

The following link provides read-only access to the published version: <https://rdcu.be/b7Cv4>

Real-time cdPCR opens a window into events occurring in the first few PCR amplification cycles
<Version: Accepted>

David L. Duewer*, Margaret C. Kline, Erica L. Romsos

Material Measurement Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899-8390 USA

* David L. Duewer, Corresponding Author
NIST, 100 Bureau Drive, Stop 8390
Gaithersburg, MD 20899-8390 USA
david.duewer@NIST.gov
(T) 301-975-3935, (F) 301-926-8671

ABSTRACT

Polymerase chain reaction (PCR) end-point limiting dilution techniques, collectively termed "digital PCR (dPCR)", have been proposed as providing a potentially primary method for DNA quantification. We are evaluating several commercially available dPCR systems for use in certifying mass concentration in human genomic DNA reference materials. To better understand observed anomalies among results from chamber- and droplet-dPCR (cdPCR and ddPCR) systems, we have developed a graphical tool for evaluating and documenting the performance of PCR assays in real-time cdPCR systems: the ogive plot, the cumulative distribution of crossing threshold values. The ogive structure appears to embed information about early amplification events. We have successfully simulated ogives observed with different assays and reaction conditions using a four-stage amplification model parameterized by the probability of creating an intact 1) first generation "long" amplicon of indeterminate length from an original DNA target, 2) second generation defined-length amplicon from a long amplicon, and 3) defined-length amplicon from another defined-length amplicon. We are using insights from this model to optimize dPCR assay design and reaction conditions and to help validate assays proposed for use in value-assigning DNA reference materials.

KEYWORDS

digital polymerase chain reaction (dPCR)
crossing threshold
cumulative distribution
ogive plot
polymerase chain reaction model

INTRODUCTION

Polymerase chain reaction (PCR) assays perform best when used with an accurately determined quantity of input DNA. This is particularly true for highly multiplexed assays, such as those used for forensic human identification. [1] We are evaluating the suitability of digital PCR (dPCR) limiting dilution end-point technologies [2] for measuring the mass concentration of human genomic DNA in aqueous buffer. If dPCR can be successfully established as a potentially primary method [3], we anticipate using commercially available implementations to value assign a suite of calibration materials designed for use by the forensic community that will be stable over time and provide results that are traceable [4] to the International System of Units (SI).

Endpoint dPCR assays estimate the number of DNA targets present in a sample (copy number) from the number of positive and negative signals present at the end of a given number of PCR cycles. Both droplet dPCR (ddPCR) and some chamber dPCR (cdPCR) systems determine counts only after amplification is complete. Real-time cdPCR systems provide additional information by monitoring the signal intensities of a fixed array of reaction chambers at every amplification cycle. The resulting amplification curves (signal intensity as a function of cycle number) enable estimating a "crossing threshold (Ct)" for each positive signal: the cycle when the signal intensity in that chamber became positive. Ct values are typically interpolated from the signal intensities on either side of some set threshold value.

The cumulative distribution ("ogive") of these Ct estimates provides insight into assay performance. These ogives are easily calculated by rank-ordering the Ct values and plotting them against either their associated rank index scaled by the total number of chambers, i/N , or its Poisson transformation, $-\ln(1 - i/N) \times (\text{scale factor})$. Plotting against the transformed value enables comparing ogives derived from samples of different DNA concentration when the relative concentrations are known. Figure 1a displays the complete ogives for three treatments of one DNA using a commercial PCR assay. "Eyeball" examination of such low-resolution curves facilitate determining whether a real-time cdPCR assay has reached completion in a given number of amplification cycles. However, the detailed shape of the ogives as displayed in Figure 1b may be of more fundamental interest.

A considerable body of work on the quantitative analysis of real-time quantitative PCR (qPCR) amplification curves has recently been reviewed. [5,6] The mechanistic models underlying these analysis methods assume that amplification during the initial cycles is qualitatively the same as in cycles near the Ct. For qPCR, the large number of targets typically present in the input DNA ensures that signals from non-representative individual targets are lost in the ensemble average. As in qPCR, the shape of amplification curves for individual real-time cdPCR reaction chambers reflect only the dominant processes during cycles near the Ct. Unlike qPCR, in well-implemented cdPCR evaluations each reaction chamber contains at most a few individual targets. Since the ogive depicts the behavior of the ensemble of these sparsely-occupied chambers, the ogive shape potentially embeds information on non-dominant reaction processes.

Figure 2 presents a simplified model of the PCR amplification process, following the forward strand of a single double-stranded DNA (dsDNA) fragment. An original fragment, t_0 , is partially replicated from where a forward primer binds to wherever primer extension ends – a variable function of reaction conditions and time. This first-generation long amplicon, t_1 , is in turn partially replicated from the location of forward primer to that of the reverse primer binding site. This second-generation defined-length amplicon, t_2 , is in turn fully replicated in subsequent cycles. For book-keeping purposes, call these later-generation defined-length amplicons t_3 . There are then four processes associated with successful amplification: t_0 to t_1 with probability p_{01} , t_1 to t_2 with probability p_{12} , t_2 to t_3 with probability p_{23} , and t_3 to another t_3 . Since there are no molecular differences between the t_2 and t_3 amplicons, the probability that a t_3 successfully amplifies is also p_{23} .

Table 1 enumerates the number of molecular entities that contain the target DNA sequence for the initial 12 PCR amplification cycles under the assumption that all amplification processes proceed with perfect efficiency. While the total number of targets doubles with each cycle, amplification during the initial few cycles is dominated by p_{01} and p_{12} . The p_{23} processes dominate replication at the fourth cycle and become the only significant contributors after the seventh. The Ct values of typical dPCR assays exceed 20 cycles.

In this report we present evidence that real-time cdPCR ogives do embed information about early-cycle amplification and investigate how the amplification probabilities influence ogive shape. These insights may be useful in optimizing the performance of both cdPCR and ddPCR assays and systems.

MATERIALS AND METHODS

dPCR systems

The work reported here was performed using the Fluidigm BioMark (San Francisco, CA) 12.765 Digital Array real time/end point limiting dilution assay system; similar results have been obtained using BioMark 48.770 arrays. The Fluidigm Digital PCR Analysis Tool provided by the manufacturer was used for all primary data reduction using assay-specific global intensity thresholds and a quality score threshold of 0.01. Cycles 1 through 60 were analyzed with the user global analysis method. Detailed results were exported into a spreadsheet for further manipulation.

