UC Irvine
UC Irvine Previously Published Works

Title
Digital-Droplet PCR to Detect Indels Mutations in Genetically Modified Anopheline Mosquito
Populations.

Permalink
https://escholarship.org/uc/item/39r9m8cj

Journal
Journal of Visualized Experiments, 2021(172)

ISSN
1940-087X

Authors
Carballar-Lejarazú, Rebeca
Pham, Thai Binh
Kelsey, Adam
et al.

Publication Date
2021

DOI
10.3791/62607

Peer reviewed

jove

Digital-Droplet PCR to Detect Indels Mutations in Genetically Modified Anopheline Mosquito Populations

Rebeca Carballar-Lejarazú[1], Thai Binh Pham[2], Adam Kelsey[1], Taylor Tushar[1], Anthony A. James[1,2]

[1] Department of Microbiology & Molecular Genetics, University of California, Irvine [2] Department of Molecular Biology & Biochemistry, University of California, Irvine

Corresponding Author

Anthony A. James

aajames@uci.edu

Abstract

Recent advances in mosquito genomics and genetic engineering technologies have fostered a need for quick and efficient methods for detecting targeted DNA sequence variation on a large scale. Specifically, detecting insertions and deletions (indels) at gene-edited sites generated by CRISPR guide RNA (gRNA)/Cas9-mediated non-homologous end-joining (NHEJ) is important for assessing the fidelity of the mutagenesis and the frequency of unintended changes. We describe here a protocol for digital-droplet PCR (ddPCR) that is well-suited for high-throughput NHEJ analysis. While this method does not produce data that identifies individual sequence variation, it provides a quantitative estimate of the sequence variation within a population. Additionally, with appropriate resources, this protocol can be implemented in a field-site laboratory setting more easily than next-generation or Sanger sequencing. ddPCR also has a faster turn-around time for results than either of those methods, which allows a more quick and complete analysis of genetic variation in wild populations during field trials of genetically-engineered organisms.

Introduction

Gene drives have immense potential to control insect populations of medical and agricultural relevance[1,2,3,4,5]. For example, gene-drive systems based on CRISPR Cas nucleases and guide RNAs (gRNAs) can be used to modify vector mosquito populations by introducing traits that confer refractoriness to malaria parasites leading to reduced transmission and less disease[1,4,5]. The gene-drive system copies itself and the associated trait from one homologous chromosome to another in the pre-meiotic germ cells, and this ensures that the majority of the offspring inherit the drive and create the potential for long-lasting and sustainable population modification in the field. However, one disadvantage of the Cas/gRNA-based methods is the possibility of generating insertion and deletion (indel) mutations through non-homologous end-joining (NHEJ) DNA repair, resulting in the generation of drive-resistant alleles, which when accumulated to a high enough frequency in the population, can stop the drive system from spreading[1,2,3,4].

This protocol details a high-throughput and reliable method that can determine the prevalence and relative quantity of indel mutations, at both the population and individual level, during Cas/gRNA-based gene drive.

Next-generation sequencing (NGS) methods provide unparalleled sequencing resolution. However, the cost and technical requirements associated with NGS prohibit routine testing and limit its use as a high-throughput method to assess indels[6,7,8]. Traditional PCR quantification methods have long been used as the standard evaluation procedure for genome indels; however, these methods are labor-intensive, take a long time to procure data, and have a high degree of variability. Digital-Droplet PCR (ddPCR) has been proven to be more sensitive at detecting mutations than Sanger sequencing in some applications and has a lower detection limit than NGS in others[6,7,8,9]. Moreover, the cost to assess a sample set and turn-around time for obtaining results is less expensive and faster, respectively, for ddPCR than either Sanger sequencing or NGS[9]. Using a dual-probe system, the Drop-Off assay identifies NHEJ alleles based on the absence of wild-type (WT) sequence at the gRNA-directed target Cas9 cut site. In this assay, a short amplicon including the predicted cut site of the Cas/gRNA-based system is amplified with a specific primer pair. One fluorescent probe is designed to bind to a conserved region of the amplicon and another fluorescent probe recognizes the WT sequence of the cut site. In the presence of an NHEJ allele, the latter will not bind to the amplicon.

