Detection and image processing after nucleic acid amplification in emulsion

A Tupik¹, L Novikov¹,² and A Evstrapov¹

¹ Institute for Analytical Instrumentation of the Russian Academy of Sciences, St. Petersburg, 198095, Russia
² Saint Petersburg State Institute of Technology (Technical University), St. Petersburg, 190013, Russia

E-mail: tunix@yandex.ru

Abstract. To analyze the results of amplification in the emulsion, special software package was developed. It was used to process images of emulsions, in which digital loop mediated isothermal amplification (LAMP) was carried out. Program units made it possible to combine a set of images of a reaction chamber into a single image file, count the number of droplets in chamber, and obtain histograms of the size distribution and fluorescence intensity of the droplets. The data obtained are necessary at the stage of equipment development for the selection of design parameters and for improving of the experiment methodology.

1. Introduction

Digital nucleic acid amplification methods are an actively developing field of research because they allow absolutely quantitative determination of DNA in a sample, which makes the need for standard curves redundant. It is possible to achieve sensitivity at the level of single molecules, which is relevant in the detection of rare cancer mutations, low-concentration pathogens, in sequencing methods, etc. [1] The development and improvement of these methods is closely related to the creation of devices for amplification in the emulsion. Droplets, insulated from the environment by a shell of oil or polymer, are peculiar analogues of reaction chambers for carrying out chemical, enzymatic or other specific reactions [2]. The advantages of droplet-based amplification of nucleic acids are: high performance studies in an array of droplets, low consumption of reagents and samples, high rates of emulsion formation. In contrast to ordinary reaction chambers, droplets are not static but manipulated within networks of channels for transportation, merge or “break up” droplets after amplification.

In digital nucleic acid amplification, the sample is first partitioned into many independent sub-reactions (droplets) such that each partition contains either a few or no target DNA. When the initial sample is diluted to an average concentration of ~ 1 molecule/droplet, the distribution of DNA in the droplets can be approximated by the Poisson distribution. After amplification the ratio of positive droplets (presence of fluorescence amplified DNA fragments) over the total number allows to determine the concentration of the target DNA in the initial sample. The more droplets, the higher the sensitivity of the method. Over a million droplets are generated in modern microfluidic systems [3]. However, this increases the complexity and duration of the detection of results, which consists in counting the total number of droplets and the fluorescent droplets.
Initially, to record the results of digital polymerase chain reaction (PCR) in the emulsion, each droplet in the flow was sequentially interrogated on equipment inherited from flow cytometry. However, the readout is characterized by relatively low speed and long duration, because droplets cannot withstand high flow rates. This limitation can be overcome by converting droplets into cytometry-compatible particles such as magnetic or agarose beads, or by using a double emulsion [3]. The second approach is to register using optical microscopy of droplets located in 2D arrays. This is faster, inexpensive and easier to implement, and can also be used to perform real-time registration. However, this requires the development of special algorithms and programs to automatically recognize, count and record the fluorescence signal from a few thousand droplets.

To analyze the results of amplification in the emulsion, special software was developed. It made it possible to combine a set of images of a single reaction chamber into a single image file, automatically calculate the number of droplets, and obtain a histogram of the size distribution and fluorescence intensity of the droplets. This program was used to process images of emulsions, in which droplets was carried out digital loop mediated isothermal amplification (LAMP). LAMP is characterized by high specificity as well as greater resistance to inaccuracies in the temperature of amplification and to the presence of inhibitors, which makes it an attractive alternative to PCR analysis.

2. Experimental techniques
For the experiments, microfluidic chips with a high-performance droplets generator were used [4]. It allows the formation of a water-in-oil emulsion with droplets diameters of about 20 µm and volumes of less than ten femtoliters per droplet. For amplification, domestic reagents for LAMP and a target DNA fragment at a final concentration of 0.01 pM were used. An emulsion was generated on the microfluidic chip from the prepared amplification mixture, and then was transferred to a flat reaction chamber and covered with a cover glass and was heated for approximately 45 minutes. The result of DNA amplification was determined by fluorescence induced by fluorescent probes with FAM dye.

A prototype of a fluorescent scanner DMC-1, developed in the IAI RAS, was used to droplets detection. The excitation wavelength was 473 nm. Fluorescence was recorded in the range of 515-535 nm. The scanner allows to receive a set of images in the size 450 × 350 µm in two modes: optical transmission and fluorescence. The resolution of the scanner is 1.4 µm per pixel. When registering a reaction chamber with a diameter of 6 mm in both modes, a set of about 500 image fragments is obtained.

