JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
Cite This: J. Am. Chem. Soc. 2018, 140, 11755−11762
pubs.acs.org/JACS

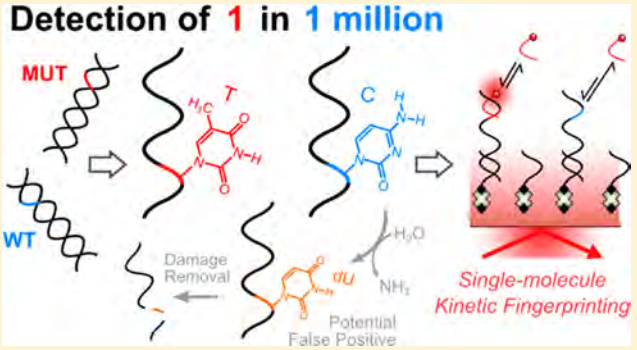
Ultraspecific and Amplification-Free Quantification of Mutant DNA by Single-Molecule Kinetic Fingerprinting

Stephen L. Hayward,[†,‡] Paul E. Lund,[§,‡] Qing Kang,[†] Alexander Johnson-Buck,[*,†,§,∥] Muneesh Tewari,[*,†,∥,#,¶,⊥] and Nils G. Walter[*,§,∥,¶]

[†]Department of Internal Medicine, Division of Hematology/Oncology, [§]Single Molecule Analysis Group, Department of Chemistry, [∥]Center for RNA Biomedicine, [#]Department of Biomedical Engineering, [¶]Center for Computational Medicine and Bioinformatics, and [⊥]Biointerfaces Institute, University of Michigan, Ann Arbor, Michigan 48109, United States

Supporting Information

ABSTRACT: Conventional techniques for detecting rare DNA sequences require many cycles of PCR amplification for high sensitivity and specificity, potentially introducing significant biases and errors. While amplification-free methods exist, they rarely achieve the ability to detect single molecules, and their ability to discriminate between single-nucleotide variants is often dictated by the specificity limits of hybridization thermodynamics. Here we show that a direct detection approach using single-molecule kinetic fingerprinting can surpass the thermodynamic discrimination limit by 3 orders of magnitude, with a dynamic range of up to 5 orders of magnitude with optional super-resolution analysis. This approach detects mutations as subtle as the drug-resistance-conferring cancer mutation *EGFR* T790M (a single C → T substitution) with an estimated specificity of 99.99999%, surpassing even the leading PCR-based methods and enabling detection of 1 mutant molecule in a background of at least 1 million wild-type molecules. This level of specificity revealed rare, heat-induced cytosine deamination events that introduce false positives in PCR-based detection, but which can be overcome in our approach through milder thermal denaturation and enzymatic removal of damaged nucleobases.



INTRODUCTION

Detection and quantification of rare mutant DNA alleles is increasingly important for a broad set of applications in biological and clinical research, including tracking the emergence of rare alleles responsible for drug resistance in populations of microbes, viruses, or cancer cells.[1,2] For example, rare fragments of circulating tumor DNA (ctDNA) are present in biofluids (e.g., urine and blood) of cancer patients and can be used to track tumor evolution or detect minimal residual disease after treatment.[2] However, the high-specificity detection of such rare DNA mutant alleles in a large wild-type background remains a critical challenge, which is further complicated by the low (and highly variable) DNA content of different biofluids (Table S1). While a wide variety of innovative methods for detection of nucleic acids have been developed,[3−8] PCR-based methods continue to dominate. Quantitative PCR-based assays primarily exploit differential hybridization thermodynamics to distinguish homologous sequences; however, such thermodynamic discrimination is necessarily finite,[9] frequently resulting in poor specificity for challenging targets such as point mutations. Next-generation-sequencing (NGS)-based protocols have been optimized to increase specificity, but in practice still have difficulty detecting rare DNA alleles with $<10^{-4}$ relative abundance without significant increases in sequencing depth.[10,11]

A significant and perhaps underappreciated confounder in all workflows involving PCR amplification—including NGS and droplet digital PCR (ddPCR)—are errors caused by heat-induced chemical damage (e.g., nucleotide deamination or oxidation) and imperfect copying fidelity of DNA polymerases.[12−14] Once present, these errors can be exponentially amplified over many cycles, yielding significant false positives in assays for rare point mutations. While amplification-free methods can in principle avoid these errors, they typically struggle to detect single-nucleotide variants at relative abundances below ∼0.1%.[15−17] This limitation is fundamental because hybridization reactions exhibit a maximal discrimination factor $Q_{max,therm} = e^{-\Delta\Delta G^{\circ}/RT}$, where $\Delta\Delta G^{\circ}$ is the difference in the Gibbs free energy of hybridization of a detection probe to a target sequence and of the same probe to a related but spurious target sequence; in practice, this translates to $Q_{max,therm}$ values ranging from about 20 to 20 000 for single-nucleotide variants.[9] Moreover, in contrast to PCR-based

Received: June 25, 2018
Published: August 20, 2018

Downloaded via UNIV OF MICHIGAN ANN ARBOR on September 20, 2018 at 19:56:47 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

© 2018 American Chemical Society

DOI: 10.1021/jacs.8b06685
J. Am. Chem. Soc. 2018, 140, 11755−11762

methods, amplification-free methods rarely achieve the ability to detect single molecules.

Previously, we reported a method for microRNA detection[18] using a PCR-free detection principle—termed SiMREPS, or single-molecule recognition through equilibrium Poisson sampling—wherein the direct visualization of the repeated binding of fluorescent probes to surface-captured single-stranded microRNAs yields a sequence-specific "kinetic fingerprint" that permits high-confidence identification and counting of single target molecules. While the theoretical potential for high specificity was discussed, the detection of single-nucleotide variants at very low abundance (i.e., <0.01%) was not experimentally realized in prior work. Here, we show that SiMREPS can detect very rare mutations in double-stranded (ds)DNA with extremely high single-base discrimination. Using a straightforward, gentle denaturation combined with enzymatic removal of DNA damage, we achieve direct, amplification-free detection of single target DNA molecules with a specificity (99.99999%) that exceeds the most advanced versions of digital PCR and NGS by at least an order of magnitude[10,19] and an apparent discrimination factor ($Q_{app} > 10^6$) exceeding $Q_{max,therm}$ by greater than 2 orders of magnitude. Using time-encoded kinetic information, SiMREPS thus realizes the ultrahigh specificity required for rare mutant DNA allele quantification and, consequently, is able to reveal spontaneous, heat-induced cytosine deamination events that render PCR-based detection of C → T mutations particularly challenging.

EXPERIMENTAL SECTION

Oligonucleotides and DNA Substrates. All unmodified DNA oligonucleotides used in this study were purchased from Integrated DNA Technologies (IDT) with standard desalting purification. Fluorescent probe oligonucleotides were purchased from IDT with either a 5′ or 3′ fluorophore modification and high-performance liquid chromatography (HPLC) purification. Biotinylated locked nucleic acid (LNA) capture probes were purchased from Exiqon with HPLC purification. See Table S2 for all oligonucleotide sequences. Fluorescent probes 5′-/Cy5/ATG TCT TG-3′ (FP2) specific for exon 19 deletion and 5′-CTG CAT GA/Cy5/-3′ (FP1) specific for T790M were used in all experiments unless otherwise noted. Double-stranded DNA substrates were prepared by combining complementary oligonucleotides (20 μM each strand) in annealing buffer (10 mM Tris-HCl [pH 8.0 at 25 °C], 50 mM NaCl, 1 mM EDTA), heating at 95 °C for 5 min, followed by slow cooling at room temperature for 15 min. Details for DNA substrates prepared by restriction enzyme digestion from plasmid DNA are provided in the Supporting Information. For additional information regarding the design of auxiliary competitor probes, see the Supplementary Note.

Prediction of Fluorescent Probe Discrimination Properties Using NUPACK. To determine the predicted discrimination between mutant (MUT) and wild-type (WT) DNA with a specific FP design, the equilibrium molar concentrations of FP-MUT and FP-WT complexes were calculated using the NUPACK Web application[20] (http://www.nupack.org/partition/new) using the following parameters: target DNA concentration = 1 nM, FP concentration = 25 nM, 25 °C, 600 mM NaCl, number of strand species = 2, maximum complex size = 2, histogram filter minimum concentration = 1×10^{-7}. The predicted discrimination was then taken as the ratio of [FP-MUT]/[FP-WT] for each specific FP design. To determine the predicted FP dimer mole fraction, the FP dimer molar concentration for each FP design was calculated and normalized by the total FP concentration [FP−FP]/25 nM.

The theoretical $Q_{max,therm}$ was calculated from ref 9, i.e., $Q_{max,therm} = e^{(-\Delta\Delta G°/RT)}$, where $\Delta\Delta G° = \Delta G°_{FP\cdot MUT} - \Delta G°_{FP\cdot WT}$.

$\Delta G°_{FP\cdot MUT}$ and $\Delta G°_{FP\cdot WT}$ are the free energies of hybridization of the MUT fluorescent probe (FP) to the MUT and WT targets, respectively, and were estimated using NUPACK with the following parameters: nucleic acid type = DNA, DNA energy parameters = SantaLucia, 1998, number of strand species = 2, maximum complex size = 2, temperature = 20 °C, [Na^+] = 0.6 M, 1 μM each strand, dangle treatment = some. Additional guidelines for SiMREPS probe and assay design can be found elsewhere.[21]

Single-Molecule Fluorescence Microscopy. SiMREPS experiments were performed on an Olympus IX-81 objective-type TIRF microscope equipped with a 60× oil-immersion objective (APON 60XOTIRF, 1.49 NA) with both CellTIRF and *z*-drift control modules, and an EMCCD camera (IXon 897, Andor, EM gain 150). For an expanded discussion of the general instrumentation requirements, we refer the reader to a recent publication by Johnson-Buck et al.[21] Sample cells were constructed as previously described.[18] Briefly, glass coverslips were passivated with a 1:10 mixture of biotin-PEG-5000 and mPEG-5000 (Laysan Bio, Inc.) and stored for up to 2 weeks in the dark under nitrogen (see Supplementary Protocol). Directly prior to single-molecule experiments, 20 μL pipet tips (Thermo Scientific #2149P05HR, low retention) were cut to an ∼1 cm length as measured from the wide end, and the noncut base was adhered to the coated glass coverslip via epoxy (Hardman Double/Bubble #04001) to form the sample cell.

SiMREPS Assay Protocol. The SiMREPS protocol used in this study was based on our previously described method,[18] with three main modifications: (1) the LNA concentration was increased from 20 nM to 100 nM during surface coating, (2) a thermal denaturing step was added directly before analyte capture to increase access to the target DNA strand, and (3) two additional oligonucleotides including an LNA blocker and WT competitor (WTC) were added to the imaging solution when probing T790M samples. All DNA handling was performed in GeneMate low-adhesion 1.7 mL micro centrifuge tubes, and all dilutions and denaturing steps were performed in the presence of 2 μM poly-T oligodeoxyribonucleotide (dT_{10}) in 1× PBS. The updated protocol was performed as follows. First, sample cells were washed with T50 buffer (10 mM Tris-HCl [pH 8.0 at 25 °C], 50 mM NaCl) and then incubated with 40 μL of 1 mg/mL streptavidin for 10 min. Cells were then washed with T50 buffer to remove excess streptavidin, incubated with 40 μL of 100 nM biotinylated LNA in 1× PBS for 10 min, and washed with 1× PBS. Target DNA oligonucleotide was denatured at 95 °C for 3 min (unless otherwise specified) in a thermocycler, cooled in room-temperature water for 5 min (via partial submersion), and immediately added to the LNA-coated surface. DNA capture was performed for 1 h at room temperature in a humidified chamber. After capture, excess DNA was removed by washing with 4× PBS. Directly before imaging, imaging buffer containing 25 nM fluorescent probe, 4× PBS, 5 mM 3,4-dihydroxybenzoate, 50 nM protocatechuate dioxygenase, and 1 mM Trolox was added to the sample cell. When probing T790M, 25 nM LNA blocker and 1 μM WTC were also added to the imaging solution (refer to Supplementary Note). Transient binding of the fluorescent probe was monitored for 10 min under TIRF illumination by 640 nM laser light with a 500 ms exposure time (1200 total frames).

Analysis of SiMREPS Data. Two different analysis pipelines were used to investigate kinetic data in this study. The first analysis pipeline is our previously optimized method, termed "diffraction limited", which employs custom Matlab code to identify sites of fluorescent probe binding and dissociation in a 14 024 μm^2 field of view and calculate intensity-versus-time trajectories at each site.[18] Hidden Markov modeling (HMM) is then applied using the QuB software suite (State University of New York at Buffalo) to determine the number of binding and dissociation events (N_{b+d}) and the median fluorescent probe bound ($\tau_{bound,median}$) and unbound ($\tau_{unbound,median}$) time for each candidate molecule.[22] Based on (i) no DNA, (ii) WT DNA only, and (iii) MUT DNA only control experiments, kinetic thresholds were defined to achieve minimal false positive frequency while retaining high sensitivity for true MUT candidates. Thresholds applied for all DNA targets include a signal-to-noise ratio of >3, mean bound state intensity >500 arbitrary fluorescence units above the mean background intensity, maximum $\tau_{bound,median}$ of 20 s, maximum $\tau_{unbound,median}$ of 30 s, and $N_{b+d} \geq 20$. For exon 19 deletion quantification, the minimum $\tau_{bound,median}$ and

DOI: 10.1021/jacs.8b06685
J. Am. Chem. Soc. 2018, 140, 11755−11762

$\tau_{unbound,median}$ thresholds were set to 2 and 2.5 s, respectively, while for T790M quantification the minimum $\tau_{bound,median}$ and $\tau_{unbound,median}$ thresholds were set to 3 s. Although $\tau_{unbound,median}$ is in principle a function of solution probe concentration, in practice we have found that minimum and maximum thresholds for $\tau_{unbound,median}$ can be used to eliminate molecule candidates that display uncharacteristically transient interactions (e.g., rapid and repeated photoblinking of the Cy5) or irreversible binding followed by photobleaching (e.g., probe interactions with surface defects) and thus can reasonably be assumed to be artifactual. The diffraction-limited analysis pipeline was used for all experiments in this study except for experiments presented in Figure 3b,c ("super-resolution" curves) and Figure 3d, where a new super-resolution analysis pipeline, which also incorporates subpixel image registration,[23] was used (see Supporting Information).

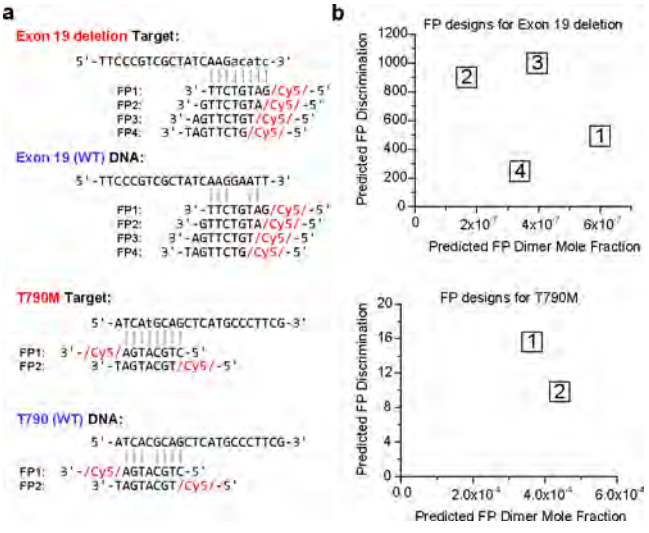


Figure 1. Fluorescent probe design. (a) Fluorescent probe (FP) sequences having various registers with respect to the *EGFR* mutations used in this study. A 3′ exogenous barcode sequence (TAGGAC) present on the DNA analytes was not considered while designing the probe sequences and is omitted for clarity. Lower-case letters represent the mutation site. Cy5, fluorophore. (b) Plots of mutant/wild-type discrimination and FP dimer formation predicted using NUPACK.

Statistical Analysis. Statistical significance of the differences between mean counts per field of view shown in Figure 4f was determined using a two-tailed, unpaired *t* test (GraphPad QuickCalcs, www.graphpad.com). The discrimination factor Q_{app} from LOD experiments was calculated as described in the Supplementary Note. Curve fitting was performed using OriginPro 8 software, and values were weighted using instrumental weighting by the respective error value during linear regression.

Code Availability. The custom Matlab scripts used in this study are available upon request.

Data Availability. The microscopy data (e.g., movie files) that support the findings of this study are available in the Deep Blue Data Repository from the University of Michigan library (https://deepblue.lib.umich.edu/data/concern/generic_works/kd17ct808 DOI: 10.7302/Z2CZ35DF.

