High-performance multiplexed fluorescence in situ hybridization in culture and tissue with matrix imprinting and clearing

Jeffrey R. Moffitt[a,b,1], Junjie Hao[a,b,1], Dhananjay Bambah-Mukku[a,c], Tian Lu[a,b], Catherine Dulac[a,c], and Xiaowei Zhuang[a,b,d,2]

[a]Howard Hughes Medical Institute, Harvard University, Cambridge, MA 02138; [b]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138; [c]Department of Molecular and Cellular Biology, Harvard University, Cambridge, MA 02138; and [d]Department of Physics, Harvard University, Cambridge, MA 02138

Contributed by Xiaowei Zhuang, October 25, 2016 (sent for review October 10, 2016; reviewed by Gaudenz Danuser and Taekjip Ha)

Highly multiplexed single-molecule FISH has emerged as a promising approach to spatially resolved single-cell transcriptomics because of its ability to directly image and profile numerous RNA species in their native cellular context. However, background—from off-target binding of FISH probes and cellular autofluorescence—can become limiting in a number of important applications, such as increasing the degree of multiplexing, imaging shorter RNAs, and imaging tissue samples. Here, we developed a sample clearing approach for FISH measurements. We identified off-target binding of FISH probes to cellular components other than RNA, such as proteins, as a major source of background. To remove this source of background, we embedded samples in polyacrylamide, anchored RNAs to this polyacrylamide matrix, and cleared cellular proteins and lipids, which are also sources of autofluorescence. To demonstrate the efficacy of this approach, we measured the copy number of 130 RNA species in cleared samples using multiplexed error-robust FISH (MERFISH). We observed a reduction both in the background because of off-target probe binding and in the cellular autofluorescence without detectable loss in RNA. This process led to an improved detection efficiency and detection limit of MERFISH, and an increased measurement throughput via extension of MERFISH into four color channels. We further demonstrated MERFISH measurements of complex tissue samples from the mouse brain using this matrix-imprinting and -clearing approach. We envision that this method will improve the performance of a wide range of in situ hybridization-based techniques in both cell culture and tissues.

tissue clearing | fluorescence in situ hybridization | multiplexed imaging | single-cell transcriptomics | brain

Single-molecule FISH (smFISH) is a powerful technique that allows the direct imaging of individual RNA molecules within single cells (1, 2). In this approach, RNAs are labeled via the hybridization of fluorescently labeled oligonucleotide probes, producing bright fluorescent spots for single RNA molecules, which reveal both the abundance and the spatial distribution of these RNAs inside cells (1, 2). The ability of smFISH to image gene expression at the single-cell level in both cell culture and tissue has led to exciting advances in our understanding of the natural noise in gene expression and its role in cellular response (3, 4), the intracellular spatial organization of RNAs and its role in posttranscriptional regulation (5, 6), and the spatial variation in gene expression within complex tissues and its role in the molecular definition of cell types and tissue functions (6, 7).

To extend the benefits of this technique to systems-level questions and high-throughput gene-expression profiling, approaches to increase the multiplexing of smFISH (i.e., the number of different RNA species that can be simultaneously quantified within the same cell) have been developed (8–13). Most of these approaches take advantage of color multiplexing, which has allowed a few tens of RNA species to be imaged simultaneously. We have recently introduced multiplexed error-robust FISH (MERFISH), a massively multiplexed form of smFISH that allows RNA imaging and profiling at the transcriptomic scale (13, 14). MERFISH achieves this level of multiplexing by assigning error-robust barcodes to individual RNA species, labeling RNAs combinatorically with oligonucleotide probes that contain a representation of these barcodes, and reading out these barcodes through sequential rounds of single-color (13) or multicolor (14) smFISH imaging (Fig. S1). Using this approach, we have imaged 140 and 1,000 RNA species in individual cells with two different encoding schemes, one of which allows both error detection and correction and the other that allows error detection (13), and the number of addressable RNA species can be further changed by using different encoding schemes. Recently, we have increased the measurement throughput of MERFISH and demonstrated the ability to profile gene expression in tens of thousands of cells in a single day-long measurement (14). A different multiplexed smFISH method using color-based barcodes and sequential imaging (seqFISH) has been independently proposed and initially demonstrated with measurements of 12 RNA species in individual cells (12). While this current paper was in review, a paper reporting an extension of seqFISH that adds the error-correction capability and demonstrates the capability of imaging 125 or 250 RNA species was published (15).

Significance

Multiplexed single-molecule FISH allows spatially resolved gene-expression profiling in single cells. However, because of off-target binding of FISH probes and cellular autofluorescence, background can become limiting in multiplexed single-molecule FISH measurements, especially when tissue samples are imaged or when the degree of multiplexing is increased. Here we report a sample clearing approach for FISH that substantially reduced these background sources by anchoring RNAs to a polymer matrix and then removing proteins and lipids. This approach allows measurements with higher detection efficiency and sensitivity across more color channels in both cell culture and tissue with no detectable loss in RNA. We anticipate that this clearing approach will greatly facilitate applications of multiplexed FISH measurements in a wide variety of biological systems.

Author contributions: J.R.M., J.H., D.B.-M., T.L., C.D., and X.Z. designed research; J.R.M., J.H., D.B.-M., and T.L. performed research; J.R.M., J.H., and T.L. analyzed data; and J.R.M., J.H., D.B.-M., T.L., C.D., and X.Z. wrote the paper.

Reviewers: G.D., University of Texas Southwestern Medical Center; and T.H., The Johns Hopkins University.

Conflict of interest statement: X.Z., J.R.M., J.H., and T.L. are inventors on patents applied for by Harvard University that cover the multiplexed error-robust FISH and matrix-imprinting–based clearing methods.

Freely available online through the PNAS open access option.

[1]J.R.M. and J.H. contributed equally to this work.

