PNAS
www.pnas.org

Supplementary Information for

High-fidelity amplified FISH for the detection and allelic discrimination of single mRNA molecules

Salvatore A. E. Marras, Yuri Bushkin and Sanjay Tyagi

Sanjay Tyagi
Email: sanjay.tyagi@rutgers.edu

This PDF file includes:

Other supplementary materials for this manuscript include the following:

Dataset S1 is provided as an Excel Spreadsheet at the online site of the article

Supplementary Materials and Methods

Probe synthesis, purification, and annealing. The sequences of all of the probes, targets, and hybridization chain reaction (HCR) hairpins are listed in the supplementary section entitled Sequences of the Probes and Targets. The HCR hairpins were based on the designs described by Choi et al. and Koos et al. (1,2). Dye modified oligonucleotides were purchased from LGC Biosearch Technologies (Petaluma, CA), and unmodified oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). For the synthesis of sm-FISH probes, pools of amino-linked oligonucleotides were coupled to fluorophores, and then purified as described earlier (3). HCR hairpins were labeled and purified using the same approach. Relatively longer oligonucleotides, including the HCR hairpins, the right acceptor probes, and the left acceptor probes, were purified further by denaturing polyacrylamide gel electrophoresis through a 10% polyacrylamide gel containing 8 M urea.

All hairpin-containing oligonucleotides were "snap cooled" to promote self-annealing prior to use. In this procedure, the oligonucleotides were dissolved in 2X SSC (300 mM sodium chloride, 30 mM sodium citrate) at a 2.5 μM concentration in volumes that ranged from 25 to 200 μL, placed in boiling water for 2 min, and then returned to room temperature for at least 5 min.

For ligations that utilized click chemistry (4), the 3′ amino groups in the oligonucleotides that were complementary to target sequences were modified by the addition of an azide group, using a 4-azidobutyrate-N-hydrosuccinimidyl ester (Sigma-Aldrich, St. Louis, MO). Separately, the amino group at the 5′ ends of the generic right acceptor hairpins, or at the 5′ ends of the generic passive tags, was conjugated to dibenzocyclooctyne (DBCO) using dibenzocyclooctyne-N-hydroxysuccinimidyl ester (Sigma-Aldrich). These reactions were carried out in 100 mM sodium bicarbonate (pH 8.3), made fresh just before use, and the mixtures were then incubated overnight at 4 ºC using at least a 10-fold molar excess of the hydroxysuccinimidyl ester compared to the concentration of the oligonucleotide. The coupled oligonucleotides were then purified by reverse-phase high-pressure liquid chromatography. The two oligonucleotides were then linked to each other by mixing them in an equimolar ratio, and then incubating them overnight at 4 ºC in a buffer composed of 50 mM KCl, 2.5 mM

$MgCl_2$, and 10 mM Tris-HCl (pH 8.3). Finally, the ligated oligonucleotides were purified using denaturing polyacrylamide gel electrophoresis through a 10% polyacrylamide gel containing 8 M urea, resuspended in water, and then their concentration was measured using a Nanodrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA). These probes were snap cooled prior to use, as described above.

Reverse transcription PCR with SuperSelective primers. In order to determine the presence or absence of the L858R point mutation in *EGFR* mRNAs expressed in different cells lines, we performed real-time reverse transcription polymerase chain reactions with SuperSelective primers that have the ability to selectively amplify sequences that differ from each other by a single nucleotide (5). RNA from either HeLa, H1975, or A431 cell lines (cultured on 10-cm plates to 50% confluence) was isolated using TRIzol RNA isolation reagents (Thermo Fisher Scientific), following the manufacture's instructions. cDNAs were then synthesized from these mRNA templates by performing reverse transcription in a 20-µL volume, initiated with 1 µg of RNA, utilizing SuperScript IV Reverse Transcriptase (Thermo Fisher Scientific) and an *EGFR* mRNA-specific reverse transcription primer following the manufacture's instructions.

For each cDNA sample, three real-time PCR assays were carried out. In the first reaction, a pair of conventional forward and reverse primers designed to bind to a region of the *EGFR* cDNA on either side of the site of the L858R mutation was used (outer primers), so that both the wild-type and the mutant sequences would be amplified equally efficiently; in the second reaction, a mutant-specific SuperSelective primer was used in combination with a conventional reverse primer to amplify only the cDNA from mutant transcripts; and in the third reaction, a wild type-specific SuperSelective primer was used in combination with the same conventional reverse primer to amplify only the cDNA from wild-type transcripts. The sequences of all of primers are listed in the supplementary section entitled Sequences of the Probes and Targets.

Real-time polymerase chain reactions were performed in 25-µL volumes containing 20 mM Tris-HCl (pH 8.3), 50 mM KCl, 3 mM $MgCl_2$, 0.5% Tween 20 (Sigma-Aldrich), 50 mM tetramethylammonium chloride (Sigma-Aldrich), 0.5X SYBR Green (Thermo Fisher Scientific), 250 mM dATP, 250 mM dCTP, 250 mM dGTP, 250 mM dTTP, and 1.5 Units of Platinum *Taq* DNA Polymerase (Thermo Fisher Scientific).

These reactions were initiated with 2.5 µL of a 100-fold diluted cDNA sample. PCR was carried out on a CFX96™ Touch Real-Time PCR Detection System (Bio-Rad Laboratories, Hercules, CA), which was programmed as follows: 2 min at 95 ºC to activate the DNA polymerase; followed by 35 cycles of denaturation for 15 sec at 95 ºC, annealing for 20 seconds at 60 ºC, and primer extension for 20 seconds at 72 ºC, while monitoring fluorescence intensity during each primer extension step.

Cell culture, mutagenesis, transfection, and cell fixation. HeLa, A431, and H1975 cells were obtained from the American Type Culture Collection (Manassas, VA), and were cultured on glass coverslips (thickness 0.17 mm) coated with 0.1% gelatin in modified Eagle's Minimum Essential Medium (Sigma-Aldrich) supplemented with 10% fetal bovine serum (Clontech, Mountain View, CA). Peripheral blood mononuclear cells (PBMCs) were isolated from the buffy coat fraction of blood samples, cultured, and stimulated as described earlier (6). Mutations were introduced into a sequence encoding green fluorescent protein (GFP) at the indicated site in plasmid pTRE-d2EGFP (Clontech) by site-directed mutagenesis, utilizing a USB® Change-IT kit (Affymetrix, Santa Clara, CA), following the manufacturer's instructions. The presence of the expected mutations was confirmed by nucleotide sequence analysis. The plasmids were transfected into HeLa cells; the cells were detached after an overnight incubation; and the cells were then re-plated onto glass coverslips coated with gelatin. For *in situ* hybridization, the coverslips were withdrawn from the culture medium, washed briefly with phosphate-buffered saline, pH 7.4 (PBS), fixed using 4% formaldehyde dissolved in PBS for 10 min, and then incubated with 70% alcohol for at least 10 min. The coverslips containing the fixed cells were either stored in 70% alcohol in a refrigerator at 4 ºC, or they were used immediately.

Probe and assay optimization for single-nucleotide discrimination. We followed two different strategies for the discrimination of single-nucleotide variations (SNVs). In the first strategy, there was no toehold sequence, the hybridization was done at 37 ºC in a single step in which both donor and acceptor probes were added together; and in the second strategy, in which there was a 5-nucleotide-long toehold, the acceptor probes were hybridized first at 50 ºC, followed by the removal of the excess probes by washing, and a second hybridization at 50 ºC with the donor probes.

The results of optimization of the length of the loop and toehold sequence in the first strategy are presented in Fig. S2. The levels of discrimination obtained with the first strategy are depicted in Fig. 2, Fig. S2, and Fig. S3.

In the second strategy, utilizing the *EGFR* L858R mutation, and depicted in Fig. 3, Fig. 4, Fig. S5, Fig. S6, Fig. S7, and Fig. S8, we used relatively longer hairpin loops in the donor probes, we used a 5-nucleotide-long toehold sequences in the acceptor probes, we increased the temperature of hybridization to 50 ºC, and we performed a two-step hybridization. In this strategy, we first hybridized the two acceptor probes for 6 hrs, and then removed excess probes by washing, and we then performed the second hybridization overnight with the pair of mutant-specific and wild-type-specific right and left arm-donating probes. The rationale behind this second strategy was that if the donor probes are added after the removal of the excess acceptor probes, then the probability of off-site interactions between the donor and acceptor probes would be reduced to a great extent. As it turned out, both hybridization protocols yielded high-specificity results.

Although toehold interactions between adjacent probes enhances the kinetics of initiation of strand displacement and the eventual unmasking of the HCR initiator, we found that toehold interactions are not essential. In the most common probe designs, the arm sequence of the donor probes was the same sequence as the toehold sequence. However, in cases in which a toehold sequence was not used, a portion of the HCR initiator sequence was utilized for this purpose (see section entitled Sequences of the Probes and Targets).

Hybridization, washing, hybridization chain reactions, and mounting of the coverslips. Cells attached to coverslips were equilibrated with hybridization wash buffer (10% formamide, 2X SSC), and then immersed in 50 µL of hybridization buffer, which consisted of 10% dextran sulfate (Sigma-Aldrich), 1 mg/mL *Escherichia coli* transfer RNA (Sigma-Aldrich), 2 mM ribonucleoside vanadyl complexes (New England Biolabs, Ipswich, MA), 0.02% ribonuclease-free bovine serum albumin (Thermo Fisher Scientific), 10% formamide, 2X SSC, and probes that differed in amount and composition. This hybridization reaction mixture was first added as a droplet onto a stretched-out piece of Parafilm (Bemis in North America, Oshkosh, WI) over a glass

plate, and then a coverslip containing the cells was placed faced down onto the droplet, followed by incubation at 37 ºC (or 50 ºC when the probes contained toehold sequences) overnight in a humid chamber. Following hybridization, the coverslips were washed twice for 10 min each in 1 mL of hybridization wash buffer at room temperature.

Following the removal of the excess probes or the excess HCR hairpins by these washes, the coverslips were equilibrated with mounting buffer (2X SCC, 0.4% glucose), and then mounted in the mounting buffer supplemented with 1 µL of 3.7 mg/mL glucose oxidase and 1 µL of catalase suspension (both from Sigma-Aldrich) for each 100 µL preparation. This mounting medium inhibits photobleaching of the probes by catalytically removing the oxygen from the medium (7), and it should be prepared immediately before use. After removing the excess mounting medium by gently blotting with a tissue, the coverslips were sealed with clear nail polish, and then imaged the same day.

HCR was performed in either of two conditions. In condition 1, we used HCR buffer 1 (1M NaCl, 0.05% Tween 20, and 100 mM NaH_2PO_4-Na_2HPO_4, pH 7.4) that also contained 25 nM of each HCR hairpin, and performed HCR at 37 ºC for 2-4 hr. In condition 2, we used HCR buffer 2 (750 mM NaCl, 75 mM sodium citrate, 10% dextran sulfate, 0.05% Tween 20) that also contained 125 nM of each HCR hairpin, and performed HCR at 25 ºC for 2-4 hr. The second condition yielded HCR signals that were about twice as bright as the first condition. The second condition was used for experiments described in Fig. S9, whereas the first condition was employed for the rest of the experiments. All hybridization, wash, and HCR solutions were filtered through a 0.2-micron filter (Thermo Fisher Scientific).