RESULTS

Optimized Fluorescent Probe Design for Improved Discrimination between Wild-Type and Mutant Alleles. To adapt the SiMREPS kinetic fingerprinting approach to rare mutant DNA allele detection, we began by optimizing the design of the fluorescent probe (FP) through a combination of theoretical and empirical methods. The two *EGFR* mutations chosen for use in this study were an in-frame deletion in exon 19 (COSMIC ID: COSM6225; c.2236_2250del15

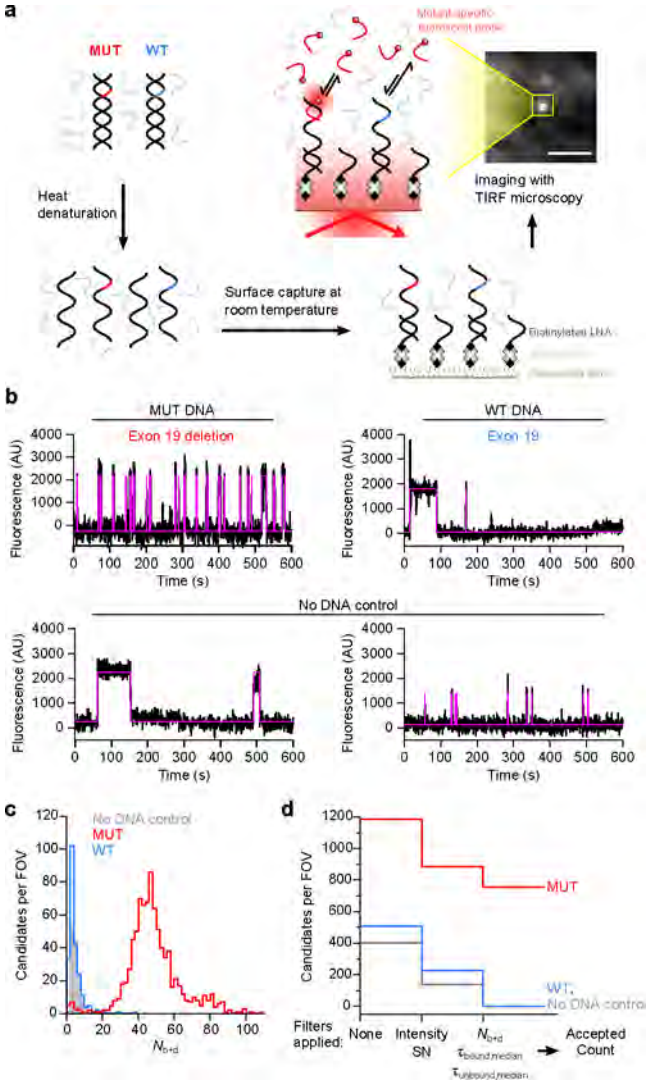


Figure 2. Direct DNA visualization and counting of mutant molecules by amplification-free single-molecule kinetic fingerprinting. (a) Schematic depicting the experimental SiMREPS approach adapted for DNA detection. Double-stranded (duplex) DNA is converted to single-stranded DNA with brief heat denaturation in the presence of high concentrations of single-stranded dT_{10} carrier to disfavor reannealing of complementary analyte strands. Single-stranded target DNA is captured by target gene-specific LNA capture strands immobilized on the slide surface, and remaining unbound DNA is washed away prior to kinetic fingerprinting with a mutant (MUT)-specific fluorescent probe (FP) with optimized binding and dissociation kinetics. Surface-immobilized molecules are seen as bright points in the accompanying 50-frame averaged image showing a subset of the entire field-of-view. Scale bar = 3 μm. (b) Representative kinetic traces using a MUT-specific FP with MUT DNA, with wild-type (WT) DNA, or without DNA (no DNA control) for *EGFR* exon 19/deletion. Fluorescence intensity trace, black; two-state idealization from hidden Markov modeling, magenta line. (c) Histograms of the number of candidate molecules per field-of-view (FOV) showing a given number of intensity transitions (N_{b+d}), after applying thresholds for fluorescence intensity and signal-to-noise (see Experimental Section), for experiments in which 500 fM *EGFR* exon 19 deletion MUT, 500 nM WT, or no DNA was used. (d) Plot for the same data set shown in (c) depicting the number of candidate molecules as a function of filtering thresholds applied to arrive at the final accepted molecule count.

DOI: 10.1021/jacs.8b06685
J. Am. Chem. Soc. 2018, 140, 11755−11762

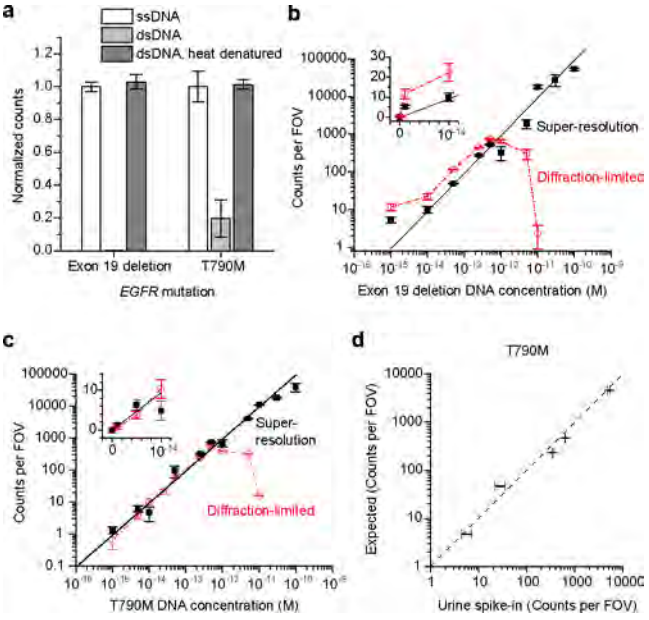


Figure 3. Dynamic range of mutant dsDNA detection and quantification in buffer and urine matrices is enhanced by super-resolution analysis. (a) Impact of DNA strandedness and brief heat denaturation on the number of detected molecules (counts) for 1 pM of the two MUT target DNA alleles used in this study. Data are presented as means ± SEM of at least three independent measurements and are normalized to the mean counts detected using ssDNA as input. (b) Performance of super-resolution and diffraction-limited analysis methods over a wide concentration range for exon 19 deletion. Data are presented as the mean ± SEM of $n = 3$ independent measurements. The red dashed line connecting points from diffraction-limited analysis is included to show the reduction in detected counts at concentrations of >1 pM. Error-weighted linear fits of the super-resolution analysis (black) were constrained to a y-intercept of 0; $R^2 = 0.948$. (c) Same as in (b), but for T790M. $R^2 = 0.935$. (d) Correlation of observed counts from a urine spike-in sample versus the number of counts expected at a given concentration of spiked-in MUT DNA allele substrate fragment. Expected counts at a given concentration are based on the standard curve constructed with super-resolution analysis of MUT DNA captured from 1× PBS. Points that appear exactly on the diagonal (dashed line) have perfect correlation. Data are presented as the mean ± SEM (x-error bars) of $n = 3$ independent measurements, y-Error bars represent the error in the number of expected counts given the error of the linear fit.

[p.E746_A750delELREA]) and a single-point mutation in exon 20 that results in the missense mutation T790M (COSMIC ID: COSM6240, c.2369C>T [p.T790M]).

First, the effect of FP length on DNA detection was examined. The differences in the kinetics of FP binding to, and dissociating from, genuine target molecules versus those that are closely related in sequence (i.e., off-target molecules) form the basis of the SiMREPS assay's ability to distinguish between target and off-target molecules.[18,21] Therefore, it is advantageous to design FP sequences that undergo many cycles of binding and dissociation within the experimental observation time. Although 9-nt DNA FPs were used predominantly in our previous work on miRNA detection,[18] we found the bound residence time to be undesirably long for a 9 nt FP binding to a DNA target, whereas an 8 nt FP (Figure S1) yielded a reliable kinetic fingerprint in a 10 min experiment. We hypothesize that this is due to the difference in binding energetics between RNA−DNA and DNA−DNA hybrids. The modeling software NUPACK[20] was then used to assess 8 nt FP sequences having different registers with respect to the mutation site (Figure 1a). In addition, FP registers were chosen that would minimize the number of possible contiguous base pairs with the wild-type sequence, which has been shown to be a key determinant of binding stability, and thus dissociation rate, for DNA duplexes on this length scale.[24,25] For exon 19 deletion, for example, the FP was chosen to span the junction of the 15 bp deletion, thus limiting the number of contiguous complementary base pairs with the wild-type sequence. FPs were evaluated based on predictions of (i) theoretical discrimination between mutant and wild-type and (ii) probability of FP dimer formation (Figure 1b).

High-Confidence Detection of Mutant dsDNA with Single-Molecule Kinetic Fingerprinting. Whereas microRNAs are easily captured from solution because they are single-stranded, extending SiMREPS to DNA required development of an effective protocol for denaturation of the targeted dsDNA. We found that thermal denaturation with addition of oligonucleotide dT_{10} as a carrier in high molar excess substantially reduced reannealing (Figure S2), thus permitting efficient surface capture of the single-stranded (ss)DNA target strand. We thus incorporated this short thermal denaturation step into the SiMREPS protocol, followed by surface capture and visualization of single target molecules via the repeated transient binding of an 8-nucleotide fluorescent probe using total internal reflection fluorescence (TIRF) microscopy (Figure 2a).

Since some FP binding to the imaging surface is observed even in absence of the mutant target—in this case, an *EGFR* exon 19 deletion (c.2236_2250del15)—simply counting fluorescent spots would typically yield hundreds of false positives per field of view (Figure 2). However, analysis of fluorescence-versus-time traces revealed clear differences between the repetitive binding of the FP to a single immobilized MUT molecule and the more sporadic binding observed in the presence of only wild-type DNA or in no-DNA controls (NDC) (Figure 2b). By requiring a minimum number of FP binding and dissociation events (N_{b+d}) for a candidate to be counted as a MUT molecule, the majority of nonspecific binding events from WT and NDC conditions could be excluded (Figure 2c).

The few remaining false positives could be excluded by applying filters based on the apparent lifetimes in the FP-bound ($\tau_{bound,median}$) and FP-unbound ($\tau_{unbound,median}$) states, resulting in essentially zero background in typical measurements (Figure 2d and Experimental Section).

Assay Validation for Clinically Relevant Mutant DNA Targets and Optional Dynamic Range Expansion by Super-Resolution Data Analysis. We applied the above kinetic analysis to two clinically relevant *EGFR* tyrosine kinase domain mutations used to validate the assay: c.2236_2250del15 (exon 19 deletion) and c.2369C>T (T790M). First, we compared the number of counts detected for dsDNA, either with or without heat denaturation, relative to the same concentration (1 pM) of ssDNA; we observed a similar number of counts only when the dsDNA had been heat denatured, confirming that the denaturation protocol enabled the same detection efficiency for dsDNA as ssDNA (Figure 3a). Next, we found that the kinetic fingerprinting assays exhibit a linear dependence on MUT dsDNA concentration over more than 2 orders of magnitude (Figure 3b,c, Figure S3) with diffraction-limited imaging. We note that at high densities of captured MUT molecules deviations from linearity occurred (>1 molecule per

DOI: 10.1021/jacs.8b06685
J. Am. Chem. Soc. 2018, 140, 11755−11762

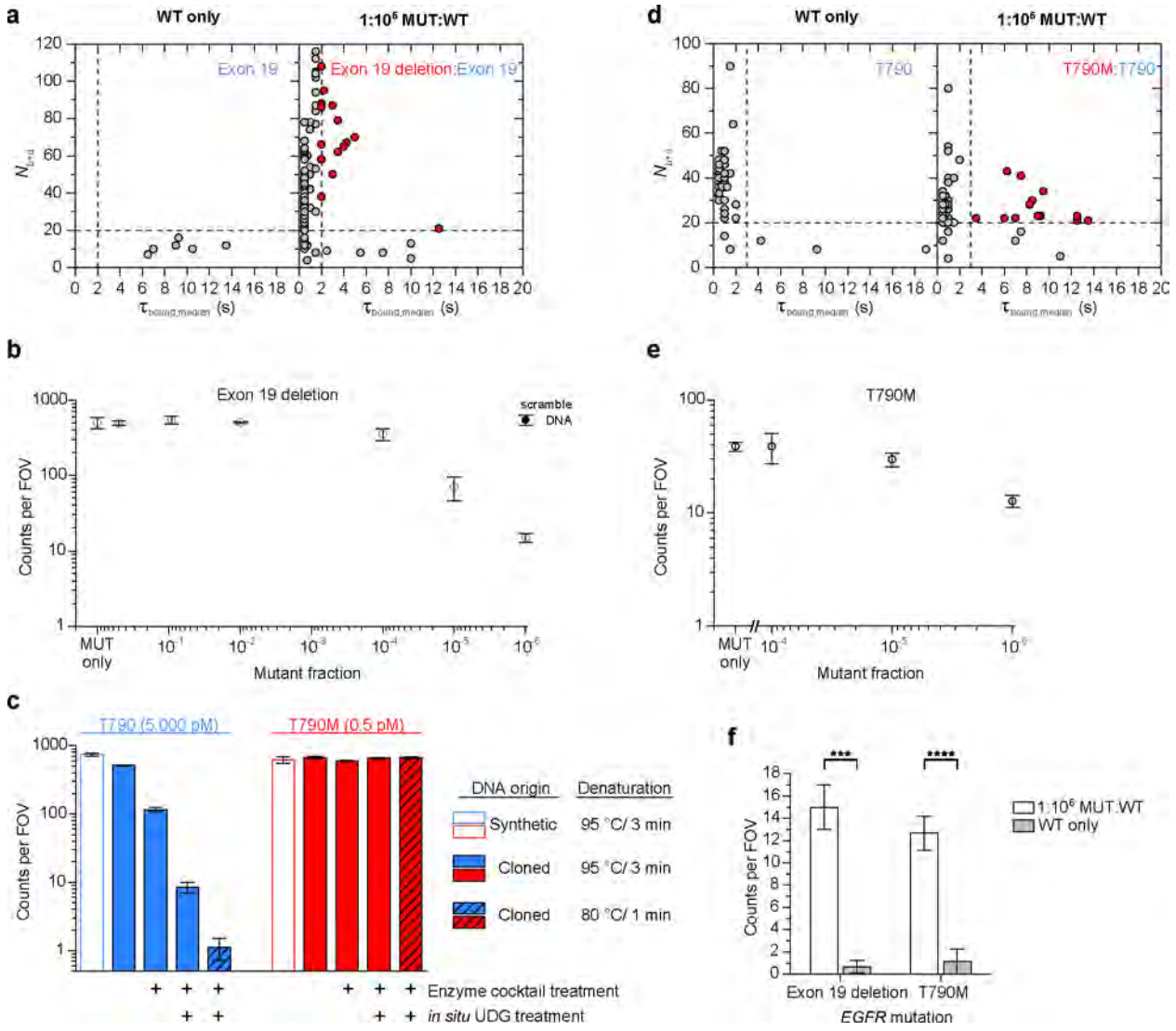


Figure 4. High MUT versus WT discrimination and mitigation of DNA damage leads to ultralow limit of detection (LOD). (a) Plot of diagnostic kinetic parameters used to identify genuine exon 19 deletion (MUT) DNA allele molecules based on their kinetic fingerprints. Plots are generated from one representative movie from each condition, 500 nM WT (WT only), and a mixture of 500 fM MUT with 500 nM WT (1:10^6 MUT:WT), after applying thresholds for fluorescence intensity, signal-to-noise, and median unbound dwell time. Each point represents an individual candidate molecule, plotted by its median bound dwell time ($\tau_{bound,median}$) and number of binding and dissociation events (N_{b+d}) during a 10 min observation. Dashed lines indicate threshold values used to discriminate genuine MUT molecules (red points, upper right quadrant) from nonspecific binding (gray points). (b) LOD experimental analysis for *EGFR* exon 19 deletion. 500 fM MUT DNA was mixed with increasing amounts of WT DNA (open circles), or a scrambled DNA sequence that is not complementary to the capture LNA (filled circle), to achieve various MUT fractions. Data are presented as the mean ± SD of n = 3 independent measurements except for 10^{-1} and 10^{-4} MUT fractions (n = 6). (c) Effect of DNA preparation method, enzymatic removal of damaged DNA bases, and heat denaturation protocol on the number of observed counts for T790 (WT, blue) and T790M (MUT, red) alleles. DNA prepared by chemical synthesis, synthetic, open bars; DNA prepared by restriction digestion from plasmid DNA, cloned, solid bars; mild heat denaturation at 80 °C for 1 min, hashed bars. Enzyme cocktail treatment: combination of uracil-DNA glycosylase (UDG), formamidopyrimidine-DNA glycosylase (FpG), and endonuclease VIII (EndoVIII); *in situ* UDG treatment: treatment of surface-immobilized DNA with UDG prior to observation via the TIRF microscope. Note the difference in concentration used between WT (5 nM) and MUT (500 fM). Data are presented as the mean ± SEM of n = 3 independent measurements. Tabulated values are presented in Table S3. (d) Plot of diagnostic kinetic parameters as in (a), but for the T790M-specific fluorescent probe and LNA capture strand: 5 nM WT (WT only) and a mixture of 50 fM MUT with 5 nM WT (1:10^6 MUT:WT). (e) LOD experimental analysis for T790M: 50 fM MUT DNA was mixed with increasing amounts of WT DNA to achieve various MUT fractions. Data are presented as the mean ± SD of n = 3 independent measurements except for 10^{-4} MUT fraction (n = 4). (f) Comparison of counts from low MUT allelic fraction and WT only conditions. Exon 19/deletion: 15 ± 2 counts (1:10^6 MUT:WT), 0.7 ± 0.6 counts (WT only), P = 0.0003; T790/M: 12.7 ± 1.5 counts (1:10^6 MUT:WT), 1.1 ± 1.1 counts (WT only), P < 0.0001. Data are presented as the mean ± SD of n = 3 independent measurements, except for T790 WT only (n = 8). Significance was assessed using a two-tailed, unpaired t test.

