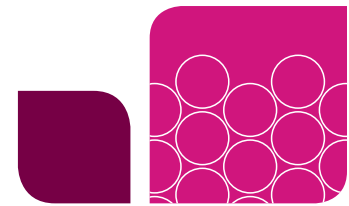


Guidelines for Phasing Variants Using Bio-Rad's Droplet Digital PCR System

Regan JF,¹ Kamitaki N,² Legler T,¹ Cooper S,¹ Klitgord N,¹ Karlin-Neumann G,¹ Wong C,³ Hodges S,¹ Koehler R,¹ Tzonev S,¹ McCarroll SA.² ¹ Digital Biology Center, Bio-Rad Laboratories Inc., Pleasanton, CA; ² Department of Genetics, Harvard Medical School, Boston, MA; Program in Medical and Population Genetics and Stanley Center for Psychiatric Research, Cambridge, MA; ³ Department of Genetics, Harvard Medical School, Boston, MA.



Droplet Digital PCR

Bulletin 6718

Abstract

We describe a Droplet Digital PCR–based method, Drop-Phase, which allows users to quickly and accurately phase genomic variants. We successfully phased pairs of *CFTR* variants and show that the methodology can measure linked species that are up to 210 kb apart. This method can be used to phase variants discovered through sequencing or to screen unsequenced individuals for a particular haplotype of interest for clinical trials.

Introduction

Droplet Digital PCR (ddPCR) enables absolute quantification of nucleic acid sequences by dividing reaction mixtures into nanoliter-sized droplets. Well-known applications for Droplet Digital PCR include sensitive rare mutation detection and accurate measurements of copy number variation (CNV). Here, we show that Droplet Digital PCR can also be used to determine which alleles are located on the same chromosome, a process known as phasing.

Drop-Phase involves performing a duplex ddPCR reaction that targets the two variants to be phased (Figure 1). DNA molecules that are long enough to preserve the physical link between the two targeted variants will result in more double-positive droplets than would be expected from random colocalization of unlinked variants (Figure 2). This overabundance of double-positive droplets is measured by QuantaSoft Software and used to calculate the concentration of linked variants in the reaction.

Unlike traditional phasing methods, which are labor-intensive, low-throughput, and limited in distance, Drop-Phase is easy to perform, high throughput, and works over large genomic distances.

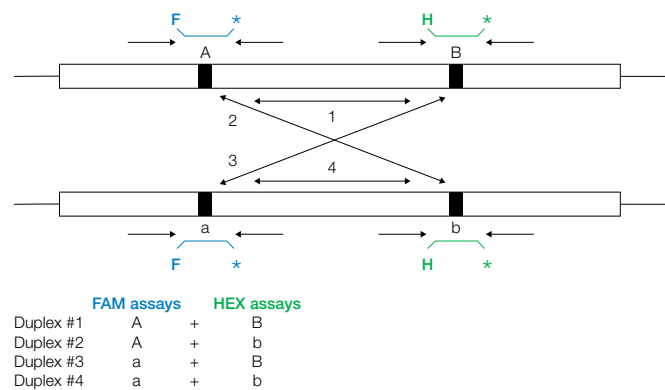


Fig. 1. Assay diagram. A and a are two alleles of the same DNA region and B and b are two different alleles of a region near to A/a on the same strand of DNA (Regan et al. 2015). * denotes quencher of the fluorophore (FAM or HEX).

Materials and Methods

Samples

All cell lines were obtained from the Coriell Institute for Medical Research. For cell line information and culture conditions please reference Regan JF et al. (2015).

Droplet Digital PCR

We performed ddPCR as instructed by the manufacturer (Bio-Rad Laboratories), but used wide-bore pipet tips (Rainin) and gentle mixing to minimize DNA fragmentation.

To highlight the importance of sample extraction for phasing more distant variants, we compared two different sample extraction methods, silica columns and polysaccharide-based precipitation. DNA extracted using these chemistries was interrogated using a series of duplex assays composed of a HEX assay targeting a conserved anchor point paired with a single FAM assay targeting a conserved region of the genome that is a known distance (1, 10, 33, 60, 100, 150, or 210 kb) from the anchor point.

ddPCR Supermix for Probes (No dUTP) (Bio-Rad Laboratories) was used for all experiments. Primers, probes, and gDNA were used at concentrations of 900 nM, 250 nM, and ~650 pg/μl, respectively. Droplets were generated using the QX200 Droplet Generator, transferred into 96-well plates (Eppendorf), sealed with a Pierceable Foil Heat Seal (Bio-Rad Laboratories), and cycled in a C1000 Thermal Cycler (Bio-Rad Laboratories). To compare sample extraction methods, reactions were cycled using the following protocol: 95°C, 10 min for 1 cycle; (94°C, 30 sec; 60°C, 1 min) for 40 cycles; 98°C, 10 min for 1 cycle. To phase *CFTR* variants, reactions were cycled as follows: 95°C, 10 min for 1 cycle; (94°C, 30 sec; 55°C, 1 min) for 40 cycles; 98°C, 10 min for 1 cycle. Droplets were read using a QX200 Droplet Reader and data were analyzed using QuantaSoft Software v.1.7.4.0917 (Bio-Rad Laboratories).

