labeling solution according to the following (volumes are given for one well and one 96-well plate):

Component	Volume for one well (µl)	Volume for one 96-well plate (60 imaged	Final quantity per well

	wells + 10 wells for pipetting margins, µl)		
RNase-free water	1.56 (till a total volume of 2 µl)	Till a total volume of 280	-
20x SSC	0.2	14	1x
LNA X at 199 ng/µl	0.13	9.1	25 ng
LNA Y at 221.1 ng/µl	0.11	7.7	25 ng
Total volume	2	140	2 µl

Prepare the LNA labeling solution fresh for each experiment. Keep this tube on ice and away from direct sources of light while preparing the hybridization mix.

42. In a nuclease-free tube, prepare the hybridization solution according to the following, volumes are given for one well and one 96-well plate (corresponding to 60 imaged wells + 10 wells as pipetting margins):

Component	Volume for one well (µl)	Volume for one 96-well plate (60 wells + 10 wells for pipetting margins, µl)	Final concentration per well (in 100 µl final hybridization volume)
10 M urea	75	5250	7.5 M
20x SSC	5	350	1x
70% Dextran sulfate	14.3	1001	10%
20 mg/ml tRNA	1.7	119	0.34 mg/ml
Total volume	96	6720	-

Prepare the hybridization solution fresh for each experiment.

CRITICAL STEP Make sure the 10 M urea solution is completely solubilized before use in the hybridization mixture. See details in the reagent setup section.

CRITICAL STEP The 70% dextran sulfate solution is very viscous. To facilitate pipetting an accurate volume, use wide bore pipette tips. Alternatively, cut around 2 mm of pipette tips with an RNAse free blade.

43. Add the LNA labeling solution to the hybridization one.

44. Mix well either by pipetting, or vortexing briefly to homogenize the mixture.

CRITICAL STEP A homogenous mixture is essential for proportionate distribution of hybridization components across wells.

45. Add 98 µl of this mixture per well in a transparent PCR 96-well plate using a multi dispenser pipette.

46. From the plate containing primary probe sets adjusted at 25ng/µl (see step 27), add 2 µl (50 ng) of a primary probe set to each well using a multi-channel pipette. This will bring the total volume per well to 100 µl.

CRITICAL STEP Pipette carefully and change tips when adding the different probe sets. Well cross contaminations will lead to incorrect signals.

47. Seal the plate using 96-Well Plate sealing tapes.

48. Vortex the plate well for at least 1 minute.

49. Centrifuge the plate for 1 minute at 100 g at room temperature.

50. Using a thermocycler, denature and anneal the primary probes to the fluorescently labeled LNA X and Y according to the following:

Number of cycle	Denaturation	Annealing	Extension

1	90°C, 3 mins		
2		53°C, 15 mins at least, leave until usage	

This will produce a hybridization mix containing fluorescently labeled probe duplexes. Pre-heat the machine lid to 90°C to avoid liquid evaporation.

51. During this time, prepare 20 ml of the pre-hybridization solution according to the following:

Component	Volume (ml)	Final concentration
RNase free water	4	-
10 M urea	15	7.5 M
20x SSC	1	1x
Total	20	-

Prepare the pre-hybridization solution fresh for each experiment and use within a day.

52. Prepare cells (from step 39) for the in-situ hybridization:

- Aspirate the 70 % ethanol solution used to permeabilize cells.

- Add 200 µl of RNase free PBS.

- Rinse the cells in PBS for 1 minute.

- Aspirate the PBS.

- Add 200 µl of the pre-hybridization solution to each well.

- Incubate the cells in the washing solution for at least 15 mins and up to 1 hour at room temperature.

53. Replace the pre-hybridization solution by 100 µl of the hybridization mixture (now containing annealed probe duplexes).

    CRITICAL STEP Aspirate the pre-hybridization solution and add the hybridization mix progressively (3-4 wells at a time for instance) to avoid drying the cells.

54. Seal the plate using 96-well plate sealing tapes and Parafilm.

55. Place the plate in a plastic or glass closable container alongside wet tissue paper. This acts as a humidifying chamber to prevent hybridization mix evaporation.

56. Place the container in an incubator at 48 °C overnight.

57. The next day, prepare a fresh volume of the washing solution as described in step X (7.5 M urea in 1x SSC, around 120 ml per plate).

58. Remove the cell plate from the 48 °C incubator and aspirate as much hybridization mixture as possible.

    CRITICAL STEP Do this progressively (3-4 wells at a time for instance) while adding the washing solution (next step) to avoid well drying.

59. Add 200 µl of the washing solution per well using a multi-channel pipette.

60. Incubate the plate at 48 °C for 20 minutes.

61. Aspirate the washing solution and replace with 200 µl fresh washing solution.

62. Repeat these washing steps around eight times for a total incubation time of around 160 minutes in pre-warmed washing solution at 48°C for optimal results.

63. Rinse 3 times with RNase free PBS at room temperature

64. At this stage, an immunofluorescence can be performed. If not required, proceed with step 65 for labeling the DNA and cytoplasm or step 71 for directly mounting the sample. The exact procedure will depend on the used antibody but below are some general guidelines:

    - Prepare a 0.1-0.2% (vol/vol) Triton X-100 solution in PBS.

- Add 200 µl of 0.1-0.2% Triton X-100 per well and incubate for 5-10 mins at room temperature to permeabilize cells.

- Wash twice with RNase-free PBS.

- Either block cells in 1% molecular biology grade BSA in PBS for one hour at room temperature or proceed directly with the antibody(ies) incubation(s). Incubate in the dark to preserve the smFISH (and antibody if fluorescently coupled) signal.

  CRITICAL STEP Although BSA blocking can improve immunofluorescence quality, we recommend skipping this step, if possible, to minimize impacting the smFISH signal (many BSA batches are contaminated with RNases). In such a case, antibodies are also diluted in 0.1-0.2% Triton X-100 -PBS.

- Wash cells 3 times (5 minutes incubation each) in RNase-free PBS to remove unbound antibodies.

- If needed, incubate the cell with the secondary antibody for the required time and in the dark.

- Wash cells 3 times (5 minutes incubation each) in RNase-free PBS to remove unbound secondary antibodies.

  TROUBLESHOOTING

65. For labeling DNA and/or the cytoplasm, prepare the staining solution as described in the reagents setup section. We routinely use DAPI and a far-red CellMask™ to stain the DNA and nucleus-cytoplasm respectively.

66. Aspirate the PBS and add 100 µl of this staining solution per well to the 96-well plate containing cells.

67. Incubate the plate at room temperature and in the dark for 20 minutes.

68. During this time, remove the homemade (or any other preferred) mounting medium from -20 °C to thaw.

69. Remove the staining solution.

70. Wash twice with 200 μl RNase free PBS.

71. Replace the PBS with the mounting medium progressively taking care not to dry the cells.

72. Incubate at room temperature for around 30 minutes.

PAUSE POINT At this point the plate can be imaged directly, or sealed using 96-Well Plate sealing tapes and stored at -20 °C for 1-2 months. Note that the fluorescence quality may decrease while cells are stored for increasing times

Imaging TIMING 1 d

CRITICAL The exact details of the imaging process will depend on the microscope, the sample, and the aim of the experiment. Below we provide details for imaging using an automated spinning disk microscope (Opera Phenix), as we routinely do for imaging 96-well plates. For general guideline on imaging, see the Experimental Design section.

73. Use a 63x water or oil objective for imaging cells in culture and ensure a small enough pixel size to resolve RNA molecules (72-103 nm provided adequate results).

74. Set an appropriate naming scheme for saving the screen image files. Plan this in accordance with the image analysis pipeline to be used.

75. For every imaged well, start the acquisition in the far-red channel (the cytoplasm-nucleus CellMask™ stain), followed by the red (RNA), green (often a protein subcellular marker), and blue (DAPI) channels to minimize photobleaching.

76. Acquire images in 3D as stacks in the z plane. Z-stack sizes will vary depending on the thickness of the sample and desired coverage, but we generally acquire either 35 slices in Z with 0.3 μm steps, or 18 slices with with 0.6 μm steps. These configurations ensure proper single molecule detection across the sample thickness while avoiding oversampling.

77. Acquire 16 tiled fields of view pew well (each as 3D stacks) to ensure enough cells for visualization or analysis.

Image analysis TIMING 1-5 d

CRITICAL There are many analyses pipelines and tools for detecting and quantifying RNAs molecules in various specimen. The choice of analysis tool will largely depend on the type, volume, and complexity of generated data. Here, we outline the main steps involved in using FISH-quant v2 for analyzing HT-smFISH experiments revealing one mRNA in cultured cells in a high-throughput screen (tutorial available at https://github.com/fish-quant/big-fish-examples).

78. Install the Big-FISH (part of FISH-quant v2) python package for the analysis of smFISH images, alongside its dependencies (available alongside a documentation at https://github.com/fish-quant/big-fish). This package contains several subpackages and tools.

79. Start by using the Detection subpackage which can identify spots corresponding to single RNA molecules in 2D or 3D images based on a threshold. A major advantage Big-FISH provides is the ability to automatically set an appropriate detection threshold regardless of the SNR in the image. This reduces the need for human intervention and allows detecting HT-smFISH RNA signal in thousands of images.

80. If secondary labels with spot-like signals are found in the image (such as centrosomes or P-bodies), detect them using the same approach.

81. Analyze areas of dense RNA accumulation (transcription sites and RNAs in P-bodies for instance) using the provided cluster decomposition tool. Skip this step if RNA cluster analysis is not needed.

82. Use the Segmentation subpackage to segment cells an nuclei. Deep-learning-based approaches work well and are readily scalable.