25 μm^2) where neighboring single-molecule kinetic fingerprints could no longer be resolved from one another.

To date, single-molecule kinetic fingerprinting has been diffraction-limited and therefore unable to separate single-molecule binding events above a critical analyte density, where signal saturation artifacts arise. For cases where a broader dynamic range of analyte concentrations without sample dilution is desired, we developed a super-resolution analytic

DOI: 10.1021/jacs.8b06685
J. Am. Chem. Soc. 2018, 140, 11755−11762

Table 1. Apparent Specificity and Discrimination Factor for *EGFR* T790M MUT versus WT as a Function of Input Material and Assay Protocol

	denaturation protocol	enzyme cocktail	*in situ* UDG	specificity	Q_{app}	$Q_{app}/Q_{max,therm}$
5 nM synthetic WT[a]	95 °C/3 min	−	−	99.975 ± 0.001%	3860	0.95
5 nM cloned WT[a]		−	−	99.9821 ± 0.0003%	5580	1.37
	95 °C/3 min	+	−	99.9960 ± 0.0003%	24 700	6.1
		+	+	99.99970 ± 0.00007%	340 000	83
	80 °C/1 min	+	+	99.99996 ± 0.00001%	2 600 000	640
50 nM cloned WT[b]	80 °C/1 min	+	+	99.99999 ± 0.00001%	11 000 000	2600

[a]Calculated from data shown in Figure 4c and Table S3. [b]Calculated from data shown in Figure 4f and Figure S12.

pipeline, inspired by stochastic reconstruction microscopy techniques.[26,27] Here, fluctuations in fluorescence intensity due to FP binding and dissociation are localized with subpixel accuracy and clustered according to position, and the resulting kinetic fingerprints analyzed analogously to the standard diffraction-limited SiMREPS approach (Figure S4). This super-resolution analysis increased the dynamic range of SiMREPS to >5 orders of magnitude for both exon 19 deletion (Figure 3b) and T790M (Figure 3c), rivaling digital PCR assays.[28]

A common application of current PCR-based assays for mutant DNA detection is for quantification of MUT ctDNA extracted from human biofluids.[1,2,29,30] In order to examine matrix effects on our assay, we tested the robustness of our super-resolution kinetic fingerprinting assay in urine, a compositionally complex biological sample matrix in which trans-renal ctDNA can be detected.[31] As proof-of-principle, we quantified 28-bp synthetic T790M DNA molecules spiked into healthy donor urine and observed a linear response that showed strong concordance with counts measured in buffer alone ($R^2 = 0.919$) (Figure 3d), supporting the potential for clinical development of our SiMREPS assay.

Enhancing the Specificity for Rare Mutant Alleles Amid Excess Wild-Type. Using our standard diffraction-limited imaging, we sought to characterize the specificity for MUT DNA in the presence of a large excess of WT sequence. Using the optimized FP for exon 19 deletion (FP2, Figure 1), a specificity of >99.99999% was achieved between the exon 19 deletion MUT and WT DNA, based on evaluation of single-molecule kinetic fingerprints of binding event frequency and duration (Figure 4a, Supplementary Note). This approach enabled a limit of detection (LOD), defined here as the lowest MUT allele fraction showing >3 SEM above background, of ≤0.0001% (10^{-6}; Figure 4b, Figure S5). Interestingly, we found that this LOD was limited not by FP specificity, but by saturation of surface capture sites with excess WT DNA, as was revealed by replacing the exon 19 WT sequence with a scrambled DNA sequence incapable of binding the capture probe (Figure 4b, filled circle).

We next aimed to achieve similarly robust detection of the T790M point mutation, a clinically actionable biomarker for which there is a specific therapy available[32] and which represents an especially challenging analytic target for current DNA detection technology.[33] Despite using the optimized FP sequence for T790M (FP1, Figure 1), we observed that our initial sensitivity for T790M detection was lower than expected. To test the hypothesis that this decreased sensitivity arose in part from sequestration of the FP due to spurious binding to the partially complementary LNA, we introduced an additional "blocker" strand and indeed observed a ∼ 3-fold improvement in observed molecule count and homogeneity of the kinetic fingerprints (Figure S6, Supplementary Note). Moreover, in negative control experiments containing only T790 (WT) DNA, we observed appreciable background noise in the fluorescence microscopy images, which we hypothesized was due to extremely transient off-target binding between the MUT-specific FP and surface-bound WT molecules that becomes significant when the WT is present at very high concentrations. Consistent with this hypothesis, inclusion of an auxiliary competitor probe designed to suppress such interactions resulted in significantly improved signal-to-noise (Figure S7, Supplementary Note).

Effects of Molecular Heterogeneity and DNA Damage on Specificity. In contrast to the exon 19 WT allele, for which false positive signal was independent of WT concentration and indistinguishable from no-DNA controls, we noticed that the chemically synthesized T790 (WT) DNA exhibited a significant false positive frequency of about 1 in 3870 (Table S3).

While the resulting apparent discrimination factor Q_{app} was already comparable to the theoretical thermodynamic maximum $Q_{max,therm}$ of 4080 (Table 1), it was surprising to observe any significant MUT-like kinetic fingerprints in the WT sample. We hypothesized that errors during solid-phase oligonucleotide synthesis could result in a minute fraction of MUT sequence in the WT sample. Therefore, high-fidelity T790 (WT) and T790M (MUT) DNA fragments were prepared by enzymatic digestion from a multi-insert bacterial plasmid (Figure S8). The sequence fidelity of substrates prepared in this way is expected to be primarily dictated by DNA replication error rates in *E. coli*, which is on the order of 10^{-9}−10^{-11} errors per base pair.[34] This high-fidelity material exhibited ∼30% fewer false positives than the corresponding synthetic DNA (Figure 4c), suggesting that, although chemical synthesis errors may contribute, they cannot account for the majority of the false positive signal.

Because T790M is a cytosine (C) to thymine (T) point mutation, we hypothesized that spontaneous deamination of cytosine to uracil (U) may cause a MUT-like kinetic fingerprint in ostensibly WT DNA. We therefore analyzed synthetic DNA bearing a C → U mutation as a mimic of deamination and found this target to have a sufficiently similar, though not identical, kinetic fingerprint to that of MUT DNA (Figure S9, Supporting Information). To remove such false positives from WT material, we pretreated WT DNA with a combination of commercially available enzymes that recognize and remove damaged DNA bases prior to SiMREPS analysis: uracil-DNA glycosylase (UDG), formamidopyrimidine-DNA glycosylase (FpG), and endonuclease VIII (EndoVIII) (Supporting Information). Consistent with our hypothesis, this "enzyme cocktail" pretreatment (designed to remove uracil as well as uracil-like products arising from oxidation of cytosine, such as uracil glycol and 5-hydroxy uracil[35,36]) decreased false positives by >80% compared to the synthetic DNA (Figure 4c), with separate experiments

DOI: 10.1021/jacs.8b06685
J. Am. Chem. Soc. 2018, 140, 11755−11762

indicating UDG as the dominant factor in the reduction (Figure S10).

In light of literature evidence that prolonged heating catalyzes C → U deamination for T790[12], we reasoned that the remaining false positives (about 1 in 24 700) may be caused by the thermal denaturation used to liberate ssDNA in the SiMREPS assay. Therefore, we also treated DNA with UDG *in situ* (i.e., after denaturation and surface capture), which reduced the rate of false positives to about 1 in 340 000. Next, we examined multiple denaturation temperatures and durations for their effect on false positive signal (Figure S11). Using a shorter and lower-temperature denaturation protocol (80 °C for 1 min) led to a reduction of false positives by nearly 8-fold compared to the standard protocol of 95 °C for 3 min, yielding a false positive frequency of approximately 1 in 2.6 million (Figure 4c, Table S3). An LOD experiment using assay conditions revised to limit the influence of DNA damage (Figure 4d,e and Figure S12) measured reliable detection down to 10^{-6} MUT allele fraction, nearly matching the result achieved for the exon 19 deletion (Figure 4f) and resulting in an estimated specificity of 99.99999% and a Q_{app} of 1.1×10^7.

DISCUSSION

Exponential amplification of input nucleic acids by PCR is integral to most assays for rare mutant alleles due to its multiplicative improvement of both sensitivity and specificity. Yet PCR itself imposes limitations on the specificity for very rare sequences, especially point mutations such as *EGFR* T790M, by introducing and amplifying errors such as single-base substitutions and deamination of cytosine. Methods utilizing isothermal amplification in lieu of PCR have also been reported, and while they show promise for point-of-care detection with limited or no instrumentation, residual signal from WT-only samples still places a lower limit on the fraction of MUT allele that can be detected.[37] Conversely, most amplification-free assays forego the multiplicative improvements provided by PCR, resulting in lower sensitivity and a sequence-dependent maximal specificity dictated by hybridization thermodynamics. Recently, simulation-guided design of isothermal exchange probes has permitted the consistent realization of specificity close to the thermodynamic limit, but a sufficiently large excess of WT sequence (200- to 3000-fold) will nonetheless give the same signal as a given concentration of MUT sequence.[15]

In contrast, the kinetic fingerprinting assay developed here exhibits both the ability to directly observe single molecules and extremely high specificity by essentially integrating over multiple, temporally resolved probing events. Our amplification-free SiMREPS assay can detect both exon 19 deletion and T790M mutations with an LOD ≤ 1 in 1 million; in the case of T790M, this is 10-fold lower than even the highest-performing ddPCR assay for that mutation.[19] The apparent discrimination factor Q_{app} for T790M (Figure 4f, Table 1) is ∼11 million in our optimized assay, which is approximately 2600 times greater than the maximum thermodynamic discrimination factor under our imaging conditions as estimated by NUPACK.[38] While kinetic fingerprinting has been shown to be capable of single-nucleotide discrimination,[18,39] the demonstrated specificity in prior studies was 2−4 orders of magnitude lower than that achieved here and in one case required asymmetric PCR amplification.[39] The performance achieved in the present work suggests that arbitrarily high specificity in an amplification-free assay can indeed be achieved by single-molecule kinetic fingerprinting, given (1) suitable assay optimization and (2) treatment of samples to remove damage that would otherwise be read as false positives. Furthermore, the >100-fold increase of assay dynamic range by super-resolution analysis should significantly broaden the potential applications of this technique.

The *absolute* LOD—defined in terms of absolute concentration of MUT analyte, rather than abundance relative to the WT sequence that was discussed above—of our kinetic fingerprinting assays for *EGFR* mutations is currently ∼1 fM (Figure 3b,c). While similar to or better than the that of leading amplification-free methods,[16,40] this absolute LOD is significantly higher than that of ddPCR, the highest-sensitivity method for nucleic acid quantification. Our assay's sensitivity is mainly a function of the efficiency of capturing the target nucleic acids within a small enough region of the coverslip to easily detect, which depends in turn on diffusion rates and both the thermodynamics and kinetics of analyte capture. However, as a single-molecule technique, there is no reason SiMREPS should not be capable of much lower absolute LODs in the future after including suitable mass transfer or other optimizations of the capture protocol.

There is a growing need for accurate, reproducible quantification of ultrarare DNA alleles for multiple applications, including, for example, identification of residual disease or emerging drug-resistant clones after cancer therapy based on detection of ctDNA in blood.[31] In addition, detection of specific DNA alleles in urine is an important approach for analyzing fetal DNA in pregnant women[41−44] and in patients with solid tumors.[45,46] On the basis of the advances described here that overcome the barriers of specificity, LOD, and ease of workflow faced by current gold standard PCR and NGS approaches, we envision that single-molecule kinetic fingerprinting by SiMREPS will be beneficial for a range of fields that require highly specific, rare sequence detection and quantification. Furthermore, since the detection principle underlying single-molecule kinetic fingerprinting does not rely on any nucleic-acid-specific enzymatic processing, it is likely that this technique will find application in the detection of analytes other than DNA and RNA with specificity surpassing the thermodynamic limit.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.8b06685.

Additional experimental details, supporting figures, tables, notes, and protocol (PDF)

AUTHOR INFORMATION

Corresponding Authors

*alebuck@med.umich.edu
*mtewari@med.umich.edu
*nwalter@umich.edu

ORCID

Alexander Johnson-Buck: 0000-0003-3748-6000
Nils G. Walter: 0000-0002-7301-1275

Author Contributions

‡S.L. Hayward and P.E. Lund contributed equally.

Notes

The authors declare the following competing financial interest(s): A.J.-B., M.T., and N.G.W. are inventors on multiple patent applications related to SiMREPS, and equity holders of

DOI: 10.1021/jacs.8b06685
J. Am. Chem. Soc. 2018, 140, 11755−11762

aLight Sciences LLC, a startup company aiming to commercialize the presented technology.

■ ACKNOWLEDGMENTS

The authors acknowledge support from a Michigan Economic Development Corporation MTRAC for Life Sciences grant to M.T., N.G.W., and A.J-B.; pilot grants from the University of Michigan MCubed 2.0 program, the James Selleck Bower Permanently Endowed Innovative Promise Funds for Cancer Research of the University of Michigan Rogel Cancer Center, and Fast Forward GI Innovation Fund to N.G.W. and M.T.; and NIH grant R21 CA204560 to N.G.W. and M.T. The authors also wish to thank the Single Molecule Analysis in Real-Time (SMART) Center of the University of Michigan, seeded by NSF MRI-R2-ID award DBI-0959823 to N.G.W., as well as J. D. Hoff for training, technical advice, and use of the objective-type TIRF microscope.

