Biomolecular Detection and Quantification 1 (2014) 8–22

Contents lists available at ScienceDirect

Biomolecular Detection and Quantification

journal homepage: www.elsevier.com/locate/bdq

Original Article

Determining lower limits of detection of digital PCR assays for cancer-related gene mutations

Coren A. Milbury*,[1], Qun Zhong[1], Jesse Lin, Miguel Williams, Jeff Olson, Darren R. Link, Brian Hutchison

RainDance Technologies, 749 Middlesex Turnpike, Billerica, MA 01821, USA

ARTICLE INFO

Article history:
Received 13 June 2014
Received in revised form 18 July 2014
Accepted 11 August 2014
Available online 20 August 2014

Keywords:
Digital PCR
Mutation detection
Limit of detection
Assay sensitivity
EGFR L858R
EGFR T790M

ABSTRACT

Digital PCR offers very high sensitivity compared to many other technologies for processing molecular detection assays. Herein, a process is outlined for determining the lower limit of detection (LoD) of two droplet-based digital PCR assays for point mutations of the epidermal growth factor receptor (*EGFR*) gene. Hydrolysis probe mutation-detection assays for *EGFR* p.L858R and p.T790M mutations were characterized in detail. Furthermore, sixteen additional cancer-related mutation assays were explored by the same approach. For the *EGFR* L8585R assay, the assay sensitivity is extremely good, and thus, the LoD is limited by the amount of amplifiable DNA that is analyzed. With 95% confidence limits, the LoD is one mutant in 180,000 wild-type molecules for the evaluation of 3.3 μg of genomic DNA, and detection of one mutant molecule in over 4 million wild-type molecules was achieved when 70 million copies of DNA were processed. The measured false-positive rate for the *EGFR* L8585R assay is one in 14 million, which indicates the theoretical LoD if an unlimited amount of DNA is evaluated. For the *EFGR* T790M assay, the LoD is one mutant in 13,000 for analysis of a 3.3 μg sample of genomic DNA, and the dPCR assay limit sensitivity approaches one mutant in 22,000 wild-type molecules.

1. Introduction

1.1. Digital PCR

Invention of the Polymerase Chain Reaction (PCR) [1] changed life science research and molecular diagnostics. Now, digital PCR (dPCR) is changing the field of PCR and re-defining expectations for mutation detection. The power of digital PCR arises from evaluating individual molecules as positive or negative for a particular parameter, such as mutation status. For many assays, dPCR sensitivity is significantly higher than traditional PCR analysis, and the accuracy and precision of the assay improves by counting larger numbers of molecules individually. The sensitivity of digital PCR will facilitate detection limits that redefine our understanding of disease onset, progression, and recurrence.

One particularly attractive application of digital PCR is the quantitative detection of a small number of mutated DNA molecules among a large number of wild-type molecules, which is relevant to cancer research, and especially for the detection of minor alleles. Cancerous tissue is often highly heterogeneous and cancer biomarkers vary across types of disease and stages of disease progression which complicates cancer detection and identification at early stages. Subclonal populations of cells within a tumor may contain a mutation that differs from the primary mutation, but the subclonal mutation could be correlated to a prognosis and/or a response to specific therapy regimens. Similarly, detecting mutations in circulating tumor DNA (i.e., a Fluid Biopsy™ sample) is a relatively simple and non-invasive approach to monitoring disease recurrence, which requires a high sensitivity of mutation detection to provide effective therapies at the earliest stage of progression [2–4]. These examples reflect the need for mutation detection tools that are qualitatively more sensitive than existing tools, enabling sensitivity better than 1 in 10,000 and maximized to enable transformational advances in cancer research.

Digital PCR technologies can be deeply sensitive and highly precise for detection of low abundance minority alleles, such as those observed in progressive cancer and metastatic samples. The

Abbreviations: PCR, Polymerase Chain Reaction; *EGFR*, epidermal growth factor receptor; LoB, limit of blank; LoD, limit of detection; N, total number of droplet events counted; N_{WT}, number of droplets with only wild-type DNA; N_{Mut}, number of droplets with only mutated DNA; λ, average number of targets "loaded" per droplet; p, fraction of PCR-positive droplets; R, ratio of mutant to wild-type molecules; Λ_{FP}, average number of false-positive events; R_{FP}, average false positive rate (Λ_{FP}/#WT).

* Corresponding author. Tel.: +1 978 495 3300.
E-mail address: milburyc@raindancetech.com (C.A. Milbury).

[1] These two authors contributed equally to the manuscript.

http://dx.doi.org/10.1016/j.bdq.2014.08.001
2214-7535/

Table 1
EGFR assay details.

Protein Variant	p.T790M	p.L858R
Nucleotide Variant	c.2369C>T	c.2573T>G
Amplicon Location	chr7: 55249045-55249095	chr7:55259486-55259563
Amplicon Length	51 base pairs	78 base pairs
Forward Primer	5′-CCTCACCTCCACCGTGCA-3′	5′-GCAGCATGTCAAGATCACAGATT-3′
Reverse Primer	5′-AGGCAGCCGAAGGGCA-3′	5′-CCTCCTTCTGCATGGTATTCTTTCT-3′
Wild-type Probe	/5TET/T+CATC+A+C+GC/ZEN/A+GCTC/3IABkFQ/	/VIC/-AGTTTGGCCAGCCCAA-MGBNFQ
Mutant Probe	/6FAM/T+CATC+A+T+GC/ZEN/A+GC+TC/3IABkFQ/	/6FAM/-AGTTTGGCCCGCCCAA-MGBNFQ
Probe Type (Vendor)	PrimeTime® LNA-ZEN (Integrated DNA Technologies)	TaqMan® MGB (Life Technologies)

epidermal growth factor receptor (*EGFR*) is frequently impacted by mutations that arise in cancer. Mutations in the *EGFR* gene (chromosomal locus 7p12.3-p12.1) are common across several cancer types, and the mutations often result in altered expression and activity. Upregulation and overexpression of *EGFR* can lead to uncontrolled cellular division, resulting in rampant tissue growth and lead to cancer. A common example is a point mutation (c.2573T>G) in exon 21 of *EGFR*, commonly known as variant L858R. This type of activating mutation occurs frequently in epithelial cancers, particularly lung cancer and glioblastomas. Non-small cell lung cancers (NSCLC) are some of the most common cancers in the world and are frequently associated with mutations in *EGFR*. Two tyrosine kinase inhibitors (TKIs), Gefitinib (Iressa®) and Erlotinib (Tarceva®), are common treatments for NSCLC. Nearly all responders to treatment possess somatic alterations in exons 18–21 of the *EGFR* gene. Approximately 90% of *EGFR* mutations in NSCLC are highly variable deletions located in exon 19 or point mutations in exon 21[5]. The L858R and T790M alterations are two of the most frequent *EGFR* mutations. The *EGFR* T790M mutation is particularly important because it is linked to known drug resistance, reinforcing the value of early detection [6].

Herein, the process of determining lower limits of detection (LoD) for assays performed using the RainDance RainDrop® digital PCR system for the detection of mutations in *EGFR* and other common cancer-related mutations is reported. The data demonstrate that digital PCR supports the potential for detecting mutated DNA in highly heterogeneous tumor samples or body fluids, which provides a broad view of all biomarkers arising from heterogeneous and highly localized tumor(s).

2. Methods

2.1. EGFR assays for digital PCR analysis

Assays comprised of primers and hydrolysis probes were developed to screen for *EGFR* T790M and L858R point mutations via digital PCR. The digital PCR system utilized for this evaluation was the RainDrop® Digital PCR System (RainDance Technologies). Table 1 presents the assay details for the *EGFR* T790M and L858R assays. Probes were labeled with 6-carboxyfluorescein (FAM, ex 494 nm/em 522 nm), VIC (from ABI, ex 528 nm/em 554 nm), or TET (from Integrated DNA Technologies, ex 522 nm/em 539 nm). The *EGFR* L858R TaqMan® MGB probes were designed using the Life Technologies online TaqMan design tool, whereas the PrimeTime® LNA-ZEN probes were designed with the assistance of Integrated DNA Technologies, as these are custom probes (non-cataloged item). The *EGFR* T790M probe sequence contains LNA nucleotides that are denoted with "+".

Table 2
Samples evaluated to measure and verify the sensitivity of each *EGFR* assay.

Sample (Mut/WT)	*EGFR* T790M	*EGFR* L858R
0% (wild-type only)	$N = 58$	$N = 71$
~0.5 to 1.0%	$N = 4$	$N = 4$
~0.05 to 0.1%	$N = 4$	$N = 4$
~0.005 to 0.01%	$N = 4$	$N = 4$
~0.0005 to 0.001%	$N = 4$	$N = 4$

2.2. Samples and reagent components

A series of experimental samples were processed to determine the LoD for each *EGFR* assay. Table 2 summarizes the replicate samples that were analyzed in this study. A two-part design was executed; one subset of samples is composed of wild-type genomic DNA samples to assess the false positive rate of the assay, and the second subset included a mutation titration series to verify the linearity and sensitivity of the assay.

The *EGFR* T790M and L858R mutation templates were synthetic GeneArt® plasmid templates (Life Technologies Inc.). Each plasmid was linearized using restriction enzyme digestion. Wild-type genomic DNA (G3041, Promega Inc.) was nebulized to approximately 3 kb in length according to manufacturer's protocols (K7025-05, Life Technologies). DNA was quantified and qualified using a NanoDrop 2000 spectrophotometer. Fragmentation length was confirmed via gel electrophoresis.

Following DNA preparation, PCR reagent components were prepared as presented in Table 3 in a pre-PCR room to limit the risk of reagent contamination. All sample PCR reactions were prepared to 50 μl volumes, containing approximately ~20,000 copies/μl of genomic DNA (~3.3 μg DNA per each 50 μl reaction). The mutant template was only incorporated into the second subset of samples, following a titration of the DNA sample, according to Table 2.

Outside of the samples defined within this evaluation, preliminary assessment, verification, and optimization of both assays was performed. The false positive mutant count is zero for both

Table 3
Digital PCR reagent components.

Reagent	Vendor	Item #	Final concentration
TaqMan® Genotyping Master Mix	Life Technologies	4371355	1×
TaqMan® MGB Probes	Life Technologies	Custom	0.2 μM
PrimeTime® LNA-ZEN qPCR Probes	Integrated DNA Technologies	Custom	0.2 μM
Oligonucleotide primers	Integrated DNA Technologies	Custom	0.9 μM
Droplet Stabilizer	RainDance Technologies	20-00803	1×
WT genomic DNA control	Promega	G3041	~20,000 copies/μl
Mutant plasmid DNA control (GeneArt® Gene Synthesis)	Life Technologies	Custom	Variable by sample type
DNase/RNase-free water	Sigma–Aldrich	W4502-1L	To volume

Table 4
Thermal cycling protocols.

A. *EGFR* T790M

Step	Temp	Time	Cycles
Polymerase activation	95 °C	10 min	1
Denaturation	95 °C	15 s	
Annealing	58 °C[a]	15 s	50
Extension	60 °C	45 s	
Incubation	98 °C	10 min	1
Final hold	12 °C	Indefinite	

B. *EGFR* L858R

Step	Temp	Time	Cycles
Polymerase activation	95 °C	10 min	1
Denaturation	95 °C	15 s	
Annealing & Extension	60 °C[a]	60 s	45
Incubation	98 °C	10 min	1
Final hold	12 °C	Indefinite	

[a] A slow ramping speed (0.5 °C) was used during cooling from the denaturation step to the annealing step.

assays, though some false-positive wild-type counts have been observed for the *EGFR* T790M assay, possibly due to non-specific probe hydrolysis.

2.3. Emulsion generation, thermal cycling, and sample reading

Samples were loaded onto the RainDrop® Source instrument (RainDance Technologies), following operating guidelines, to emulsify the sample. The RainDrop® Source instrument uses real-time closed-loop image control to ensure uniformity of droplet creation both within and across runs. Each 50 μl sample was emulsified into 5 picoliter droplet volumes, partitioning single molecule of DNA into approximately 10 million droplets.