When cells were prepared for flow cytometry (HeLa cells or PBMCs), fixation, permeabilization, hybridization, and washing were carried out in 1.5 mL Eppendorf test tubes. For each wash cycle, the cells were centrifuged at 1,000 rpm in a swinging-bucket microcentrifuge (Thermo Fisher Scientific) for 2 min, and then the supernatant was removed by gentle pipetting. In order to ensure that cells were not lost due to their adherence to the walls of these test tubes, the hybridization wash buffer was supplemented with 10% tissue culture medium (including fetal bovine serum) and 2 mM ribonucleoside vanadyl complexes. Flow cytometry was carried out on cells suspended

in 2X SSC in an Accuri 6C™ (Fig. S1) or FACSCelesta™ flow cytometer (BD Biosciences, San Jose, CA) (Fig. 5, Fig. S9 and S10). In order to cover a wider dynamic range, the sensitivity of the flow cytometer was set at a lower voltage in the case of the data shown in Fig. S10, compared to the data shown in Fig. S9 (for the GFP channel 200 vs 400 volts and for the Cy5 channel 250 vs 400 volts).

Although the *in situ* hybridizations were carried out with a vast excess of probes over targets, the concentration of the probes had a significant effect on both signal intensity and on the generation of background signals (for example see Table S2). We therefore carried out careful optimizations of the probe concentrations, and we used equivalent concentrations of probes when comparing passively tagged probes to amp-FISH probes. For each 50-µL hybridization reaction volume, we used 25 ng of probes for the directly labeled probe sets that contained 24 probes or 48 probes, 5 ng of each probe for single amp-FISH probe pairs, and 40 to 125 ng of each probe set containing 24 amp-FISH probe pairs. The stock solutions of each multiprobe set were prepared by mixing together an equimolar amount of each probe.

Image acquisition and analysis. Cellular images were acquired using an Axiovert 200M inverted fluorescence microscope (Zeiss, Oberkochen, Germany), using either a 100x oil-immersion objective (numerical aperture 1.3) or using a 63x oil-immersion objective (numerical aperture 1.4), and using either a CoolSNAP HQ camera or a Prime sCMOS camera (both from Photometrics, Tucson, AZ), controlled by either OpenLAB image acquisition software (Perkin-Elmer, Waltham, MA) or Metamorph image acquisition software (Molecular Devices, San Jose, CA). We acquired images of 16-20 optical sections separated from each other by 0.2 µm at 100- to 1,000-millisecond exposure times in each fluorescence color channel.

Spots corresponding to single mRNA molecules were detected within the optical z-stacks utilizing a custom image-processing computer program implemented in MATLAB (MathWorks, Natick, MA), and described earlier (7). For molecular co-localization analysis, z-stacks in two or three channels were analyzed using the same algorithm to locate spots in each channel, and to determine their coordinates in three dimensions. Thereafter, we identified spots in each channel that had a counterpart in another channel that occurred within a distance of 500 nm, and these spots were then classified

as being co-localized. The efficacy of this approach for the identification of spots visible in different channels, but arising from the same mRNA molecules, has been described earlier (8,9). To compare the intensities of the sm-FISH spots to the intensity of co-localized HCR spots, we developed a new computer algorithm that accurately measures the maximum intensity within identified spots. In this algorithm, the average pixel intensity of ring-shaped bands of pixels that lie immediately outside the spot was determined. This provides a measure of the local background, which was then subtracted from the maximum intensity recorded within each spot. This procedure compensates for local differences in background levels, which is a natural feature of sm-FISH imaging, and provides a more accurate measure of spot intensities.

Statistical Analysis

All error bars represent 95% confidence intervals obtained from an analysis of 35-90 cells. Reported p-values are the probabilities of obtaining at least as large a difference in the means of two populations shown by the indicated bars by random chance alone. The p-values were $<1 \times 10^{-5}$, unless otherwise indicated over the horizontal brackets. The p-values were determined by a permutation test performed on the empirical data sets themselves (without assuming any idealized distribution) using a custom script written in MATLAB. In this algorithm, we first combined all the values from the first and the second data series that were being compared, randomized their order, segregated them into two bins of the same size as the original data series, and then computed the difference between their averages. This was repeated 100,000 times and then the probability of finding a difference as large as the experimental difference was determined and reported.

Supplementary Discussion

During the imaging of *EGFR* mRNAs, we noticed that clusters of mRNAs, that are typically visible at the gene locus of actively transcribed genes when utilizing directly labeled sm-FISH probes, were not as clearly detectable with amp-FISH probes. This likely reflects the lower efficiency of HCR within the matrix of the cell nucleus, which is denser than in the cytoplasm. Underscoring the impediments posed by the high density within the nucleus, a similar phenomenon was also reported by Battich et al. in 2013 (10), when they amplified FISH signals using branched-DNA technology.

In our experiments with *EGFR* mRNAs, wild-type and mutant cells could be distinguished from each other without using multiple sm-FISH probes as guides, despite the presence of a few non-specific spots in each cell. This was possible in the case of *EGFR* mRNA, because, on average, this mRNA is expressed at 62 copies/cell in H1975 cells, and at 27 copies/cell in HeLa cells. Other RNAs that are expressed at these or higher levels will also be amenable to detection by the same procedure. However, the levels of expression of different mRNAs in mammalian cells are sometimes quite low. For example, in NIH3T3 cells the median level of expression is 17 molecules/cell (11), and consequently quite a few mRNAs are expressed at even lower levels. Such targets will require the aid of sm-FISH probes to serve as guides to identify legitimate targets (12).

It is important to point out that the level of background signals in *in situ* hybridizations generally increases with the number and the concentration of the probes that are used (Table S2). With our *EGFR* target system, when the number of passively tagged probes was increased from one to 24, and then to 48, the number of background spots increased dramatically (also noted by Choi et al. (13)); however, this did not occur with the binary amp-FISH probes with sequestered HCR initiator. A similar pattern was observed with the GFP mRNA targets (Fig. S10).

When we increased the number of probes, the signals increased for up to 10 probes, but then did not increase any further (Fig. S10). This plateau was observed for both amp-FISH, as well as for, passively tagged probes. Although the reasons why this plateau exists is not clear, it provides an explanation as to why signals from 24 probe pairs were not 24 times as intense as the signals from a single probe pair. A second reason for the sublinear increase in signal intensity with increasing numbers of probes is the fact that individual probes bind to RNA with low and variable efficiencies (Fig. S8). Nonetheless, with 24 pairs of probes, all the RNAs present in the cell are detected (Fig. 4).

Sequences of the Probes and Targets

In the sequences of the probes and the HCR hairpin, the hairpins stems are indicated by underlines, toehold sequences by green shading, regions complementary to the target by cyan

shading, and the discriminating or discriminated nucleotides by purple shading. All sequences are written in the 5´ to 3´ direction.

HCR Hairpins

HCR hairpin H1
GGCGGTTTACTGGATGATTGATGAGGATTTACGAGGAGCTCAGTCCATCCTCGTAAATCCTCATCAATCATC-TMR or Cy3

HCR hairpin H2
TMR or Cy3-CCTCGTAAATCCTCATCAATCATCCAGTAAACCGCCGATGATTGATGAGGATTTACGAGGATGGACTGAGCT

HCR hairpin H3
Cy5-ACAGACGACTCCCACATTCTCCAGGTGGGAGTCGTCTGTAACATGAAGTA

HCR hairpin H4
CTGGAGAATGTGGGAGTCGTCTGTTACTTCATGTTACAGACGACTCCCAC-Cy5

GFP target sequence variants (d2EGFP)

Only a portion of the coding sequence is shown

G variant
UCGUGACCACCCUGACCUACGGCGUGCAGUGCUUCAGCCGCUACCCCGAC

C variant
UCGUGACCACCCUGACCUACGGCGUCCAGUGCUUCAGCCGCUACCCCGAC

A variant
UCGUGACCACCCUGACCUACGGCGUACAGUGCUUCAGCCGCUACCCCGAC

T variant
UCGUGACCACCCUGACCUACGGCGUTCAGUGCUUCAGCCGCUACCCCGAC

GFP Probes for mutation detection

Right donor probe RD5

Used for the experiments shown in Fig. 1. The left underlined sequence serves a dual role of forming an intramolecular stem and serving as a complement of the toehold sequence in the right acceptor probe.
GTTACAGACGACTCCCACAGACGAATACAGCGAGCACTGCACGCCGTTGTATTCGTCT

Right acceptor probe RA5.0

Used for the experiments shown in Fig. 1. The 11-nucleotide toehold sequence is shown in green.

GCTGCTTCATGTGGTCGGGGTAGCGGCGCTGTATTCGTCTGTGGGAGTCGTCTGTAACTACTTCATGTTACAGACGACTCCCAC

Passively tagged right probe

Used for the experiments shown in Fig. S1.

GTCGGGGTAGCGGCTGAAGAAAAATACTTCATGTTACAGACGACTCCCAC

Right donor probe variants

Used for the experiments shown in Fig. 2, Fig. S1, Fig. S3, and Table S1.

C variant RD6.6C
GTTACAGACGACTCCCACCACTGCACGCCGTGGGA

T variant RD6.6T
GTTACAGACGACTCCCACCACTGTACGCCGTGGGA

A variant RD6.6A
GTTACAGACGACTCCCACCACTGAACGCCGTGGGA

G variant RD6.6G
GTTACAGACGACTCCCACCACTGGACGCCGTGGGA

Left donor probe variants

Used for the experiments shown in Fig. 2, Fig. S1, Fig. S3, and Table S1.

T variant LD6.1T
ACGAGGCACTGTACGCCCCTCGTAAATCCTCATCAATCATC

C variant LD6.1C
ACGAGGCACTGCACGCCCCTCGTAAATCCTCATCAATCATC

A variant LD6.1A
ACGAGGCACTGAACGCCCCTCGTAAATCCTCATCAATCATC

G variant LD6.1G
ACGAGGCACTGGACGCCCCTCGTAAATCCTCATCAATCATC

Right acceptor probe RA6.3
Used for the experiments shown in Fig. 2, Fig. S1, Fig. S3, and Table S1.

GTCGGGGTAGCGGCTGAAGGTGGGAGTCGTCTGTAACTACTTCATGTTACAGACGACTCCCAC

Left acceptor probe LA6.1
Used for the experiments shown in Fig. 2, Fig. S1, Fig. S3, and Table S1.

CCTCGTAAATCCTCATCAATCATCCAGTAAACCGCCGATGATTGATGAGGATTTACGAGG
GTAGGTCAGGGTGGTCACGA

Probes used for optimization of discrimination of single-nucleotide variations in GFP mRNA

Used for the experiments shown in Fig. S2.

First probe pair, 15-nucleotide probe sequence with a 5-nucleotide toehold sequence

Right donor probe RD6.3

The left underlined sequence serves the dual role of forming an intramolecular stem and serves as a complement of the toehold sequence in the right acceptor probe.

GTTACAGACGACTCCCACAGTCCAGCACTGCACGCCGTGGACTG

Right acceptor probe RA6.2

ATGTGGTCGGGGTAGCGGCTGAGGACTGTGGGAGTCGTCTGTAACTACTTCATGTTACAGACGAC
TCCCAC

Second probe pair, 15-nucleotide probe sequence, no toehold sequence

Right donor probe RD6.1
GTTACAGACGACTCCCACAGCACTGCACGCCGTGTGGGA

Right acceptor probe RA6.0
ATGTGGTCGGGGTAGCGGCTGAGTGGGAGTCGTCTGTAACTACTTCATGTTACAGACGACTCCCAC

Third probe pair, 12-nucleotide probe sequence, no toehold sequence

Right donor probe RD6.5
GTTACAGACGACTCCCACAGCACTGCACGCGTGGGA

Right acceptor probe RA6.0
ATGTGGTCGGGGTAGCGGCTGAGTGGGAGTCGTCTGTAACTACTTCATGTTACAGACGACTCCCAC

Fourth probe pair, 9-nucleotide probe sequence, no toehold sequence

Right donor probe RD6.7
GTTACAGACGACTCCCACACTGCACGCGTGGGA

Right acceptor probe RA6.0
ATGTGGTCGGGGTAGCGGCTGAGTGGGAGTCGTCTGTAACTACTTCATGTTACAGACGACTCCCAC

Generic right acceptor hairpin

Sequestered initiator for HCR with hairpins H3 and H4 and used for ligating to the right probes to create right acceptor probes. Used for the experiments shown in Figs. 4 and 5.

