Research paper

Three-color crystal digital PCR

J. Madic a, 1, A. Zocevic a, 1, V. Senlis b, E. Fradet c, B. Andre a, S. Muller a, R. Dangla a, M.E. Droniou a, * a Stella Technologies, 1 Mail du Professeur Georges Mathé, 94800 Villejuif, France
b Ecole Polytechnique, Route de Saclay, 91128 Palaiseau, France

A R T I C L E   I N F O

Article history:
Received 9 July 2016
Received in revised form 23 October 2016
Accepted 24 October 2016
Available online xxx

Keywords:
Multiplex digital PCR
Crystal Digital PCR
Naica
Multicolor PCR
EGFR
NSCLC

A B S T R A C T

Digital PCR is an exciting new field for molecular analysis, allowing unprecedented precision in the quantification of nucleic acids, as well as the fine discrimination of rare molecular events in complex samples. We here present a novel technology for digital PCR, Crystal Digital PCR TM, which relies on the use of a single chip to partition samples into 2D droplet arrays, which are then subjected to thermal cycling and finally read using a three-color fluorescence scanning device. This novel technology thus allows three-color multiplexing, which entails a different approach to data analysis. In the present publication, we present this innovative workflow, which is both fast and user-friendly, and discuss associated data analysis issue, such as fluorescence spillover compensation and data representation. Lastly, we also present proof-of-concept of this three-color detection system, using a quadruplex assay for the detection of EGFR mutations L858R, L861Q, and T790M.

© 2016 Published by Elsevier GmbH. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Digital PCR has been emerging in recent years as a powerful new addition to the molecular biologist's toolbox, yielding quantitative information of unparalleled precision about nucleic acids [1–3]. Its applications are diverse, ranging from clinical specialties such as oncology, for example for the genotyping and monitoring of lung cancer, an area with high research activity [4–6]; organ transplantation [7–9], microbiology [10–12], virology [13–15] or non-invasive prenatal testing [16–20]; to environmental studies [21–23], or health and safety monitoring for food and feed products [24].

This technology, developed independently by different groups in the early 90's [25–29], was at first technically challenging: for a single sample, end-point dilutions had to be carried out in dozens of 96 or 384 well plates, making sample partitioning, the distinguishing feature of digital PCR, fastidious and unpractical as a standard lab technique. In recent years however, the availability of commercial digital PCR platform has rendered possible the implementation of digital PCR in standard biology labs, allowing proof-of-concept or even large-scale studies to be performed for an ever increasing range of applications. A more in-depth account of the birth and development of digital PCR is related in the excellent review by A. A. Morley [30].

In many domains, whether in biology or medicine, there is a quest to extract as much information from a given sample, ideally from a single experiment. This need for multiplex analysis can be driven by multiple reasons: rare availability of sample, or an effort to decrease costs of analysis for example. Hence, there is a need in digital PCR for detecting more than one or two targets at once. Currently, most available commercial platforms only offer two-color detection, but have developed a multiplexing strategy which involves manipulating probe and primer concentrations, in order to generate populations of different fluorescence amplitudes for a given detection channel [31,32]. While undoubtedly elegant, this approach, termed amplitude-based multiplexing, may not be entirely reliable in the presence of inhibitors, or of nucleic acids of poor quality. These factors may affect differently the multiple targets [33], leading to smears, displacements or fusions of partitions populations not observed for the controls onto which the assay was calibrated. An alternative multiplexing strategy that ensures robustness of results is to use multiple distinct detection channels, assign one target per detection channel and take advantage of the many fluorophores available for probes and/or primers. Although partition-specific effects such as competition and differential amplification efficiencies described in [34] apply to both amplitude-based multiplexing and multicolor multiplexing, these effects will be more easily identified and accounted for using a multicolor-based approach.

* Corresponding author.
E-mail address: magali.droniou@stilla.fr (M.E. Droniou).
1 Both authors contributed equally to the work presented in this article.

http://dx.doi.org/10.1016/j.bdq.2016.10.002
2214-7355/© 2016 Published by Elsevier GmbH. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
The Naica System™ for Crystal Digital™ PCR was developed with this consideration in mind. This new system relies on a three color detection instrument, allowing the user to perform multicolor multiplex digital PCR. This present paper first presents the Naica System and the Crystal Digital PCR workflow and then discusses specific issues associated with multicolor digital PCR experiments: data representation, spillover compensation and population identification.

2. Presentation of the Naica system and crystal digital PCR workflow

2.1. General description

Currently available commercial digital PCR platforms rely on two distinct approaches. The first to be developed was chamber digital PCR (cdPCR), which relies on 2D arrays of microchambers to partition the sample [35]. Once filled with the PCR mix, the microchambers are thermocycled on a flat-block thermocycler, then imaged using fluorescence to reveal the amplified positive partitions. The second approach is droplet digital PCR (ddPCR™), in which the sample is instead partitioned in a bulk emulsion of microdroplets using platform-specific consumables. The emulsion is transferred to a PCR tube or plate for thermocycling. After PCR amplification, data acquisition is performed in a process analogous to flow cytometry, whereby droplets are fluorescently read one by one as they pass in front of a single laser excitation source [36].

