

1 **Autonomous combinatorial color barcoding for multiplexing**  
2 **single molecule RNA visualization**

3 Yong-Sheng Cheng<sup>1\*</sup>, Yue Zhuo<sup>1</sup>, Katharina Hartmann<sup>1</sup>, Peng Zou<sup>1</sup>,  
4 Gözde Bekki<sup>1</sup>, Heike Alter<sup>1</sup> & Hai-Kun Liu<sup>1\*</sup>

5  
6 <sup>1</sup>Division of Molecular Neurogenetics, German Cancer Research Center (DKFZ),  
7 DKFZ-ZMBH Alliance, Im Neuenheimer Feld 280, Heidelberg, 69120, Germany.

8 \*Correspondence: [yongsheng.cheng@dkfz.de](mailto:yongsheng.cheng@dkfz.de) and [l.haikun@dkfz.de](mailto:l.haikun@dkfz.de)

9  
10 **Abstract**

11  
12 Single molecular fluorescence in situ hybridization (smFISH) detects RNA  
13 transcripts with spatial information and digital molecular counting. However,  
14 the broad usage of smFISH is still hindered by the complex chemical probe  
15 conjugation or microscopy set-up, especially for investigating multiple gene  
16 expression. Here we present a multiple fluorophore enzymatical labeling  
17 method (termed HuluFISH) for smFISH probes to achieve flexible  
18 combinatorial color barcoding in single hybridization step. The multiplex  
19 capacity of HuluFISH follows an exponential growth with the increase of the  
20 number of fluorophore types. We demonstrate that this method can be used  
21 to detect cellular heterogeneity in embryonic mouse brain on single cell level.

22  
23  
24  
25  
26  
27  
28

## 29 Introduction

30

31 Since the invention of *in situ* hybridization<sup>1</sup>, it has been continuously advancing  
32 our understanding of gene expression with spatial information. The smFISH  
33 technology pioneered by Robert Singer's lab<sup>2</sup> and further developed by Raj  
34 et al.<sup>3</sup>, brings *in situ* RNA quantification into single molecular and digital  
35 manner. Nevertheless, the limited choices of single fluorophore on probes  
36 cannot cope with the increasing demand of simultaneous multiple gene  
37 detection. Although sequential hybridization has been employed to achieve  
38 multiplex gene detection using smFISH<sup>4,5</sup>, sophisticated experimental  
39 settings hinder its broad applications in the biomedical community. One  
40 alternative strategy for increasing the multiplexity beyond fluorophore limit is  
41 using combinatorial color barcoding via spectral or spatial separated groups  
42 of smFISH probe<sup>6,7</sup>. Current combinatorial barcoding either needs a long gene  
43 target for mRNA<sup>6</sup> or only targets intrinsically non-stable introns<sup>7</sup>, and thus  
44 restricts their applications from detecting the majority of the transcripts (median  
45 mouse coding DNA sequence (CDS) length is 1026 bp). Therefore, a  
46 combinatorial color barcoding on individual smFISH probe will empower the  
47 conventional smFISH with massive color combinations with the same  
48 number of probes.

49

50

## 51 **Results**

52

53 Conventional smFISH probes or its derivatives are using chemical labeling for  
54 conjugating a fluorophore to the internal, 5' or 3' end of an unlabeled  
55 oligonucleotide pre-equipped with an amine group, which is readily reacting  
56 with fluorophores functionalized with a N-Hydroxysuccinimide (NHS) ester<sup>2,8,9</sup>.  
57 We apply a novel enzymatic fluorophore labeling method, which is based on  
58 the usage of T4 DNA ligase (T4DL), for HuluFISH 1.0. It does not require  
59 any amine modification of unlabeled gene-specific oligonucleotides (GSO)  
60 for HuluFISH probe. As a consequence, it is now possible to cost-effectively  
61 synthesize single fluorophore labeled smFISH probes (Figure 1a). Comparing  
62 with other enzymatic labeling methods we have tested, T4DL has the most cost-  
63 effective design (Suppl. Figure1a). In the T4DL based labeling strategy, only a  
64 standard PCR primer quality oligonucleotide is required, and the free 3'  
65 hydroxyl group from the GSO is enzymatically conjugated with a common  
66 pre-fluorescently-labeled oligonucleotide (termed Hulu), mediated by an  
67 adaptor with 4 bp 3' degenerative sequence to facilitate duplex formation  
68 (Figure 1a). This new T4DL based chemistry also abolishes the necessity of  
69 HPLC purification of smFISH probe (Suppl. Figure1b and 1c). The  
70 polyacrylamide gel electrophoresis (PAGE) purified mouse Gapdh probe  
71 has comparable detection sensitivity with the commercially available one  
72 (Figure 1b).