The Bio-Rad QX100 Droplet Digital PCR System (Hercules, CA) system was used to investigate whether our real-time cdPCR insights were applicable to ddPCR systems. The manufacturer's software was used to determine the number of positive and negative droplets at the end of 40 cycles using assay-specific intensity thresholds. These results were exported into a spreadsheet for further manipulation.

Sample materials

All experimental ogives displayed here were obtained using a commercially obtained human genomic dsDNA that was component #16 of the discontinued SRM 2390 RFLP Profiling Standard. Each unit of SRM 2390 provided approximately 25 μL of 200 ng/ μL extracted single-donor male human genomic DNA in TE^{-4} buffer (10 mmol/L tris(hydroxymethyl)aminomethane HCl, 0.1 mmol/L ethylenediaminetetraacetic acid). These solutions had been stored at $-80\text{ }^{\circ}\text{C}$ from the time they were vialled in the late 1980's. Following discontinuation of the SRM, the solution in the remaining vials was pooled, diluted with TE^{-4} pH 8.0 buffer to have an absorbance of 1.0 at 260 nm at a path length of 1 cm, and stored in perfluoroalkoxy fluoropolymer (PFA) containers at $4\text{ }^{\circ}\text{C}$. This stock solution has a nominal mass concentration of 50 ng/ μL .

A heat-denatured version of the DNA stock was prepared as a single-strand DNA (ssDNA) control material using an Applied Biosystems GeneAmp PCR System 9700 thermal cycler (Thermo Fisher Scientific Inc.) to minimize mechanical shearing and achieve reproducible thermal conditions. The solution was heated to and held at $98\text{ }^{\circ}\text{C}$ for 15 min, flash cooled, and stored in a PFA container at $4\text{ }^{\circ}\text{C}$ until use.

Working solutions having nominal mass concentrations of 2 ng/ μL were prepared from these stocks by 1 \rightarrow 25 volumetric dilution into TE^{-4} pH 8.0 buffer.

A third version of the stock was prepared by selective fragmentation with the PstI restriction enzyme [7]. PstI cuts human dsDNA into fragments that average about 7,000 basepairs (bp) in length. PstI does not cut ssDNA. For every 50 μL reaction volume of the PstI cutting experiment, the following materials were combined in a 0.2 mL PCR reaction tube: 2 μL of 20,000 U/mL enzyme (New England BioLabs Inc., Ipswich, MA), 5 μL of 10X Buffer#3 (New England BioLabs, Inc), 0.5 μL of 10 mg/mL bovine serum albumin (100X BSA, New England BioLabs Inc), 32.5 μL of sterile DI water, and 10 μL of the 50 ng/ μL stock DNA. After preparing the solutions, the DNA was enzymatically cut by heating to $37\text{ }^{\circ}\text{C}$ for 1 h in a 9700 thermal cycler, then cooled and stored at $4\text{ }^{\circ}\text{C}$ until used. An additional 1 \rightarrow 5 dilution of the cut material with TE^{-4} buffer resulted in the 1 \rightarrow 25 dilution (2 ng/ μL).

Since there is no consensus on the infectious status of extracted DNA, all solutions were handled as biosafety level 1 materials capable of transmitting disease [8].

PCR assays

Assays developed for other PCR-based technologies can in principle be adapted, with suitable modification of reagents and conditions, for use with dPCR. Four PCR assays were used in this study, the commercial Quantifiler Human DNA Quantification Kit (Life Technologies, Foster City, CA) [9] and three NIST-developed PCR assays that probe conserved sequences adjacent to known short-tandem repeat loci: D6S474, D9S2157, and D14S1434. Electronic Supplementary Material (ESM) Table S1 lists the primers and probes for these three assays. These assays are presented in this report for their pedagogic, not quantitation, utility.

ESM Table S2 lists the sample composition for the three NIST-developed assays as used in the BioMark 12.765 (and 48.770) cdPCR systems. Samples for the Quantifiler assay combined 8.2 μL of the Quantifiler Human Primer Mix, 9.8 μL of the Quantifiler Human PCR Reaction Mix, 1.0 μL of 20X GE Loading Reagent, and 2.0 μL of diluted DNA. Eight μL of these mixtures were added to the appropriate sample inlet for each replicate panel of a 12.765 array. The arrays were filled using the BioMark IFC Controller MX and placed into the BioMark System for amplification and detection. Amplification conditions were 95 $^{\circ}\text{C}$ for 10 min, followed by 60 cycles of 15 s at 95 $^{\circ}\text{C}$ and 1 min at 60 $^{\circ}\text{C}$. The ramp speed between temperature set points was 2 $^{\circ}\text{C}/\text{s}$.

The Bio-Rad ddPCR system requires use of a proprietary master mix, therefore the composition of samples prepared for Quantifiler ddPCR assessment was, per ddPCR reaction: 10 μL of the Bio-Rad Supermix for Probes (No dUTP), 8.4 μL of the Life Technologies Quantifiler primer-probe mix, and 1.6 μL of ≈ 8 ng/ μL DNA solution. For each replicate assessment, 20 μL of the resulting solutions were loaded into a cell of a Bio-Rad DG8 droplet generation cartridge. Generated droplets were transferred to a 96-well plate, heat-sealed with foil, and PCR amplified. Amplification on a 9700 Thermal Cycler was as follows: 95 $^{\circ}\text{C}$ for 10 min, followed by 40 cycles of 15 s at 95 $^{\circ}\text{C}$ and 1 min at 60 $^{\circ}\text{C}$. Thermal cycler ramp rates were set at 70 % between temperatures. After the 40 cycles endpoint there was a 98 $^{\circ}\text{C}$ hold for 10 min, followed by a 4 $^{\circ}\text{C}$ hold until the samples were removed from the thermal cycler and put onto the QX100 droplet reader. At the end of amplification, the 96-well plate was transferred to the QX100 Droplet reader and the number of negative and positive droplets determined.

Computation

Ogive visualization, simulation, and other analysis of instrument-provided results was accomplished with spreadsheet-environment programs developed at NIST.