Amino-GGACT GTGGGAGTCGTCTGTAACTACTTCATGTTACAGACGACTCCCAC

Generic passive tag

Initiator for HCR with hairpins H3 and H4, ligated to right probes. Used for the experiments shown in Figs. 4 and 5. Residues AAA serve as a spacer sequence.

Amino-AAATACTTCATGTTACAGACGACTCCCAC

GFP probes used in the probe tiling experiments

Gray residues are the targets of the right acceptor probes, the yellow residues are the targets of the right donor probes, and both kinds of shaded areas are the targets of the 48 sm-FISH probes. These probes relate to the experiments shown in Fig. S9. In addition, we prepared a set of 24 pairs of amp-FISH probes and a set of 24 passively tagged probes (Dataset S1) for the experiments described in Fig. S10. The target binding regions of the passively tagged probes was the same as the target binding regions of the acceptor probes of the amp-FISH probe sets.

In the tiled probe sets for GFP, IFNγ, and *EGFR* mRNAs, the sequences of the 24 donors were identical, except for their hairpin loop sequences, which were each complementary to different sequences within the target mRNA. Similarly, the sequences of the 24 acceptors were identical, except for their linear 5'-terminal segments, which were each complementary to a sequence in the target mRNA that was adjacent to the sequence where one of the donors bound.

ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGA
CGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCT
ACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCA
CCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGA
AGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCT
TCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACC
CTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGG
GCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAA
GAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCT
CGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACA
ACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACA
TGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACA
AGAAGCTTAGCCATGGCTTCCCGCCGGAGGTGGAGGAGCAGGATGATGGCACGCTGCCC
ATGTCTTGTGCCCAGGAGAGCGGGATGGACCGTCACCCTGCAGCCTGTGCTTCTGCTAGG
ATCAATGTGTAG

***EGFR* probes**

The sequence of the human *EGFR* transcript is shown below, where the target regions of various probes are shown by different color emphases. The cyan region towards the middle represents the target of the amp-FISH probes used for discriminating the L858R mutation. The purple guanidine residue indicates the site of the mutation in this transcript. In the wild-type sequence below, the bold purple nucleotide is a thymidine. In the mutant sequence this nucleotide is a guanosine. Underlined residues are the targets of the 48-directly labeled probes, gray residues are the targets of the right acceptor probes, and the yellow residues are the

targets of the right donor probes. This sequence was used for the experiments shown in Fig. 3, Fig. 4, Fig. S4, Fig. S5, Fig. S6, Fig. S8, and Table S2.

ATGCGACCCTCCGGGACGGCCGGGGCAGCGCTCCTGGCGCTGCTGGCTGCGCTCTGCCC
GGCGAGTCGGGCTCTGGAGGAAAAGAAAGTTTGCCAAGGCACGAGTAACAAGCTCACGCA
GTTGGGCACTTTTGAAGATCATTTTCTCAGCCTCCAGAGGATGTTCAATAACTGTGAGGTGG
TCCTTGGGAATTTGGAAATTACCTATGTGCAGAGGAATTATGATCTTTCCTTCTTAAAGACCA
TCCAGGAGGTGGCTGGTTATGTCCTCATTGCCCTCAACACAGTGGAGCGAATTCCTTTGGA
AAACCTGCAGATCATCAGAGGAAATATGTACTACGAAAATTCCTATGCCTTAGCAGTCTTATC
TAACTATGATGCAAATAAAACCGGACTGAAGGAGCTGCCCATGAGAAATTTACAGGAAATCC
TGCATGGCGCCGTGCGGTTCAGCAACAACCCTGCCCTGTGCAACGTGGAGAGCATCCAGT
GGCGGGACATAGTCAGCAGTGACTTTCTCAGCAACATGTCGATGGACTTCCAGAACCACCT
GGGCAGCTGCCAAAAGTGTGATCCAAGCTGTCCCAATGGGAGCTGCTGGGGTGCAGGAGA
GGAGAACTGCCAGAAACTGACCAAAATCATCTGTGCCCAGCAGTGCTCCGGGCGCTGCCG
TGGCAAGTCCCCCAGTGACTGCTGCCACAACCAGTGTGCTGCAGGCTGCACAGGCCCCCG
GGAGAGCGACTGCCTGGTCTGCCGCAAATTCCGAGACGAAGCCACGTGCAAGGACACCTG
CCCCCCACTCATGCTCTACAACCCCACCACGTACCAGATGGATGTGAACCCCGAGGGCAAA
TACAGCTTTGGTGCCACCTGCGTGAAGAAGTGTCCCCGTAATTATGTGGTGACAGATCACG
GCTCGTGCGTCCGAGCCTGTGGGGCCGACAGCTATGAGATGGAGGAAGACGGCGTCCGC
AAGTGTAAGAAGTGCGAAGGGCCTTGCCGCAAAGTGTGTAACGGAATAGGTATTGGTGAAT
TTAAAGACTCACTCTCCATAAATGCTACGAATATTAAACACTTCAAAAACTGCACCTCCATCA
GTGGCGATCTCCACATCCTGCCGGTGGCATTTAGGGGTGACTCCTTCACACATACTCCTCC
TCTGGATCCACAGGAACTGGATATTCTGAAAACCGTAAAGGAAATCACAGGGTTTTTGCTGA
TTCAGGCTTGGCCTGAAAACAGGACGGACCTCCATGCCTTTGAGAACCTAGAAATCATACG
CGGCAGGACCAAGCAACATGGTCAGTTTTCTCTTGCAGTCGTCAGCCTGAACATAACATCCT
TGGGATTACGCTCCCTCAAGGAGATAAGTGATGGAGATGTGATAATTTCAGGAAACAAAAAT
TTGTGCTATGCAAATACAATAAACTGGAAAAAACTGTTTGGGACCTCCGGTCAGAAAACCAA
AATTATAAGCAACAGAGGTGAAAACAGCTGCAAGGCCACAGGCCAGGTCTGCCATGCCTTG
TGCTCCCCCGAGGGCTGCTGGGGCCCGGAGCCCAGGGACTGCGTCTCTTGCCGGAATGT
CAGCCGAGGCAGGGAATGCGTGGACAAGTGCAACCTTCTGGAGGGTGAGCCAAGGGAGTT
TGTGGAGAACTCTGAGTGCATACAGTGCCACCCAGAGTGCCTGCCTCAGGCCATGAACATC
ACCTGCACAGGACGGGGACCAGACAACTGTATCCAGTGTGCCCACTACATTGACGGCCCC
CACTGCGTCAAGACCTGCCCGGCAGGAGTCATGGGAGAAAACAACACCCTGGTCTGGAAG
TACGCAGACGCCGGCCATGTGTGCCACCTGTGCCATCCAAACTGCACCTACGGATGCACTG
GGCCAGGTCTTGAAGGCTGTCCAACGAATGGGCCTAAGATCCCGTCCATCGCCACTGGGA
TGGTGGGGGCCCTCCTCTTGCTGCTGGTGGTGGCCCTGGGGATCGGCCTCTTCATGCGAA
GGCGCCACATCGTTCGGAAGCGCACGCTGCGGAGGCTGCTGCAGGAGAGGGAGCTTGTG
GAGCCTCTTACACCCAGTGGAGAAGCTCCCAACCAAGCTCTCTTGAGGATCTTGAAGGAAA
CTGAATTCAAAAAGATCAAAGTGCTGGGCTCCGGTGCGTTCGGCACGGTGTATAAGGGACT
CTGGATCCCAGAAGGTGAGAAAGTTAAAATTCCCGTCGCTATCAAGGAATTAAGAGAAGCAA
CATCTCCGAAAGCCAACAAGGAAATCCTCGATGAAGCCTACGTGATGGCCAGCGTGGACAA
CCCCCACGTGTGCCGCCTGCTGGGCATCTGCCTCACCTCCACCGTGCAGCTCATCACGCA
GCTCATGCCCTTCGGCTGCCTCCTGGACTATGTCCGGGAACACAAAGACAATATTGGCTCC
CAGTACCTGCTCAACTGGTGTGTGCAGATCGCAAAGGGCATGAACTACTTGGAGGACCGTC
GCTTGGTGCACCGCGACCTGGCAGCCAGGAACGTACTGGTGAAAACACCGCAGCATGTCA
AGATCACAGATTTTGGGC(T)GGCCAAACTGCTGGGTGCGGAAGAGAAAGAATACCATGCAG
AAGGAGGCAAAGTGCCTATCAAGTGGATGGCATTGGAATCAATTTTACACAGAATCTATACC
CACCAGAGTGATGTCTGGAGCTACGGGGTGACTGTTTGGGAGTTGATGACCTTTGGATCCA
AGCCATATGACGGAATCCCTGCCAGCGAGATCTCCTCCATCCTGGAGAAAGGAGAACGCCT
CCCTCAGCCACCCATATGTACCATCGATGTCTACATGATCATGGTCAAGTGCTGGATGATAG

ACGCAGATAGTCGCCCAAAGTTCCGTGAGTTGATCATCGAATTCTCCAAAATGGCCCGAGA
CCCCCAGCGCTACCTTGTCATTCAGGGGGATGAAAGAATGCATTTGCCAAGTCCTACAGAC
TCCAACTTCTACCGTGCCCTGATGGATGAAGAAGACATGGACGACGTGGTGGATGCCGACG
AGTACCTCATCCCACAGCAGGGCTTCTTCAGCAGCCCCTCCACGTCACGGACTCCCCTCCT
GAGCTCTCTGAGTGCAACCAGCAACAATTCCACCGTGGCTTGCATTGATAGAAATGGGCTG
CAAAGCTGTCCCATCAAGGAAGACAGCTTCTTGCAGCGATACAGCTCAGACCCCACAGGCG
CCTTGACTGAGGACAGCATAGACGACACCTTCCTCCCAGTGCCTGAATACATAAACCAGTC
CGTTCCCAAAAGGCCCGCTGGCTCTGTGCAGAATCCTGTCTATCACAATCAGCCTCTGAAC
CCCGCGCCCAGCAGAGACCCACACTACCAGGACCCCCACAGCACTGCAGTGGGCAACCCC
GAGTATCTCAACACTGTCCAGCCCACCTGTGTCAACAGCACATTCGACAGCCCTGCCCACT
GGGCCCAGAAAGGCAGCCACCAAATTAGCCTGGACAACCCTGACTACCAGCAGGACTTCTT
TCCCAAGGAAGCCAAGCCAAATGGCATCTTTAAGGGCTCCACAGCTGAAAATGCAGAATAC
CTAAGGGTCGCGCCACAAAGCAGTGAATTTATTGGAGCATGA

Right donor probe *EGFR* 3.1 (mutant)
GTTACAGACGACTCCCACAGTCCGTTTGGCC(C)GCCCAAAATGGACT

Right donor probe *EGFR* 3.1 (wild type)
GTTACAGACGACTCCCACAGTCCGTTTGGCC(A)GCCCAAAATGGACT

Left acceptor probe *EGFR* 3.0 (wild type)
TAGGTGTTTGGCC(A)GCCCAAAATACCTACCTCGTAAATCCTCATCAATCATC

Right acceptor probe *EGFR* 3.0
CTCCTTCTGCATGGTATTCTTTCTCTTCCGCACCCAGCAGGACTGTGGGAGTCGTCTGTAACTA
CTTCATGTTACAGACGACTCCCAC

Left acceptor probe *EGFR* 3.0
CCTCGTAAATCCTCATCAATCATCCAGTAAACCGCCGATGATTGATGAGGATTTAGGAGGTAG
GTCTGTGATCTTGACATGCTGCGGTGT

Passive right probe *EGFR* 3.1 mutant

The sequence on the 3´ side of the probe sequence is an initiator for HCR hairpins H3 and H4, and AAAAA serves as a spacer.