73 Currently, the smFISH probe selection is based on melting temperature  
74 ( $T_m$ )<sup>3</sup> or Gibbs free energy<sup>10</sup>, which are not very indicative of probe  
75 hybridization efficiency. We developed a pipeline based on Primer3<sup>11</sup> and  
76 DECIPHER<sup>12</sup> to design and filter for GSO with high hybridization efficiency,  
77 which is a more tangible indicator (Figure 1c). Comparing with the  
78 conventional  $T_m$  based method, our probe design has better signal-to-noise  
79 ratio (SNR) and higher contrast (Figure 1d and 1e). With this new approach,  
80 we still have short probe (17-21 bp) to minimize the off-target effect, and a  
81 good balance between hybridization capacity and the number of probes we  
82 could design for smaller RNA (minimally 24 GSOs). This can be used for  
83 customized probe design for any other smFISH methods.

84  
85 In principle, our T4DL based labeling method also enables multiple  
86 fluorophore labeling if the Hulu oligonucleotide is pre-synthesized with multiple  
87 fluorophores. However, the technical complexity increases with the number  
88 of fluorophores to be incorporated into a single oligonucleotide. Therefore,  
89 we extended the T4DL based labeling to multiple-way ligation for  
90 incorporating multiple single fluorophore labeled Hulu oligonucleotides  
91 (HuluFISH 2.0, Figure 2a). Ligation control experiment shows that HuluFISH  
92 2.0 has a specific ligation product for Gapdh HuluFISH probes, and higher  
93 yield compared with the HuluFISH 1.0 (Suppl. Figure 2a). The HuluFISH 2.0  
94 is insensitive to ligation conditions (Suppl. Figure 2b), which demonstrates the

95 robustness of the its probe preparation over temperature, reaction time, etc.  
96 One critical challenge for using multi-colored probe is that when multiple  
97 fluorophores are close to each other, they could be quenched by multiple  
98 mechanisms, for example self-quenching and Förster resonance energy  
99 transfer (FRET)<sup>13,14</sup>. Considering the size limitation of the Hulu  
100 oligonucleotide, here we use 15 bp spacing for the individual dye, and an  
101 adaptor oligonucleotide annealed with the Hulu oligonucleotide in order to  
102 rigidify the ssDNA backbone for dyes (Figure 2b).

103

104 Gapdh probe staining without the adaptor masks the FISH signal by fusing  
105 dots with high background in all channels for Atto488, Atto565, and  
106 Atto647N (Figure 2c). And these dot-like signals 3 channels are not co-  
107 localized very well. With the stabilization by the adaptor oligonucleotide,  
108 individual clear dots can be obtained in all 3 channels and well co-localized  
109 within every channel for Gapdh probes (Figure 2d). Without GSO, the Hulu-  
110 adaptor duplex does not generate any dot like signal (Suppl. Figure 2c). With  
111 the multiple labeling capacity of our method, we could assign various color  
112 combinations to a panel of genes, and decode the dots by counting their  
113 appearance in channels (Suppl. Figure 2d). The evolved multiple fluorophore  
114 labeling capability with HuluFISH 2.0 extends the conventional smFISH with  
115 an autonomous combinatorial color barcoding mechanism. Fluorophores in  
116 each color combination are covalently linked with individual probe, therefore

117 the fluorophore stoichiometry is invariable between probes. During imaging  
118 acquisition, the intensity ratio between fluorophores will be independent of  
119 the brightness of FISH dots. The barcoding capacity simply increases with  
120 the exponentials of the channel (fluorophore choice) number  $n$  (the  
121 theoretical number of combinations is the sum of all color combinations:  
122  $\sum_{k=1}^n \binom{n}{k} = 2^n - 1$ ). If the relative ratio of the maximal intensities of each FISH  
123 dot among channels can be precisely determined, the number of  
124 combinations can be higher.

125  
126 One of the most interesting applications for smFISH is exploring the multiple  
127 gene expression patterns in tissue samples. Just with 3 base colors, the  
128 color combinations can be used to detect 7 genes in one round of  
129 hybridization. Here we use embryonic day 12.5 (E12.5) mouse telencephalon  
130 cryo-section samples to visualize the tissue heterogeneity of these 7 genes  
131 (Figure 3a). Simultaneous 7-gene detection shows the molecular  
132 heterogeneity of fetal brain neural progenitors *in vivo* (Figure 3b).  
133 Hierarchical cluster analysis reveals subgroups of mouse telencephalon  
134 neural progenitors on single cell resolution (Figure 3c and 3d).