The use of ddPCR provides the ability to design primers to target deletions, single base-pair differences and insertions, which will allow for NHEJ profiling in mosquito population analyses[9]. Given these attractive features, we created a protocol for ddPCR for high-throughput detection of indels

generated from a Cas/gRNA-based gene-drive system in mosquitoes.

Protocol

1. DNA extraction

1. Prepare EDTA/Nuclei Lysis Buffer (EDTA/NLS) with the ratio of 500 µL of NLS and 120 µL of EDTA per sample. Scale-up for multiple samples. Chill the mixture on ice.
   NOTE: The solution will turn cloudy in 2-5 min when chilled depending on the volume.

2. Homogenize the mosquito sample using a mechanical homogenizer for 10-15 s in a 1.5 mL microcentrifuge tube filled with 600 µL of chilled EDTA/NLS; mix thoroughly.

3. Add 17.5 µL of 20 mg/mL of Proteinase K to the tube and mix thoroughly.

4. Incubate overnight at 55 °C. Alternately, incubate the sample at 55 °C for 3 h with shaking and vortex the sample every 1 h.

5. Add 200 µL of Protein Precipitation Solution to the room temperature sample and vortex vigorously for 20 s.

6. Chill the sample for 5 min on ice.

7. Centrifuge the sample to pellet proteins at 15,890 RCF for 4 min.

8. Carefully aspirate the supernatant which contains the DNA and transfer it to a clean 1.5 mL microcentrifuge tube containing 600 µL of isopropanol.

9. Gently mix the solution by inverting the tube 5-10 times. Centrifuge for 1 min at 15,890 RCF. Carefully decant the supernatant while preserving the DNA pellet.

10. Add 600 µL of room-temperature 70% ethanol. Wash the pelleted DNA by gently inverting the tube.

11. Centrifuge for 1 min at 15,890 RCF. Carefully remove the supernatant by aspiration using a glass pipette tip.

12. Invert the tube onto clean absorbent paper and air-dry the pellet for 10-15 min.

13. Resuspend the DNA with PCR-grade water. Use 20 µL per individual mosquito sample or 100 µL for 10 pooled mosquitoes.

    NOTE: DNA extraction methods for mosquito samples using a commercially-available kit (see **Table of Materials**) are adapted from the manufacturer's Isolating Genomic DNA from Tissue Culture Cells and Animal Tissue protocol.

2. ddPCR reactions and droplet generation preparation

1. Quantify DNA using a fluorometer.

   NOTE: For the drop-off assay, it is recommended to use a range of 3,000-30,000 haploid genome copies per reaction, which is designed to detect NHEJ events with a HEX-labeling probe that binds to a WT sequence of the targeted cut site and will not anneal (drop-off) if there is a deletion or insertion at the target site, indicating the presence of an NHEJ variant.

2. Calculate the copy number using the haploid genome weight and concentration of DNA in the extract. This is done by multiplying the concentration of the extracted DNA by the volume used to obtain the total DNA mass, then dividing it by the haploid genome weight. Ensure that the volume added is between 1-10 µL. Dilute as necessary to be in the recommended haploid genome copy range.

   NOTE: One *Anopheles gambiae* haploid genome is estimated to be 0.27 pg per adult mosquito[10].

3. Design primers and probes. Design forward and reverse oligonucleotide primers with a Primer Melting Temperature (Tm) in the range of 55-60 °C that flank the 5'- and 3'-ends of the gRNA target site producing an amplicon of 150-400 bp.

   1. HEX (Hexachloro-fluorescein)-labeled probe for NHEJ detection: Design an oligonucleotide of ~15-20 bp in length complementary to the target site and add the HEX-probe to the 5'-end and BHQ1 (BLack Hole Quencher 1) to the 3' end. The Tm of the hydrolysis probe should be 3-10 °C higher than the Tm of the primers.

   2. FAM (6-carboxyfluorescein)-labeled probe for reference WT: Design an oligonucleotide ~15-20 bp in length complementary to a conserved genome site distant (about 25 bp) from the target site and add the FAM-probe to the 5'-end and BHQ1 to the 3' end. The Tm of the hydrolysis probe should be 3-10 °C higher than the Tm of the primers.

3. PCR reaction preparation

1. Prepare 25 µL of the ddPCR Sample Mix with the following components: ddPCR supermix for probes (no UTP): 12.5 µL, forward and reverse primers (10 µM): 1 µL each, HEX/FAM probes (10 µM): 0.625 µL each, DNA: 1-5 µL (3,750-37,500 haploid genome copies) and water: up to 25 µL.