3. Image processing
The software package for the analysis images obtained as a result of scanning the reaction chamber is developed in MATLAB [5] and consists of several program units:

- first unit for forming a complete image of the reaction chamber from a set of fragments of its image, each 240 × 310 pixels in size;
- second program unit for detecting droplets on the image;
- third unit for comparing the total number and number of fluorescent droplets and their distributions by size and fluorescence intensity.

3.1. First program unit
The first program unit forms a single image taking into account a known number of image fragments in one line and the number of these lines, and also corrects the background. The result is a single rectangular image in the 0: X_{max} and 0: Y_{max} coordinate system. In the center of the image is a round-shaped reaction chamber with a contrasting background around. Given that a differentiation procedure is used to detect droplets in the future, it is necessary to remove the sharp contrast on the edges of the chamber. For this, the shape of the edges of the chamber is approximated by an ellipsoidal function as follows.
Denote the position of the left and right edges of the chamber as \( x_1 \) and \( x_2 \), and the top and bottom as \( y_1 \) and \( y_2 \). Then the image size is equal to \( X = |x_1 - x_2| \) on the \( x \)-axis, and \( Y = |y_1 - y_2| \) on the \( y \)-axis. Suppose that the chamber image is stretched horizontally, i.e. \( X > Y \). The coordinates of the center of the ellipse (the center of the reaction chamber): \( X_0 = x_1 + 0.5 \cdot X \), \( Y_0 = y_2 + 0.5 \cdot Y \). The distance from the center to any focus of the ellipse is \( r = \sqrt{(0.25 \cdot X^2 - 0.25 \cdot Y^2)^2} \) and the focus coordinates are \( (X_0 - c, Y_0) \) for the first focus \( F_1 \) and \( (X_0 + c, Y_0) \) for the second focus \( F_2 \). Then the distance from any point \( (x, y) \) to \( F_1 \) is equal to \( r_1 = \sqrt{[x - (X_0 - c)]^2 + (y - Y_0)^2} \), and for \( F_2 \) is equal to \( r_2 = \sqrt{[x - (X_0 + c)]^2 + (y - Y_0)^2} \). The condition for finding this point inside the ellipse is the inequality \( r_1 + r_2 < X \). This condition is used to equalize the contrast of the image: if it is not, the point (pixel) is set to a value equal to the background level.

As a result of the work of the first program unit, two images of the reaction chamber (in the transmission and fluorescence mode) with background correction are obtained.

### 3.2. Second program unit

The second program unit solves the problem of detecting droplets and estimating their parameters. For the image in transmission mode, the number of droplets and their diameter are determined, and for the image in the fluorescence mode, the number of drops, their diameter and the fluorescence intensity.

It is possible in the dialogue mode to set the thresholds for the minimum size of the droplets, the minimum and maximum of fluorescence intensity. Thresholds are selected by the distribution histograms displayed, taking into account a priori data on the characteristics of the droplet generator and the dye.

The program for parameters estimation of droplets in optical transmission mode uses commands from the MATLAB Image Processing Toolbox. First, on the grayscale image, the droplets borders are determined using the Canny method [5] (its parameters for detecting the borders of all the droplets have been chosen empirically). In the generated binary image, all areas with closed loop are filled, and the thin lines remained from unclosed areas are removed by morphological opening and closing. Then, the coordinates of the centers and the diameters of all the droplets are determined.

Next, the diameters are compared with given thresholds in order to discard random small point noise from further processing. Then a histogram of the distribution of the number of droplets by their diameters is shown and the first table containing the coordinates of the centers of the droplets detected in the transmission mode and their diameters is formed.

The program for estimation the parameters of droplets in the fluorescence mode uses the same MATLAB Toolbox. Similarly a binary image with a given threshold is formed, thin lines remained from unclosed areas are removed, and the centers and diameters of the fluorescent droplets are determined. Then all the droplets are sequentially analyzed and, if the diameter is greater than the minimum threshold, the intensity is estimated. Then the intensity of each droplet is compared with the minimum and maximum intensity thresholds to remove too dim or bright fluorescent artifacts. According to the results, histograms of the distribution of the number of fluorescent droplets by their diameters and intensities are displayed. A second table is also formed, containing the coordinates of the centers of the detected droplets, their diameters and fluorescence intensities.