## 135 Discussion

136

137 Here, we present the HuluFISH as a new framework for smFISH. HuluFISH  
138 has the capability to enzymatically ligate multiple fluorophores to probes,  
139 which are designed by filtering for their hybridization efficiency. And this new  
140 approach allows us to simultaneously detect genes with the multiplexity that  
141 increases exponentially with the number of available microscopy lasers and  
142 fluorophore types. With 4 to 5 color channels, it is possible to image 15 to  
143 31 genes in one round of hybridization, which will fulfill a large number of  
144 experimental needs in detecting multiple RNA species, without resorting to  
145 multiple-step sequential hybridizations or super-resolution microscopy.  
146 HuluFISH labeling method is compatible with any other FISH related  
147 techniques. In particular, SeqFISH<sup>4</sup> or MERFISH<sup>5</sup> could employ HuluFISH  
148 labeling to either reduce the number of hybridization steps for fixed  
149 multiplexity or increase the multiplexity within their operational steps.  
150 Multiplexing *in situ* quantification of gene expression has become the next  
151 frontier in many fields for biomedical research. We believe the broad  
152 application of HuluFISH and its derivatives will greatly facilitate the  
153 discovery processes like cellular heterogeneity and precise gene expression  
154 regulation, in particular for project like the Human Cell Atlas Initiative.

155

156

## 157 **Methods**

158

### 159 **Cell culture and tissue section preparation**

160

161 Mouse Hepa 1-6 cells were cultured in DMEM medium with 10% fetal  
162 bovine serum and 1 x penicillin/streptomycin. Hepa 1-6 cells were directly  
163 grown on coverslip without coating. Embryonic mouse brain tissue cryo-  
164 sections were cut at 6 to 10  $\mu\text{m}$  from embryonic day 12.5 C57BL/6J mouse  
165 embryo embedded in Tissue-Tek O.C.T. (Sakura, 4583). Adherent Hepa 1-  
166 6 cells or cryo-sections were fixed with 4% formaldehyde in PBS for 10 min  
167 and then quenched with 135 mM glycine in PBS for 10 min at room  
168 temperature. Fixed cells were then washed once with PBS and  
169 permeabilized in 70% ethanol overnight at 4 °C. All water used for FISH  
170 related buffers was diethyl pyrocarbonate (DEPC) treated. After  
171 permeabilization cells were stored in cryo-protectant (25% glycerol, 25%  
172 ethylene glycol, 0.1M phosphate buffer, pH 7.4) at -20 °C until FISH  
173 staining.

174

### 175 **Probe design**

176

177 smFISH probes based on the conventional design<sup>3</sup> were implemented in a R  
178 script to select GSOs first with Primer3<sup>11</sup> to get all possible GSOs without strong  
179 secondary structure from the input mRNA sequence using the standard  
180 condition for selecting the right\_primer in Primer3. Then non-overlapping  
181 GSOs were selected with minimally 2 bp gap. For HuluFISH 1.0 probes, all

182 GSOs from Primer3 were additionally calculated for their hybridization  
183 efficiency with DECIPHER package in R<sup>12</sup> under the condition used for  
184 staining. And the GSOs were filtered to have hybridization efficiency above 0.9  
185 (maximally 1) and then non-overlapping HuluFISH GSOs were selected as  
186 before. For HuluFISH 2.0 GSOs, additional tag sequence was added to  
187 their 3' end after their selection. Adaptor, tag for GSO and Hulu sequences  
188 were randomly generated and controlled for strong secondary structure by  
189 UNAFold<sup>15</sup>. Passed sequences were blasted against a local mouse and  
190 human transcript database (ensemble release 87) for less than or equal to  
191 15 bp exact match.

## 192 **HuluFISH probe labeling and purification**

193  
194  
195 HuluFISH was initially an acronym for Helix-stabilized, unbiased and ligated  
196 uni/multi-color probe for FISH. In search of a multicolor object such as  
197 rainbow and confetti to name this technology in an imagery fashion, we got  
198 the inspiration from a famous Chinese cartoon, Hulu Brothers (húlú is  
199 calabash in Chinese), where each of the seven protagonists was  
200 transformed from a calabash with a distinct rainbow color, much like the  
201 base color in HuluFISH multiplexing. Besides, calabash fruits on a winding  
202 vine resemble fluorophores on a helical HuluFISH probe. Additionally, húlú  
203 bears the image of life in Chinese culture: it is a container for elixir and a  
204 symbol for reproduction, coinciding with the intended use of HuluFISH in

205 medicine and life science.