2. Thoroughly mix the reactions by vortexing or reflux pipetting (up and down) (20x).

   NOTE: If the reactions are in a 96-well plate, pipette the entire volume up and down 20 times rather than vortexing to avoid bubble formation.

3. Briefly centrifuge the samples to settle the mixture at the bottom of the tube or well.

   NOTE: Ensure the reactions are at room temperature for the droplet generation. Prepare 1x of ddPCR mix for extra/unused wells in each cartridge (each cartridge has 8 wells).

4. Droplet generation

1. Using a 50 µL multichannel pipette, load 20 µL of the ddPCR sample mix into the middle row of the cartridge (**Figure 1A**, top).

2. Load 70 µL of the oil into the bottom row. Load 20 µL of 1x ddPCR supermix into unused wells.

   NOTE: Do not introduce bubbles.

3. Place the gasket touching only the edges, avoiding the center concaved area (**Figure 1B**).

4. Place the plate securely in the droplet generator and close the cover to start the run.

5. Using the multichannel pipette, transfer 40 µL of the emersion mix from the top row of the cartridge (**Figure 1A**, bottom) into the 96-well plate.

   1. Draw liquid sample for 3-5 s at a 45-30° angle. Expel the mixture slowly for over 3 s at a 45° angle into the side of the well, allowing it to drip down the side. It is okay to go to the second stop (complete expulsion) of the pipette to expel all the liquid.

6. Using foil heat seals, seal the plates for 5 s at 180 °C.

5. PCR

1. Place the sealed plate into the thermocycler (**Figure 1C**) and set the recommended PCR conditions if following the NHEJ Drop-Off guidelines as follows:

1. Initial denaturation at 95 °C for 10 min.

2. Set 40 cycles of 94 °C for 30 s to denature, 55 °C for 1 min to anneal, and 60 °C for 2 min to extend.

3. Hold at 98 °C for 10 min.

4. Hold at 4 °C.

   NOTE: The annealing temperature for specific primers and probe sets may be optimized using a thermal gradient. Use a ramp rate of 2 °C/s for all the steps. PCR conditions should be adjusted depending on each experimental design and set-up.

6. Droplet reading

1. Place the plate securely in the droplet reader with well A-1 at the top left (smoothed corner, the other three are edged) (**Figure 1D**).

2. Set up the plate in the program: Designating FAM as the known reference channel and HEX as the unknown one (**Figure 2A**).

3. Run the Droplet reading experiment as direct quantification. After the run finishes, change the Experiment type to **Drop-Off** (DOF) for the analysis (**Figure 2A**).

7. Analysis

1. Designate the correct experimental parameters (**Figure 2A**): **Sample Information**, **SuperMix**, **Target Name** (WT or NHEJ), **Target Type** (Ref or Unknown), **Signal Ch1** (HEX or FAM), **Signal Ch2** (HEX or FAM), and manually set the threshold for droplet count (recommend above 10,000 for reliable results). The software will perform most of the analysis with the designated parameter.

2. Check the droplet count in the **Droplet** tab; ensure that all are above 10,000 (**Figure 2B**).

3. Check 1 D amplitude for efficient signal separation from negatives.

4. In the **Plate Editor**, highlight the entire plate and set the **Experiment Type** to **Drop off**.

5. Set the WT target as a **Reference**, designating channel one for FAM and channel two for HEX.

6. Set NHEJ target as unknown, designating channel one for FAM and channel two as none.

7. In the 2D amplitude tab, set the **Cluster Thresholds** with the graph tools for each sample. Consider the tail associated with the WT cluster; this is normal for NHEJ assays (**Figure 3A**).

   NOTE: Under the **Ratio** tab, click on the **Gear Icon** from the top-right of the graph. Select the **Fractional Abundance**. The graph will now plot a point corresponding to the percentage of NHEJ events (**Figure 3B**).

