Clinical Chemistry 69:9
976–990 (2023)

Review

Digital PCR Partition Classification

Matthijs Vynck,[a,b,†] Yao Chen,[a,b,c,†] David Gleerup,[a,b] Jo Vandesompele,[a,d,e,f,g] Wim Trypsteen,[a,d,e,f] Antoon Lievens,[a,h] Olivier Thas,[a,c,i,j] and Ward De Spiegelaere[a,b,*]

BACKGROUND: Partition classification is a critical step in the digital PCR data analysis pipeline. A range of partition classification methods have been developed, many motivated by specific experimental setups. An overview of these partition classification methods is lacking and their comparative properties are often unclear, likely impacting the proper application of these methods.

CONTENT: This review provides a summary of all available digital PCR partition classification approaches and the challenges they aim to overcome, serving as a guide for the digital PCR practitioner wishing to apply them. We additionally discuss strengths and weaknesses of these methods, which can further guide practitioners in vigilant application of these existing methods. This review provides method developers with ideas for improving methods or designing new ones. The latter is further stimulated by our identification and discussion of application gaps in the literature, for which there are currently no or few methods available.

SUMMARY: This review provides an overview of digital PCR partition classification methods, their properties, and potential applications. Ideas for further advances are presented and may bolster method development.

Background

The polymerase chain reaction (PCR), invented in the 1980s (1), and the subsequent development of real-time quantitative PCR (qPCR) in the 1990s (2) have had a transformative impact on in vitro biomolecular studies. These techniques have become a cornerstone in laboratories studying organisms from all of the branches of the tree of life (3–8).

Digital PCR (dPCR) is a method allowing absolute quantification of target nucleic acids, with foundations dating back to the 1980s (9). It has been heralded for its absolute quantification without the need for a calibration curve and its high and scalable precision, but it has several other advantages such as a higher resistance to suboptimal amplification efficiency, for example due to PCR inhibitors (10). On the other hand, some limitations of dPCR are its higher costs and lower throughput capabilities when compared to qPCR. It has a lower dynamic range and is not always more sensitive than qPCR. Because of this, qPCR is still the workhorse methodology for many applications (11). Yet, motivated by its potential advantages (12), dPCR has been implemented across many different domains. Examples include genetically modified organism detection (6); molecular pathology and hematology for oncology, such as liquid biopsy analyses (13, 14); and clinical and environmental microbiology, such as viral load monitoring (11, 12). The recent introduction of instruments allowing up to 6 color measurements, facilitating simultaneous analysis of a larger number of target nucleic acids and an increased sample throughput, will likely further increase dPCR's adoption (11).

Another often mentioned advantage of dPCR is its ease of reaction readout. Unlike qPCR, which requires real-time reaction readout for quantification, quantification in dPCR mostly relies only on an end-point fluorescence detection. In a dPCR experiment, the sample and PCR assay constituents are distributed across many partitions (e.g., droplets, microplates containing capillaries or channels), creating isolated reactions. The end readout captures the individual partitions' fluorescence intensities. In its simplest implementation, these intensities form 2 clusters: partitions containing no target nucleic acid with a low fluorescence intensity and partitions containing the target nucleic acid with a higher fluorescence intensity (Fig. 1). By counting the fraction of partitions with high intensities ("positive partitions") and given knowledge of the partition volumes, the target nucleic acid concentration can be determined (10). Thus, a digital absence-presence signal suffices for absolute quantification.

An essential step in dPCR analysis is the conversion of the partitions' fluorescence intensities to this digital signal. It

[a]Digital PCR Consortium, Ghent University, Ghent, Belgium; [b]Department of Morphology, Imaging, Orthopedics, Rehabilitation and Nutrition, Faculty of Veterinary Medicine, Ghent University, Ghent, Belgium; [c]Department of Applied Mathematics, Computer Science and Statistics, Faculty of Sciences, Ghent University, Ghent, Belgium; [d]OncoRNALab, Cancer Research Institute Ghent, Ghent, Belgium; [e]Department of Biomolecular Medicine, Ghent University, Ghent, Belgium; [f]Center for Medical Genetics, Ghent University, Ghent, Belgium; [g]CellCarta, Zwijnaarde, Belgium; [h]BASF Innovation Center Ghent, Zwijnaarde, Belgium; [i]Data Science Institute, I-BioStat, Hasselt University, Hasselt, Belgium; [j]National Institute for Applied Statistics Research Australia, School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, Australia.
*Address correspondence to this author at: Salisburylaan 133, Entrance 78, Merelbeke 9820, Belgium. E-mail ward.despiegelaere@ugent.be.
[†]These authors contributed equally.
Received January 27, 2023; accepted April 19, 2023.
https://doi.org/10.1093/clinchem/hvad063

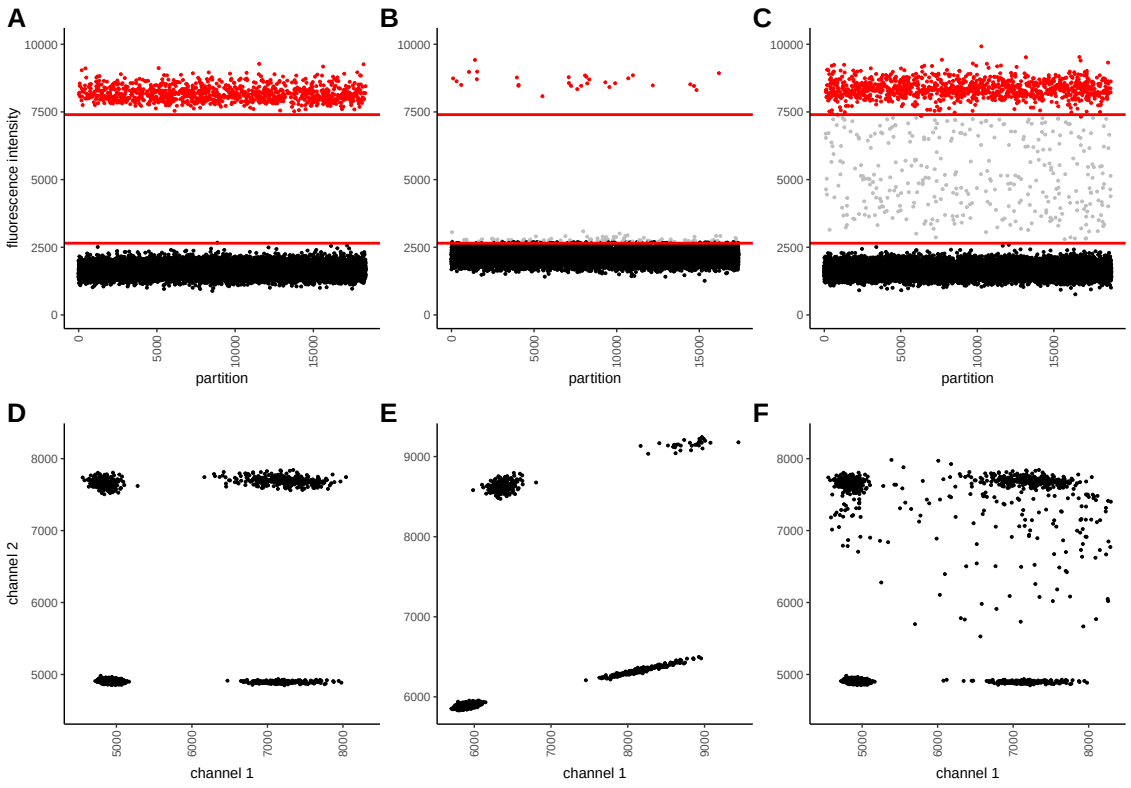

Fig. 1. Examples of typical dPCR intensity measurements (simulated data). (A–C), 1-color plots; (D–F), 2-color plots. Red lines represent relaxed (high) or stringent (low) thresholds. (A), No rain; (B), Rare population, baseline shift compared to (A); (C), Rain and/or low resolution; (D), No rain, no bleed through; (E), Rare population, bleed through; (F) Rain, no bleed through.