■ REFERENCES

(1) Rybicka, M.; Stalke, P.; Bielawski, K. P. *Rev. Med. Virol.* **2016**, *26* (5), 369−81.
(2) Diaz, L. A., Jr.; Bardelli, A. *J. Clin. Oncol.* **2014**, *32* (6), 579−86.
(3) Li, Z.; Hayman, R. B.; Walt, D. R. *J. Am. Chem. Soc.* **2008**, *130* (38), 12622−3.
(4) Johnson, R. P.; Richardson, J. A.; Brown, T.; Bartlett, P. N. *J. Am. Chem. Soc.* **2012**, *134* (34), 14099−107.
(5) Ge, Z.; Lin, M.; Wang, P.; Pei, H.; Yan, J.; Shi, J.; Huang, Q.; He, D.; Fan, C.; Zuo, X. *Anal. Chem.* **2014**, *86* (4), 2124−30.
(6) Das, J.; Ivanov, I.; Sargent, E. H.; Kelley, S. O. *J. Am. Chem. Soc.* **2016**, *138* (34), 11009−16.
(7) Chen, S. X.; Seelig, G. *J. Am. Chem. Soc.* **2016**, *138* (15), 5076−86.
(8) Smith, S. J.; Nemr, C. R.; Kelley, S. O. *J. Am. Chem. Soc.* **2017**, *139* (3), 1020−1028.
(9) Zhang, D. Y.; Chen, S. X.; Yin, P. *Nat. Chem.* **2012**, *4* (3), 208−214.
(10) Newman, A. M.; Lovejoy, A. F.; Klass, D. M.; Kurtz, D. M.; Chabon, J. J.; Scherer, F.; Stehr, H.; Liu, C. L.; Bratman, S. V.; Say, C.; Zhou, L.; Carter, J. N.; West, R. B.; Sledge, G. W., Jr; Shrager, J. B.; Loo, B. W., Jr.; Neal, J. W.; Wakelee, H. A.; Diehn, M.; Alizadeh, A. A. *Nat. Biotechnol.* **2016**, *34* (5), 547−555.
(11) Aravanis, A. M.; Lee, M.; Klausner, R. D. *Cell* **2017**, *168* (4), 571−574.
(12) Chen, G.; Mosier, S.; Gocke, C. D.; Lin, M. T.; Eshleman, J. R. *Mol. Diagn. Ther.* **2014**, *18* (5), 587−93.
(13) Potapov, V.; Ong, J. L. *PLoS One* **2017**, *12* (1), e0169774.
(14) Chen, L.; Liu, P.; Evans, T. C., Jr.; Ettwiller, L. M. *Science* **2017**, *355* (6326), 752−756.
(15) Wang, J. S.; Zhang, D. Y. *Nat. Chem.* **2015**, *7* (7), 545−53.
(16) Cohen, L.; Hartman, M. R.; Amardey-Wellington, A.; Walt, D. R. *Nucleic Acids Res.* **2017**, *45* (14), e137.
(17) Li, Z.; Zhou, X.; Li, L.; Liu, S.; Wang, C.; Yu, C.; Su, X. *Anal. Chem.* **2018**, *90* (11), 6804−6810.
(18) Johnson-Buck, A.; Su, X.; Giraldez, M. D.; Zhao, M.; Tewari, M.; Walter, N. G. *Nat. Biotechnol.* **2015**, *33* (7), 730−732.
(19) Watanabe, M.; Kawaguchi, T.; Isa, S.; Ando, M.; Tamiya, A.; Kubo, A.; Saka, H.; Takeo, S.; Adachi, H.; Tagawa, T.; Kakegawa, S.; Yamashita, M.; Kataoka, K.; Ichinose, Y.; Takeuchi, Y.; Sakamoto, K.; Matsumura, A.; Koh, Y. *Clin. Cancer Res.* **2015**, *21* (15), 3552−60.
(20) Zadeh, J. N.; Steenberg, C. D.; Bois, J. S.; Wolfe, B. R.; Pierce, M. B.; Khan, A. R.; Dirks, R. M.; Pierce, N. A. *J. Comput. Chem.* **2011**, *32* (1), 170−3.
(21) Johnson-Buck, A.; Li, J.; Tewari, M.; Walter, N. G. A guide to nucleic acid detection by single-molecule kinetic fingerprinting. Methods 2018 (in press, accepted manuscript), DOI: 10.1016/j.ymeth.2018.08.002. Available online: Aug 10, 2018 (accessed Aug 11, 2018).
(22) Blanco, M.; Walter, N. G. *Methods Enzymol.* **2010**, *472*, 153−78.
(23) Guizar-Sicairos, M.; Thurman, S. T.; Fienup, J. R. *Opt. Lett.* **2008**, *33* (2), 156−8.
(24) Cisse, I. I.; Kim, H.; Ha, T. *Nat. Struct. Mol. Biol.* **2012**, *19* (6), 623−7.
(25) Dupuis, N. F.; Holmstrom, E. D.; Nesbitt, D. J. *Biophys. J.* **2013**, *105* (3), 756−66.
(26) Hess, S. T.; Girirajan, T. P.; Mason, M. D. *Biophys. J.* **2006**, *91* (11), 4258−72.
(27) Rust, M. J.; Bates, M.; Zhuang, X. *Nat. Methods* **2006**, *3* (10), 793−5.
(28) Jones, G. M.; Busby, E.; Garson, J. A.; Grant, P. R.; Nastouli, E.; Devonshire, A. S.; Whale, A. S. *Biomol. Detect. Quantif* **2016**, *10*, 31−33.
(29) Cohen, J. D.; Li, L.; Wang, Y.; Thoburn, C.; Afsari, B.; Danilova, L.; Douville, C.; Javed, A. A.; Wong, F.; Mattox, A.; Hruban, R. H.; Wolfgang, C. L.; Goggins, M. G.; Dal Molin, M.; Wang, T. L.; Roden, R.; Klein, A. P.; Ptak, J.; Dobbyn, L.; Schaefer, J.; Silliman, N.; Popoli, M.; Vogelstein, J. T.; Browne, J. D.; Schoen, R. E.; Brand, R. E.; Tie, J.; Gibbs, P.; Wong, H. L.; Mansfield, A. S.; Jen, J.; Hanash, S. M.; Falconi, M.; Allen, P. J.; Zhou, S.; Bettegowda, C.; Diaz, L. A., Jr.; Tomasetti, C.; Kinzler, K. W.; Vogelstein, B.; Lennon, A. M.; Papadopoulos, N. *Science* **2018**, *359* (6378), 926−930.
(30) Heitzer, E.; Ulz, P.; Geigl, J. B. *Clin. Chem.* **2015**, *61* (1), 112−23.
(31) Husain, H.; Melnikova, V. O.; Kosco, K.; Woodward, B.; More, S.; Pingle, S. C.; Weihe, E.; Park, B. H.; Tewari, M.; Erlander, M. G.; Cohen, E.; Lippman, S. M.; Kurzrock, R. *Clin. Cancer Res.* **2017**, *23* (16), 4716−4723.
(32) Wang, S.; Cang, S.; Liu, D. *J. Hematol. Oncol.* **2016**, *9*, 34.
(33) Milbury, C. A.; Zhong, Q.; Lin, J.; Williams, M.; Olson, J.; Link, D. R.; Hutchison, B. *Biomol. Detect. Quantif* **2014**, *1* (1), 8−22.
(34) Fijalkowska, I. J.; Schaaper, R. M.; Jonczyk, P. *FEMS Microbiol. Rev.* **2012**, *36* (6), 1105−21.
(35) Kreutzer, D. A.; Essigmann, J. M. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95* (7), 3578−82.
(36) Purmal, A. A.; Lampman, G. W.; Bond, J. P.; Hatahet, Z.; Wallace, S. S. *J. Biol. Chem.* **1998**, *273* (16), 10026−35.
(37) Gootenberg, J. S.; Abudayyeh, O. O.; Kellner, M. J.; Joung, J.; Collins, J. J.; Zhang, F. *Science* **2018**, *360* (6387), 439−444.
(38) Wolfe, B. R.; Porubsky, N. J.; Zadeh, J. N.; Dirks, R. M.; Pierce, N. A. *J. Am. Chem. Soc.* **2017**, *139* (8), 3134−3144.
(39) Su, X.; Li, L.; Wang, S.; Hao, D.; Wang, L.; Yu, C. *Sci. Rep.* **2017**, *7*, 43824.
(40) Geiss, G. K.; Bumgarner, R. E.; Birditt, B.; Dahl, T.; Dowidar, N.; Dunaway, D. L.; Fell, H. P.; Ferree, S.; George, R. D.; Grogan, T.; James, J. J.; Maysuria, M.; Mitton, J. D.; Oliveri, P.; Osborn, J. L.; Peng, T.; Ratcliffe, A. L.; Webster, P. J.; Davidson, E. H.; Hood, L.; Dimitrov, K. *Nat. Biotechnol.* **2008**, *26* (3), 317−25.
(41) Koide, K.; Sekizawa, A.; Iwasaki, M.; Matsuoka, R.; Honma, S.; Farina, A.; Saito, H.; Okai, T. *Prenatal Diagn.* **2005**, *25* (7), 604−607.
(42) Stephanie, C.; Lee, S. W.; Jiang, P.; Leung, T. Y.; Chan, K. A.; Chiu, R. W.; Lo, Y. D. *Clin. Chem.* **2013**, *59* (8), 1228−1237.
(43) Shekhtman, E. M.; Anne, K.; Melkonyan, H. S.; Robbins, D. J.; Warsof, S. L.; Umansky, S. R. *Clin. Chem.* **2009**, *55* (4), 723−729.
(44) Tsui, N. B.; Jiang, P.; Chow, K. C.; Su, X.; Leung, T. Y.; Sun, H.; Chan, K. A.; Chiu, R. W.; Lo, Y. D. *PLoS One* **2012**, *7* (10), e48319.
(45) Su, Y.-H.; Wang, M.; Brenner, D. E.; Norton, P. A.; Block, T. M. *Ann. N. Y. Acad. Sci.* **2008**, *1137* (1), 197−206.
(46) Su, Y. H.; Wang, M.; Block, T. M.; Landt, O.; Botezatu, I.; Serdyuk, O.; Lichtenstein, A.; Melkonyan, H.; Tomei, L. D.; Umansky, S. *Ann. N. Y. Acad. Sci.* **2004**, *1022* (1), 81−89.

DOI: 10.1021/jacs.8b06685
J. Am. Chem. Soc. 2018, 140, 11755−11762

Supporting Information

Ultra-specific and Amplification-free Quantification of Mutant DNA by Single-molecule Kinetic Fingerprinting

Stephen L. Hayward[†‡], Paul E. Lund[§‡], Qing Kang[†], Alexander Johnson-Buck[†§‖*], Muneesh Tewari[†‖#Ø⊥*], Nils G. Walter[§‖Ø*]

[†]Department of Internal Medicine, Division of Hematology/Oncology, [§]Single Molecule Analysis Group, Department of Chemistry, [‖]Center for RNA Biomedicine, [#]Department of Biomedical Engineering, [Ø]Center for Computational Medicine and Bioinformatics, [⊥]Biointerfaces Institute, University of Michigan, Ann Arbor, Michigan 48109, United States.

‡These authors contributed equally to this work.

* Correspondence should be addressed to A.J.B. (alebuck@med.umich.edu), M.T. (mtewari@med.umich.edu), and N.G.W. (nwalter@umich.edu).

CONTENTS

Experimental Details

Figures S1-12

Tables S1-3

Supplementary Note

Supplementary Protocol for Cleaning and Surface Passivation of Coverslips

Supplementary References

EXPERIMENTAL DETAILS

Preparation of cloned T790 and T790M DNA substrates for use in SiMREPS assays. Double-stranded DNA substrates were prepared by annealing chemically synthesized oligonucleotides ("Synthetic" substrates, see above), or by restriction enzyme digestion of plasmid DNA ("Cloned" substrates). The cloning strategy used was such that the resulting Cloned DNA substrates were identical in sequence to the Synthetic substrates, save that sequence fidelity in the Cloned substrates is expected to be primarily dictated by DNA replication error rates in *E. coli*, which is on the order of 10^{-9}-10^{-11} errors per base pair[1]. All of the plasmids used in this study are available through Addgene (www.addgene.org). Plasmids containing 4 copies of the desired wild-type or mutant fragment were prepared using standard molecular biology techniques using enzymes purchased from New England Biolabs. Briefly, partially complementary forward and reverse oligonucleotides (5'-NNN GAG TCH BHV WAT CAC GCA GCT CAT GCC CTT CGT AGG AC-3', 5'-NNN GAC GTC CTA CGA AGG GCA TGA GCT GCG TGA T-3', wild-type T790; 5'-NNN GAG TCH BHV WAT CAT GCA GCT CAT GCC CTT CGT AGG AC-3', 5'-NNN GAC GTC CTA CGA AGG GCA TGA GCT GCA TGA T-3', mutant T790M) were purchased from IDT and phosphorylated using T4 polynucleotide kinase (#M0201S) according to the manufacturer's recommendations. Phosphorylated oligonucleotides were annealed by combining complementary strands at 1:1 stoichiometry and heating and slow cooling using a thermocycler. Annealed oligonucleotides were then filled in using either T4 DNA polymerase (#M0203S) or *E. coli* DNA Polymerase I (large Klenow fragment, #M0210S) to create fully-complementary, blunt-ended fragments, then ligated into pUC19 vector linearized with SmaI, and transformed into the JM109 strain of *E. coli* (Promega, P9751). Colonies were screened by PCR (forward primer: 5'-TTT CCC AGT CAC GAC GTT-3', reverse: 5'-TGT GGA ATT GTG AGC GGA-3') and clones containing ≥4 copies of the insert were selected. The plasmid sequences for all clones were confirmed by Sanger sequencing (**Figure S8a**-**b**). Wild-type T790 and mutant T790M Cloned substrates were prepared from their respective multi-insert plasmids (pUC19_4xT790_28bp and pUC19_4xT790M_28bp) by digestion with MlyI and ZraI (#R0610S and #R0659S) as described below.

Restriction digestion and enzymatic treatment to remove damaged DNA. Cloned DNA substrates used in experiments to assess the impact of deamination and other forms of DNA damage were prepared by restriction enzyme digestion of the multi-insert plasmid DNA, following a modified protocol. All enzymes were purchased from New England Biolabs. Plasmid DNA was first treated with a combination of Uracil-DNA glycosylase (UDG, #M0280S), Formamidopyrimidine-DNA glycosylase (Fpg, # M0240S), and Endonuclease VIII (EndoVIII, #M0299S), to remove a variety of damaged DNA bases such as uracil, 5-hydroxycytosine, 5-hydroxyuracil, and uracil glycol.

A typical reaction contained 20-80 µg of multi-insert plasmid DNA (pDNA), 1.7 U of UDG per µg pDNA, 5 U Fpg per µg pDNA, and 5 U EndoVIII per µg pDNA, in 1X NEBuffer 1 (10 mM Bis-Tris-Propane-HCl [pH 7.0 at 25 °C], 10 mM $MgCl_2$, 1 mM DTT) supplemented with 100 µg/mL BSA at a final pDNA concentration of 67 ng/µL (pre-digestion enzyme cocktail treatment). The reaction mixture was incubated at 37 °C for 1 hr 15 min. The reaction was then divided into 75 µL aliquots and each aliquot was purified on a separate column using the QIAquick PCR Purification Kit (Qiagen), with the modification the PB + reaction mixture was passed through the column two times before washing with buffer PE. The eluted pDNA was pooled, and then digested with MylI and ZraI. A typical digestion reaction contained 10-40 µg of pDNA and 5 U of each restriction enzyme per µg pDNA, in 1X NEB CutSmart® Buffer at a final DNA concentration of 45 ng/µL. The reaction mixture was incubated at 37 °C for 1 hr, and then stopped by chelating the available Mg^{2+} by adding 1.5 equivalents of EDTA (pH 8.0).

Where indicated, a second enzyme cocktail treatment (post-digestion) was performed by adding 1.7 U of UDG per µg pDNA, 5 U Fpg per µg pDNA, and 5 U EndoVIII per µg pDNA to the reaction mixture and incubating at 37 °C for 1 hr. The concentration of the resulting treated DNA fragments was determined by gel electrophoresis as described above.

In situ UDG treatment of samples was performed as follows: following the 1-hr capture step incubation, the samples cells were washed 3 times with 100 µL of 1X NEBuffer 3. The surface-immobilized DNA fragments were then treated by adding 50 µL 1X NEBuffer 3 containing 1 U of UDG to each sample cell and incubating at 37 °C in a humidified chamber for 30 min. The sample cells were then washed 3 times with 4X PBS and imaged following the standard SiMREPS assay protocol described in the **Experimental Section** of the main text.

Quantification of restriction digested cloned DNA substrates. After restriction enzyme digestion, the concentration of liberated DNA fragments was determined by comparison with a standard curve after gel electrophoresis and SYBR gold staining on a 20% native polyacrylamide gel (**Figure S8c**). The standard curve was constructed by serial dilution of the synthetic dsDNA substrate in annealing buffer (10 mM Tris-HCl [pH 8.0 at 25 °C], 50 mM NaCl, 50 mM EDTA). Loading dye (New England Biolabs, #B7021S) was added to each sample and samples were electrophoresed on a 20% native-PAGE gel (20 cm × 18 cm × 1.5 mm) in 0.5X TBE at 12 W for 4 hr at 4 °C. The gel was then stained with 1X SYBR Gold nucleic acid gel stain (Invitrogen) in milliQ water for 25 min, and then imaged on a Typhoon™ 9410 Variable Mode Imager (GE Healthcare) operating in Fluorescence mode, with 488 nm laser excitation and 560 LP emission filter set. The resulting gel images were quantified in ImageQuant v5.2 (Molecular Dynamics). Only the full-length duplex was considered during quantification.

Limit of Detection (LOD) experiments. The Exon 19 deletion LOD assay was performed with synthetic 28 nt Exon 19 deletion (MUT) and Exon 19 (WT) dsDNA. MUT DNA was kept constant at 500 fM, while WT DNA was varied between 0 and 500 nM. The standard denaturation protocol (95 °C, 3 min) was employed. The T790M LOD assay was performed with multi-insert plasmid cloned T790M (MUT) and T790 (WT) treated with UDG, FpG, and EndoVIII pre- and post-digestion (see above for details). Furthermore, the milder thermal denaturing condition (80 °C, 1 min) and *in situ* UDG treatment (see above) was also implemented. MUT DNA was kept constant at 50 fM, while WT was varied between 0 to 50 nM.

Super-resolution data analysis. Movies of fluorescent probe binding and dissociation were subjected to super-resolution analysis as follows. First, lateral microscope stage drift was corrected using custom MATLAB code that incorporates subpixel image registration through the function dftregistration[2]. Next, the intensity of each movie frame *n* was subtracted from that of each corresponding frame *n*-2 (to account for the fact that many binding and dissociation events occur partway through a movie frame), resulting in a new movie whose pixel intensities reflect frame-to-frame changes in fluorescence intensity arising from

probe binding or dissociation (**Figure S4.**). These changes in fluorescence intensity were better resolved than the raw fluorescence intensity signal, and were localized with subpixel accuracy (s.d. typically <100 nm) using the MATLAB function FastPeakFind (Adi Natan). The resulting set of localizations was subjected to agglomerative hierarchical clustering by (*x*,*y*) coordinates with a distance cutoff of 2 pixels (524 nm). Finally, each cluster of localizations was subjected to filtering according to the time intervals between consecutive binding (signal increase) and dissociation (signal decrease) events, as well as the overall event frequency. A cluster of events was considered to have arisen from repeated fluorescent probe binding to a single target DNA molecule if it satisfied the following criteria, which were empirically chosen to give rise to the highest sensitivity and specificity for the probes and targets examined:

- >15 binding and/or dissociation events within the observation window (10 min)
- Median dwell time in the probe-bound state >3 s
- Median dwell time in the probe-unbound state >10 s
- Maximum dwell time in the probe-unbound state <300 s
- Standard deviation of both *x*- and *y*-position <0.75 pixels (<196.5 nm)

To expedite clustering in fields of view extremely dense with localizations (DNA concentrations $\geq$ 1 pM), only the central 200-by-200 pixel (~2746 μm^2) region was analyzed, and the counts per 14,024-μm^2 field of view were estimated by linear extrapolation.

T790M spiked into healthy donor urine. Urine was collected from a healthy donor under supervision of the University of Michigan IRB and was immediately mixed with EDTA, a nucleic acid preservative, to achieve a final concentration greater than 100 mM. Cellular debris was removed with two sequential 10-minute centrifugation spins (1,600 × *g* then 3,000 × *g*) and then the urine was frozen at -80 °C. On the day of the experiment, the urine was thawed at room temperature, and T790M synthetic 28 bp dsDNA was spiked in at specified concentrations. Both control samples and spike-in samples were denatured using milder denaturing condition (80 °C, 1 min), incubated for 1 hour at room temperature to promote capture, and treated *in situ* with UDG. Data were analyzed by the super-resolution analysis pipeline.

