



# Low copy target detection by Droplet Digital PCR through application of a novel open access bioinformatic pipeline, 'definetherain'



Mathew Jones<sup>a</sup>, James Williams<sup>a</sup>, Kathleen Gärtner<sup>a,b</sup>, Rodney Phillips<sup>a,c,d</sup>, Jacob Hurst<sup>a,c</sup>, John Frater<sup>a,c,d,\*</sup>

<sup>a</sup> Peter Medawar Building for Pathogen Research, Nuffield Department of Clinical Medicine, Oxford, UK

<sup>b</sup> IUH, Hôpital St. Louis, Paris, France

<sup>c</sup> Institute for Emerging Infections, The Oxford Martin School, Oxford, UK

<sup>d</sup> Oxford National Institute of Health Research Biomedical Research Centre, Oxford, UK

## ABSTRACT

### Article history:

Received 11 October 2013

Received in revised form 13 February 2014

Accepted 21 February 2014

Available online 2 March 2014

### Keywords:

Quantitative PCR

Droplet Digital PCR

HIV-1

k means clustering

Droplet Digital PCR (ddPCR) represents a new and alternative platform to conventional quantitative-PCR (qPCR) for the quantitation of DNA templates. However, the proposed improvement in sensitivity and reproducibility offered by ddPCR is not yet fully proven, partly because the delineation between positive and negative responses is not always clear.

Data are presented demonstrating the sensitivity of the ddPCR system to both reagent concentrations and choice of cut-off for defining positive and negative results. By implementing *k*-nearest clustering, cut-offs are produced that improve the accuracy of ddPCR where target DNA is present at low copy numbers, a key application of ddPCR. This approach is applied to human albumin and HIV-1 proviral DNA ddPCR quantitative protocols. This tool is coded in JavaScript and has been made available for free in a web browser at <http://www.definetherain.org.uk>. Optimisation of the analyses of raw ddPCR data using 'definetherain' indicates that low target number detection can be improved by its implementation. Further application to patient samples will help define the clinical utility of this approach.

© 2014 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

## 1. Introduction

Droplet Digital PCR (ddPCR) (Bio-Rad, CA, USA) is a novel platform designed to provide greater sensitivity and precision for the detection and quantitation of DNA target molecules (Hindson et al., 2011, 2013). The system partitions the PCR into approximately 20,000 droplets, each of which acts as an individual reaction. Each individual droplet is defined on the basis of fluorescent amplitude as being either positive or negative. By then quantifying the proportion of positive droplets, Poisson statistics are used to call the number of target templates from which absolute DNA levels can be calculated precisely without the need for a standard curve (Pinheiro et al., 2012). The potential advantage of ddPCR over conventional qPCR is sensitivity and accuracy at low template copy numbers (Strain et al., 2013). This is of particular interest in the field of HIV-1 'reservoir' research, where accurate quantitation of proviral DNA in

CD4T cells is critical to HIV-1 cure strategies (Eriksson et al., 2013; Finzi et al., 1997; Wong et al., 1997).

Quantitative real-time PCR (qPCR) is used widely to quantify HIV-1 DNA levels. Since qPCR uses an indirect measurement (cycle threshold) and any assay noise is amplified exponentially, there is the potential for limitations in data accuracy and reproducibility (Miotke et al., 2014). Also, the reliance on logarithmic standard curves may result in a trade-off between accuracy and reproducibility and high dynamic ranges, especially where standards are not optimised between laboratories.

In quantitation of HIV-1 copy number in clinical samples, ddPCR has been reported to be both as sensitive (Kiselinova et al., 2014) and more sensitive (Strain et al., 2013) than qPCR, but less specific at low HIV-1 template numbers (Henrich et al., 2012). In addition, ddPCR was found to be less sensitive than qPCR in quantitation of cytomegalovirus (Hayden et al., 2013) and *Chlamydia trachomatis* (Schachter, 2013) in clinical samples, suggesting that further methodological optimisation was needed for ddPCR to match the sensitivity of qPCR.

The manufacturer's software interprets the fluorescent amplitudes of each droplet to generate both a positive and a negative cluster. A proprietary method is then used to define a fluorescence

\* Corresponding author at: Peter Medawar Building for Pathogen Research, University of Oxford, South Parks Road, Oxford OX1 3SY, UK. Tel.: +44 01865 281230. E-mail address: [john.frater@ndm.ox.ac.uk](mailto:john.frater@ndm.ox.ac.uk) (J. Frater).

threshold, thus allowing each droplet to be classified as either positive or negative for target DNA. The Quantasoft software (Bio-Rad, CA, USA) provided by the manufacturer also allows a user-defined cutoff to be entered to define the threshold between the positive and negative clusters.

The droplet amplitude files generated by the machine can be exported in a format which allows users to generate their own droplet classification schemes to analyse the output. An open-access freely available web-based JavaScript program is presented, based on *k*-nearest neighbour clustering called 'definetherain'. This can be run within most web browsers to call more accurately the positive ddPCR droplets, with particular utility at low template numbers. The performance of 'definetherain' is compared with the rare event detection mode setting of the Quantasoft package, and the software is applied to control standards for analysis of HIV-1 proviral DNA and the human 'housekeeping' albumin gene.