is a continuous measure that is a proxy of the presence or absence of the target nucleic acid(s) (Fig. 1). Digitization of the fluorescence intensities is typically achieved by utilizing a fluorescence intensity threshold, so that a partition with a fluorescence readout below this threshold is considered to be devoid of target nucleic acid and a partition with a fluorescence readout above this threshold to contain one or more target nucleic acids (15–18). While often straightforward, partition classification can be complicated by many factors that alter a partition's fluorescence intensity. Such factors include random noise, partition volume variability (19), mismatches in primer/probe sequences (16), low-quality nucleic acids (20), suboptimal amplification efficiency and template accessibility (21), and heterogeneous illumination in instruments that rely on imaging (22). Such undesirable variability in partition intensities is likely multifactorial, and while sample treatment and assay optimization often reduce this variability, a significant proportion of it may remain (17, 23). This often presents in the form of a decreased resolution, or of so-called rain: partitions with intermediate fluorescence intensities not clearly clustering with the clearly negative or positive partitions. Such variability may complicate partition classification (Fig. 1, C and F).

The improper placement of a threshold ultimately results in partition misclassification. Misclassification has been demonstrated to be a major contributor to quantification bias (24). Consequently, an objective and correct partition classification is essential, especially for applications for which minor misclassifications may have a major impact. In addition, there is a threat of investigator-specific bias upon manual thresholding, especially in experiments with substantial amounts of rain or with a low resolution between partition clusters (16) (Fig. 1, C and F). An additional motivation for adopting automated classification methods is that for experiments with up to 6 color measurements and the resulting 6-dimensional data, data visualization and visualization-based manual thresholding becomes cumbersome and laborious, as well as complicated for less experienced users.

Data-driven partition classification procedures have been developed. In this review, we provide a comprehensive overview of such methods as well as a glossary (Table 1) and commentary on proprietary software. Univariate partition classification methods that aim to obtain a partition classification for a single color are described. Next, a description of bivariate partition classification methods for the analysis of experiments using 2 colors is given. We comment on the application of these methods and discuss their underlying assumptions, potential strengths, and weaknesses.

Methods in dPCR Instrument Software

All instrument manufacturers provide proprietary software with their instruments that allow for partition classification. The partition classification software provided by the instrument manufacturers have several advantages: they are often intuitive given their well-designed user interfaces, and they are a one-stop shop to go from the raw intensity measurements to final results. This is in stark contrast with some methods proposed in the literature that do not always provide user interfaces or that perform partition classification only (described later).

Nonetheless, proprietary software comes with its own drawbacks: these proprietary partition classification methods can result in biased, i.e., under- or overestimated, quantities. For example, Bio-Rad's QuantaSoft software, perhaps the most used software currently, requires data with a sufficiently high resolution [e.g., Lievens et al. (17) discuss what can be considered an appropriate resolution] and concentration. With a too-low resolution or a low concentration, results are often not satisfactory (15, 26–28). Indeed, QuantaSoft tends to set the threshold marginally above or even within the negative population (16, 29). Additional problems have been described for assays displaying some or a significant amount of rain. Here, erroneous threshold determination may result in underestimation of the target concentration (30) or high false-positive rates (28).

While some of these issues may have been (partially) resolved since they were pointed out in the literature, details on the methods used within the proprietary software remain undisclosed. This implies that changes or improvements to proprietary software are difficult to verify or benchmark by the user community. Such changes may furthermore impact the replicability of results.

An additional problem is that manual tuning of the automatically determined thresholds is often performed to address erroneous automated partition classification (31). Such manual tuning introduces subjectivity and may lead to a lack of reproducibility due to investigator bias.

Many of these issues can be addressed by adopting independently developed partition classification methods, and some of the aforementioned issues are already addressed by methods discussed in the literature (Tables 2 and 3). Further advantages include peer review of these methods and, often, their open-source availability. Still, the user-friendly sample-to-result graphical interfaces of proprietary software often trump those of the methods and software developed by the research community. Addressing this issue may further incentivize their adoption.

Univariate Methods

Univariate partition classification methods are methods designed to classify 1-color partition intensity measurements. Note that while higher-order multiplexing is possible, none of the univariate methods proposed to date are designed to detect more than 2 partition classes.

definetherain (29) classifies partitions into a negative or positive cluster using k-means clustering (per the method's implementation, though the article mentions a k-nearest neighbor approach) in a 2-step procedure. In a first step, a positive control with distinct negative and positive clusters is required to determine the negative and positive cluster center and the SDs of the intensities in these clusters. In a second step, partitions in other samples are assigned to the negative or positive cluster when the partition's intensity is less than the negative cluster center plus 3 SD or higher than the positive cluster center minus 3 SD, respectively. Partitions with intermediate intensities are classified as rain and omitted from further calculations.

Dreo et al. (15) report on 2 partitioning methods. In the first method, similar to *definetherain*, they also use a mean plus/minus 3 SD approach for delimiting clusters. However, unlike *definetherain*, they require a preliminary threshold for cluster separation, which may be a manual threshold, or a threshold provided by another method, for example a threshold calculated by the instrument manufacturer's software. They suggest a second approach, which they term "manual global threshold," that consists of determining the mean of partition intensities in no-template controls (NTC) or negative controls and add a given number of SDs until not more than one partition has an intensity above this threshold.

ddpcRquant (16) performs classification by determining an upper bound on NTC or negative control partition intensities. First, the data are resampled into groups of a specific size, which are called blocks. The upper bound is determined by fitting a generalized extreme value distribution to the maxima of these blocks. A quantile of this distribution determines the negative

Table 1. Glossary: Overview of frequently used digital PCR terminology, extended from the digital MIQE consensus guidelines (25).

Term	Description	Alternative name
baseline	the fluorescence of the negative partitions	fluorescence noise, background
bleed through	fluorescence of a fluorophore intended to be detected in a given channel that is also detected in one or more other channels	fluorescence bleed through, crosstalk, spectral overlap, spillover
channel	part of the light spectrum used to detect signal, typically annotated as emission wavelength	color, detection channel, emission channel
cluster	group of partitions that display similar fluorescence intensities	(partition) population
duplex or multiplex	assay uses 2 or more primer pairs to amplify multiple target sequences	
fluorescence intensity	the fluorescence of a partition	fluorescence amplitude, end-point fluorescence, relative fluorescence unit
higher-order multiplexing	assays designed to detect more targets than the number of detection channels	intensity multiplexing
negative control	complex biological specimen that does not contain the target	
negative population	partition group that contains no target	negative cluster
no-template control	samples that contain no targets and are used as a general control for extraneous nucleic acid contamination or nonspecific amplification	NTC, blanco, blank sample
partition	the subreaction used for limiting dilution and subsequently measured as positive or negative post reaction	droplet, chamber
positive control	sample that contains target and is used to test if the assay is performing correctly	
positive population	partition group that contains 1 or more targets. In (non-higher-order) multiplex assays, there can be as many as 2^n to 1 positive clusters, where n is the number of targets	positive cluster
rain	the partitions that are located within the space between the positive and negative clusters	
resolution	a measure of the separation in fluorescence between positive and negative partitions	peak resolution, separability score
rotation	cluster principal components (directions of largest variance) may not be parallel to the axes, apparent as a so-called "rotated" cluster	
singleplex	assay used to detect one target sequence of DNA or cDNA	
threshold	the line that separates the partition clusters based on fluorescence intensity	

Table 2. Univariate partition classification methods and their properties.