DNA Extraction

For polysaccharide-based precipitation, 2.8×10^5 cells/ml in 1x PBS were extracted using the PrepFiler Forensic DNA Extraction Kit (Life Technologies) following the manufacturer's instructions with the following modifications: the lysis incubation was reduced to 10 min in the absence of shaking and all vortexing steps were replaced with gentle end-over-end inversion except for the mixing step following isopropyl alcohol addition. Samples were centrifuged only briefly and transferred using wide-bore pipet tips.

For silica column-based extraction, 2.8×10^5 cells in PBS were extracted using the DNeasy Blood and Tissue Kit (Qiagen) following manufacturer's instructions with the following modifications: 160 μl of PBS were added to the 40 μl sample to bring the sample volume to 200 μl before addition of 200 μl buffer AL + proteinase K, and samples were mixed by gentle end-over-end inversion and transferred using wide-bore pipet tips.

Linkage Calculation

Linked targets result in double-positive droplets. The percentage of linked targets in a sample was calculated by determining the excess of double-positive droplets over that expected due to chance colocalization of unlinked targets (Figure 2). For the mathematical calculations regarding linkage, please refer to Regan JF et al. (2015). Linkage is calculated by QuantaSoft Software in the Table view as copies/μl of linked molecules for each well.

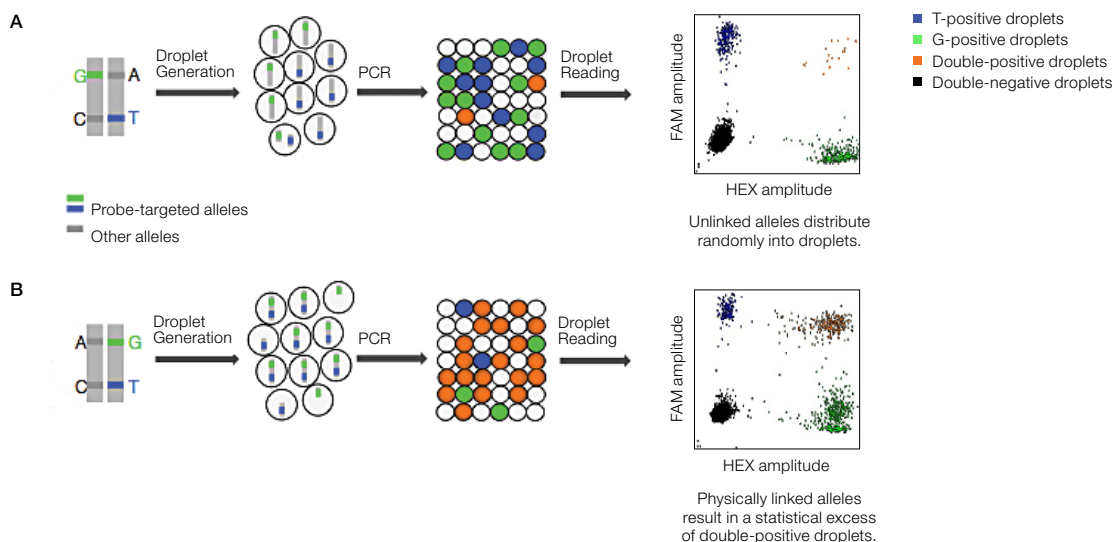


Fig. 2. Drop-Phase method. Variant-specific fluorescent probes (FAM: ■, HEX: ■) are used to detect the variants of interest in undigested genomic DNA. Following PCR, droplets are positive for one fluorophore (■ or ■), both fluorophores (■), or neither (■), depending on whether the assayed variants are linked or unlinked. Unlinked variants that are present on different chromosomes partition independently into droplets (**A**). Variants on the same unbroken strand cosegregate into the same droplets because they are physically linked (**B**).