Following emulsion generation on the RainDrop® Source instrument, the samples were thermal cycled following the protocol outlined in Table 4. A Bioer GenePro thermal cycler was used. Thermal cycled samples were loaded onto the RainDrop® Sense instrument (RainDance Technologies). The RainDrop® Sense instrument uses a 488 nm laser to read the FAM and VIC/TET fluorescence intensity per droplet.

2.4. Assessment of additional cancer assays

The approach to determine assay sensitivity was employed to assess sixteen cancer-related assays in addition to *EGFR*. The assay designs are listed in Appendix Table A1. Analysis of wild-type DNA, as well as a mutation-titration series, was evaluated for each of these assays. Assay designs included TaqMan® MGB probes as well as custom PrimeTime® LNA-ZEN qPCR probes (similar to those used for the two *EGFR* assays). Reagent composition and sample layout followed the methods presented above, and limits of detection were determined by the process outlined in the following sections.

2.5. Data analysis with the RainDrop Analyst data analysis software

After evaluating all samples using the RainDrop® digital PCR System, data from cluster plots were spectrally-compensated and analyzed using the RainDrop Analyst data analysis software, following standard procedures. The sample containing the highest mutant titration was used as a control sample to define gates around the cluster of droplet events displaying signal from only wild-type copies (identified by VIC/TET fluorescence) and the mutant cluster (identified by FAM signal). These gates were applied across all samples evaluated within each assay. The same mutant gate was set within all wild-type only samples that were evaluated. For all wild-type only samples, any droplets with mutant signal (droplet events that are counted within the mutant gate) are considered mutant-positive, and are therefore false-positive events.

Examples of gated dPCR cluster plots for the *EGFR* T790M and L858R assays are presented in Figs. 1 and 2, respectively. For the *EGFR* T790M data in this report, the red target gates (Fig. 1) were relatively conservative (e.g., larger), whereas the gates for the *EGFR* L858R data relatively stringent (e.g., narrowly defined regions; shown in red, Fig. 2). Fig. 3 demonstrates the difference between smaller elliptical gates (stringent) and larger rectangular (more conservative) gates. Setting larger gates provides a more inclusive count of droplet events, but also leads to a higher probability for counting false-positive events due to gates overlapping with neighboring clusters or ambiguous single events being captured within the gated region. Though the gates for *EGFR* L858R are more stringent gates, they still capture approximately 98% or more of events associated with each cluster. The difference in droplet count between the stringent and conservative gates is typically quite small. Establishing a balance between stringent and conservative gate setting is defined by the user, based on tolerance for

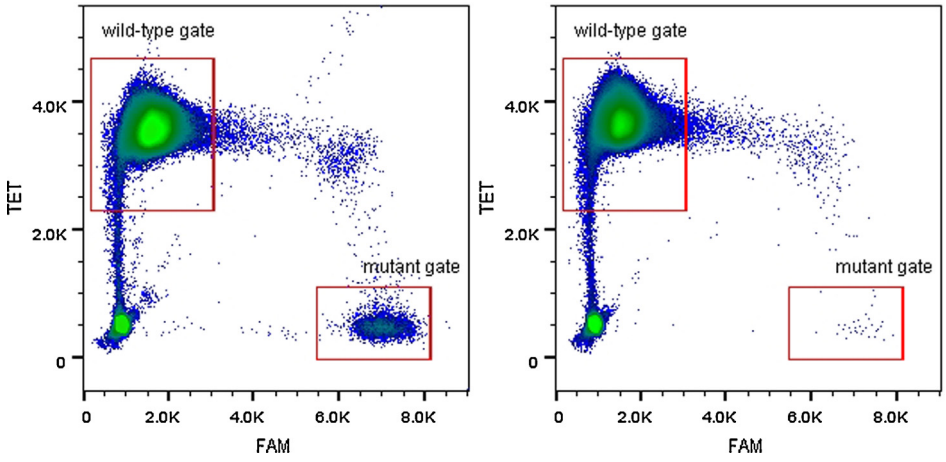

Fig. 1. *EGFR* T790M. The two dPCR plots are from a sample containing wild-type and mutant DNA (left) and a wild-type only sample (right) with the maximum recommended DNA loading. A small number of false-positive samples are routinely observed with this assay. Rectangular gates were applied for *EGFR* T790M, as this assay can exhibit some sample-to-sample variability.

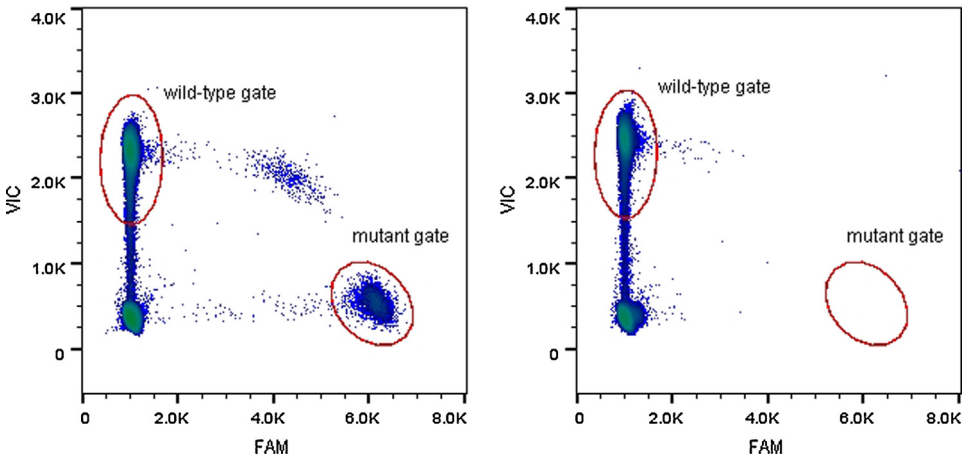

Fig. 2. *EGFR* L858R. The two dPCR plots are from a sample containing wild-type and mutant DNA (left) and a wild-type only sample (right) with the maximum recommended DNA loading. 67 of 71 wild-type only samples analyzed with this assay yielded no false-positives, and the other 4 samples each yielded 1 false-positive droplet. Stringent gates were applied for the *EGFR* L858R assay, as clusters are typically tight and exhibit minimal sample-to-sample variability.

false-positives (leading to lower sensitivity) versus risk of missing a small fraction of true-positive events.

3. Statistical calculations

3.1. Assessing sensitivity of dPCR assays

The sensitivity of a mutation detection assay is defined by a threshold for making positive calls about mutational status on individual samples with a known level of confidence. For quantitative real-time PCR (qPCR), the baseline for mutation detection is generated by evaluating negative controls (i.e., samples that do not contain target DNA). The quantification cycle (C_q) is the cut-off for qPCR that defines a concentration below which a mutation-containing sample cannot be called with confidence. Similarly, for dPCR, the baseline level of false-positive molecule counts determines what concentration of true mutant-positives can be detected with statistical confidence. The following sections describe the

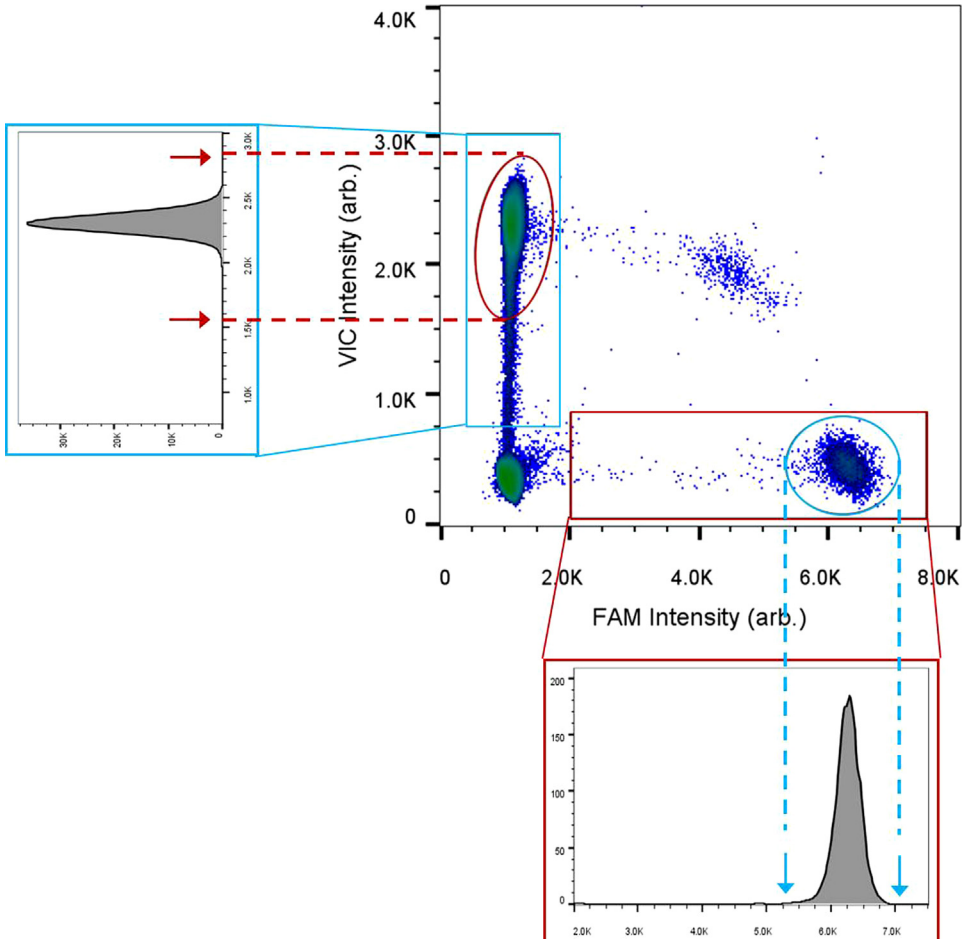

Fig. 3. Comparison of stringent (small elliptical) gates to conservative (rectangular) gates for one of the *EGFR* L858R samples. The number of droplet events counted in the elliptical gate is at least 98% of the population in rectangular gates. A smaller gate will miss a small percentage of true-positives but will minimize false-positives.

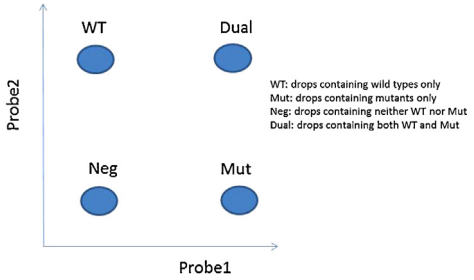

Fig. 4. Two-dimensional dPCR cluster plot landscape illustrating a heterogeneous sample analyzed with two probes (i.e., a duplex assay).

analysis and process for evaluating the limit of blank (LoB) and limit of detection (LoD) for mutation detection with digital PCR. This approach is not limited to mutation detection, and should be applicable for other minor allele detection approaches (i.e., pathogen and viral detection).

3.2. Conversion of droplet event counts to molecule counts

When a bulk solution of homogeneously-distributed target molecules, with an *average* concentration of less than one molecule per partition volume, is partitioned into droplets, some droplets will contain more than one target molecule. However, a droplet partition that initially contains more than one identical target amplified by PCR is not distinct from a droplet initially containing only one target molecule. Therefore, the actual number of target molecules must be calculated from the measured number of target-containing partitions (i.e., droplets).

The partitioning of molecular targets into droplets is described by a Poisson distribution [7–9], where $P(k;\lambda)$ is the probability for there to be k targets in one droplet when the population of all droplets contains an average "loading" of λ targets per droplet. For a system in which 50 molecules are loaded into 100 droplets, λ is ½. Alternatively, if 500 molecules are loaded into 100 droplets, then λ is 5.

$$P(k;\lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{1}$$

The fraction of PCR-negative drops (p^-) is:

$$p^- = P(0;\lambda) = e^{-\lambda} \tag{2}$$

and the fraction of PCR-positive drops (p), which is inclusive of all droplets with one or more molecules of target DNA is:

$$p = 1 - p^- = 1 - e^{-\lambda} \tag{3}$$

A duplex assay enables detection of two independent molecular targets, e.g., wild-type DNA (*WT*) and mutant DNA (*Mut*). When a sample containing both mutant and wild-type DNA is analyzed, four clusters of dPCR data are evident, as depicted in the two-dimensional histogram of the fluorescence intensities in Fig. 4.