Method	Methodology	Required input parameters	Output	Sample requirements	Software availability	Note	Reference
definetherain	k-means clustering (2 clusters)	—	• estimate of average number of copies per partition, number of positive, negative, and rain partitions • visualization of results	positive control sample	Javascript-based web application	rain is excluded from calculations	(29)
manual global threshold	mean + *x* SD	—	—	no-template controls or negative controls	—	—	(15)
definetherain adaptation	mean +/– 3 SD	temporary threshold	• concentration estimate • positive, negative, rain clusters upper and lower bounds • number of positive, negative, and rain partitions • quality of cluster separation (Bhattacharyya distance, Jeffries-Matusita distance)	sample with well-discriminated positive and negative partition clusters, manual threshold, or a threshold provided by another method	R	—	(15)
ddpcRquant	extreme value theory	tail probability, resampling parameters	• concentration estimate with CI • number of positive, negative, and total partitions • visualization of results	no-template controls or negative controls	R, Rshiny	performs a baseline correction	(16)
Cloudy	kernel density estimation	—	• concentration estimate with CI • mean number of copies per partition with CI • number of positive, negative, and rain partitions • quality parameters (number of populations identified, resolution,…) • visualization of results	—	R (GitHub), Rshiny	determines reaction-specific thresholds	(17)

Continued

Table 2. (continued)

Method	Methodology	Required input parameters	Output	Sample requirements	Software availability	Note	Reference
Umbrella	nonparametric mixture model	—	• concentration estimate with CI • partition-level class probabilities • number of positive and total partitions • visualization of results	no-template controls or negative controls	R (GitHub)	method includes 2 estimators, performs a baseline correction	(30)
PoDCall	gaussian mixture model	outlier distance, resampling parameters	• concentration estimate • number of positive and total partitions • visualization of results	—	R/Bioconductor, Rshiny	implementation allows threshold determination in 2 channels independently, determines reaction-specific thresholds	(18, 32)
Porco et al.	mean +/− 3 s.d.	negative and positive cluster limits	—	—	—	determines reaction-specific thresholds, performs a baseline correction	(28)

partitions' upper bound: all partitions with an intensity above this bound are classified as positive. This process is repeated several times to obtain a more stable upper bound estimate. *ddpcRquant* additionally performs a baseline correction, that is, a correction for shifts in partition intensities between reactions. This makes use of the robust mode of the negative cluster in each of the samples.

Cloudy (17) relies on kernel density estimation for identification of clusters. The outermost peaks of the density are considered to represent the centers of the negative and positive partition clusters. The method allows for accounting for skewness in the intensities using an empirical formula that provides a good fit to skewness observed in empirical data. *Cloudy* aims to robustly estimate location and scale using the median and median absolute deviation, respectively. Partitions within a certain number of deviations of the median are assigned to the negative or positive cluster, while partitions in between are said to be rain. The fraction of such rain partitions is reported. A threshold is eventually placed at the upper limit of the negative partition cluster.

Umbrella (30) utilizes a model-based clustering approach. The method nonparametrically estimates the density function of the intensities observed in one or multiple NTCs. Like *ddpcRquant* (16), it implements a baseline shift correction for aligning NTCs. The NTC density function, along with an estimate of the density of intensities observed in a sample, can then be used to estimate a mixture distribution, which is a combination of intensities deriving from the negative and positive partitions. This mixture distribution is then used for assigning to each partition a class probability of being positive or negative, based on the partition's intensity. Two concentration estimators are proposed: the first relies on the class probabilities (not imposing a hard threshold), and the second assigns partitions a positive or negative status based on a class probability cutoff (akin to hard thresholding).

PoDCall (18, 32) introduces a gaussian mixture model approach for partition classification. First, a significance test for unimodality (the presence of a single mode) is performed to verify the presence of a single cluster, which is presumed to be the negative cluster. In case the unimodality hypothesis is not rejected, *PoDCall* continues to check for outliers, defined as partitions with intensities larger than the upper quartile plus 9 times the interquartile range. If no outliers are detected, the threshold is placed at the maximum partition intensity observed. If outliers are detected, *PodCall* uses Mclust (33) to identify outliers and sets its threshold at the minimum partition intensity of outliers detected by Mclust or the interquartile approach. In case the unimodality hypothesis is rejected, an equal-variances gaussian mixture model is fitted to a random subset of the partition intensities. The threshold is subsequently placed at the average of the first 2 gaussian components' means.

Motivated by amplification inhibition due to matrix effects in environmental samples, Porco et al. (28) describe a procedure with which they aim to better delineate the positive cluster. Their procedure aims to reduce the impact of rain and "stars" (partitions with very high fluorescence intensities) on concentration estimation. They calculate group-specific upper and lower thresholds to delineate the positive cluster, where the grouping is motivated by the experimental design of the experiment. For example, samples from the same location are assumed to yield similar results and are assigned the same thresholds. These thresholds are obtained by what is conceptually a "mean plus/minus 3 SD" approach but taking into account the difference in locations of the negative and positive cluster at the sample level. This is akin to a within-group baseline correction.

Bivariate Methods

Bivariate partition classification methods are methods designed to classify 2 color partition intensity measurements. As opposed to the univariate partition classification methods described earlier, these methods do not all have the same goal. Analysis of "classical" 2-target, 4-cluster assays is addressed by Strain et al. (26), Chiu et al. (34), Lau et al. (35), and Zhu et al. (36) Attali et al. (27) describe analysis of drop-off assays, while Dobnik et al. (37) and Brink et al. (38) develop methods for higher-order multiplexing analysis.

Strain et al. (26) introduce a "triage" classification method. Their method first removes partitions with atypical intensities (rain, or very low intensities ["hail"]). Large clusters are then identified and modeled using bivariate normal distributions. The probability of each partition to arise from one of these bivariate normal distributions is calculated, and partitions unlikely to belong to any of the distributions labeled as ambiguous or outlier partitions are then disregarded in the subsequent steps. For nonoutlier partitions, independent assortment of targets across partitions in both channels is verified and a violation of the independent assortment assumption is flagged. Details on the identification of large clusters and the verification of independent assortment are not disclosed, and source code is not publicly available.

twoddPCR (34) uses k-means clustering to identify the 4 partition clusters. It combines several reactions to attain a better performance. A refinement of the clusters identified by k-means is achieved by calculating a Mahalanobis distance from the cluster centers: partitions within this distance are considered as part of the cluster, while other partitions are identified as rain and excluded

Table 3. Multivariate partition classification methods and their properties.