206

207 FISH GSOs and adaptor oligonucleotides were synthesized from Sigma with  
208 lowest quality for purification (desalting). For individual gene, GSOs were  
209 pooled together to have 100  $\mu\text{M}$  total oligonucleotide concentration.  
210 Fluorescent Hulu oligonucleotides were purchased from Eurofins Genomics  
211 with various dyes, including Atto dyes, Alexa dyes or Cy dyes. For HuluFISH  
212 1.0, ligation was performed in T4 DNA ligase buffer (NEB, B0202S), with 30  
213  $\mu\text{M}$  adaptor for HuluFISH 1.0, 3  $\mu\text{M}$  GSOs and Hulu oligonucleotide, 25%  
214 PEG8000, 30 U/ $\mu\text{L}$  T4 DNA ligase (NEB, M0202M). Ligation reaction mix was  
215 then incubated in a thermocycler, with 12 cycles of 37 °C 10 seconds / 16 °C 5  
216 minutes. For HuluFISH 2.0, ligation reaction mix was prepared as HuluFISH  
217 1.0 with some modifications, such as 16.7  $\mu\text{M}$  of GSOs, adaptor for  
218 HuluFISH 2.0 and Hulu 2.0 oligonucleotides, 50 U/ $\mu\text{L}$  T4 DNA ligase. Then  
219 the ligation mix was left in the dark at room temperature for 2 hours. The  
220 ligation product was concentrated with 9 volumes of butanol and centrifuged  
221 as pellet at 20,000 g, 15 minutes at 4 °C. colorful labeled oligonucleotide  
222 pellet was washed once with 100% ethanol and spin down to remove  
223 ethanol, then resolubilized in loading buffer (8M Urea, 1 x TBE (Carl Roth,  
224 A118.1), 0.01% bromophenol blue and xylene cyanol). With 5 minute dena-  
225 turing at 90 °C, oligonucleotides were loaded onto 15% Urea-PAGE gel (8M

226 Urea, 1 x TBE, 15% Rotiphorese Gel 30 (Carl Roth, 3029.2), 0.05%  
227 ammonium persulfate, 0.05% tetramethylethylenediamine) pre-run at 300 V  
228 for 30 minutes. Running condition was usually 300 V, 30 minutes, or until the  
229 bromophenol blue reached the end. Gel bands with fluorescent dye-  
230 oligonucleotide conjugates were excised under the ambient light. Gel pieces  
231 were homogenized manually by microtube pestle (Sigma, Z359947-100EA),  
232 and then extracted with 500  $\mu$ L 10 mM TE buffer (pH 8.5, 10 mM  
233 tris(hydroxymethyl)aminomethane (Tris), 1 mM Ethylenediaminetetraacetic  
234 acid (EDTA)) at room temperature overnight, protected from light by  
235 wrapping in aluminium foil. The extracted oligonucleotides in TE were  
236 concentrated again by butanol, and washed once by ethanol like before. The  
237 final pellets were dried in dark at room temperature for 5-10 minutes, and  
238 then resolubilized in H<sub>2</sub>O. The concentration was determined by nanodrop one  
239 (ThermoFisher) as ssDNA.

240

## 241 **FISH probe staining and imaging**

242

243 HuluFISH probe mix was adjusted to 10 nM for each single oligonucleotide  
244 in hybridization buffer (2 x SSC (saline-sodium citrate), 10 % (w/v) dextran  
245 sulfate, 10% (v/v) formamide, 1 mg/mL tRNA (Roche, 10109541001), 2 mM  
246 ribonucleoside vanadyl complex (NEB, S1402S), 0.2 mg/mL BSA). Gapdh-  
247 Quasar570 probe was purchased from Biosearch Technology, resuspended  
248 and used for the staining as instructed from the manufacturer. Hybridization

249 was performed in a water bath at 30 °C overnight, with the sample faced  
250 down on the parafilm. Cells on coverslip or tissue sections on glass slide  
251 were washed with washing buffer (2 x SSC, 10 % (v/v) formamide, 0.1 %  
252 (w/v) Tween-20) at 37 °C for 6 × 10 minutes. The last washing step included  
253 0.5 µg/mL DAPI (4',6-diamidino-2-phenylindole) for nuclei staining. The  
254 sample was mounted in  
255 ProLong Gold Antifade (ThermoFisher, P10144), and cured overnight. The  
256 sample then was either imaged on a widefield microscope (Zeiss Cell  
257 Observer) with 200 ms, 950 ms and 5000 ms for 405nm, 488 nm and 561  
258 nm channel, or on a confocal microscope with Airyscan (Zeiss LSM800,  
259 equipped with 405, 488, 561, and 640 nm laser) with maximal laser power  
260 (c.a. 5%) in each channel. The sample was scanned with Airyscan  
261 technology with the optimal settings provided by ZEN software.