## 2. Materials and methods

### 2.1. Preparation of standards

8E5/LAV cells (which each contain a single HIV-1 provirus) were cultured in RPMI 1640 medium supplemented with 10% fetal calf serum, antibiotics and L-glutamine. This cell line was obtained through the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH from Dr. Thomas Folks. Cells were assessed for their viability and counted using a disposable hemocytometer. The cells were then extracted using a QiaAmp DNA Blood Mini Kit (Qiagen, Hilden, Germany) and eluted into water. A master stock of  $10^5$  cell-equivalent genomes was made and subsequent serial dilutions were made to achieve a  $10^4$ ,  $10^3$ ,  $10^2$ ,  $10^1$  and  $10^0$  dilution series. 25  $\mu$ l aliquots of each concentration were made and stored at  $-20^\circ\text{C}$ .

### 2.2. Measurement of cell number and HIV-1 copy number using albumin and HIV-1 qPCR assays

Cell copy number was first quantified in triplicate at two dilutions using an adapted albumin qPCR (Schachter, 2013). The master mix contained  $2\times$  Lightcycler 480 Probes Master Mix (Roche, Welwyn Garden City, UK), 200 nM Probe (FAM – CCT GTC ATG CCC ACA CAA ATC TCT CC – BHQ-1), 250 nM Albumin.F (GCT GTC ATC TCT TGT GGG CTG T) and 250 nM Albumin.R (AAA CTC ATG GGA GCT GCT GGT T) (Eurofins MWG Operon, Ebersberg, Germany) with 10  $\mu$ l DNA sample in a total volume of 25  $\mu$ l. 10  $\mu$ l of DNA sample was assayed in triplicate for HIV-1 using 500 nM Probe (FAM – AGT RGT GTG TGC CCG TCT GTT G – BHQ-1), 500 nM LTR.OS (GRA ACC CAC TGC TTA ASS CTC AA) and 500 nM LTR.AS (TGT TCG GGC GCC ACT GCT AGA GA) (Eurofins MWG Operon) and  $2\times$  LightCycler 480 probes Master Mix, in a total volume of 25  $\mu$ l. Both qPCR amplifications were performed using the following program: one cycle of  $95^\circ\text{C}$  for 10 min; 45 cycles of  $95^\circ\text{C}$  for 15 s and  $60^\circ\text{C}$  for 1 min. The data were analysed using proprietary Roche LightCycler software. 8E5 cells (NIH AIDS reagent Program, Germantown, USA), that contain one integrated copy of HIV-1 per cell, were used in duplicate as qPCR standards, with cell and HIV-1 copy numbers ranging in serial 10-fold dilutions from  $1\times 10^5$  to  $1\times 10^0$  DNA copies per reaction. 8E5 cells were either diluted in water or spiked into a background of  $1\times 10^5$  PBMC DNA equivalents.

### 2.3. Droplet Digital PCR (ddPCR)

The two qPCR assays described above were optimised to work in the ddPCR system. The ddPCR mixture contained 8  $\mu$ l of  $2\times$  ddPCR Supermix (Bio-Rad, Hercules, USA), 400 nM of both the albumin and HIV-1 forward and reverse primers and 125 nM probe in each 20  $\mu$ l

reaction. The entire 20  $\mu$ l reaction was loaded into a droplet cartridge (Bio-Rad, Hercules, USA), a gasket placed over the cartridge according to the Bio-Rad protocol and the cartridge placed in the droplet generator (Bio-Rad #186-3002). Once inside the droplet generator a vacuum was applied to the cartridge. This draws both the PCR reagents and oil through a flow-focusing nozzle where around 20,000 individual droplets approximately 1 nl in size are formed, suspended in an emulsion (Strain et al., 2013).

The emulsion was transferred into a 96 well plate (Eppendorf, Hamburg, Germany) and sealed using a foil lid and a thermal plate sealer (Eppendorf, Hamburg, Germany). Sealed plates were cycled using a C-1000 thermal cycler (Bio-Rad, Hercules, USA) under the following conditions: 10 min hold at  $95^\circ\text{C}$ , 45 cycles of  $95^\circ\text{C}$  for 15 s then  $60^\circ\text{C}$  for 60 s. After amplification, the plate was transferred to a Bio-Rad droplet reader from which raw fluorescence amplitude data was extracted from the Quantasoft software for downstream analysis.

### 2.4. Development and application of 'definetherain' data analysis software

The fluorescence amplitude of each analysed droplet was exported and analysed using 'definetherain', a bespoke software package designed to improve calling of positive droplets when the template is of low copy number. The software is written in Javascript, and is available to use at <http://www.definetherain.org.uk>. A simple movie is also on-line demonstrating how to use the software. Full source and the data used to develop the program are also available at github: <https://github.com/jacobhurst/definetherain>.

The system uses a control well of known input copy number to define the amplitudes of a valid positive and negative response. The *k*-nearest neighbour algorithm is used with *k* set to 2. After defining the positive and negative clusters within this control sample, cutoffs are determined using each cluster's mean and standard deviation. When analysing subsequent test samples the amplitude of each droplet is examined. If the value is larger than the mean of the negative cluster plus three times the negative standard deviation (SD), and smaller than the mean of the positive cluster minus three times the positive SD, the droplet is defined as 'rain', and is removed from any further calculations. The concentration is then calculated using the following formula:

$$c = -\ln\left(\frac{N_{\text{neg}}/N}{V_{\text{droplet}}}\right)$$

Here,  $N_{\text{neg}}$  is the number of negative droplets,  $N$  is the total number of valid droplets and  $V_{\text{droplet}}$  is the volume per droplet (0.91  $\mu$ l). Having defined a positive and a negative response, the software provides the facility of loading and analysing other datasets using the cut-offs defined using the control.