As a result of the second program unit, two tables are formed containing the coordinates of the centers of the detected droplets and their diameter in the transmission mode, and in the fluorescence mode – the coordinates of the centers, diameters and intensities.

### 3.3. Third program unit

The third program block in the dialogue mode allows to analyze the results. This program:
• sequentially displays the images of the reaction chamber in the transmission and fluorescence mode, as well as their joint image with indicating the total droplets number and number of fluorescent droplets;
• determines the proportion of fluorescent droplets in the total number of droplets;
• after selecting any range of intensity on the histogram of the distribution of the intensity of the fluorescence of droplets, allows you to display the diameters of these droplets on the histogram of the distribution of the diameters of the fluorescent droplets;
• marks droplets from a selected range of fluorescence intensity on a chamber image in fluorescence mode;
• after selecting any diameter range on the histogram of the diameters distribution of fluorescent droplets, allows you to determine the intensity of selected droplets on the histogram of the intensity distribution of fluorescent droplets;
• marks fluorescent droplets from a selected range of diameters to the camera image in fluorescence mode.

The data obtained during the work of the third program unit is necessary at the development stage for the selection of design parameters and for improvement of experimental procedures.

An example of detecting non-fluorescent (bright rings) and fluorescent (dark rings) droplets on a fragment of the chamber image in transmission mode after amplification is shown in figure 1.

![Figure 1](image)

**Figure 1.** An example of detecting fluorescent and non-fluorescent droplets in an emulsion image.

3.4. **Testing on a simulated image**

It is obvious that the reliability of the findings obtained by the researcher on the results of the analysis in the third program unit is determined by the quality of processing the source data in first and second program unit. The quality check of the applied algorithms and programs was carried out on the model of the reaction chamber image in the fluorescence mode. The model is synthesized as follows. The researcher enters the range of fluorescent intensity and diameter of the droplets, as well as their number. The program randomly generates the positions of the droplets on the background of the real chamber image as substrate, also their diameters and intensity in a given range. Figure 2 shows a fragment of the image of the reaction chamber in the fluorescence mode and the simulated model of the reaction chamber image. All droplets are round objects that do not intersect with each other, with diameters ranging from 5 to 100 µm.

As a result of processing the model with the preceding two program units, it turned out that the probability of detecting droplets is one, the probability of false detection and skipping is zero when the intensities of the objects is more than 2% of the background level of fluorescence. The maximum error in estimating the diameters of the droplets is no more than ten percent, and the intensity is no more than five percent. Further research and expansion of the range of sizes and brightness of objects is possible to adapt the software for detecting and estimating parameters of contrasting objects with closed borders for other similar purposes.
Figure 2. Image the fragments: a) of the reaction chamber in fluorescence mode and b) the simulated test-model (the number of droplets, diameter and fluorescent intensity are randomly set).

4. Results
Based on the results of image processing of three reaction chambers, it was found that the number of droplets in one reaction chamber ranges from 10 000 to 25 000. The number of fluorescent droplets in this case is tenths of a percent. Figure 3 shows the histograms of the droplet number distribution by diameter.

Figure 3. Histograms of the distribution of the number of droplets by diameters after carrying out isothermal amplification: a) the fluorescent droplets, b) all the droplets.

Microfluidic droplet generation methods make it possible to obtain relatively good monodispersity of droplets with an error within several percent [3, 6]. The presence on the histograms of a decreasing front indicates that some of the droplets merge during transportation and during heating for amplification. These should be considered when developing a method for the experiment.

5. Conclusion
The considered software package provides registration and evaluation of droplet parameters when solving problems of digital amplification. This allows the developers of a serial device to choose the optimal parameters of the reaction chamber, the droplets generator and methods of conducting the experiment, test the device control and data processing algorithms, and determine the requirements for the user interface.
References
[1] Karlin-Neumann G and Bizouarn F 2018 Digital PCR (NY: Humana Press) p 553
[2] Guo M T, Rotem A, Heymanab J A and Weitz D A 2012 Lab Chip 12 2146
[3] Quan P, Sauzade M and Brouzes E 2018 Sensors 18 1271
[4] Tupik A, Posmitnaya Y, Rudnitskaya G, Varlamov D and Evstrapov A 2018 J. Phys.: Conf. Ser. 1135 012110
[5] Gonsales R, Woods R and Eddins S 2009 Digital Image Processing using MATLAB (Gatesmark Publishing) p 827
[6] Schuler F, Siber C, Hin S, Wadle S, Paust N, Zengerleab R and Stetten F 2016 Anal Methods 8 2750