## 262 **Image analysis**

263  
264 Except for the nuclear outline manually defined in ImageJ, all the image  
265 analysis was performed in R, and majorly based upon the package  
266 EBImage<sup>16</sup>. All intensity threshold values were based on the arbitrary units  
267 generated from Zeiss Airyscan and thus not specified in the following  
268 description. FISH dot identification relied on 2D local maxima identification  
269 and alignment. Initially for each frame, 2D maxima above a low threshold  
270 value were identified. Each 2D local maximum regarded its projection on the  
271 neighboring z-slices for alignment: those that fall within 0.08 µm were

272 assigned to the same FISH dot. The pixels with maximal intensities  
273 (pseudo-3D-maxima) for identified FISH dots were extracted for further  
274 analysis.

275 Signal-to-noise ratio (SNR) and contrast were generated adaptively for each  
276 individual FISH dot. To this end, pixel values (local background) were taken  
277 from a square centered around the pseudo-3D-maxima, excluding all  
278 circular regions covering the PSF (point spread function) for 2D maxima on  
279 the same plane. Contrast is defined as the ratio of the maximal intensity and  
280 the mean of its local background values; SNR, as traditionally defined,  
281 equals to maximal intensity divided by the standard deviation of local  
282 background values.

283 For color decoding in samples with Hulu-probe for multiple genes, the  
284 presence of fluorophore on each channel was initially separately  
285 determined. Dual or triple color coding was assigned when FISH dots from  
286 different channels co-localized within 0.08  $\mu\text{m}$ . Single color assignment  
287 required thresholding with a higher intensity, given there were three copies  
288 of fluorophores in the single-color Hulu-probe. Nuclei were manually  
289 segmented on the maximum intensity projected image in ImageJ. Without  
290 the assistance of membrane immunostaining, each identified FISH dot was  
291 assigned to its closest nuclei.

## 292 **Statistical analysis**

293  
294 Wilcoxon two samples test was used for evaluating the significance of our

295 probe design based on HybEff and the conventional one based on Tm. p-  
296 value is indicated in corresponding figure legend. Single cell gene set  
297 expression data were hierarchically clustered using Euclidean distance and  
298 shown as heatmap. 5 clusters were retrieved from 58 cells by cutting the  
299 dendrogram tree.

### 300 **Author Contributions**

301 Y.S.C and H.K.L conceived and designed the project. Y.S.C performed  
302 most of the experiments. Y.Z helped with Image acquisition and analysis,  
303 K.H, P.Z. G.B and H.A helped with FISH staining H.K.L supervised the  
304 project. Y.S.C, Y.Z, and H.K.L analyzed the results and wrote the  
305 manuscript.

### 306 **Acknowledgement**

307 We thank Sven Poppelreuther Carl Zeiss Imaging center in the DKFZ for  
308 providing critical suggestions in using Airyscan technology. We thank Bernd  
309 Fisher for generously provide R code for local maxima. This work was  
310 supported by the Helmholtz Association (VH-NG-702), the Deutsche  
311 Forschungsgemeinschaft (LI 2140/1-1), the Deutsche Krebshilfe (110226),  
312 the Helmholtz Alliance Preclinical Comprehensive Cancer Center (PCCC),  
313 the Chica Heinz Schaller foundation, the EMBO young investigator program,  
314 and the ERC (European Research Council) consolidator grant (647055) to  
315 H-K.L.

316

317 **Competing Financial Interests**

318 Y.S.C and H.K.L are inventors on two provisional patent applicatons which present  
319 the HuluFISH.

320 **Material & Correspondence**

321 Correspondence and requests for materials should be addressed to Y.S.C  
322 (yongsheng.cheng@dkfz.de) or H.K.L (email: l.haikun@dkfz.de).

323

324

## 325 **References**

326

327 1. Gall, J. G. & Pardue, M. L. Formation and Detection of Rna-Dna Hybrid

328 Molecules in Cytological Preparations. *Proc. Natl. Acad. Sci.* 63, 378–383

329 (1969).

330 2. Femino, A. M., Fay, F. S., Fogarty, K. & Singer, R. H. Visualization of Single

331 RNA Transcripts in Situ. *Science* 280, 585–590 (1998).

332 3. Raj, A., van den Bogaard, P., Rifkin, S. A., van Oudenaarden, A. & Tyagi, S.

333 Imaging individual mRNA molecules using multiple singly labeled probes. *Nat*

334 *Meth* 5, 877–879 (2008).

335 4. Lubeck, E., Coskun, A. F., Zhiyentayev, T., Ahmad, M. & Cai, L. Single-cell in

336 situ RNA profiling by sequential hybridization. *Nat. Methods* 11, 360–361

337 (2014).

338 5. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially

339 resolved, highly multiplexed RNA profiling in single cells. *Science* aaa6090

340 (2015). doi:10.1126/science.aaa6090

341 6. Lubeck, E. & Cai, L. Single-cell systems biology by super-resolution imaging

342 and combinatorial labeling. *Nat. Methods* 9, 743–748 (2012).