[2]To whom correspondence should be addressed. Email: zhuang@chemistry.harvard.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1617699113/-/DCSupplemental.

smFISH measurements typically benefit from high signal-to-background ratios, resulting in the detection of individual RNA molecules with high accuracy and detection efficiency (1, 2): In many cases, the bright fluorescent signals that arise from the tens of fluorescently labeled probes bound to each copy of an RNA substantially exceed the background that arises from probes binding off target or from cellular autofluorescence. However, we have observed that as the degree of multiplexing is increased, the background level also tends to increase. The resulting decrease in the signal-to-background ratio makes a number of important applications and extensions of multiplexed smFISH challenging. For example, efforts to further increase the degree of multiplexing, to thousands or potentially tens of thousands of RNAs, will likely be limited by increased background. In addition, many RNAs are not long enough to accommodate tens of oligonucleotide probes, limiting the ability to measure relatively short RNAs and to discriminate many different RNA isoforms. Finally, background is typically more pronounced in complex tissues, making multiplexed smFISH measurements in tissues more challenging.

Here we report a sample clearing approach aimed at improving the signal-to-background ratio in RNA FISH measurements by substantially reducing background fluorescence signal. Many modern tissue-clearing approaches are designed to preserve the protein content of the sample while reducing scattering and autofluorescence background by extracting lipids and matching refractive index (16–23). For example, embedding and cross-linking tissues to hydrogels provides a powerful approach to tissue clearing, minimizing sample distortion during lipid removal and index matching while maintaining the protein content of the sample (19, 20). These approaches have also been made compatible with RNA FISH by stabilizing RNA molecules, for example through cross-linking of RNAs to proteins, without removing the protein content of the cell (20, 24). Here we show that a major source of background in RNA FISH measurements is the nonspecific binding of FISH probes to cellular components other than RNAs, such as proteins. For this reason, a clearing method that preserves RNAs while removing proteins and lipids is desired for RNA FISH imaging. In the recently developed expansion microscopy method, proteins (25) and, more recently, RNAs (26, 27) are physically anchored to a solvent-expandable and clearable poly-electrolyte matrix, effectively imprinting signals of these components on this matrix and allowing these molecular signals to be expanded along with the matrix for increasing image resolution. Inspired by this approach, we anchored RNA molecules to a nonswellable polyacrylamide (PA) matrix and then removed unwanted, non-RNA components, such as proteins and lipids, with the aim to remove their contribution to background fluorescence. We demonstrated that this matrix-imprinting and -clearing approach substantially reduced the background because of off-target binding of FISH probes and cellular autofluorescence. By comparing the copy number of 130 RNAs measured via MERFISH in uncleared and cleared cultures of human cells, we demonstrated that this matrix-imprinting–based clearing approach improves the detection efficiency and detection limit of MERFISH with no detectable loss in RNAs. Moreover, the reduction in autofluorescence, in particular in the blue–green spectral range, facilitated extension of MERFISH imaging from two to four distinct color channels with no reduction in performance. This improvement reduced the number of hybridization rounds needed for MERFISH measurements, which should further increase the MERFISH measurement speed and throughput. Finally, we demonstrated that this clearing approach substantially reduces the background in tissue, facilitating high-performance MERFISH measurements in cryosections of adult mouse brain tissues. Given the simplicity and efficacy of this matrix-imprinting–based clearing method, we envision that this approach could be used to substantially improve the performance of a wide range of in situ hybridization methods for both RNA and DNA in cell cultures and tissues.

Results

A Matrix-Imprinting and -Clearing Approach to Reduce the Background for smFISH Measurements. Our first step in the development of a sample clearing method for smFISH was to determine the physical origin of off-target binding of oligonucleotide probes: are these probes binding to the incorrect RNA or other cellular components, such as proteins or lipids? To address this question, we stained human lung fibroblast (IMR-90) cells using FISH probes targeting the Filamin A (FLNA) mRNA. As expected, we observed both bright fluorescence spots marking individual molecules of FLNA mRNA (Fig. 1*A*, *Left*) and a diffuse background because of off-target probe binding (Fig. 1*A*, *Center*) that was not present in samples not stained with FISH probes (Fig. S2). We then measured the RNase sensitivity of both the foreground RNA spots and the diffuse background, reasoning that if the background arose from off-target binding to incorrect RNAs, both the foreground spots and the background should be RNase-sensitive. We found that a brief RNase A treatment completely removed the bright foreground spots, but produced little if any reduction in the background (Fig. 1*A*, *Right*). Thus, we conclude that the vast majority of off-targeting binding of smFISH probes arose from binding to cellular components other than RNA, such as proteins and lipids.

Hence, we reasoned that one way to reduce the background would be to remove the non-RNA components, such as proteins and lipids, from the sample. Moreover, because these components are also a major source of autofluorescence, the autofluorescence background might be reduced by such an approach as well. To this end, we fixed the sample and hybridized it with oligonucleotide probes as in standard smFISH (1, 2) or MERFISH measurements (13, 14), and then embedded the sample in an inert, nonfluorescence matrix to which RNA molecules were anchored, effectively imprinting the desired RNA signal onto this matrix (Fig. 1*B*). Once RNAs were anchored, cellular proteins and lipids were removed (for example, by digestion and extraction) without, in principle, affecting the number and localization of RNAs within the sample. smFISH probes bound