## 3. Results

### 3.1. Adaptation of qPCR protocols to optimise ddPCR

Where ddPCR is to be used to enhance sensitivity and reproducibility of pre-existing qPCR protocols, it is likely that reagent concentrations and reaction conditions will need to be titrated to optimise the new platform. The manufacturers recommend initially running a gradient ddPCR, and subsequently selecting the temperature that provides a sufficient separation of the negative and positive droplets for a threshold to be set. The transfer of some assays may result in 'polydiversity'. This is where an unacceptable number of droplets fluoresce in the region between the defined positive and negative cut-offs, and as a result, these assays cannot

be used reproducibly in ddPCR (personal communication with Bio-Rad). In our hands, the optimisation of a number of conditions, in addition to annealing temperature was necessary.

The ddPCR platform was applied to two assays that were optimised for the quantitation of the human albumin gene (Fig. 1a) and HIV-1 proviral DNA (Fig. 1b). Initial testing of the qPCR protocols in the ddPCR workflow, indicated that the fluorescent probe concentration would have to be reduced to perform optimally on the ddPCR platform. A probe titration was performed with the HIV-1 assay and applying the QuantaSoft software's (QS) 'auto' threshold showed that at both the 250 and 500 nM probe concentrations, droplets that clustered with the negative population (green dots in Fig. 1c) were classified falsely as positive (red dots within the blue shaded regions in Fig. 1c). Probe concentration impacted considerably on both the sensitivity and specificity of the assay, with the concentration used for the optimised qPCR (500 nM) performing poorly in both regards. To assess the HIV-1 ddPCR's specificity at different probe concentrations, water controls were assayed in duplicate to determine if positive droplets would be detected by ddPCR. In this assay, false positives were virtually eliminated at both the 125 and 250 nM concentrations compared to the 500 nM concentration (Fig. 1d). The sensitivity of the assay using different probe concentrations was tested by assaying a known target input of 33,000 HIV-1 copies (Fig. 1e). Probe concentrations of 125, 250 and 500 nM were tested, and measured 25,480, 20,700 and 15,258 HIV-1 copies, respectively, with 125 nM concentration providing the closest result. At probe concentrations below 125 nM, discrimination between positive and negative droplets was impaired (data not shown) and consequently, all subsequent assays used 125 nM of probe per 20  $\mu$ l ddPCR. This effect was observed for both the albumin and HIV-1 assay, suggesting that reducing the probe concentration may aid in the successful transfer of protocols from qPCR to ddPCR.

Other parameters requiring optimisation included primer concentration and Bio-Rad Supermix concentration, with 400 nM of both the forward and reverse primers (compared with 250–500 nM for the qPCR), and 8  $\mu$ l of Supermix per 20  $\mu$ l reaction providing the optimised outcome for both the albumin and HIV assays (data not shown).

Following optimisation of the assays, the albumin and Bio-Rad control assay of 'Alien TLX' DNA (DNA, primers and probes of unknown sequence provided by Bio-Rad as an optimised ddPCR) showed equivalent separation between the negative and positive droplets. However this was reduced markedly for the HIV-1 assay. The difference in mean fluorescence between the positive and negative clusters for the HIV assay was 1951.5 fluorescent units (fU) compared to 4547.3fU for the Bio-Rad control (Fig. 1f). This relative loss of signal impairs the software's ability to demarcate negative and positive droplets, limiting low copy number quantitation.

### 3.2. Optimisation of positive droplet classification

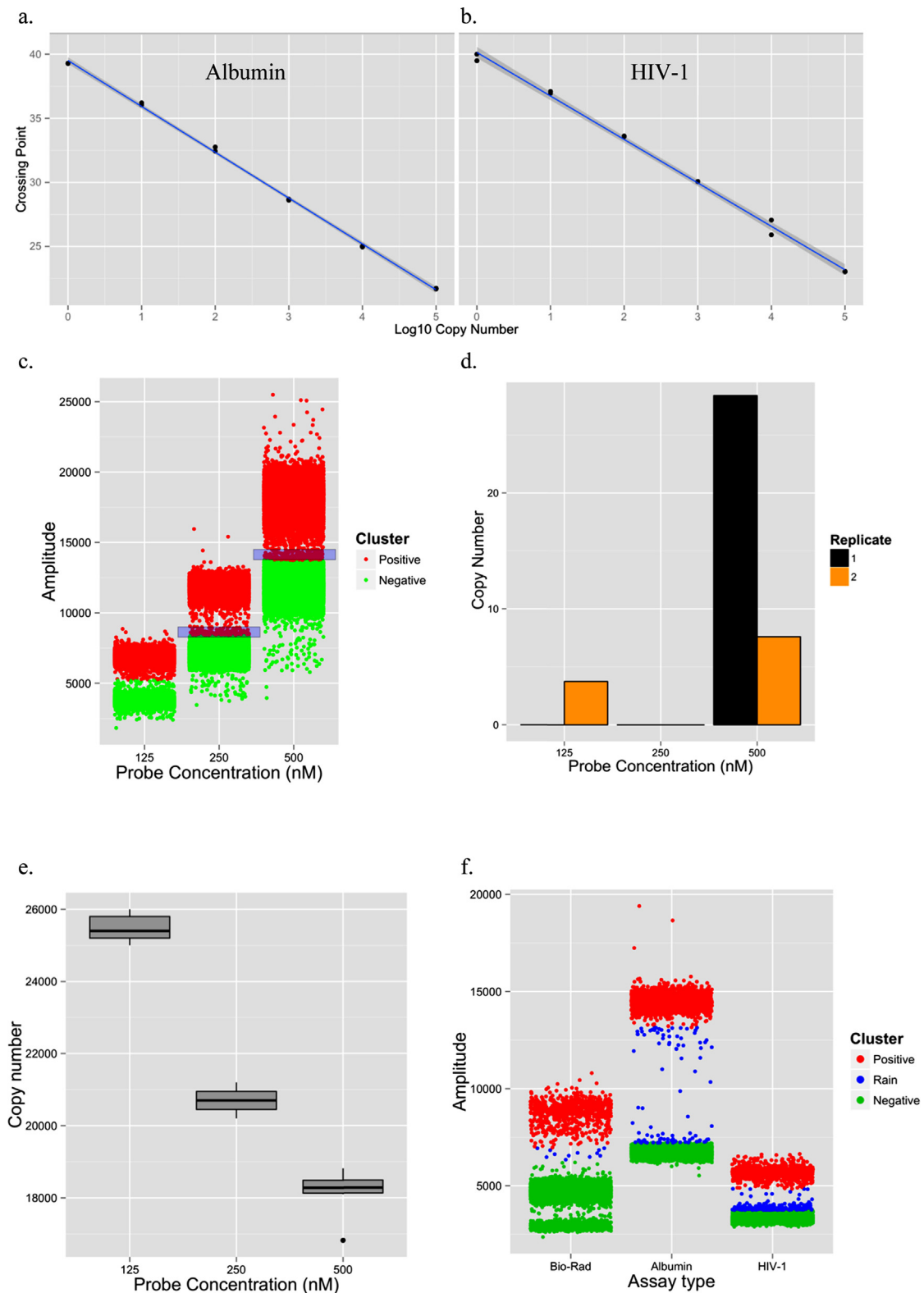
The classification of positive and negative droplets in ddPCR requires the imposition of a cut-off to facilitate the binary read-out. However, not all droplets emit a fluorescent signal that can be classified easily as either positive or negative. In the albumin ddPCR data in Fig. 2, these droplets are positioned graphically in the indeterminate region between the two distinct positive and negative clusters; droplets that lie in this region are colloquially known as 'rain' (Fig. 2 – red dots in panels a and b). Whether these droplets are damaged positive droplets emitting a reduced fluorescent signal, damaged negative droplets emitting an increased background fluorescence or a mixture of the two is not clear, and therefore their allocation to positive or negative values is problematic, especially when dealing with a low copy number target.

