Supplementary Information

clampFISH 2.0 enables rapid, scalable amplified RNA detection *in situ*

Contents

Supplementary Methods

Supplementary methods: cell culture and tissue processing

Details of culturing and plating cell lines

The WM989 A6-G3 human melanoma cell line, first described in (Shaffer et al. 2017) was derived from WM989 cells (a gift from the lab of Dr. Meenhard Herlyn) that were twice isolated from a single cell and expanded. WM989 A6-G3 H2B-GFP cells were derived by transducing WM989 A6-G3 cells with 60µL Lenti_EFS (https://benchling.com/s/seq-6Jv3Rmebv1nIevxPfYQ6/edit), isolating a single cell, and expanding this clone (Clone A11). Both lines were cultured in Tu2% media (80% MCDB 153, 10% Leibovitz's L-15, 2% FBS, 2.4mM $CaCl_2$, 50 U/mL penicillin, and 50 µg/mL streptomycin). WM989 A6-G3 RC4 cells were derived by treating WM989 A6-G3 cells with 1µM vemurafenib in Tu2%, isolating a single drug-resistant colony, and culturing these cells in 1µM vemurafenib in Tu2% (Goyal et al. 2021) for several months. All cell lines were passaged with 0.05% trypsin-EDTA (Gibco, 25300120).

For the amplifier screen and pooled amplification experiment, WM989 A6-G3 H2B-GFP and WM989 A6-G3 RC4 cells were mixed together and plated on coverslips (VWR, 16004-098, 24x50mm, No. 1 coverglass) with 24-well silicone isolators (Grace Bio-Labs, 665108). For the readout probe stripping experiment, conventional single-molecule RNA FISH comparison experiment, amplification characterization experiment, and one-pot amplification experiment, we plated WM989 A6-G3 or WM989 A6-G3 RC4 cells into separate wells of an 8-well chambers (Lab-tek, 155411, No. 1 coverglass). For the high-throughput profiling experiment, we plated WM989 A6-G3 cells into 5 wells and WM989 A6-G3 RC4 cells into 1 well of a 6-well plate (Cellvis, P06-1.5H-N, No. 1.5 coverglass), and allowed them to grow out for 6 days (2-3 cell divisions for WM989 A6-G3 cells) before fixation.

Details of tissue experiment sample preparation

For the fresh frozen tissue experiment, we performed clampFISH 2.0 in a melanoma xenograft tumor taken from experiments described in (Torre et al. 2021). Briefly, human WM989-A6-G3-Cas9-5a3 cells (without a genetic knockout), derived by isolating and expanding a single WM989 A6-G3 cell, were injected into 8-week-

old male NOD/SCID mice (Charles River Laboratories) and fed AIN-76A chow (mouse #8947) or AIN-76A chow containing 417mg/kg PLX4720 (mouse #8948). The animal holding rooms were maintained at 70–73°F, a humidity of 30–35%, and had a 12-h on/off light schedule with lights on from 6AM to 6PM. Once the tumor reached 1,500mm^3 the mouse was euthanized, and the tumor tissue was dissected and placed in a cryomold with optimal cutting temperature compound (TissueTek, 4583), frozen in liquid nitrogen, and then stored at -80°C. Tumors were then sectioned on a cryostat to 6µm thickness, placed onto a microscope slide (Fisher Scientific, 6776214), fixed and permeabilized with the same protocol used for cell lines while in LockMailer slide jars (Fisher Scientific, 50-340-92), and then stored at 4°C.