RESULTS AND DISCUSSION

Ogive structure

Figure 1a presents the complete Quantifiler assay ogives for the untreated, heat-treated, and PstI cut materials. We attribute the extremely slow rise of the ogive for the untreated DNA to the relative inaccessibility of a significant fraction of the proprietary Quantifiler target sequences. Cutting the input DNA with the PstI restriction endonuclease reduces the size of the t_0 targets from greater than 48000 basepairs (bp) to fragments averaging about 7000 bp. This makes the target sequences accessible to the assay, albeit apparently reducing their number. Heat treatment, denaturing dsDNA to ssDNA, increases the number of positive chambers by doubling the number of independently sorting fragments that carry the target sequence or its complement. However, the residual slow "tail" of the ogive for the heat-treated material suggests that the material retains a small fraction of relatively inaccessible targets.

Figure 1b displays the early portion of the ogives at higher visual resolution, revealing a striking "staircase" structure with variable-height raisers alternating with variable-length treads. In addition to requiring an additional amplification cycle to reach the longest tread, the ogive for the heat-treated material has treads corresponding to reaction chambers containing four, three, two, and one ssDNA t_0 . That is, there are treads that are $\ln_2(4) = 2$, $\ln_2(3) = 1.58$, and $\ln_2(2) = 1$ Cts lower than the longest tread. The ogives for the untreated and PstI cut materials have treads corresponding to chambers containing six ($\ln_2(6) = 2.58$ Cts lower), four, and two ssDNA t_0 ; that is, three, two, and one dsDNA t_0 . This is fully compatible with the assumption that the t_0 are independently and randomly dispersed as intact entities into the reaction chambers, a critical requirement for the mathematical transformation of the fraction of positive chambers into the number of t_0 present.

The ogives for the untreated and PstI cut materials also have a tread corresponding to chambers containing one ssDNA t_0 . This could indicate that a small fraction of the sample DNA in the original reaction mixture is denatured to ssDNA. However, the presence of a small tread in the ogive for the heated material at about one cycle above that corresponding to one ssDNA t_0 per chamber suggests this cannot be a complete explanation: there cannot be one-half of an ssDNA t_0 . The presence of this minor tread and the several small treads in the ogive for the untreated material suggest that a fraction of Quantifiler-accessible t_0 targets do not successfully create a t_1 during the first amplification cycle but rather during the second or following cycles. Thus the staircase structure embeds information about at least imperfect p_{01} events during the early cycles.

ESM Figure S1 displays high-resolution ogives for the three NIST-developed PCR assays. The staircase structure of Figure 1b is present, if less distinct, in the ogives of the D6S474 and D9S2157 assays. The ogive for D6S474 is about 1 Ct earlier than that for D9S2157 and the two curves are not fully parallel. There is only a hint of structure in the D14S1434 ogive.

Simulation as a decoder ring

Since ogive shape is not consistent across PCR assays, decoding embedded information is likely to require modeling the entire amplification process. The amplification model outlined in Figure 2 is oversimplified in that it ignores various non-target reactions, [10] assumes that the three target amplification processes are independent, and assumes that the processes do not change as products accumulate and reactants decrease. Nonetheless, we have developed a simulation system based upon it as a first step in decoding how different values for p_{01} , p_{12} , and p_{23} affect ogive shape.

The simulation proceeds by randomly allocating a given number of t_0 into a given number of reaction chambers and then determining the amplification success for all the t_0 , t_1 , t_2 or t_3 in each chamber during each cycle over a given number of amplification cycles. For each target present, a random number is drawn from the uniform $U(0,1)$ distribution (that is, there is an equal chance for the value to be in the range 0 to 1) and compared it to the appropriate amplification probability, p . Values less than or equal to p signify successful amplification, values larger signify failure. The counts of successfully amplified targets of each type are recorded at the end of each cycle. After the final cycle, the Ct value for each chamber is interpolated from the count of total targets present ($t_0 + t_1 + t_2 + t_3$, see Table 1) at each cycle. The entire set of estimated Ct values is then sorted by increasing magnitude and recorded. Results from multiple runs using the same set of parameter values can be combined to better define the expected ogive shape and its variability.

While the amplification probabilities are the parameters of direct interest, a number of others must be defined. All required parameters are listed in Table 2, along with the default values used in the present work. These parameters can be divided into three groups.

Group 1, amplification probability: Rather than requiring the amplification probabilities to always be exactly the same in all reaction chambers, the probabilities in each i^{th} chamber can be allowed to vary. For convenience, in the current simulation system the probabilities are drawn from bounded $N(\mu_{xy}, \sigma_{xy})$ Gaussian kernels where “ μ_{xy} ” defines the mean location for one of the three processes, “ σ_{xy} ” defines the dispersion (as a standard deviation), draws less than 0 are set to 0, and draws greater than 1 are set to 1. No other kernel distributions have been investigated.

Group 2, platform and assay specification: The number of reaction chambers and amplification cycles followed enable mimicking specific cdPCR platforms and assay practice. The number of t_0 randomly dispersed into the reaction chambers and the number of total targets at threshold enable tuning the model to mimic specific cdPCR analyses. Adjusting the threshold target number, N_{Ct} , is a convenient way of accounting for differences in probe “brightness” (fluorescence per released probe) and crossing threshold settings.

Group 3, simulation implementation: The number of independent simulation runs enables evaluating the expected dispersion of ogives across replicates. Each replicate is obtained using identical model parameters but with different initial random allocation of the t_0 into the reaction chambers and different probability values drawn from their defining kernels. In practice two replicates suffice for preliminary visualization, 20 for a well-defined estimate of median performance, and 100 for a well-defined estimate of the 95 % confidence interval. The only parameter that is required just for computational efficiency is the number of t_3 needed to be present before invoking ensemble averaging. Before this number is exceeded, the simulation applies the p_{23} of each chamber to individual t_3 targets; the number of t_3 produced is the count of the successful amplifications. After the number is exceeded, the probability is applied once to the total number of t_3 present; the number of t_3 produced is $\text{INT}((2^{p_{23,i}} - 1) \cdot (\text{number of } t_3 \text{ present}))$ where the function INT truncates the value to have integer value. Our implementation arbitrarily sets the value of this parameter at 1000, a value exceeded at about the 10th cycle for p_{23} values close to 1. No visually appreciable differences in ogive shape were observed using a value of 500, but the sensitivity of the simulation to smaller values has not been investigated.

Influence of p_{01}

Figure 3a depicts simulated ogives for a series of $p_{01} = N(\mu_{01}, 0)$ values with $p_{12} = p_{23} = N(1, 0)$. These ogives embody the assumptions 1) that the amplification probabilities are identical in all reaction chambers and 2) during every cycle every t_1 produces a t_2 and every t_2 or t_3 produces a t_3 .