The QuantaSoft software (QS) provided with the Bio-Rad ddPCR system, classifies droplets by first determining a fluorescence threshold and subsequently, all droplets with a fluorescence value greater than this threshold are considered positive. To calculate this threshold, every droplet is allocated to either a positive or negative cluster, and a proprietary method is applied to the data to define the fluorescence threshold. For the QS 'Rare Event Detection' setting recommended for low copy number quantitation, a different threshold is set. The result of this method of classifying droplets is a threshold that is set generally much closer to the negative population than the positive population resulting in the majority of rain droplets being classified as positive. For example in Fig. 2, all droplets above the hashed line in the high ( $10^5$  albumin copies) and low ( $10^2$  albumin copies) template input assays (2a and 2b, respectively) would be classified by QuantaSoft as positive, including the red droplets which might be considered as rain. When dealing with standard quantitation assays, the small proportion of positive droplets in the rain region is of negligible consequence as the total number of positive droplets is high. Although the high frequency of positive droplets in these assays results in rain droplets potentially being scored as false positives, this might be considered an acceptable trade-off to optimise sensitivity. When the target DNA is of low copy number however (Fig. 2b), the impact of potential false positive droplets is proportionately greater. Here, if rain droplets are classified as positive, they comprise a relatively higher proportion of total positive droplets than for high copy assays as the total denominator is lower, and would result in falsely high quantitation.

### 3.3. Application of 'definetherain' to improve ddPCR droplet calling

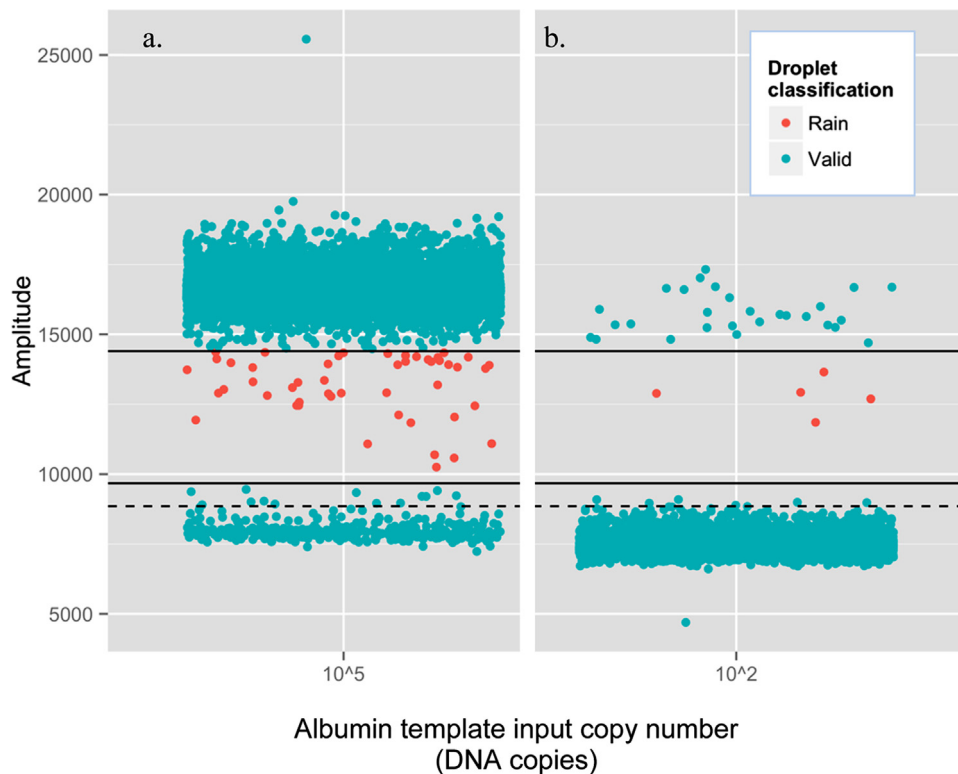
To address this problem a JavaScript program, 'definetherain', that can run within a web browser was devised. 'definetherain' defines indeterminate rain droplets using the positive population rather than the negative population, and a  $k$ -nearest neighbour algorithm. Using droplet amplitude raw data extracted from each well, the software forms two clusters, from which the extreme members are defined. A control is used to define the position of the positive and negative clusters. This control can then be used to define the extreme members of rain from other wells. By excluding the rain from the Poisson calculation more accurate copy number calls can be made in low copy number situations. Fig. 3 shows screen shots from 'definetherain' to illustrate how it is applied. In Fig. 3a, a control sample containing  $10^5$  copies of albumin DNA was analysed to define those droplets which are positive (red) and those which are negative (blue). From the mean and the standard deviation of these two clusters, cutoffs were defined and applied to test samples. Values that lay between these two clusters (highlighted by the shaded grey region) were classified as rain and removed from further analysis. Fig. 3b and c shows how these definitions of positive and negative clusters can be applied subsequently to test samples. Here two templates (low (3b) and high (3c) copy numbers, respectively) were analysed, and the interpretation of positive droplets presented in red. The rain has been removed from the diagram by the software, but would have been positioned in the region between the two clusters.