The Naica System used herein performs digital PCR using a hybrid approach named crystal digital PCR, combining the 2D array format of cdPCR and the use of droplet partitions as implemented in ddPCR. First, the samples are partitioned into 2D monolayer arrays of monodisperse droplets, called droplet crystals because of the spontaneous periodic arrangement of the droplets within the monolayer, similar to that of atoms within crystals. These droplet crystals are then thermocycled on a flat-block thermocycler before being transferred onto a fluorescence microscope and imaged to reveal the amplified partitions.

The whole process takes place inside a specifically designed microfluidic chip, the Sapphire chip shown in Fig. 1a, and requires the use of two instruments: i) the Naica Geode, which performs the sample partitioning and the thermal cycling of the droplet crystals, and ii) the Naica Prism3, an automated fluorescence microscope equipped with 3 distinct fluorescence channels.

2.2. Step1: preparing the Sapphire chips for crystal digital PCR

The Sapphire chip is an injection molded plastic microfluidic chip that incorporates 4 identical microfluidic networks, each capable of processing one sample or PCR mix. The main feature of the microfluidic networks is a thin rectangular chamber (2 cm × 1 cm × 120 μm) connected in one corner to 33 droplet production nozzles that function using confinement gradients as described in Dangla et al. [37], and are designed to produce droplets with a diameter of 94 μm, which equates to a mean droplet volume of 0.43 ± 0.03 nL (see Supplemental Method S2 for more details on droplet size assessment). A distribution channel connects the 33 nozzles to the male Luer input port, while a waste channel connects the opposite corner of the chamber to the male Luer outlet port.

The Sapphire chips (Stilla Technologies) are delivered primed with an emulsion oil containing droplet stabilizing surfactants (Stilla Technologies, proprietary mixture), such that the entire microfluidic networks and part of the ports are filled with oil. The inlet and outlet ports of primed chips are sealed with removable Luer caps.

To prepare the Sapphire chips, the caps are removed from the inlet ports and 20 μL of PCR mix are pipetted into each of the 4 inlet ports. Pressure-permeable caps (Stilla Technologies) are then positioned onto the loaded ports.

2.3. Step 2: combined partitioning and PCR amplification using the Naica Geode

The Naica Geode (Stilla Technologies) consists in a pressure chamber atop a flat-block thermocycler. A touch-screen user interface and on-board computer allow control of both the pressure in the chamber and the temperature of the bottom of the chamber.

2.3.1. Mechanisms of droplet generation

First, the prepared Sapphire chips are positioned onto the flat-block thermocycler and the lid of the pressure chamber is closed tight. Then, as the combined partitioning and PCR program is launched, the pressure within the chamber is progressively increased to 1 bar. This overpressure is transmitted to the aqueous PCR mix in the inlet port through the pressure sensitive cap, but not directly transmitted to the outlet port which is sealed with a solid Luer cap. This pressure imbalance thus creates a flow of the PCR mix from the inlet port into the microfluidic network. When the aqueous PCR mix passes through the droplet injection nozzles, it spontaneously breaks down into monodisperse droplets which are then propelled by surface tension into the oil-filled microfluidic chamber. As more and more droplets are produced, they self-arrange into a compact hexagonal 2D monolayer.

At the same time, the flow of PCR mix into the oil-filled chamber forces oil to flow through the waste channel and into the sealed outlet port, compressing the air remaining in the port. The flow stops...
when pressure in the outlet port equilibrates with the overpressure applied in the chamber.

2.3.2. Thermal cycling

The second part of the program, PCR amplification, follows immediately after, according to user-defined thermal cycling program, whether for PCR or RT-PCR. The overpressure is maintained within the chamber throughout the PCR cycles.

After PCR amplification, the overpressure in the Naica Geode is slowly decreased down to atmospheric pressure. As the overpressure is decreased, a pressure imbalance between the inlet and outlet ports drives a back-flow of emulsion oil into the microfluidic network. However, the confinement gradients near the droplet injection nozzles prevent the droplet crystal from flowing out of the chamber. Instead, emulsion oil flows back into the inlet port replacing the PCR mix that was initially deposited in the port.

After completion of the PCR program, the chips are transferred into the Naica Prism3 for fluorescence imaging.

2.4. Step 3: 3-color fluorescence image acquisition and analysis of the droplet crystal

The Naica Prism3 (Stillia Technologies) is an automated fluorescence microscope comprising a CCD camera equipped with a tri-band emission filter (495–520 nm)/560–610 nm/655–720 nm), a motorized x-y-z stage for sample positioning and 3 sets of independently actionable high power LEDs with band pass filters of 415–480 nm (blue channel), 530–550 nm (green channel) and 615–645 nm (red channel) for fluorescence excitation. This 3-band fluorescence spectrum allows the detection of many different sets of 3 fluorophores, as illustrated in Table 1.