The sharp-cornered staircase of $p_{01} = 1$ depicts an ideal assay where the total number of targets exactly doubles at every cycle; see Table 1. Under this assumption, the tread lengths are exactly proportional to the number of reaction chambers that contain the same number of t_0 and the height of the riser to that tread is proportional to the logarithm of that number of t_0 . Decreasing the p_{01} values introduces increasing curvature to the right-edge of each tread. For $p_{01} = 0.86$, the number of amplification cycles required for all chambers containing at least one t_0 to cross the signal threshold and be considered positive (that is, for the assay to reach completion) is increased by about two. The left edge of the treads continues to be sharp until p_{01} is less than about 0.70.

Figure 3b depicts simulated ogives for a series of $p_{01} = N(0.86, \sigma_{01})$ values, again with $p_{12} = p_{23} = N(1, 0)$. These ogives embody the assumptions that 1) the probability of a t_0 successfully generating a t_1 varies among the reaction chambers, with the probability greater than 0.86 in about half of the chambers and lower than 0.86 in the other half and 2) during every cycle every t_1 produces a t_2 and every t_2 or t_3 produces a t_3 . Under the assumption that the p_{01} variability is symmetric about its mean value, their variability does not much effect the staircase structure even when σ_{01} is large. The proportion of chambers with a less-than average p_{01} appears to be largely balanced by the symmetric proportion of chambers with a greater-than average p_{01} .

Influence of p_{12}

The assumptions embodied in the ogives depicted in Figure 4 are similar to those of Figure 3, except p_{12} is varied and p_{01} is held constant. The consequences of those assumptions on the staircase structure are also similar, except that 1) the ogives are considerably smoother and largely without minor treads and 2) a given decrease in p_{12} has somewhat less effect on assay completion than does the same decrease in p_{01} . For $p_{12} = 0.86$, completion is delayed by less than one cycle.

Influence of p_{23}

Figure 5a depicts simulated ogives for a series of $p_{23} = N(\mu_{23}, 0)$ values with $p_{01} = p_{12} = N(1, 0)$. These ogives embody the assumptions 1) that the amplification probabilities are identical in all reaction chambers and 2) during every cycle every t_0 produces a t_1 and every t_1 produces a t_2 . As p_{23} becomes smaller, the staircase structure emerges at later amplification cycles but is otherwise relatively unchanged.

Figure 5b depicts simulated ogives for a series of $p_{23} = N(0.94, \sigma_{23})$ values, again with $p_{01} = p_{12} = N(1, 0)$. These ogives embody the assumptions 1) that the probability of a t_2 or t_3 successfully generating a t_3 varies among the reaction chambers, with the probability greater than 0.94 in about half of the chambers and lower than 0.94 in the other half and 2) during every cycle every t_0 produces a t_1 and every t_1 produces a t_2 . Unlike variability in p_{01} or p_{12} , fairly small differences in p_{23} among the reaction chambers can dramatically impact ogive shape. Chambers in which p_{23} is smaller than the mean result in those chambers crossing the threshold after more cycles and thus having a larger Ct. Chambers in which p_{23} is larger than the mean result in smaller Ct. However, since the probability of amplification cannot be greater than 1, as p_{23} is made more variable the ogives become somewhat similar to those in Figure 3, albeit lacking the minor treads.

Modeling ogives

Figures 6a and 6b compare the observed Quantifiler ogives for the PstI cut and untreated materials displayed in Figure 1 with simulation results of models "tuned" to mimic the observed structure. The tuning required adjusting up to eight parameters: N_{Ct} , N_{t0} , μ_{01} , σ_{01} , μ_{12} , σ_{12} , μ_{23} , and σ_{23} . Rather than attempting to automate the tuning process before establishing the model's utility, values were iteratively adjusted until a close visual match was obtained. The starting values were assigned from: 1) N_{Ct} , the observed location of the one-dsDNA tread; 2) N_{t0} , the number of t_0 estimated from the endpoint count of positive chambers; 3) μ_{01} and σ_{01} , inspection of the riser shape (Figure 3); 4) μ_{12} and σ_{12} , inspection of the tread slope and curvature (Figure 4); and 5) μ_{23} and σ_{23} , inspection of the "roundedness" of the curve where the risers meet the treads (Figure 5).

Figure 6a displays a model for the PstI cut material. The wide confidence intervals around the risers relative to the treads is intrinsic to the staircase structure: random allocation of t_0 into the reaction chambers creates variability only in the number of chambers containing given numbers of t_0 , not in the Ct expected for chambers containing that number of t_0 .

Figure 6b displays a model for the untreated material. No single set of parameter values provided an adequate fit to the observed ogive. However, a satisfactory fit was achieved by combining two models differing only in their p_{01} values. The model for the ogive to the left of the second tread has a reasonably large μ_{01} value. The model for the right edge has a very small μ_{01} value. The N_{t_0} value estimated for the left-side model is well-supported by the evidence. However, using the same N_{t_0} value for the right-side is convenient but arbitrary: values from one-half to many times that value provide similar structure.

Figures 7a through 7c display models for the untreated material with the three NIST-developed PCR assays. While not constrained to do so, in all cases $\mu_{23} \geq \mu_{12} \geq \mu_{01}$. This agrees with the intuitive premise that the probability of a successful amplification should be higher for smaller fragments.

Application to ddPCR

If the Fluidigm real-time cdPCR results reflect intrinsic PCR processes and not platform-specific artifacts, the Quantifiler assay of PstI cut and untreated DNA should give somewhat different ddPCR results. SEM Figure S2 displays droplet fluorescence intensity at the end of 40 amplification cycles for PstI cut and untreated DNA. In keeping with the cdPCR observations, the number of droplets with fluorescence in the gap between the negative droplet background and the main positive droplet distribution (“rain”) is greater for the untreated DNA (about 1.5% of the total positive count) than for the PstI cut (about 0.8 %). However, the fraction of positive ddPCR droplets at 40 cycles for the untreated DNA is at least as large as that for the PstI cut whereas the fraction of positive cdPCR chambers for the untreated is much less than for the cut even at 60 cycles. This suggests that model parameters determined from real-time cdPCR ogives may not be descriptive of ddPCR or qPCR processes.