Using known copy number templates for both albumin and HIV-1 proviral DNA ranging from  $10^5$  to  $10^0$  copies, ddPCR was carried out using both 'definetherain' and QuantaSoft software. Fig. 4 shows data for the albumin (panels a and b) and HIV (panels c and d) ddPCRs. For each target, results are presented as reported by QuantaSoft (panels a and c) and 'definetherain' (panels b and d). For 'definetherain', the distribution of the data points and the associated linear regression of those points (with 95% confidence intervals shown as shaded regions) approximates much better to the 'x=y' black line in the figure, which represents unity between



**Fig. 1.** Optimisation of ddPCR conditions from qPCR reaction protocols. Optimisation of ddPCR conditions from qPCR protocols. (a) Standard curve plot from the albumin qPCR assay. (b) Standard curve plot from the HIV-1 qPCR assay. (c) Dot plot showing fluorescent amplitude of droplets from the HIV-1 DNA assay at differing probe concentrations. Red droplets were classified as positive by QuantaSoft Auto threshold and green droplets were classified as negative. Shaded blue regions indicate areas where QS has identified incorrectly droplets in the negative cluster as positive. (d) Plot showing the HIV-1 copy number in duplicate (orange and black columns) as a surrogate for the frequency of false positives for three probe concentrations assayed using water controls in ddPCR. (e) Box and whiskers plot showing the calculated HIV-1 copy number for differing probe concentrations: 125 nM ( $n=5$ ), 250 nM ( $n=6$ ), 500 nM ( $n=6$ ). Expected copy number was 33,000 for each probe concentration. (f) Dot plot showing a typical droplet distribution for the Bio-Rad test assay, the albumin assay and the HIV-1 assay, with positive, negative and rain droplets indicated. It should be noted that the Bio-Rad test assay is a two channel (FAM and VIC) assay, hence the presence of two distinct negative clusters. (For interpretation of the references to color in this legend, the reader is referred to the web version of the article.)





**Fig. 2.** Defining 'rain' in the ddPCR platform. Results of ddPCR from (a) high ( $10^5$ ) and (b) low ( $10^2$ ) copy number samples assayed in the albumin quantitation assay. Droplets that do not appear to be distinguishable as either positive or negative (rain) are coloured red. The dashed horizontal line indicates the threshold above which Quantasoft determines a droplet to be positive. The full horizontal lines indicate limits that have been applied using 'definetherain' on a positive control sample. Droplets that fall between these cut-offs are defined as rain and are not included in the calculation to determine copy number. (For interpretation of the references to color in this legend, the reader is referred to the web version of the article.)

expected and observed data (correlation coefficients,  $p$  values and gradients of the line are shown for each plot). This is particularly true at lower copy numbers, where 'definetherain' results in much closer approximation to the expected values than QuantaSoft. In a linear regression model combining data from both assays, 'definetherain' has a statistically significant smaller absolute difference between the copy number and the expected value when compared to the QuantaSoft method ( $p = 0.0056$ ) at the lower dilutions (expected value <3000 copies). This relationship is lost when including all the data ( $p = 0.16$ ), supporting the benefit conferred by 'definetherain' at low template concentrations.

#### 4. Discussion

Quantitation of DNA templates by qPCR has become a standard technique in most molecular laboratories. The introduction of ddPCR as a potential improvement in regards to sensitivity, reproducibility and accuracy is of interest, especially if the platform can be proven at low input copy numbers (Hindson et al., 2013). It is proposed that the digital nature of the reaction through the creation of approximately 20,000 individual PCRs, each of which can be scored as positive or negative, should result in improved sensitivity. This would be applicable to a number of research fields, particularly that of HIV-1 cure research, for which a gold-standard molecular assay that accurately represents the HIV-1 latent reservoir is an important objective (Eriksson et al., 2013).