The Prism3 instrument acquires a set of 3 high resolution images for each droplet crystals through the transparent bottom surface of the Sapphire chips, with one image for each fluorescence acquisition channel (blue, green and red). High resolution images are obtained by stitching 15 sub-images locally acquired with the use of the CCD camera. Exposure times for each channel can be changed by the user (range: 0–1000 ms) depending on assay requirements. Droplet crystals generated using the Naica system are very stable and can be read for up to two weeks after PCR.

An image analysis algorithm then identifies and characterizes each droplet in the droplet crystal to extract the fluorescence intensity of the droplet in each channel, its size, position, shape and texture. The main steps in the algorithm are background subtraction, local maxima detection and watershed segmentation around identified maxima. Artefacts such as dust particles and droplets of inappropriate size are filtered out using size, shape, texture and position criteria in order to keep only the fluorescence values of non-coalesced droplets for data analysis. The Crystal Digital PCR process yields between 25,000 and 30,000 analyzable droplets, with only a few hundred droplets additionally present in the chamber removed due to our quality control procedure. This equates to a theoretical dynamic range of 5 logs, from 0.2 copies/ul to 20,000 copies/ul. It is also possible to re-analyze images acquired using the Naica system using a customizable configuration file allowing advanced users to change parameters such as droplet size or circularity values.

3. Materials and methods

3.1. DNA

Heterozygote Human genomic DNA containing the *EGFR* mutations L858R, L861Q and T790 M were obtained from Horizon Diagnostics (Cambridge, UK). The Universal Exogenous qPCR Positive Control (IPC) DNA was purchased from Eurogentec (Liège, Belgium).

Genomic DNA from Human peripheral blood mononuclear cells (PBMC) was extracted using QIAamp DNA Blood kit (Qiagen, Hilden, Germany). DNA preparation purity was assessed by absorbance using a Nanodrop spectrophotometer (ThermoFischer, Waltham, MA, USA), prior to quantification by fluorescent method using the Qubit fluorimeter (ThermoFischer, Waltham, MA, USA). Only DNA with absorbance ratios of A260/A280 > 1.8 and A260/A230 > 1.8 was used in this study. All DNA were stored at −20 °C for less than 6 months prior to performing experiments.

3.2. Oligonucleotides

Oligonucleotide primers and hydrolysis probes were synthesized by Eurogentec (Liège, Belgium) in RP-HPLC-grade.

<table>
<thead>
<tr>
<th>Name</th>
<th>Oligo Type</th>
<th>5’Fluorophore</th>
<th>Sequence</th>
<th>3’Quencher</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALB–Cy5P</td>
<td>Hydrolysis probe</td>
<td>Cy5</td>
<td>TGTCTGAAACATTACACCTTCCATGCAG</td>
<td>BHQ-2</td>
<td>[38]</td>
</tr>
<tr>
<td>ALB F4</td>
<td>Primer</td>
<td>–</td>
<td>GGATGAGAAAGCCCAGTAACTG</td>
<td>–</td>
<td>\emptyset</td>
</tr>
<tr>
<td>ALB R2</td>
<td>Primer</td>
<td>–</td>
<td>GCCAACACCTCAACTTTCTCT</td>
<td>–</td>
<td>\emptyset</td>
</tr>
<tr>
<td>BRAF FVII600P</td>
<td>Hydrolysis probe</td>
<td>FAM</td>
<td>CTAAGCTAGA + G + T + GAATTCGCC</td>
<td>BHQ-1</td>
<td>[39]</td>
</tr>
<tr>
<td>BRAF 51F</td>
<td>Primer</td>
<td>–</td>
<td>TACGTGGTTTCTTCACTTACACCTCAG</td>
<td>–</td>
<td>N/A</td>
</tr>
<tr>
<td>BRAF 176R</td>
<td>Primer</td>
<td>–</td>
<td>CCAGAACACTGTCGAACTGAG</td>
<td>–</td>
<td>N/A</td>
</tr>
<tr>
<td>EGFR L858 WT P</td>
<td>Hydrolysis probe</td>
<td>FAM</td>
<td>AGTTGCC + C + CA + GCCCAA</td>
<td>BHQ-1</td>
<td>[6]</td>
</tr>
<tr>
<td>EGFR L858R</td>
<td>Hydrolysis probe</td>
<td>Cy5</td>
<td>AGTTGCC + CC + GCCCAA</td>
<td>BHQ-2</td>
<td>[6]</td>
</tr>
<tr>
<td>EGFR L861Q</td>
<td>Hydrolysis probe</td>
<td>Cy5</td>
<td>ACCGC + C + GTTGGCCCA</td>
<td>–</td>
<td>N/A</td>
</tr>
<tr>
<td>EGFR T790M P</td>
<td>Hydrolysis probe</td>
<td>Cy3</td>
<td>ATGACCT + G + CA + T + CATGAG</td>
<td>BHQ-2</td>
<td>[6]</td>
</tr>
</tbody>
</table>

The ‘+’ sign is a prefix designating a Locked Nucleic Acid (LNA) base.