GTTTGGCC(C)GCCCAAAATAAAAATACTTCATGTTACAGACGACTCCCAC

Passive left probe *EGFR* 3.0 wild type

The sequence on the 5´ side of the probe sequence is an initiator for HCR hairpins H1 and H2, and AAAAA serves as a spacer.

CCTCGTAAATCCTCATCAATCATCCAGTAAACCGCCAAAAAGTTTGGCC(A)GCCCAAAAT

Right acceptors with a click-chemistry linkage

NNNN…NNNN-Click link-GGACTGTGGGAGTCGTCTGTAACTACTTCATGTTACAGACGACTCCCAC

The N's identify the probe sequences that are complementary to the gray residues in the target shown above. Oligonucleotides corresponding to these probes were synthesized with a 3´-amino

group, pooled, and used to link to a generic right acceptor hairpin, or to a generic passive tag, shown above, via click chemistry.

Right donor probes

GTTACAGACGACTCCCACAGTCCNNN….NNNGGACTG

The N's identify the probe sequences that are complementary to the yellow residues in the target shown above. The right donor probes were synthesized as one continuous molecule by direct synthesis, and then they were pooled together.

Five amp-FISH probe pairs used in the experiments described in Fig. S8

***EGFR*-RA1**
AGCTGTTTTCACCTCTGTTGGGACTGTGGGAGTCGTCTGTAACTACTTCATGTTACAGACGACTCCCAC

***EGFR*-RA2**
AGAAGGTTGCACTTGTCCACGGACTGTGGGAGTCGTCTGTAACTACTTCATGTTACAGACGACTCCCAC

***EGFR*-RA3**
ACTGTATGCACTCAGAGTTCGGACTGTGGGAGTCGTCTGTAACTACTTCATGTTACAGACGACTCCCAC

***EGFR*-RA4**
CACACTGGATACAGTTGTCTGGACTGTGGGAGTCGTCTGTAACTACTTCATGTTACAGACGACTCCCAC

***EGFR*-RA4**
CAGTTTGGATGGCACAGGTGGGACTGTGGGAGTCGTCTGTAACTACTTCATGTTACAGACGACTCCCAC

***EGFR*-right donor probe 1**
GTTACAGACGACTCCCAC AGTCC CTTATAATTTTGGTTTTCTG GGACT

***EGFR*-right donor probe 2**
GTTACAGACGACTCCCAC AGTCC GCATTCCCTGCCTCGGCTGA GGACT

***EGFR*-right donor probe 3**
GTTACAGACGACTCCCAC AGTCC TCCACAAACTCCCTTGGCTC GGACT

***EGFR*-right donor probe 4**
GTTACAGACGACTCCCAC AGTCC GGTCCCCGTCCTGTGCAGGT GGACT

***EGFR*-right donor probe 5**
GTTACAGACGACTCCCAC AGTCC GCACACATGGCCGGCGTCTG GGACT

EGFR 32 sm-FISH probes

The probes in this set were complementary to the 32 underlined sequences that are situated towards the 3´ end of the *EGFR* mRNA sequence listed above.

IFNγ Probes

The sequence of the IFNγ transcript is shown below, where the target regions of the various probes are shown by different emphases. Gray residues are the targets of the right acceptor probes, yellow residues are the targets of the right donor probes, and the underlined residues are the targets of the 48 directly labeled probes (where the underlined region is longer than 25 nucleotides, and the junction between the yellow and gray residues represents the break point between adjacent probes). The probes in this set were used for the experiments shown in Fig. 5.

GTGCAGCACATTGTTCTGATCATCTGAAGATCAGCTATTAGAAGAGAAAGATCAGTTAAGTC
CTTTGGACCTGATCAGCTTGATACAAGAACTACTGATTTCAACTTCTTTGGCTTAATTCTCTC
GGAAACGATGAAATATACAAGTTATATCTTGGCTTTTCAGCTCTGCATCGTTTTGGGTTCTC
TTGGCTGTTACTGCCAGGACCCATATGTAAAAGAAGCAGAAAACCTTAAGAAATATTTTAAT
GCAGGTCATTCAGATGTAGCGGATAATGGAACTCTTTTCTTAGGCATTTTGAAGAATTGGAA
AGAGGAGAGTGACAGAAAAATAATGCAGAGCCAAATTGTCTCCTTTTACTTCAAACTTTTTA
AAAACTTTAAAGATGACCAGAGCATCCAAAAGAGTGTGGAGACCATCAAGGAAGACATGAA
TGTCAAGTTTTTCAATAGCAACAAAAAGAAACGAGATGACTTCGAAAAGCTGACTAATTATT
CGGTAACTGACTTGAATGTCCAACGCAAAGCAATACATGAACTCATCCAAGTGATGGCTGA
ACTGTCGCCAGCAGCTAAAACAGGGAAGCGAAAAAGGAGTCAGATGCTGTTTCGAGGTCG
AAGAGCATCCCAGTAATGGTTGTCCTGCCTGCAATATTTGAATTTTAAATCTAAATCTATTTA
TTAATATTTAACATTATTTATATGGGGAATATATTTTTAGACTCATCAATCAAATAAGTATTTA
TAATAGCAACTTTTGTGTAATGAAAATGAATATCTATTAATATATGTATTATTTATAATTCCTA
TATCCTGTGACTGTCTCACTTAATCCTTTGTTTTCTGACTAATTAGGCAAGGCTATGTGATTA
CAAGGCTTTATCTCAGGGGCCAACTAGGCAGCCAACCTAAGCAAGATCCCATGGGTTGTG
TGTTTATTTCACTTGATGATACAATGAACACTTATAAGTGAAGTGATACTATCCAGTTACTGC
CGGTTTGAAAATATGCCTGCAATCTGAGCCAGTGCTTTAATGGCATGTCAGACAGAACTTG
AATGTGTCAGGTGACCCTGATGAAAACATAGCATCTCAGGAGATTTCATGCCTGGTGCTTC
CAAATATTGTTGACAACTGTGACTGTACCCAAATGGAAAGTAACTCATTTGTTAAAATTATCA
ATATCTAATATATATGAATAAAGTGTAAGTTCACAACTA

The right acceptor probes with click links and the right donor probes for IFNγ mRNA were synthesized using the same strategy, and their general structures are the same as described above for *EGFR* mRNA.

Primers for reverse transcription and SuperSelective PCR assays

These primers were used for the experiments shown in Fig. S4.

The outer reverse conventional primer was used both for the reverse transcription reactions and for real-time PCR amplifications (as the conventional reverse primer in combination with a SuperSelective primer).

Outer forward conventional primer
AGGAACGTACTGGTGAAAACACCGCAGCAT

Outer reverse conventional primer
TCTGCATGGTATTCTTTCTCTTC

Wild-type SuperSelective primer (underlines depict the anchor sequence and the discriminating nucleotide in the sequences below)
AGGAACGTACTGGTGAAAACACCGCAGCATAAGAATCTACCGACACCAC TTTTGGGCT

Mutant SuperSelective primer
AGGAACGTACTGGTGAAAACACCGCAGCATAAGAATCTACCGACACCAC TTTTGGGCG

Supplementary Figures and their Legends

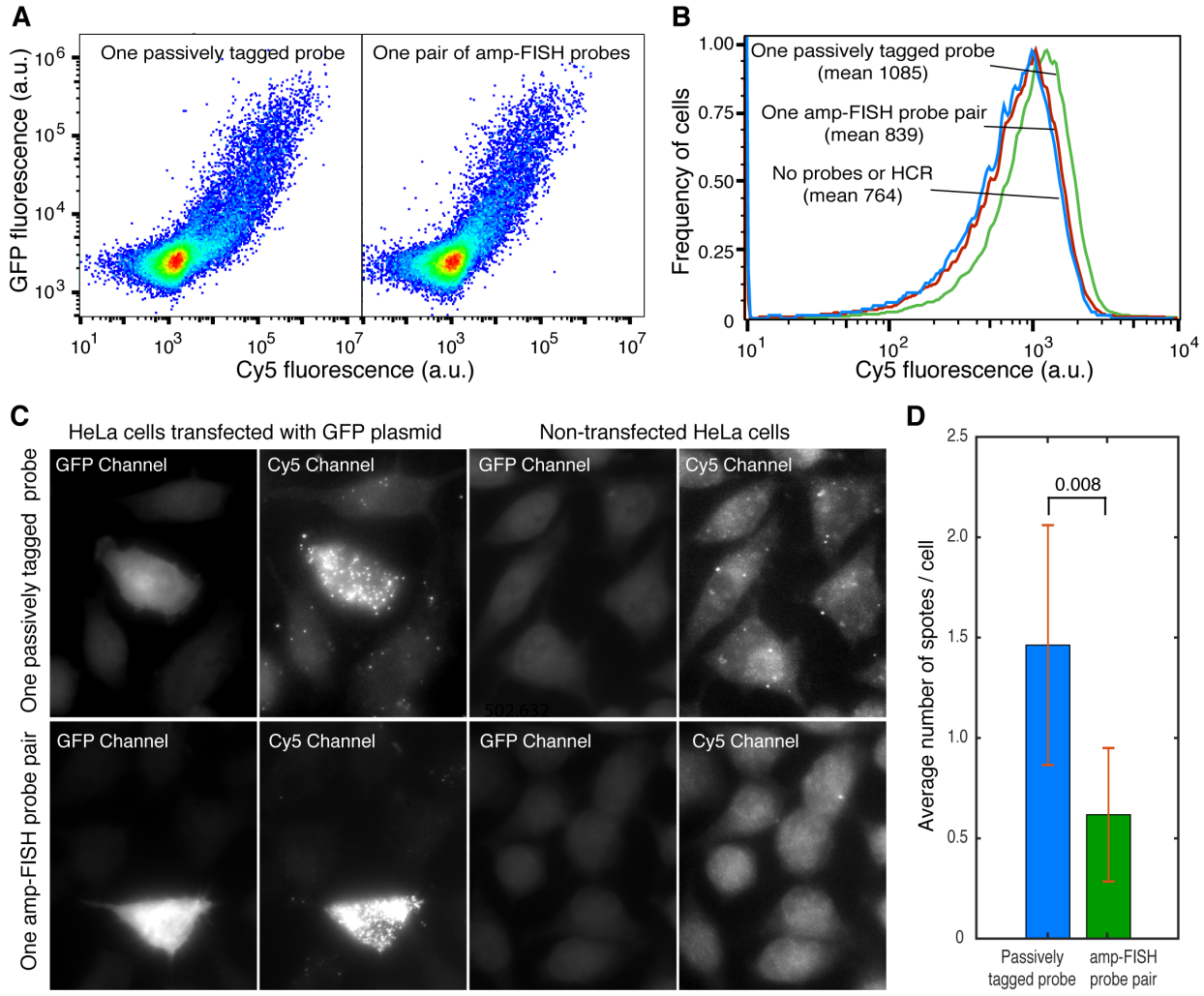

Fig. S1. Comparison of HCR signals and background signals generated with either a pair of high-fidelity amp-FISH probes or with passively tagged probes, both of which are specific for the same target site in mRNAs generated by GFP plasmids transfected into HeLa cells. (*A*) Flow cytometric analysis of transfected HeLa cells, hybridized with either amp-FISH probes or passively tagged probes and then subjected to HCR, demonstrating that the resulting target-specific signal intensities are quite similar. (*B*) Flow cytometry analysis of non-transfected HeLa cells hybridized with these two kinds of probes and subjected to HCR. The data, presented as a histogram for better comparison, demonstrate that passively tagged probes produce background signals, whereas the background that is produced from amp-FISH probes cannot be distinguished from the

background that is produced when no probes were used at all. (*C*) Image-based analysis of signals and backgrounds. The four top panels show images of transfected and non-transfected cells that were hybridized with passively tagged probes. Non-specific background signals arose in the Cy5 channel of the non-transfected cells. On the other hand, the four lower images of transfected and non-transfected cells that were hybridized with amp-FISH probes, showed only one background spot in a non-transfected cell. *(D)* Average number of spots observed in non-transfected cells using the two kinds of probes. Error bars represent 95% confidence intervals. The p-value shown (0.008) is the probability of obtaining at least as large a difference in the means of two populations as indicated by the bars by random chance alone. It was calculated by bootstrapping, using a custom MATLAB script. The number of spots were measured from 80 cells in each category.