Following the optimisation of two qPCRs (albumin and HIV-1 proviral DNA), methodological data are presented showing the protocol adjustments required to convert these to a ddPCR format. As well as changes in cycling conditions and primer concentration, ddPCR was found to be particularly sensitive to probe concentration for these two assays, which could lead to increased calling of

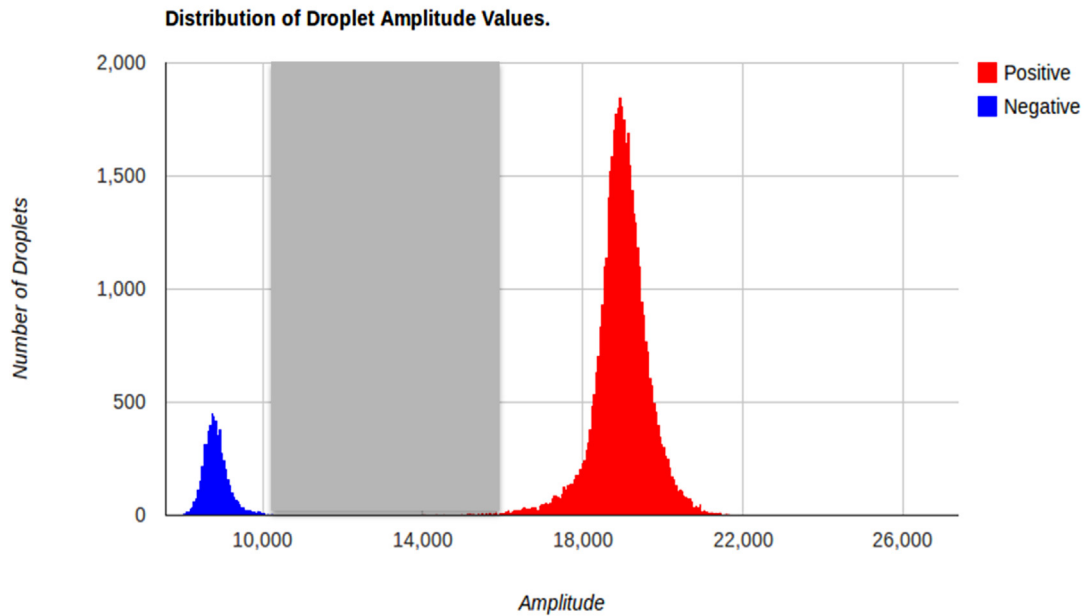
false positive droplets. One explanation for this may lie in the different fluorescence detection systems used in ddPCR and standard qPCR. qPCR instruments detect total fluorescence from a reaction volume of approximately 20  $\mu$ l. In comparison, ddPCR sequentially passes a single 2  $\mu$ l droplet past the detector, therefore it is possible that a reduced amount of fluorescence is lost due to the smaller volume.

The cut-off used to call positive and negative droplets is critical to obtaining results comparable to qPCR, when using ddPCR to quantify known and unknown low copy number DNA targets. Newly developed open source software, 'definetherain', is presented to address this. By defining the cut-off for positive droplets from the positive as opposed to the negative cluster and removing ambiguous rain droplets, the software provides improved accuracy at low template copy numbers. This is particularly relevant to the HIV-1 reservoir where the proviral target may be present in as few as one per million PBMCs.

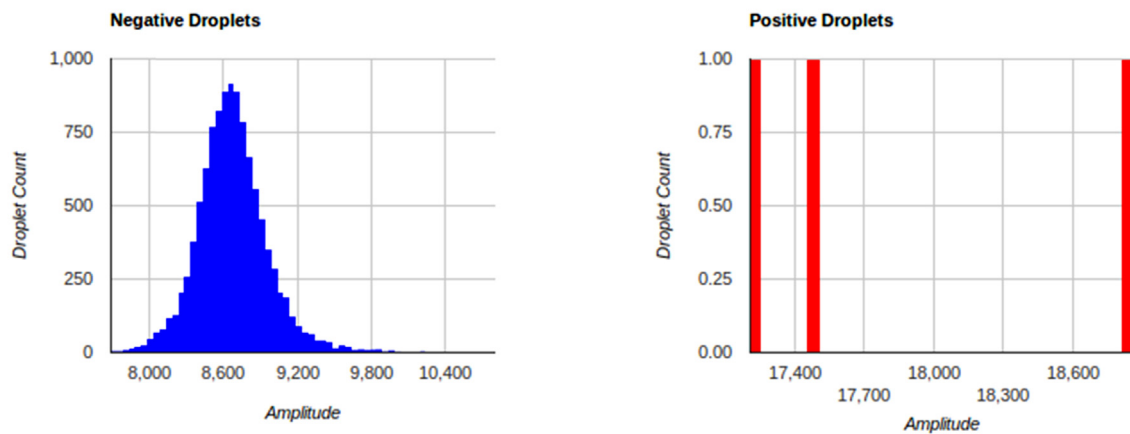
The HIV-1 assay presented here indicates that despite extensive optimisation, certain primer–probe combinations that perform well in qPCR do not provide a sufficient shift in fluorescent amplitude to delineate reliably between positive and negative populations in ddPCR. It is, of course, possible that untested primer–probe combinations would work better in ddPCR (Henrich et al., 2012; Kiselinova et al., 2014; Strain et al., 2013) but it is worth noting that effective reagents for one platform do not transfer automatically to the other.