Method	Methodology	Input parameters	Output	Requirements	Software availability	Note	Reference
Strain et al.	bivariate normal distributions	preliminary clustering	—	—	—	outliers are excluded from calculations	(26)
ddPCRmulti	—	—	• visualization of results • target concentration	positive control with all the major clusters present	Rshiny (not available anymore)	—	(37)
ddPCR	kernel density estimation, gaussian mixture models	—	• target concentration • visualization of results	—	R (GitHub), R package, Rshiny	outliers and rain are excluded from the analysis, only 3 cluster data (1 negative, 2 positive clusters) are allowed	(27)
twoddPCR	k-means clustering, Ma halanobis distance	maximum distance	• concentration estimate • number of positive, negative, and total partitions • visualization of results	—	R package, Rshiny	—	(34)
Calico	Gridding, k-means clustering, Mahalanobis distance	—	• partition-level class probabilities • classification results • quality scores	positive control	R (GitHub)	—	(35)
ddPCRclust	density estimation, spectral clustering, k-means clustering, gaussian mixture model	—	• number of positive and total partitions • visualization of results	—	R (GitHub), R package, Rshiny	—	(38)
Zhu et al.	gridding, watershed algorithm	number of grids	—	—	—	—	(36)

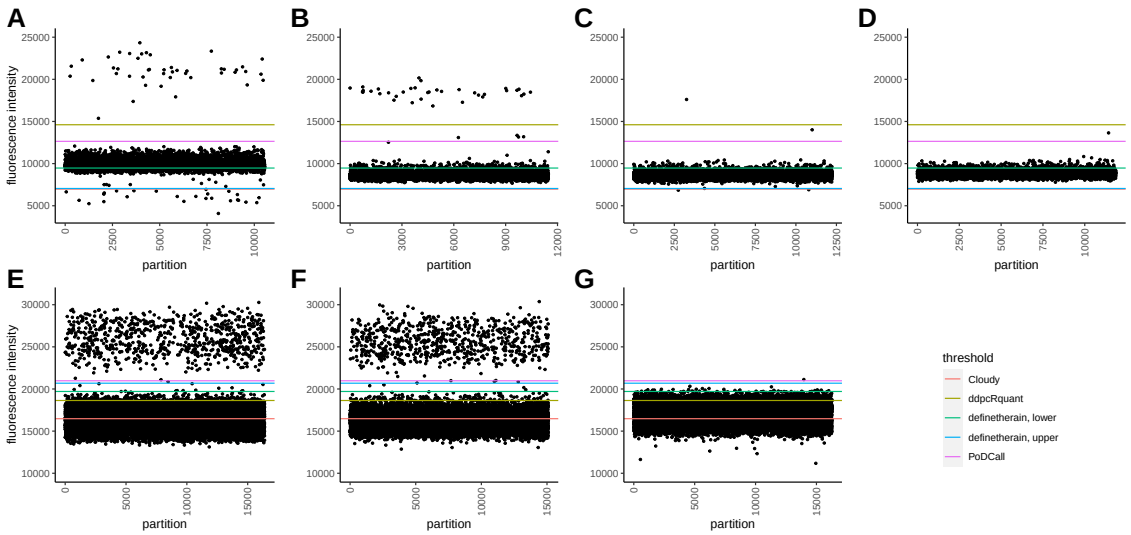

Fig. 2. Thresholds for single-color experiments as determined by *definetherain*, *Cloudy*, *ddpcRquant*, and *PoDCall*. Albumin dataset from Jones et al. (29) (A–D); in-house bovine papillomavirus dataset (E–G). Failure may arise from baseline shifts (A, B), or nonnormality (e.g., multimodality) of the intensity distribution (E–G). See Supplemental Figs. 1 and 2, Supplemental Tables 1 and 2, and the main text for more details.

from further calculations. This refinement takes into account cluster rotations. The Mahalanobis distance should be adjusted by the user to ensure that the ellipses are large enough to include the majority of partitions in a cluster but small enough to exclude ambiguous partitions. *twoddPCR* further introduces clustering by the k-nearest neighbor algorithm, where an initial k-means clustering serves as a training set.

Calico (35) refines a standard k-means clustering approach by implementing sample space gridding. By gridding, differences in density across the 2-color intensity measurement space are reduced. These densities are a proxy for the cluster sizes: by implementing gridding, the k-means algorithm becomes less susceptible to differences in cluster sizes. After obtaining cluster centers of the gridded data, a second round of k-means clustering is performed by initializing the algorithm with the previously obtained cluster centers. *Calico* further refines its partition classification by removing outliers by taking into account the rotation of clusters and by removing partitions with low probabilities of belonging to any cluster.

Zhu et al. (36) propose the use of a density-watershed algorithm for analysis of assays with 2 or 4 expected clusters. First, data are gridded and data densities within each of the grids obtained. These data densities are then used in the watershed algorithm to assign each of the grids to distinct regions. The total distance of observations within these regions to so-called benchmark points is calculated and used to determine the best-fitting number of clusters. After the determination of an optimal cluster number, the regions are selected to represent one of the clusters and the unselected regions are merged to the closest cluster, ultimately resulting in 2 to 4 clusters.

ddPCR (27) is a method for a specific class of single-target genotyping, drop-off assays using 2 color-readouts. An initial quality control step excludes failed reactions and outlying partitions. Next, the negative partition cluster is identified by fitting a 2-component gaussian mixture model to one of the colors. The lower component is assumed to represent negative partitions and the upper component the positive partitions. Assuming a normal distribution of both components, data is assigned a negative, positive, or rain class. The method continues using the positive partitions only. Subsequently, the partition intensities in the second readout color are modeled using kernel density estimation and heuristic optimization, aiming to identify 2 clusters that represent the wildtype and mutant partitions. A final threshold is placed at the local minimum of the density between both clusters.

ddPCRmulti (37) is an interactive data analysis tool for 2-color higher-order multiplexing data, aiming to identify 4 targets giving rise to 16 clusters. Specifically, *ddPCRmulti* automatically calculates positions of 3 thresholds in the horizontal and in the vertical direction each,