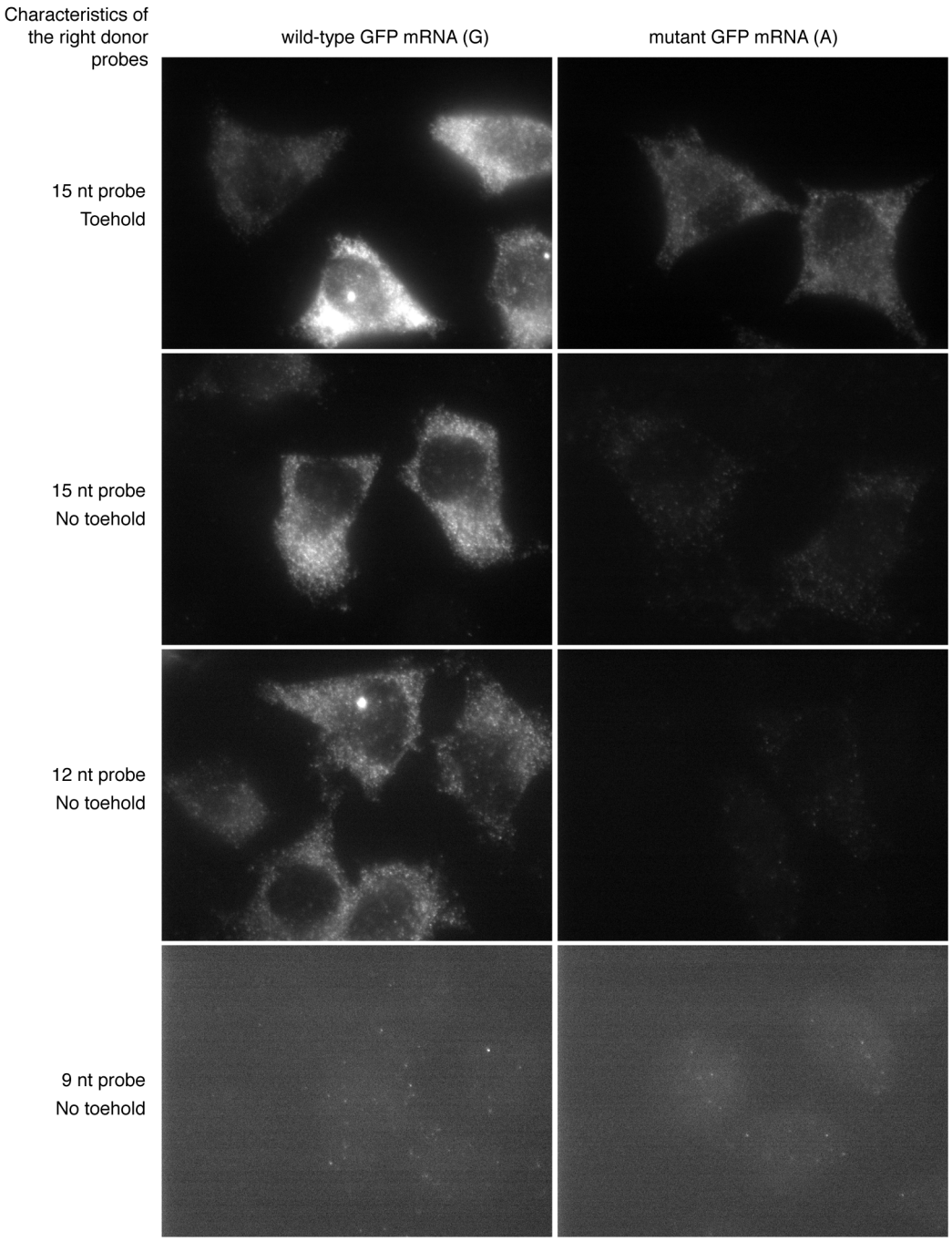

Fig. S2. Optimization of the design of donor probes for discrimination of an SNV within GFP mRNA. HeLa cells transfected with either a wild-type or a mutant version of a plasmid encoding GFP were probed at 37 ºC using a pair of amp-FISH probes designed to detect GFP wild-type mRNA, followed by HCR with hairpins H3 and H4 labeled with Cy5. All images in the figure are compressed z-stacks acquired in the Cy5 channel and they

show two or three cells expressing GFP mRNA. The length of the target-specific hairpin loop of the right donor probe was varied (to be either 15, 12, or 9 nucleotides in length), but the target-specific sequence of the corresponding right acceptor probes was kept at a length of 22 nucleotides. Two different pairs of amp-FISH probes, both of whose donors contained a 15-nucleotide-long hairpin loop were tested. One pair contained toehold sequences to initiate the interaction between the donor and the acceptor (as shown in Fig. 1), while the other pair lacked toehold sequences, and the interaction between the donor and the acceptor was based on the arm of the donor (freed up as a result of the hybridization of the donor to the target mRNA) displacing the arm of the hairpin in the acceptor due to their close proximity. Like the 15-nucleotide-long donor that lacked a toehold sequence, the donors that contained shorter hairpin loops (12 or 9 nucleotides in length) also did not possess toehold sequences. The bottom two images in this figure are displayed at a higher contrast in order to reveal the few spots that are present. A comparison of the resulting images shows that 15-nucleotide-long target-specific hairpin loops in donors that lack toehold sequences are better at discriminating wild-type mRNAs than are corresponding donors that have the same length hairpin loops and do possess toehold sequences. Moreover, a comparison of the resulting images also shows that discrimination improves even more when the length of the target-specific hairpin loop in the donor is decreased to 12 nucleotides. However, a further decrease in the length of the target-specific hairpin loop results in the loss of signals from wild-type mRNAs (as well as from mutant mRNAs). As a consequence of this optimization process, we decided to use donors that possessed 11-nucleotide-long target-specific hairpin loops that lacked toehold sequences for the experiments whose results are shown in Fig. 2*B*.

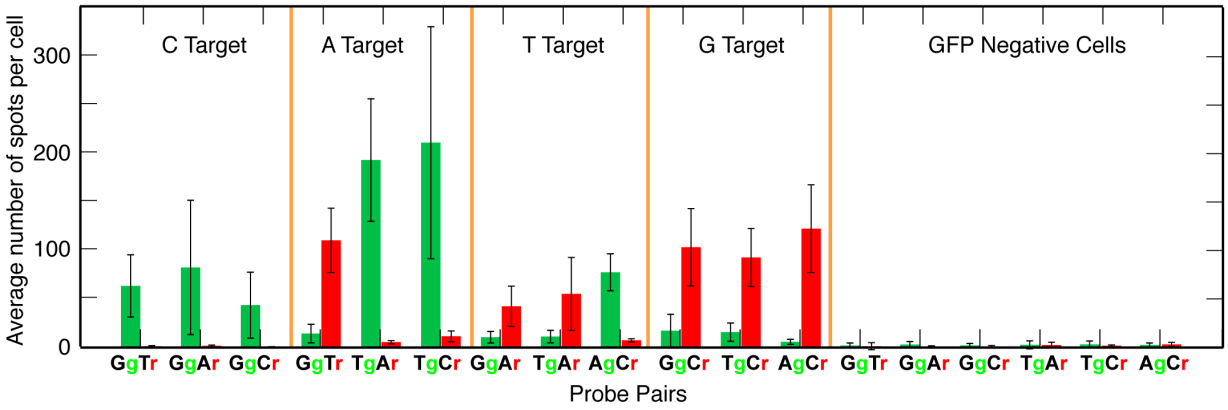

Fig. S3. Average number of TMR (green) and Cy5 (red) HCR spots per cell in HeLa cells transfected with plasmids encoding GFP variants differing as to the identity of a single nucleotide (indicated in each panel). The panel on the right shows the results obtained with non-transfected cells. The allele-discriminating donor probe pairs that were used are indicated at the bottom along with the color of the HCR product that each probe in the pair was expected to produce. Out of the six possible probe pairs, the only three that were used included one probe that was perfectly complementary to the target. Another version of this data is presented in Fig. 2*C*, in which the relative percentages of spots in each color was determined in single cells, and then averaged over the population of cells. The number of spots averaged across six probe pairs and four mutations was 107 ± 31 spots for GFP-positive cells and 3 ± 0.9 spots for GFP-negative cells.

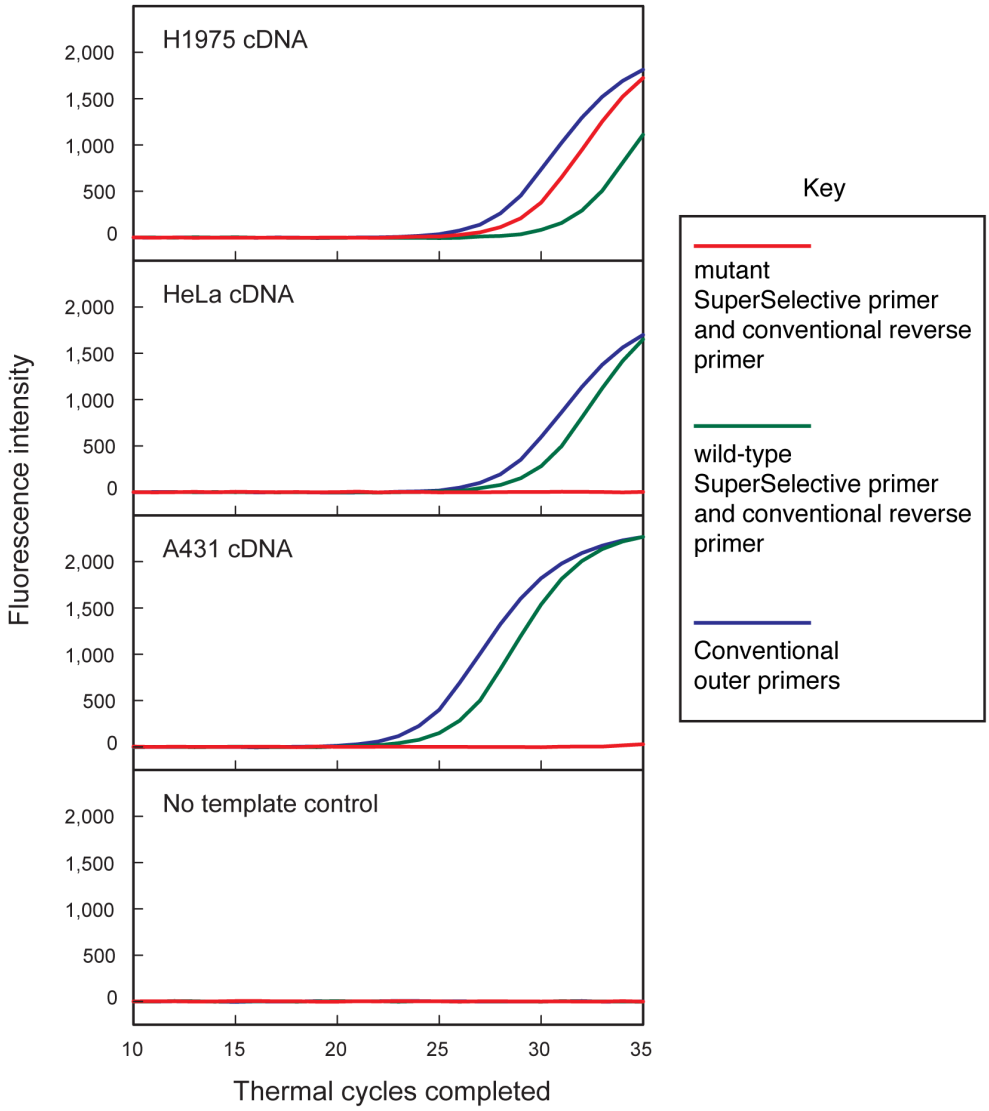

Fig. S4. Confirmation of the genotypes of *EGFR* mRNAs with respect to mutation L858R in three different cell lines by real-time reverse transcription PCR. Conventional primers initiate the amplification of both wild-type and mutant mRNAs, because they are designed to bind to either side of the site containing the mutation (outer primers). The mutant SuperSelective primer only initiates amplification on mRNAs possessing the *EGFR* L858R mutation, and the wild-type SuperSelective primer only initiates amplification on mRNAs possessing the *EGFR* wild-type sequence. The results of these assays confirm that the H1975 cell line is heterozygous for the *EGFR* L858R mutation, whereas, the HeLa and A431 cell lines are pure wild-type with respect to this mutation.