CONCLUSIONS

The simple PCR model described in Figure 2 accommodates the so-far observed range of ogive staircase structures. The success of this model suggests that these shapes embed information about the cdPCR process, in particular early-amplification events. Long sloping tails (Fig. 1a) or well-defined treads in the staircase above the one t_0 per chamber tread (Fig. 3a) indicate that the t_0 targets are not fully accessible; pre-treating the sample to improve accessibility or increasing the number of amplification cycles may be necessary to avoid undercounting the number of targets in a sample. Ogives lacking distinct treads (Fig. 4a) indicate poor t_0 or t_1 target-amplification efficiency; redesigning the assay to more efficiently amplify these targets or increasing the number of amplification cycles may be necessary. Ogives of typical staircase structure but delayed relative to those for other assays (Fig. 5a) may indicate relatively poor efficiency in amplifying the t_2 and t_3 short amplicon; this may require increasing the number of cycles to ensure complete amplification but otherwise has little impact on the endpoint determination. We postulate that the probability of amplifying the t_2 and t_3 targets, p_{23} , may correlate with the usual qPCR assay efficiency values. Of potential interest to designers of cdPCR platforms, the presence of non-zero values for the σ_{01} , σ_{12} , and σ_{23} dispersion parameters in the descriptive models (Figs. 3b, 4b, and 5b) may quantify among-chamber variability, perhaps arising in slight differences in chamber morphology or thermal properties that could be further optimized during manufacture. In addition to helping diagnosing potential problems, the ogive plot provides a convenient mechanism for documenting the effect of assay modifications.

While descriptive, we recognize the proposed probabilistic model is at best incomplete. We hope that our observations and initial development will spur others to more fully mine the information that is embedded in real-time cdPCR ogive structure.

DISCLAIMER

Certain commercial equipment, instruments, or materials are identified in this report to specify adequately experimental conditions or reported results. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the equipment, instruments, or materials identified are necessarily the best available for the purpose.

ACKNOWLEDGMENTS

This work was supported in part by the National Institute of Justice through an interagency agreement with the NIST Special Programs Office.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

- 1 Kline MC, Duewer DL, Redman JW, Butler JM (2003) NIST Mixed Stain Study #3: DNA quantitation accuracy and its influence on short tandem repeat multiplex signal intensity. *Anal Chem* 75(10):2463-2469
- 2 Sykes PJ, Neoh SH, Brisco MJ, Hughes E, Condon J, Morley AA (1992) Quantitation of targets for PCR by use of limiting dilution. *Biotechniques* 13(3):444-9
- 3 JCGM 200:2012 International vocabulary of metrology –Basic and general concepts and associated terms (VIM). Joint Committee for Guides in Metrology. Sèvres, France. <http://www.bipm.org/en/publications/guides/vim.html>
- 4 De Bièvre P, Dybkaer R, Fajgelj A, Hibbert DB (2011) Metrological traceability of measurement results in chemistry: Concepts and implementation. *Pure Appl Chem* 83(10):1873-1935
- 5 Tellinghuisen J, Spiess, A-N (2014) Comparing real-time quantitative polymerase chain reaction analysis methods for precision, linearity, and accuracy of estimating amplification efficiency. *Anal Biochem* 449:76-82
- 6 Tellinghuisen J, Spiess, A-N (2014) Statistical uncertainty and its propagation in the analysis of quantitative polymerase chain reaction data: Comparison of methods. *Anal Biochem* 464:94-102
- 7 Fuchs R, Blakesley R (1983) Guide to the use of type II restriction endonucleases. *Methods Enzymol.* 100:3-38
- 8 CDC/NIH: Biosafety in Microbiological and Biomedical Laboratories, 5th ed.; HHS publication No. (CDC) 21-1112; Chosewood, L.C.; Wilson, D.E.; Eds.; US Government Printing Office: Washington, DC (2009); available at <http://www.cdc.gov/biosafety/publications/bmbl5/index.htm>
- 9 Applied Biosystems (2003) Quantifiler human DNA quantitation kit and Quantifiler Y human male DNA quantitation kit user's manual, Part Number 4344790 Rev A. Foster City, CA
- 10 Stolovitzky G, Cecchi G (1996) Efficiency of DNA replication in the polymerase chain reaction. *Proc Natl Acad Sci USA* 93:12947-12952

TABLES

Table 1. Target numbers in the first 12 cycles of a "perfect" PCR amplification

Cycle	t ₀		t ₁		t ₂		t ₃		Σ(t ₀ ,t ₁ ,t ₂ ,t ₃) Present	% (t ₂ + t ₃)
	Present	Produced	Present	Produced	Present	Produced	Present	Produced		
0	1								1	
1	1	1	1						2	0.0
2	1	1	2	1	1				4	25.0
3	1	1	3	2	3	1	1		8	50.0
4	1	1	4	3	6	4	5		16	68.8
5	1	1	5	4	10	11	16		32	81.3
6	1	1	6	5	15	26	42		64	89.1
7	1	1	7	6	21	57	99		128	93.8
8	1	1	8	7	28	120	219		256	96.5
9	1	1	9	8	36	247	466		512	98.0
10	1	1	10	9	45	502	968		1024	98.9
11	1	1	11	10	55	1013	1981		2048	99.4
12	1	1	12	11	66	2036	4017		4096	99.7

Table 2. Model Parameters

Symbol	Use	Default
$N(\mu_{01},\sigma_{01})$	$p_{01,i} = \text{MIN}(1, \text{MAX}(0, N(\mu_{01},\sigma_{01})))^a$	$N(1,0)$
$N(\mu_{12},\sigma_{12})$	$p_{12,i} = \text{MIN}(1, \text{MAX}(0, N(\mu_{12},\sigma_{12})))^a$	$N(1,0)$
$N(\mu_{23},\sigma_{23})$	$p_{23,i} = \text{MIN}(1, \text{MAX}(0, N(\mu_{23},\sigma_{23})))^a$	$N(1,0)$
N_{chmbr}	number of reaction chambers	765
N_{cycle}	number of amplification cycles	60
N_{Ct}	number of targets at threshold crossing	2^{27}
N_{t_0}	number of t ₀	300
N_{reps}	number of independent simulations to evaluate and combine	2 to 100
N_{avg}	number of t ₃ required to use an ensemble average	1000

a The functions MIN and MAX are to be interpreted as “Take the {minimum, maximum} of the values within the parenthesis.”