83. If needed, use post-processing tools to optimize the segmentation (filling out gaps for instance).

84. Use the Cell matching subpackage to combine results of RNA detection with cell segmentation. This provides RNA expression and spatial information at the single cell level.

85. Finally, use the Localization feature extraction subpackage to quantify subcellular RNA patterns, compute RNA localization in a cell population, and perform statistical testing.

Troubleshooting

Troubleshooting suggestions are found in Table 2 below.

Timing

Step 1, design and ordering of an oligo pool: 4 d

Steps 2-7, PCR 1 to amplify specific probe sets: 1-2 d

Steps 8-13, PCR 2 to add a T7 promoter and remove barcodes: 1-2 d

Step 14, column-purifying PCR 2: 0.5-1 d

Steps 15-27, preparing, conducting, dosing, and diluting the in vitro transcription reaction: 2-3 d

Steps 28-39, growing, plating, fixing, and permeabilizing cells: 2-3 d

Steps 40-56: preparing and carrying out the in-situ hybridization: 1 d

Steps 57-63, washing the in-situ hybridization: 0.5-1 d

Step 64, optional immunofluorescence: 0.5-1 d

Steps 65-72 nuclear/ cytoplasmic labeling and mounting: 0.5 d

Step 73, Imaging: 1d

Steps 74-85, Image analysis: 1-5 d

Box 4, preparing mouse tissue sections for HT-smFISH: 3 d

<u>Anticipated results</u>

This procedure describes a cost-effective method to generate and hybridize RNA smFISH probes in a scalable format. This protocol has the unique ability to generate a large number of high-quality homemade sets of smFISH probes, which are directly and independently usable in flexible experimental designs. Examples of fluorescence images illustrating single mRNA molecules localized to a variety of subcellular locations revealed using HT-smFISH are presented in Figure 5. Such images allow a large-scale quantitative investigation of (i) RNA expression levels in single cells, (ii) inter-cellular heterogeneity in gene expression, (iii) and the subcellular distribution of RNA molecules. HT-smFISH can also be combined with secondary in-situ labeling protocols (immunofluorescence, protein tagging etc), further extending the utility of this method. Moreover, HT-smFISH probes can also be used to detect RNA in primary cells and tissue sections using the exact same hybridization conditions described in this protocol (Figure 6).

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The main data discussed in this protocol were generated as part of the studies published in the supporting primary research papers[6,13]. Example barcodes, barcode primers, and probe sets are provided in Supplementary Table 1 and 2.

Code availability

Oligostan-HT is available at https://hub.docker.com/r/oligostan/oligostan_ht_rna alongside a documentation and test data.

FISH-quant v2 for image analysis are available at https://github.com/fish-quant.

Author contributions

The HT-smFISH methodology was conceived by E.B. and developed by A.T., C.H., C.L., E.B., E.C., F.L., M.P., T.G., and V.G. Oligostan-HT was developed by T.G. with input from E.B., and C.L. for barcoding and target selection. The RNA localization analysis pipeline FISH-quant v2 was developed by A.I., F.M., and T.W. S.S. performed experiments in tissue. A.S. wrote the manuscript with help from E.B. A.S. made the figures. All authors reviewed the manuscript.

Competing interests

The authors declare competing financial interests

Related links

Key references using this protocol:

- Safieddine, A., Coleno, E., Salloum, S. et al. A choreography of centrosomal mRNAs reveals a conserved localization mechanism involving active polysome transport. Nat Commun 12, 1352 (2021). https://doi.org/10.1038/s41467-021-21585-7

- Kwon, O.S., Mishra, R., Safieddine, A. et al. Exon junction complex dependent mRNA localization is linked to centrosome organization during ciliogenesis. Nat Commun 12, 1351 (2021). https://doi.org/10.1038/s41467-021-21590-w

Figure legends

Figure 1: Overview of the HT-smFISH pipeline. *The pipeline starts with the desired RNA target selection and sequence extraction (Box 1). Sequences then serve as input for a dedicated code (Oligostan-HT) used to generate thermodynamically optimized probe sets, as well as barcodes (Box2). Probe sets are then ordered as single stranded DNA oligonucleotides in a custom oligo pool (Box 3). These oligonucleotides are used to generate transcript-specific RNA probe sets via PCR amplification and in vitro transcription (Steps 2-27). For high-throughput RNA screening experiments, cells are grown in 96-well plates (Steps 28-39). In situ hybridization is also carried out (Steps 40-72) and imaged (Steps 73-77) directly in 96-well plates. Images can be directly visualized and/or analyzed with various tools such as FISH-quant v2, RS-FISH, StarFISH, deepBlink, and DypFISH (Steps 78-85).*

Figure 2: Oligo pool design. *Oligo pools are complex mixtures of probe sets, each targeting an individual RNA species (colored boxes, 100-200 or 1000-2000 probe sets per oligo pool in total). Every DNA oligo of the pool contains a hybridization sequence that is flanked by two sequences (readout X and Y) common to all oligos in the pool. Different oligos belonging to the same probe set (targeting the same RNA) are flanked by identical RNA-specific unique barcodes (BC 1 and BC 2).*

Figure 3: RNA probe set generation from an oligo pool. *(A) A simplified example oligo pool containing 3 probe sets (red, blue, and green) is shown. Probe set A, presented in blue, is amplified using barcode-specific primers during PCR 1 (BC 1 and BC 2 primers differentially amplify probe set A only). PCR 1 is done in 96-well plates with one well containing one amplified probe sets (probe sets in red and green will be amplified in different wells in this example) (B) A second PCR (PCR 2) using a T7-Readout Y primer is then performed to add a T7 promoter (red)*

48

to all probe sets. PCR 2 is also done in 96-well plates and on all probe sets using the same pair of primers (since Readout X and Y are common across the entire oligo pool) . (C) Finally, T7 driven in vitro transcription done in 96-well plates is used to generate single stranded RNA probes. This produces single stranded primary RNA probes where each well in a 96-well plate contains a probe set specific for one RNA target. This will be used for the in-situ hybridization.

Figure 4: In situ hybridization and imaging: (A) A labeling step uses the Readout X and Y sequences to pre-hybridize locked nucleic acids oligos (LNAs) labeling primary RNA probes with TYE 563. This results in fluorescent duplexes. (B) Duplexes are then hybridized to their RNA targets in situ using mild hybridization conditions in 96-well plates. HT-smFISH is compatible with additional labeling steps such as immunofluorescence.(C) Imaging can be done using an automated spinning disk microscope. Images are acquired in 3D and then projected to produce 2D images covering the full thickness of the cells. Mosaics spanning several fields of view can be constructed to facilitate visual inspection.

Figure 5: HT-smFISH images highlighting the localization of mRNAs to various subcellular locations. The subcellular localization as well as the mRNA name are indicated above each panel. Experiments were done in either HeLa Kyoto cells (HSPA5, KIF1C, DDX11), HeLa cells stably expressing Centrin1-GFP to label centrosomes (NIN), or HEK293 cells stably expressing LSM14A-GFP to label processing bodies (P-bodies, DYNC1H1 and AURKA). The mRNA is shown in green and secondary labels in red. NIN mRNAs accumulating on centrosomes are indicated with yellow arrowheads. Green arrowheads indicate cytoplasmic mRNA accumulations in foci that are distinct from P-bodies (DYNC1H1 mRNA in translation factories), while red ones represent labeled P-bodies. Co-localization of the mRNA and the P-body marker is indicated with a yellow arrowhead (AURKA mRNAs). DNA is in blue and scale bars correspond to 10 microns.

Figure 6: HT-smFISH probe sets hybridized in mouse colon tissue sections and primary neurons. *(A) The schematics represents mouse colon tissue extraction and reparation for HT-smFISH labeling. Intestinal crypt cells are depicted in which the mRNA is labeled by probes shown in green. Green dots represent single mRNA molecules. The cartoon was created with BioRender.com (B) Micrographs of longitudinal and transversal mouse colon sections imaged by widefield microscopy. Left and green, dots corresponding to single KIF1C mRNA molecules labeled with TYE563; middle and red, autofluorescence of the sample obtained via imaging the mock GFP channel. Scale bars are 10 microns and DAPI stained nuclei are shown in blue. (C) micrographs of primary mouse neurons derived from embryonic brains and imaged with spinning disk confocal microscopy. Left, a global view showing NSMF mRNA labeled with labeled with TYE563 (shown in green) revealed by HT-smFISH, as well as HOMER (shown in red, imaged with the green channel) and MAP2 (shown in grey, imaged with the far-red channel) proteins revealed by double immunofluorescence acting as synaptic and dendritic markers respectively. Middle and right panels highlight areas of mRNA expression in neuronal cell bodies and extensions. Scale bars are 10 microns and DAPI stained nuclei are shown in blue.*

Figure 7: Images of negative and positive control wells routinely included in HT-smFISH experiments. *Schematics depict the hybridization mixture probe composition in negative and positive control wells. Images are maximum projections for confocal images of HeLa cells stained with the HT-smFISH hybridization buffer containing neither primary probe sets nor labeled LNAs, or labeled LNAs alone as a negative control; or hybridized with a complete mixture containing probe sets against KIF1C or DYNC1H1 mRNAs accumulating in cell protrusions and translation factories respectively as positive controls. DNA is stained with DAPI and presented in green, scale bars correspond to 10 microns.*

Figure 8: Examples of capillary gel electrophoresis used to quality control various steps during the probe set generation. PCR 1 (A), PCR 2 (B) and T7 in vitro transcription (IVT, C) examples showing primer and template sequences, alongside the resulting capillary electrophoresis with the expected bands indicated by the red arrow. Molecular size markers are in green. All gels were obtained using a LabChip GX system equipped with the appropriate assay chip and kit.