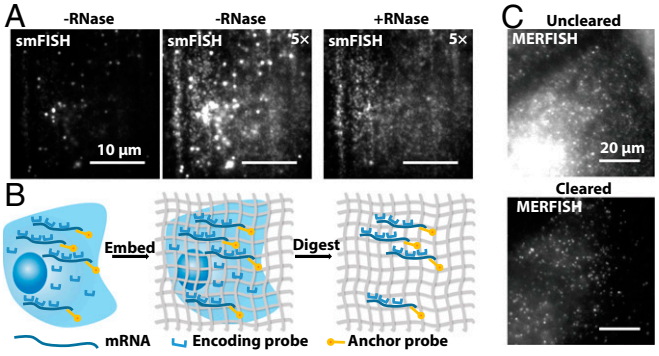

Fig. 1. Matrix imprinting and clearing reduces background in smFISH measurements. (*A*) A human fibroblast cell (IMR-90) stained with smFISH probes targeting the FLNA mRNA before (*Left* and *Center*) and after (*Right*) treatment with RNase A. The contrast of the *Center* and *Right* panels has been increased fivefold from that of the *Left* panel to better visualize the background from probes bound off-target. (Scale bars, 10 μm.) (*B*) Schematic diagram of a matrix-imprinting and -clearing approach to reduce background in smFISH measurements. Cells are stained with smFISH probes or encoding probes for MERFISH measurements, and a poly-dT anchor probe, which targets the polyA tail of mRNAs. Cells are then embedded in a PA matrix, to which the poly-dT anchor probes are covalently linked via a terminal acrydite moiety. Proteins and lipids are then digested and extracted, freeing off-target bound smFISH probes to diffuse out of the PA matrix and removing cellular components that contribute to autofluorescence. (*C*) U-2 OS cells labeled with MERFISH-encoding probes targeting 130 RNAs followed by staining with a readout probe conjugated to Cy5 that binds to the encoding probes in an uncleared sample (*Upper*) and a sample treated with the matrix-imprinting and -clearing protocol (*Lower*). (Scale bars, 20 μm.)

SYSTEMS BIOLOGY

off-target to these components should then be free to diffuse from the matrix. We used PA as the inert matrix and a 15-nt poly-dT oligonucleotide to bind and anchor polyadenylated (polyA) RNAs to the PA matrix, although other methods may also be used to link RNAs to the matrix. This anchor probe was comprised of 50% locked-nucleic acid bases to stabilize the hybridization to polyA tails of the RNAs (28) and additionally contained a terminal acrydite moeity, which can be covalently incorporated into the PA matrix as it polymerizes.

To test whether this clearing approach led to a reduction in off-target binding, we first measured the efficacy of protein and lipid removal and observed that this protocol efficiently removed cellular proteins and lipids from embedded cultured human osteosarcoma (U-2 OS) cells (Fig. S3). Next, we performed labeling as in MERFISH experiments and tested whether off-target probe binding was indeed reduced by clearing. In a MERFISH measurement, we typically stain cells first with a complex library of "encoding" oligonucleotide probes (13, 14). These encoding probes are not themselves fluorescently labeled. Instead, each encoding probe contains a targeting sequence that directs its binding to a cellular RNA and multiple readout sequences. Multiple encoding probes are targeted to each RNA, and the set of readout sequences contained within these encoding probes form a specific barcode that is unique to that RNA species (Fig. S1 *B* and *C*). These barcodes are then measured in a series of hybridizations, each round of hybridization using either one fluorescently labeled "readout" probe complementary to a specific readout sequence, reading out one bit per round in the single-color imaging mode (13), or multiple readout probes labeled with spectrally distinct dyes, reading out multiple bits simultaneously in the multicolor imaging mode (14) (Fig. S1 *B* and *C*). One advantage of this two-step labeling approach with encoding hybridization followed by readout hybridization is that it substantially reduces the time required for each hybridization round, because hybridization of the readout probes to encoding probes (including all fluid handling and sample washing) requires <30 min (13, 14) compared with the overnight hybridization typically required for direct hybridization of FISH probes to cellular RNAs, because the readout sequences on the encoding probes do not form a secondary structure and are not occluded by cellular proteins.

To demonstrate the clearing efficacy, we stained U-2 OS cells with encoding probes used for a MERFISH measurement of 130 RNAs at a total concentration of 300 μM, which is threefold higher than typically used in our MERFISH experiments, to generate high background. We then embedded and cleared the sample in the PA matrix as described above, and stained the RNA-imprinted matrix with a readout probe labeled with a Cy5 dye. Fig. 1*C* shows that the cleared samples contained visible smFISH spots but substantially lower background than uncleared samples, demonstrating that this approach indeed reduced the background as a result of off-target probe binding.

As an aside, MERFISH measurements require repeated sample staining with a series of readout probes and, in cases where the FISH signal is removed by chemical cleavage of the fluorophores (14), the efficient removal of cleaved fluorophores. To facilitate the rapid penetration of readout probes as well as the rapid removal of cleaved dyes, we embedded samples in 50- to 100-μm-thick PA films. These films were thick enough to cover cultured cells or moderately sized tissue slices, yet thin enough that the rate of readout probe hybridization and the rate of dye cleavage/removal were not substantially changed from those observed in uncleared samples (Fig. S4).

RNA Is Preserved During Matrix Imprinting and Clearing. To determine if any RNAs were lost during matrix imprinting and clearing, we used MERFISH to determine the copy number of 130 RNAs in a cleared sample of U-2 OS cells and compared these numbers to that derived previously from an uncleared sample (14). We used our previously published 16-bit, modified Hamming-distance-4 (MHD4) code to encode RNAs (13). In this encoding scheme, all valid binary barcodes used to encode RNAs are separated by a Hamming distance of at least 4, which means that at least four bits must be read incorrectly to change one valid barcode to another, drastically reducing the probability of misidentifying RNAs. Furthermore, this scheme also allows us to correct single-bit errors because every single-bit error produces a barcode uniquely close to a single valid barcode. This specific MHD4 code contains 140 valid barcodes (13); we only used 130 of them to encode RNAs, leaving the remaining 10 barcodes to serve as "blank" controls to determine the rate of spurious RNA detection and estimate misidentification rates.