Representative Results

An application of this procedure appears in Carballar-Lejarazú et al.[9]. The ddPCR Drop-off assay utilizes two fluorescent probes to discern WT and indel sequences: A FAM probe binds to a conserved sequence within the amplicon, whereas the HEX probe targets the WT sequence of the targeted site (**Figure 4A**). In the presence of an indel, the HEX probe will not bind. Representative results can be found in **Figure 2**, Table 1, and Table 2 of Carballar-Lejarazú et al.[9]. Using this protocol, ddPCR has been proven to detect a wide variety of CRISPR-Cas9 induced NHEJ events and quantify the NHEJ frequency in an individual or pooled sample. Fifteen different pooled samples of 10 mosquitoes each contained various NHEJ alleles (Table 2 of Carballar-Lejarazú et al.[9]). These were analyzed with ddPCR using the protocol and parameters presented here. Results from Table 1[9] show that all 15 samples carried 100% indel alleles as identified by the Drop-off assay (**Figure 4B**). In another experiment, 11 pooled samples of WT mosquitoes and NHEJ mosquitoes with different NHEJ percentages (0%, 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, and 100%) were examined with this ddPCR protocol, and the results (**Figure 2**; Carballar- Lejarazú et al.[9]) showed that the identified percentage is close to the compared technique of Indel Detection by Amplicon Analysis (**Figure 4C**).