The partitioning of fragments of homoduplex DNA into droplets is independent of mutational status,[2] and the joint probability of two independent events is equal to the product of the individual probabilities. Therefore, the fraction of wild-type-only droplets is:

$$p_{WT} = p(WT) \cdot p(noMut) = (1 - e^{-\lambda_w}) \cdot e^{-\lambda_m} \tag{4}$$

where λ_m is the average "loading" of mutant targets per droplet, and λ_w is the average "loading" of wild-type targets per droplet. The number of wild-type only droplets is:

$$N_{WT} = N \cdot (1 - e^{-\lambda_w}) \cdot e^{-\lambda_m} \tag{5}$$

Similarly, the numbers of droplet events corresponding to other combinations of mutant and wild-type DNA are:

$$N_{Mut} = N \cdot e^{-\lambda_w} \cdot (1 - e^{-\lambda_m}) \tag{6}$$

$$N_{Neg} = N \cdot e^{-\lambda_w} \cdot e^{-\lambda_m} \tag{7}$$

$$N_{Dual} = N \cdot (1 - e^{-\lambda_w}) \cdot (1 - e^{-\lambda_m}) \tag{8}$$

This series of four equations (Eqs. (5)–(8)) contains only two unknown quantities (λ_w and λ_m) and these values can be calculated by measurement of at least three quantities from the set of (N, N_{WT}, N_{Mut}, N_{Neg}, and N_{Dual}). In most cases, it is recommended that N, N_{WT}, and N_{Mut} are used to calculate λ_w and λ_m following Eqs. (9) and (10). Alternative methods for selecting and processing information from gated clusters of droplet events are discussed in the Appendix.

$$\lambda_w = -\ln\left(\frac{1 + \frac{N_{Mut}-N_{WT}}{N} + \sqrt{\left(1 + \frac{N_{Mut}-N_{WT}}{N}\right)^2 - \frac{4 \cdot N_{Mut}}{N}}}{2}\right) \tag{9}$$

$$\lambda_m = -\ln\left(\frac{1 + \frac{N_{WT}-N_{Mut}}{N} + \sqrt{\left(1 + \frac{N_{WT}-N_{Mut}}{N}\right)^2 - \frac{4 \cdot N_{WT}}{N}}}{2}\right) \tag{10}$$

Finally, the ratio of mutant to wild-type molecules in the sample is given by:

$$R = \frac{\lambda_m}{\lambda_w} \tag{11}$$

Appendix Tables A3–A6 contain the droplet event counts and calculated wild-type and mutant DNA molecule counts for the *EGFR* T790M and L858R samples evaluated.

3.3. Estimation of 95% confidence intervals

Confidence intervals (CI) define a range of values which is likely to include the true value of the sample. Confidence intervals are constructed at a confidence level (e.g., 95%), which defines the likelihood that the true value is within the range.

When a DNA sample is partitioned, the probability of a droplet being either PCR-positive or PCR-negative follows a binomial distribution [8,10]. When the number of PCR-positive droplets (Np) is fairly large (>20), the binomial distribution is approximated with a normal distribution with a mean of Np and a variance of $Np(1-p)$ [10,11] so the 95% confidence interval for Np is approximated by $\{Np \pm 1.96\sqrt{Np(1-p)}\}$. In the case of rare mutation detection, p is small enough that the confidence interval for Np simplifies to $\{Np \pm 1.96\sqrt{Np}\}$.

The normal approximation is not appropriate for very small values of Np (<20). For such a case, a Poisson distribution with Np as the expected value is a better approximation [12]. Table 5 presents Excel function (Microsoft) formulas for $Np < 20$ and $Np > 20$, such that one can calculate the confidence limits in an Excel spreadsheet. Alternatively, web-based calculators that use binomial distributions are available. Online tools include uCount[SM] from The Wittwer Lab at the University of Utah (https://dna.utah.edu/ucount/uc.php) [13] or the Exact Binomial and Poisson Confidence Intervals calculator by John C. Pezzullo (http://statpages.org/confint.html) [14].

With the RainDrop® dPCR System, 25–50 μl samples are partitioned into 5-picoliter droplets, which yields 5–10 million individual droplets. The recommended maximum loading of total

[2] This section of the report is applicable when the sample DNA has been processed to retain the native homoduplex character. For cases in which the DNA has been fragmented by heating or subject to temperature above the melting temperature, the double-stranded DNA will likely denature and re-anneal into heteroduplex fragments, and the analysis approach would not be valid.

Table 5
Confidence intervals for various mutant-positive droplet event counts (*Np*). *Np* is the count of mutant-positive droplets, which includes droplets that contain only mutant molecules and droplets that contain both mutant and wild-type molecules.

Confidence limits	*Np* = 0 and up	*Np* > 20
99% CI_lower limit	CHIINV(0.995, 2*Np*)/2	*Np* − NORMSINV(0.995)*sqrt(*Np*)
99% CI_upper limit	CHIINV(0.005, 2*Np* + 2)/2	*Np* + NORMSINV(0.995)*sqrt(*Np*)
95% CI_lower limit	CHIINV(0.975, 2*Np*)/2	*Np* − NORMSINV(0.975)*sqrt(*Np*)
95% CI_upper limit	CHIINV(0.025, 2*Np* + 2)/2	*Np* + NORMSINV(0.975)*sqrt(*Np*)
90% CI_lower limit	CHIINV(0.95, 2*Np*)/2	*Np* − NORMSINV(0.95)*sqrt(*Np*)
90% CI_upper limit	CHIINV(0.05, 2*Np* + 2)/2	*Np* + NORMSINV(0.95)*sqrt(*Np*)

human genomic DNA is 10% (i.e., one genome equivalent per 10 droplets), and for most digital PCR applications, the amount of mutant DNA within the sample is normally present at a small fraction of the wild-type DNA. At this low fraction, very few mutant-positive droplets are expected to contain more than one copy of the mutated DNA fragment, so the number of *mutant-positive droplets* is no different from the number of *mutant molecules* ($\lambda_m = p_m$). For example, in the mutant titration series, wild-type DNA was loaded into ~10% of the drops, and the highest mutant loading was present in only ~0.06% of the drops, so $\lambda_m = 0.0006$, and $p_m = 1 - e^{-\lambda m} = 0.9997\lambda_m$, so the difference between the two is within 0.03%. Therefore, the quantity, *Np*, is used interchangeably to refer to mutant positive droplets and mutant DNA molecules in this paper.

For mutation detection applications, the number of wild-type targets is usually much greater than the number of mutant targets, and the uncertainty in the ratio of the two will be dominated by the mutant counting. The 95% CI of the ratio (R) of mutant to wild type are given as below:

$$R_{low} = \frac{p_{m\_low}}{\lambda_w} \tag{12}$$

$$R_{high} = \frac{p_{m\_high}}{\lambda_w} \tag{13}$$

3.4. Assessing the frequency of false-positives

To determine application and assay sensitivity using any molecular detection technology, one must account for signal-to-noise limitations. Similarly, for mutation detection assays using digital PCR, false-positives might arise from fluorescence detection ("system") noise and molecular biology ("assay") noise. Comparing samples to negative controls is typically necessary to classify mutational status as positive or negative. The Clinical and Laboratory Standards Institute (CLSI) guidelines recommend measurement of at least 60 negative samples to assess the probability distribution function of false-positive events detected in negative control samples.[15]

In this work, collections of wild-type only samples were analyzed with the *EGFR* T790M and L858R assays. The data from *EGFR* T790M in Fig. 5 reveals an average false-positive count (Λ_{FP}) of 39 mutant molecules per sample, which was measured from 58 negative control samples that each contained, on average, 862,147 wild-type molecules. For purposes of assessing the false-positive distribution displayed in Fig. 5, the mutant count from each individual wild-type sample was normalized to the average number of wild-type counts using the average false positive ratio (R_{FP}, the average false-positive ratio of all negative control samples) of 4.5E−5 to compensate for differences in the absolute number of molecules analyzed in each sample.

In contrast to the *EGFR* T790M assay, the data from L858R yields only 4 samples out of 71 with one false-positive event detected ($\Lambda_{FP} = 0.06$) and a false positive ratio of 7.1E−8. Fig. 5 reveals that the number of false-positive mutant counts follows a Poisson distribution in both cases, confirming that the Poisson model is a valid approximation of the measured data. (Note that applying the *Poisson model fit* to the false-positive data is not related to the *Poisson loading correction* described in previous sections.)

A dPCR practitioner often evaluates and reports the confidence level when assigning mutational status to a sample containing a number of positive events that is very close to the false-positive baseline. For this purpose, a hypothesis (significance) test is executed and a *p*-value is calculated. The *p*-value is the probability that the observed mutant count or higher falls within the distribution of false-positive counts measured from negative control samples. In this case, the null hypothesis is: "there is no difference between the sample and negative controls." The null hypothesis is rejected when the *p*-value is less than a specified significance level. The choice of significance level at which to reject the null hypothesis depends on the application and user preference. For clinical molecular analysis, common choices for *p*-value thresholds are 0.05 (95%) and 0.01 (99%); the analysis described in this report is based on a 95% confidence threshold.

Two *EGFR* T790M mutant-positive samples are highlighted below to demonstrate the calculation of *p*-values. $P(k;\Lambda_{FP})$ represents the Poisson distribution that estimates the frequency of

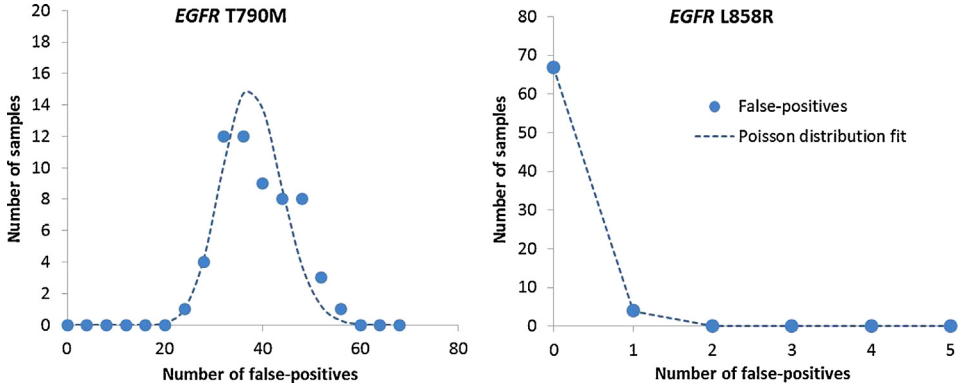

Fig. 5. False-positives measured in *EGFR* T790M (left, $\Lambda_{FP} = 39$) and L858R (right, $\Lambda_{FP} = 0.06$) negative controls.

false-positives. Because these positive samples have very similar wild-type load to the negative control samples, it is assumed that the false positive ratio is consistent with negative samples. From the ratio, Λ_{FP} is calculated for each sample, followed by calculating the p-value.

Sample 1 (#1 in Appendix Table A5): measured mutants = 45, $\Lambda_{FP} = 39$,

$$p\text{-value} = \sum_{k=45}^{\infty} P(k; \Lambda_{FP}) = 1 - \sum_{k=0}^{44} P(k; 39) = 0.17$$

Sample 2 (#5 in Appendix Table A5): measured mutant events = 81, $\Lambda_{FP} = 40$,

$$p\text{-value} = \sum_{k=81}^{\infty} P(k; \Lambda_{FP}) = 1 - \sum_{k=0}^{80} P(k; 40) = 8.7\text{E} - 9$$

In the first sample, a measurement of 45 mutant molecules results in a p-value of 0.17, which is greater than 0.05 (i.e., 5% significance level for p-value), and therefore, the sample is not significantly different than the wild-type only controls. In contrast, the second sample has a p-value of 8.7E−9, which indicates that this sample is significantly resolved from the negative controls. Appendix Tables A5 and A6 list the p-values calculated for the rest of the mutant-positive titration samples for the *EGFR* T790M and L858R assays. An Excel equation could be used to calculate the p-value as shown below:

$$p\text{-value} = 1 - POISSON.DIS(N_{Mut} - 1, \Lambda_{FP}, TRUE) \quad (14)$$

3.5. Estimating assay sensitivity and the limit of detection (LoD)

The limit of detection (LoD) and limit of blank (LoB) are metrics that describe the sensitivity of an analytical procedure [15–18]. Fig. 6 depicts how the LoB and LoD are defined at stated error rates. The α error rate indicates the fraction of all measurements of true negative samples that will be greater than the LoB (i.e., false-positive mutation call). The β error rate indicates the probability of a false-negative mutation call (i.e., measurement less than LoB) from a sample with a true value equal to the LoD. For the analyses in this paper, α error (false positive) and β error (false negative) rates are 5%, which is a common approach. In practice, the selection of α and β values is largely dictated by the application. For example, there may be applications in which both false positive and false negative mutation calls are unacceptable, and one would decide to lower the α and β rates (resulting in a more conservative LoD). Some applications may dictate a relatively high mutational abundance threshold (i.e., well-separated from analytical performance) to avoid unnecessary and expensive validation tests. Conversely, the occurrence of false negatives is not acceptable in some applications. In such cases, the β-error value would be conservative to reflect the needs of the application.