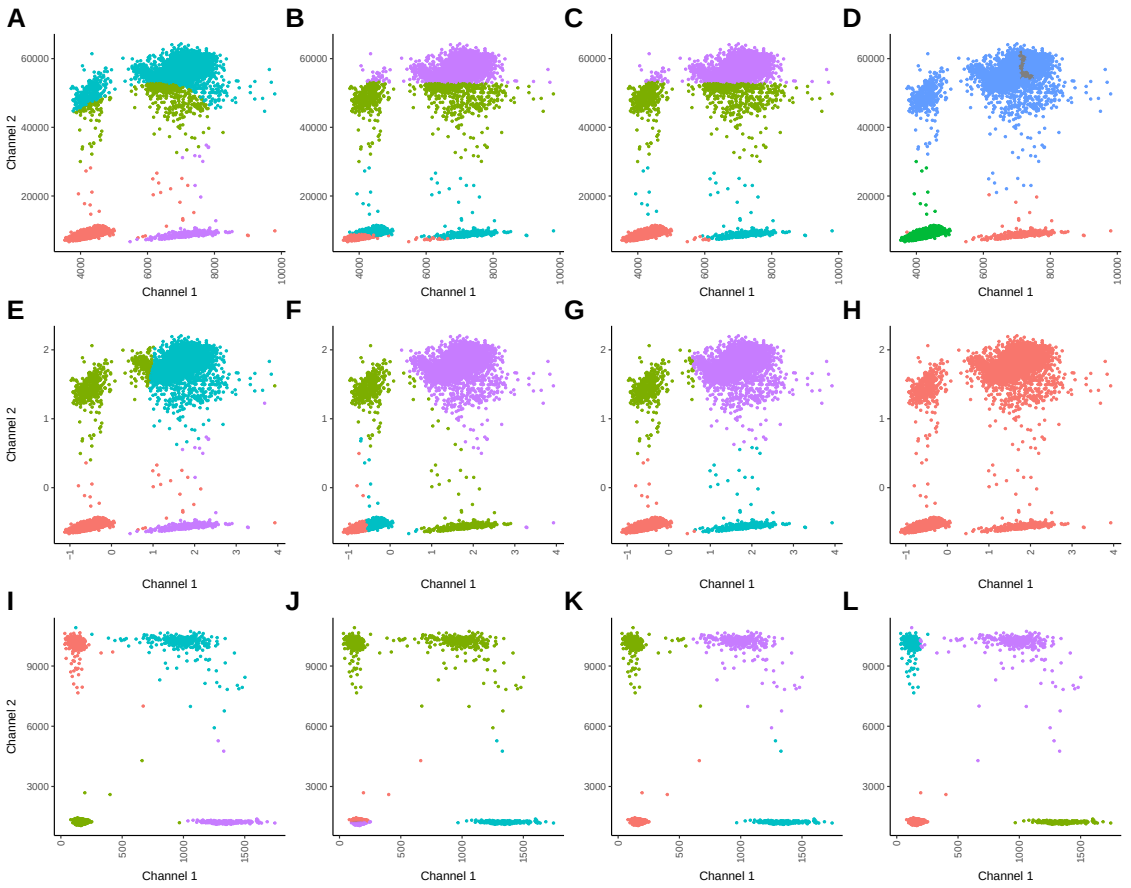

Fig. 3. Clusters for 2-color experiments as determined by *calico* (A, E, I), *twoddpcr* without providing centroids (B, F, J) or with centroids provided (C, G, K), and *ddPCRclust* (D, H, L). In-house HIV intactness dataset (A–D, original measurements assay 1; E–H, scaled measurements assay 1; I–L, assay 2). See Supplemental Fig. 3, Supplemental Table 3, and the main text for more details.

orthogonal to one another. The aim is to set the thresholds in the middle between the clusters. Dobnik et al. (37) acknowledge susceptibility to nonorthogonal cluster positions, uneven cluster size and cluster shape. Thresholds may be tuned interactively in the accompanying web application. The algorithm used for threshold selection is not disclosed, and the source code is not publicly available.

ddPCRclust (38) is designed to deal with higher-order nonorthogonal multiplexed assays. *ddPCRclust* finds the cluster centers by 3 independent methods: flowDensity, SamSPECTRAL, and flowPeaks. For each of the methods, first-order clusters are identified based on the angle between the negative population and the cluster centers. The positions of higher-order clusters are subsequently estimated, taking into account the location of the lower-order clusters. Partitions that are within a specific Mahalanobis distance from the centers are assigned to that cluster, while remaining partitions are assigned to the cluster for which the partitions' Mahalanobis distance is minimal. In a final step, results of the 3 independent approaches are combined by creating a cluster ensemble.

Method Comparison and Assumptions, and Their Implications

A total of 15 distinct partition classification methods have been developed. It is a natural consequence that all these methods have some potential strengths and weaknesses.

First, many of these methods make use of positive, negative, or no-template controls to incorporate prior knowledge for improving partition classification (15, 16, 29, 30, 35, 37). While this can improve classification, relying on control samples introduces the assumption that such samples are representative of the samples analyzed next. For example, many of the early

methods calculate an intensity threshold based on positive or NTC samples. Still, there may be significant baseline variability, for example due to sample- (such as matrix), run-, and lot-specific effects (25). Not taking baseline variability into account may result in biased quantification (Fig. 2, A–D). Another issue is exemplified in (31), where a discrepancy in baseline fluorescence between NTCs and subsequently analyzed samples resulted in biased quantification. The use of a negative control rather than an NTC can in such cases prove useful.

Second, partition classification methods can be divided into parametric and nonparametric methods. Many of the parametric methods assume that intensities follow a normal distribution (15, 18, 26–29, 34, 35). This assumption imposes a constraint on the intensity data distribution, which may be a poor fit and lead to partition misclassification (16). Indeed, partition intensities may show rain or can be ill-separated and the intensity distributions not well described by common distributions (Fig. 2, E and G, and Fig. 3, A and D) (16). Contrarily, nonparametric methods do not impose distributional assumptions on the intensities. Instead, it is often assumed that intensity distributions are similar to negative or positive control samples. For example, *ddpcRquant* (16) determines the upper bound of negative partition intensities from NTCs. *Umbrella* (30) estimates the parameters of a nonparametric mixture model from negative control samples. Such approaches remain sensitive to differences in the intensity distribution between controls and samples, although this also applies to parametric methods relying on control samples.

Another drawback of some of the methods is that they require the selection of tuning parameters such as the number of SDs away from the mean (15, 29, 34, 35) or a distribution's quantile to use as a threshold (16). Similarly, many of the methods rely on k-means clustering (15, 29, 34, 35), a method that leans on proper cluster center initialization and often does not correctly identify the correct clusters when these are of varying sizes and density (Fig. 3, A and B, and Fig. 3, E and F) (39), as encountered frequently in rare event quantification experiments. To improve the performance of k-means, Lau et al. (35) grid data before performing k-means clustering. Again, the results heavily depend on a tuning parameter, that is, the chosen grid size: a too large grid size may not capture the small clusters and a too small grid size will not be very helpful to the follow-up k-means clustering. A third example is encountered in the methods that rely on density estimation (17, 27): improper bandwidth selection may result in the identification of fewer or more clusters than expected due to over- and under-smoothing, respectively, with the risk of placing a threshold within the negative or positive cluster (Fig. 2, E and G).

Yet another divide between methods is in the handling of rain. A minority of methods remove rain (26, 27, 29). Removal may induce a negative bias, as partitions with intermediate intensities often arise because of reduced amplification efficiency or template sequence variability and are likely to contain target sequences (16, 21, 25).

Several methods provide quality control measures for the cluster separation resolution (15, 17) or for identifying discrepancies between control and other samples (30). In addition, some methods are accompanied by interactive graphical user interfaces, allowing easy application of these methods and allowing users to interactively tune parameters and visualize results (16–18, 27, 34, 38). Such quality control measures and visualization capabilities may help in identifying classification failure, while user interfaces allow adoption of the methods by users with no or little programming affinity.

Discussion

Accurate partition classification is a critical step in the dPCR data analysis pipeline. Partition classification may be straightforward in many cases: a manually placed threshold or an automatically calculated one will often give a similar result; i.e., the estimated quantity is agnostic to the chosen (appropriate) partition classification method (Supplemental Figs. 4 and 5, Supplemental Tables 4 and 5). However, in other circumstances, classification may not be straightforward. An example is the quantification of low amounts of target nucleic acid, where a low false-positive rate is essential, such as residual disease estimation using liquid biopsies (13, 14). Also, in the presence of rain, for example due to sequence variability in fast-mutating viruses (Supplemental Fig. 6, Supplemental Table 6) (16), inhibited amplification in environmental samples (11), or when clusters have low resolution (separability), automated partition classification may improve the accuracy of the concentration estimates (16, 17). Other recommended use cases include assays for which subjective manual thresholding is not accepted. Companion diagnostics for use in clinical practice or clinical trials or assay developers wishing to obtain Conformité Européenne-In Vitro Diagnostic Devices Regulation certification may well be obliged to adopt (partly) automated partitioning methods. The same holds true for laboratory-developed tests that need to pass accreditation, for example under the ISO15189, ISO17025, or College of American Pathologists/Clinical Laboratory Improvement Amendments standards, or conducted under good laboratory practices.