Please cite this article in press as: J. Madic, et al., Three-color crystal digital PCR, Biomol Detect Quantif (2016), http://dx.doi.org/10.1016/j.bdq.2016.10.002
Table 3

<table>
<thead>
<tr>
<th>Reagents</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>PerfeCta Multiplex qPCR ToughMix</td>
<td>1X</td>
</tr>
<tr>
<td>FITC</td>
<td>40 nM</td>
</tr>
<tr>
<td>BRAF 51F</td>
<td>500 nM</td>
</tr>
<tr>
<td>BRAF 176R</td>
<td>500 nM</td>
</tr>
<tr>
<td>BRAF WT600 P - FAM</td>
<td>250 nM</td>
</tr>
<tr>
<td>ALB F4</td>
<td>250 nM</td>
</tr>
<tr>
<td>ALB R2</td>
<td>250 nM</td>
</tr>
<tr>
<td>ALB P - Cy5</td>
<td>250 nM</td>
</tr>
<tr>
<td>IPC primers and Y labelled probe</td>
<td>3X</td>
</tr>
</tbody>
</table>

Oligonucleotide sequences and associated references are presented in Table 2.

Oligonucleotides without references were designed using the online Primer3 design tool (http://simgene.com/Primer3).

The Universal Exogenous qPCR Positive Control DNA (IPC) was detected using a set of primers and probe purchased from Eurogentec (Catalog Number: RT-IPCY-B02). The IPC-specific probe is labelled with Yakima Yellow fluorophore on its 5’ end, and quenched using TAMRA.

3.3. PCR Mix

PCR reactions were carried out using PerfeCta Multiplex qPCR ToughMix (Quanta Biosciences, Gaithersburg, MD, USA). In order to allow adequate imaging of all droplets for software analysis, FITC (Saint Louis, MO, USA) was added to each reaction to a final concentration of 40 nM. Final concentrations of PCR reagents are detailed in Tables 3 and 4.

3.4. Crystal digital PCR instrument parameters

The Naica Geode was programmed to perform the sample partitioning step, followed by the PCR thermal cycling program: 95 °C for 10 minutes, followed by 45 cycles of 95 °C for 10 seconds and 60 °C for 15 seconds.

Image acquisition was performed using the Naica Prism3 reader using the following exposure times: blue channel: 100 ms; green channel: 50 ms; red channel: 50 ms.

Total droplet enumeration and droplet quality control, enabled by the detection of the reference dye FITC in the blue channel, was performed by the Crystal Reader software. Extracted fluorescence values for each droplet were then further analyzed using the Crystal Miner software (Stilla Technologies, Villejuif, France). Thresholds were set using the automated tool available in the Crystal Miner software.

Table 4

<table>
<thead>
<tr>
<th>Reagents</th>
<th>Final concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>PerfeCta Multiplex qPCR ToughMix</td>
<td>1X</td>
</tr>
<tr>
<td>FITC</td>
<td>40 nM</td>
</tr>
<tr>
<td>EGFR L858F</td>
<td>500 nM</td>
</tr>
<tr>
<td>EGFR L858R</td>
<td>500 nM</td>
</tr>
<tr>
<td>EGFR L858WT P - FAM</td>
<td>250 nM</td>
</tr>
<tr>
<td>EGFR L858R P - Cy5</td>
<td>250 nM</td>
</tr>
<tr>
<td>EGFR L858Q P - Cy5</td>
<td>250 nM</td>
</tr>
<tr>
<td>EGFR T790 M F</td>
<td>1 nM</td>
</tr>
<tr>
<td>EGFR T790 M R</td>
<td>1 nM</td>
</tr>
<tr>
<td>EGFR T790 P - Cy5</td>
<td>250 nM</td>
</tr>
</tbody>
</table>

3.5. Analysis and calculations

3.5.1. Determination of the limit of blank

The limits of blank (LoB) for the EGFR assays were assessed using series of wild-type only controls (see Table S1), and recording the number of false positive events for each of the probes detecting either EGFR L858R, EGFR L861Q or EGFR T790M. A Poisson model was fit to the data using a single parameter, representing the mean of false positive droplets and evaluated for goodness-of-fit using the online calculator developed by H. Arsham at the University of Baltimore (http://home.ubalt.edu/ntsbarsh/Business-stat/otherapplets/PoissonTest.htm). Where data was found to fit a Poisson model, the 95% upper confidence limit of the Poisson distribution was used to calculate the LoB. Where data were found not to fit a Poisson model, the limit of blank was derived from the 95% percentile value.

3.5.2. Determination of target concentration values

In order to calculate accurate concentration values, the limit of blank (number of false positive droplets) was subtracted from the number of positive droplets found for each sample assayed. The concentration in the final reaction mix for each sample (expressed as copies per microliter), as well as the associated uncertainty, were then calculated as previously described [40].