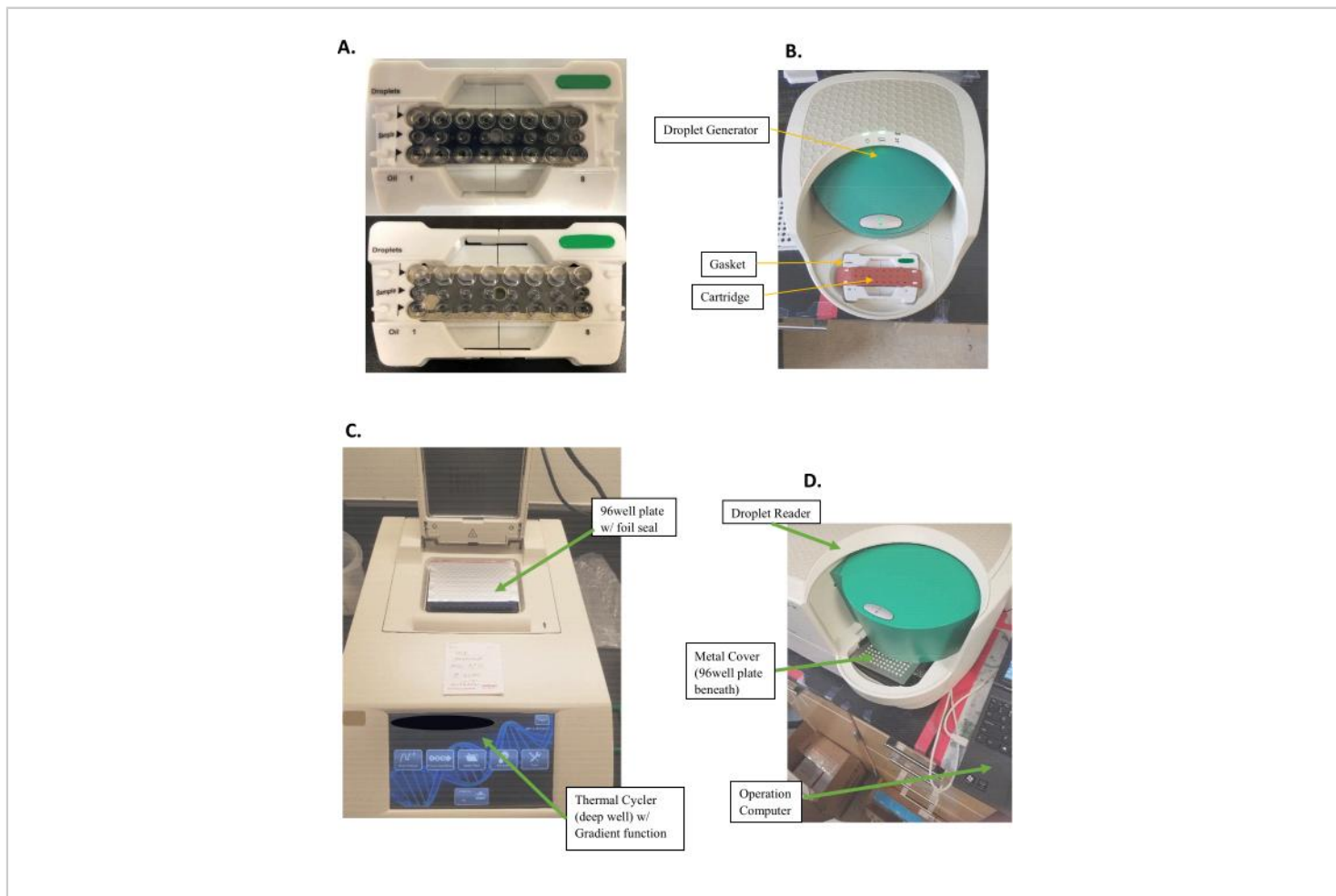
Figure 1: Experimental set-up and procedure. (**A**) Cartridge preparation for Droplet generation. (Top) Samples are filled in the middle row of the cartridge, while oil is filled in the bottom row. (Bottom) Top row filled with emulsified droplets after droplet generation. (**B**) Droplet generator with a cartridge filled with sample and covered by a gasket in place. (**C**) 96-well plate covered with foil seal in a Thermo-Cycler. (**D**) Droplet Reader with 96-well-plate in place with a metal cover latched over the plate to secure it. Please click here to view a larger version of this figure.
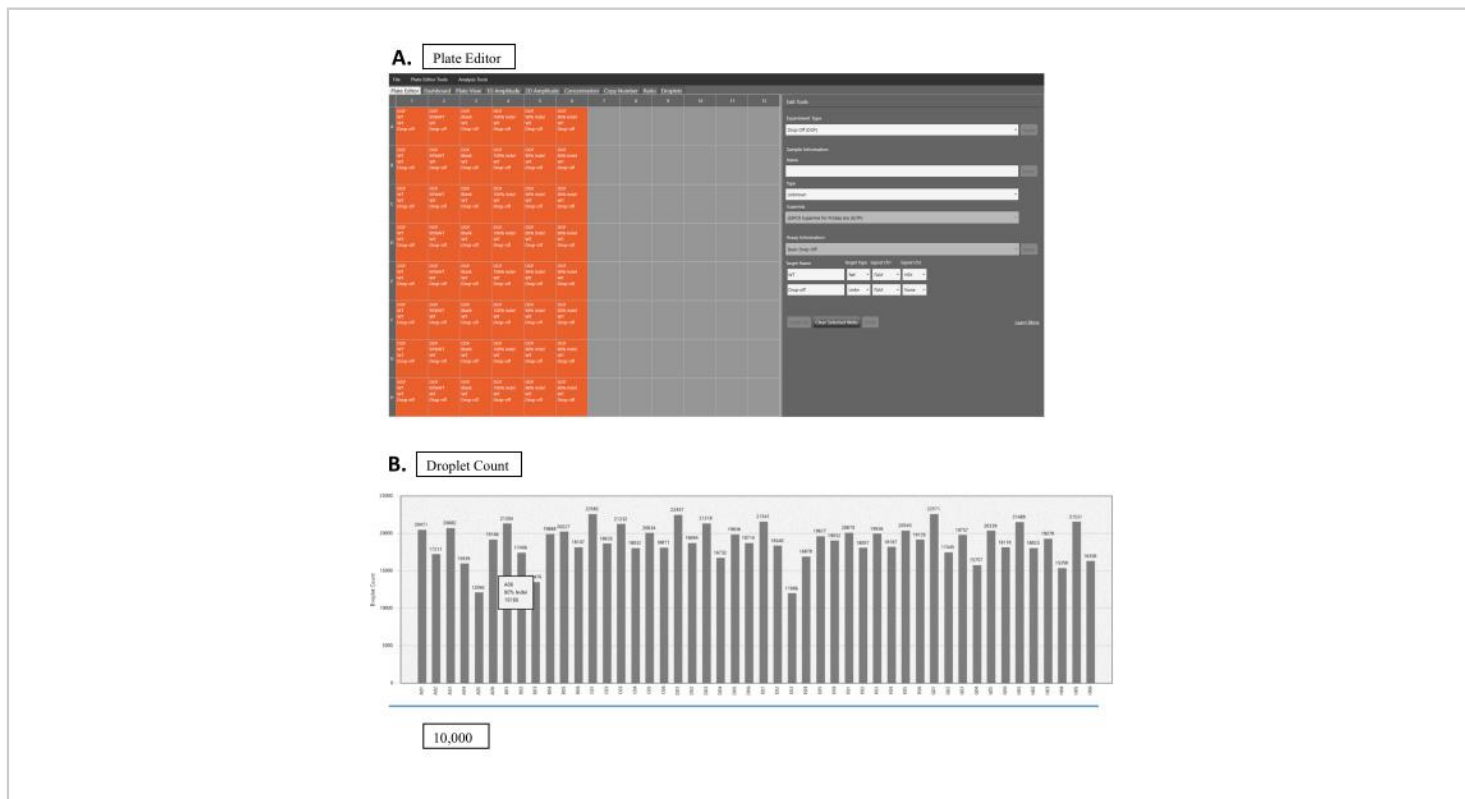
Figure 2: Droplet reading. (**A**) Software interface for droplet reading. Orange boxes show wells with samples. Gray boxes are empty wells. Experimental parameters are set up in the **Edit Tools** panel (right-hand side). Each sample can be edited by clicking on the respective sample box. Select **Drop Off** (DOF) for **Experimental Type**. In sample information, fill in the appropriate information for the sample's **Name** and **Type**, as well as **SuperMix**. Choose the **Basic Drop-Off** for the **Assay Information**. For