Results

We demonstrated the utility of Drop-Phase by successfully phasing pathogenic variants in the *CFTR* gene in samples derived from six different cystic fibrosis patients (Figure 3). Our results matched the clinical outcome, where a disease phenotype is observed only when pathogenic variants are present in both copies of the *CFTR* gene.

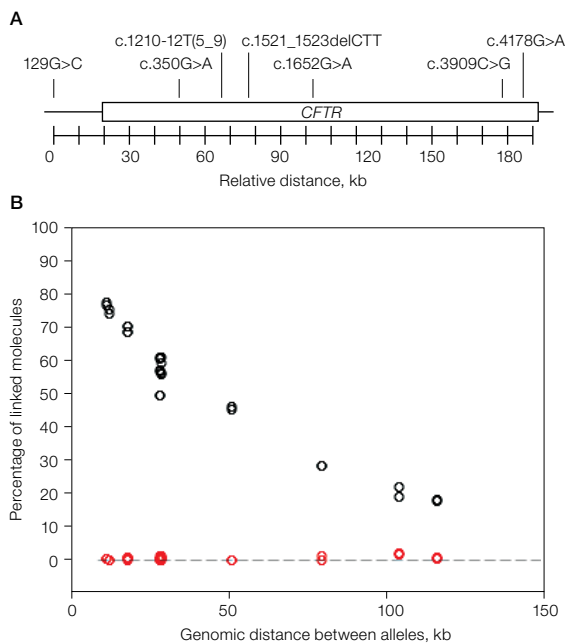


Fig. 3. Phasing *CFTR* alleles in cell lines of cystic fibrosis patients. **A**, locations of mutations (alleles); **B**, experimentally determined percentage of physically linked *CFTR* alleles in cell lines from six cystic fibrosis patients. Four duplex assays were performed per targeted locus. Alleles were classified as *cis*- or *trans*-configured based on linkage or lack of linkage, respectively. All calls matched those expected for the disease status. *cis*-alleles (○); *trans*-alleles (◐).

We next evaluated the influence of different sample extraction chemistries on the ability to measure linked DNA molecules and found that silica-based columns permitted linkage measurements only up to ~60 kb, whereas a polysaccharide-based precipitation method enabled linkage measurements out to 210 kb (Figure 4).

Correctly interpreting Drop-Phase plots can be easily achieved by understanding the factors (competition and cross-reactivity) that contribute to the creation of additional clusters (Figure 5). At high enough DNA loads, some droplets will contain both targeted and nontargeted markers. An example of a nontargeted marker is the wild-type sequence of the targeted gene that is present on the other chromosome. When both targeted and nontargeted markers are present in the same droplet, these templates compete for PCR reagents, resulting in a reduction in the fluorescence amplitude for these droplets. At high enough DNA loads, competition can result in as many as three additional clusters near each major FAM+, HEX+, and FAM+/HEX+ cluster,

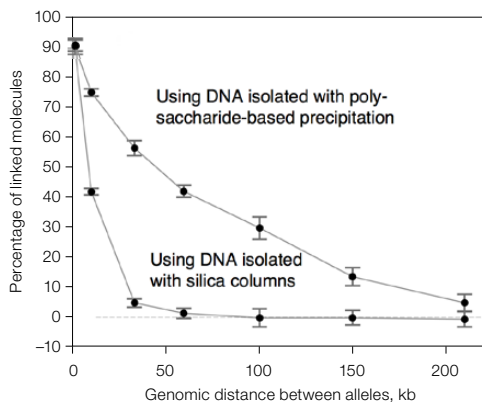


Fig. 4. DNA extraction methods influence distance across which chromosomes can be phased. FAM-labeled assays were designed to target sequences at different distances (1–210 kb) from the *RPP30* locus, which was targeted using a HEX-labeled assay. Means of triplicate measurements and 95% confidence intervals are shown.

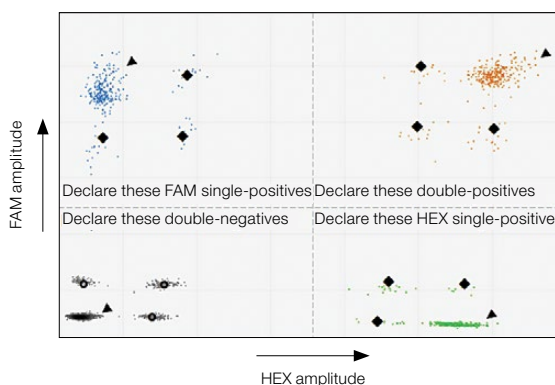


Fig. 5. Droplet cluster identification and classification in the presence of competition and cross-reacting fluorescent probes. Although fluorescent probes are designed to detect only the targeted allele, they can bind at a lower frequency to the nontargeted allele, creating additional clusters on a 2-D plot. Major clusters (▲); minor competition clusters (◆); cross-reactive clusters (◐).

for a maximum of nine additional clusters. The only way to avoid these additional clusters is to load less DNA.