In this report, false-positive droplet event probability (i.e., data shown in Fig. 5 and represented as the blue curve in Fig. 6 panel A) was determined from a series of samples that each contained approximately the same level of wild-type DNA. The false-positive rate is likely dependent (but not necessarily linearly proportional) with respect to wild-type load. Unless all samples are expected to contain the same approximate level of DNA, evaluating false positives with varying wild-type loads is recommended for assays of interest. Minimally, one should consider the desired LoD, and how much DNA should be evaluated to achieve that sensitivity. For example, detection of 1 mutant in 1000 wild-type molecules will require a substantially smaller amount of total DNA than what would be required to attempt detection of 1 mutant in 1 million wild-type molecules. The total amount of sample DNA does not guarantee a level of sensitivity, but sufficient input DNA is necessary to support the desired LoD.

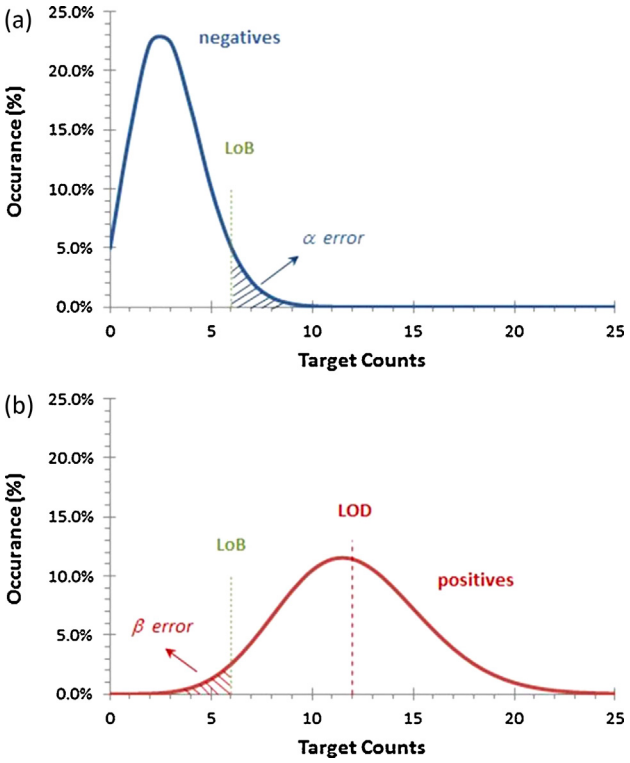

Fig. 6. Relationship between the limit of blank (LoB) and the limit of detection (LoD). (a) The blue curve is the distribution of measurements from negative samples; α-error refers to false-positive mutation calls. (b) The red curve is the expected distribution of measurements from positive samples at the LoD concentration; β error refers to false-negative mutation calls. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 6
Calculation of LoB and LoD from the average false-positive counts.

Λ_{FP}	LoB	LoD
0	0	3
0–0.05	1	5
>0.05	$\Lambda_{FP} + 1.645\sqrt{\Lambda_{FP}} + 0.8$	$(1.645 + \sqrt{1.645^2 + 4LoB})^2/4$

Note: non-integers should be rounded up to the nearest integer.

In addition to characterizing the false-positive distribution by processing wild-type only samples, CLSI guidelines [15] recommend characterization of the distribution of measured droplet events for positive samples with concentrations at the limit of detection. However, in this report, it has been demonstrated that the distribution of false-positive counts follows a distribution that arises from counting uncertainty (i.e., a Poisson distribution) and it has been assumed that the dominant variability in true positive counting also arises from counting uncertainty. Therefore, "true positive" distribution (i.e., the red curve in Fig. 6 panel B) is considered a Poisson function as well, for the purpose of determining the LoD. Alternatively, one could estimate the LoB and LoD following a non-parametric approach with a sufficient number of negative and low-level positive control samples to adequately represent the probability distributions with empirical data [17].

Table 6 reports the LoB and LoD at various average false-positive levels, Λ_{FP}, assuming a Poisson probability model fit to empirical data. In this analysis, α-error and β-error are set at 5%. For Λ_{FP} greater than 0.05, a normal approximation with a correction is used

Table 7
Droplet event and molecule counts from 58 *EGFR* T790M wild-type samples.

	Average	Standard deviation	Total
Total droplets processed (*N*)	9,049,729	683,030	524,884,260
Mutant only droplets	35.5	7.1	2059
Mutant molecules (false-positives)	39	7.8	2265
Wild-type only droplets	822,325	68,465	47,694,859
Wild-type molecules	862,147	72,261	50,004,523

to describe the Poisson distribution to simplify the calculation of LoB and LoD with an analytical expression.

4. Results and discussion

4.1. *EGFR T790M*

4.1.1. *EGFR T790M limit of detection*

Appendix Table A3 contains the raw data from 58 wild-type only samples that were processed to calculate the LoB and LoD for the *EGFR* T790M assay; Table 7 contains the summarized data from negative control samples. Based on the wild-type only controls, and the equations described in Table 6, the LoB for *EGFR* T790M mutations in a single sample with an average measured count of 862,147 copies of genomic DNA is 1 mutant in 17,000 wild-type molecules, and the LoD is 1 in 13,000. This represents the highest sensitivity that can be achieved with a single sample processed with one assay reaction. One might split a single sample across multiple dPCR reactions if an extremely sensitive LoD is desired and the DNA is plentiful. For cases with enough sample DNA to split across 8 samples at maximum loading (i.e., 26 μg of human genomic DNA processed in an 8-channel microfluidic chip), then the analysis described above results in a demonstrated LoD of 1 in 18,000. Furthermore, if data from all 58 wild-type samples are pooled together (representing a single replicate of 200 μg of DNA negative control DNA), then the LoD is 1 in 20,000. Finally, we extrapolate to determine that the LoD would continue to improve as a function of the wild-type load, and reach a plateau at the measured false-positive rate of 0.0045% (or 1 in 22,000). Fig. 7 displays the LoD assessed from collections of empirical data (blue points) as well as the extrapolated sensitivity that approaches a limit equal to the measured false-positive rate. We do not envision practical applications for human genomic DNA analysis that would yield sufficient DNA and justify splitting a sample over many chips or channels. However, the demonstration and extrapolation highlight the transformational sensitivity that the assay and technology can deliver as well as emphasizing that the practical sensitivity is likely to be limited by the amount of available sample DNA. Furthermore, applications probing non-human genomes (e.g., viral load or pathogen detection and microbial genomics) are more likely to take advantage of ultra-high sensitivity workflows due to the possibility for large quantities of input DNA with relatively small genomes.

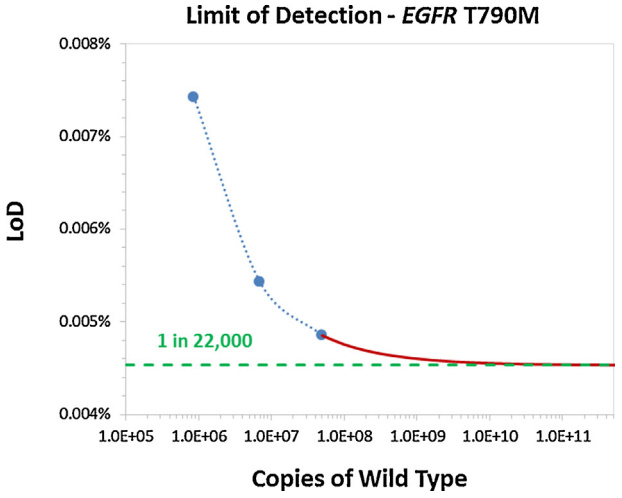

Fig. 7. LoD performance model for *EGFR* T790M. Blue points represent the empirical data collected from 58 negative samples pooled to different levels. The red curve is an extrapolation of the data to represent evaluation of more DNA. The green dashed line is the LoD if the amount of DNA available for analysis is unlimited. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The data from the wild-type samples that is reported in Fig. 5 presents a clearly-defined probability function related to the false-positives. No-template control samples do not yield false-positives, and multiple lots of wild-type control DNA (i.e., negative controls) yield the same level of false-positives. This evidence suggests that the false-positive droplet events are not due to contamination of the reagents or wild-type samples, and the false-positive background is expected to be present in all samples with this assay. For the purposes of demonstrating this analysis approach, the average background level of false-positive mutant molecules can be considered as a systematic offset, and the offset can be subtracted from the number of mutant molecules detected in any unknown samples. The result is a more *accurate* measurement of the true level of mutant molecules present in the sample, albeit with no improvement to the *precision* of the measurement. The limit of detection, which is defined as the true mutant concentration within a sample that will be detected within stated α and β error limits, can be recalculated after removal of the systematic false-positive offset; see Appendix Table A7 for the associated calculations. The LoD for the *EGFR* T790M assay, after correction for the false-positive background, is 1 in 34,000 when one 50ul sample at ~10% wild-type loading is processed.

The key to this analysis approach is establishing confidence that the false-positive background is universal for all samples. For the *EGFR* T790M assay, this approach is justified by processing commercially-available blended human cell-line DNA and no-template controls. Additional testing, outside the scope of this document, has evaluated several individual DNA samples (Coriell Specimen Collections), as well as synthetic plasmid templates, wherein comparable numbers of false-positive events have been observed. The approach would be solidified by sound explanation of the reason (other than contamination) for the level of false-positives observed with this assay. A subsequent section of this report presents some possible reasons for false-positive events and recommends further exploration of the root cause of systematic false-positive events. Further development of this approach, understanding reasons for false-positives, and establishing metrics for justification to subtract a background of false-positives will emerge from the community of dPCR practitioners.

4.1.2. *EGFR T790M – verification of linearity and sensitivity*

A mutation titration series was evaluated in quadruplicate, and the data are displayed in Fig. 8. The expected mutant to wild-type ratio (solid red line) and 95% confidence interval (dotted red lines) are plotted along with the measured data. The data is congruent with the expected concentration at the high values; it deviates from the expected value at low mutant concentrations due to the contribution of false-positive counts. Notice that the last titration sample ($R = 0.00065\%$) is a mutant concentration with an expected value that is below the LoB. Indeed, none of the four replicates at this concentration are significantly different from the negative controls.

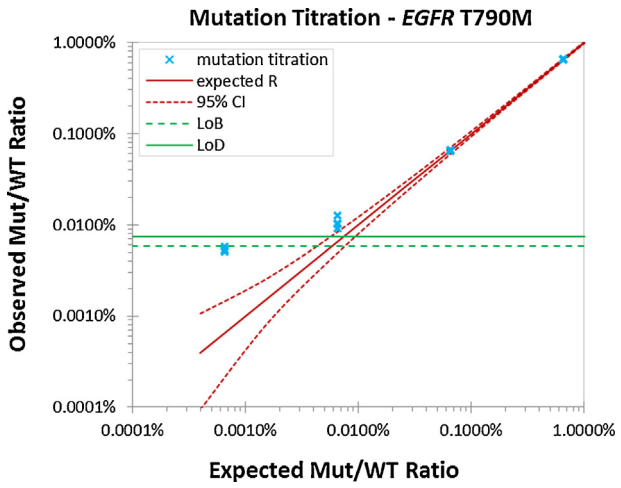

Fig. 8. *EGFR* T790M mutant titration series with four replicates of each concentration.