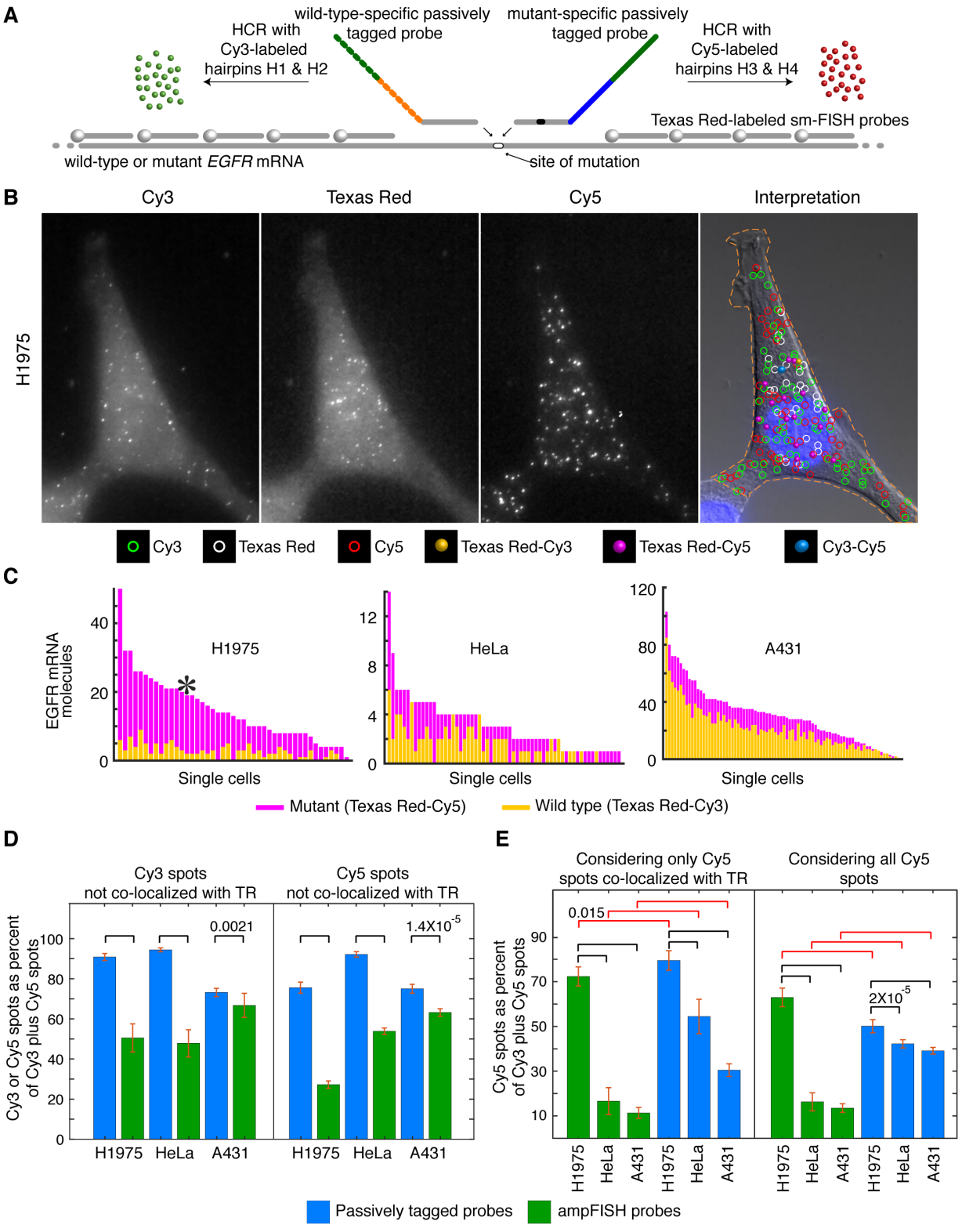
A
wild-type-specific passively tagged probe
mutant-specific passively tagged probe
HCR with Cy3-labeled hairpins H1 & H2
HCR with Cy5-labeled hairpins H3 & H4
Texas Red-labeled sm-FISH probes
wild-type or mutant EGFR mRNA
site of mutation
B
Cy3
Texas Red
Cy5
Interpretation
H1975
Cy3
Texas Red
Cy5
Texas Red-Cy3
Texas Red-Cy5
Cy3-Cy5
C
EGFR mRNA molecules
H1975
HeLa
A431
Single cells
Mutant (Texas Red-Cy5)
Wild type (Texas Red-Cy3)
D
Cy3 spots not co-localized with TR
Cy5 spots not co-localized with TR
Cy3 or Cy5 spots as percent of Cy3 plus Cy5 spots
0.0021
1.4X10-5
H1975
HeLa
A431
E
Considering only Cy5 spots co-localized with TR
Considering all Cy5 spots
Cy5 spots as percent of Cy3 plus Cy5 spots
0.015
2X10-5
Passively tagged probes
ampFISH probes

Fig. S5. Demonstration that mutations are detected with higher fidelity using interacting hairpin amp-FISH probes than with linear passively tagged probes. (*A*) Schematic depiction of a pair of wild-type-specific and mutant-specific passively tagged probes whose target-specific portions were of the same length and sequence as the target-specific portions of the left donor and right donor probes in the assay described in Fig. 3, but that lacked hairpin stems, and that were used in the absence of arm-acceptor hairpin probes. The wild-type-specific probe was tagged with an HCR initiator that elicited signals from Cy3-labeled H1 and H2 HCR hairpins, and the mutant-specific probe was tagged with an HCR initiator that elicited signals from Cy5-labeled H3 and H4 HCR hairpins. The probe mixture also contained a set of 48 sm-FISH probes labeled with Texas Red (TR). (*B*) A single H1975 cell imaged in three different channels with this set of control probes. All *EGFR* mRNA molecules present in the cell were expected to be visible in the TR image, whereas, *EGFR* mRNA molecules possessing the wild-type sequence were expected to also be visible in the Cy3 image and mRNA molecules possessing the mutant sequence were expected to also be visible in the Cy5 image. Those Cy3 and Cy5 spots that are not also visible in the TR image are likely to have arisen from non-specific sources. The three images on the left of the figure are merged z-stacks in the indicated color channel contrasted to display the spot-like signals that each kind of probe created. The panel on the right shows a DIC-DAPI image overlaid with the locations of spots that were detected in a single channel, or in various combinations of the three channels identified by the codes shown below the panels. (*C*) Histograms showing the number of TR-positive *EGFR* mRNA molecules that were identified as mutant (purple) or wild type (yellow) in each analyzed cell (sorted in order of the sum of the two kinds of spots in the cell). The cell line H1975 expresses both mutant and wild-type RNAs whereas, whereas, the other two cell lines express only the wild-type version. The bar with an asterisk on top represents the cell shown in *B*. (*D*) Passively tagged probes yield more spots that are not co-localized with TR spots than amp-FISH probes. These extra spots likely arise from nonspecific sources, and they reduce the accuracy of the assays. (*E*) Comparison of the performance of passively tagged probes and amp-FISH probes for discrimination between the heterozygote cell line (H1975) and the homozygous wild-type cell lines (HeLa and A431). Cy5 spots as a percent of Cy5 and Cy3 spots combined are plotted. The data in the left panel is for those Cy5 and Cy3 spots that are co-localized with TR spots, whereas,

the data in the right panel is for all Cy5 and Cy3 spots, whether they are co-localized with TR, or not. The left panel shows that passively tagged probes create more nonspecific spots than the amp-FISH probes (the first three bars represent the same data as in Fig. 3*C*). This degrades the quality of discrimination between heterozygote mutant and wild-type cells, particularly when co-localization with TR is not used as an aid to identify target-bound probes (right panel). The bars in the pairs denoted by the horizontal brackets were compared to each other to determine the probability that the differences in their values can be obtained by chance alone (p-values). P-values were $<1 \times 10^{-5}$, unless indicated otherwise. The red brackets denote comparisons between passively tagged probes and amp-FISH probes, whereas the black brackets denote comparisons between different cell lines using the same kind of probes.

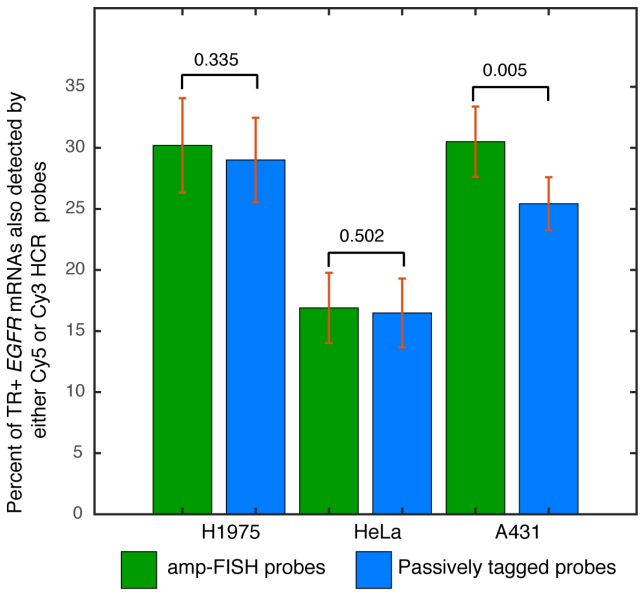

Fig. S6. Efficiency of HCR signal generation by an amp-FISH probe pair is similar to an analogous passively tagged probe designed to bind to the same site. This data is derived from the *EGFR* detection experiment described in Figs. 3 and S5. The efficiency of detection is the fraction of TR-positive *EGFR* spots that are also detected by either Cy5 or Cy3 amp-FISH probes. Although 16-28% of the *EGFR* mRNAs are detected by each of the two kinds of probes, there was no significant difference in their detection efficiencies (except for cell line A431, which may result from the crowding of spots due to the high expression in this cell line), indicating that the probability of HCR initiation by an amp-FISH probe pair bound to the target is similar to the probability of HCR initiation the passively tagged probe. If a fraction of amp-FISH probes were able to bind but failed to progress to the release of the HCR initiator due to unfavorable energetics, a smaller fraction of *EGFR* RNA would have been detected by them than by the passively tagged probes.