The albumin ddPCR was reproducible when assaying between 10 and 100,000 template copies. The inter-assay coefficient of variance (CV) calculated for three assays performed independently was 10.8% and 8.1% for 10 and 100,000 template copies, respectively. The CVs for the HIV-1 ddPCR were only <15% for templates >10 copies (data not shown). Low copy number quantification is likely

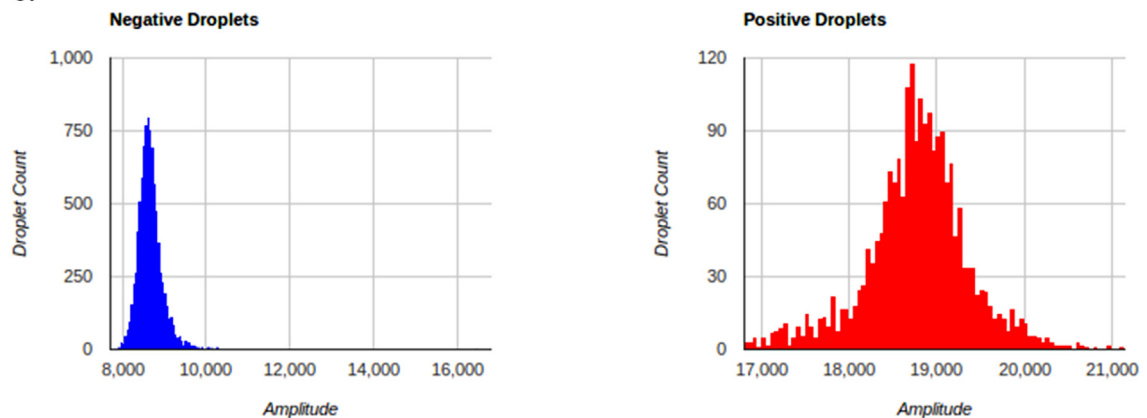
a.



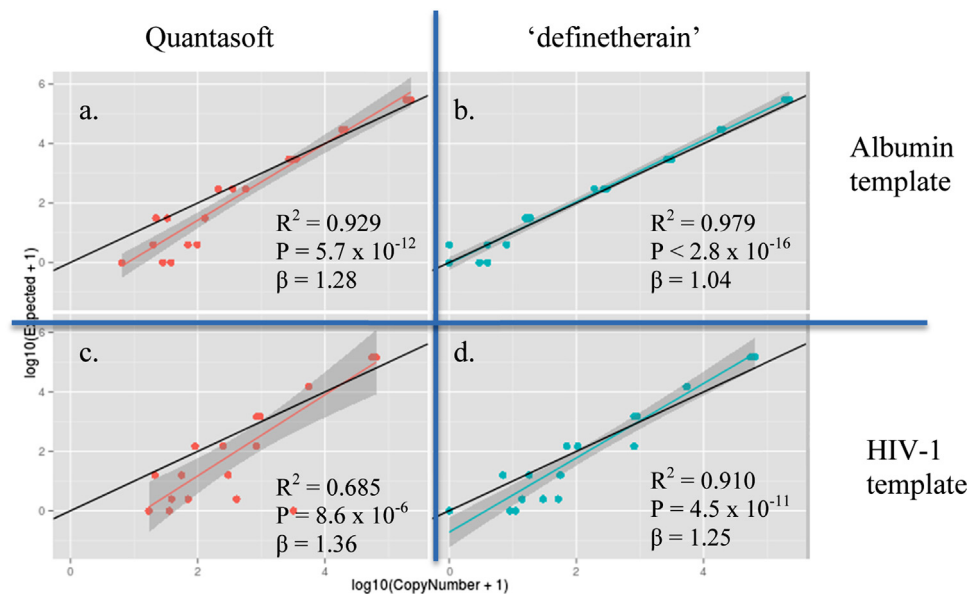
b.



c.



**Fig. 3.** Screenshots illustrating the application of 'definetherain'. Screenshots illustrating the application of 'definetherain'. (a) A modified screenshot from the 'definetherain' javascript program illustrating the  $k$ -nearest neighbour clustering of albumin positive control sample data. The control sample has been used to automatically define the cut-offs for a positive or a negative response. Intermediary values between the positive and negative clusters are classified as rain (the shaded grey region) and are excluded from further analysis. Droplets in any subsequent samples with amplitudes that fall within this range will not be included in the copy number calculation. (b) An example graphical output from 'definetherain' from an experiment sample with a low copy number. (c) An example graphical output from 'definetherain' from an experimental sample with high copy number. (For interpretation of the references to color in the text, the reader is referred to the web version of the article.)



**Fig. 4.** Comparison of ‘definetherain’ and QuantaSoft to report albumin and proviral HIV-1 DNA quantification by ddPCR. Comparison of ‘definetherain’ and QuantaSoft to report albumin and proviral HIV-1 DNA quantitation by ddPCR. Top row (panels a & b), data from the albumin dilution assay with copy number called by QuantaSoft and ‘definetherain’, respectively. Bottom row (panels c & d), data from the HIV-1 proviral dilution assay with copy number calculated by QuantaSoft and ‘definetherain’. The black line would indicate a 1:1 correspondence between the digital droplet copy number and the expected number. The blue or red line is a linear regression line based on the data points. The shaded area indicates the 95% confidence intervals for the regression. *p* Values, line gradients and correlation coefficients for each fit are shown. (For interpretation of the references to color in this legend, the reader is referred to the web version of the article.)

to be a benefit of ddPCR over qPCR where a marked increase in the coefficient of variation is seen at lower copy numbers. With the current bioinformatic interpretation of intermediate droplets, there are however limitations to ddPCR at the lower limit of detection, compared with conventional qPCR. Using ‘definetherain’ to define positive and negative droplets and remove rain droplets reduces the frequency of false positives at lower input target numbers, thus allowing for a lower limit of detection to be attained from the same data set. In Table 1, the methods and relative advantages and disadvantages of QuantaSoft and ‘definetherain’ are detailed.