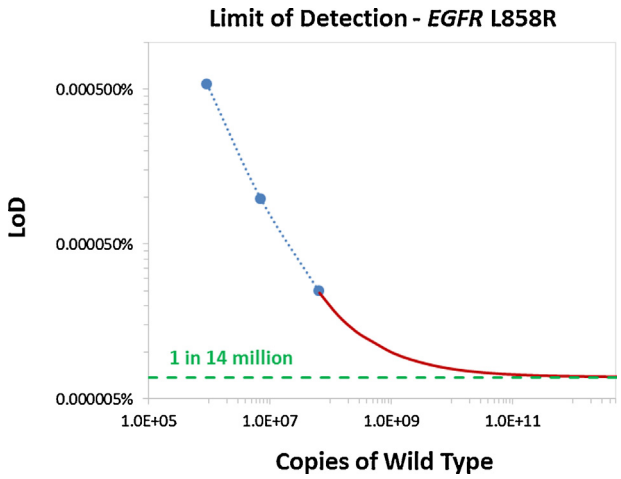

Fig. 9. LoD performance model for *EGFR* L858R. Blue points represent the empirical data collected from 71 negative samples pooled to different levels. The red curve is an extrapolation of the data to represent evaluation of more DNA. The green dashed line is the LoD if the amount of DNA available for analysis is unlimited. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 8
Droplet event and molecule counts from 71 *EGFR* L858R wild-type samples.

	Average	Standard deviation	Total
Total droplets processed (N)	8,405,878	622,029	596,817,340
Mutant only droplets	0.06	0.23	4
Mutant molecules (false positives)	0.06	0.26	4.4
Wild-type only droplets	861,260	82,956	61,149,444
Wild-type molecules	908,851	89,513	64,528,410

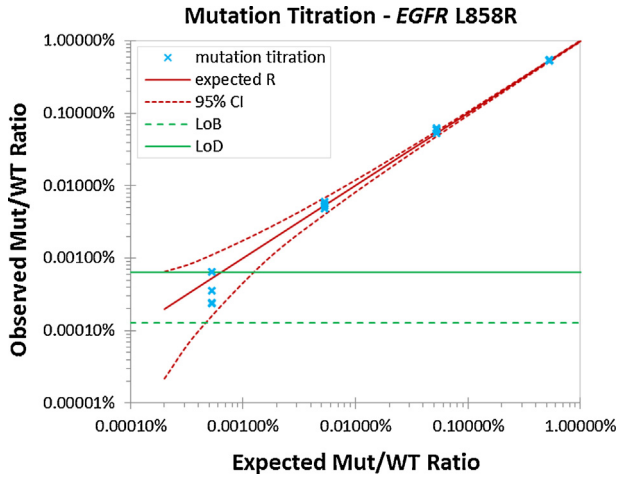

Fig. 10. *EGFR* L858R mutant titration series with four replicates of each concentration.

4.2. EGFR L858R assay characterization

4.2.1. EGFR L858R limit of detection

Following the same analysis described in the previous sections, data from 71 wild-type samples and a mutant titration series was processed to assess the sensitivity of the *EGFR* L858R assay. Table 8 summarizes the data collected from wild-type samples (data for each individual sample are reported in Appendix Table A4).

Even though there are very few false-positives detected by this assay, the LoD is modulated by the conservative β-error of 5%. For example, as displayed in Table 6, if an assay yields zero false-positives ($\Lambda_{FP} = 0$), at least three true mutants must be present to limit the probability of making a false negative call to be below 5%. Similarly, a single false-positive ($\Lambda_{FP} = 1$) among 1 million wild-type alleles dictates that a minimum of 9 true mutants must be present to maintain α and β errors of 5% or better, and the resulting LoD of this hypothetical example is 1 in 111,000.

Fig. 9 presents the assay LoD relative to the amount of DNA that is evaluated. The blue data points display the demonstrated LoD that is achieved for a single sample (1 in 180,000), a sample with enough DNA to be processed with eight assay reactions (1 in 1 million), or by dividing a very large amount of DNA (230 μg) into 71 individual samples for analysis, then recombination of data (1 in 4 million) (see Appendix Table A8 for associated calculations). Similar to the T790M assay, more DNA yields higher detection sensitivity. In contrast to the T790M assay, the false-positive rate is extremely low, so the LoD is almost exclusively limited by the amount of DNA processed and the LoD does not plateau until 1 in 14 million if one were to process hundreds of billions of wild-type molecules.

4.2.2. EGFR L858R – verification of linearity and sensitivity

A mutant titration series was evaluated in quadruplicate and the data are displayed in Fig. 10. Unlike the *EGFR* T790M assay, the data is linear throughout the whole titration range. Notice that the last titration sample ($R = 0.00053\%$) has a mutant concentration above the LoB but below the LoD. For this experiment, all four replicates are significantly different than the negative controls, but if this sample were to be processed numerous times, it is anticipated from the negative control data that more than 5% of the measurements would fall below the LoB, leading to false-negative calls. This ambiguity, which arises from measuring a single or small number of replicates, demonstrates the importance of calculating and understanding the LoB and LoD parameters, even for extremely sensitive assays.

4.3. Comparing and contrasting EGFR assay performance

The results reveal a significantly greater sensitivity associated with the *EGFR* L858R assay in contrast to the *EGFR* T790M assay. Many factors possibly contribute to the performance of an assay. For example, the complexity of the sequence composition within the targeted region, the primer and probe oligonucleotide design and quality of synthesis, reagent quality, polymerase efficiency

and fidelity, and sample preparation might influence the measured assay sensitivity.

Many commercially-available quantitative PCR (qPCR) assays fail to enable highly sensitive mutation detection of *EGFR* T790M due to a number of compounding factors that reduce the specificity of primers and probes leading to a relatively high level of signal background (i.e., a high C_q). It is difficult to design sensitive assays for the *EGFR* T790M (c.2369C>T) point mutation. This region of *EGFR* (chr7:55259486-55259563) is relatively rich in guanine and cytosine and exhibits sequence homology (~83%) with the proto-oncogene receptor tyrosine-protein kinase erbB-2 (HER2/neu). In addition to the high sequence similarities with *HER2*, this region of *EGFR* T790M contains a repetitive sequence element adjacent to the mutation, which creates a challenge to design robust probes with high specificity. Further adding to this challenge, there is a single nucleotide polymorphism located eight nucleotides upstream from the mutation, which limits primer design options. Some or all of the challenges associated with qPCR assay design might be reflected in the lower sensitivity of the *EGFR* T790M dPCR assay.

Another hypothesis for the source of false-positives is that mutant DNA fragments are generated during the PCR reaction from wild-type DNA fragments as a result of polymerase error. In principle, one of the advantages of single-molecule partitioning with the RainDrop™ dPCR System is that the sensitivity of mutation detection is better than the polymerase error rate [9]. Polymerase errors typically occur as purine/purine (A ↔ G) or pyrimidine–pyrimidine (C ↔ T) transitions. Pyrimidine–purine transversions (A/G ↔ C/T) are less common [19]. If a polymerase error occurs at or near the nucleotide of interest, and results in an amplifiable fragment that hybridizes with the mutant probe, then, due to the presence of the initial two complementary strands of wild type target, the resulting droplet will produce fluorescence signal from both wild-type and mutant probes. The level of polymerase error resulting in droplet events counted in the "dual" cluster can be determined by comparing the actual count of events in the dual cluster to the expected count based on the Poisson loading calculations. Evidence of polymerase error with specific sequence targets has been observed, which supports this hypothesis. However, for the *EGFR* T790M assay, the false-positive background arises from droplet events that are characterized as mutant-only ("Mut" cluster). To satisfy the hypothesis of polymerase error, another process or reaction must eliminate the amplification and/or generation of fluorescence signal from the original two wild-type template strands, and this factor is not yet explained.

4.4. Additional cancer-based assays

Following the same methodologies presented herein, each of the additional sixteen assays was evaluated. Data from these assays are presented in Appendix Table A2. As with *EGFR* T790M, some assays present an apparent level of low-level positives in the presumably wild-type controls. As with *EGFR* T790M, it is difficult to assign the source of these positives that reside within the mutant control gates. There are a couple assays that exhibit very few false-positives, as observed with *EGFR* L858R; namely *PIK3CA* H1047R and *KRAS* G12C. For these assays, data indicate that the sensitivity of the assays is largely dependent upon the amount of DNA that is evaluated. For both *PIK3CA* H1047R and *KRAS* G12C, fewer molecules of wild-type DNA were evaluated than those evaluated in the *EGFR* L858R assay. Similar to *EGFR* T790M, however, several assays demonstrated assay LoDs of around 1 mutant in 20,000 wild-type molecules, as they too exhibited false positives in the mutant clusters gates. This screening of an additional sixteen assays emphasizes the importance of evaluating assays with controls prior to initiating a full analysis of critical samples.

As discussed in the section above, polymerase errors are more likely to present as transitions than transversions. Within these sixteen assays (as well as the *EGFR* assays), less sensitive LoDs were observed for those assays whose targets are transitions. Out of 17 assays tested herein, all eight assays with G/C >A/T mutations have much lower LoDs than the other mutations assed. Though further evaluation is warranted, it is likely that the false positive rate and assay LoD are influenced by polymerase errors.

5. Conclusion

The performance of the *EGFR* assays in the RainDrop® Digital PCR System demonstrates unprecedented sensitivity. The data presented demonstrate that the sensitivity of digital PCR assays is often limited by the amount of DNA that is evaluated, and it is possible to achieve sensitivity for detection of mutant alleles that is orders of magnitude better than most alternative technologies.

A LoD for the *EGFR* T790M assay has been determined at 1 in 13,000 for a single sample at maximum loading, and up to 1 in 22,000 with unlimited DNA (or 1 in 34,000 for a single sample if false-positives are qualified to be a systematic offset and subtracted). For this assay, the primer and probe design challenges or polymerase errors are factors that impact the ultimate sensitivity but single molecule dPCR overcomes many sensitivity limitations that arise from environments containing many target molecules.

Finally, for the *EGFR* L858R assay, we have empirically demonstrated sensitivity of 1 in 180,000 for a single sample. Moreover, by extrapolation, we have determined that the *EGFR* L858R assay can reach a LoD that far surpasses empirical data (i.e., 1 in 14 million), and, in practice, the sensitivity is limited only by the amount of sample DNA that is processed.

This report provides guidance on the analysis of digital PCR data and it reveals some topics worthy of deeper exploration. Firm understanding of the molecular biology and technology contributions to false-positive droplet events is crucial for improved sensitivity. Also, validation of objective automated gating of droplet event clusters is likely necessary for dPCR practitioners to take advantage of the full potential sensitivity of the technology for routine applications. Finally, availability of DNA reference materials will enable validation within and across platforms.

Acknowledgements

Kerry Emslie and Leo Pinheiro of the Australian National Measurement Institute (NMIA) and Ross Haynes of the National Institute of Standards and Technology (NIST) reviewed this document. We are grateful for their constructive feedback regarding the experimental results, analysis of data, and associated discussion. Participation in the review of this report does not imply recommendation or endorsement of RainDance Technologies by NMIA or NIST.

Appendix A.

A.1. Alternative Poisson correction approach for duplex assays

As discussed in the main text, four distinct clusters of droplet events could be gated and counted in a typical duplex assay (i.e., N_{WT}, N_{Mut}, N_{Dual}, and N_{Neg}), and their relationships to λ_m and λ_w are described in Eqs. (5)–(8). Mathematically, there are more than one combination of (5)–(8) that could lead to the solution for λ_m and λ_w. Although the expressions are more complex, Eqs. (9)–(10) are recommended for most duplex assay applications because the accuracy of the cluster counts are less susceptible to assay or system noise, which is often located near and between the PCR-negative cluster and the dual-occupancy cluster.