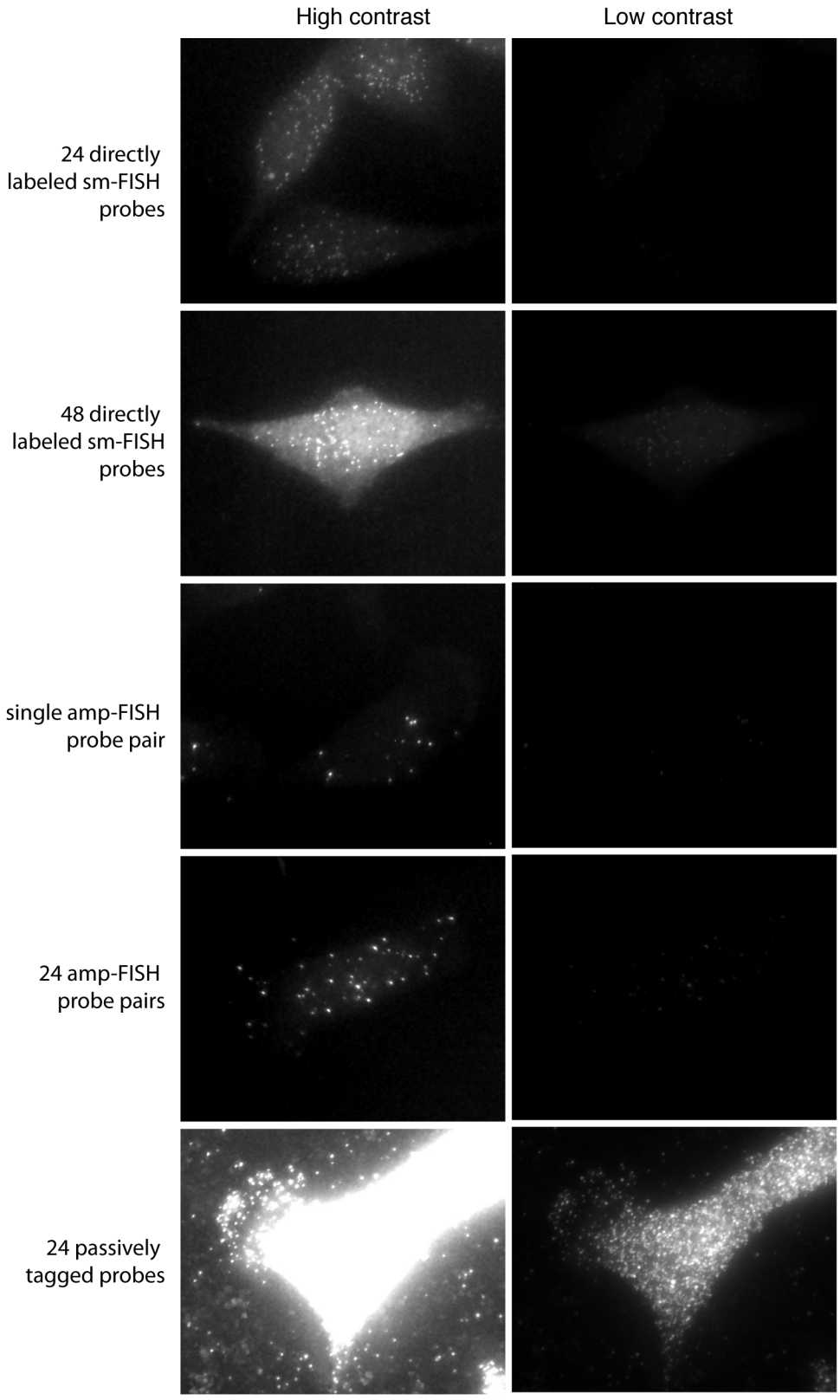
High contrast
Low contrast
24 directly labeled sm-FISH probes
48 directly labeled sm-FISH probes
single amp-FISH probe pair
24 amp-FISH probe pairs
24 passively tagged probes

Fig. S7. Comparison of the number and intensity of *EGFR* mRNA spots created by different probe sets in H1975 cells. For each of the five probe sets, the same image is presented at two different levels of contrast. The images on the left are displayed utilizing a common contrast that is higher than the common contrast used to display the images on the right. The two directly labeled probe sets are expected to illuminate all *EGFR* mRNA molecules in the cell, while the single amp-FISH probe pair should only light up the portion of *EGFR* mRNAs that possess the wild-type version of the L858R mutation.

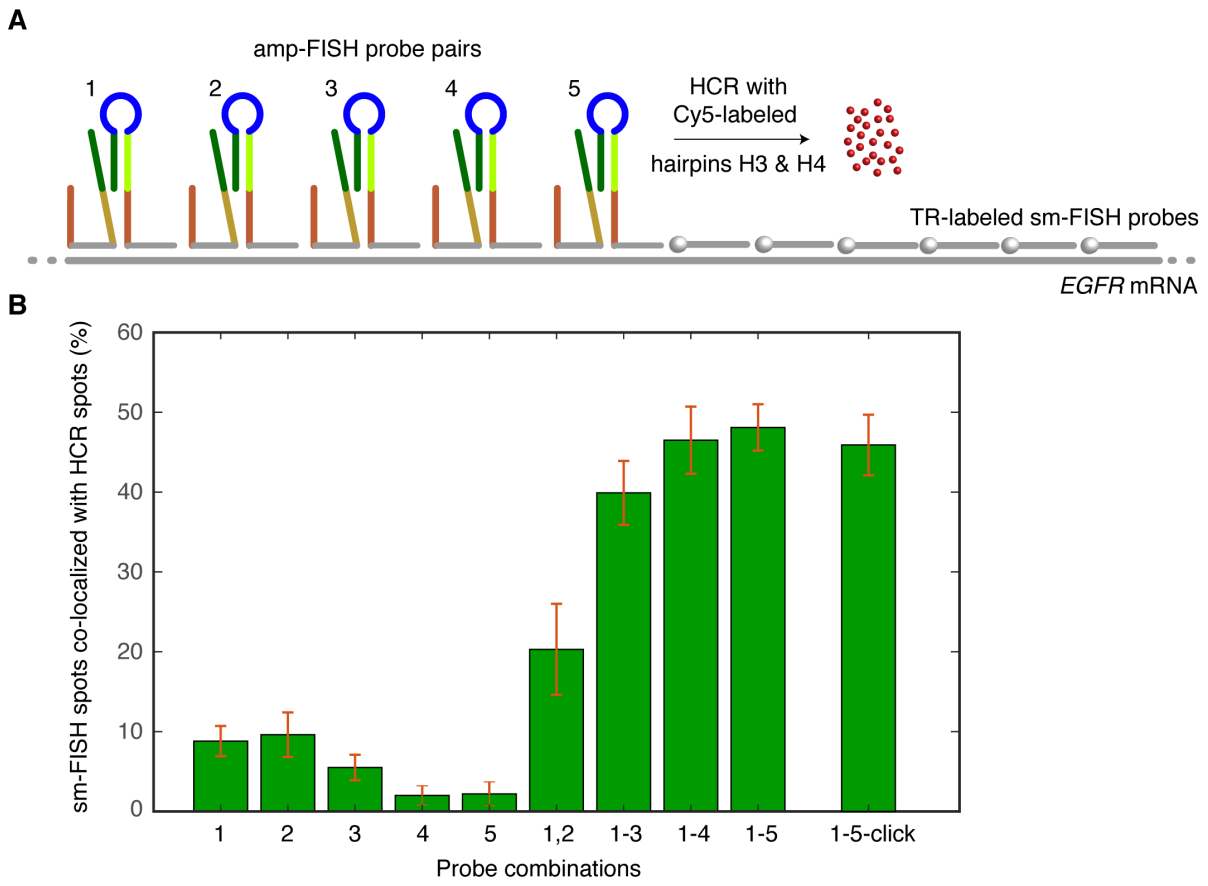

Fig. S8. Efficiency of detection of *EGFR* mRNAs in HeLa cells utilizing different amp-FISH probe pairs singly and in combinations, and the impact on detection efficiency of a "click chemistry" link between the right acceptor hairpin and the right probe sequence. (*A*) Five pairs of amp-FISH probes designed to hybridize to a 5´ region of *EGFR* mRNA (each signaling in Cy5) were hybridized to HeLa cells, along with a set of 32 sm-FISH probes (each labeled with Texas Red), which were designed to hybridize to a different region of *EGFR* mRNA. A two-step hybridization protocol was performed: first, the right acceptor probes were hybridized, and then after the removal of excess donor probes, the right donor probes were hybridized. The sm-FISH probes were present during each step. Hybridization was followed by HCR, imaging in both the Texas Red and Cy5 channels, detection of spots in each channel, and identification of spots that fluoresced in just one or both (co-localized) channels. (*B*) Efficiency of detection determined for five different probe pairs, singly or in combinations. Assuming sm-FISH probes are able to identify all *EGFR* mRNA molecules in the cells, the fraction of Texas Red-positive sm-FISH spots that co-

localize with Cy5-positive HCR spots reflects the efficiency of detection of this mRNA by a given amp-FISH probe pair or combination. 3 ng of each probe in each pair was used for each hybridization reaction. See Supplementary Table S2 for the impact of the concentrations of the probes. All the data, except the data shown in the right-most bar, was obtained by employing right acceptor probes produced by automated DNA synthesis. On the other hand, the right acceptor probes employed for the data on the right-most bar were constructed by joining the acceptor hairpin with a target-specific sequence by click chemistry. The error bars represent the 95% confidence interval obtained by analysis of 35 to 90 cells.

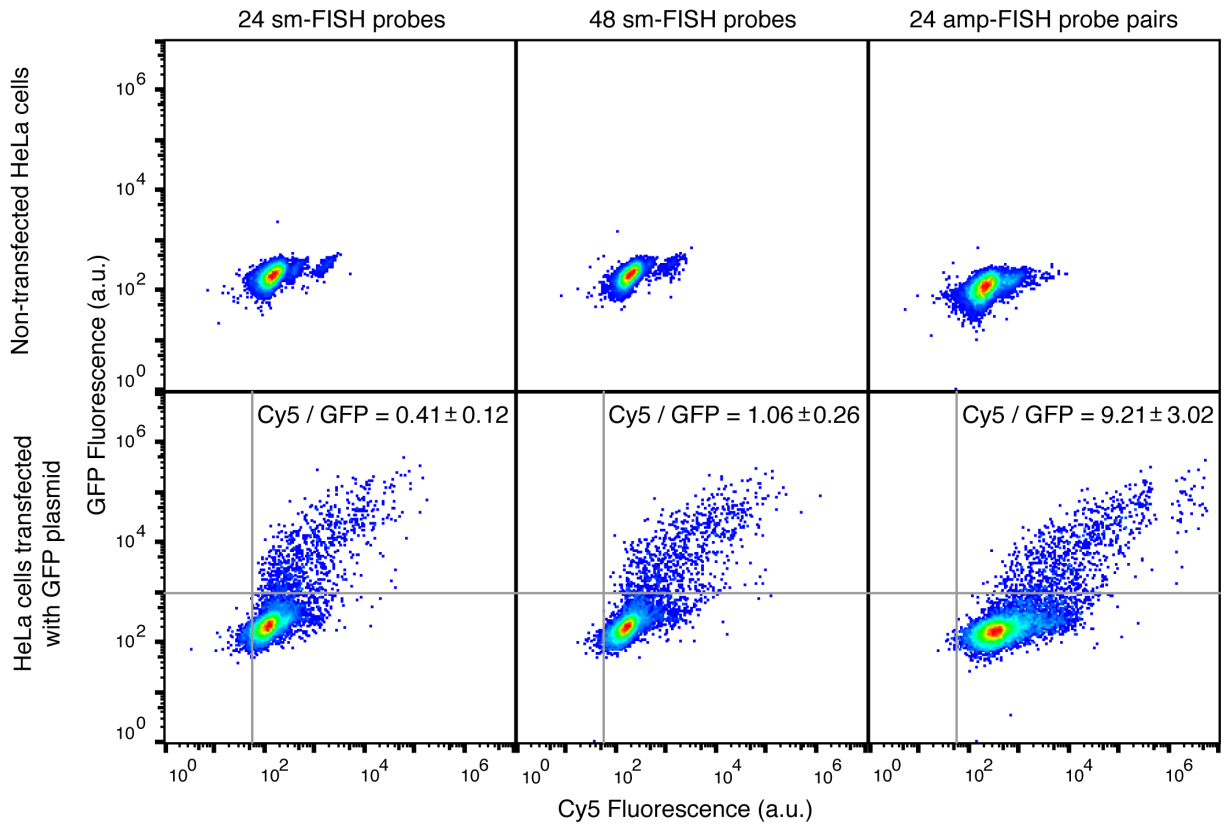

Fig. S9. Enhancement in flow cytometric signal intensity obtained through the use of amp-FISH probes compared to the use of sm-FISH probes. GFP mRNA was probed utilizing the indicated numbers of sm-FISH probes and amp-FISH probe pairs in HeLa cells, with and without transfection of a plasmid encoding GFP. The cells were then analyzed in the Cy5 and GFP channels in a flow cytometer. The average ratios of Cy5 to GFP fluorescence from single GFP-positive cells (that appear in the upper right quadrants) are listed in the bottom panels. These measurements indicate that when the amp-FISH probes were used, the signal intensity increased 22-fold compared to the signal intensity when the 24-probe sm-FISH set was used, and 9-fold compared to the signal intensity when the 48-probe sm-FISH set was used. The background signal intensity remained the same. The errors in these measurements represent the 95% confidence interval obtained from the Cy5/GFP fluorescence ratios, measured from the 770 to 850 cells that were present in the upper right quadrants.