For the formalin-fixed paraffin embedded (FFPE) tissue experiment, we performed clampFISH 2.0 in two patient-derived xenografts (PDXs), with sample identifiers WM4505-1 (used in replicates 1 and 2) and WM4298-2 (used in replicate 2). The PDXs were each derived from a tumor from a metastatic site of a male patient diagnosed with AJCC Stage IV melanoma. PDX WM4505-1 was derived from an unknown metastatic site in a patient previously treated with combination dabrafenib and trametinib with a mixed response, and whose primary tumor site was the scalp. PDX WM4298-2 was derived from a left back metastatic site in a patient previously treated with vemurafenib, which was discontinued due to an allergic reaction, and whose primary tumor site is unknown. Each PDX was grown out in male NSG mice that were 6-8 weeks old at the time of implantation, with passages performed via subcutaneous implantation of a fragment of the PDX into another mouse. The PDXs were grown for a total of 4 passages (for WM4505-1) or 3 passages (for WM4298-2), where after the first passage, the mice were continuously fed chow containing BRAF/MEK inhibitors (PLX4720 200ppm + PD-0325901 7ppm, chemical additive diet, Research Diets, New Brunswick, NJ). Finally, a piece of about 3x3x3mm^3 of each PDX tumor was implanted into an 6-8 week old male NSG mouse that, once the tumor was palpable, was fed chow containing the BRAF/MEK inhibitors. Tumor size was assessed once weekly by caliper measurements (length x width^2/2). When the tumors reached 1,000mm^3 or when necessary for animal welfare, the tumor was harvested and immediately placed in 10% neutral buffered formalin overnight (less than 48hrs), washed once with 1xPBS, and stored in 70% ethanol at room temperature. Following the Wistar Institute Histotechnology facility's standard protocol, the fixed tumor samples were embedded in paraffin, sectioned to 5µm thickness, and placed on a microscope slide. To avoid exposure to the air, the samples were sealed with a thin layer of paraffin, then stored at room temperature.

For both the fresh frozen tissue and the FFPE tissue samples, we next placed the samples' slides in 2X SSC for 1 - 5 minutes, in 8% sodium dodecyl sulfate (Sigma-Aldrich, 75746-250G; dissolved in nuclease-free water) for 2 minutes, and then into 2X SSC for up to 2 hours, after which we began the primary probe steps. We performed the clampFISH 2.0 steps in parallel for both types of samples (fresh frozen and FFPE) in two separate experimental replicates (replicate 1: fresh frozen mouse #8948 and FFPE PDX WM4505-1; replicate 2: fresh frozen samples #8948 and #8947 and FFPE samples WM4505-1 and WM4298-2).

Supplementary methods: clampFISH 2.0 protocol

Details of clampFISH 2.0 primary probe steps

We hybridized only a single primary probe set per well with the amplifier screen experiment (GFP or *EGFR* probe sets) and the pooled amplification experiment (GFP probe set). We hybridized primary probe sets for *FN1*, *MITF*, and *NGFR* in an experiment assessing a one-pot amplification protocol. For all other experiments we hybridized 10 primary probe sets together (see Supplementary Table 1 for list of primary probe sets).

Details of clampFISH 2.0 amplification steps

For the amplifier screen experiment, we used half the amplifier probe concentrations (typically ~10nM, with ~5nM used in this experiment) and did not have Triton-X in the hybridization buffer.

For the amplifier screen experiment and the pooled amplification experiment, after the last amplification round we proceeded with conventional single-molecule RNA FISH per (Raj et al. 2008) by first rinsing briefly with 10% wash buffer, adding GFP or *EGFR* probes as well as a 20 nucleotide secondary-targeting readout probe at 4nM final concentration in 10% hybridization buffer (10% formamide, 10% dextran sulfate, 2X SSC), covering with a coverslip, placing in a humidified container and incubating overnight in at 37°C, adding 10% wash buffer to remove the coverslip, washing 2 x 30 minutes in 10% wash buffer in a 37°C incubator, while adding 50 ng/mL of the nuclear stain 4′,6-diamidino-2-phenylindole (DAPI) to the second wash, after which we did not carry out further readout probe steps. For the wash and click steps that use a hotplate, in these two experiments we instead used a 37°C incubator or bead bath, with the sample in a LockMailer slide jar submerged in the appropriate buffer.

For the high-throughput profiling experiment in a 6-well plate, we replaced the use of a hotplate with a 37°C incubator and increased the incubation time of the 10 minute wash in 10% wash buffer, the 10 minute click reaction, and all steps in 30% wash buffer by an additional 4 minutes to accommodate the longer time to warm up.