the WT sample, choose WT for the **Target Name**, Ref for **Target Type**, and both FAM and HEX for **Signal Ch1** and **Ch2**, respectively. For NHEJ samples, fill in the appropriate name for the **Target Name**, choose Unknown for the **Target Type**, and choose FAM for **Signal Ch1**. Leave **Signal Ch2** at None. (**B**) Droplet count results for multiple samples. Please click here to view a larger version of this figure.
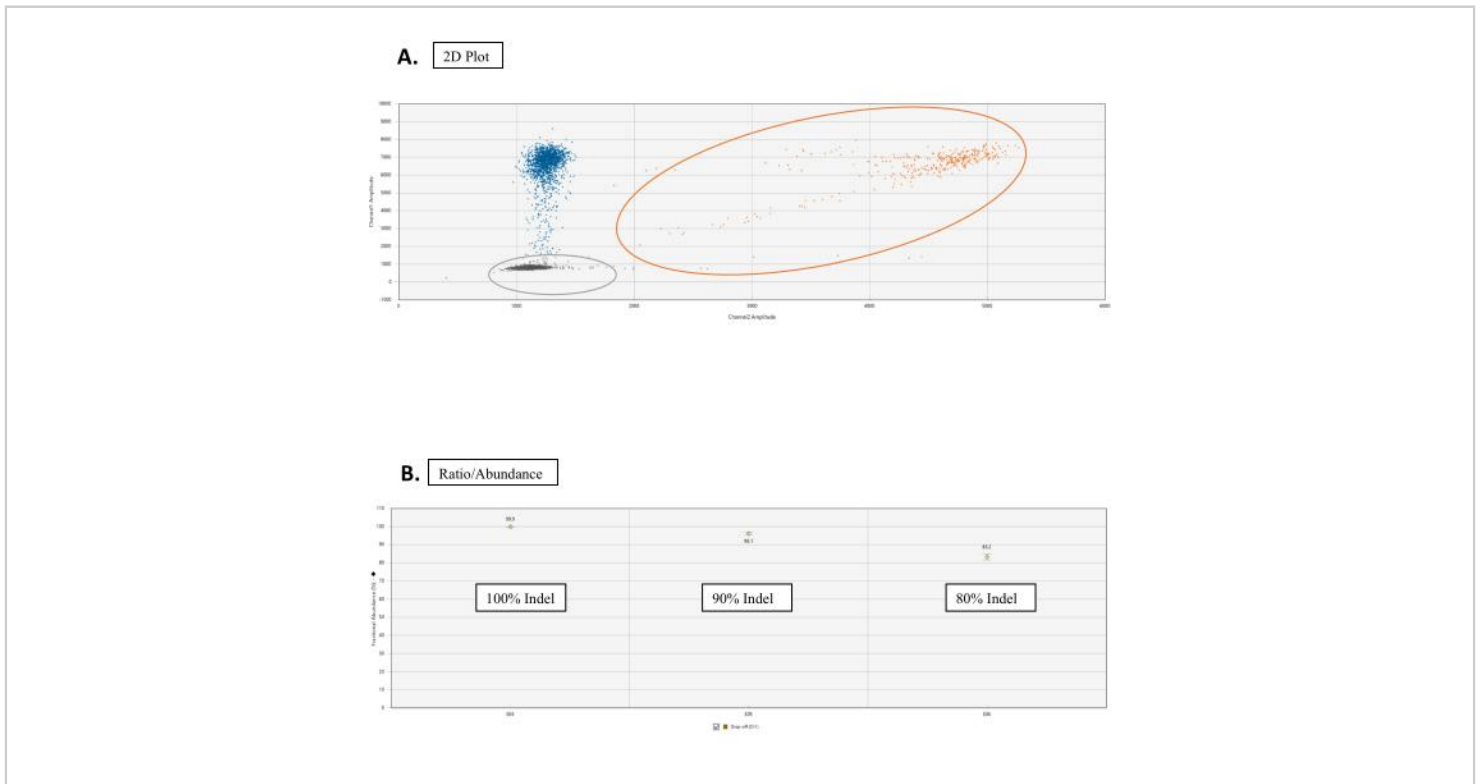
Figure 3: Drop-Off assay analysis. (**A**) Cluster 2D plot for the droplet count of the WT and NHEJ alleles. In the 2D Amplitude tab, all droplets are unclassified by default. In this figure, colors are manually assigned for distinguishing. The orange dots cluster are WT allele counts obtained by binding of both FAM and HEX probes at the reference sequence and target site sequence, respectively. Blue dots represent scores of droplets that have FAM binding to the reference sequence but no HEX binding at the target site sequence (hence drop-off of HEX). Gray dots are empty droplets that don't have either FAM or HEX binding. (**B**) Ratio/Abundance graphs of NHEJ events. Under the **Ratio** tab, select **Fractional Abundance** for a graph with the correspondent percentage of NHEJ events. Please click here to view a larger version of this figure.