343 7. Levesque, M. J. & Raj, A. Single-chromosome transcriptional profiling reveals

344 chromosomal gene expression regulation. *Nat. Methods* 10, 246–248 (2013).

345 8. Raj, A. & Tyagi, S. in *Methods in Enzymology* 472, 365–386 (Elsevier, 2010).

346 9. Lyubimova, A. *et al.* Single-molecule mRNA detection and counting in

347 mammalian tissue. *Nat. Protoc.* 8, 1743–1758 (2013).

- 348 10. Tsanov, N. *et al.* smiFISH and FISH-quant – a flexible single RNA detection  
349 approach with super-resolution capability. *Nucleic Acids Res.* gkw784 (2016).  
350 doi:10.1093/nar/gkw784
- 351 11. Untergasser, A. *et al.* Primer3—new capabilities and interfaces. *Nucleic Acids*  
352 *Res.* 40, e115–e115 (2012).
- 353 12. Wright, E. S., Yilmaz, L. S., Corcoran, A. M., Ökten, H. E. & Noguera, D. R.  
354 Automated Design of Probes for rRNA-Targeted Fluorescence In Situ  
355 Hybridization Reveals the Advantages of Using Dual Probes for Accurate  
356 Identification. *Appl. Environ. Microbiol.* 80, 5124–5133 (2014).
- 357 13. Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nat.*  
358 *Methods* 5, 507–516 (2008).
- 359 14. Zhuang, X. *et al.* Fluorescence quenching: A tool for single-molecule protein-  
360 folding study. *Proc. Natl. Acad. Sci.* 97, 14241–14244 (2000).
- 361 15. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction.  
362 *Nucleic Acids Res.* 31, 3406–3415 (2003).
- 363 16. Pau, G., Fuchs, F., Sklyar, O., Boutros, M. & Huber, W. EBImage—an R  
364 package for image processing with applications to cellular phenotypes.  
365 *Bioinformatics* 26, 979–981 (2010).