A simpler solution comes from employing the PCR-negative cluster (N_{Neg}) in Eq. (7), by dividing Eq. (5) or (6) by (7), and λ_m and λ_w follow:

$$\lambda_w = \ln\left(1 + \frac{N_{WT}}{N_{Neg}}\right)$$

$$\lambda_m = \ln\left(1 + \frac{N_{Mut}}{N_{Neg}}\right)$$

Another alternative solution from using the combination of Eqs. (5)–(7) is by adding (5) and (7) or (6) and (7), to yield λ_w and λ_m as follows:

$$\lambda_w = \ln\left(\frac{N}{N_{Neg} + N_{Mut}}\right)$$

$$\lambda_m = \ln\left(\frac{N}{N_{Neg} + N_{WT}}\right)$$

Finally, if no cluster counts are known to be susceptible to poor resolution and/or count accuracy, then one might wish to take all counts into consideration. This approach would yield an over-determined system of equations that would be solved by a regression method.

The best approach is to apply the various combinations for analysis of *known* samples to determine which approach yields the most accurate final result. This statement and the discussion throughout this report highlight the importance of development, distribution, and maintenance of DNA reference materials to enable digital PCR developers and practitioners to recommend appropriate measurement and analysis methods, define performance specifications accurately, and validate data – especially clinical data – from multiple instruments and technology platforms.

See Tables A1–A8.

Table A1
Sixteen cancer assay sequences and probe chemistries evaluated in addition to *EGFR* L858R and T790M.

Target	Oligonucleotide	Sequence	Probe type	Vendor
Epidermal growth factor receptor (*EGFR*) p.E746_A750del, c.2235del15	WT PROBE	VIC-ATTAAGAGAAGCAACATCTCCGA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	MUTANT PROBE	6FAM-TCGCTATCAAAACATCTCCGA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	FORWARD PRIMER	5′-CTGGATCCCAGAAGGTGAGA-3′		
	REVERSE PRIMER	5′-CCACACAGCAAAGCAGAAAC-3′		
Phosphoinositide-3-kinase, catalytic, alpha polypeptide (*PIK3CA*) p.H1047R, c.3140A>G	WT PROBE	/5TET/C+CATG+A+T+GT/ZEN/G+CAT/3IABkFQ/	LNA-ZEN	Integrated DNA Technologies
	MUTANT PROBE	/56-FAM/C+CATG+A+C+GT/ZEN/GCAT/3IABkFQ/	LNA-ZEN	Integrated DNA Technologies
	FORWARD PRIMER	5′-GAGCAAGAGGCTTTGGAGTA-3′		
	REVERSE PRIMER	5′-ATGCTGTTTAATTGTGTGGAAGA-3′		
Phosphoinositide-3-kinase, catalytic, alpha polypeptide (*PIK3CA*) p.E545K, c.1633G>A	WT PROBE	/5TET/CCT +GC+T +C+AG/ZEN/TG+AT/3IABkFQ/	LNA-ZEN	Integrated DNA Technologies
	MUTANT PROBE	/56-FAM/TC+C TGC +T+T+A/ZEN/+GT+G/3IABkFQ/	LNA-ZEN	Integrated DNA Technologies
	FORWARD PRIMER	5′-GACAAAGAACAGCTCAAAGCA-3′		
	REVERSE PRIMER	5′-GCACTTACCTGTGACTCCAT-3′		
Murine sarcoma viral oncogene homolog B1 (*BRAF*) p.V600E, c.1799T>A	WT PROBE	/5TET/TCGAGAT+TTC+ACT+GTAGCT/3IABkFQ/	LNA-ZEN	Integrated DNA Technologies
	MUTANT PROBE	/56-FAM/TCGAGAT+TTC+TCT+GTAGCT/3IABkFQ/	LNA-ZEN	Integrated DNA Technologies
	FORWARD PRIMER	5′-ACCTCAGATATATTTCTTCATG-3′		
	REVERSE PRIMER	5′-CCAGACAACTGTTCAAAC-3′		
Anaplastic lymphoma receptor tyrosine kinase (*ALK*) p.R1275Q, c.3824G>A	WT PROBE	VIC-TGGCCCGAGACATC-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	MUTANT PROBE	6FAM-TGGCCCAAGACATC-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	FORWARD PRIMER	5′-GGAAGAGTGGCCAAGATTGGA-3′		
	REVERSE PRIMER	5′-TGAGGCAGTCTTTACTCACCTGTA-3′		
Kirsten rat sarcoma viral oncogene homolog (*KRAS*) p.G12D, c.35G>A	WT PROBE	VIC-TTGGAGCTGGTGGCGTA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	MUTANT PROBE	6FAM-TTGGAGCTGATGGCGTA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	FORWARD PRIMER	5′-CTGAAAATGACTGAATATAAACTTGTGG-3′		
	REVERSE PRIMER	5′-TAGCTGTATCGTCAAGGCACTC-3′		
Kirsten rat sarcoma viral oncogene homolog (*KRAS*) p.G12C, c.34G>T	WT PROBE	VIC-TTGGAGCTGGTGGCGTA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	MUTANT PROBE	6FAM-TTGGAGCTTGTGGCGTA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	FORWARD PRIMER	5′-CTGAAAATGACTGAATATAAACTTGTGG-3′		
	REVERSE PRIMER	5′-TAGCTGTATCGTCAAGGCACTC-3′		
Kirsten rat sarcoma viral oncogene homolog (*KRAS*) p.G12V, c.35G>T	WT PROBE	VIC-TTGGAGCTGGTGGCGTA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	MUTANT PROBE	6FAM-TTGGAGCTGTTGGCGTA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	FORWARD PRIMER	5′-CTGAAAATGACTGAATATAAACTTGTGG-3′		
	REVERSE PRIMER	5′-TAGCTGTATCGTCAAGGCACTC-3′		
Kirsten rat sarcoma viral oncogene homolog (*KRAS*) p.G12S, c.34G>A	WT PROBE	VIC-TTGGAGCTGGTGGCGTA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	MUTANT PROBE	6FAM-TTGGAGCTAGTGGCGTA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	FORWARD PRIMER	5′-CTGAAAATGACTGAATATAAACTTGTGG-3′		
	REVERSE PRIMER	5′-TAGCTGTATCGTCAAGGCACTC-3′		
Kirsten rat sarcoma viral oncogene homolog (*KRAS*) p.G12A, c.35G>C	WT PROBE	VIC-TTGGAGCTGGTGGCGTA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	MUTANT PROBE	6FAM-TTGGAGCTGCTGGCGTA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	FORWARD PRIMER	5′-CTGAAAATGACTGAATATAAACTTGTGG-3′		
	REVERSE PRIMER	5′-TAGCTGTATCGTCAAGGCACTC-3′		
Kirsten rat sarcoma viral oncogene homolog (KRAS) p.G12R, c.34G>C	WT PROBE	VIC-TTGGAGCTGGTGGCGTA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	MUTANT PROBE	6FAM-TTGGAGCTCGTGGCGTA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	FORWARD PRIMER	5′-CTGAAAATGACTGAATATAAACTTGTGG-3′		
	REVERSE PRIMER	5′-TAGCTGTATCGTCAAGGCACTC-3′		

Table A1 (*Continued*)

Target	Oligonucleotide	Sequence	Probe type	Vendor
Kirsten rat sarcoma viral oncogene homolog (*KRAS*) p.G13D, c.38G>A	WT PROBE	VIC-CTGGTGGCGTAGGC-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	MUTANT PROBE	6FAM-CTGGTGACGTAGGC-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	FORWARD PRIMER	5′-CTGAAAATGACTGAATATAAACTTGTGG-3′		
	REVERSE PRIMER	5′-TAGCTGTATCGTCAAGGCACTC-3′		
Human Epidermal Growth Factor Receptor 2 (*HER2*) p.L755S, c.2264T>C	WT PROBE	6FAM-TCCCTCAACACTTTG-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	MUTANT PROBE	VIC-TCCCTCGACACTTTG-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	FORWARD PRIMER	5′-CCTGATGGGGAGAATGTGAA-3′		
	REVERSE PRIMER	5′-GTGGAGGGGCTTACGTCTAA-3′		
Guanine Nucleotide Binding Protein (G Protein), Alpha Stimulating Activity Polypeptide (*GNAS*) p.R201C, c.601C>T	WT PROBE	/5TET/tcgc+Tgc+Cgt+Gtcc/3IABkFQ/	LNA-ZEN	Integrated DNA Technologies
	MUTANT PROBE	/56-FAM/tcgc+Tgc+Tgt+Gtcc/3IABkFQ/	LNA-ZEN	Integrated DNA Technologies
	FORWARD PRIMER	5′-ACCTCAGATATATTTCTTCATG-3′		
	REVERSE PRIMER	5′-CCAGACAACTGTTCAAAC-3′		
Guanine Nucleotide Binding Protein (G Protein), Alpha Stimulating Activity Polypeptide (*GNAS*) p.R201H, c.602G>A	WT PROBE	/5TET/tcgc+Tgc+Cgt+Gtcc/3IABkFQ/	LNA-ZEN	Integrated DNA Technologies
	MUTANT PROBE	/56-FAM/cgct+Gcc+Atg+Tcct/3IABkFQ/	LNA-ZEN	Integrated DNA Technologies
	FORWARD PRIMER	5′-ACCTCAGATATATTTCTTCATG-3′		
	REVERSE PRIMER	5′-CCAGACAACTGTTCAAAC-3′		
Tumor Suppressor Protein P53 (*TP53*) p.R273H, c.818G>A	WT PROBE	VIC-ACAAACACGCACCTCA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	MUTANT PROBE	6FAM-ACAAACATGCACCTCA-MGBNFQ	TAQMAN MGB	LIFE TECHNOLOGIES
	FORWARD PRIMER	5′-TGGTAATCTACTGGGACGGAACAGC-3′		
	REVERSE PRIMER	5′-GGAGATTCTCTTCCTCTGTG-3′		

Table A2
Sixteen cancer-based assays evaluated, in addition to the *EGFR* L858R and T790M assays, following the approach described herein. The best empirical LoD was calculated for using the number of wild-type DNA molecules evaluated.

Gene	Protein variant	Nucleotide variant	Best empirical LoD	# WT DNA molecules evaluated	Estimated LoD with unlimited DNA
EGFR	p.E746_A750del	c.2235del15	1:648,000	1.7E+07	1:1.5 million (0.00006%)
BRAF	p.V600E	c.1799T>A	1:340,000	1.6E+07	1:615,000 (0.00016%)
PIK3CA	p.H1047R	c.3140A>G	1:1.3 million	2.0E+07	1:5 million (0.00009%)
PIK3CA	p.E545K	c.1633G>A	1:66,000	2.1E+07	1:80,000 (0.0012%)
ALK	p.R1275Q	c.3824G>A	1:26,000	2.1E+07	1:30,000 (0.003%)
KRAS	p.G12C	c.34G>T	1:940,000	2.5E+07	1:2.2 million (0.00004%)
KRAS	p.G12D	c.35G>A	1:45,000	3.5E+07	1:46,500 (0.002%)
KRAS	p.G12S	c.34G>A	1:59,000	2.3E+07	1:70,000 (0.0014%)
KRAS	p.G12V	c.35G>T	1:425,000	2.4E+07	1:720,000 (0.00014%)
KRAS	p.G12A	c.35G>C	1:290,000	2.2E+07	1:455,000 (0.0002%)
KRAS	p.G12R	c.34G>C	1:290,000	2.2E+07	1:450,000 (0.0002%)
KRAS	p.G13D	c.38G>A	1:19,000	2.0E+07	1:21,000 (0.005%)
HER2	p.L755S	c.2264C>T	1:185,000	2.4E+07	1:255,000 (0.0004%)
GNAS	p.R201C	c.601C>T	1:19,000	2.0E+07	1:21,000 (0.0048%)
GNAS	p.R201H	c.602G>A	1:17,000	2.0E+07	1:18,000 (0.0055%)
TP53	p.R273H	c.818G>A	1:19,600	2.0E+07	1:22,000 (0.0045%)

Table A3
Poisson correction for 58 negative samples for *EGFR* T790M assay. N is the total measured droplet count, N_{WT} is the count of wild-type only drops; N_{Mut} is the count of mutant only drops; λ_w is the measured wild-type loading, λ_m the mutant loading, calculated following Eq. (9)–(10); R is the ratio of mutant to wild type in the sample; R_{FP} is the average R of the 58 negative controls, $R_{FP} = 4.5e^{-5}$.