3.5.3. Determination of ratios

The ratios were expressed as a percentage of mutant sequences amongst total DNA present. Since the EGFR L858 probe was used to characterize concentration of wild-type DNA, the ratio for EGFR L858R was calculated as Ratio EGFR L858R = Concentration EGFR L858R *100/(Concentration WT + Concentration EGFR L858R), whereas both ratios for EGFR T790 M and EGFR L861 Q used the following equations Ratio = Concentration Mutant *100/Concentration WT. In order to calculate theoretical ratios and associated uncertainty, the WT and mutant stock concentration were first assessed by digital PCR, then diluted serially to yield a range of ratios from 3% to 0.3%. The uncertainty associated with the ratios was calculated as described by Whale and coworkers [40].

4. Results & discussion

4.1. Visualization of three-color experiments

4.1.1. Quality control of the analytic process

A key feature of the Naica system is the ability to access raw data, i.e. the images of the droplet crystals generated by the Prism3 device, which allows the user to perform a preliminary qualitative assessment of the experiment, and check that the acquisition parameters, such as exposure times, are adequate for each channel (Fig. 2A–C). Moreover, advanced features within the Crystal Miner analysis software also allow users to visualize which objects have been classified as droplets, and have been included into the analysis. This means that the user has access to the entire analytical chain, including steps preceding the representation of extracted fluorescence data on scatterplots.

4.1.2. Graphical display

Scatterplots have proved a very useful medium to display raw fluorescence data from digital PCR experiments. These types of graph enable the instant visual capture of various data features, such as the segregation of partitions into populations of varying fluorescence intensities, the mean fluorescence values and spread of these populations, or the co-encapsulation of detected sequences within partitions. Experimental data can be represented using different types of scatterplots. In a 1D plot, fluorescence values for a given detection
channel, expressed as relative fluorescence units (RFU), are plotted on the vertical axis, whereas the values displayed on the horizontal axis represent droplet index, i.e. the order in which the droplets have been analyzed (Fig. 3A). These plots enable the immediate visualization of the distribution of droplet fluorescence intensities detected for a single channel.

In a 2D dot plot, fluorescence intensity values for one of the three detection channels are plotted onto the vertical axis, while fluorescence intensity values for one of the other 2 detection channels are plotted onto the horizontal axis (Fig. 3B). The use of these 2D scatterplots thus additionally allows the characterization of potential correlations in fluorescence for the different channels. Different
mechanisms underlie these correlations: co-amplification of different sequences in a single partition, which may or may not have biological relevance; cross-reactivity of the probes with different templates; or fluorescence spillover (see Section 4.2).

Until recently, commercially available digital PCR only offered up to two detection channels, so a single 2D dot plot was sufficient to display fluorescence data for a given sample. However, in 3-color experiments, the output for a single channel can be visualized either as one 3D dot plots, three 2D dot plots or with a single 3D dot plot (Fig. 3C). A 3D dot plot is evidently the best way to capture all information at once for the experiment, however it is not an easy format to manipulate for the human eye. On the contrary, 1D and 2D dot plots can be more easily apprehended, and can thus be used for further data analysis, such as threshold setting to discriminate negative from positive partitions.

4.1.3. The RGB additive color system

In a most simple experimental setting, for each detection channel, there is one negative population of droplets, in which either amplification and/or probe hybridization and cleavage did not occur, and one positive population, whose increased level of fluorescence indicates that the aforementioned events did take place. Such binary output means, for n detection channels, droplets can be categorized into 2^n classes. In a two-color experiment, droplets can thus be grouped into 4 classes: double negative, positive for channel 1 and negative for channel 2, negative for channel 1 and positive for channel 2 and finally double positive; whereas three-color experiments imply that droplets can be categorized into 8 classes (Table 5).

In order to enable immediate capture of each droplet status with regard to these 8 classes on scatterplots, an additive color system based on the primary colors Red, Green and Blue (RGB) is used for display (Fig. 4A). This color scheme enables the projection of classifications in all three channels onto one- and two-dimensional scatterplots, thus adding depth by including contributions of parameters not plotted on the axes (Fig. 4B–C). It also allows immediate visualization of the different populations or clusters present for a given sample when displaying data using a three-dimensional scatterplot (Fig. 4D).

4.2. Fluorescence spillover and compensation

Fluorescence experiments using multiple fluorophores are commonly set up to detect a unique fluorophore per acquisition channel. However, some combinations of fluorochrome spectral properties and detection channels specifications may result in more than one fluorophore contributing to the signal recorded for a given acquisition channel, a phenomenon commonly designated as fluorescence spillover (also called fluorescence crosstalk or bleedthrough) [41], [42].

As described in Section 2.4., the Naica Prism3 is equipped with three LED for excitation and a tri-band filter for detection of fluorescence. The chambers are illuminated sequentially from shorter to longer wavelengths (Blue > Green > Red LED), and each image is acquired separately for each detection channel (Blue image, Green image, Red Image). Illumination with a specific LED excites fluorophores whose excitation spectrum overlaps with that of the light passing through the excitation filter. These fluorophores then re-emit light within a range of different wavelengths as fluorescence. Finally, only the light that passes through the tri-band emission filter contributes to the final image for each detection channel (Fig. 5).