366

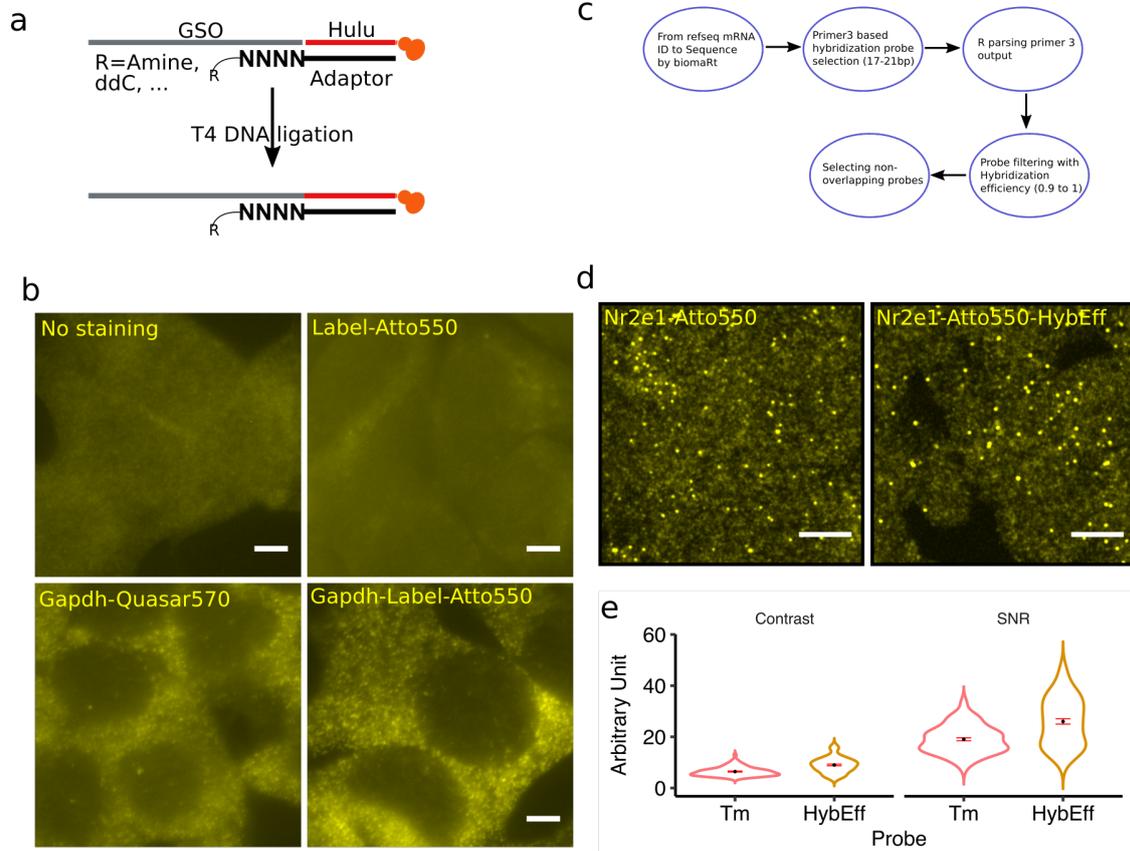
367

368

369

370

371 **Figures and Figure Legends**



372  
373

374 **Figure 1: HuluFISH 1.0 probe's enzymatic labeling and improved probe**

375 **design.** (a) T4DL based labeling scheme for HuluFISH probe with standard

376 3' end -OH group of GSO. (b) *in situ* staining with HuluFISH probe (Gapdh-

377 Label-Atto550) and commercial mouse Gapdh probe (Gapdh\_Quasar570) in

378 Hepa 1-6 cell. (c) smFISH probe selection pipeline used for all following

379 probes in this paper. Scale bar: 10  $\mu$ m. (d) smFISH detection of low-

380 expressing gene, Nr2e1 in embryonic brain tissue with conventional Tm

381 based or our new hybridization efficiency based probe design. Scale bar: 5

382  $\mu$ m. (e) Contrast and SNR analysis for Nr2e1's conventional (Tm) and our

383 (HybEff) design. Between Tm and HybEff, welch two sample t-test's p-value

384 for contrast and SNR are  $2.6 \times 10^{-10}$  and  $3.0 \times 10^{-8}$ .