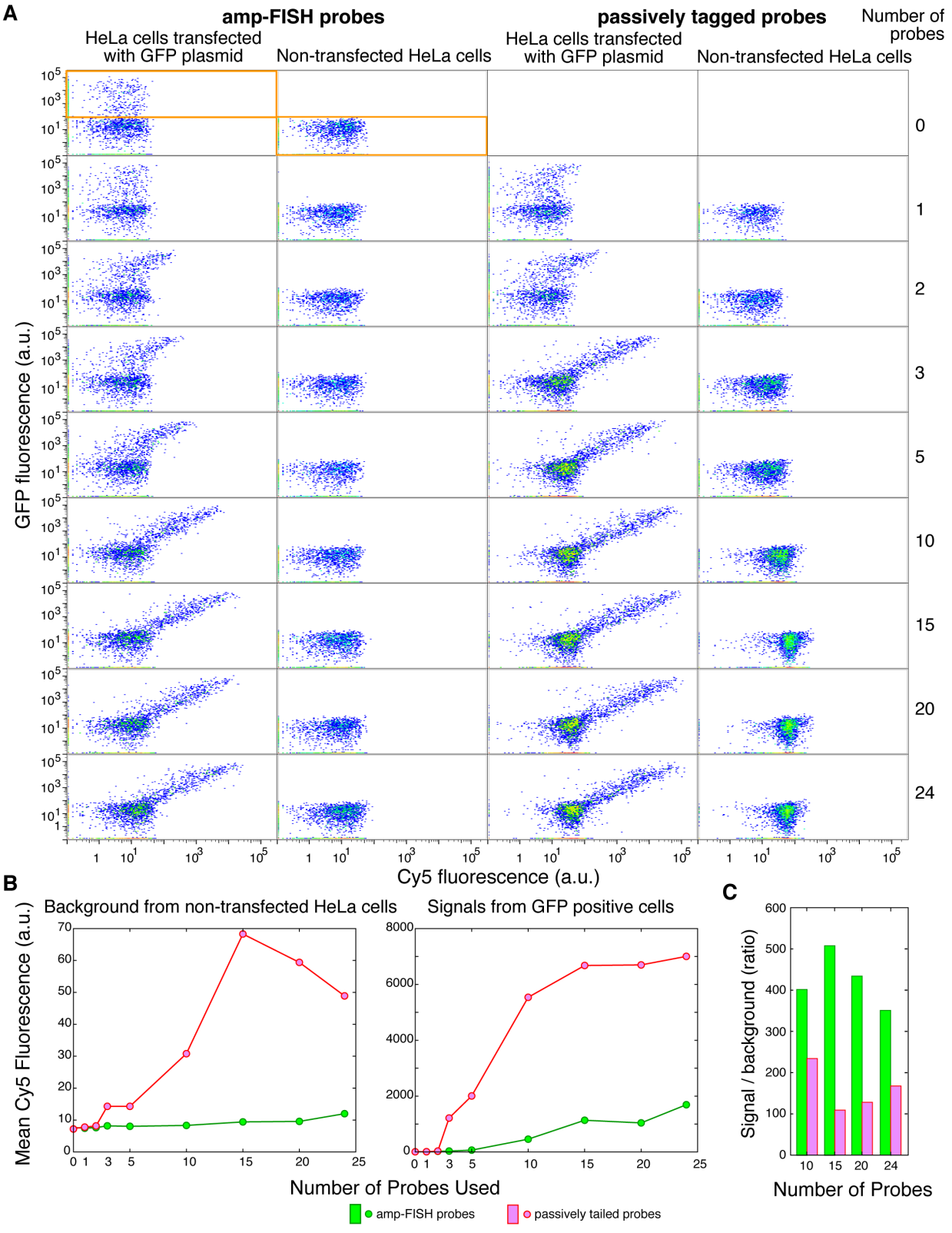
A
amp-FISH probes
passively tagged probes
Number of probes
HeLa cells transfected with GFP plasmid
Non-transfected HeLa cells
HeLa cells transfected with GFP plasmid
Non-transfected HeLa cells
GFP fluorescence (a.u.)
Cy5 fluorescence (a.u.)
0
1
2
3
5
10
15
20
24
B
Background from non-transfected HeLa cells
Signals from GFP positive cells
Mean Cy5 Fluorescence (a.u.)
Number of Probes Used
amp-FISH probes
passively tailed probes
C
Signal / background (ratio)
Number of Probes

Fig. S10. Changes in the mRNA specific and background signals as the number of amp-FISH probe pairs and passively tagged probes are increased. (*A*) HeLa cells transfected with a plasmid encoding GFP and non-transfected HeLa cells were probed with various combinations of up to 24-pairs of amp-FISH probes pairs and 24 passively tagged probes all designed to recognize GFP mRNA. After hybridization with the primary probes, signals were developed using Cy5-labeled HCR hairpins H3 and H4 and the cells were analyzed by flow cytometry in the GFP and Cy5 channels. For each set of probes, we used all probes up to the number indicated at the right margin starting from the 5' end of GFP mRNA (as demarcated in Dataset S1). A shift of GFP-positive cells to the right compared to no-probe controls (top two panels) on the scatter diagram signifies a staining of GFP mRNA, and a shift of the non-transfected cells to the right is indicative of the generation of background signals. (*B*) Changes in the average Cy5 fluorescence generated by non-transfected HeLa cells (background) and by GFP positive cells (signals) as the number of probes in the pools increased. The average Cy5 fluorescence was calculated in each case from cell populations that fall within the orange rectangular zones corresponding to the ones shown in the top panels in *A*. As the number of probes is increased, an increase in the signals is observed. However, the signals appear to saturate with both kinds of probes when the number of probes reaches 10 to 15. The reasons for this saturation are not clear, but it may arise from the local scarcity of HCR-substrate hairpins or may be due to limits on HCR polymerization imposed by the crosslinked cellular matrix. (*C*) Signal-to-background ratios for both kinds of probe sets for sets containing ten or more probes. Although, the passively tagged probes yield higher signals than the amp-FISH probes, they also produce high background levels (*A* and *B*). Overall amp-FISH probes yield higher signal-to-background ratios than passively tagged probes.

Supplementary Tables

Table S1. Number of spots for *EGFR* mRNAs visible in one or more of three different fluorescence channels in single cells in three different cell lines (relates to Fig. 3).

Classification of Spots	H1975	HeLa	A439
All *EGFR* RNAs (TR, Cy3-TR, Cy5-TR) (molecules)	63.8 ± 7.6	27.2 ± 4.6	121.3 ± 19.9
Detected *EGFR* RNAs (Cy3-TR, Cy5-TR) (molecules)	16.6 ± 2.6	4.3 ± 0.9	33.9 ± 5.8
Mutant fraction (Cy5-TR/detected *EGFR* RNAs) (%)	72.3 ± 4.4	16.3 ± 6.6	11.6 ± 2.1
TR (spots)	47.1 ± 6.4	22.9 ± 3.6	86.8 ± 17.8
Cy3 (spots)	4.6 ± 0.7	3.3 ± 0.6	67.9 ± 12.9
Cy5 (spots)	3.9 ± 0.7	0.7 ± 0.2	10.5 ± 2.6
Cy3-TR (spots)	4.2 ± 0.7	3.7 ± 0.8	30.3 ± 6.5
Cy5-TR (spots)	12.4 ± 2.1	0.5 ± 0.2	3.6 ± 0.8
Cy3-Cy5 (spots)	0.2 ± 0.1	0.0 ± 0.0	1.7 ± 0.6
Cy3-Cy5-TR (spots)	0.1 ± 0.1	0.0 ± 0.0	0.6 ± 0.3

Mutant fraction was calculated for single cells and then averaged. The errors represent the 95% confidence interval obtained by analysis of 40 to 60 cells.

Table S2. Effects of probe concentration on average number of specific (co-localized with sm-FISH spots) and non-specific HCR spots (not co-localized) per cell when five pairs of amp-FISH probes are used for the detection of *EGFR* mRNA (relates to Fig. S8).

Amount of each probe (ng/50 µL reaction)	sm-FISH spots (all *EGFR* molecules)	sm-FISH spots co-localized with HCR spots	HCR spots not co-localized with sm-FISH spots
3.0	21.8 ± 2.9	9.6 ± 1.5	7.2 ± 0.9
1.0	21.1 ± 3.4	6.5 ± 1.1	2.5 ± 0.4
0.3	30.0 ± 5.7	4.7 ± 1.1	1.4 ± 0.4
3.0 donor, but no acceptor	25.4 ± 3.7	0.1 ± 0.1	1.3 ± 0.3

The experimental design is described in the legend of Fig. S5. The errors represent the 95% confidence interval obtained by analyzing 35 to 90 cells.

References to Supplementary Information

1. Choi HM, Beck VA, Pierce NA (2014) Next-generation *in situ* hybridization chain reaction: higher gain, lower cost, greater durability. *ACS Nano* 8:4284-4294.
2. Koos B*,* et al. (2015) Proximity-dependent initiation of hybridization chain reaction. *Nat Commun* 6:7294.
3. Raj A, Tyagi S (2010) Detection of individual endogenous RNA transcripts *in situ* using multiple singly labeled probes. *Methods Enzymol* 472:365-386.
4. Jewett JC, Bertozzi CR (2010) Cu-free click cycloaddition reactions in chemical biology. *Chem Soc Rev* 39:1272-1279.
5. Vargas DY, Kramer FR, Tyagi S, Marras SAE (2016) Multiplex real-time PCR assays that measure the abundance of extremely rare mutations associated with cancer. *PLoS One* 11:e0156546.
6. Bushkin Y*,* et al. (2015) Profiling T cell activation using single-molecule fluorescence *in situ* hybridization and flow cytometry. *J Immunol* 194:836-841.
7. 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* 5:877-879.
8. Batish M, van den Bogaard P, Kramer FR, Tyagi S (2012) Neuronal mRNAs travel singly into dendrites. *Proc Natl Acad Sci USA* 109:4645-4650.
9. Vargas DY*,* et al. (2011) Single-molecule imaging of transcriptionally coupled and uncoupled Splicing. *Cell* 147:1054-1065.
10. Battich N, Stoeger T, Pelkmans L (2013) Image-based transcriptomics in thousands of single human cells at single-molecule resolution. *Nat Methods* 10:1127-1133.
11. Schwanhausser B*, et al.* (2011) Global quantification of mammalian gene expression control. *Nature* 473:337-342.
12. Levesque MJ, Ginart P, Wei Y, Raj A (2013) Visualizing SNVs to quantify allele-specific expression in single cells. *Nat Methods* 10:865-867.
13. Choi HMT*,* et al. (2018) Third-generation *in situ* hybridization chain reaction: multiplexed, quantitative, sensitive, versatile, robust. *Development* 145: dev165753.