Box 1: Considerations for target selection

In HT-smFISH, one set of probes is designed per gene of interest. These probes recognize at least one of the gene transcript isoforms. Therefore, the design of a probe set starts with the selection of one transcript isoform per gene. Four criteria are used for this: (i) Reference Sequence (RefSeq) annotation, (ii) 'basic' tag in GENCODE annotation, (iii) median expression in GTEx V7 data and (iv), for coding genes, Coding DNA Sequence (CDS) length as indicated in ENSEMBL database.

The RefSeq collection provides a comprehensive, integrated, non-redundant and well-annotated set of transcript sequences. Similarly, the 'basic' tag of GENCODE transcripts prioritises full-length protein coding transcripts over partial or non-protein coding transcripts within the same gene, and intends to highlight those transcripts that will be useful to the majority of users.

We favour RefSeq annotations and, for each gene, retain the transcript associated with a RefSeq ID. If several RefSeq IDs are available, we select the RefSeq transcript with maximum median expression. If no RefSeq ID is available, we select transcripts with a 'basic' GENCODE tag and, if several 'basic' tagged transcripts are available, we select the basic transcript with maximum median expression. If no basic transcript is available, we select the transcript with maximum median expression.

For coding genes, in addition to median expression, we also consider CDS length and select RefSeq or basic transcripts with maximum median expression and maximum CDS length. If the transcript with maximum median expression is distinct from the one with maximum CDS length, we select both transcripts. Otherwise, we select only one single transcript corresponding to the transcript with maximum median expression and maximum CDS length. Once ENSTs have been identified, cDNA sequences are downloaded from ENSEMBL Biomart and a Multi-FASTA file containing all transcript sequences is created. For each sequence, the multifasta headings

must indicate an ENST and ENSG number and a gene name (for instance >ENSG00000119638|ENST00000238616|NEK9).

Box 2: Barcode and probe set design

Barcodes used to PCR-amplify probe sets are selected from a library of 240,000 orthogonal 25mer DNA sequences generated by Xu et al[55]. These sequences are filtered based on melting temperature (excluded Tm < 58 °C and Tm > 68 °C), repetitive sequences, restriction enzyme sites, uniqueness, and secondary structure formation. A list of barcodes is available at the Elledge lab website (https://elledge.hms.harvard.edu). For convenience, we use always the same barcodes for different oligo pools and the primer plates to be ordered are provided in Supplementary Table 1. In this file, each barcode has an ID and is identified in a precise plate and well.

Probes are designed using Oligostan-HT (available on Docker Hub, https://hub.docker.com/r/oligostan/oligostan_ht_rna), which uses a multifasta target sequence file as an entry point (file described in Box 1) and provides optimal hybridization sequences and also the entire probe sequences that should be ordered (BC 1-Readout Y-Hyb-Readout X-BC 2, Supplementary Table 2). The output file further indicates which barcode to use to amplify a given probe set, and the probe list is organized as to follow the barcode primer plates given in Supplementary Table 1: the first probe set is amplified with barcode primers BCID 1 and BCID 2, located on well A1 and A2 of primer plate one, the second probe set is amplified with barcode primers BCID 3 and BCID 4, located on well A3 and A4 of primer plate one, and so on.

Oligostan-HT primarily takes into consideration: (i) the Gibbs free energy of RNA-RNA hybridization at 48 °C (DG) and assigns it a score (dGScore, the higher the score the closer to the desired DG, with a default cutoff at 0.7 and max value of 0.99); (ii) and the GC content of probes (as detailed below). This script provides all possible probes meeting the requested criteria

for all sequences of the multifasta entry sequence file. Note that the script excludes probes which cross hybridize with repeated sequences found in the RNA target using the RepeatMasker tool. This is important since probes recognizing repeated sequences are a source of background. The length of the probe's hybridization sequence is 28 nt on average but will often vary between 23-38 nts to balance variation in GC content and obtain the desired DG. Probes are designed according to four successive criteria, with the aim of getting at least 50 probes per target RNA. The script starts with stringent criteria (avoiding extreme GC content compositions as probes with too low GC may not bind strongly enough to their target, and with too high GC may hybridize to each other and to off-target RNA in the cell), and then allows less stringency and probe overlap if less than 50 probes are found:

1. Design 1#: 0.35 < GC < 0.65, without probe overlap (with a distance of >1 nt between probes). The target sequences that do not return at least 50 probes in design #1 are next submitted to:

2. Design #2: 0.25 < GC < 0.7, without overlap (with a distance of >1 nt between probes). The target sequences that do not return at least 50 probes in design #2 are next submitted to:

3. Design #3: 0.35 <GC < 0.65, with probe overlap allowed (less than 15 nt). The target sequences that do not return at least 50 probes in design #3 are next submitted to:

4. Design #4: 0.25 < GC < 0.7, with overlap allowed (less than 15 nt).

The target sequences that do not return at least 50 probes are not included in the output file (this is modifiable). The script then combines the output of these four designs, and subjects them to 5 additional criteria described on Xu et al[55] for optimal hybridization properties:

1. aCompFilter: whether the probe contains an A nucleotide composition< 28%

2. aStackFilter: whether the probe does not contain AAAA stacks

3. cCompFilter: whether the probe contains an C nucleotide composition between 22 and 28%

4. cStackFilter: whether the probe does not contain CCCC stacks

5. cSpecStackFilter: whether the probe does not contain 4 non-consecutive Cs in any 6 consecutive nucleotides in the first 12 positions

The script will prioritize probes satisfying the most criteria possible (NbOfPNAS in the output file is the number of passed filters) while selecting the 50 probes for every RNA target. A strict cut-off for filters can also be selected. Based on our experience, we found filter 3 to be the most stringent and suggest bypassing that first when needing to obtain 50 probes (which is the default setting of Oligostan-HT). Next, Oligostan-HT adds the gene specific barcodes (starting with barcode ID1 of Supplementary Table 1) and Readout sequences X and Y (Table 1) to every oligo. Finally, the script adds extra nucleotides to equalize the length of all probes to facilitate oligo pool synthesis. Oligostan-HT provides 50 to 100 oligonucleotide probes per target sequence, and this is adjustable (Supplementary Table 2 for an example of Oligostan-HT output file).

To run Oligostan-HT, install the docker image from docker hub and run a container. In the terminal window of the container, type " cp -r /lib/Oligostan/export/ /". This will copy some files to the local folder you selected when starting the container. Open the ReadMeEBL file and follow the guidelines. Test sample files are provided.

In some cases, a single oligo pool may contain sequences for different experiments, as this is more cost effective than ordering several smaller oligo pools. In this case, the probes for the different target sets in the different experiments are designed separately, with each experiment having an Oligostan-HT file (which also allows to fine tune Oligostan-HT's parameters separately for each experiment). Oligostan-HT files are then and combined together in a single oligo pool in a second step. However, the files use the same barcodes across different experiments. Thus, combining them requires reallocating barcodes to ensure that one probe set is amplified with one barcode pair. To do this, we create an accompanying csv Set file describing

these target sets (one target set corresponds to all probe sets for a given experiment). The Set file should have four columns with the following names: "ENSG", "GeneName", "Set", and "ENST". This Set file hence allows merging different probe sets designed separately for different experiments into one oligo pool. To do this, use the script Merge_MultipleSets_RNAProbes.R, provided in the Oligostan-HT docker container (see the ReadMeEBL.txt file).

Box 3: Barcode primers and Oligo pool organization and ordering

We routinely order oligo pools from Genescript using 150 mer custom array synthesis. We use two oligo pool sizes of either 12,000 or 92,000 oligonucleotides. Oligo pools are shipped in 80 µl Tris EDTA with a concentration ranging from 70-150 ng/µl.

The barcode primers for the first PCR are ordered in plates containing 25 nmol of each barcode (adjusted to 100 µM, smaller quantities can be ordered as well) and a file containing sufficient primers for >1600 RNA targets is provided in Supplementary Table 1. Note that the plates have a number and the barcodes have an ID number, which is used in the same order by Oligostan-HT (i.e. the first probe sets in the output list of Oligostan-HT must be amplified with primers ID1 and ID2 located in well A1 and A2 of the primer plate n°1, the second probe set with primers ID3 and ID4 located in well A3 and A4 of the primer plate n°1, and so on).

Box 4: Preparing frozen tissue sections for HT-smFISH

Here, we use mouse colon tissue as an example, but the method extends to other tissues.

Additional materials

Biological materials

Mice (C57B6 strain, 10 months old, males). Mice were housed in temperature-controlled ventilated cages (20–22 °C) with a 12 h light-dark cycle, with percentage of humidity between 45 and 55%, and maintained in pathogen-free conditions in the institute animal facility.

CAUTION Any experiments involving live mice must conform to relevant institutional and national regulations. Here, mouse experiments were performed in strict accordance with the guidelines of the European Community (86/609/EEC) and the French National Committee (87/848) for care and use of laboratory animals, comply the ARRIVE guidelines and were approved by the French Ministry of Higher Education, Research and Innovation (reference APAFIS#18685) to be performed in the institute animal facility (agreement # F3417216).

Reagents

- Sucrose (Sigma, cat. no. S0389).

- Tissue embedding medium for cryotomy (Fisher Scientific, cat. no. 23-730-571, or any another suitable medium).

- Liquid nitrogen

  CAUTION Handle using proper safety equipment.

- Isopentane (Sigma, cat. no. 59060-10ML).

  CAUTION Very volatile and extremely flammable at room temperature and pressure. Read the associated safety sheets and handle using appropriate equipment.

- Cryostat (Thermo scientific, Microm hm 525, or a comparable system).

- SuperFrost Plus™ adhesive slides (Fisher Scientific, cat. no. 12-550-15, or any preferred slide).