Sample	Droplet count			Loading		R	Target count	
	N	N_{Mut}	N_{WT}	λ_m	λ_w		#WT	#Mut
1	9,292,320	39	857,246	4.6E−06	9.7E−02	0.0048%	899,406	43
2	8,751,560	38	792,909	4.8E−06	9.5E−02	0.0050%	831,161	42
3	9,661,960	45	870,560	5.1E−06	9.4E−02	0.0054%	912,312	49
4	8,520,880	24	762,983	3.1E−06	9.4E−02	0.0033%	799,332	26
5	8,272,480	32	733,229	4.2E−06	9.3E−02	0.0046%	767,785	35
6	8,697,580	26	762,866	3.3E−06	9.2E−02	0.0036%	798,419	28
7	8,765,590	37	763,351	4.6E−06	9.1E−02	0.0051%	798,658	41
8	7,788,450	28	683,666	3.9E−06	9.2E−02	0.0043%	715,555	31
9	9,399,530	43	862,734	5.0E−06	9.6E−02	0.0052%	904,934	47
10	9,857,770	32	886,833	3.6E−06	9.4E−02	0.0038%	929,294	35
11	8,913,880	28	809,859	3.5E−06	9.5E−02	0.0036%	849,043	31
12	7,878,070	35	688,037	4.9E−06	9.1E−02	0.0053%	719,958	38
13	8,629,560	22	764,766	2.8E−06	9.3E−02	0.0030%	800,801	24
14	8,733,070	33	778,805	4.1E−06	9.3E−02	0.0044%	815,748	36
15	8,223,940	36	750,831	4.8E−06	9.6E−02	0.0050%	787,350	40
16	7,189,280	28	683,011	4.3E−06	1.0E−01	0.0043%	717,672	31
17	8,628,230	32	778,454	4.1E−06	9.5E−02	0.0043%	815,841	35
18	8,682,560	29	806,395	3.7E−06	9.7E−02	0.0038%	846,338	32
19	9,258,960	43	808,417	5.1E−06	9.1E−02	0.0056%	845,913	47
20	9,074,190	44	793,753	5.3E−06	9.2E−02	0.0058%	830,641	48

Table A3 (*Continued*)

Sample	Droplet count			Loading		*R*	Target count	
	N	N_{Mut}	N_{WT}	λ_m	λ_w		#WT	#Mut
21	8,833,630	23	778,397	2.9E−06	9.2E−02	0.0031%	814,853	25
22	8,650,790	32	775,301	4.1E−06	9.4E−02	0.0043%	812,273	35
23	9,359,870	34	867,254	4.0E−06	9.7E−02	0.0041%	910,104	37
24	8,749,980	49	807,034	6.2E−06	9.7E−02	0.0064%	846,716	54
25	9,924,810	36	904,048	4.0E−06	9.6E−02	0.0042%	947,911	40
26	9,273,750	44	852,782	5.2E−06	9.6E−02	0.0054%	894,579	48
27	9,238,320	33	828,852	3.9E−06	9.4E−02	0.0042%	868,423	36
28	10,142,200	40	891,111	4.3E−06	9.2E−02	0.0047%	932,718	44
29	9,452,310	49	831,402	5.7E−06	9.2E−02	0.0062%	870,267	54
30	10,307,200	46	901,286	4.9E−06	9.2E−02	0.0053%	943,155	50
31	8,824,910	45	836,557	5.6E−06	1.0E−01	0.0057%	878,912	50
32	10,015,800	38	935,156	4.2E−06	9.8E−02	0.0043%	981,740	42
33	9,594,880	38	881,745	4.4E−06	9.6E−02	0.0045%	924,931	42
34	9,626,790	42	906,768	4.8E−06	9.9E−02	0.0049%	952,365	46
35	8,497,390	40	767,334	5.2E−06	9.5E−02	0.0055%	804,222	44
36	9,338,540	46	834,294	5.4E−06	9.4E−02	0.0058%	873,946	51
37	8,948,330	27	802,349	3.3E−06	9.4E−02	0.0035%	840,629	30
38	9,194,420	29	824,632	3.5E−06	9.4E−02	0.0037%	863,986	32
39	9,665,900	33	909,850	3.8E−06	9.9E−02	0.0038%	955,568	36
40	8,631,750	41	813,778	5.2E−06	9.9E−02	0.0053%	854,739	45
41	8,451,160	31	784,877	4.0E−06	9.7E−02	0.0041%	823,754	34
42	10,825,000	41	1,020,960	4.2E−06	9.9E−02	0.0042%	1,072,370	45
43	10,332,800	45	950,164	4.8E−06	9.6E−02	0.0050%	996,733	50
44	9,414,420	29	857,467	3.4E−06	9.5E−02	0.0035%	899,065	32
45	9,909,550	32	900,658	3.6E−06	9.5E−02	0.0037%	944,253	35
46	9,023,920	28	832,170	3.4E−06	9.7E−02	0.0035%	873,079	31
47	9,110,150	36	823,009	4.3E−06	9.5E−02	0.0046%	862,591	40
48	8,620,500	25	784,679	3.2E−06	9.5E−02	0.0033%	822,721	28
49	8,238,160	31	749,670	4.1E−06	9.5E−02	0.0043%	786,005	34
50	8,225,370	29	734,577	3.9E−06	9.4E−02	0.0041%	769,475	32
51	8,585,660	29	772,546	3.7E−06	9.4E−02	0.0039%	809,543	32
52	9,058,930	43	859,634	5.2E−06	1.0E−01	0.0053%	903,205	48
53	10,044,100	30	957,633	3.3E−06	1.0E−01	0.0033%	1,006,415	33
54	9,342,810	40	868,684	4.7E−06	9.8E−02	0.0048%	911,765	44
55	8,217,480	43	757,698	5.8E−06	9.7E−02	0.0060%	794,942	47
56	9,048,170	44	823,129	5.3E−06	9.5E−02	0.0056%	863,012	48
57	8,718,730	29	798,698	3.7E−06	9.6E−02	0.0038%	837,684	32
58	9,225,920	35	829,971	4.2E−06	9.4E−02	0.0044%	869,709	38

Table A4
Poisson correction for 71 negative samples for *EGFR* L858R assay. $R_{FP} = 7.1e^{-8}$.

Sample	Droplet count			Loading		*R*	Target count	
	N	N_{Mut}	N_{WT}	λ_m	λ_w		#WT	#Mut
1	7,221,390	0	680,170	0.0E+00	9.9E−02	0.0E+00	714,367	0
2	8,099,920	0	774,847	0.0E+00	1.0E−01	0.0E+00	814,456	0
3	8,498,880	0	815,968	0.0E+00	1.0E−01	0.0E+00	857,841	0
4	8,436,800	0	830,430	0.0E+00	1.0E−01	0.0E+00	874,196	0
5	7,943,180	0	729,851	0.0E+00	9.6E−02	0.0E+00	765,589	0
6	7,454,740	0	685,925	0.0E+00	9.7E−02	0.0E+00	719,562	0
7	9,147,660	1	809,388	1.2E−07	9.3E−02	1.3E−06	847,459	1.1
8	9,288,860	0	834,079	0.0E+00	9.4E−02	0.0E+00	873,931	0
9	8,437,430	0	828,280	0.0E+00	1.0E−01	0.0E+00	871,808	0
10	6,545,580	0	636,378	0.0E+00	1.0E−01	0.0E+00	669,477	0
11	8,117,250	0	810,659	0.0E+00	1.1E−01	0.0E+00	854,053	0
12	8,710,990	0	854,799	0.0E+00	1.0E−01	0.0E+00	899,702	0
13	7,509,460	0	719,880	0.0E+00	1.0E−01	0.0E+00	756,762	0
14	8,328,420	0	798,532	0.0E+00	1.0E−01	0.0E+00	839,451	0
15	8,602,610	0	781,659	0.0E+00	9.5E−02	0.0E+00	819,480	0
16	9,974,820	0	894,498	0.0E+00	9.4E−02	0.0E+00	937,177	0
17	8,236,180	0	787,741	0.0E+00	1.0E−01	0.0E+00	828,001	0
18	7,667,560	0	748,526	0.0E+00	1.0E−01	0.0E+00	787,629	0
19	8,703,550	0	854,263	0.0E+00	1.0E−01	0.0E+00	899,149	0
20	8,519,800	0	836,964	0.0E+00	1.0E−01	0.0E+00	880,982	0
21	7,512,580	0	695,574	0.0E+00	9.7E−02	0.0E+00	729,912	0
22	8,167,940	0	772,241	0.0E+00	9.9E−02	0.0E+00	811,224	0
23	8,137,700	1	739,786	1.4E−07	9.5E−02	1.4E−06	775,600	1.1
24	6,794,800	0	766,088	0.0E+00	1.2E−01	0.0E+00	812,823	0
25	8,053,820	0	856,069	0.0E+00	1.1E−01	0.0E+00	905,071	0
26	8,350,220	0	943,007	0.0E+00	1.2E−01	0.0E+00	1,000,637	0

Table A4 (*Continued*)

Sample	Droplet count			Loading		*R*	Target count	
	N	N_{Mut}	N_{WT}	λ_m	λ_w		#WT	#Mut
27	8,712,660	0	969,456	0.0E+00	1.2E−01	0.0E+00	1,027,759	0
28	6,968,880	0	739,934	0.0E+00	1.1E−01	0.0E+00	782,239	0
29	8,043,370	0	834,480	0.0E+00	1.1E−01	0.0E+00	881,016	0
30	8,311,230	0	878,618	0.0E+00	1.1E−01	0.0E+00	928,616	0
31	9,628,460	0	999,325	0.0E+00	1.1E−01	0.0E+00	1,055,077	0
32	8,712,960	0	930,813	0.0E+00	1.1E−01	0.0E+00	984,384	0
33	9,316,300	0	930,892	0.0E+00	1.1E−01	0.0E+00	980,750	0
34	8,620,290	0	928,510	0.0E+00	1.1E−01	0.0E+00	982,424	0
35	8,928,080	0	963,637	0.0E+00	1.1E−01	0.0E+00	1,019,715	0
36	8,603,120	0	890,080	0.0E+00	1.1E−01	0.0E+00	939,568	0
37	8,377,540	0	875,373	0.0E+00	1.1E−01	0.0E+00	924,565	0
38	7,993,710	0	834,634	0.0E+00	1.1E−01	0.0E+00	881,499	0
39	9,129,150	0	975,721	0.0E+00	1.1E−01	0.0E+00	1,031,904	0
40	8,453,960	0	897,837	0.0E+00	1.1E−01	0.0E+00	949,183	0
41	9,034,090	0	908,822	0.0E+00	1.1E−01	0.0E+00	957,853	0
42	8,289,420	0	896,566	0.0E+00	1.1E−01	0.0E+00	948,858	0
43	8,441,650	0	921,643	0.0E+00	1.2E−01	0.0E+00	975,945	0
44	7,925,780	0	814,854	0.0E+00	1.1E−01	0.0E+00	859,854	0
45	8,879,910	0	919,898	0.0E+00	1.1E−01	0.0E+00	971,115	0
46	7,887,010	0	819,608	0.0E+00	1.1E−01	0.0E+00	865,396	0
47	9,008,630	0	942,661	0.0E+00	1.1E−01	0.0E+00	995,716	0
48	7,895,160	0	852,831	0.0E+00	1.1E−01	0.0E+00	902,503	0
49	8,621,540	0	875,833	0.0E+00	1.1E−01	0.0E+00	923,582	0
50	8,792,080	0	937,430	0.0E+00	1.1E−01	0.0E+00	991,268	0
51	8,656,130	1	941,022	1.3E−07	1.2E−01	1.1E−06	996,210	1.1
52	8,269,540	0	852,238	0.0E+00	1.1E−01	0.0E+00	899,424	0
53	8,578,770	0	876,470	0.0E+00	1.1E−01	0.0E+00	924,547	0
54	8,464,580	0	857,687	0.0E+00	1.1E−01	0.0E+00	904,318	0
55	8,835,740	0	911,090	0.0E+00	1.1E−01	0.0E+00	961,564	0
56	8,568,860	0	936,126	0.0E+00	1.2E−01	0.0E+00	991,319	0
57	8,713,740	0	890,786	0.0E+00	1.1E−01	0.0E+00	939,680	0
58	8,449,810	0	921,728	0.0E+00	1.2E−01	0.0E+00	975,984	0
59	8,726,290	0	941,135	0.0E+00	1.1E−01	0.0E+00	995,858	0
60	7,648,540	0	812,500	0.0E+00	1.1E−01	0.0E+00	858,978	0
61	8,378,460	1	881,857	1.3E−07	1.1E−01	1.2E−06	931,803	1.1
62	7,780,120	0	823,354	0.0E+00	1.1E−01	0.0E+00	870,261	0
63	8,732,820	0	911,819	0.0E+00	1.1E−01	0.0E+00	963,019	0
64	8,778,120	0	977,732	0.0E+00	1.2E−01	0.0E+00	1,036,598	0
65	8,157,080	0	863,753	0.0E+00	1.1E−01	0.0E+00	912,993	0
66	8,359,880	0	928,319	0.0E+00	1.2E−01	0.0E+00	984,026	0
67	9,002,200	0	999,617	0.0E+00	1.2E−01	0.0E+00	1,059,601	0
68	8,851,780	0	941,824	0.0E+00	1.1E−01	0.0E+00	995,793	0
69	9,687,240	0	977,081	0.0E+00	1.1E−01	0.0E+00	1,029,943	0
70	8,771,750	0	954,042	0.0E+00	1.2E−01	0.0E+00	1,010,022	0
71	8,368,770	0	924,206	0.0E+00	1.2E−01	0.0E+00	979,337	0