We performed this MERFISH measurement of these 130 RNA species, as described previously (14), using two-color imaging to read out 16 bits in 8 rounds of hybridization and imaging (two bits per round), as well as reductive cleavage of disulfide bonds to remove the fluorophores linked to the readout probes between consecutive rounds of smFISH imaging (Fig. S1*C*). Fig. 2*A* shows that individual RNA molecules could be clearly detected in each of the eight hybridization and imaging rounds, allowing their identity to be decoded. As described previously (14), we used the depletion of RNAs near cell boundaries to perform cell segmentation. Fig. 2*B* shows that the copy number per cell observed for these RNAs measured in the cleared sample correlated strongly with those measured in an uncleared sample with a Pearson correlation coefficient of 0.94 between the $\log_{10}$ copy numbers ($\rho_{10} = 0.94$ for the 116 RNA species whose measured copy numbers were larger than that observed for the largest blank barcode count). On average, the ratio between the copy numbers measured in the cleared sample to those measured in the uncleared sample was 1.12 ± 0.04 (SEM, n = 116 RNAs), and this ratio was largely independent of the length of the RNAs (Fig. 2*C*). Although we conservatively used only

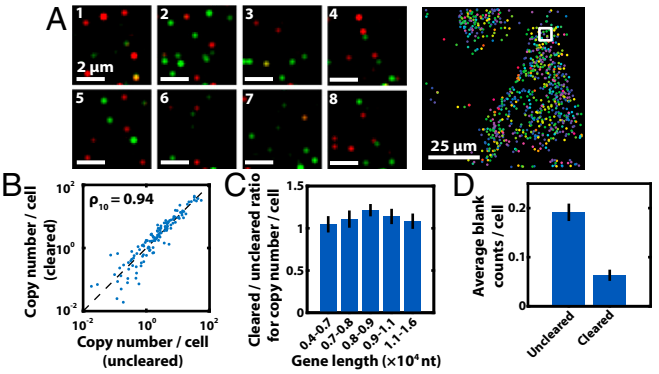

Fig. 2. Matrix imprinting and clearing improves MERFISH performance with no loss in RNA. (*A*, *Left*) Two-color smFISH images from each of the eight rounds of hybridization and imaging in a MERFISH measurement of 130 RNA species in matrix-imprinted and -cleared U-2 OS cells using readout probes labeled with Cy5 (green) or Alexa750 (red). Yellow represents the overlay between the two dyes. Only a small portion of the MERFISH imaging FOV is shown. (Scale bars, 2 μm.) (*Right*) All identified RNAs (colored markers) detected in a single FOV with the barcodes of the RNAs represented by the colors of the markers. The white box represents the portion of this FOV displayed on the *Left*. (Scale bar, 25 μm.) (*B*) The average copy numbers per cell observed for these RNA species in matrix-imprinted and -cleared U-2 OS cells versus the copy numbers obtained from previously published measurements in an uncleared sample (14). Copy numbers were corrected by subtracting the average copy number observed for the blank barcodes. Uncorrected copy numbers are displayed in Fig. S5*B*. The $\log_{10}$ counts correlate with a Pearson correlation coefficient of 0.94 (P value: 10^{-54}). The dashed line represents equality. (*C*) The average ratio of the copy number per cell for a sample that was matrix imprinted and cleared to that observed for an uncleared sample for RNAs within the specified RNA length range. Error bars represent SEM (n = 26 genes for each bin). (*D*) Average copy number per cell of the blank barcodes (i.e., barcodes not assigned to an RNA) in an uncleared sample and in a matrix-imprinted and -cleared sample. Error bars represent SEM (n = 10 blank barcodes).

those RNAs with copy numbers greater than the largest blank count for analysis here, the results were similar when all 130 RNAs were used.

These measurements showed that several aspects of MERFISH performance were improved with matrix imprinting and clearing. Previously, we observed a MERFISH detection efficiency of ∼90% (14); thus, a copy number ratio of ∼1.1 between the cleared and uncleared samples suggested that clearing increased the detection efficiency to near 100%. Second, we observed that the average frequency at which the blank barcodes were observed per cell in the cleared samples dropped substantially relative to that observed in the uncleared samples (Fig. 2*D*). The average level of blank barcode counts observed in the uncleared sample (Fig. 2*D*) was comparable to the observed copy number for the lowest abundance RNAs measured here, leading to the possibility that the copy number observed for these low abundance RNAs might have been biased by a background rate of spurious RNA counts in uncleared samples. Indeed, we observed an excess of these low-abundance RNAs in uncleared samples relative to that expected from bulk RNA-seq (Fig. S5*A*), whereas this bias was substantially reduced in cleared samples (Fig. S5). Thus, we conclude that the increased signal-to-background in cleared samples results in an improvement in both the detection efficiency and the detection limit in MERFISH measurements.

Four-Color MERFISH Imaging. In addition to providing a substantial decrease in the background because of off-target binding of FISH probes, the removal of proteins and lipids from the sample may also reduce the level of autofluorescence. To quantify this decrease, we measured the fluorescence of unlabeled U-2 OS cells in uncleared and cleared samples with four excitation wavelengths: 750, 647, 561, and 488 nm. Consistent with the expectation that cell autofluorescence is substantially higher in the blue–green spectral range than in the red range, the clearing protocol had little effect on the already low autofluorescence background in the 750- and 647-nm channels, but produced a substantial reduction in the autofluorescence observed in the 561- and 488-nm channels (Fig. 3*A*).

With this reduction in the autofluorescence, we explored the possibility of using all four excitation channels to read out four different bits of the 16-bit code simultaneously in each round of imaging during MERFISH measurements. We again stained U-2 OS cells with the same MERFISH-encoding probe set for the 130 RNA species as described above and performed MERFISH measurements in which each round of hybridization used four different readout probes, conjugated respectively to Alexa750, Cy5, ATTO565, or Alexa488 via a disulfide bond (Fig. 3*B* and Table S1). With such four-color imaging, the 16-bit MERFISH measurement only required four rounds of hybridization and imaging.