Procedure

Tissue extraction and freezing TIMING 1d

1. Anesthetize the mouse properly (we use inhalant gas) and sacrifice it via cervical dislocation.

2. Extract the colon, wash it with PBS, and fix it with 4% PFA in PBS at room temperature for 3 hours.

3. Transfer the colon to a 4%PFA 30% sucrose in PBS solution and leave overnight at 4°C. We recommend preparing the diluted PFA and sucrose solutions fresh.

4. The next day, the tissue becomes fixed. Wash it once in PBS and then embed it in a suitable Tissue embedding medium for cryotomy according to the manufacturer's instructions. Immerse the tissue for 2-3 min in isopentane cooled to -80°C in liquid nitrogen.

   PAUSEPOINT Frozen embedded tissue can be stored at -80°C for up to 2 months.

Cryo-sectioning, section fixation and pemeabilization TIMING 2d

5. Generate sections of 6 µm thickness using a cryostat with a sharp razor and collect them on adherent slides (such as Superfrost™ Plus microscope slides). Preferably, collect one section per slide.

6. Further fix collected tissue sections (to better conserve structures) using 4%PFA in PBS for 20 min at room temperature and then wash 3 times with PBS.

7. Permeabilize sections with 70% ethanol overnight at 4°C.

PAUSEPOINT Slides can be stored for 2-3 weeks in ethanol at 4°C, but we recommend using fresh sections for minimal autofluorescence.

8. Wash tissue sections three times with PBS and immerse them in hybridization solution for at least 30 mins at room temperature just before carrying out the *in-situ*. Hybridization conditions, temperature, buffers, washing and LNA/probe set ratios are identical to those used for cell lines. Make sure used volumes are sufficient to cover the tissues sections well. We use a coplin jar.

Oligonucleotide or sequence name	Sequence (5' to 3')	Use
T7- readout Y primer	TAATACGACTCACTATAGGGTTACACTCGGACCTCGTCGACATGCATT	To remove barcodes and add a T7 promoter during PCR 2
Readout X primer	CACTGAGTCCAGCTCGAAACTTAGGAGG	To remove barcodes during PCR 2
Labeled LNA X	/5TYE563/CACTGAGTCCAGCTCGAAACTTAGGAGG/3TYE563/	To hybridize with readout X
Labeled LNA Y	/5TYE563/CATGTCGACGAGGTCCGAGTGTAA/3TYE563/	To hybridize with readout Y
Readout X (used in this protocol)	CCTCCTAAGTTTCGAGCTGGACTCAGTG	To hybridize with LNA X
Readout Y(used in this protocol)	TTACACTCGGACCTCGTCGACATGCATT	To hybridize with LNA Y
Readout alternative 1	ACACAACGCTAACGGGCGATTCTATAA	Can be used as a readout if needed (with Readout alternative 2, and the appropriate LNA sequence)
Readout alternative 2	TTACGAAGAGTCGCCGCGTTTCATTTA	Can be used as a readout if needed (with Readout alternative 1, and the appropriate LNA sequence)
Readout alternative 3	TTCTATGCGCGGCAATACGGATAACCTG	Can be used as a readout if needed (with Readout alternative 4, and the appropriate LNA sequence)

Readout alternative 4	TAGTGTAGCTCCGCGTTATGCCAG AAC	Can be used as a readout if needed (with Readout alternative 3, and the appropriate LNA sequence)

Table 1: List of oligonucleotides and readout sequences used across all oligo pools

Table 2: Troubleshooting table grouped by problem

Step	Problem	Possible reason	Possible solution
7	No bands appear after PCR 1	Insufficient amplification due to low cycle number	Increase the number of PCR cycles using a small number of probe sets
		Low quality oligo pool	Run a small amount of the oligo pool on capillary electrophoresis
		Mistake in barcode primers	Check carefully that the barcode sequences match the oligo pool sequences
	High molecular weight smears appear after PCR 1	Non-specific amplification due to high cycle number	Decrease the number of PCR cycles using a small number of test probe sets
26	The in vitro transcription reaction did not yield sufficient amounts of RNA	Quantity of commercial or homemade T7 RNA polymerase was not optimal	Perform small-scale reactions to determine the exact amount of enzyme needed for the desired amplification
		Insufficient amount of purified PCR 2 product	Check PCR 2 on gel electrophoresis
		RNA concentration is not within the instrument dosage range	Adjust the in vitro transcription dilution so that the final RNA concentration falls within detection range
41	Overall single molecule FISH signal quality is poor with high background	Ratio between the primary probes and fluorescent LNAs is not accurate / optimal	Make sure the primary probes and LNAs are properly dosed and diluted
			Decrease the amount of LNA (30% and 60% less as a start)
		Less efficient duplex formation with LNAs labelled with different fluorophores	Decrease the amount of LNA (30% and 60% less as a start) to have less free LNA in the hybridization mix
		Too much primary RNA probe	Decrease the amount of primary probes keeping the readout ratio constant
		Not enough washing time	Increase the number of washing steps and make

			sure the washing solution is pre-warmed
		High auto-fluorescence	Make sure the cell growth media and used mounting medium do not contribute to auto-fluorescence. Prepare fresh cells and make sure they do not dry.
		Non-specific primary probe binding	Possible cross-hybridization with an abundant RNA. We recommend a BLAST search for problematic probe sets.
			Raise the hybridization temperature up to 52°C
			Test increasing hybridization and washing temperature (50°C and 52°C)
		Irregular distribution of hybridization mixture components	Make sure the hybridisation mixture is well homogenised throughout the well
	Single molecule FISH signal quality is overall poor, faint signal	Ratio between the primary probes and fluorescent readout sequences is not accurate / optimal	Make sure the primary and LNAs are properly dosed and diluted
			Increase the amount of fluorescent LNAs (we recommend 30% and 60% more as a start)
		Check the number of probes and whether they come from the less favourable fourth design in Oligostan-HT	Try using a large target RNA sequence, for instance by including 5' and 3' UTRs
		Primary probes not efficiently hybridized to the fluorescent LNAs	Make sure the annealing temperatures and duration are correct
		Irregular distribution of hybridisation mixture components	Make sure the hybridisation mixture is well homogenised throughout the in-situ
64	Immunofluorescence lowering the quality of the FISH signal	The extra incubation time and/or reagents	Minimize antibody incubation times and

		reduces FISH signal quality	avoid the blocking step(s) when possible .
		Antibody or IF reagents contain RNAse	Change antibody or reagents

Supplementary Information

Supplementary Table 1: Barcodes and PCR barcode primers to be ordered. The primers are organized by plate and ID. Oligostan-HT uses this order such that the first two primers of plate 1 (wells A1 and A2) are used to PCR the first probe set in the Oligostan-HT output file, the next two (wells A3 and A4) for the second probe set, and so on.

Supplementary Table 2: A sample of Oligostan-HT's output. The output is shown in an excel sheet with several columns. ENST: Ensemble transcript identifier; ENSGx: Ensemble gene identifier; GeneName: name of gene or RNA; SET: the experiment to which the probe set belongs; dGOpt: ΔG37°C reference value (yielding the highest number of probes); Start/End Pos: the sart and end position of the probe with respect to the target transcript's coding sequence; ProbeSize: the length of the probe's hybridization sequence; Seq: the probe's hybridization sequence; dGScore and dG37: ΔG37°C score (higher is better) and value; GCpc: percentage of GC in the hybridization sequence; GCFilter: whether the probe satisfies the specified % GC filter; aCompFilter: whether the probe contains an A nucleotide composition< 28%; aStackFilter: whether the probe does not contain AAAA stacks; cCompFilter: whether the probe contains an C nucleotide composition between 22 and 28%, cStackFilter: whether the probe does not contain CCCC stacks; cSpecStackFilter: whether the probe does not contain 4 non-consecutive Cs in any 6 consecutive nucleotides in the first 12 positions; NbOfPNAS: the number of filters satisfied (from the PNAS paper of Xu et al[55]); PNASFilter: whether the probe satisfies the assigned number of PNAS filters; RSESeqFilter: whether probes contain a restriction enzyme site, filter not used; InsideUTR : whether the probe is inside a UTR; BC1ID: ID of the first barcode to be used to amplify the probe; BC1PN: the number of plate the barcode is stored in; BC1WP: position of the barcode in the 96-well plate; BC1: sequence of barcode 1; BC2ID, BC2PN, BC2WP, BC2: same as barcode 1, but for barcode 2; BC1-Y-Hyb-X-BC2: sequence resulting from concatenation of barcode1- Readout Y-Hybridization sequence- Readout X-barcode2; WithComplementSeq:

same but with a few additional nucleotides to normalize all oligonucleotide length to facilitate synthesis. This column has the oligo pool sequences to be ordered.
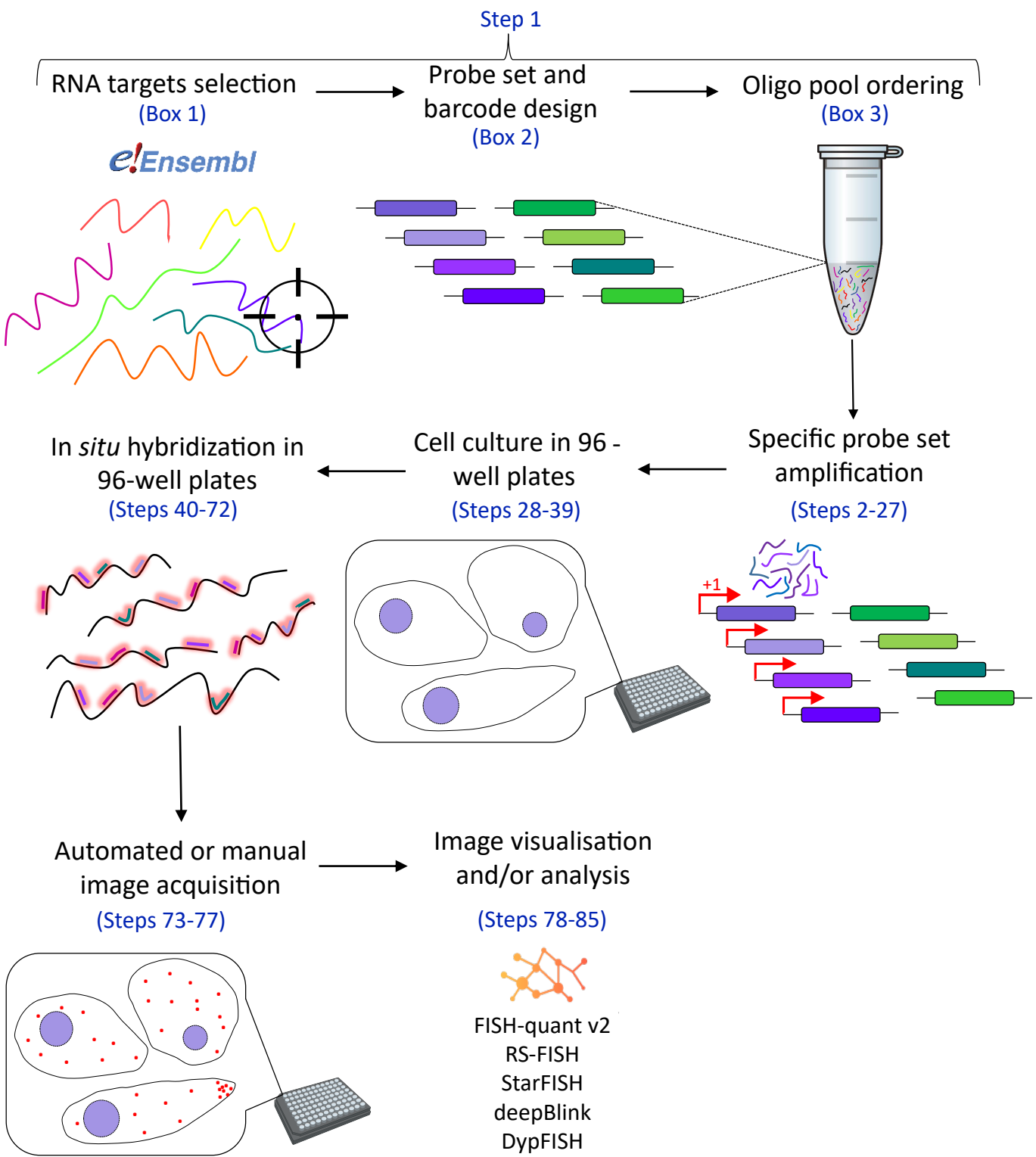
References:

1. Chin, A. & Lécuyer, E. RNA localization: Making its way to the center stage. *Biochim. Biophys. Acta Gen. Subj.* **1861**, 2956–2970 (2017).

2. Buxbaum, A. R., Haimovich, G. & Singer, R. H. In the right place at the right time: visualizing and understanding mRNA localization. *Nat. Rev. Mol. Cell Biol.* **16**, 95–109 (2015).

3. Kejiou, N. S. & Palazzo, A. F. mRNA localization as a rheostat to regulate subcellular gene expression. *Wiley Interdiscip. Rev. RNA* **8**, (2017).

4. Lee, C.-D. & Tu, B. P. Glucose-Regulated Phosphorylation of the PUF Protein Puf3 Regulates the Translational Fate of Its Bound mRNAs and Association with RNA Granules. *Cell Rep.* **11**, 1638–1650 (2015).

5. Böckler, S. & Westermann, B. Mitochondrial ER contacts are crucial for mitophagy in yeast. *Dev. Cell* **28**, 450–458 (2014).

6. Safieddine, A. *et al.* A choreography of centrosomal mRNAs reveals a conserved localization mechanism involving active polysome transport. *Nat. Commun.* **12**, 1352 (2021).

7. Lécuyer, E. *et al.* Global analysis of mRNA localization reveals a prominent role in organizing cellular architecture and function. *Cell* **131**, 174–187 (2007).

8. Yoon, Y. J. *et al.* Glutamate-induced RNA localization and translation in neurons. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E6877–E6886 (2016).

9. Singer-Krüger, B. & Jansen, R.-P. Here, there, everywhere. mRNA localization in budding yeast. *RNA Biol.* **11**, 1031–1039 (2014).

10. Chang, P. *et al.* Localization of RNAs to the mitochondrial cloud in Xenopus oocytes through entrapment and association with endoplasmic reticulum. *Mol. Biol. Cell* **15**, 4669–4681 (2004).

11. Medioni, C., Mowry, K. & Besse, F. Principles and roles of mRNA localization in animal development. *Dev. Camb. Engl.* **139**, 3263–3276 (2012).

12. Chouaib, R. *et al.* A Dual Protein-mRNA Localization Screen Reveals Compartmentalized Translation and Widespread Co-translational RNA Targeting. *Dev. Cell* **54**, 773-791.e5 (2020).

13. Kwon, O. S. *et al.* Exon junction complex dependent mRNA localization is linked to centrosome organization during ciliogenesis. *Nat. Commun.* **12**, 1351 (2021).

14. Sepulveda, G. *et al.* Co-translational protein targeting facilitates centrosomal recruitment of PCNT during centrosome maturation in vertebrates. *eLife* **7**, e34959 (2018).

15. Chen, L.-L. Linking Long Noncoding RNA Localization and Function. *Trends Biochem. Sci.* **41**, 761–772 (2016).

16. Cabili, M. N. *et al.* Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biol.* **16**, 20 (2015).

17. Kloc, M., Dallaire, P., Reunov, A. & Major, F. Structural messenger RNA contains cytokeratin polymerization and depolymerization signals. *Cell Tissue Res.* **346**, 209–222 (2011).

18. Garcia-Jove Navarro, M. *et al.* RNA is a critical element for the sizing and the composition of phase-separated RNA-protein condensates. *Nat. Commun.* **10**, 3230 (2019).

19. Femino, A. M., Fay, F. S., Fogarty, K. & Singer, R. H. Visualization of Single RNA Transcripts in Situ. *Science* **280**, 585–590 (1998).

20. Pichon, X., Lagha, M., Mueller, F. & Bertrand, E. A Growing Toolbox to Image Gene Expression in Single Cells: Sensitive Approaches for Demanding Challenges. *Mol. Cell* **71**, 468–480 (2018).

21. Tsanov, N. *et al.* smiFISH and FISH-quant - a flexible single RNA detection approach with super-resolution capability. *Nucleic Acids Res.* **44**, e165 (2016).

22. Battich, N., Stoeger, T. & Pelkmans, L. Image-based transcriptomics in thousands of single human cells at single-molecule resolution. *Nat. Methods* **10**, 1127–1133 (2013).

23. Sinnamon, J. R. & Czaplinski, K. RNA detection in situ with FISH-STICs. *RNA* **20**, 260–266 (2014).

24. Marx, V. Method of the Year: spatially resolved transcriptomics. *Nat. Methods* **18**, 9–14 (2021).

25. Method of the Year 2020: spatially resolved transcriptomics. *Nat. Methods* **18**, 1–1 (2021).

26. Wang, G., Moffitt, J. R. & Zhuang, X. Multiplexed imaging of high-density libraries of RNAs with MERFISH and expansion microscopy. *Sci. Rep.* **8**, 4847 (2018).

27. Lubeck, E., Coskun, A. F., Zhiyentayev, T., Ahmad, M. & Cai, L. Single-cell in situ RNA profiling by sequential hybridization. *Nat. Methods* **11**, 360–361 (2014).

28. Eng, C.-H. L. *et al.* Transcriptome-scale super-resolved imaging in tissues by RNA seqFISH+. *Nature* **568**, 235 (2019).

29. Shah, S. *et al.* Dynamics and Spatial Genomics of the Nascent Transcriptome by Intron seqFISH. *Cell* **174**, 363-376.e16 (2018).

30. Eng, C.-H. L., Shah, S., Thomassie, J. & Cai, L. Profiling the transcriptome with RNA SPOTs. *Nat. Methods* **14**, 1153–1155 (2017).

31. Goh, J. J. L. *et al.* Highly specific multiplexed RNA imaging in tissues with split-FISH. *Nat. Methods* **17**, 689–693 (2020).

32. Chen, X., Teichmann, S. A. & Meyer, K. B. From Tissues to Cell Types and Back: Single-Cell Gene Expression Analysis of Tissue Architecture. *Annu. Rev. Biomed. Data Sci.* **1**, 29–51 (2018).

33. Close, J. L., Long, B. R. & Zeng, H. Spatially resolved transcriptomics in neuroscience. *Nat. Methods* **18**, 23–25 (2021).

34. Zhuang, X. Spatially resolved single-cell genomics and transcriptomics by imaging. *Nat. Methods* **18**, 18–22 (2021).

35. Larsson, L., Frisén, J. & Lundeberg, J. Spatially resolved transcriptomics adds a new dimension to genomics. *Nat. Methods* **18**, 15–18 (2021).