4.2.1. Identification of fluorescence spillover

In order to demonstrate the effect of fluorescence spillover on experimental data, a duplex Crystal Digital PCR experiment was designed using two fluorophores with overlapping spectral characteristics, FAM, to be detected using the Blue Detection Channel, and Yakima Yellow, to be detected using the Green Detection Channel.

As detailed in Fig. 6, there is an overlap between FAM and Yakima Yellow excitation spectra, part of which falls within the boundaries of the Blue LED excitation light. Thus, it can be expected that both fluorophores will be excited upon illumination by the Blue LED, while only Yakima Yellow will be excited upon illumination by the Green LED, and none of these fluorophores will be excited by the Red LED (Fig. 6A–C). This therefore would result in fluorescent emissions from both fluorophores contributing to the signal acquired for the Blue Detection channel, whereas only Yakima Yellow fluorescence would be detected for the Green Detection channel, and no fluorescent signal detected for the Red channel (Fig. 6A–C).

This duplex assay was developed using the FAM-labelled hydrolys probe to detect the human BRAF gene, and the Yakima Yellow-labelled probe to hybridize to a universal exogenous qPCR control sequence (IPC). Reactions were optimized using individual templates, and specificity of each detection system was confirmed by assaying the IPC-specific probe and primers onto human genomic DNA, and conversely, the BRAF-specific probe and primers onto the IPC template (data not shown). The duplex probe system was then used to assay a mixture of the human genomic DNA and IPC templates by Crystal Digital PCR.

We observed, as expected, two distinct populations for the Green Detection channel, one of low fluorescence intensity, corresponding to the droplets which did not contain any IPC, and one of higher fluorescence intensity values, corresponding to amplification and detection of the IPC template (Green positive). However, in the Blue Detection channel, the droplets segregated into four different populations of increasing fluorescence intensity, resulting from the fluorescence spillover of the Yakima Yellow probe into the Blue detection channel, as shown in Fig. 7.

The Naica system is uniquely designed to accommodate up to three different fluorophores. In order to demonstrate the full capacities of the Naica system, we have combined the previously described assay, using a FAM-labelled probe targeting BRAF and a
Fig. 4. The RGB additive system.
A. Schematic of the color-code used by the Crystal Miner software to display droplet populations. 1-D (B), 2-D (C) and 3-D (D) scatterplots of droplets. The 8-color mode enables the user to project the distribution of fluorescence in all three channels on 1D and 2D dot plots. For example, the population with the higher fluorescence values in (B) is composed of red (+,−,−); yellow (+,+;−); magenta (+,−,+); and pale grey (+,+,+); whereas the population with lower fluorescence values is composed of blue (−,−,+); green (−,+,−); cyan (−,+,+); and dark grey (−,−,−) dots.

Fig. 5. Light path for fluorescence excitation and emission within the Naica Prism3.
Samples (Droplets) are illuminated sequentially by the Blue, Green and Red LED, and re-emit light in the form of fluorescence. Light emitted passes through tri-band filter to generate the resulting image for the corresponding channel.
Fig. 6. Spectral overlap for FAM and Yakima Yellow (YY) fluorophores and its detection using the Prism3 device. On each graph, wavelength (nm) is plotted on the x axis, and normalized signal intensity on the y axis. A. Both FAM and Yakima Yellow contribute to the image for the Blue detection channel. B. The image for the Green detection channel is based solely on the signal from the Yakima Yellow fluorophore. C. Neither FAM, nor Yakima Yellow are excited by the light emitted by the Red LED, there is no signal for the Red detection channel.

Fig. 7. Identification of additional droplet populations due to fluorescence spillover. Arrows indicate the corresponding populations in the 1D and 2D scatterplots. The suffixes negative and positive designate negative and positive droplet populations respectively for the FAM-labelled probes detecting the BRAF gene and the Yakima Yellow (YY) – labelled probe detecting the exogenous control IPC.
Yakima Yellow–labelled probe targeting IPC, with a Cy5-probe targeting the human albumin (ALB) gene. The specificity of the probe was assessed as previously described. Analysis of the spectral properties of Cy5 predicts that there may be some fluorescence spillover into the Green detection channel (Fig. 8).

This additional fluorescence spillover of the Cy5 fluorophore can indeed be readily observed in the Green channel, leading to the presence of three populations of varying fluorescence in the 1D dot plot for the Green channel. Visualization of this raw data in 2D dot plots is even more confusing since the Green versus Blue 2D dot plots shows the segregation of the lower intensity value droplets into 2 populations, whereas the Blue versus Red 2D dot plot displays a total of 8 populations, due the spillover of the Yakima Yellow fluorophore into the Blue channel (Fig. 9).

This example illustrates the impossibility to analyze raw (or uncompensated) data, since fluorescence spillover, by generating multiple populations of varying intensity prevents classification of the droplets into either negative or positive.

4.2.2. Compensation for fluorescence spillover

It is therefore crucial for an adequate interpretation of the results to correct this fluorescence spillover, and this is achieved through compensation. Fluorescence spillover compensation is a well-known feature in flow cytometry analysis, where experiments up to 12-color can be run [43].