In an experiment assessing a one-pot amplification protocol (adding secondary probes, tertiary probes, and the click reagents simultaneously), we first added one of two buffers: a buffer with dextran sulfate and formamide (10% formamide, 10% dextran sulfate, 2X SSC, 0.25% Triton-X, 10% DMSO) or without those reagents (2X SSC, 0.25% Triton-X, 10% DMSO) to the sample in a well of an 8-well chamber. We next added the secondary probe and circularizer oligonucleotide mixture (containing 10 secondary probes), added a tertiary probe mixture (containing 10 tertiary probes), mixed the sample using a pipette tip, added a pre-mixed copper sulfate and BTTAA mixture, added freshly-dissolved ascorbic acid, and again mixed the sample using a pipette tip (with these reagents at approximately the same final concentrations as described above). After incubation of the one-pot mixtures at 37°C for 30 minutes, we continued with the standard 10% wash buffer and 30% wash buffer washes. In parallel, and with the same batches of reagents, we performed clampFISH 2.0 in the standard manner to round 1 and to round 4 as a positive control.

Supplementary methods: imaging and image analysis

Details of image analysis

For the amplifier screen experiment, we segmented cells in rajlabimagetools from 60X magnification z-stacks at 0.3μm z-steps. For each cell, we manually selected a minimum single-molecule RNA FISH (smFISH) spot intensity threshold. To generate the specificity and sensitivity curves (Supplementary Fig. 2), we compared the positions of these smFISH spots with the positions of clampFISH 2.0 spots whose intensities were above a single threshold for all cells, calling a spot ‘colocalized’ if it was within 2.8 pixels in XY and within 1 z-plane of a spot of the opposite type (Ie. comparing clampFISH 2.0 spots to smFISH spots and vice-versa). For cells in which the smFISH spot count was 20 or greater, we took an equivalent number of the highest-intensity clampFISH 2.0 spots from that cell and used this list of clampFISH 2.0 spot intensities for plotting boxplots (Supplementary Fig. 5) and for calculation of the median intensity (Supplementary Fig. 6).

For the pooled amplification experiment (Supplementary Fig. 7), in order to quantify the typical spot intensity we used rajlabimagetools to extract the 10,000 highest-intensity GFP clampFISH 2.0 spots from 60X z-stacks of 40 segmented cells per condition (an average of 250 spots per cell). The highest-intensity spots were chosen to eliminate potential biases associated with manually-chosen thresholds.

For the readout probe stripping experiment (Extended Data Fig. 7), we segmented 39-48 cells in the before-stripping 20X images, chose gene-specific clampFISH 2.0 spot intensity thresholds, aligned the same segmentations to the post-stripping images, and extracted spot counts from the post-stripping images.

For the amplification characterization experiment, we used Cellpose (Stringer et al. 2021) to automatically segment cells using cellular background fluorescence in the YFP channel (with the DAPI channel also included as a Cellpose input), and excluded abnormally small or large cells. For each of the 4 probed genes we used rajlabimagetools to extracted the top N spots from each round of amplification, where: N = (number of cells)×k, and k is the assumed average number of spots per cell (k = 120 , 1, 20, and 80 spots/cell for *UBC*, *ITGA3*, *FN1*, and *MITF*, respectively). To avoid saturating the camera's photon-collecting capacity at higher rounds of amplification, we extracted spots from longer exposure times on amplification rounds 1,2, and 4 (1000, 1000, 500, and 500 milliseconds for each gene, respectively) and shorter exposure times on amplification rounds 6, 8, and 10 (all were 100 milliseconds), and scaled these intensities by the ratio of median spot intensities between the two exposure times at round 6. For all no-click conditions, we used the longer exposure times to extract spot intensities. We then normalized the data by dividing all intensity values by the median value from round 1, using these in Fig. 1e, Extended Data Fig. 4, and Supplementary Fig. 3). Coefficients for the displayed exponential curve fit were calculated using a least-squares linear regression of log2-transformed median intensity values from rounds 2, 4, 6, 8, and 10.

To generate plots where spot size is depicted (Extended Data Figs. 5-6) we imaged stacks with 15 z-planes at 0.2μm spacing using a 100X objective (1x1 camera binning, 65nm width per pixel), manually segmented 7-12 cells, and manually selected minimum spot intensity thresholds, where a single uniform threshold was chosen for a given clampFISH 2.0 condition whereas conventional single-molecule RNA FISH thresholds were chosen for each cell individually. We then performed a least-squares fit of the above-threshold spots at their maximum-intensity z-plane to a 2D gaussian distribution with an allowable standard deviation between 0 and 227.5nm (0 to 3.5 pixels). To calculate a median full width at half maximum spot size, we multiplied the median standard deviation of the gaussian fit by 2.355.