We then compared the measured copy numbers derived from this four-color measurement to those determined with a two-color (750 and 647 nm) measurement in the cleared sample. Fig. 3*C* demonstrates that these copy numbers correlated strongly with a ρ_{10} of 0.99 and had an average ratio of 1.01 ± 0.02 (SEM, $n = 109$ RNAs with copy numbers greater than that observed for the largest blank count). To confirm that imaging in the new color channels did not introduce additional error, we determined the "1" to "0" or "0" to "1" error rates per bit and found that these error rates did not vary substantially with the color channel (Fig. 3*D*).

Finally, to confirm that the improved performance that we observed with cleared samples was reproducible, we performed additional two-color and four-color MERFISH measurements in cleared samples. Fig. S6 shows that the copy numbers derived from all of these measurements correlated strongly (all ρ_{10} are 0.94 or greater). By comparing each of these datasets to the previously determined detection efficiency of MERFISH measurements in uncleared samples (14), we estimated an average MERFISH detection efficiency of 96 ± 7% (SEM, $n = 4$ replicates) for cleared samples. Furthermore, we observed an ∼fourfold reduction in the average rate of blank barcode detection: 0.08 ± 0.03 counts per cell (SEM, $n = 4$ replicates) for cleared samples versus 0.30 ± 0.07 counts per cell (SEM, $n = 7$ replicates) (14) for uncleared samples.

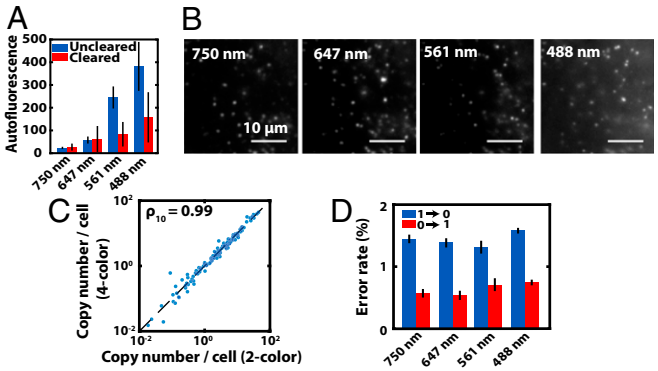

Fig. 3. Autofluorescence reduction by matrix imprinting and clearing facilitates four-color MERFISH. (*A*) The average autofluorescence observed for unstained U-2 OS cells before (blue) and after (red) matrix imprinting and clearing when excited with 750-, 647-, 561-, or 488-nm light. Error bars represent SEM ($n = 3$ replicates). (*B*) Images of cleared U-2 OS cells stained with MERFISH-encoding probes targeting 130 RNAs and the first four readout probes each conjugated to one of the following dyes: Alexa750, Cy5, ATTO565, or Alexa488. Samples were imaged with excitation light listed in *A*. (Scale bars, 10 μm.) (*C*) Average copy number per cell determined via four-color MERFISH to that determined with two-color MERFISH, both in cleared samples. The copy numbers have been corrected by subtracting the average rate of blank barcode detection as in Fig. 2*B*. The dashed line represents equality. The Pearson correlation coefficient between the $\log_{10}$ abundances is 0.99 (*P* value: 10^{-98}). (*D*) The average rate of observing a "1" to "0" error (blue) or a "0" to "1" error (red) per bit for bits that are read out with each of the four color channels, as indicated by the excitation wavelength. Each error rate ("1" to "0" or "0" to "1") was calculated for each individual bit using the frequency at which errors were corrected at that bit, as described previously (13), and then these per bit error rates were averaged over the bits that were detected in the same color channel (Table S1). Error bars represent SEM ($n = 4$ bits read out with each color channel).

MERFISH Measurements of Brain Tissue. To explore whether clearing can overcome the increased background that we have observed in tissues, we performed MERFISH measurements of 130 RNA species on four cryosections taken from adult mouse hypothalamus, each ∼2 mm × 2 mm wide and 10 μm thick (Fig. 4 *A* and *B*). We performed 3D imaging with seven ∼1.5-μm-thick optical sections measured per field-of-view (FOV). These RNAs were again encoded with the 16-bit MHD4 code and read out with eight rounds of hybridization using two-color imaging per round. These samples were matrix-imprinted and cleared as described above but with the addition of a brief treatment with 4% (wt/vol) SDS before PA embedding, which further improved tissue clearing. Fig. 4 *C* and *D* illustrate that this matrix-imprinting and -clearing approach substantially reduced the background observed in these tissue slices. smFISH spots representing individual RNA molecules were clearly observable in the cleared sample in each round of imaging with low "1" to "0" or "0" to "1" error rates, similar to those detected in cultured cells, allowing individual RNAs to be decoded (Fig. 4 *E*–*G*).

We compared the average RNA density determined via MERFISH from these four tissue slices with the abundance determined via bulk RNA-seq data derived from the same region of the hypothalamus (29), and observed a strong correlation ($\rho_{10} = 0.84$) between our MERFISH results and the bulk RNA-seq results (Fig. 4*H*). At the very low abundance range corresponding to those RNAs that are expressed poorly in the hypothalamus [<0.5 fragments per kilobase of transcript per million (FPKM)], the correlation between MERFISH and bulk RNA-seq results reduced substantially, suggesting that the abundance of these RNAs was