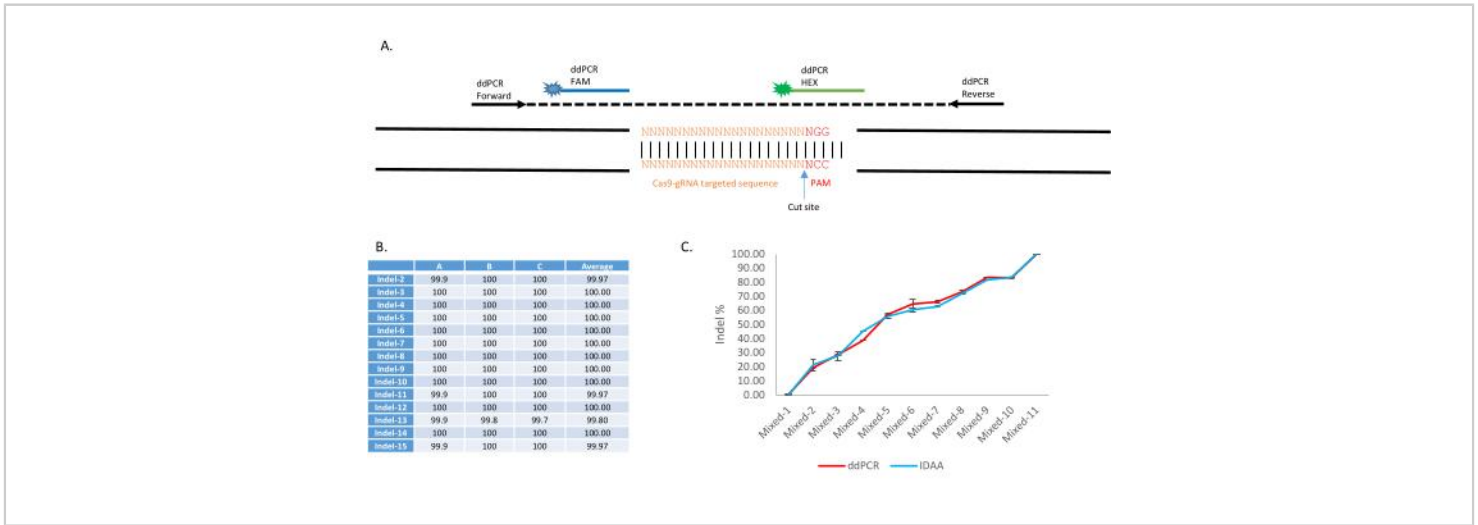
Figure 4: Application of Drop-Off assay with ddPCR for non-homologous end-joining identification and quantification in the transgenic *Anopheles stephensi* line, AsMCRkh1. (**A**) Schematic presentation of the ddPCR Drop-Off assay to detect mutations at a targeted DNA site with a dual-probe system. An amplicon of 150-400 bp is amplified with the forward and reverse primers. A FAM-labeled probe is designed to bind to a conserved sequence of the amplicon, whereas a HEX-labeled probe is designed to bind to the WT gRNA targeted site. (**B**) Detection of various types of indels with ddPCR. Fifteen pools of 10 AsMCRkh1 mosquitoes each containing various types of indel, including insertion, deletion, and substitution, were analyzed with the ddPCR Drop-Off assay. Details of mutations and sequences can be found in Table 2 and Table S3 of Carballar *et al.*[9]. (**C**) Quantification of NHEJ in mixed samples of AsMCRkh1 and WT mosquitoes with various ratios (10:0, 9:1, 8:2, 7:3, 6:4, 5:5, 4:6, 3:7, 2:8, 1:9, and 0:10) using ddPCR and a compared technique of Indel Detection by Amplicon Analysis[9]. Images adapted from Carballar-Lejarazú et al. *Biotechniques*. 68(4):172-179 (2020)[9]. Please click here to view a larger version of this figure.