First, the contribution of each fluorophore to the fluorescent signal in each detection channel must be evaluated, which is achieved through the use of mono-color controls, ie controls where only one fluorescent channel will have positive droplets. From these mono-color controls, the interferent signal in other channels can thus be calculated, as a percentage of the fluorescence intensity in the primary detection channel. Once calculated, this matrix is inverted and applied to the extracted fluorescence values, yielding the “true” fluorophore signal, unencumbered by spillover contribution. As a result, the various population clusters regain an orthogonal positioning, when viewed on a 2D dot plot (Fig. 10). For reviews on fluorescence spillover, and compensation algorithm, please refer to [41] for additional details.

The amount of spillover is dependent on the signal intensity of the probe, and will need to be calibrated using specific experimental conditions (reaction efficiency, exposure times, . . . ). However, once an assay is calibrated, the compensation matrix can be applied to all subsequent samples using the same assay.

4.3. Droplet classification in three-color digital PCR experiments

Once appropriate spillover compensation has been applied to the raw fluorescence data, the fluorescence values displayed on scatterplots represent the signal originating from a unique probe for each detection channel. Thresholds or gates can therefore be set to discriminate between positive and negative populations of droplets and thus allow quantification of targets.

In order to separate negative and positive droplet populations, a simple threshold, i.e. a unique fluorescence intensity value, can be set independently for each detection channel. Such thresholds can readily be set and displayed on either one- or two-dimensional scatterplots, as a line with the chosen threshold value as an invariant coordinate, while the other coordinate tends to infinity. In a three-dimensional setting, these thresholds then equate to two-dimensional planes.

Droplet populations may also be selected by gating using 2-D geometrical shapes or free-hand drawing tools. However, caution must be exercised during interpretation when using such tools on multi-dimensional data. As described in Section 4.1.2., all data from a duplex experiment can be readily displayed on a single 2-D scatterplot to determine the single-positive and double-positive populations for each channel. However, in a triplex experiment, these populations are single- or double-positive only for the referential used. In fact, they represent the projection of both single positives and double-positives, as well as both double-positives and triple-positives onto the 2D referential. Such a selection is therefore not wrong, but care must be taken when interpreting data. In order to ease visualization of contributions from the third channel, the Naica analysis software offers an 8-color visualization mode, using the RGB additive color system described in Section 4.1.3.
4.4. **Assay performance using a three-color approach**

Oncology is one of the leading fields for application of digital PCR, addressing mainly, yet not exclusively, genotyping of cell-free DNA present in bodily fluids, such as blood. In cases of lung cancer specifically, tumors may be hard to biopsy, with up to 25% of patients suffering from advanced lung cancer not being able to be biopsied [44].

Cell-free DNA has therefore now been accepted as a viable alternative to biopsy, as shown by the recent recommendation by Health...
authorities in the USA and Europe to allow circulating DNA to be used to determine the presence of absent of specific cancerous mutations, and therefore assess the therapeutic pertinence of targeted therapies [45], [46].

Approximately 10–35% of patients with non-small cell lung cancer have mutations in the epidermal Growth Factor Receptor (EGFR) gene [47–49]. Some of these mutations may predict response to tyrosine kinase inhibitor (TKI), and can be indicative of response to treatment such as EGFR L858R or L861Q mutations, or resistance, such as EGFR T790M. Cell-free tumor DNA in the patient plasma is typically present in very small proportion to wild-type DNA [32]. It is therefore of the utmost importance to yield the maximum information from each sample, and not split samples for analysis. Multiplexing does allow monitoring of multiple mutations within a single test.

In order to detect activating and resistance mutations in a single assay, a quadriplex assay was developed to detect activating mutations EGFR L858R or L861Q, using two Cy5-labelled probes, as well as the resistance mutation EGFR T790M using a Cy3-labelled probe, while quantifying wild-type EGFR L858 using a FAM-labelled probe. In order to characterize the performances of the assay, we first evaluated the specificity of the probes detecting the mutations against wild-type DNA.

This preliminary characterization allowed us to calculate a limit of blank (LoB), defined as the number of events measured where no mutations are present, i.e. the false positive rate (results available in Table S1, please see Section 2.5.1 for details on calculations). The assays were then characterized by testing serial dilutions of reference DNA carrying the EGFR L858R, L861Q or the T790 M mutations in a background of control DNA at a final concentration of 500 copies/μl, at ratios varying from 0.03% to 3% (Fig. 11). A regression curve was then fitted to the observed values and showed good linearity of the data (R2 values>0.9). Moreover, theoretical values and associated uncertainties, representing quantification of ratios in an ideal digital PCR system, were also plotted. As can be observed in Fig. 11A–C, there is excellent correlation between “ideal” and observed values for each probe within this quadriplex assay.

Human genomic DNA bearing either EGFR L858R (A), L861Q (B) or T790 M (C) mutations were serially diluted in a background of 500 copies/μl of wild-type human genomic DNA. The resulting DNA mixtures were used to calibrate the EGFR quadriplex assay performance using Crystal Digital PCR.