FIGURE CAPTIONS

- Fig 1. Observed cumulative distribution of Ct values as functions of the fraction of positive chambers (ogives) for the Quantifiler assay of untreated, heat-treated, and PstI endonuclease cut human genomic DNA. Each thick curve combines results from two Fluidigm 12.765 panels. Segment a) displays the complete ogives; Segment b) is a higher-resolution view of the ogives for the first six amplification cycles after amplification onset. The thin horizontal lines mark the Ct expected for a chamber containing a given number of single-stranded targets. The number of such targets is at the right edge of each line; the relative location of each line is proportional to the logarithm of that number.
- Fig 2. Schematic of PCR amplification. The horizontal lines represent a forward (5' to 3') strand of a dsDNA fragment that contains the target sequence or the forward strand of an amplicon derived from that fragment. The arrows represent the probabilities that an intact strand will be produced in one amplification cycle. See text for definitions.
- Fig 3. Simulated ogives for various values of p_{01} for $p_{12} = p_{23} = N(1,0)$. Each ogive is the median of 20 independent simulation runs. Figure 3a displays ogives for the series $p_{01} = N(\{1, 0.98, 0.94, 0.86, 0.70, 0.38\}, 0)$. Figure 3b displays ogives for the series $p_{01} = N(0.86, \{0.00, 0.10, 0.20, 0.40\})$ as well as the reference $N(1,0)$.
- Fig 4. Simulated ogives for various values of p_{12} for $p_{01} = p_{23} = N(1,0)$. Each ogive is the median of 20 independent simulation runs. Figure 4a displays ogives for the series $p_{12} = N(\{1, 0.98, 0.94, 0.86, 0.70, 0.38\}, 0)$. Figure 4b displays ogives for the series $p_{12} = N(0.86, \{0.00, 0.10, 0.20, 0.40\})$ as well as the reference $N(1,0)$.
- Fig 5. Simulated ogives for various values of p_{23} where $p_{01} = p_{12} = 1$. Each ogive is the median of 20 independent simulation runs. Figure 5a displays ogives for $p_{23} = N(\{1, 0.98, 0.96, 0.94, 0.92, 0.90\}, 0)$. Figure 5b displays ogives for $p_{23} = N(0.94, \{0.0, 0.005, 0.02, 0.005\})$ as well as the reference $N(1,0)$.
- Fig 6. Comparison of observed and simulated ogives for the Quantifiler assay of untreated and PstI cut DNA. The thick curves combine results from two Fluidigm 12.765 panels. The thinner solid curves are the median of 100 simulated ogives; the dotted curves are empirical 95 % confidence intervals. Simulation parameter values are displayed in the lower right of each segment. Figure 6a displays results for the PstI cut material. Figure 6b displays results for the untreated material. The simulated ogives in Figure 6b combine two models, identical except that one-half of the t_0 have $p_{01} = N(0.86, 0)$ and the other half have $p_{01} = N(0.012, 0)$.
- Fig 7. Comparison of observed and simulated ogives for three NIST-developed assays of untreated DNA. Each observed curve combines results from two Fluidigm 12.765 panels of untreated human genomic DNA. Figure 7a displays results for D6S474; 7b for D9S2157, and 7c for D14S1434. The segments are formatted as in Figure 6.