For the conventional single-molecule RNA FISH comparison experiment (Fig. 2a-b), we manually segmented cells from 60X images using rajlabimagetools, manually selected minimum spot intensity thresholds for the conventional single-molecule RNA FISH data for each cell individually, and counted spots in each cell from 11 z-planes at 0.5μm spacing. We scaled and aligned these segmentations to the 20X and 10X images, and extracted clampFISH 2.0 spots exceeding a gene-specific threshold for 20X (3 z-planes at 1μm spacing) and 10X (3 z-planes at 2μm spacing) images. To calculate the detection efficiency for a given gene, we divided the number of clampFISH 2.0 spots detected across all cells at 20X magnification by the number of conventional single-molecule RNA FISH spots detected across all cells at 60X magnification, finding detection efficiencies of (format: replicate 1, replicate 2): *AXL* (73%, 63%), *EGFR* (49%, 53%), and *DDX58* (49%, 65%). To quantify sensitivity and specificity on the lowly-expressed gene *DDX58*, we denoted cells with 3 or more spots as

'DDX58 high' and with 2 or fewer spots as '*DDX58* low', and did so using conventional single-molecule RNA FISH at 60X magnification (the gold standard) and using clampFISH 2.0 at 20X magnification. In two biological replicates (different passages of WM989 A6-G3 cells), we found clampFISH 2.0 at 20X magnification could identify *'DDX58* low' cells with a specificity of 97% (32/33 cells, replicate 1) and 99% (86/87 cells, replicate 2) and *'DDX58* high' cells with a sensitivity of 41% (35/86 cells, replicate 1) and 53% (10/19 cells, replicate 2).

For the high-throughput profiling experiment, we stitched and registered the tiled scans from multiple imaging cycles at 20X magnification using the custom pixyDuck repository and then divided the scan into smaller subregions. We imaged 5 wells (replicate 1) and 1 well (replicate 2) of WM989 A6-G3 cells, dividing those scans into 10x10 subregions, and 1 well (replicates 1 and 2) of WM989 A6-G3 RC4 cells, dividing those scans into 6x6 subregions. We used dentist2 to choose spot intensity thresholds, extract spots, and then assign those spots to cellular segmentations generated by Cellpose based on cellular background fluorescence (eg. autofluorescence) in the YFP channel (using the diameter parameter of 90 pixels for WM989 A6-G3 cells and 350 pixels for WM989 A6-G3 RC4 cells). We used the housekeeping gene *UBC*, for which a readout probe was hybridized on every readout cycle, for the following quality control steps. First, we kept only subregions where there was an average of at least 25 *UBC* spots per cell for all readout cycles (we observed that near the edges of the wells, fewer spots above our chosen thresholds were detected, presumably because the coverslip used to spread out all probe-containing solutions were smaller than the full well). We then took only cells where, for all readout cycles, the *UBC* spot count was: at least 4, at least 0.025/um^2×cell area, always within 50% of the median count from all readout cycles. Out of the initial 1,297,062 (replicate 1) and 253,662 (replicate 2) WM989 A6-G3 cells segmented, 722,298 (replicate 1) and 234,410 (replicate 2) cells passed all quality control metrics and were included in downstream analyses. To analyze only cells expressing high levels of one or more of 8 marker genes, we chose, for each gene, the following minimum spot count thresholds (format: minimum spot count to be considered high-expressing, percentage of cells high-expressing in replicate 1): *WNT5A* (>=15, 0.59%), *DDX58* (>=10, 0.56%), *AXL* (>=25, 3.56%), *NGFR* (>=30, 1.07%), *FN1* (>=100, 2.79%), *EGFR* (>=5, 1.40%), *ITGA3* (>=50, 2.31%), *MMP1* (>=40, 1.48%). For the 5.93% of cells (42,802 out of 722,298) in replicate 1 and the 10.5% of cells (24,685 out of 234,410) in replicate 2 expressing high levels of one or more marker genes, we used MATLAB's clustergram function to perform hierarchical clustering using all 10 genes' normalized spot counts (replicate 1: Fig. 3c; replicate 2: Supplementary Fig. 16), where each gene's spot counts were transformed such that the mean is 0 and the standard deviation is 1.