To verify that multiplexing does not come at the cost of sensitivity, we also set up duplex assays, in which each probe detecting one of the three mutations was assayed in association with the probe detecting the wild-type EGFR. As shown in Table 6, both

**Table 6**

<table>
<thead>
<tr>
<th>Duplex Assay</th>
<th>Cmax</th>
<th>Cmin</th>
<th>% Mutant: WT</th>
<th>Cmax</th>
<th>Cmin</th>
<th>% Mutant: WT</th>
</tr>
</thead>
<tbody>
<tr>
<td>T790M N + 1</td>
<td>0.8</td>
<td>0.3</td>
<td>0.04</td>
<td>0.5</td>
<td>0.2</td>
<td>0.03</td>
</tr>
<tr>
<td>T790M N + 2</td>
<td>0.5</td>
<td>0.2</td>
<td>0.03</td>
<td>0.6</td>
<td>0.2</td>
<td>0.03</td>
</tr>
<tr>
<td>T790M N + 3</td>
<td>0.7</td>
<td>0.3</td>
<td>0.05</td>
<td>0.8</td>
<td>0.4</td>
<td>0.07</td>
</tr>
<tr>
<td>L861Q N - 1</td>
<td>1.1</td>
<td>0.6</td>
<td>0.07</td>
<td>0.6</td>
<td>0.3</td>
<td>0.05</td>
</tr>
<tr>
<td>L861Q N + 1</td>
<td>1.1</td>
<td>0.6</td>
<td>0.07</td>
<td>0.6</td>
<td>0.3</td>
<td>0.05</td>
</tr>
<tr>
<td>L861Q N + 2</td>
<td>0.3</td>
<td>0.1</td>
<td>0.02</td>
<td>1.2</td>
<td>0.4</td>
<td>0.03</td>
</tr>
<tr>
<td>L861Q N + 3</td>
<td>0.3</td>
<td>0.1</td>
<td>0.02</td>
<td>0.9</td>
<td>0.4</td>
<td>0.02</td>
</tr>
</tbody>
</table>

C: Concentration of sample in copies per microliter. Cmin and Cmax respectively designate the minimum and maximum concentrations of the 95% confidence interval.

Please cite this article in press as: J. Madic, et al., Three-color crystal digital PCR, Biomol Detect Quantif (2016), http://dx.doi.org/10.1016/j.bdq.2016.10.002
systems show similar performances both for absolute quantification, ie the concentration of mutant alleles observed, and for the ratio of mutant alleles to wild-type.

This assay thus highlights that multiplexing does not inherently decrease the sensitivity of an assay and furthermore provides an elegant demonstration of the power afforded by increased color capacities.

5. Towards higher color-multiplexing capabilities

There is a strong case to increase even further multiplexing capacities in digital PCR, in order to decrease costs while increasing throughput, while making the most of scarce samples. Additionally, higher multiplexing would also enable more complex experiments to be performed, for example for linkage or single cell analyses. Finally, higher multiplexing digital PCR could also be used to ensure data reliability by allowing the inclusion of multiple controls within a single assay.

In this publication, we have characterized the first three-color digital PCR, and discussed the specific considerations, intrinsic to this type of experiment, such as fluorescence spillover, data representation, and population assignment. Some of these considerations can readily be extended to n-color digital PCR, for n > 3, for example the calibration and compensation of a measuring matrix. Others, such as data representation and population identification, may however need ad hoc developments.

Data originating from n-color experiments can be represented by a serie of (n(n-1))/2 two-dimensional scatterplots, meaning that data from a 4-color experiment can be displayed using 6 2D-scatterplots, while 10 scatterplots will be needed for a 5-color experiment, and 21 for a 6-color experiment! While it is true that if the only requirement for the experiment is the absolute quantification of each target, regardless of co-occurrence in the same droplet, thresholds can easily be checked and set using one-dimensional scatterplots. However, we would still advise to verify appropriate fluorescence spillover compensation, and lack of probe cross-reactivity on 2-D scatterplots.

In such case, or for more complex experimental designs where co-occurrence of signals within droplets may need to be characterized, analyzing 21 scatterplots for each sample would certainly prove unmanageable. Thus, in order to expand to n > 3-color digital PCR, and render it user-friendly (or even just usable), we believe two key aspects of data analysis must be tackled: the first regarding the development of algorithms for automated cluster identification in n-dimensions, and the second on envisioning new paradigms for the representation of n-dimensional data, this last issue having potential applications beyond digital PCR.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bdq.2016.10.002.

References

[20] N. Raki, T. Dreo, I. Gutierrez-Aguirre, A. Blejec, M. Ravnikar, Reverse transcriptase droplet digital PCR shows high resilience to PCR inhibitors from plant, soil and water samples, Plant Methods 10 (December (1)) (2014).


Please cite this article in press as: J. Madic, et al., Three-color crystal digital PCR, Biomol Detect Quantif (2016), http://dx.doi.org/10.1016/j.bdq.2016.10.002