Accurate and automated partition classification becomes increasingly important with the advent of higher multiplexing capabilities. With more colors comes increased complexity, as the number of clusters increases exponentially as the number of colors increases linearly.

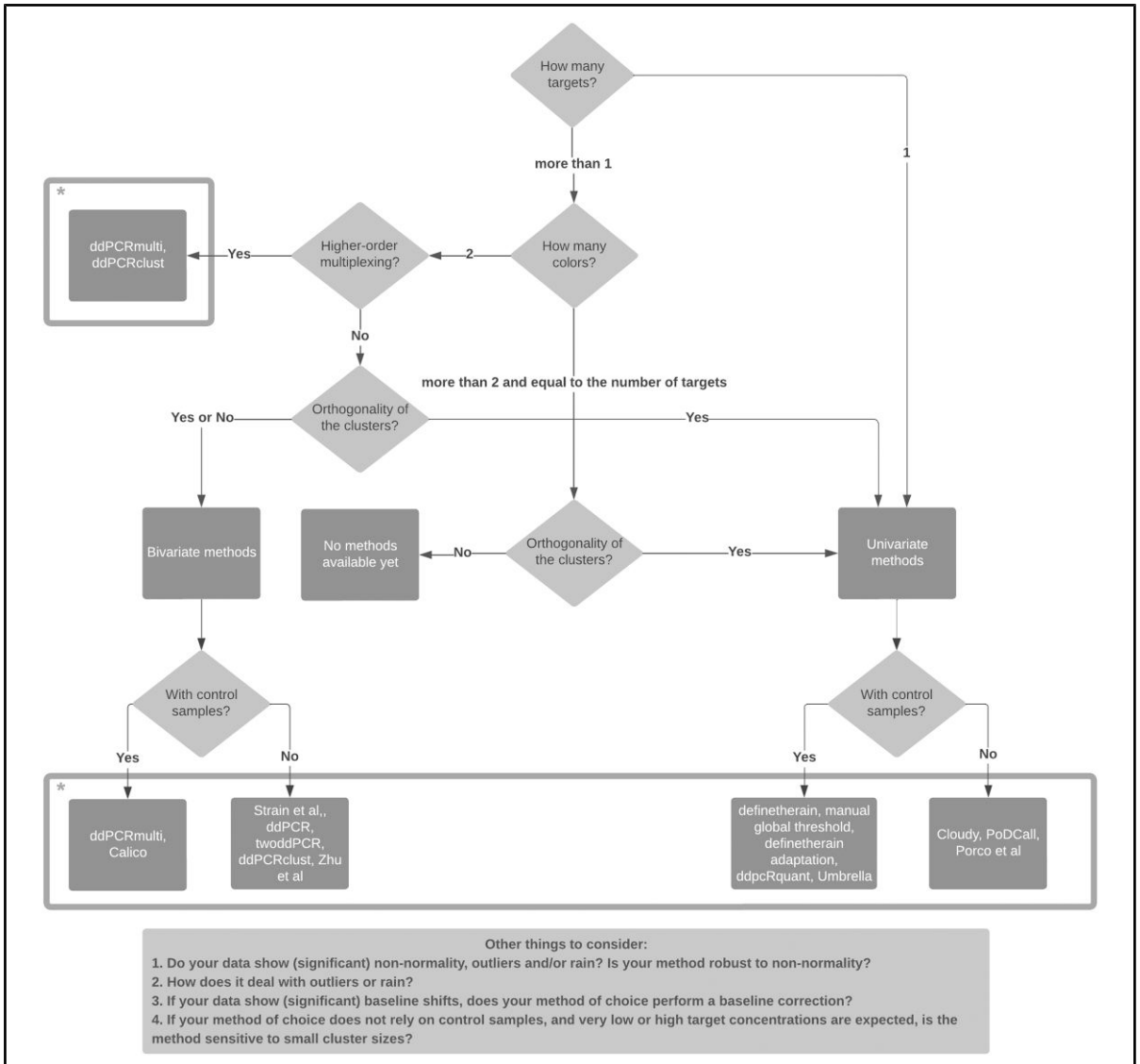

Fig. 4. A guide to method selection and data and method properties to consider before moving on to partition classification.
*** These methods can be used in the given setting but are not necessarily appropriate or optimal (notice the considerations).**

While recent instruments offer up to 6 colors, the number of partitions has not increased substantially. Consequently, the number of partitions in the clusters will, on average, decrease with an increasing number of colors. This will make proper partition classification increasingly complex as smaller clusters are harder to delineate. Further complexity arises in higher-order multiplexing.

Partition classification methods beyond 2 colors have not been discussed in the literature. While for 2 or more colors all methods for univariate analysis can be applied repeatedly to each of the channels individually, some difficulties arise. Due to bleed through of fluorophores, clusters are often not in an orthogonal constellation, necessitating compensation (40). Compensation is often imperfect, and adjustment may be needed (30). Remaining nonorthogonality may result in a multimodal distribution of partition intensities and affect partition classification (Fig. 2, E and G). Some of the methods proposed for 2 colors extend readily to settings with more colors. However, vigilance is needed as performance of both univariate and bivariate methods in settings beyond the one for which they were designed remains unexplored, and some threats

exist. For example, partition cluster sizes will decrease with increasing numbers of colors. This can result in method failure. The exploitation of the expected cluster locations (38) may prove useful to attain a better partition classification performance when moving toward analysis of more colors.

Despite the publication of numerous methods, comparative method evaluation has been limited. One study compared *ddPCR*, *ddPCRclust*, and *twoddpcr* (Table 3), noting that 2 of the 3 methods failed to yield results (20). Kokkoris et al. (31) compared manual classification with the manufacturer's software, *definetherain* and *ddpcRquant* (Table 2) for inhibited environmental samples. They found that *ddpcRquant* fails due to a discrepancy of the fluorescence intensity distribution between NTCs and that of the negative partitions in samples. *definetherain* did not deal with the rain present in their samples. The manufacturer's automatic partitioning failed for samples with nonspecific amplification. These findings may raise questions concerning the applicability of the proposed methods, though the data provided by Colozza-Gama et al. (20) were not in accordance with the *ddPCR* method requirements, and application of *ddpcRquant* on the more representative negative samples instead of NTCs could have provided a better partition classification for Kokkoris et al. (31). This emphasizes the need for a thorough understanding of automated partition classification methods before moving on to their application. Yet, given the different requirements, assumptions, and design considerations of the partition classification methods and the assays on which they are applied, it will be worthwhile to show which methods perform well or fail and under which scenarios they do so. An elaborate method evaluation study seems warranted and should include an assessment of the robustness of these methods against frequently occurring issues such as bleed through, suboptimal resolution, baseline shifts, and so on. This can be achieved with the development of a benchmarking framework, including real and simulated datasets reflecting the levels of difficulty among different applications, along with suitable evaluation criteria (41).

An issue related to evaluation and applicability is that of method reproducibility. While most authors provide access to source code or to web applications (Tables 2 and 3), there has been a lack of appropriate dissemination for some methods: source code is lacking or the method is not described in sufficient detail to allow replication (26, 28, 36), while for others the sole dissemination was a web application that is no longer available (37). Such reproducibility or availability issues are a significant barrier to the adoption of methods.