SYSTEMS BIOLOGY

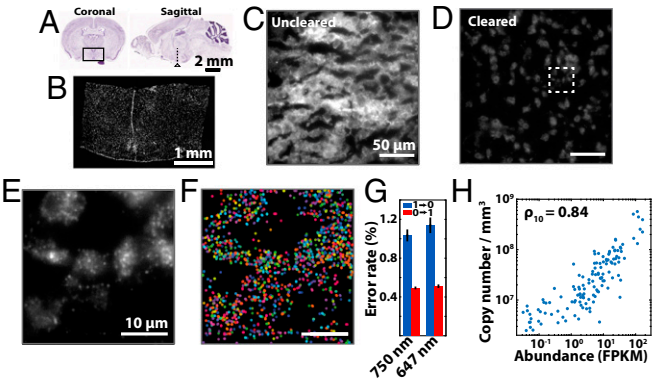

Fig. 4. MERFISH measurements of adult mouse brain tissue. (*A*) Nissl-stained images of coronal and sagittal slices of an adult mouse brain taken from the Allen brain atlas (42). The black box and dashed line represent the region of the mouse hypothalamus studied. (Scale bar, 2 mm.) (*B*) Image of a single, 10-μm-thick cryosection of the mouse hypothalamus stained with DAPI. The complete volume of the central 2-mm × 2-mm region of this slice was imaged with MERFISH using seven 1.5-μm-thick optical sections per FOV. (Scale bar, 1 mm.) (*C* and *D*) Images of a small portion of a mouse hypothalamus slice stained with an encoding probe set used for a MERFISH measurement of 130 RNAs and a readout probe conjugated to Cy5. (*C*) Single optical section image of an uncleared sample. (*D*) Single optical section image of a matrix-imprinted and -cleared sample. (Scale bars, 50 μm.) (*E*) Zoom-in of the region of *D* marked with the white dashed box. (*F*) Decoded RNAs (different colors represent different barcodes) from all seven optical sections of the region shown in *E*. Not all RNA molecules shown in *F* are observed in *E* because *E* represents only one of the seven optical sections and one of the 16 bits. (Scale bars, 10 μm.) (*G*) The average rate of observing a "1" to "0" error (blue) or a "0" to "1" error (red) per bit for bits that are read out with each of the two color channels, as indicated by the excitation wavelength. Error rates were calculated as in Fig. 3*D*. Error bars represent SEM ($n = 8$ bits read out with each color channel). (*H*) The density of 130 RNA species as determined via MERFISH versus the abundance as determined via bulk RNA-seq for the region of the mouse hypothalamus shown in *A*. The Pearson correlation coefficient between the $\log_{10}$ abundances is 0.84 (*P* value: 10^{-35}).

near or below our current detection limit, a conclusion supported by the similarity between the copy numbers of these RNAs and the average copy number observed for the blank barcodes [$6 \times 10^6 \pm 2 \times 10^6$/mm^3 (SEM, $n = 10$ blank barcodes)].

Discussion

Massively multiplexed smFISH allows spatially resolved gene-expression profiling within single cells. However, a number of important applications of and advances to this approach are limited by the fluorescence background encountered in these experiments. Here we described a clearing approach that substantially reduced several background sources in RNA FISH measurements by effectively imprinting the desired RNA signal onto an inert, nonfluorescent, PA matrix and then removing unwanted cellular components that give rise to background because of off-target probe binding and autofluorescence. The reduction in fluorescence background provided by this approach led to improvement in both the detection efficiency and the detection limit in MERFISH measurements. Moreover, this matrix-imprinting and -clearing approach produced a substantial reduction in the background observed for measurements in tissue samples, allowing high-performance MERFISH measurements in brain tissue sections, which should facilitate spatial mapping of distinct cell types in the brain.

This matrix-imprinting and -clearing approach complements several existing methods that have been used to improve the signal or reduce nonspecific binding background for RNA detection in FISH experiments (15, 24, 26, 30, 31) or in situ sequencing experiments (32, 33). For example, signal amplification techniques, such as branched DNA (34), rolling-circle amplification (35, 36), and hybridization chain reaction (37) have been used to increase the signal associated with each RNA molecule (15, 24, 26, 30–34), thereby increasing the ratio between RNA signals and autofluorescence background, although background as a result of off-target probe binding may also be amplified concurrently by these approaches. Alternatively, proximity-dependent approaches (38, 39), in which a fluorescent signal is only produced when two separate probes or two ends of the same probe are in close proximity, have been used to reduce the effect of nonspecific probe binding (30, 32, 36, 38, 39). Because matrix imprinting and clearing is compatible with each of these approaches, it may be combined with these techniques to further improve the performance of both conventional and highly multiplexed smFISH measurements.

In addition, the reduction in autofluorescence in the blue and green color channels provided by matrix imprinting and clearing allowed us to extend MERFISH measurements from two colors to four colors with no loss in performance, which should substantially increase the measurement speed and throughput of MERFISH. We previously reported a high-throughput MERFISH imaging platform and demonstrated the ability to profile 140 RNA species in 40,000 human cells in ∼18 h using a 16-bit MHD4 code, two-color imaging, and eight hybridization rounds (14). This platform also allowed us to profile a similar number of RNAs across an ∼16-mm^2 × 10-μm tissue volume in ∼18 h here. This throughput was facilitated by our ability to rapidly bind FISH probes (through hybridization of readout probes to encoding probes, which requires <30 min) and then rapidly remove FISH signals (through the use of chemical cleavage to remove fluorophores from readout probes, which requires <20 min) in each round of hybridization. With the increase to four color channels, we can now read out the 16-bit code in just four rounds of hybridization with four bits detected per round, further reducing the measurement durations and allowing tens of thousands of cells to be profiled in <10 h with MERFISH. Similarly, we also anticipate that our previously demonstrated MERFISH measurement of 1,001 RNA species using a 14-bit MHD2 encoding scheme (13) can also be completed in just four hybridization rounds. While this current paper was in review, another paper demonstrated the profiling of >100 RNA species in tissue using a different multiplexed smFISH method (seqFISH) with a similar number of hybridization rounds: for example, 125 RNA species could be measured with a four-letter, five-color error-correcting barcoding scheme using four rounds of hybridization, but the slower approaches used there for probe hybridization (overnight hybridization of FISH probes directly to cellular RNA) and signal removal (4 h of DNase I digestion of FISH probes) in each round made the sample processing and measurement time per round longer (15).