Analysis of FP binding and dissociation rates. SiMREPS experiments were performed essentially as described above, except that only the forward strands (ssDNA) were used at a final concentration of 1 pM. The binding kinetics of FP1 (specific for T790M) to the T790M forward target strand and T790 Synthetic Deaminated forward target strand were measured, as well as the binding kinetics of an FP for T790 (wild-type) to the T790 forward target strand (**Table S2**). Data were analyzed using the diffraction-limited analysis pipeline, and a more stringent signal-to-noise ratio >3.5 was used to create a high-quality set of single-molecule traces from which to determine binding and dissociation rate constants. Bound and unbound dwell times were extracted from the single-molecule traces after idealization in QuB and used to calculate the average τ_{bound} and $\tau_{unbound}$ for each molecule (**Figure S9d, f**). The cumulative distribution functions generated from the full set of all bound and unbound dwell times were fit with single-exponential association functions in OriginPro 8 software with $y_o = 0$ to calculate the FP dissociation rate (k_{off}, **Figure S9e**), and the FP binding rate (k_{on}, **Figure S9g**), respectively. The dissociation constant, K_d, was determined from the following relationship:

$$K_{\mathrm{d}} = \frac{k_{\mathrm{off}}}{k_{\mathrm{on}}}$$

Measurement of DNA re-annealing kinetics in bulk solution. To determine the kinetics of re-annealing of short DNA targets following heat-denaturation, the Qubit HS dsDNA assay kit (Thermo Fisher) was used in conjunction with a Qubit 3.0 fluorimeter following the manufacturer's protocol. A dsDNA-specific dye was used to measure the re-annealing kinetics of complementary DNA strands in solution as a function of added carrier molecule (oligo dT_{10}). The DNA used in these experiments was a 26 bp scramble DNA (refer to **Table S2** for sequence information). This model DNA was kept constant at 15 nM, and oligo dT_{10} was varied between 0 and 1 µM. Each sample was denatured at 95 °C for 3 minutes, and re-annealing kinetics were measured at various time points. Each concentration was performed with three independent measurements and control experiments consisting of single species ssDNA or fully annealed dsDNA without heat-denaturation (as described in the main text, **Experimental Section - Oligonucleotides and DNA substrates**) were used to determine the fraction dsDNA as a function of time.

SUPPORTING FIGURES

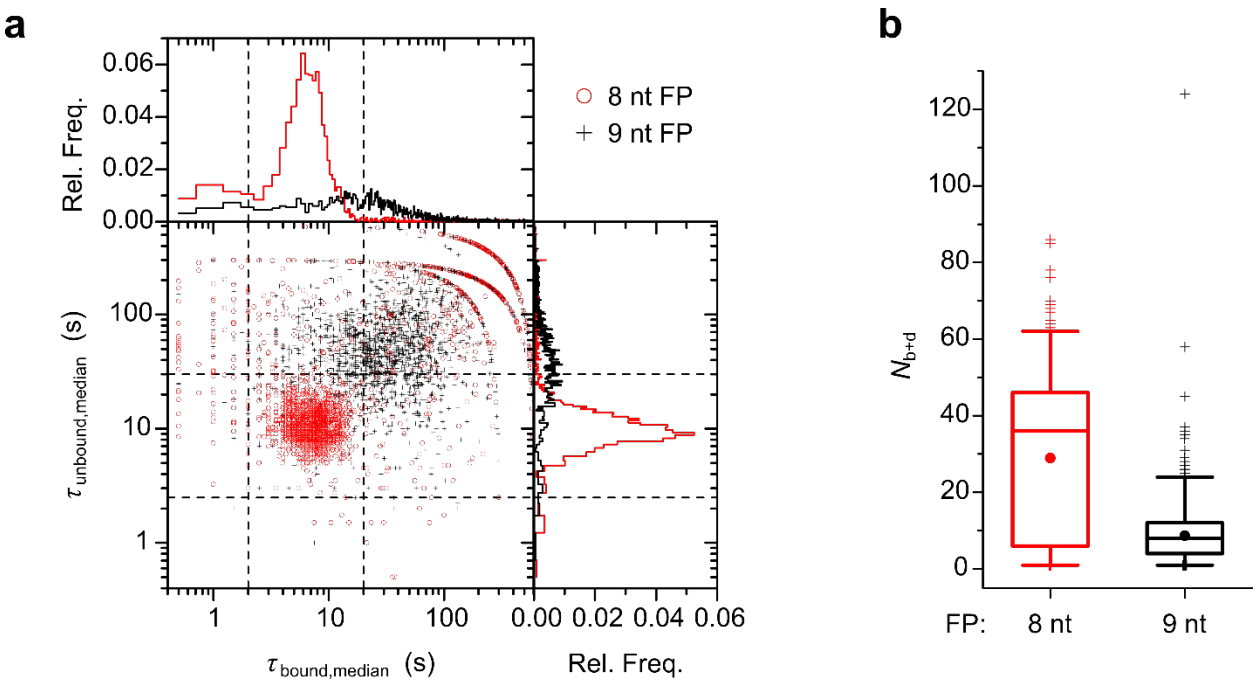
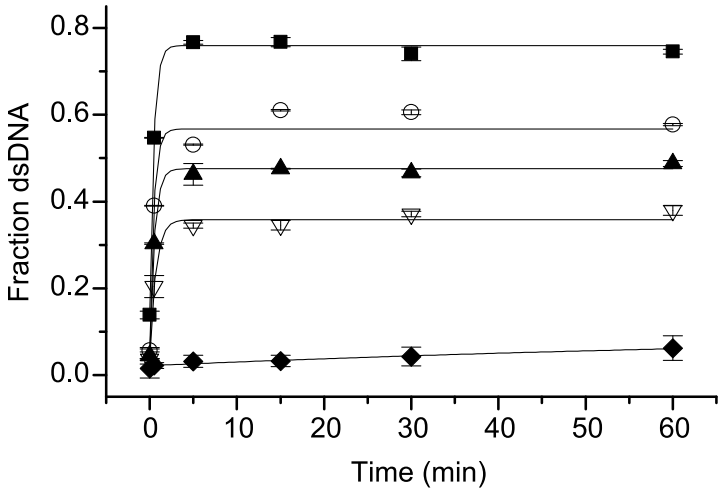


Figure S1 | Effect of FP length on binding kinetics with an *EGFR* Exon 19 Deletion DNA target. (**a**) Kinetic data for the 8 nt (red circle) or 9 nt (black cross) fluorescent probe binding to the Exon 19 Deletion DNA target, which show that reducing the FP length from 9 nt to 8 nt decreases the median FP bound ($\tau_{bound,median}$) and median unbound ($\tau_{unbound,median}$) time. Molecule candidates were plotted by their $\tau_{bound,median}$ and $\tau_{unbound,median}$ values after applying thresholds for fluorescence intensity (>500 a.u.), signal-to-noise (>3). Dashed lines indicating the threshold values ultimately chosen for filtering are shown for reference. While it is possible that such a shift in $\tau_{unbound,median}$ could arise from increased probe concentration in solution or the potential influence of photobleaching on apparent unbound- and bound-state lifetimes, it should be noted that there is good experimental precedent for the observed relationship between k_{on} and short (≤ 10 bp) duplex length such as those used here. This weak dependence (*i.e.*, changes typically one order-of-magnitude or less) of k_{on} (and thus of $\tau_{unbound,median}$) on duplex length has been observed previously in literature, both by our group and by others[3-6]. Rel. Freq., Relative Frequency.(**b**) Box and whisker plot of N_{b+d} values observed for molecule candidates shown in panel **a** after applying thresholds for fluorescence intensity (>500 a.u.), signal-to-noise (>3). Primarily as a result of the shorter lifetime in the bound state, over threefold more binding and dissociation events (N_{b+d}) were observed per 10-minute measurement when using an 8 nt FP, significantly increasing the ability to distinguish signal from background noise. Box boundaries represent the 25- and 75-percentiles; whiskers represent 1- and 99-percentiles; the mean is represented by the filled circle • ; outliers are represented by +.



μM oligo dT_{10}	k_{obs} (min^{-1})	Plateau
■ 0	2.14 (0.07)	0.76
○ 0.015	2.1 (0.4)	0.57
▲ 0.06	1.8 (0.1)	0.48
▽ 0.1	1.4 (0.7)	0.36
◆ 1	0.01 (0.02)	0.10

Figure S2 | Optimizing the thermal denaturation protocol. Bulk solution re-annealing kinetics for a model DNA (26 bp scrambled DNA; 15 nM) were investigated in the presence of oligo dT_{10} carrier at various concentrations using a fluorimeter (**Experimental Details**, above). DNA denaturation was performed at 95 °C for 3 min prior to the experimental analysis. Error bars represent the standard deviation of 3 independent measurements. Fit lines are single-exponential association curves of the form: $y = y_0 + (Plateau - y_0)e^{-k_{obs}t}$, where *t* is time in minutes. Standard error of the fit is shown in parenthesis.

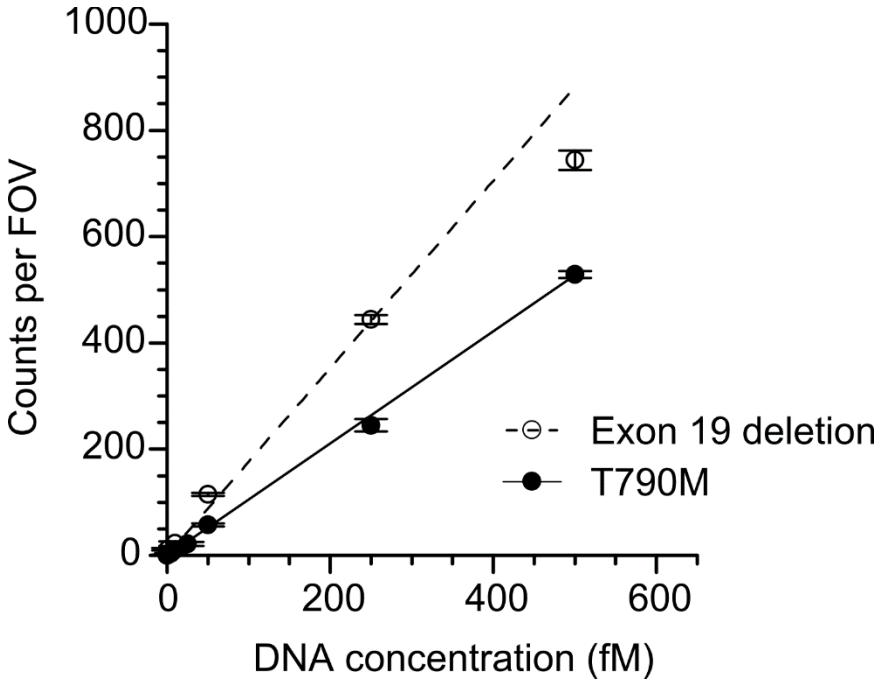


Figure S3 | Diffraction limited analysis

Standard curves from SiMREPS assays using diffraction limited analysis of 0–500 fM *EGFR* Exon 19 deletion and *EGFR* Exon 20 T790M DNA. Data points 0–500 fM are the same as those presented in **Figure 3b-c.** Error-weighted linear fits were constrained to a *y*-intercept of 0; R^2 = 0.973 (Exon 19 deletion), 0.999 (T790M). Data are presented as mean ± s.e.m. of *n* = 3 independent measurements. FOV, field of view.

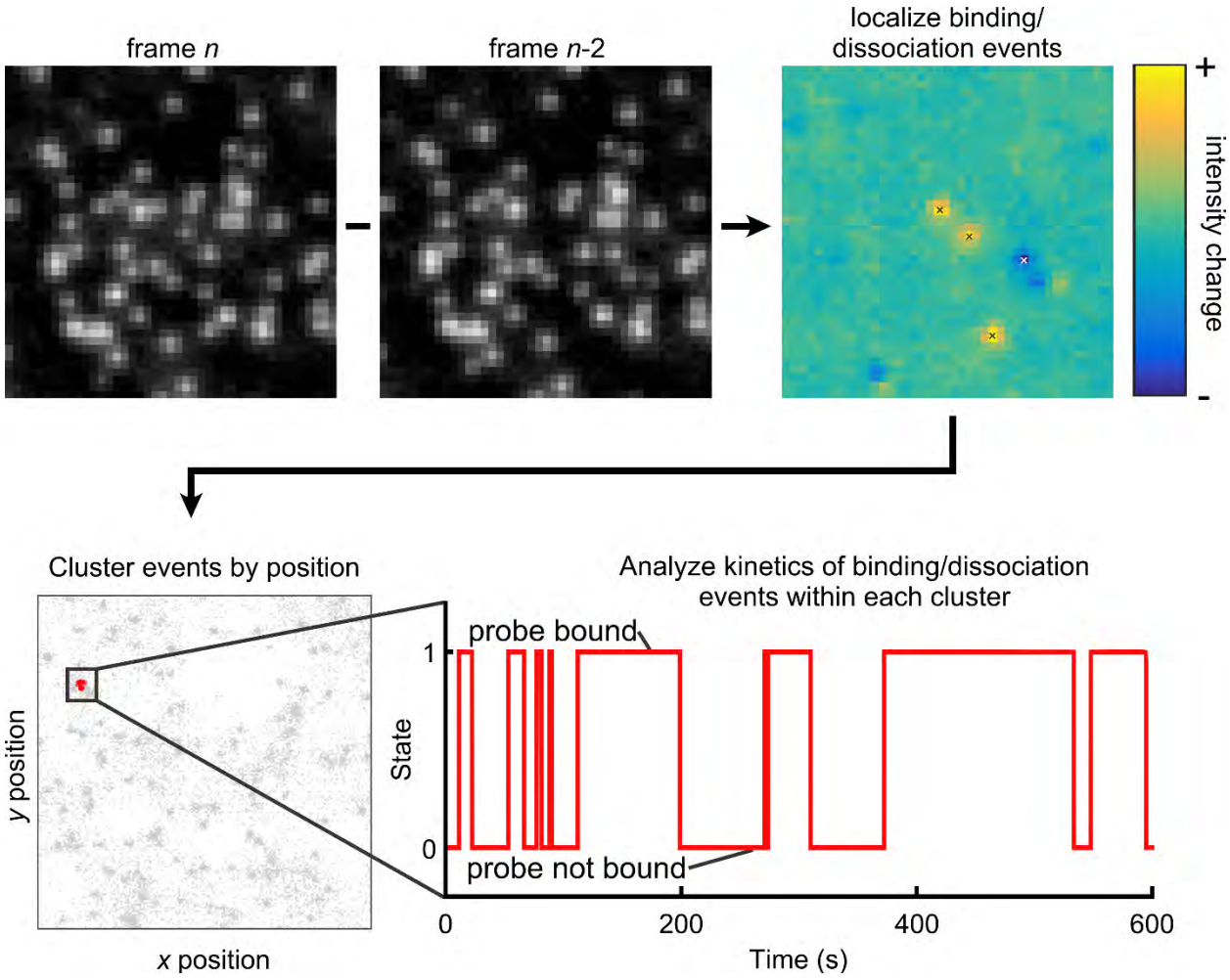
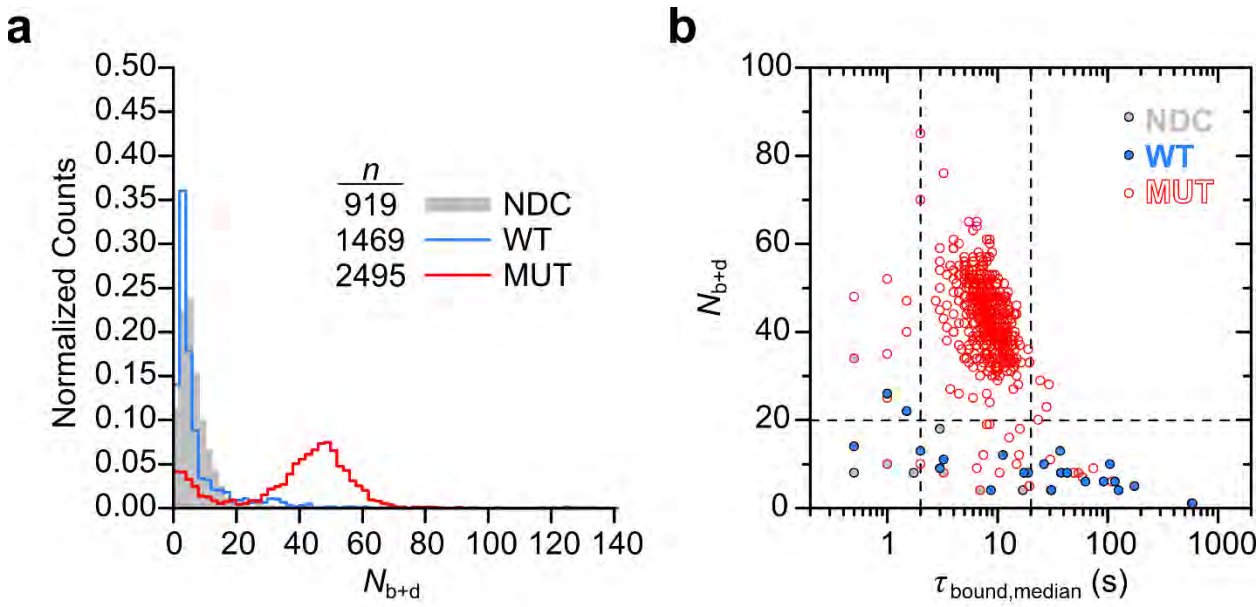


Figure S4 | Schematic of the super-resolution analysis workflow

Frame-by-frame subtraction of fluorescence intensity isolates local fluctuations due to probe binding and dissociation, which are subjected to sub-pixel localization, positional clustering, and kinetic analysis to distinguish true and false positives.



c

Molar Ratio (Exon 19 del : Exon 19 WT)	Counts ± std. dev.
1:0	500 ± 82
1:1	496 ± 35
1:10	544 ± 64
1:100	504 ± 14
1:10,000	352 ± 65
1:100,000	71 ± 25
1:1,000,000	15 ± 2
0:1	0.7 ± 0.6

Figure S5 | Kinetic Fingerprinting analysis of *EGFR* Exon 19 (WT) and *EGFR* Exon 19 deletion (Exon 19 Del, MUT). (**a**) Histogram comparing number of binding and dissociation events (N_{b+d}) between no DNA control (NDC), *EGFR* Exon *EGFR* Exon 19 (WT), and *EGFR* Exon 19 Deletion (MUT) per 10 minute experimental observation. No filtering criteria have been applied. *n*, number of molecule candidates in histogram, three movies per condition. (**b**) Kinetic data from a sample containing only MUT at 500 fM and a sample containing only WT at 500 nM establishing filtering criteria for the experimental LOD analysis using the MUT-specific fluorescent probe. Molecule candidates were plotted after applying thresholds for fluorescence intensity (>500 a.u.), signal-to-noise (>3), and $\tau_{unbound,median}$ between 2.5 and 30 s (inclusive). Dashed lines indicate the threshold values chosen for $\tau_{bound,median}$ and N_{b+d} (**c**) Measured MUT counts at varying MUT:WT molar ratios.