FIGURES

Figure 1

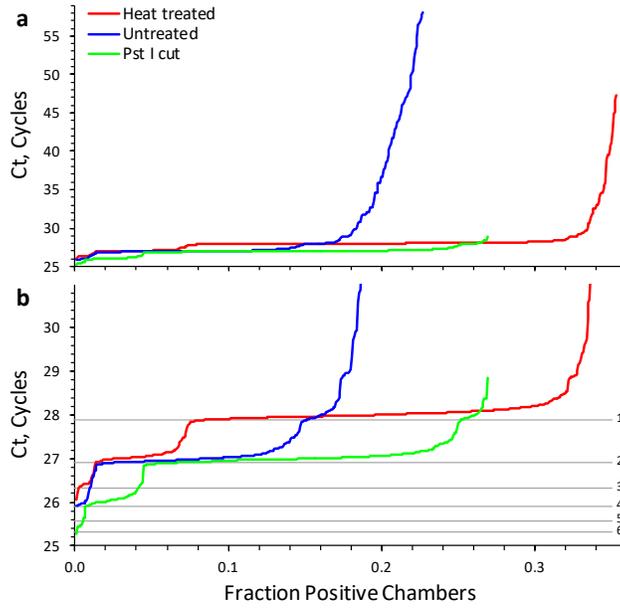


Figure 2

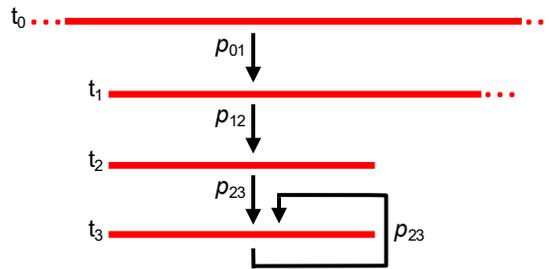


Figure 3

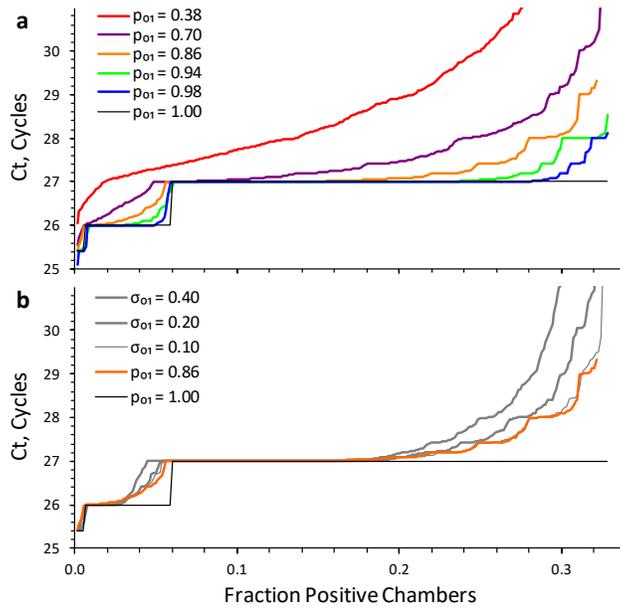


Figure 4

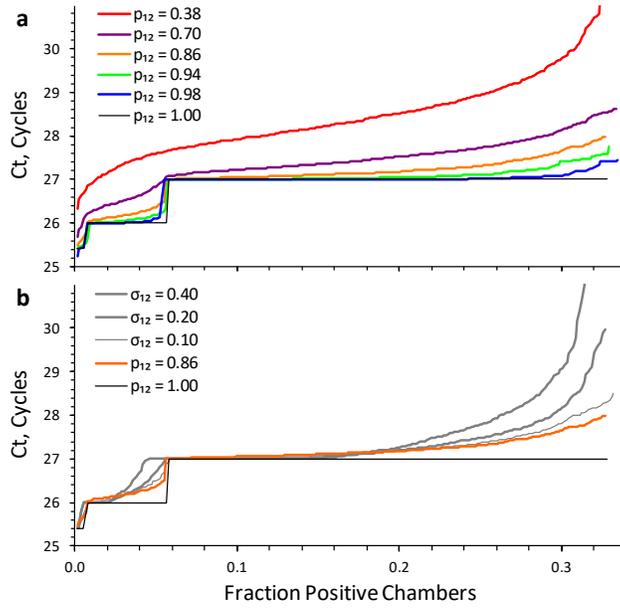


Figure 5

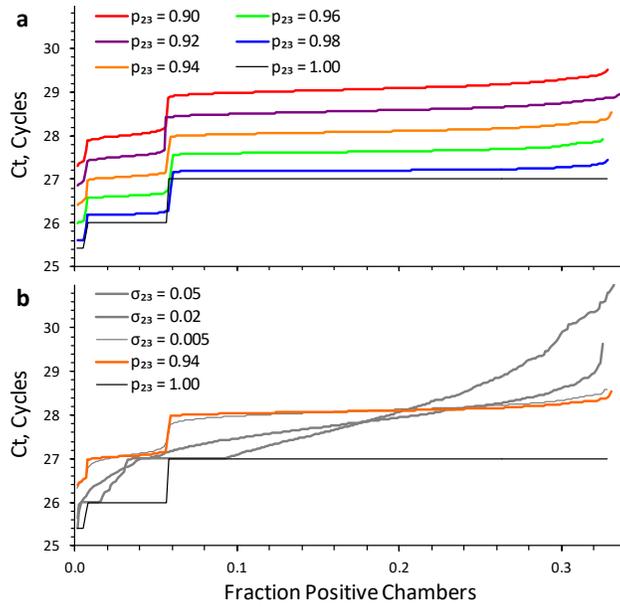


Figure 6

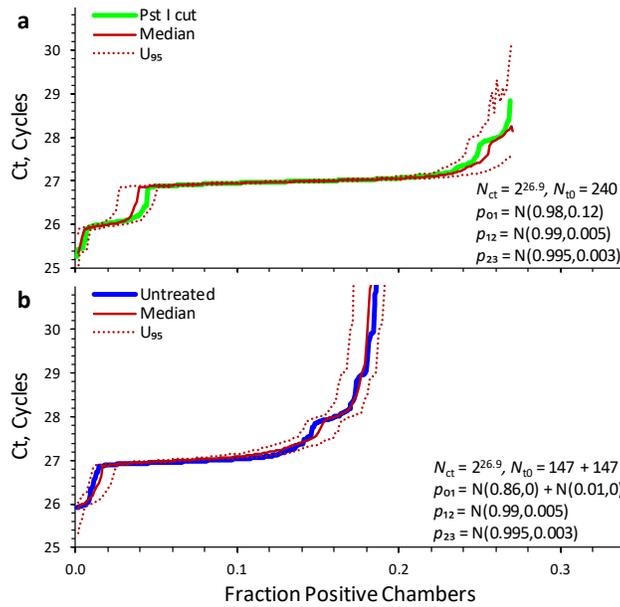
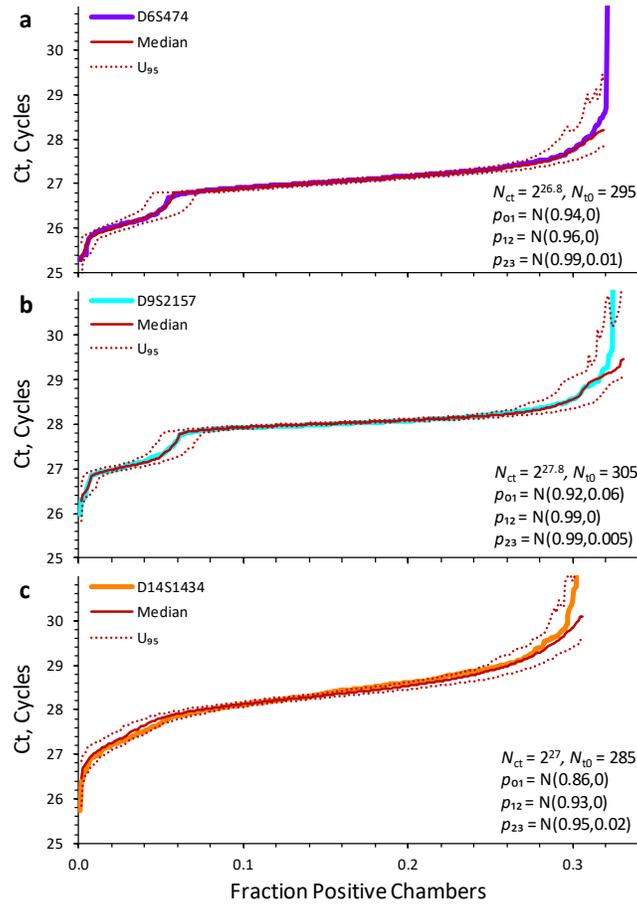


Figure 7



Electronic Supplementary Material for

Real-time cdPCR opens a window into events occurring in the first few PCR amplification cycles

David L. Duewer, Margaret C. Kline, Erica L. Romsos

Material Measurement Laboratory
National Institute of Standards and Technology
Gaithersburg, MD 20899-8390 USA