So far, there has been a reasonably limited uptake of third-party automated partition classification methods. Most authors use the automated partitioning algorithms available in the software provided by the manufacturer or threshold manually (31). Yet, automated partitioning in commercial software has been shown to yield questionable quantification in some cases and to fail in others (16, 30, 36). In addition, with higher multiplexing and higher throughput instruments now available, manual thresholding is becoming increasingly cumbersome.

A final note: a majority of currently available commercial dPCR instruments rely on end-point fluorescence detection for quantification (11, 42). Real-time dPCR instruments have also been described (11, 42). Data analysis strategies different from those used for endpoint dPCR instruments may prove useful (43, 44). We did not consider specific data analysis or data quality filtering procedures for real-time dPCR to be within the scope of this review. The methods described in this review can still be applied to real-time dPCR data, considering the end point fluorescence only.

Conclusion

While many methods have been proposed for a variety of often specific applications, many are potentially applicable to a wider set of applications. Consequently, there often is a choice between several partition classification methods for a given experiment. Comparative method performance is lacking and complicating the selection of an appropriate method. To aid in vigilant method application, Fig. 4 provides an overview of method aspects and potential issues to consider. Some preliminary recommendations can be made based on theoretical grounds. Quantification of low concentrations may be best approached with methods modeling the negative partitions, such as *Umbrella* or *ddpcRquant*, which will not be susceptible to changes in the fluorescence intensities of the positive partitions. This will also be true for assays or samples that show a substantial amount of rain (Fig. 2, Supplemental Fig. 6). In such cases, methods that exclude rain (*definetherain*, *ddpcr*) may be biased and should likely not be used. *Umbrella* is likely a good candidate for analysis of severely inhibited samples, as it allows to model the rain, even when it overlaps with the negative cluster. Care should be taken that controls mimic the actual samples as closely as possible. Still, baseline shifts may occur, and methods that (*a*) do not rely on control samples (*Cloudy*, *PoDCall*, *twoddpcr*, *ddPCRclust*) or (*b*) incorporate a baseline correction (*Umbrella*, *ddpcRquant*) may be better suited to such scenarios (Figs. 2 and 3). Methods that do not rely on control samples may set their threshold within a cluster (Figs. 2 and 3). In contrast, when a large variation between samples occurs, control samples will be of little value and methods that do not rely on such samples may provide better results (*PoDCall*, *Cloudy*, *twoddpcr*, *ddPCRclust*). Again, it requires emphasis that these recommendations are purely based on reasoning and have not been validated by comprehensive, numerical method evaluation. Additionally,

manual thresholding can yield reproducible results when there is a limited amount of baseline shift, rain, and a clear separation of the positive and negative cluster.

Supplemental Material

Supplemental material is available at *Clinical Chemistry* online.

Nonstandard Abbreviations: PCR, polymerase chain reaction; qPCR, quantitative polymerase chain reaction; dPCR, digital polymerase chain reaction; NTC, no-template control.

Author Contributions: *The corresponding author takes full responsibility that all authors on this publication have met the following required criteria of eligibility for authorship: (a) significant contributions to the conception and design, acquisition of data, or analysis and interpretation of data; (b) drafting or revising the article for intellectual content; (c) final approval of the published article; and (d) agreement to be accountable for all aspects of the article thus ensuring that questions related to the accuracy or integrity of any part of the article are appropriately investigated and resolved. Nobody who qualifies for authorship has been omitted from the list.*

Matthijs Vynck (Conceptualization-Equal, Formal analysis-Equal, Software-Equal, Visualization-Equal, Writing—original draft-Equal, Writing—review & editing-Equal), Yao Chen (Conceptualization-Equal, Formal analysis-Equal, Software-Equal, Visualization-Equal, Writing—original draft-Equal, Writing—review & editing-Equal), David Gleerup (Writing—review & editing-Equal), Jo Vandesomple (Writing—review & editing-Equal), Wim Trypsteen (Writing—review & editing-Equal), Antoon Lievens (Writing—review & editing-Equal), Olivier Thas (Conceptualization-Equal, Funding acquisition-Equal, Writing—review & editing-Equal), and Ward De Spiegelaere (Conceptualization-Equal, Funding acquisition-Equal, Visualization-Equal, Writing—review & editing-Equal)

Authors' Disclosures or Potential Conflicts of Interest: *Upon manuscript submission, all authors completed the author disclosure form. Disclosures and/or potential conflicts of interest*:

Employment or Leadership: J. Vandesompele, cofounder of pxlence, which develops dPCR assays.
Consultant or Advisory Role: J. Vandesompele, CellCarta, which provides dPCR services, and scientific advisory board dPCR of Qiagen; W. Trypsteen, digital PCR advisory board member for Qiagen.
Stock Ownership: J. Vandesompele, CellCarta and pxlence.
Honoraria: W. Trypsteen, Qiagen.
Research Funding: O. Thas and W. De Spiegelaere obtained grant 01IO0420 from Ghent University's Special Research Fund (BOF) on which Y. Chen and D. Gleerup are enrolled. O. Thas and W. De Spiegelaere obtained research funding from Stilla Technologies on which M. Vynck is enrolled. W. De Spiegelaere, Flanders Innovation & entrepreneurship—VLAIO, free loan of dPCR equipment for research from Stilla Technologies, Optolane, Thermo Fisher.
Expert Testimony: None declared.
Patents: None declared.
Other Remuneration: W. Trypsteen, support for attending meetings and/or travel from Qiagen.