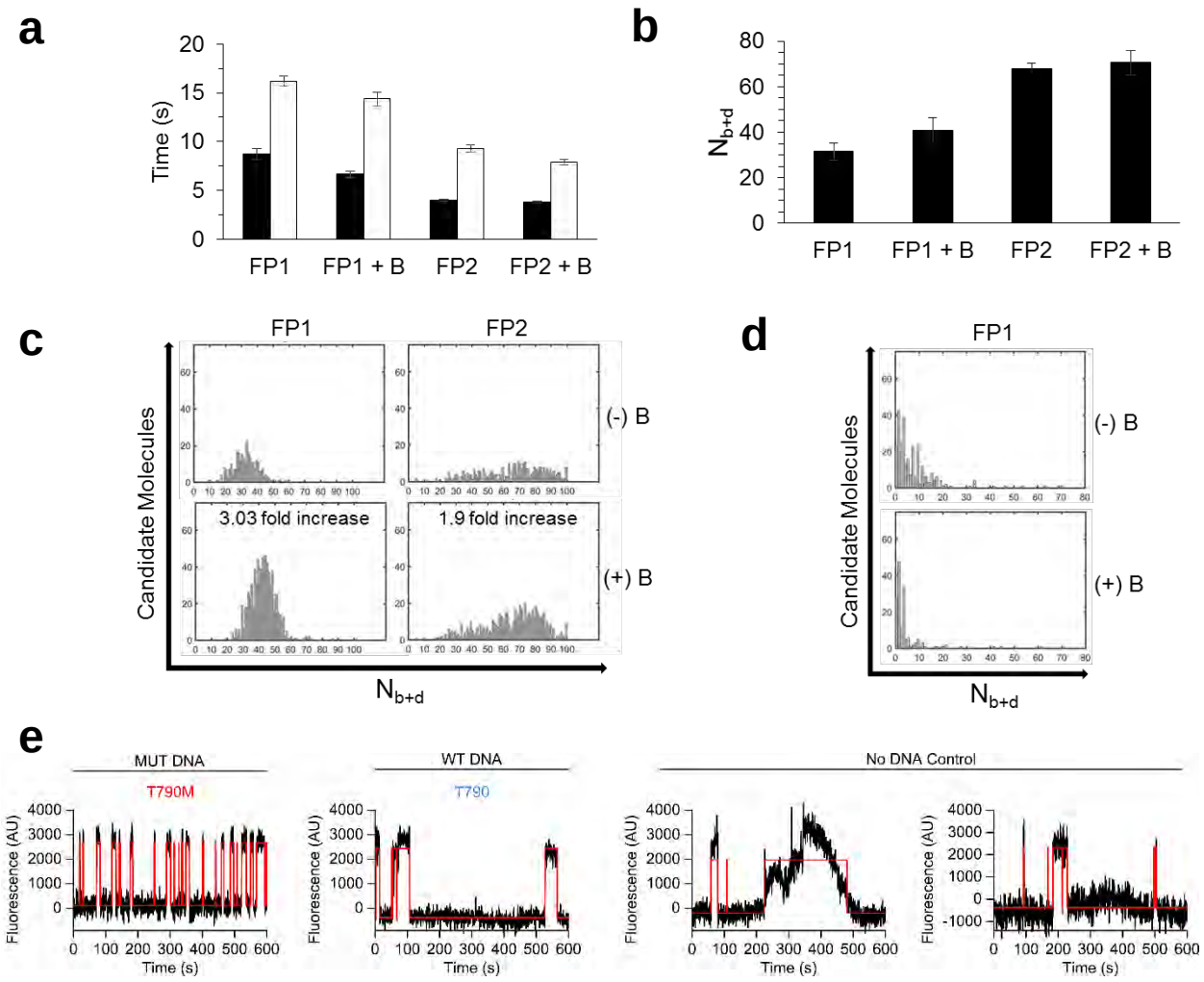


Figure S6 | Optimization of T790M kinetic fingerprinting. (**a**) Both FP1 and FP2 were tested experimentally to determine the effect of FP register on median bound time ($\tau_{bound,median}$, black bar), median unbound time ($\tau_{unbound,median}$, white bar), and (**b**) number of binding and dissociation events (N_{b+d}) per 10-minute measurement. The use of auxiliary probes during imaging were also investigated to suppress rare off-target binding events due to partly homologous sequences. A 10 nt oligo complementary to the LNA capture sequence (LNA blocker, or "B"), was added to sequester transient FP-LNA interactions. The addition of the LNA blocker sequence (**c**) increased candidate molecules/FOV for T790M-containing samples ("fold increase" refers to the increase in counts upon addition of LNA blocker for each respective FP design), while reducing background noise due to spurious interactions as exemplified by no DNA control experiments (**d**) (Refer to **Supplementary Note** for more information). The combination of FP Design #1 with LNA blocker molecule was chosen due to its reliable kinetic fingerprint

for MUT molecules, and reduction of background noise. (**e**) Representative example traces using mutant-specific FP #1 with mutant DNA (MUT), with wild-type DNA (WT), or without DNA (No DNA control) for *EGFR* Exon 20 T790/M (c.2369C>T).

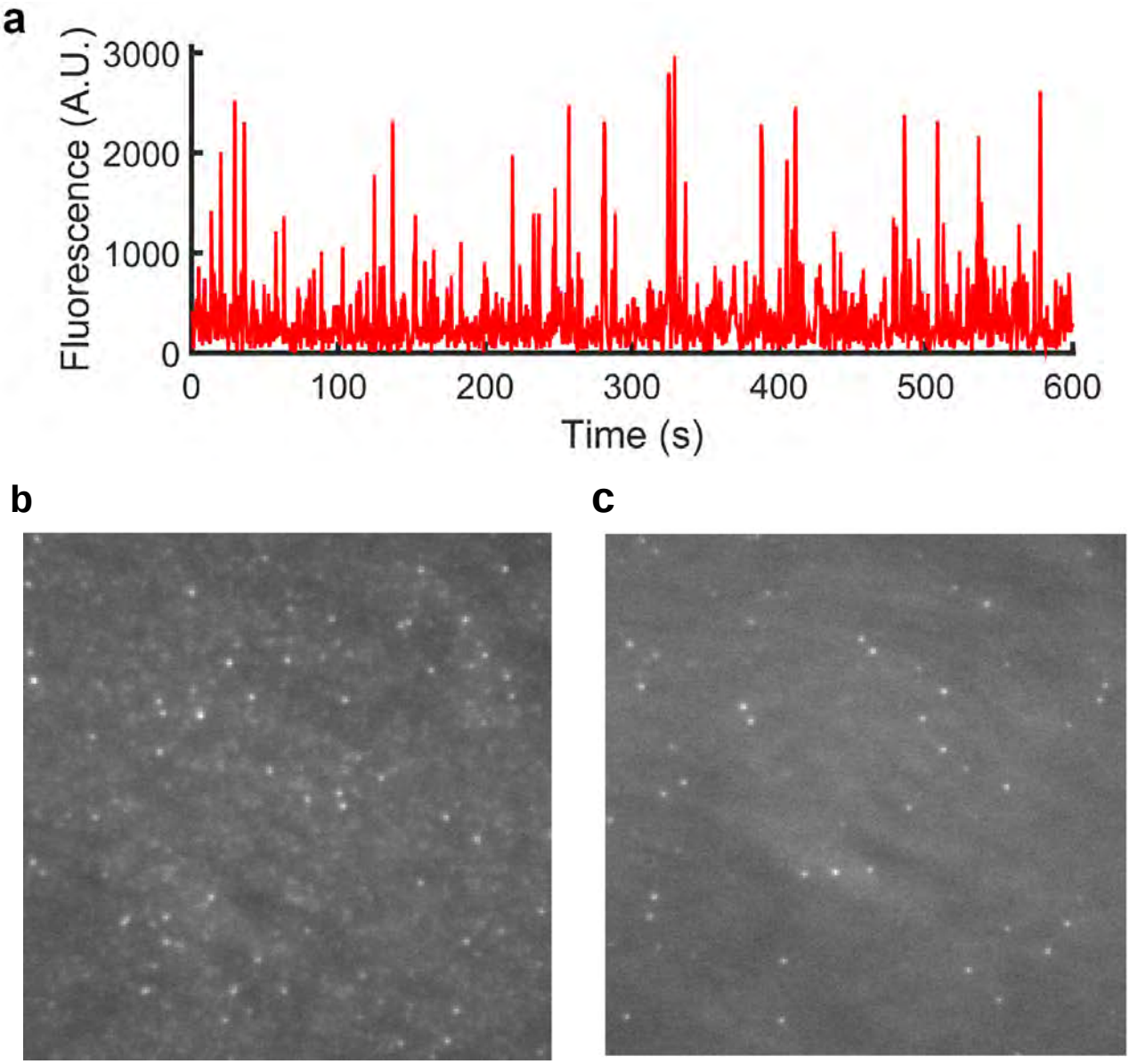


Figure S7 | Competitor probe suppresses background binding of T790M fluorescent probe to WT sequence. (**a**) Binding of fluorescent probe to WT sequence captured at high concentration (5 nM) yields brief binding events, resulting in spurious background signal that can interfere with SiMREPS analysis. These background probe binding events (faint spots in **b**) are suppressed upon addition of 1 µM WT competitor sequence (WTC, **Table S2**) (**c**).

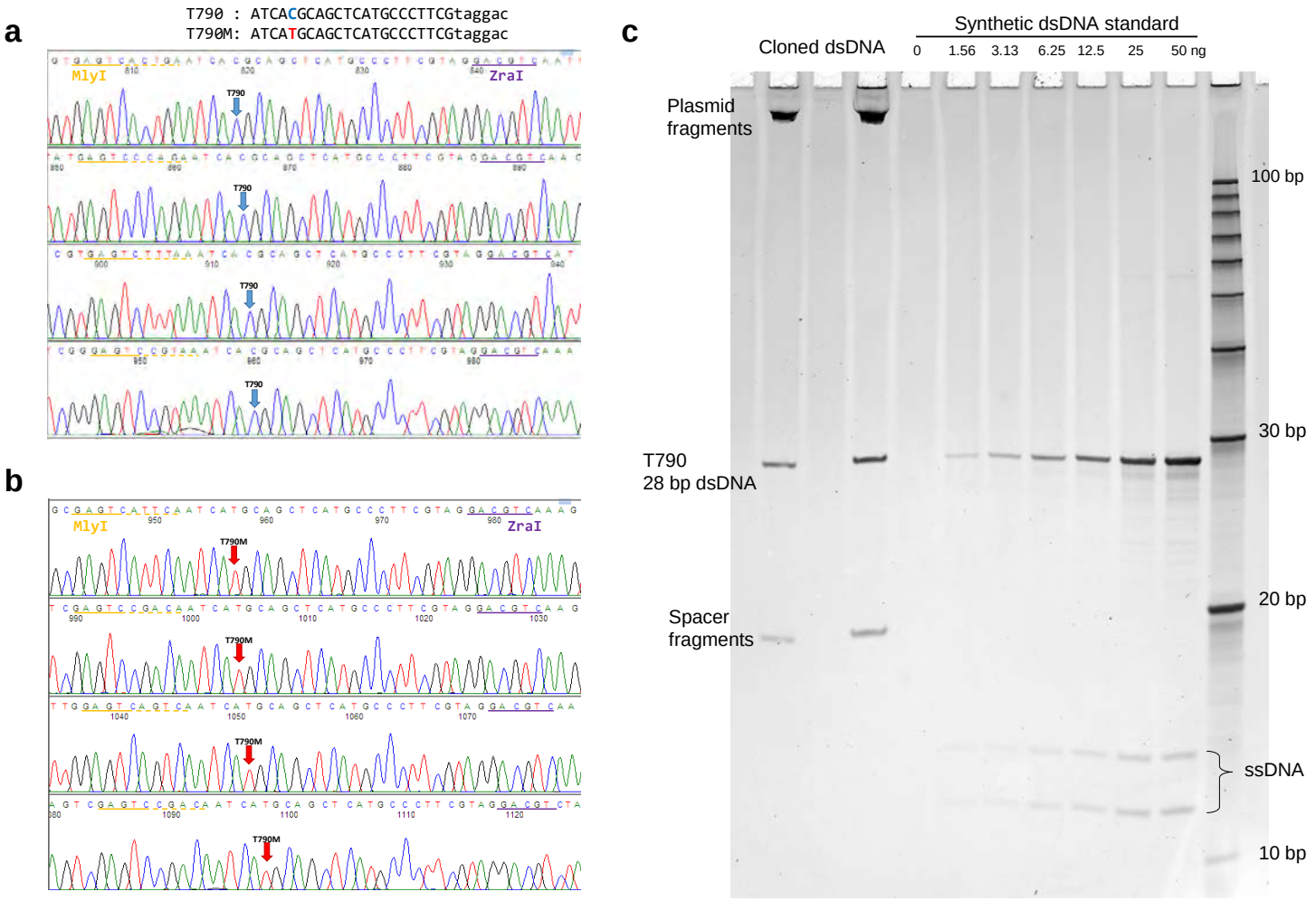


Figure S8 | Generation of high fidelity T790 and T790M dsDNA through a multi-insert plasmid cloning approach in *E. coli* followed by oligo recovery with restriction enzyme digest. Sequencing chromatograms for (**a**) pUC19_4xT790_28bp and (**b**) pUC19_4xT790M_28bp plasmids. Arrows indicate the mutation position. Recognition sequences for MlyI and ZraI are indicated in gold and purple, respectively. (**c**) Representative example of gels used to determine the concentration of liberated DNA substrates after digestion.