Finally, we envision that the substantial reduction in background provided by this matrix-imprinting and -clearing approach will facilitate several additional extensions of MERFISH. First, an increase in the degree of multiplexing—to the simultaneous measurement of several thousand or tens of thousands of RNAs—would likely require substantially higher encoding probe concentrations than are currently used and, thus, would benefit from the much lower off-target probe binding achieved here in cleared samples. Second, we have performed MERFISH using tens of encoding probed per RNA [e.g., 92 encoding probes for RNAs that are 3 kb or longer (13, 14)]. With the dramatic decrease in background enabled by this clearing approach, it should be possible to detect RNAs that are much shorter, potentially with as few as <10 encoding probes per RNA. This advance would facilitate the detection of relatively short messenger and long-noncoding RNAs, and possibly some small RNAs. The ability to detect RNA molecules with relatively few FISH probes will also substantially improve the ability to distinguish RNA isoforms. Third, the combination of expansion microscopy (25) with MERFISH may be facilitated by a common matrix-imprinting approach, which may help RNA profiling in RNA-dense regions of cells and further increase in the

degree of multiplexing. We also anticipate that the alternative RNA anchoring approach reported for expansion microscopy (26)—in which RNAs are alkylated with a cross-linker that is covalently incorporated into the PA gel—may be used for background reduction by matrix imprinting and clearing as well. Finally, whereas our current implementation of the matrix-imprinting and -clearing approach removes cellular proteins and, thus, information regarding the protein content of the sample, this information could be restored by labeling samples with antibodies conjugated to oligonucleotides before embedding. These oligonucleotides could then be anchored to the matrix, followed by digestion of the antibodies along with cellular proteins, allowing the original location of the antibodies to be determined via FISH imaging of these oligonucleotides (25). Such labeling approaches could be used to label cell boundary markers to facilitate cell segmentation in cases where a depletion of RNA at the edge of the cell is not sufficient to identify cell boundaries. Moreover, such approaches can convert protein identities into oligonucleotide signals, which may be used to perform highly multiplexed proteomic and transcriptomic measurements simultaneously in single cells.

Materials and Methods

Detailed descriptions of all protocols are provided in the *SI Materials and Methods*. All animal experiments were approved by the Institutional Animal Care and Use Committee of Harvard University.

Human osteosarcoma cells (U-S OS, American Type Culture Collection) and Human fibroblasts (IMR-90, American Type Culture Collection) were cultured, fixed, permeabilized, and stained with smFISH probes or MERFISH-encoding probes, as described previously (13, 14, 40). Mouse hypothalamus tissue was freshly frozen, cryosectioned into 10-μm-thick slices, postfixed onto coverslips, cleared with 4% (wt/vol) SDS, permeabilized with 70% (vol/vol) ethanol, and then stained with encoding probes. Cells or tissue samples were embedded in a 4% (vol/vol) solution of a 19:1 ratio of acrylamide to bis-acrylamide containing 50 mM Tris HCl (pH 8), 300 mM NaCl, 0.03% (wt/vol) ammonium persulfate, and 0.15% (vol/vol) *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (TEMED). Protein and lipids were removed with a >12 h, 37 °C digestion with proteinase K in 0.8 M guanidine HCl, 0.5% (vol/vol) Triton-X 100, 50 mM Tris pH 8, and 1 mM EDTA.

MERFISH measurements with U-2 OS cells were performed with a published encoding probe set (14). The encoding probe set for measurements in mouse brain tissue was designed as described previously (14). Readout probes were purchased from Biosynthesis and are described in Table S1. Encoding probes were constructed by modifying the Oligopaints approach (41) with a high-yield enzymatic amplification protocol and a high-speed probe design algorithm (13, 14, 40). Samples were imaged on custom setups with readout hybridization, buffer exchange, and reductive cleavage procedures detailed in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank Hazen Babcock, Guiping Wang, and Chenglong Xia for helpful discussions and assistance in experiments. This work was supported in part by NIH Grants R01HD082131 (to C.D.) and R01MH113094 and R01MH111502 (to C.D. and X.Z.). C.D. and X.Z. are Howard Hughes Medical Institute investigators.