5. Conclusion

The ddPCR is an elegant adaptation of the current qPCR format and has the potential to be applied widely. However, investigators should be aware that reaction conditions are not interchangeable between the two formats and – where the target is present at low copy numbers – the implementation of ‘definetherain’ might decrease the number of false positive droplets. This software is available for free at <http://www.definetherain.org.uk> with instructions on how to use it.

Acknowledgement

MJ and JW contributed equally. JH and JF contributed equally.

**Table 1**  
Comparison of methods to define positive droplets in ddPCR.<sup>a</sup>

Method	QuantaSoft	Definetherain
Positive and negative droplet definition.	The method for determining negative and positive droplets is not in the public domain however, it results in the majority of rain droplets being classified as positive.	A control sample's amplitude data is used to form two clusters. The derived cut-offs use the control cluster's mean and standard deviations to define which droplets should and should not be used in the calculations.
Advantages	Automatic system. Produced by the manufacturer	Automatic system. Can call low level variants. Open source and available from <a href="http://definetherain.org.uk">definetherain.org.uk</a> . Will run within Chrome or Firefox web browsers.
Disadvantages	The problem of “rain”. Windows program only, without open-source methodology or bioinformatics	Dependent on the use of a control to define cluster positions. Assumes the control is representative.

<sup>a</sup> The methods (where published in the public domain) and key advantages and disadvantages of both the QuantaSoft and the ‘definetherain’ programs.

References

Eriksson, S., Graf, E.H., Dahl, V., Strain, M.C., Yuki, S.A., Lysenko, E.S., Bosch, R.J., Lai, J., Chioma, S., Emad, F., Abdel-Mohsen, M., Hoh, R., Hecht, F., Hunt, P., Somsouk, M., Wong, J., Johnston, R., Siliciano, R.F., Richman, D.D., O'Doherty, U., Palmer, S., Deeks, S.G., Siliciano, J.D., 2013. Comparative analysis of measures of viral reservoirs in HIV-1 eradication studies. *PLoS Pathog.* 9 (2), e1003174.

Finzi, D., Hermankova, M., Pierson, T., Carruth, L.M., Buck, C., Chaisson, R.E., Quinn, T.C., Chadwick, K., Margolick, J., Brookmeyer, R., Gallant, J., Markowitz, M., Ho, D.D., Richman, D.D., Siliciano, R.F., 1997. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science* 278, 1295–1300.

Hayden, R.T., Gu, Z., Ingersoll, J., Abdul-Ali, D., Shi, L., Pounds, S., Caliendo, A.M., 2013. Comparison of droplet digital PCR to real-time PCR for quantitative detection of cytomegalovirus. *J. Clin. Microbiol.* 51, 540–546.

Henrich, T.J., Gallien, S., Li, J.Z., Pereyra, F., Kuritzkes, D.R., 2012. Low-level detection and quantitation of cellular HIV-1 DNA and 2-LTR circles using droplet digital PCR. *J. Virol. Methods* 186, 68–72.

Hindson, B.J., Ness, K.D., Masquelier, D.A., Belgrader, P., Heredia, N.J., Makarewicz, A.J., Bright, I.J., Lucero, M.Y., Hiddessen, A.L., Legler, T.C., Kitano, T.K., Hodel, M.R., Petersen, J.F., Wyatt, P.W., Steenblock, E.R., Shah, P.H., Bousse, L.J., Troup, C.B., Mellen, J.C., Wittmann, D.K., Erndt, N.G., Cauley, T.H., Koehler, R.T., So, A.P., Dube, S., Rose, K.A., Montesclaros, L., Wang, S., Stumbo, D.P., Hodges, S.P., Romine, S., Milanovich, F.P., White, H.E., Regan, J.F., Karlin-Neumann, G.A., Hindson, C.M., Saxonov, S., Colston, B.W., 2011. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. *Anal. Chem.* 83, 8604–8610.

Hindson, C.M., Chevillet, J.R., Briggs, H.A., Gallichotte, E.N., Ruf, I.K., Hindson, B.J., Vessella, R.L., Tewari, M., 2013. Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nat. Methods* 10, 1003–1005.

- Kiselinova, M., Pasternak, A.O., De Spiegelaere, W., Vogelaers, D., Berkhout, B., Vandekerckhove, L., 2014. Comparison of droplet digital PCR and seminested real-time PCR for quantification of cell-associated HIV-1 RNA. *PLoS ONE* 9, e85999.
- Miotke, L.K., Lau, B.T., Rumma, R.T., Ji, H.P., 2014. High sensitivity detection and quantitation of DNA copy number and single nucleotide variants with single color droplet digital PCR. *Anal. Chem.* 86 (5), 2618–2624.
- Pinheiro, L.B., Coleman, V.A., Hindson, C.M., Herrmann, J., Hindson, B.J., Bhat, S., Emslie, K.R., 2012. Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Anal. Chem.* 84, 1003–1011.
- Schachter, J., 2013. Will droplet digital PCR become the test of choice for detecting and quantifying ocular *Chlamydia trachomatis* infection? Maybe not. *Expert Rev. Mol. Diagn.* 13, 789–792.
- Strain, M.C., Lada, S.M., Luong, T., Rought, S.E., Gianella, S., Terry, V.H., Spina, C.A., Woelk, C.H., Richman, D.D., 2013. Highly precise measurement of HIV DNA by droplet digital PCR. *PLoS ONE* 8, e55943.
- Wong, J.K., Hezareh, M., Gunthard, H.F., Havlir, D.V., Ignacio, C.C., Spina, C.A., Richman, D.D., 1997. Recovery of replication-competent HIV despite prolonged suppression of plasma viremia. *Science* 278, 1291–1295.