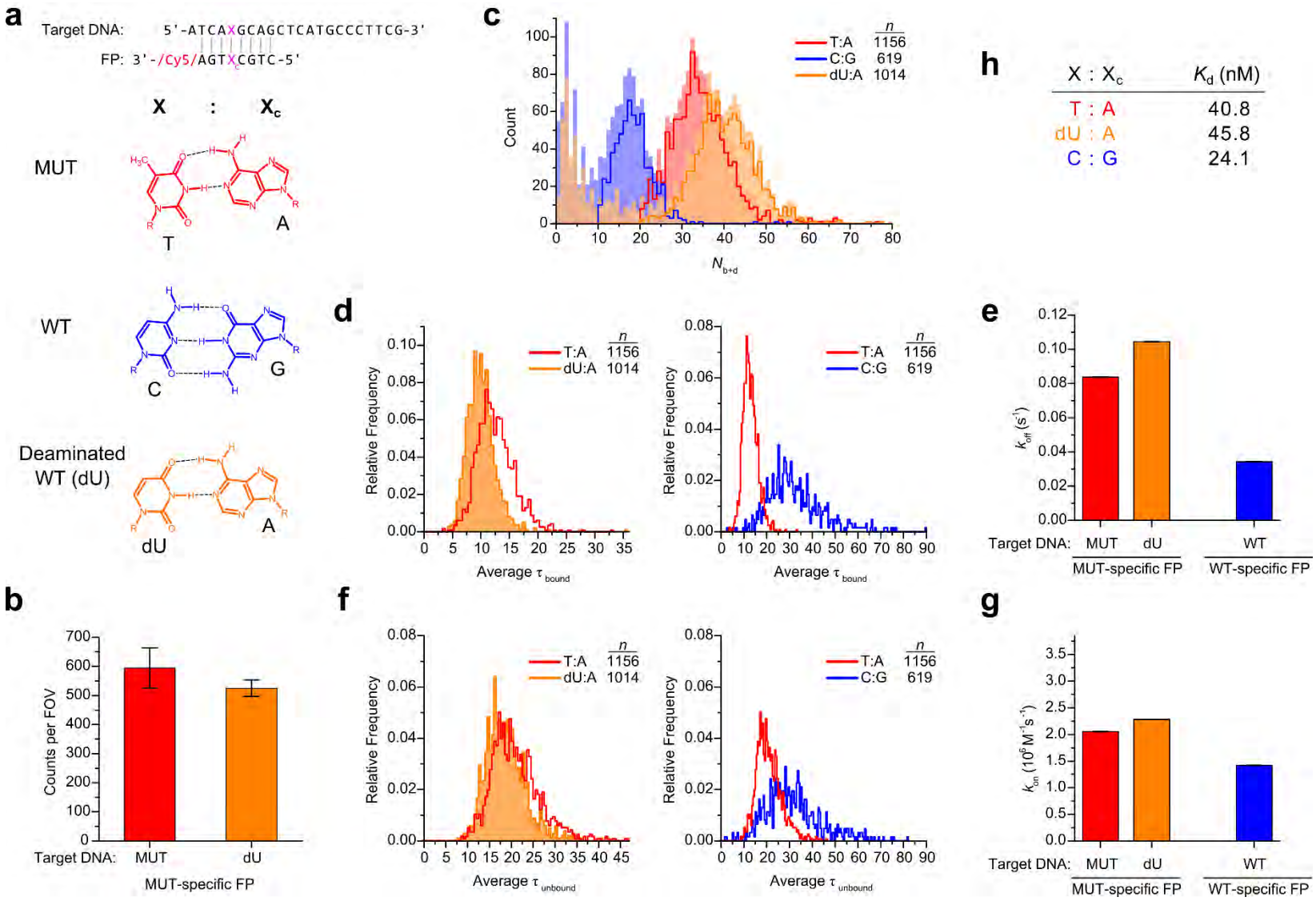


Figure S9 | Characterization of the effects of single-base pair changes on FP binding kinetics. (**a**) Diagram showing the various Target DNA-FP combinations tested, as well as expected base-pairing interactions. The 3' barcode sequence (TAGGAC) present on the

Target DNA is omitted for clarity. dU, deoxyuracil (**b**) Measured counts using the standard kinetic threshold parameters for the T790M-specfic probe (FP1) binding to genuine T790M mutant DNA target (MUT, red bar), or deaminated T790 (dU, orange bar). Data are presented as the mean ± s.e.m. of two independent measurements. (**c**) Histogram of N_{b+d} for different combinations of FP and target DNA per 10-minute experimental observation. The shaded regions indicate the N_{b+d} distribution for all molecule candidates prior to kinetic filtering, and solid lines represent the distribution of candidates that pass filtering (*i.e.*, apparently genuine target molecules). *n*, number of apparently genuine target molecules that pass filtering (**d**) Distribution of per-molecule fluorescent probe average bound dwell time (average τ_{bound}). Left plot: The average τ_{bound} for T790M FP1 (left plot) with T790M (red line) is longer than for the deaminated T790 (orange filled bars). Right plot: Substituting a C:G base pair (blue) for a T:A base pair (red, reproduced from left plot) results in significantly longer average τ_{bound}. *n*, number of molecules represented in the distribution. (**e**) Comparison of the FP dissociation rate constants (k_{off}) calculated from single-exponential fitting of the cumulative distribution function of all bound dwell times in a given condition. Error bars represent the standard error of the fit. (**f**). Same as in (**d**), but for the average time between fluorescent probe binding events (average $\tau_{unbound}$) at 25 nM FP. (**g**) Same as in (**e**), but for the fluorescent probe binding rate constants, calculated from fitting of the cumulative distribution function of unbound dwell times. (**h**) Dissociation equilibrium constants for the Target DNA-FP combinations tested, calculated from binding and dissociation rate constants.

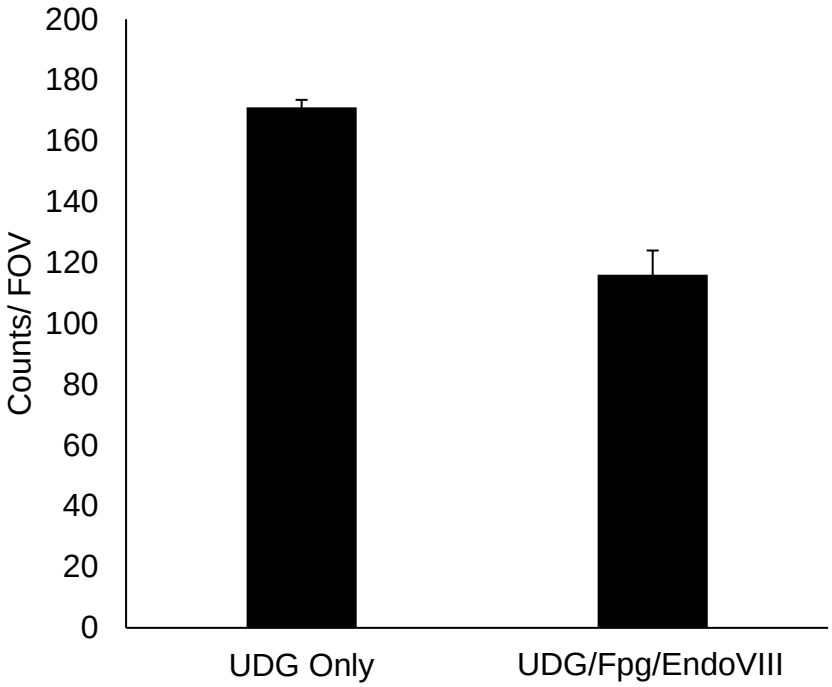


Figure S10 | Quantifying the amount of T790 false positive frequency that is driven by oxidative damage. Cloned multi-insert T790 was treated post digestion with either Uracil DNA glycosylase (UDG) alone or an “enzyme cocktail” containing UDG, formamidopyrimidine-DNA glycosylase (Fpg), and Endonuclease VIII (EndoVIII).

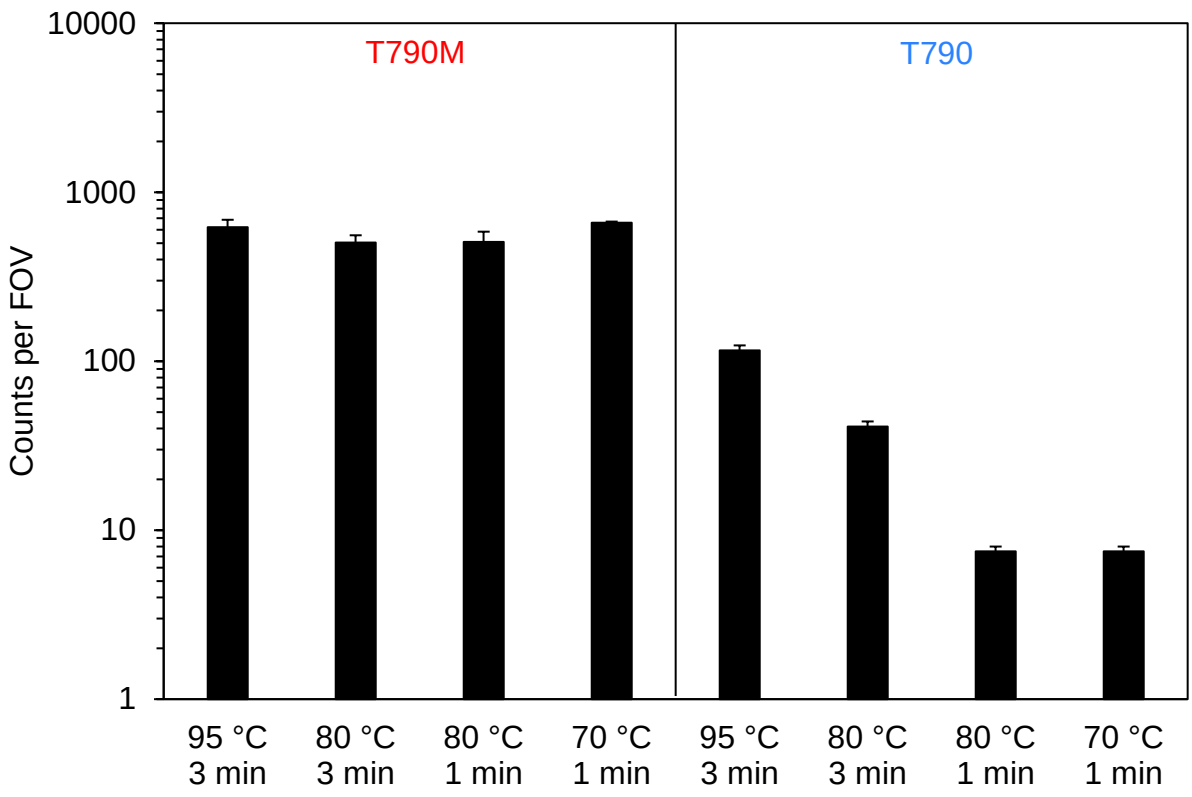
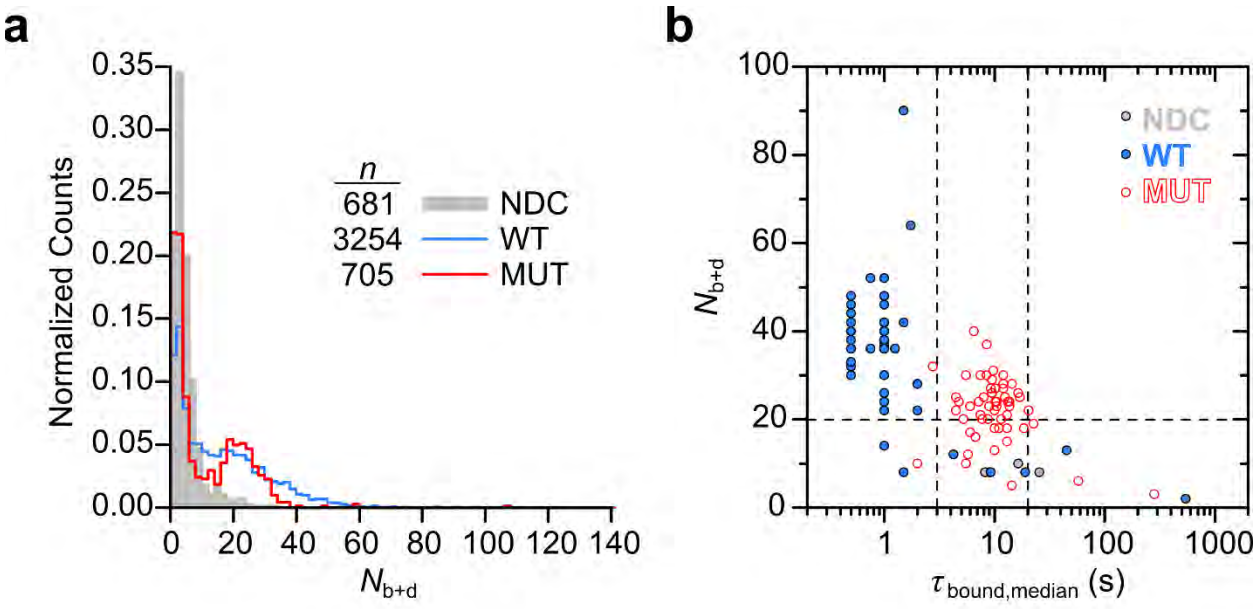


Figure S11 | Effect of denaturing conditions on T790 (WT) false positive frequency. High fidelity cloned 28 bp T790 (5 nM) and T790M (500 fM) DNA fragments were denatured at various temperatures and durations prior to SiMREPS analysis. All DNA in this experiment was pre-treated with an enzyme cocktail containing Uracil DNA glycosylase, formamidopyrimidine-DNA glycosylase, and Endonuclease VIII. Data are presented as the mean ± s.e.m. of 2 independent measurements.



c

Molar Ratio (T790M : T790)	Counts ± std. dev.
1:0	38.7 ± 3.5
1:10,000	38.8 ± 11.5
1:100,000	29.7 ± 4.0
1:1,000,000	12.7 ± 1.5
0:1	1.1 ± 1.1

Figure S12 | Kinetic Fingerprinting analysis of T790 (WT) and T790M (MUT). (**a**) Histogram comparing number of binding and dissociation events (N_{b+d}) between no DNA control (NDC), T790 (WT, 50 nM), and T790M (MUT, 50 fM) per 10–minute No filtering criteria have been applied. *n*, number of molecule candidates in histogram, three movies for NDC and MUT conditions, eight movies for WT condition. (**b**) Kinetic data from a MUT-only sample at 50 fM and a WT-only sample at 50 nM establishing filtering criteria for the experimental LOD analysis using the MUT-specific fluorescent probe. Molecule candidates were plotted after applying thresholds for fluorescence intensity (>500 a.u.), signal-to-noise (>3), and $\tau_{unbound,median}$ between 3 and 30 s (inclusive). Dashed lines indicate the threshold values chosen for $\tau_{bound,median}$ and N_{b+d} (**c**) Measured MUT counts at varying MUT:WT molar ratios.

Table S1 | Characteristics of cell-free DNA in various example biofluids

Biofluid		Cell-free DNA in biofluid (ng/mL)	GE[a]/mL	Molar GE	Size (bp)	Reference
Blood						
	Range	0 to >1,000	0 to >280,000	0 to >4.6E-16	70 – 200,	7
	Average	180 (disease)	50,000	8.3E-17	up to many	
		30 (healthy)	8,000	1.4E-17	kb	
Urine						
	Range	0.9 – 9.46	250 – 2,600	4.2E-19 – 4.4E-18	50 – 200; up to many kb	8-11
Cerebrospinal Fluid						
	Range	2 – 2,400	560 – 670,000	9.3E-19 – 1.1E-15	100 – 300	12-13
	Average	155		7.2E-17		

[a]GE, genome equivalent. The weight of the human genome is estimated to be 3.59 x 10^{-12} g.
https://www.idtdna.com/pages/education/biotech-basics, Molecular Facts and Figures

Table S2 | DNA, LNA, and FP sequences for both Exon 19/deletion and T790/M.

EGFR Exon 19 Deletion 28 nt Synthetic DNA forward	**5'**/TTC CCG TCG CTA TCA AGA CAT CTA GGA C/**3'**
EGFR Exon 19 Deletion 28 nt Synthetic DNA reverse	**5'**/GTC CTA GAT GTC TTG ATA GCG ACG GGA A/**3'**
EGFR Exon 19 28 nt Synthetic DNA forward	**5'**/TTC CCG TCG CTA TCA AGG AAT TTA GGA C/**3'**
EGFR Exon 19 28 nt Synthetic DNA reverse	**5'**/GTC CTA AAT TCC TTG ATA GCG ACG GGA A/**3'**
Capture LNA for *EGFR* Exon 19	**5'**/+AG+CG+ACG+GG+AA/Biotin TEG/**3'**
9 nt FP for *EGFR* Exon 19 Deletion	**5'**/Cy5/GAT GTC TTG/**3'**
8 nt FP for *EGFR* Exon 19 Deletion	**5'**/Cy5/ATG TCT TG/**3'**
26 bp Scramble DNA forward	**5'/**GAT TAC TTG TAT CGT AGT AAG GGC TA/**3'**
26 bp Scramble DNA reverse	**5'**/TAG CCC TTA CTA CGA TAC AAG TAA TC/**3'**
T790M 28 nt Synthetic DNA forward	**5'**/ATC ATG CAG CTC ATG CCC TTC GTA GGA C/**3'**
T790M 28 nt Synthetic DNA reverse	**5'**/GTC CTA CGA AGG GCA TGA GCT GCA TGA T/**3'**
T790 Synthetic 28 nt DNA forward	**5'**/ATC ACG CAG CTC ATG CCC TTC GTA GGA C/**3'**
T790 Synthetic 28 nt DNA reverse	**5'**/GTC CTA CGA AGG GCA TGA GCT GCG TGA T/**3'**
T790 Synthetic Deaminated 28 nt DNA forward	**5'**/ATC A/ideoxyU/G CAG CTC ATG CCC TTC GTA GGA C/**3'**
Capture LNA for T790M	**5'**/Biotin TEG /C+GAA+GGGCA+TG/**3'**
Capture LNA Blocker for T790M	**5'**/ATG CCC TTC G/**3'**
FP1 for T790M	**5'**/CTG CAT GA/Cy5/**3'**
FP2 for T790M	**5'**/Cy5/TGC ATG AT/**3'**
T790 8 nt Dark Competitor (Wild-type competitor, "WTC")	**5'**/CTG CGT GA/**3'**
FP for T790	**5'**/CTG CGT GA/Cy5/**3'**

The unique barcode identifier is displayed in red. '+' precedes any locked nucleic acid (LNA) nucleotide.

Table S3 | Effect of DNA preparation method, enzymatic removal of damaged DNA bases, and heat denaturing protocol on the number of observed counts for 0.5 pM T790M and false positive frequency for 5,000 pM T790.