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

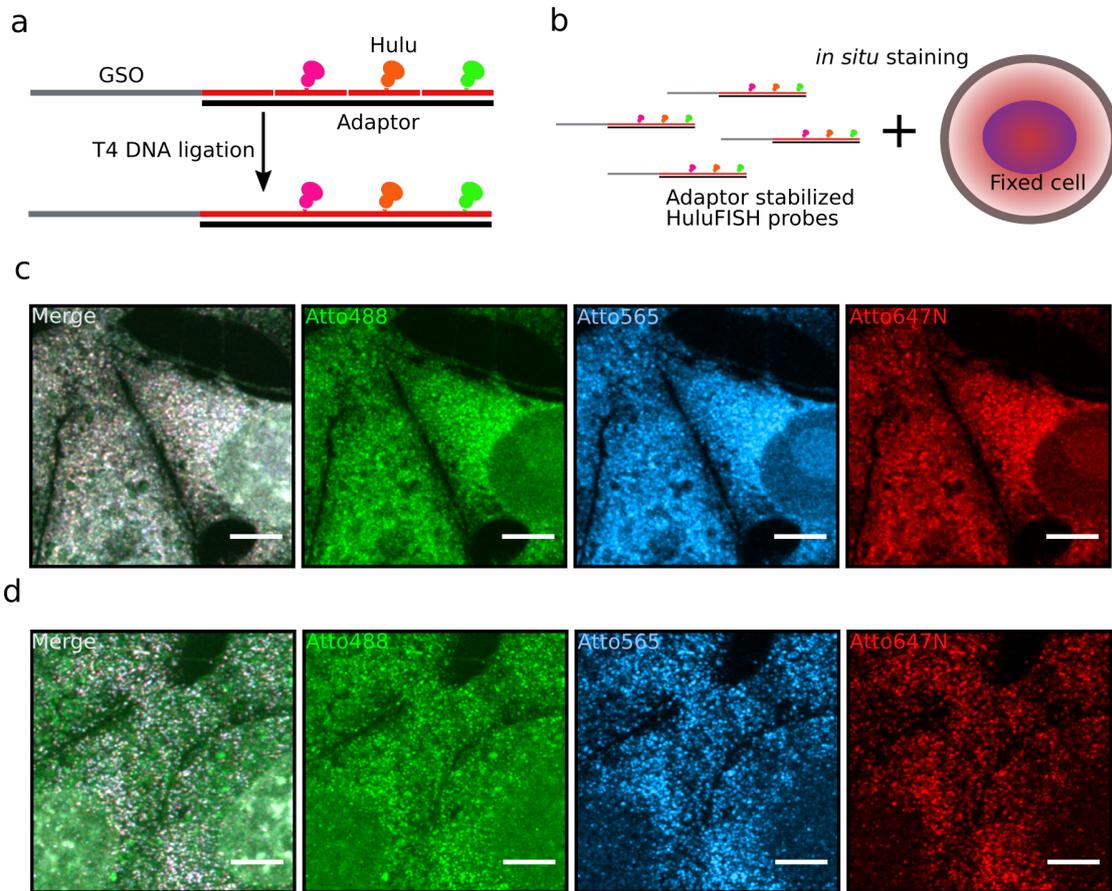
415

416

417

418

419

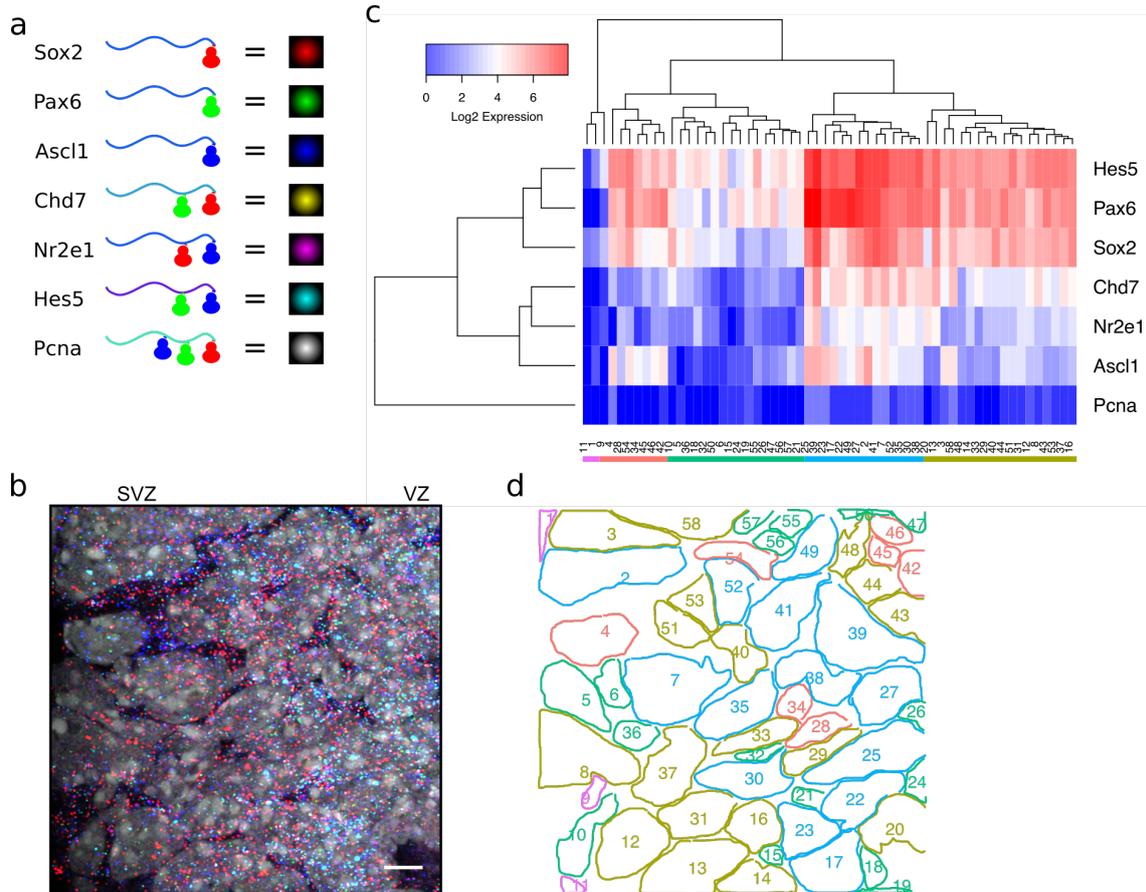


420  
421

422 **Figure 2: Multiple fluorophore labeling based on HuluFISH 2.0.**

423 (a) Multiple-way ligation based fluorophore labeling of HuluFISH probes. (b) *in*  
424 *situ* staining with HuluFISH probes pre-annealed with the adaptor to avoid  
425 multiple fluorophore quenching. (c) Gapdh expression in Hepa 1-6 with the  
426 Gapdh HuluFISH probe conjugated with Atto488, Atto565, and Atto647N,  
427 without adaptor stabilization. (d) Gapdh mRNA visualized as individual dots by  
428 adaptor pre-annealed HuluFISH probe. Scale bar is 5  $\mu$ m.

429  
430  
431



432

433

**Figure 3: HuluFISH detection of 7 genes in mouse embryonic brain.**

434

(a) Color coding scheme for HuluFISH from 3 base colors. (b) 7-gene

435

detection in E12.5 mouse embryonic brain ventricular zone (VZ) and

436

subventricular zone (SVZ). Scale bar 5 μm. (c) Hierarchical clustering of all

437

single cells in (b) based on the Log<sub>2</sub> transformation of mRNA transcript

438

counts for these 7 genes. (d) Spatial illustration of molecular subgroups in

439

mouse telencephalon neural progenitors identified in (c). Cluster color

440

scheme is the same as it in (c).

441