Table A5
Poisson correction for 16 mutant titration samples for *EGFR* T790M assay. *[WT]* is the measured concentration of wild type DNA and *[Mut]* the concentration of mutant DNA in the sample. *#Mut target* is the measured count of mutant DNA molecules. $\Lambda_{FP} = R_{FP}{}^{*}N^{*}\lambda_w$; *p*-value is calculated following Eq. (14). $R_{FP} = 4.5e^{-5}$.

Sample	Droplet count			Loading		Mut/WT	[Target] (copies/μl)		*p*-Value calculation		
	N	#WT	#Mut	λ_w	λ_m	*R*	[WT]	[Mut]	#Mut target	λ_{FP}	*p*-Value
1	9,142,560	814,578	41	9.3E−02	4.9E−06	0.0053%	1.9E+04	9.8E−01	45	39	1.7E−01
2	9,838,690	903,463	44	9.6E−02	4.9E−06	0.0051%	1.9E+04	9.8E−01	48	43	2.4E−01
3	9,573,960	870,985	44	9.5E−02	5.1E−06	0.0053%	1.9E+04	1.0E+00	48	41	1.7E−01
4	8,994,840	789,122	44	9.2E−02	5.4E−06	0.0058%	1.8E+04	1.1E+00	48	37	5.4E−02
5	9,447,200	843,816	74	9.4E−02	8.6E−06	0.0092%	1.9E+04	1.7E+00	81	40	8.7E−09
6	9,121,840	810,395	80	9.3E−02	9.6E−06	0.010%	1.9E+04	1.9E+00	88	38	1.3E−11
7	9,583,900	839,461	82	9.2E−02	9.4E−06	0.010%	1.8E+04	1.9E+00	90	40	1.4E−11
8	9,861,590	897,266	108	9.5E−02	1.2E−05	0.013%	1.9E+04	2.4E+00	119	43	0.0E+00
9	8,282,410	745,192	470	9.4E−02	6.2E−05	0.066%	1.9E+04	1.2E+01	516	35	0.0E+00
10	9,669,580	845,055	524	9.1E−02	5.9E−05	0.065%	1.8E+04	1.2E+01	574	40	0.0E+00
11	9,470,450	863,718	549	9.6E−02	6.4E−05	0.067%	1.9E+04	1.3E+01	604	41	0.0E+00
12	10,087,300	917,273	566	9.5E−02	6.2E−05	0.065%	1.9E+04	1.2E+01	623	44	0.0E+00
13	9,257,860	826,623	5211	9.4E−02	6.2E−04	0.66%	1.9E+04	1.2E+02	5724	39	0.0E+00
14	9,830,510	871,282	5362	9.3E−02	6.0E−04	0.64%	1.9E+04	1.2E+02	5886	41	0.0E+00
15	9,690,850	867,214	5362	9.4E−02	6.1E−04	0.65%	1.9E+04	1.2E+02	5891	41	0.0E+00
16	10,503,100	920,328	5710	9.2E−02	6.0E−04	0.65%	1.8E+04	1.2E+02	6261	44	0.0E+00

Table A6
Poisson correction for 16 mutation titration samples for *EGFR* L858R. $R_{FP} = 7.1e^{-8}$.

Sample	Droplet Count			Loading		Mut/WT	[target] (copies/μl)		*p*-Value calculation		
	N	#WT	#Mut	λ_w	λ_m	*R*	[WT]	[Mut]	#Mut target	λ_{FP}	*p*-Value
1	8,368,910	884,438	2	1.1E−01	2.7E−07	2.4E−06	2.2E+04	5.3E−02	2.2	0.07	2.1E−03
2	8,402,990	836,013	5	1.0E−01	6.6E−07	6.3E−06	2.1E+04	1.3E−01	5.6	0.06	7.3E−09
3	8,680,540	918,607	3	1.1E−01	3.9E−07	3.5E−06	2.2E+04	7.7E−02	3.4	0.07	5.1E−05
4	8,916,470	906,360	2	1.1E−01	2.5E−07	2.3E−06	2.1E+04	5.0E−02	2.2	0.07	2.2E−03
5	8,349,320	910,776	46	1.2E−01	6.2E−06	5.4E−05	2.3E+04	1.2E+00	52	0.07	0.0E+00
6	8,428,520	890,651	49	1.1E−01	6.5E−06	5.8E−05	2.2E+04	1.3E+00	55	0.07	0.0E+00
7	8,641,460	928,568	45	1.1E−01	5.8E−06	5.1E−05	2.3E+04	1.2E+00	50	0.07	0.0E+00
8	8,393,990	923,295	41	1.2E−01	5.5E−06	4.7E−05	2.3E+04	1.1E+00	46	0.07	0.0E+00
9	8,213,490	851,897	460	1.1E−01	6.2E−05	5.7E−04	2.2E+04	1.2E+01	513	0.06	0.0E+00
10	8,832,850	916,347	495	1.1E−01	6.3E−05	5.7E−04	2.2E+04	1.3E+01	552	0.07	0.0E+00
11	8,079,670	848,026	483	1.1E−01	6.7E−05	6.0E−04	2.2E+04	1.3E+01	540	0.06	0.0E+00
12	8,715,960	903,618	440	1.1E−01	5.6E−05	5.1E−04	2.2E+04	1.1E+01	491	0.07	0.0E+00
13	7,638,330	797,689	3985	1.1E−01	5.8E−04	5.3E−03	2.2E+04	1.2E+02	4451	0.06	0.0E+00
14	8,604,530	886,040	4346	1.1E−01	5.6E−04	5.2E−03	2.2E+04	1.1E+02	4847	0.07	0.0E+00
15	8,386,260	851,848	4256	1.1E−01	5.7E−04	5.3E−03	2.1E+04	1.1E+02	4739	0.06	0.0E+00
16	7,999,900	844,015	4293	1.1E−01	6.0E−04	5.4E−03	2.2E+04	1.2E+02	4801	0.06	0.0E+00

Table A7
Calculation of LoB and LoD from *EGFR* T790M assay data. R_{FP} is the average false positive ratios of negative control samples, #WT_avg is the average of the wild-type load per sample, Λ_{FP} is the average of the false mutant counts. LoB and LoD are limit of blank and lower limit of detection of mutant counts at #W_avg load, R_LoB and R_LoD are the limit of blank and lower limit of detection expressed in *R*.

R_{FP}	#WT_avg	Λ_{FP}	LoB	LoD	R_LoB	R_LoD
4.5E−05	8.6E+05	39	50.1	64.0	0.0058%	0.0074%

Table A8
Calculation of LoB and LoD for *EGFR* L858R.

R_{FP}	#WT_avg	Λ_{FP}	LoB	LoD	R_LoB	R_LoD
7.1E−08	9.1E+05	0.06	1.3	5.0	0.00014%	0.00054%

References

[1] Mullis KFF, Scharf S, Saiki R, Horn G, Erlich H. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. Cold Spring Harb Symp Quant Biol 1986;LI:263–73.

[2] Taly V, Pekin D, Abed AE, Laurent-Puig P. Detecting biomarkers with microdroplet technology. Trends Mol Med 2012;18(7):405–16, http://dx.doi.org/10.1016/j.molmed.2012.05.001.

[3] Taly V, Pekin D, Benhaim L, Kotsopoulos SK, Le Corre D, Li X, et al. Multiplex picodroplet digital PCR to detect KRAS mutations in circulating DNA from the plasma of colorectal cancer patients. Clin Chem 2013;59(12):1722–31, http://dx.doi.org/10.1373/clinchem.2013.206359.

[4] Chen WW, Balaj L, Liau LM, Samuels ML, Kotsopoulos SK, Maguire CA, et al. BEAMing and droplet digital PCR analysis of mutant IDH1 mRNA in glioma patient serum and cerebrospinal fluid extracellular vesicles. Mol Ther Nucleic Acids 2013;2:e109, http://dx.doi.org/10.1038/mtna.2013.28.

[5] Ladanyi M, Pao W. Lung adenocarcinoma: guiding EGFR-targeted therapy and beyond. Mod Pathol 2008;21(2):3801018.

[6] Bell DW, Gore I, Okimoto RA, Godin-Heymann N, Sordella R, Mulloy R, et al. Inherited susceptibility to lung cancer may be associated with the T790M drug resistance mutation in EGFR. Nat Genet 2005;37(12):1315–6.

[7] Pohl G, Shih L-M. Principle and applications of digital PCR. Expert Rev Mol Diagn 2004;4(1):41–7.

[8] Dube S, Qin J, Ramakrishnan R. Mathematical analysis of copy number variation in a DNA sample using digital PCR on a nanofluidic device. PLoS ONE 2008;3(8):e2876, http://dx.doi.org/10.1371/journal.pone.0002876.

[9] Heyries KA, Tropini C, Vaninsberghe M, Doolin C, Petriv OI, Singhal A, et al. Megapixel digital PCR. Nat Methods 2011;8(8):649–51.

[10] Motulsky HJ. Intuitive biostatistics. 2nd ed. New York, New York, USA: Oxford University Press; 2009.

[11] Box GH, Hunter WG, Hunter JS. Statistics for experimenters: an introduction to design, data analysis, and model building. Wiley Series in Probability and Statistics; 1978.

[12] National Institute of Standards and Technology. Engineering Statistics Handbook; 2012. Available from: http://www.itl.nist.gov/div898/handbook/pmc/section3/pmc331.htm

[13] Wittwer DNA Lab. uCount for Digital PCR: University of Utah; 2013. Available from: https://www.dna.utah.edu/ucount/uc.php

[14] Pezzullo JC. Exact Binomial and Poisson Confidence Intervals; 2014. Available from: http://statpages.org/confint.html

[15] Clinical and Laboratory Standards Institute. Protocols for determination of limits of detection and limits of quantitation, approved guideline. Wayne, PA, USA: CLSI; 2004. Contract No.: EP17.

[16] Armbruster DPT. Limit of blank, limit of detection and limit of quantitation. Clin Biochem Rev 2008;29(Suppl. 1):S49–52.

[17] Linnet K, Kondratovich M. Partly nonparametric approach for determining the limit of detection. Clin Chem 2004;50(4):732–40, http://dx.doi.org/10.1373/clinchem.2003.029983.

[18] Currie L. Nomenclature in evaluation of analytical methods including detection and quantification capabilities (IUPAC Recommendations 1995). Pure Appl Chem 1995;67(10):1699–723.

[19] Keohavong P, Thilly WG. Fidelity of DNA polymerases in DNA amplification. Proc Natl Acad Sci USA 1989;86(23):9253–7.