References

1. Mullis K, Faloona F, 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;51:263–73.
2. Higuchi R, Dollinger G, Walsh PS, Griffith R. Simultaneous amplification and detection of specific DNA sequences. Biotechnology (NY) 1992;10:413–7.
3. Reed JL, Walker ZJ, Basu D, Allen V, Nicol MP, Kelso DM, McFall SM. Highly sensitive sequence specific qPCR detection of Mycobacterium tuberculosis complex in respiratory specimens. Tuberculosis 2016; 101:114–24.
4. Vogels CB, Brito AF, Wyllie AL, Fauver JR, Ott IM, Kalinich CC, et al. Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT–qPCR primer–probe sets. Nat Microbiol 2020;5:1299–305.
5. Mittelberger C, Obkircher L, Oberkofler V, Ianeselli A, Kerschbamer C, Gallmetzer A, et al. Development of a universal endogenous qPCR control for eukaryotic DNA samples. Plant Methods 2020;16:53.
6. Querci M, Van den Bulcke M, Žel J, Van den Eede G, Broll H. New approaches in GMO detection. Anal Bioanal Chem 2010;396:1991–2002.
7. van der Velden VH, Hochhaus A, Cazzaniga G, Szczepanski T, Gabert J, van Dongen JJ. Detection of minimal residual disease in hematologic malignancies by real-time quantitative PCR: principles, approaches, and laboratory aspects. Leukemia 2003; 17:1013–34.
8. van Eijk R, Licht J, Schrumpf M, Talebian Yazdi M, Ruano D, Forte GI, et al. Rapid KRAS, EGFR, BRAF and PIK3CA mutation analysis of fine needle aspirates from non-small-cell lung cancer using allele-specific qPCR. PLoS One 2011;6:e17791.
9. Morley AA. Digital PCR: A brief history. Biomol Detect Quantif 2014;1:1–2.
10. Salipante SJ, Jerome KR. Digital PCR—an emerging technology with broad applications in microbiology. Clin Chem 2020;66: 117–23.
11. Tiwari A, Ahmed W, Oikarinen S, Scherchan SP, Heikinheimo A, Jiang G, et al. Application of digital PCR for public health-related water quality monitoring. Sci Total Environ 2022;837:155663.
12. Hindson BJ, Ness KD, Masquelier DA, Belgrader P, Heredia NJ, Makarewicz AJ, et al. High-throughput droplet digital PCR system for absolute quantitation of DNA copy number. Anal Chem 2011;83:8604–10.
13. Olmedillas-López S, Olivera-Salazar R, García-Arranz M, García-Olmo D. Current and emerging applications of droplet digital PCR in oncology: an updated review. Mol Diagn Ther 2022;26:61–87.
14. Coccaro N, Tota G, Anelli L, Zagaria A, Specchia G, Albano F. Digital PCR: a reliable tool for analyzing and monitoring hematologic malignancies. Int J Mol Sci 2020;21:3141.
15. Dreo T, Pirc M, Ramšak Ž, Pavšič J, Milavec M, Žel J, Gruden K. Optimising droplet digital PCR analysis approaches for detection and quantification of bacteria: a case study of fire blight and potato brown rot. Anal Bioanal Chem 2014;406:6513–28.
16. Trypsteen W, Vynck M, De Neve J, Bonczkowski P, Kiselinova M, Malatinkova E, et al. Ddpcrquant: threshold determination for single channel droplet digital PCR experiments. Anal Bioanal Chem 2015;407:5827–34.
17. Lievens A, Jacchia S, Kagkli D, Savini C, Querci M. Measuring digital PCR quality: performance parameters and their optimization. PLoS One 2016;11:e0153317.
18. Pharo HD, Andresen K, Berg KCG, Lothe RA, Jeanmougin M, Lind GE. A robust internal control for high-precision DNA methylation analyses by droplet digital PCR. Clin Epigenet 2018;10:24.
19. Vynck M, Thas O. Reducing bias in digital PCR quantification experiments: the importance of appropriately modeling volume variability. Anal Chem 2018;90:6540–47.
20. Colozza-Gama GA, Callegari F, Bešič N, Paniza ACJ, Cerutti JM. Machine learning algorithm improved automated droplet classification of ddPCR for detection of

BRAF V600E in paraffin-embedded samples. Sci Rep 2021;11:12648.
21. Berden P, Wiederkehr RS, Lagae L, Michiels J, Stakenborg T, Fauvart M, Van Roy W. Amplification efficiency and template accessibility as distinct causes of rain in digital PCR: Monte Carlo modeling and experimental validation. Anal Chem 2022;94:15781–9.
22. Zhang H, Yan Z, Wang X, Gaňová M, Korabečná M, Zahradník P, et al. Digital PCR system development accelerator—a methodology to emulate dPCR results. Sens Actuators B Chem 2022;358:131527.
23. Witte AK, Mester P, Fister S, Witte M, Schoder D, Rossmanith P. A systematic investigation of parameters influencing droplet rain in the Listeria monocytogenes prfA assay—reduction of ambiguous results in ddPCR. PLoS One 2016;11:e0168179.
24. Jacobs BK, Goetghebeur E, Clement L. Impact of variance components on reliability of absolute quantification using digital PCR. BMC Bioinformatics 2014;15:283.
25. dMIQE Group, Huggett JF. The digital MIQE guidelines update: Minimum information for publication of quantitative digital PCR experiments for 2020. Clin Chem 2020;66:1012–29.
26. Strain MC, Lada SM, Luong T, Rought SE, Gianella S, Terry VH, et al. Highly precise measurement of HIV DNA by droplet digital PCR. PLoS One 2013;8:e55943.
27. Attali D, Bidshahri R, Haynes C, Bryan J. Ddpcr: an R package and web application for analysis of droplet digital PCR data. F1000Res 2016;5:1411.
28. Porco D, Hermant S, Purnomo CA, Horn M, Marson G, Colling G. Getting rid of "rain" and "stars": mitigating inhibition effects on ddPCR data analysis, the case study of the invasive crayfish Pacifastacus leniusculus in the streams of Luxembourg. PLoS One 2022;17:e0275363.
29. Jones M, Williams J, Gärtner K, Phillips R, Hurst J, Frater J. Low copy target detection by droplet digital PCR through application of a novel open access bioinformatic pipeline, "definetherain." J Virol Methods 2014;202:46–53.
30. Jacobs BKM, Goetghebeur E, Vandesompele J, De Ganck A, Nijs N, Beckers A, et al. Model-based classification for digital PCR: your umbrella for rain. Anal Chem 2017;89:4461–67.
31. Kokkoris V, Vukicevich E, Richards A, Thomsen C, Hart MM. Challenges using droplet digital PCR for environmental samples. Appl Microbiol 2021;1:74–88.
32. Jeanmougin M, Brodal HP, Dietrichson Pharo H, Vedeld HM, Lind GE. PoDCall: positive droplet calling and normalization of droplet digital PCR DNA methylation data. Bioinformatics 2023; 39:btac766.
33. Scrucca L, Fop M, Murphy TB, Raftery AE. Mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. R J 2016;8:289–317.
34. Chiu A, Ayub M, Dive C, Brady G, Miller CJ. Twoddpcr: an R/Bioconductor package and Shiny app for droplet digital PCR analysis. Bioinformatics 2017;33:2743–5.
35. Lau BT, Wood-Bouwens C, Ji HP. Robust multiplexed clustering and denoising of digital PCR assays by data gridding. Anal Chem 2017;89:11913–7.
36. Zhu X, Su S, Fu M, Peng Z, Wang D, Rui X, et al. A density-watershed algorithm (DWA) method for robust, accurate and automatic classification of dual-fluorescence and four-cluster droplet digital PCR data. Analyst 2019;144: 4757–71.
37. Dobnik D, Štebih D, Blejec A, Morisset D, Žel J. Multiplex quantification of four DNA targets in one reaction with Bio-Rad droplet digital PCR system for GMO detection. Sci Rep 2016;6:35451.
38. Brink BG, Meskas J, Brinkman RR. ddPCRclust: an R package and Shiny app for automated analysis of multiplexed ddPCR data. Bioinformatics 2018;34: 2687–9.
39. Fränti P, Sieranoja S. K-means properties on six clustering benchmark datasets. Appl Intell 2018;48:4743–59.
40. Madic J, Zocevic A, Senlis V, Fradet E, Andre B, Muller S, et al. Three-color crystal digital PCR. Biomol Detect Quantif 2016; 10:34–46.
41. Weber LM, Saelens W, Cannoodt R, Soneson C, Hapfelmeier A, Gardner PP, et al. Essential guidelines for computational method benchmarking. Genome Biol 2019;20:125.
42. Tan LL, Loganathan N, Agarwalla S, Yang C, Yuan W, Zeng J, et al. Current commercial dPCR platforms: technology and market review. Crit Rev Biotechnol 2022;43: 433–64.
43. Miglietta L, Xu K, Chhaya P, Kreitmann L, Hill-Cawthorne K, Bolt F, et al. Adaptive filtering framework to remove nonspecific and low-efficiency reactions in multiplex digital PCR based on sigmoidal trends. Anal Chem 2022;94: 14159–68.
44. Yao J, Luo Y, Zhang Z, Li J, Li C, Li C, et al. The development of real-time digital PCR technology using an improved data classification method. Biosens Bioelectron 2022; 199:113873.