Table S1. Primers and Probe for NIST-Developed Assays

Locus	Type: Sequence	T _m , °C	Amplicon Length, bp
D6S474	F: ggccccagaaccaaggaa	62.2	150
	R: gcagcctcagggttctcaaa	62.4	
	P: atggatgatgaaccctc	69.0	
D9S2157	F: ggcttgctgggtactgctt	62.4	60
	R: ggaccacagcacatcagtcact	64.5	
	P: cagggcacatgaat	68.0	
D14S1434	F: cctcctattggtctacagttatttttaa	60.5	87
	R: caaggaccctgggtccat	62.3	
	P: ctccggagccagagc	69.0	

F: Forward primer, R: Reverse primer, P: probe, T_m: estimated melt temperature, bp: basepair

Table S2. Composition of Samples Prepared for the NIST-Developed Assays

Reaction component	D6S474 & D9S2157	D14S1434
TaqMan Universal PCR Master Mix, no AmpErase UNG	10.0 µL	10.0 µL
Forward primer, 5 µmol/L	1.2 µL	1.2 µL
Reverse primer, 5 µmol/L	1.2 µL	1.2 µL
Probe, 4 µmol/L	1.0 µL	1.0 µL
20X GE Reagent	1.0 µL	1.0 µL
Bovine serum albumin	0.0 µL	1.0 µL
Distilled water	3.6 µL	2.6 µL
DNA extract ^a	2.0 µL	2.0 µL

^a Diluted with TE⁻⁴ or distilled water to ≈2 ng/µL for 12.765 arrays, ≈8 ng/µL for 48.770 arrays

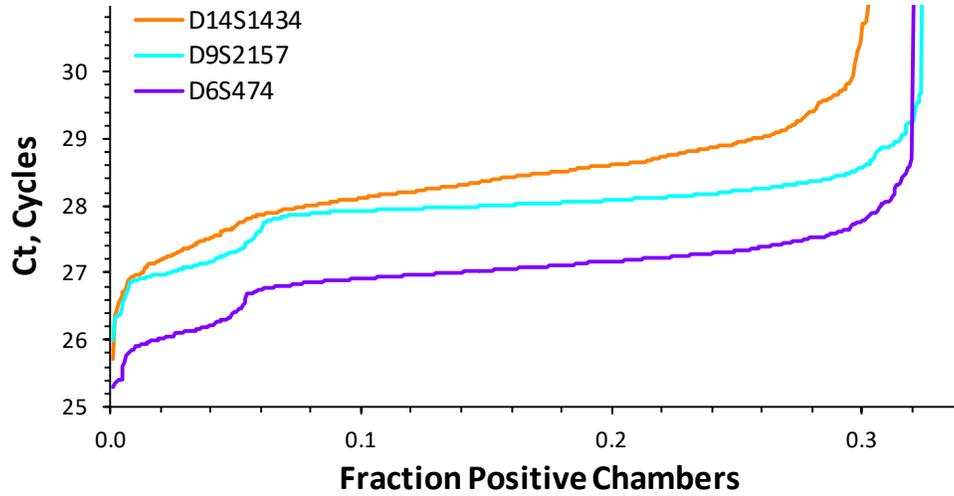


Figure S1: Observed ogives for three NIST-developed PCR assays: D14S1434, D9S2157, and D6S474. Each curve combines results from two Fluidigm 12.765 panels of untreated single-donor human genomic dsDNA extract in TE⁻⁴ buffer.

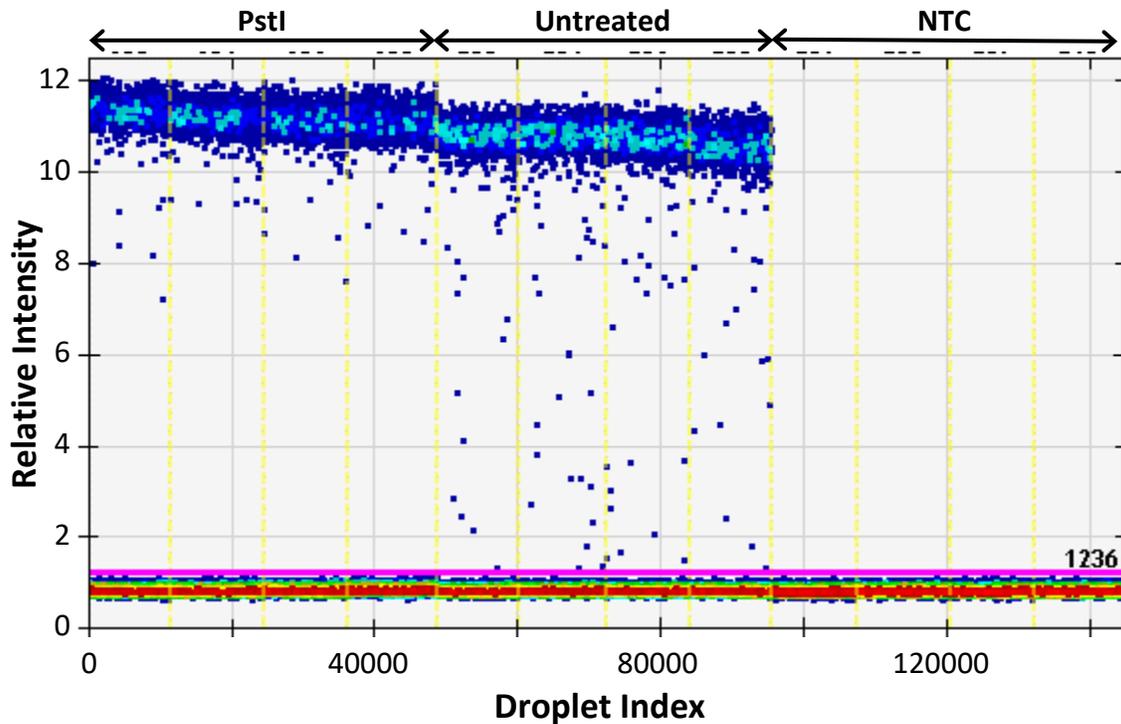


Figure S2. Comparison of PstI cut and untreated DNA ddPCR signals for the Quantifiler assay. Results for four replicate samples of the two materials are displayed, along with results for four non-template control (NTC) replicates. The horizontal axis reflects the droplet counting sequence; the vertical axis the relative droplet fluorescence intensity at the end of 40 amplification cycles. The droplets are color-coded by density at a given amplitude, with red for the most compact distribution and dark blue for the least common values. The horizontal magenta line represents the intensity threshold used to differentiate negative from positive droplets. The vertical yellow dotted lines separate droplets from different samples.