36. Lee, J. H. *et al.* Fluorescent in situ sequencing (FISSEQ) of RNA for gene expression profiling in intact cells and tissues. *Nat. Protoc.* **10**, 442–458 (2015).
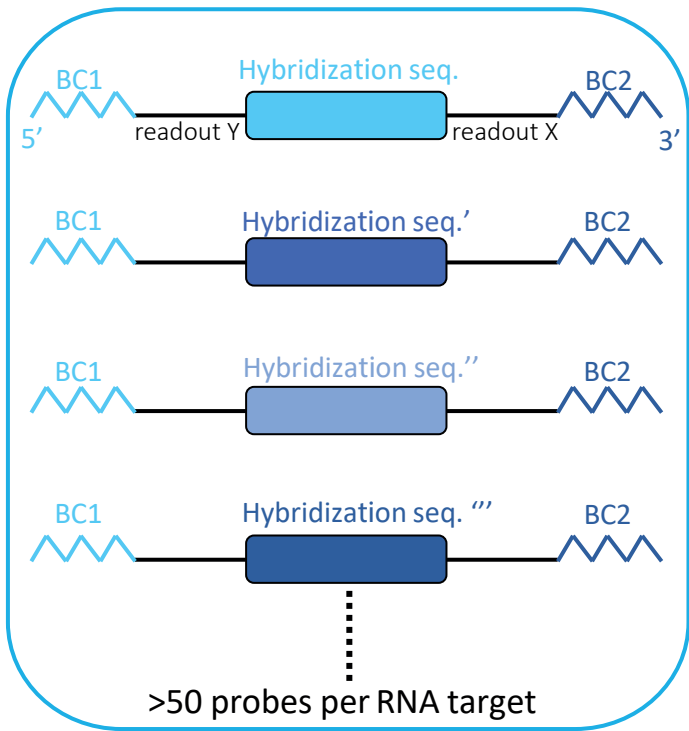
37. Ke, R. *et al.* In situ sequencing for RNA analysis in preserved tissue and cells. *Nat. Methods* **10**, 857–860 (2013).

38. Chen, X., Sun, Y.-C., Church, G. M., Lee, J. H. & Zador, A. M. Efficient in situ barcode sequencing using padlock probe-based BaristaSeq. *Nucleic Acids Res.* **46**, e22 (2018).

39. Liu, S. *et al.* Barcoded oligonucleotides ligated on RNA amplified for multiplexed and parallel in situ analyses. *Nucleic Acids Res.* **49**, e58 (2021).

40. Alon, S. *et al.* Expansion sequencing: Spatially precise in situ transcriptomics in intact biological systems. *Science* (2021) doi:10.1126/science.aax2656.

41. Wang, X. *et al.* Three-dimensional intact-tissue sequencing of single-cell transcriptional states. *Science* (2018) doi:10.1126/science.aat5691.

42. Mueller, F. *et al.* FISH-quant: automatic counting of transcripts in 3D FISH images. *Nat. Methods* **10**, 277–278 (2013).

43. Samacoits, A. *et al.* A computational framework to study sub-cellular RNA localization. *Nat. Commun.* **9**, 4584 (2018).

44. Imbert, A. *et al.* FISH-quant v2: a scalable and modular tool for smFISH image analysis. *RNA* **28**, 786–795 (2022).

45. Bahry, E. *et al. RS-FISH: Precise, interactive, fast, and scalable FISH spot detection*. 2021.03.09.434205 https://www.biorxiv.org/content/10.1101/2021.03.09.434205v2 (2021) doi:10.1101/2021.03.09.434205.

46. Perkel, J. M. Starfish enterprise: finding RNA patterns in single cells. *Nature* **572**, 549–551 (2019).

47. Eichenberger, B. T., Zhan, Y., Rempfler, M., Giorgetti, L. & Chao, J. A. deepBlink: threshold-independent detection and localization of diffraction-limited spots. *Nucleic Acids Res.* **49**, 7292–7297 (2021).

48. Savulescu, A. F. *et al.* Interrogating RNA and protein spatial subcellular distribution in smFISH data with DypFISH. *Cell Rep. Methods* **1**, 100068 (2021).

49. Rouhanifard, S. H. *et al.* ClampFISH detects individual nucleic-acid molecules using click chemistry based amplification. *Nat. Biotechnol.* (2018) doi:10.1038/nbt.4286.

50. Yaroslavsky, A. I. & Smolina, I. V. Fluorescence imaging of single-copy DNA sequences within the human genome using PNA-directed padlock probe assembly. *Chem. Biol.* **20**, 445–453 (2013).

51. Choi, H. M. T. *et al.* Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nat. Biotechnol.* **28**, 1208–1212 (2010).

52. Pichon, X. *et al.* The kinesin KIF1C transports APC-dependent mRNAs to cell protrusions. *RNA* rna.078576.120 (2021) doi:10.1261/rna.078576.120.

53. Ouyang, W., Mueller, F., Hjelmare, M., Lundberg, E. & Zimmer, C. ImJoy: an open-source computational platform for the deep learning era. *Nat. Methods* **16**, 1199–1200 (2019).

54. Rio, D. C. Expression and Purification of Active Recombinant T7 RNA Polymerase from E. coli. *Cold Spring Harb. Protoc.* **2013**, pdb.prot078527 (2013).

55. Xu, Q., Schlabach, M. R., Hannon, G. J. & Elledge, S. J. Design of 240,000 orthogonal 25mer DNA barcode probes. *Proc. Natl. Acad. Sci.* **106**, 2289–2294 (2009).

RNA targets selection
(Box 1)

Probe set and
barcode design
(Box 2)

Oligo pool ordering
(Box 3)

e!Ensembl

In *situ* hybridization in
96-well plates
(Steps 40-72)

Cell culture in 96 -
well plates
(Steps 28-39)

Specific probe set
amplification
(Steps 2-27)

+1

Automated or manual
image acquisition
(Steps 73-77)

Image visualisation
and/or analysis
(Steps 78-85)

FISH-quant v2
RS-FISH
StarFISH
deepBlink
DypFISH

Probe set A for RNA A

BC1　Hybridization seq.　BC2
5'　readout Y　readout X　3'

BC1　Hybridization seq.'　BC2

BC1　Hybridization seq.''　BC2

BC1　Hybridization seq. '''　BC2

>50 probes per RNA target

Probe set B for RNA B

BC3　Hybridization seq.　BC4
5'　readout Y　readout X　3'

BC3　Hybridization seq.'　BC4

BC3　Hybridization seq.''　BC4

BC3　Hybridization seq.'''　BC4

Probe set C for RNA C

BC5　Hybridization seq.　BC6
5'　readout Y　readout X　3'

BC5　Hybridization seq.'　BC6

BC5　Hybridization seq.''　BC6

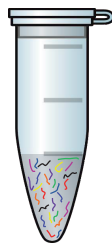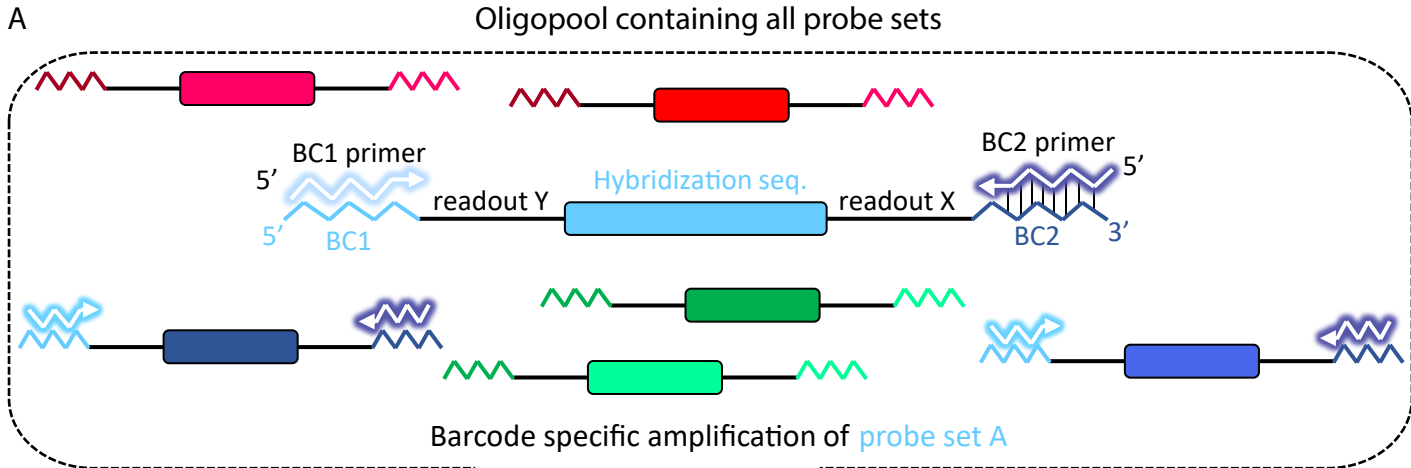BC5　Hybridization seq.'''　BC6

+ other probe sets against
100-200 or 1,000-2,000 targets,
all with common readouts X and Y
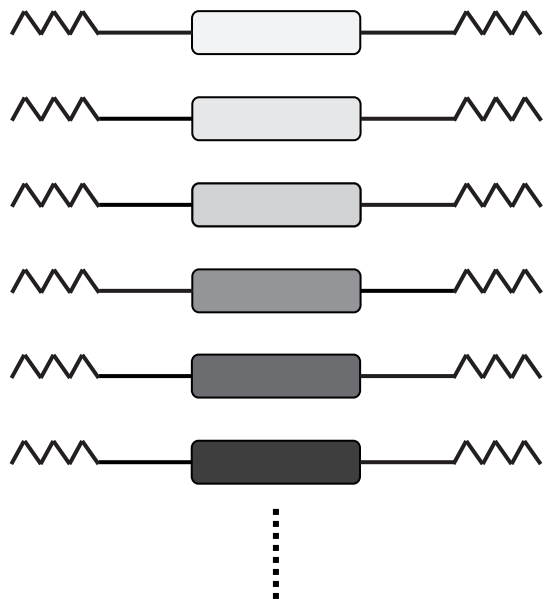
All probe sets mixed in an
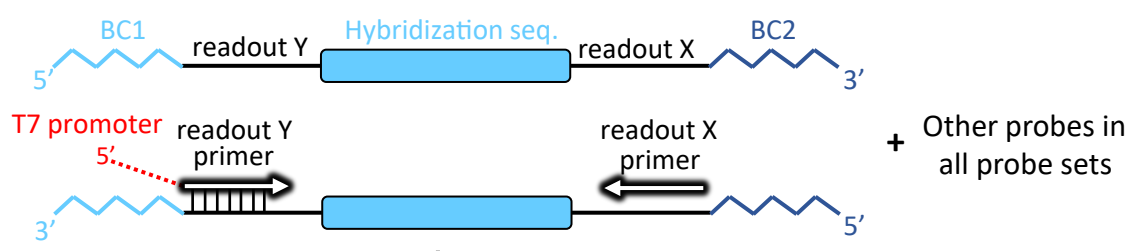oligo pool of 12,000 or 92,000 oligos