In addition to competition, the probes used in Drop-Phase reactions often cross-react with the nontargeted allele, which can add up to three extra clusters to the 2-D plots. Reducing DNA load will not remove probe cross-reactivity, but the addition of nonfluorescent competitor probes that bind the nontargeted allele will greatly reduce cross-reactivity and often drive the cross-reactive clusters into the gray cluster in the lower left corner of the plot that contains neither targeted nor nontargeted markers. This simplifies the plot and minimizes the chance of improperly classifying droplets.

To phase, regardless of whether competition or cross-reactivity exists, it is necessary only to distinguish the droplets that are positive for the targeted markers from those that lack the targeted markers. To achieve this, the minor competition clusters are grouped with their respective major cluster

containing only the targeted marker. Any additional clusters present due to cross-reactivity are classified as double-negative droplets. This results in a quadrant plot that is treated the same way as a traditional Droplet Digital PCR 2-D plot by QuantaSoft Software, which uses the number of droplets in each quadrant to determine whether linked species are present.

Conclusions

Drop-Phase reactions are quick and scalable (96 samples can be processed in 5 hours). These features make Drop-Phase an ideal choice for phasing known variants discovered through sequencing. Additionally, Drop-Phase can be used to screen a population to identify individuals with a particular haplotype. This method will benefit genome engineering, allele-specific expression analysis, disease genetics, and de novo mutation investigation.

Additional details about this application can be found in "A rapid molecular approach for chromosomal phasing," by Regan JF et al. (2015).

References

Regan JF et al. (2015). A rapid molecular approach for chromosomal phasing. PLoS One 10, e0118270.

Visit bio-rad.com/digitalPCR for more information.

Bio-Rad, Droplet Digital PCR, and ddPCR are trademarks of Bio-Rad Laboratories, Inc. in certain jurisdictions.

The QX200 Droplet Digital PCR System and/or its use is covered by claims of U.S. patents, and/or pending U.S. and non-U.S. patent applications owned by or under license to Bio-Rad Laboratories, Inc. Purchase of the product includes a limited, non-transferable right under such intellectual property for use of the product for internal research purposes only. No rights are granted for diagnostic uses. No rights are granted for use of the product for commercial applications of any kind, including but not limited to manufacturing, quality control, or commercial services, such as contract services or fee for services. Information concerning a license for such uses can be obtained from Bio-Rad Laboratories. It is the responsibility of the purchaser/end user to acquire any additional intellectual property rights that may be required.

Bio-Rad's thermal cyclers and real-time thermal cyclers are covered by one or more of the following U.S. patents or their foreign counterparts owned by Eppendorf AG: U.S. Patent Numbers: 6,767,512 and 7,074,367.

All trademarks used herein are the property of their respective owner.



**Bio-Rad
Laboratories, Inc.**

Life Science
Group

Web site bio-rad.com **USA** 1 800 424 6723 **Australia** 61 2 9914 2800 **Austria** 43 01 877 89019 **Belgium** 32 03 710 53 00 **Brazil** 55 11 3065 7550
Canada 1 905 364 3435 **China** 86 21 6169 8500 **Czech Republic** 36 01 459 6192 **Denmark** 45 04 452 10 00 **Finland** 35 08 980 422 00
France 33 01 479 593 00 **Germany** 49 089 3188 4393 **Hong Kong** 852 2789 3300 **Hungary** 36 01 459 6190 **India** 91 124 4029300
Israel 972 03 963 6050 **Italy** 39 02 49486600 **Japan** 81 3 6361 7000 **Korea** 82 2 3473 4460 **Mexico** 52 555 488 7670 **The Netherlands** 31 0 318 540 666
New Zealand 64 9 415 2280 **Norway** 47 0 233 841 30 **Poland** 36 01 459 6191 **Portugal** 351 21 4727717 **Russia** 7 495 721 14 04
Singapore 65 6415 3188 **South Africa** 36 01 459 6193 **Spain** 34 091 49 06 580 **Sweden** 46 08 555 127 00 **Switzerland** 41 0617 17 9555
Taiwan 886 2 2578 7189 **Thailand** 66 2 651 8311 **United Arab Emirates** 971 4 8187300 **United Kingdom** 44 01923 47 1301