1. Femino AM, Fay FS, Fogarty K, Singer RH (1998) Visualization of single RNA transcripts in situ. *Science* 280(5363):585–590.
2. Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S (2008) Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* 5(10):877–879.
3. Raj A, van Oudenaarden A (2008) Nature, nurture, or chance: Stochastic gene expression and its consequences. *Cell* 135(2):216–226.
4. Sanchez A, Golding I (2013) Genetic determinants and cellular constraints in noisy gene expression. *Science* 342(6163):1188–1193.
5. Holt CE, Schuman EM (2013) The central dogma decentralized: New perspectives on RNA function and local translation in neurons. *Neuron* 80(3):648–657.
6. Buxbaum AR, Haimovich G, Singer RH (2015) In the right place at the right time: Visualizing and understanding mRNA localization. *Nat Rev Mol Cell Biol* 16(2):95–109.
7. Crosetto N, Bienko M, van Oudenaarden A (2015) Spatially resolved transcriptomics and beyond. *Nat Rev Genet* 16(1):57–66.
8. Levsky JM, Shenoy SM, Pezo RC, Singer RH (2002) Single-cell gene expression profiling. *Science* 297(5582):836–840.
9. Lubeck E, Cai L (2012) Single-cell systems biology by super-resolution imaging and combinatorial labeling. *Nat Methods* 9(7):743–748.
10. Levesque MJ, Raj A (2013) Single-chromosome transcriptional profiling reveals chromosomal gene expression regulation. *Nat Methods* 10(3):246–248.
11. Jakt LM, Moriwaki S, Nishikawa S (2013) A continuum of transcriptional identities visualized by combinatorial fluorescent in situ hybridization. *Development* 140(1):216–225.
12. Lubeck E, Coskun AF, Zhiyentayev T, Ahmad M, Cai L (2014) Single-cell in situ RNA profiling by sequential hybridization. *Nat Methods* 11(4):360–361.
13. Chen KH, Boettiger AN, Moffitt JR, Wang S, Zhuang X (2015) RNA imaging. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* 348(6233):aaa6090.
14. Moffitt JR, et al. (2016) High-throughput single-cell gene-expression profiling with multiplexed error-robust fluorescence in situ hybridization. *Proc Natl Acad Sci USA* 113(39):11046–11051.
15. Shah S, Lubeck E, Zhou W, Cai L (2016) In situ transcription profiling of single cells reveals spatial organization of cells in the mouse hippocampus. *Neuron* 92(2):342–357.
16. Dodt H-U, et al. (2007) Ultramicroscopy: Three-dimensional visualization of neuronal networks in the whole mouse brain. *Nat Methods* 4(4):331–336.
17. Hama H, et al. (2011) Scale: A chemical approach for fluorescence imaging and reconstruction of transparent mouse brain. *Nat Neurosci* 14(11):1481–1488.
18. Ertürk A, et al. (2011) Three-dimensional imaging of the unsectioned adult spinal cord to assess axon regeneration and glial responses after injury. *Nat Med* 18(1):166–171.
19. Chung K, et al. (2013) Structural and molecular interrogation of intact biological systems. *Nature* 497(7449):332–337.
20. Yang B, et al. (2014) Single-cell phenotyping within transparent intact tissue through whole-body clearing. *Cell* 158(4):945–958.
21. Susaki EA, et al. (2014) Whole-brain imaging with single-cell resolution using chemical cocktails and computational analysis. *Cell* 157(3):726–739.
22. Renier N, et al. (2014) iDISCO: A simple, rapid method to immunolabel large tissue samples for volume imaging. *Cell* 159(4):896–910.
23. Richardson DS, Lichtman JW (2015) Clarifying tissue clearing. *Cell* 162(2):246–257.
24. Sylwestrak EL, Rajasethupathy P, Wright MA, Jaffe A, Deisseroth K (2016) Multiplexed intact-tissue transcriptional analysis at cellular resolution. *Cell* 164(4):792–804.
25. Chen F, Tillberg PW, Boyden ES (2015) Optical imaging. Expansion microscopy. *Science* 347(6221):543–548.
26. Chen F, et al. (2016) Nanoscale imaging of RNA with expansion microscopy. *Nat Methods* 13(8):679–684.
27. Tsanov N, et al. (September 5, 2016) smiFISH and FISH-quant—A flexible single RNA detection approach with super-resolution capability. *Nucleic Acids Res*, 10.1093/nar/gkw784.
28. Jacobsen N, et al. (2004) Direct isolation of poly(A)+ RNA from 4 M guanidine thiocyanate-lysed cell extracts using locked nucleic acid-oligo(T) capture. *Nucleic Acids Res* 32(7):e64.
29. Gregg C, et al. (2010) High-resolution analysis of parent-of-origin allelic expression in the mouse brain. *Science* 329(5992):643–648.
30. Battich N, Stoeger T, Pelkmans L (2013) Image-based transcriptomics in thousands of single human cells at single-molecule resolution. *Nat Methods* 10(11):1127–1133.
31. Shah S, et al. (2016) Single-molecule RNA detection at depth by hybridization chain reaction and tissue hydrogel embedding and clearing. *Development* 143(15):2862–2867.
32. Ke R, et al. (2013) In situ sequencing for RNA analysis in preserved tissue and cells. *Nat Methods* 10(9):857–860.
33. Lee JH, et al. (2014) Highly multiplexed subcellular RNA sequencing in situ. *Science* 343(6177):1360–1363.
34. Kern D, et al. (1996) An enhanced-sensitivity branched-DNA assay for quantification of human immunodeficiency virus type 1 RNA in plasma. *J Clin Microbiol* 34(12):3196–3202.
35. Zhong XB, Lizardi PM, Huang XH, Bray-Ward PL, Ward DC (2001) Visualization of oligonucleotide probes and point mutations in interphase nuclei and DNA fibers using rolling circle DNA amplification. *Proc Natl Acad Sci USA* 98(7):3940–3945.
36. Larsson C, et al. (2004) In situ genotyping individual DNA molecules by target-primed rolling-circle amplification of padlock probes. *Nat Methods* 1(3):227–232.
37. Dirks RM, Pierce NA (2004) Triggered amplification by hybridization chain reaction. *Proc Natl Acad Sci USA* 101(43):15275–15278.
38. Nilsson M, et al. (1994) Padlock probes: Circularizing oligonucleotides for localized DNA detection. *Science* 265(5181):2085–2088.
39. Fredriksson S, et al. (2002) Protein detection using proximity-dependent DNA ligation assays. *Nat Biotechnol* 20(5):473–477.
40. Moffitt JR, Zhuang X (2016) RNA imaging with multiplexed error-robust fluorescence in situ hybridization (MERFISH). *Methods Enzymol* 572:1–49.
41. Beliveau BJ, et al. (2012) Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc Natl Acad Sci USA* 109(52):21301–21306.
42. Lein ES, et al. (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445(7124):168–176.
43. Buxboim A, Rajagopal K, Brown AEX, Discher DE (2010) How deeply cells feel: Methods for thin gels. *J Phys Condens Matter* 22(19):194116.
44. Rasnik I, McKinney SA, Ha T (2006) Nonblinking and long-lasting single-molecule fluorescence imaging. *Nat Methods* 3(11):891–893.
45. Babcock H, Sigal YM, Zhuang X (2012) A high-density 3D localization algorithm for stochastic optical reconstruction microscopy. *Opt Nanoscopy* 1(6):6.
46. Walz S, et al. (2014) Activation and repression by oncogenic MYC shape tumour-specific gene expression profiles. *Nature* 511(7510):483–487.