Sample ID	Preparation Method; Concentration	DNA Repair Enzyme Treatment[a]	Denaturing Protocol	Counts/ FOV ± s.e.m	False Positive Frequency[b]	Fold Reduction in False Positive Frequency
T790	Synthetic; 5,000 pM	None	95 °C, 3 min	741 ± 34	1 in 3,870	N/A
T790	Cloned; 5,000 pM	None	95 °C, 3 min	513 ± 8	1 in 5,590	1.44
T790	Cloned; 5,000 pM	1. Enzyme Cocktail	95 °C, 3 min	116 ± 8	1 in 24,700	6.39
T790	Cloned; 5,000 pM	1. Enzyme Cocktail 2. *In situ* UDG treatment	95 °C, 3 min	8.5 ± 2.0	1 in 340,000	87
T790	Cloned; 5,000 pM	1. Enzyme Cocktail 2. *In situ* UDG treatment	80 °C, 1 min	1.1 ± 0.4	1 in 2,600,000	673
T790M	Synthetic; 0.5 pM	None	95 °C, 3 min	622 ± 65	N/A	N/A
T790M	Cloned; 0.5 pM	None	95 °C, 3 min	672 ± 29	N/A	N/A
T790M	Cloned; 0.5 pM	1. Enzyme Cocktail	95 °C, 3 min	596 ± 13	N/A	N/A
T790M	Cloned; 0.5 pM	1. Enzyme Cocktail 2. *In situ* UDG treatment	95 °C, 3 min	654 ± 18	N/A	N/A
T790M	Cloned; 0.5 pM	1. Enzyme Cocktail 2. *In situ* UDG treatment	80 °C, 1 min	668 ± 20	N/A	N/A

a. *Enzyme cocktail contains Uracil DNA glycosylase (UDG), formamidopyrimidine-DNA glycosylase, and Endonuclease VIII*

b. *The False Positive Frequency is the reciprocal of* Q_{app} (**Table 1**, main text) and *was calculated as the number of accepted counts divided by the estimated number of WT molecules bound per field of view in an assay conducted with 5 nM WT (2. million), calculated from the* $1{:}10^5$ *MUT:WT experiment as [True Positives (29.7) - False Positives (1.1)]/[False positives (1.1)] x* 10^5 *= 2.6 million.*

SUPPLEMENTARY NOTE

Addition of Auxiliary Competitor Probes

In addition to FP optimization, additional auxiliary competitor probes were included for T790M detection to suppress rare off-target binding events due to partly complementary sequences. These off-target binding events include: 1) FP interaction with unoccupied capture LNAs, and 2) mutant-specific FP interaction with excess WT molecules. To reduce the transient FP-LNA interaction, a 10 nt oligonucleotide complementary to the LNA capture strand (LNA Blocker), was added during the probing step. By adding the LNA Blocker, we observed a more than threefold increase in mutant counts (**Figure S6e**), as well as fewer interactions between the FP and unoccupied capture LNA (**Figure S6f**). We hypothesize that the increase in mutant counts in the presence of the LNA Blocker is driven by an improved ability to resolve individual binding events due to fewer local transient FP-LNA interactions (*i.e.*, local density does not reach the critical detection density of the diffraction limited approach; refer to **Figure 3b-c**), as well as a surface-wide increase in signal-to-noise ratio leading to fewer true mutant molecules lost during analysis.

To reduce transient interactions between the T790M (MUT) specific FP1 (refer to **Figure S6**) and T790 (WT) DNA, especially at conditions where WT DNA is at great excess, we introduced a WT competitor ("WTC"), an unlabeled 8 nt oligo with specific complementarity to the WT DNA sequence, during the probing step (**Figure S7**). By adding the WTC during probing, spurious background signal (**Figure S7b**) due to transient FP1-WT DNA interactions were significantly reduced, leading to increased signal-to-noise ratio (**Figure S7c**) and thus more reliable kinetic analysis. Notably, the WTC did not suppress long-lived false positive signal driven by deamination of cytosine in the WT DNA. These false positives were only removed by the incorporation of DNA repair enzymes into the SIMREPS protocol (**Figure 4c**, **Table S3**) and/or by using lower-temperature denaturation conditions. Sequences of all oligonucleotides used during these optimization experiments are located in **Table S2**.

Calculation of Specificity and Discrimination Factor

The specificity of an assay is calculated from the number of true negatives (*TN*) and false positives (*FP*) according to the relationship

$$Specificity = \frac{TN}{TN+FP} \quad (\mathbf{1})$$

In our SiMREPS assay for ctDNA, *TN* is equal to the number of WT molecules within the field of view that are not detected as MUT sequences, and *FP* is equal to the number of false positives in a WT-only experiment.

$$TN = (Number\ of\ WT\ molecules\ in\ FOV) - FP \quad (\mathbf{2})$$

We may estimate the number of WT molecules per field of view by assuming that the kinetics of capture are identical for MUT and WT molecules and, hence, the ratio of MUT to WT molecules on the surface will be equal to the ratio of MUT concentration (C_{MUT}) to WT concentration (C_{WT}). The number of WT molecules on the surface in an experiment with a concentration ratio $\frac{C_{WT}}{C_{MUT}}$ = 10^6 is then

$$Number\ of\ WT\ molecules\ in\ FOV = TP\ \times \frac{C_{WT}}{C_{MUT}} \quad (\mathbf{3})$$

$$= TP\ \times\ 10^6 \quad (\mathbf{4})$$

and the number of true negatives is calculated as

$$TN = TP \times 10^6 - FP \quad (\mathbf{5})$$

By substituting equation (**5**) into equation (**1**), we obtain

$$Specificity = \frac{TP \times 10^6 - FP}{TP \times 10^6 - FP + FP} \quad (\mathbf{6})$$

$$= \frac{TP \times 10^6 - FP}{TP \times 10^6} \quad (\mathbf{7})$$

$$= 1 - \frac{FP}{TP \times 10^6} \quad (\mathbf{8})$$

Example: If an assay with 500 fM MUT + 5 nM WT yields 12.7 counts, and an assay with 5 nM WT only yields 1.1 count, then *TP* = 12.7 – 1.1 = 11.6, *FP* = 1.1, Specificity = 0.9999999. This assay then has 99.99999% analytical specificity.

The apparent discrimination factor Q_{app} (**Table 1**, main text) is calculated as

$$Q_{\mathrm{app}} = \frac{Estimated\ WT\ molecules\ bound\ per\ FOV}{Apparent\ MUT\ counts\ in\ WT\ sample} \quad (\mathbf{9})$$

The number of WT molecules bound per field of view (FOV) at a given concentration was estimated from the MUT:WT titration data (**Figure 4e** and **Figure S12c**) under the assumption that the rate constant of binding to the capture probe is identical for MUT and WT molecules.

For instance, at 5 nM WT + 50 fM MUT, the number of true-positive MUT counts is estimated as (29.7 – 1.1) = 28.6, and the number of WT molecules bound to the surface at 5 nM WT is thus 28.6 × 100,000 = 2.9 million. In 50 nM WT + 50 fM MUT, the number of true-positive MUT counts is estimated as (12.7 – 1.1) = 11.6, and the number of WT molecules bound to the surface at 50 nM WT is thus 11.6 × 1,000,000 = 11.6 million.

SUPPLEMENTARY PROTOCOL

Cleaning and Surface Passivation of Coverslips

Cleaning:

1) Sonicate in 1 M KOH 30 min. Rinse 5X with milliQ H_2O

2) Make 70 mL base pirahana (50:10:10 milliQ H_2O, NH_4OH, H_2O_2) and heat until just bubbling

   Carefully pour hot pirahana onto coverslips in coplin jar

   Place coplin jar in beaker of hot water and place on hot plate on low setting (2 - 4) to reach ~65°C. Heat for $\geq$ 30 min.

3) Discard pirahana and rinse 5X with milliQ H_2O.

4) Rinse with MeOH using squirt bottle. Sonicate 10 min in MeOH.

   Meanwhile, prepare **0.1 M $NaHCO_3$:** 84 mg $NaHCO_3$ in 10 mL milliQ H_2O.

   Place m-PEG and biotin-PEG powders in drawer and let come to RT while preparing acetone.

5) Discard MeOH. Rinse once with acetone, then sonicate in acetone 10 min.

Passivation:

Reagent volumes are scaled 5 coverslips

6) Combine 49 mL acetone + 1 mL 3-aminopropyltriethoxysilane (APTES).

   Incubate 10 min, sonicate 1 min, then incubate 10 min.

   Meanwhile, weigh out 8 mg biotin-PEG and 80 mg* m-PEG into EP tubes.

   ** Can be scaled down to 7 mg and 70 mg, respectively. 80 mg is almost a full EP tube.*

7) Discard APTES solution, then rinse 5X with milliQ H_2O. Blow dry coverslips using house nitrogen.

8) Dissolve 8 mg biotin-PEG in 320 µL **0.1 M** $NaHCO_3$ buffer, then combine w/ 80 mg m-PEG and vortex.

   Spin down at 10k x *g* for 1 min at RT to remove bubbles.

   Create sandwich of coverslips with >60 µL PEG solution between each layer

9) Let sit in humidified chamber 2-3 hrs or O/N. Protect from light.

10) After PEGylation step, wash coverslips thoroughly with milliQ H_2O and blow dry with N_2.

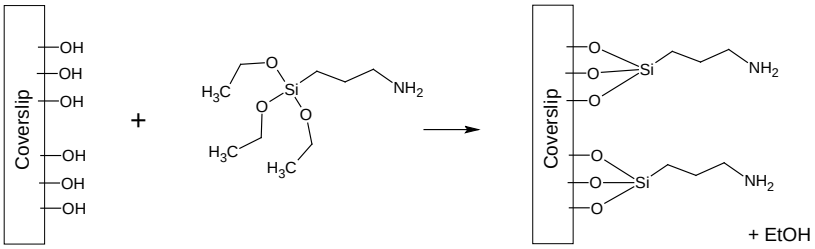
    Meanwhile, prepare **1 M $NaHCO_3$:** 84 mg $NaHCO_3$ in 1 mL milliQ H_2O.

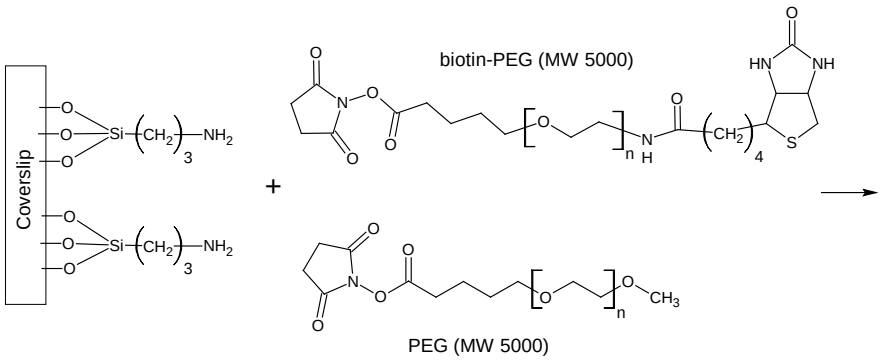
11) Incubate with Disulfosuccinimidyltartrate (DST) in **1 M** $NaHCO_3$ for 30 min protected from light. Use 16 mg DST in 560 µL buffer.

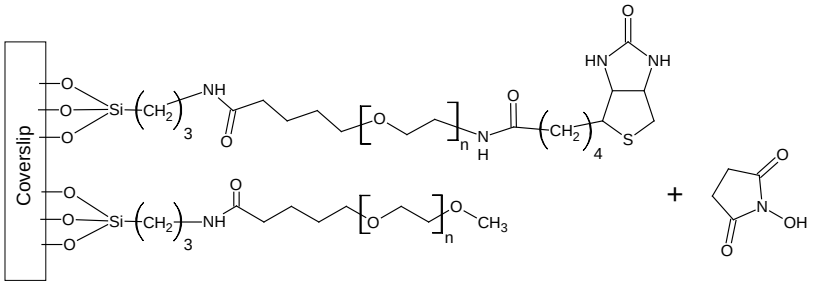
After incubation, thoroughly wash coverslips with milliQ H_2O, blow dry with N_2. Store in coverslip case parafilmed under N_2.

Example Reagents	Vendor	Catalog #
Methanol (HPLC grade)	Fisher Scientific	A452-4
Acetone (HPLC grade)	Fisher Scientific	A949-4
Potassium Hydroxide	Fisher Scientific	P250-1
Ammonium Hydroxide, 28-30 wt. % Solution of NH_3 in Water	Fisher Scientific	AC20584-0025
Hydrogen Peroxide, 35 wt. % Solution in Water, Stabilized	Fisher Scientific	AC20246-0010
(3-Aminopropyl)triethoxysilane, >= 98%	Sigma-Aldrich	A3648-100ML
Sodium bicarbonate 99+%, Extra Pure, ACROS Organics	Fisher Scientific	AC12336-0010
mPEG-Succinimidyl Valerate, MW 5,000 - 1 gram	Laysan Bio Inc.	MPEG-SVA-5000-1g
Biotin-PEG-SVA, MW 5,000 - 500 mg	Laysan Bio Inc.	Biotin-PEG-SVA-5000-1g
Disulfosuccinimidyltartrate, 100 mg	Soltec Ventures	CL107
Micro Cover Glasses, Rectangular, No. 1 ½; 24 x 50 mm	VWR	48393-241

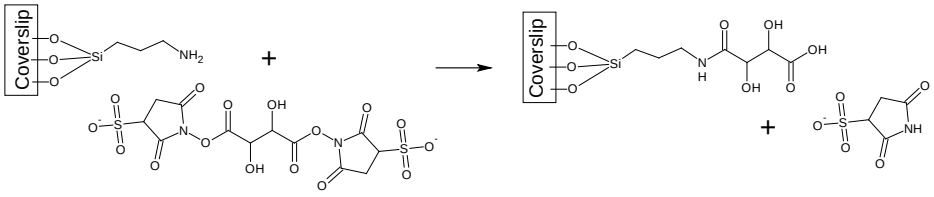
Reaction scheme:







Scavenge any remaining free silane-amines with DST:



SUPPLEMENTARY REFERENCES

1. Fijalkowska, I. J.; Schaaper, R. M.; Jonczyk, P., DNA replication fidelity in *Escherichia coli*: a multi-DNA polymerase affair. *FEMS Microbiol. Rev.* **2012,** *36* (6), 1105-21.
2. Guizar-Sicairos, M.; Thurman, S. T.; Fienup, J. R., Efficient subpixel image registration algorithms. *Opt. Lett.* **2008,** *33* (2), 156-8.
3. Dupuis, N. F.; Holmstrom, E. D.; Nesbitt, D. J., Single-molecule kinetics reveal cation-promoted DNA duplex formation through ordering of single-stranded helices. *Biophys. J.* **2013,** *105* (3), 756-66.
4. Cisse, I. I.; Kim, H.; Ha, T., A rule of seven in Watson-Crick base-pairing of mismatched sequences. *Nat. Struct. Mol. Biol.* **2012,** *19* (6), 623-7.
5. Johnson-Buck, A.; Su, X.; Giraldez, M. D.; Zhao, M.; Tewari, M.; Walter, N. G., Kinetic fingerprinting to identify and count single nucleic acids. *Nat. Biotechnol.* **2015,** *33* (7), 730-732.
6. Whitley, K. D.; Comstock, M. J.; Chemla, Y. R., Elasticity of the transition state for oligonucleotide hybridization. *Nucleic Acids Res.* **2017,** *45* (2), 547-555.
7. Schwarzenbach, H.; Hoon, D. S. B.; Pantel, K., Cell-free nucleic acids as biomarkers in cancer patients. *Nat. Rev. Cancer* **2011,** *11*, 426.
8. Bryzgunova, O. E.; Laktionov, P. P., Extracellular Nucleic Acids in Urine: Sources, Structure, Diagnostic Potential. *Acta Naturae* **2015,** *7* (3), 48-54.
9. Shekhtman, E. M.; Anne, K.; Melkonyan, H. S.; Robbins, D. J.; Warsof, S. L.; Umansky, S. R., Optimization of transrenal DNA analysis: detection of fetal DNA in maternal urine. *Clin. Chem.* **2009,** *55* (4), 723-9.
10. Xia, Y.; Huang, C. C.; Dittmar, R.; Du, M.; Wang, Y.; Liu, H.; Shenoy, N.; Wang, L.; Kohli, M., Copy number variations in urine cell free DNA as biomarkers in advanced prostate cancer. *Oncotarget* **2016,** *7* (24), 35818-35831.
11. Su, Y. H.; Wang, M.; Brenner, D. E.; Ng, A.; Melkonyan, H.; Umansky, S.; Syngal, S.; Block, T. M., Human urine contains small, 150 to 250 nucleotide-sized, soluble DNA derived from the circulation and may be useful in the detection of colorectal cancer. *J. Mol. Diagn.* **2004,** *6* (2), 101-7.
12. Wang, Y.; Springer, S.; Zhang, M.; McMahon, K. W.; Kinde, I.; Dobbyn, L.; Ptak, J.; Brem, H.; Chaichana, K.; Gallia, G. L.; Gokaslan, Z. L.; Groves, M. L.; Jallo, G. I.; Lim, M.; Olivi, A.; Quinones-Hinojosa, A.; Rigamonti, D.; Riggins, G. J.; Sciubba, D. M.; Weingart, J. D.; Wolinsky, J. P.; Ye, X.; Oba-Shinjo, S. M.; Marie, S. K.; Holdhoff, M.; Agrawal, N.; Diaz, L. A., Jr.; Papadopoulos, N.; Kinzler, K. W.; Vogelstein, B.; Bettegowda, C., Detection of tumor-derived DNA in cerebrospinal fluid of patients with primary tumors of the brain and spinal cord. *Proc. Natl. Acad. Sci. U.S.A.* **2015,** *112* (31), 9704-9.
13. Momtaz, P.; Pentsova, E.; Abdel-Wahab, O.; Diamond, E.; Hyman, D.; Merghoub, T.; You, D.; Gasmi, B.; Viale, A.; Chapman, P. B., Quantification of tumor-derived cell free DNA(cfDNA) by digital PCR (DigPCR) in cerebrospinal fluid of patients with BRAFV600 mutated malignancies. *Oncotarget* **2016,** *7* (51), 85430-85436.