For Extended Data Fig. 8-9, Supplementary Figs. 11-12 and 14-15, we ran the same pipeline on a smaller imaged area in Well A1 that, in addition to the three readout cycles included previously, also included a re-imaging of readout cycle 1 and readout cycles 4 and 5 (both of which re-used the same readout probes from readout cycle 1). To define spots, a single minimum spot intensity threshold was chosen for each gene on each round. Thresholds for readout cycle 4 images were made the same as those in cycle 1. For readout cycle 5 (performed after storing the sample at 4°C for 4 months on replicate 1 only), the thresholds were increased by 67% to 83% (the cycle 5 signal presumably appeared brighter due to changes in the microscope's optical path, i.e. greater sample illumination or increased transmission to the sensor). To calculate each cell's mean background levels for Supplementary Figs. 11-12, we generated a background image for each gene and imaging cycle by selecting 100 random image tiles with clampFISH 2.0 signal, generating an image using the 5th percentile of the 100 values for each pixel position, performing gaussian-smoothing on this image, and then averaging these smoothed values in the cellular segmentation. Mean

background-subtracted fluorescent intensity was calculated by averaging the pixel values in the cellular segmentation and subtracting the mean background level. To further correct for background contributed by autofluorescence and any residual fluorescence from previous readout cycles, in Supplementary Fig. 12 we take the mean background-subtracted fluorescent intensity of the clampFISH 2.0 signal and subtract the mean background-subtracted fluorescent intensity derived from images taken after the previous readout probes have been stripped off but before the new clampFISH 2.0 readout probes are introduced.

Supplementary methods: RNA sequencing

Details of RNA sequencing

We conducted standard bulk paired-end (37:8:8:38) RNA sequencing using RNeasy Micro (Qiagen, 74004) for RNA extraction, NEBNext Poly(A) mRNA Magnetic Isolation Module (NEB E7490L), NEBNext Ultra II RNA Library Prep Kit for Illumina (NEB, E7770L), NEBNext Multiplex Oligos for Illumina (Dual Index Primers Set 1) oligos (NEB, E7600S), and an Illumina NextSeq 550 75 cycle high-output kit (Illumina, 20024906), as previously described (Mellis et al. 2021; Shaffer et al. 2017). Prior to extraction and library preparation, the samples were randomized to avoid any experimental and human biases. We aligned RNA-seq reads to the human genome (hg19) with STAR v2.5.2a and counted uniquely mapping reads with HTSeq v0.6.1 (Dobin et al. 2013; Mellis et al. 2021; Shaffer et al. 2017) and outputs count matrix. The counts matrix was used to obtain tpm and other normalized values for each gene using scripts provided at: https://github.com/arjunrajlaboratory/RajLabSeqTools/tree/master/LocalComputerScripts

References from Supplementary Information

Acheampong, Kofi K., Dylan L. Schaff, Benjamin L. Emert, Jonathan Lake, Sam Reffsin, Emily K. Shea, Courtney E. Comar, et al. 2022. “Subcellular Detection of SARS-CoV-2 RNA in Human Tissue Reveals Distinct Localization in Alveolar Type 2 Pneumocytes and Alveolar Macrophages.” *mBio*, February, e0375121.

Dobin, Alexander, Carrie A. Davis, Felix Schlesinger, Jorg Drenkow, Chris Zaleski, Sonali Jha, Philippe Batut, Mark Chaisson, and Thomas R. Gingeras. 2013. “STAR: Ultrafast Universal RNA-Seq Aligner.” *Bioinformatics* 29 (1): 15–21.

Mellis, Ian A., Hailey I. Edelstein, Rachel Truitt, Yogesh Goyal, Lauren E. Beck, Orsolya Symmons, Margaret C. Dunagin, et al. 2021. “Responsiveness to Perturbations Is a Hallmark of Transcription Factors That Maintain Cell Identity in Vitro.” *Cell Systems* 12 (9): 885–99.e8.