A Oligopool containing all probe sets

BC1 primer

BC2 primer

readout Y   Hybridization seq.   readout X

Barcode specific amplification of probe set A

One probe set amplified per well

PCR 1

BC1   readout Y   Hybridization seq.   readout X   BC2   + Other probes in probe set A

B Adding a T7 promoter

BC1   readout Y   Hybridization seq.   readout X   BC2

T7 promoter   readout Y primer   readout X primer   + Other probes in all probe sets

PCR 2

T7 promoter   readout Y   Hybridization seq.   readout X   + Other probes in all probe sets
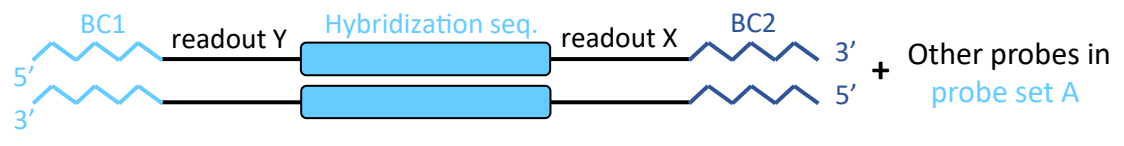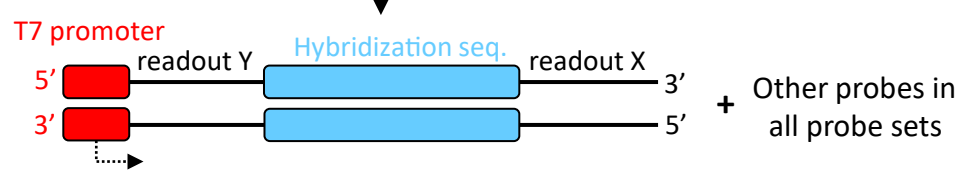
C In vitro transcription

Single stranded primary RNA probes

Hybridization seq.

readout Y   readout X

One probe set per well

probe set A, well A1

A

readout X   Hybridization seq.   readout Y

3'   5'

+

TYE 563

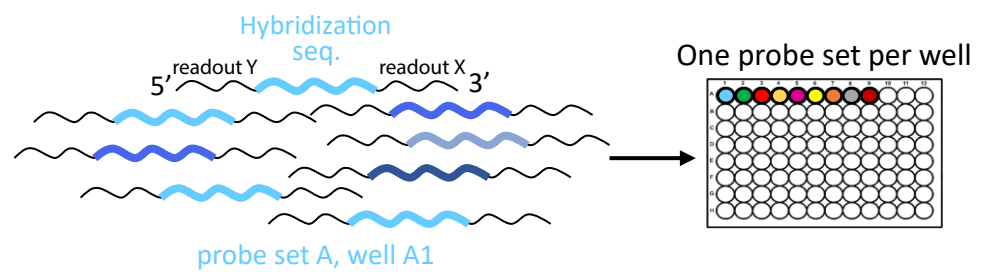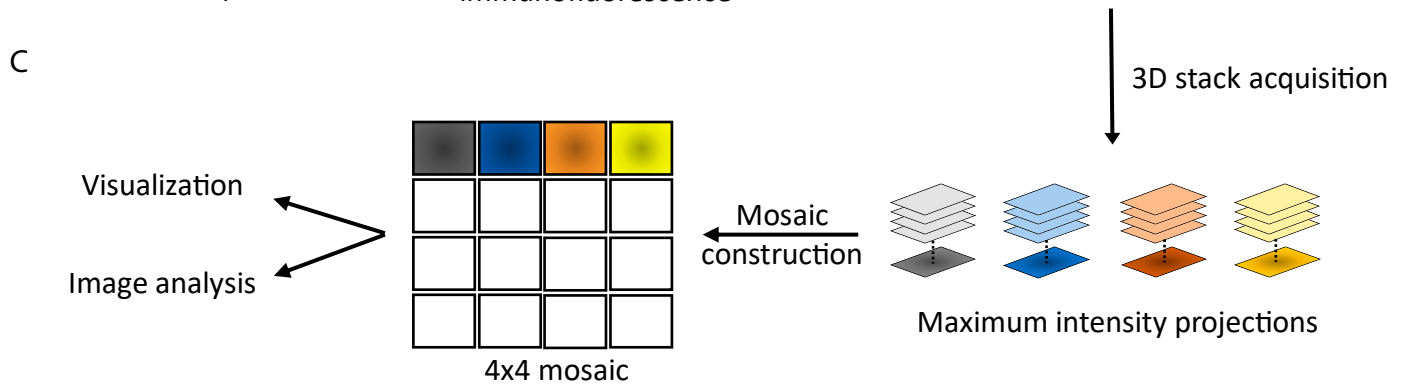Fluorescent LNA X   Fluorescent LNA Y

Pre-hybridization →

3' 5' 3'   5' 3' 5'

Fluorescent duplexes of probeset A

In-situ hybridization

B

3' 5' 3'   5' 3' 5'

5'   mRNA target   3'

Hybridization in 96 well plates

One mRNA per well — Washing → Optional immunofluorescence — Washing and mounting → Automated spinning disk confocal imaging

3D stack acquisition

C

Visualization

Image analysis

← Mosaic construction

Maximum intensity projections

4x4 mosaic

Centrosomal mRNA, *NIN*

Centrin1-GFP

Perinuclear mRNA, *HSPA5*

Intranuclear mRNA, *DDX11*
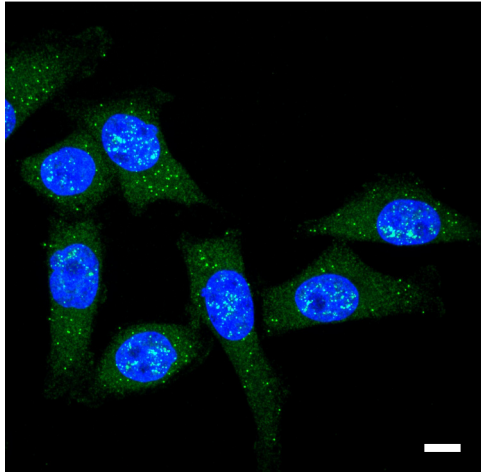
Protrusion mRNA, *KIF1C*
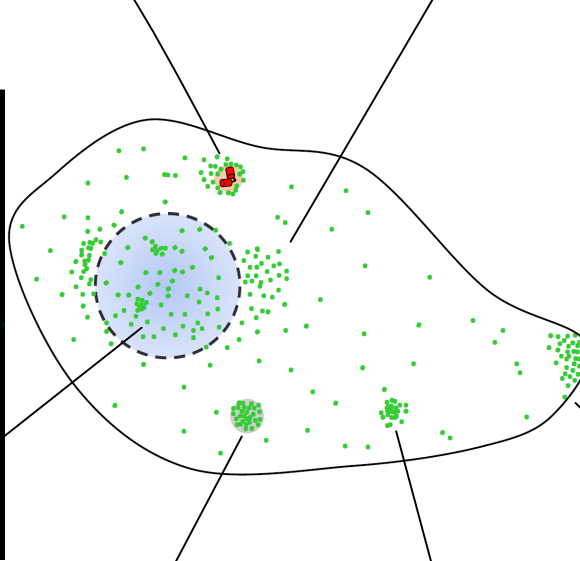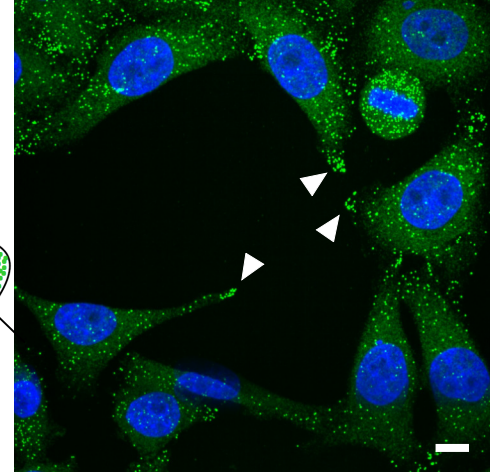
P-body mRNA, *AURKA*

LSM14A-GFP
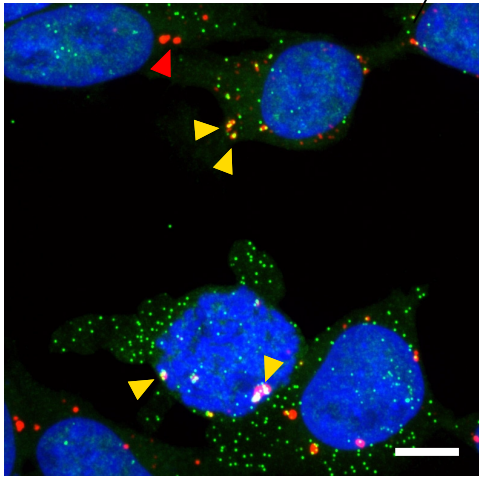
Foci mRNA, *DYNC1H1*

LSM14A-GFP

A Colon tissue preparation   HT-smFISH probes and hybridization conditions   Coverslip imaging

B

| *KIF1C* mRNA | Autofluorescence | Merge |

Longitudinal

Transversal

C

NSMF mRNA
Homer and MAP2 proteins | *NSMF* mRNA | Merge
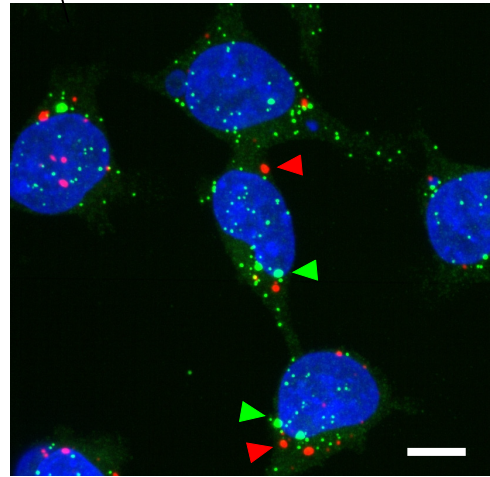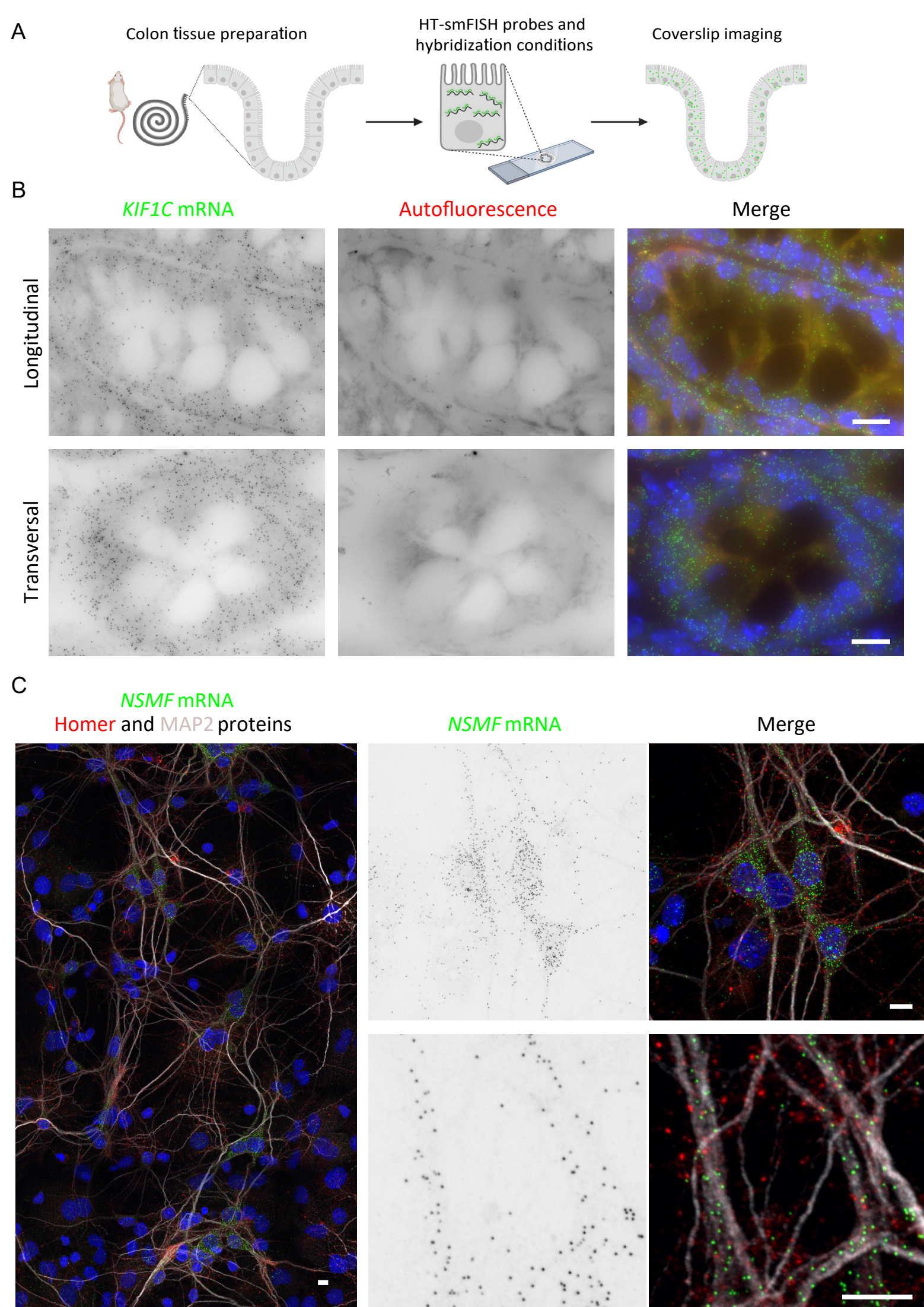
Negative control:
no probe set
no LNAs

Negative control:
no probe set
with LNAs

Positive control:
DYNC1H1 probes
with LNAs

Positive control:
KIF1C probes
with LNAs

A

Barcode 1 primer

5′ GCTCCCCATGCTAGAGGGACAAGGT

Hybridization sequence

Barcode 2 primer

GGGGTTTGGTTTACTCCTGGTTACC5′

5′GCTCCCCATGCTAGAGGGACAAGGTTTACACTCGGACCTCGTCGACATGCATTCACATTTCCAAGGAACCGTCCTCTTGTACGCCTCCTAAGTTTCGAGCTGGACTCAGTGCCCCAAACCAAATGAGGACCAATGG3′

Barcode 1   readout Y   readout X   Barcode 2

PCR 1 capillary electrophoresis

A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A12

1500   ← Markers
1000
700
500   500
400   400
300   300

200   200
135 b.p. amplicon →   150   150
                      100   100
50   50
25   25   ← primers

0   0   ← Markers

B

5′TAATACGACTCACTATAGGG

T7-readout Y primer

TTACACTCGGACCTCGTCGACATGCATT

Hybridization sequence

readout X primer

GGAGGATTCAAAGCTCGACCTGAGTCAC5′

3′CGAGGGGTACGATCTCCCTGTTCCAAATGTGAGCCTGGAGCAGCTGTACGTAAGTGTAAAGGTTCCTTGGCAGGAGAACATGCGGAGGATTCAAAGCTCGACCTGAGTCACGGGGTTTGGTTTACTCCTGGTTACC5′

Barcode 1   readout Y   readout X   Barcode 2

PCR 2 capillary electrophoresis

A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A12

1500   1500   ← Markers
1000   1000
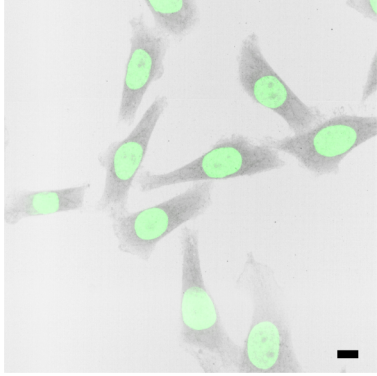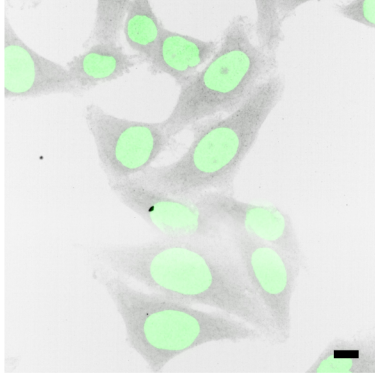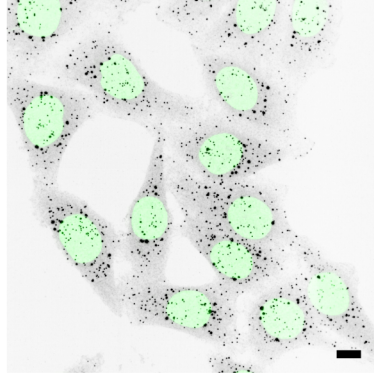700   700
500   500
400   400

300   300

200   200
106 b.p. amplicon →   150   150
                      100   100
50   50
25   25

0   0   ← Markers

C

T7 promoter   **readout Y**   **readout X**

5′ TAATACGACTCACTATAGGGTTACACTCGGACCTCGTCGACATGCATTCACATTTCCAAGGAACCGTCCTCTTGTACGCCTCCTAAGTTTCGAGCTGGACTCAGTG3′

3′ ATTATGCTGAGTGATATCCCAATGTGAGCCTGGAGCAGCTGTACGTAAGTGTAAAGGTTCCTTGGCAGGAGAACATGCGGAGGATTCAAAGCTCGACCTGAGTCAC5′

+1   5′GGGTTACACTCGGACCTCGTCGACATGCATTCACATTTCCAAGGAACCGTCCTCTTGTACGCCTCCTAAGTTTCGAGCTGGACTCAGTG3′

5′   readout Y   Hybridization sequence   readout X   3′

RNA primary probe

IVT capillary electrophoresis

A1 A2 A3 A4 A5 A6 A7 A8 A9 A10 A11 A12 B1 B2 B3 B4 B5 B6 B7 B8 B9 B10 B11 B12

6000 –

4000 –
3000 –
2000 –
1500 –
1000 –

500 –

200 –
89 nt RNA →

50 –   ← Markers