Primer/Probe	Sequence (5' » 3')
ddPCR Forward Primer	ATGATCAAATGTCGACCG
ddPCR Reverse Primer	ACCGTACTGGTTGAACA
ddPCR HEX Probe (BHQ1)	[HEX]-TTCTACGGGCAGGGC-[BHQ1]
ddPCR FAM Probe (BHQ1)	[6FAM]-CCACGTGGGATCGAAGG-[BHQ1]
HEX: Hexachloro-fluorescein, FAM: 6-carboxyfluorescein, BHQ: Black Hole Quencher	

Table 1: Sequences of primers and probes.

Discussion

Digital-droplet PCR is an efficient method to determine the presence of indel alleles resulting from NHEJ events in a Cas/gRNA-based gene-drive system and allows for quantification of the frequency of these alleles in individuals or populations. Some steps of the protocol need to be followed with special care to achieve reliable results. Firstly, the genomic DNA extraction needs to be performed carefully to ensure high quality and sufficient quantity. A good extraction will allow accurate determination of the haploid genome copies per reaction. In our experience, a commercially-available kit (see **Table of Materials**) provided consistently high-quality DNA extractions. However, extractions of individual mosquitoes can prove to be particularly challenging as the DNA pellet becomes hard to visualize and can easily be sucked up with the supernatant if not careful. Secondly, primers and probes must be designed carefully. Before completing the ddPCR experiment, ensure that the designed primers result in a single PCR product by first performing a traditional PCR and visualizing a single product via gel electrophoresis. The reference FAM probe also must be designed so that it is complementary to a highly-conserved sequence. This will ensure accurate detections of WT alleles throughout a diverse population. The primer/probe combinations for each unique experiment will have different thermocycler conditions, and it is recommended to optimize those conditions using a thermal gradient.

Other methods for identifying indels exist, such as Sanger sequencing or NGS. Sanger sequencing is limited because it has a lower limit of detection and low discovery power to identify novel variants. Sanger sequencing also is labor-intensive and is not high-throughput. Compared to Sanger sequencing, NGS does not have the same limitations of low sensitivity, discovery power, and throughput. Another benefit of NGS is its ability to detect a variety of mutations from single nucleotide polymorphism (SNPs) to rearrangements. However, NGS is a more costly and time-consuming method in the application of determining Cas9/gRNA-associated indels because there is only one target region of interest, and it is best suited for larger genome-wide analyses. Compared to the aforementioned methods, ddPCR is high throughput and has a quick turn-around time. If the ddPCR materials and instruments are available in-house, 96 samples can be processed within 1-2 days, making it well suited for quick analysis of large trials of Cas9/gRNA -modified organisms.

While many benefits exist for ddPCR there are also limitations. Firstly, the ddPCR equipment is not frequently available in an independent laboratory environment. ddPCR equipment may be available communally at larger research institutions, but this does not allow for ease of data generation and analysis outside of the institution. Secondly, unlike the alternatives, ddPCR does not provide the individual unique sequences of identified indel mutations. Digital-droplet PCR will provide the frequency of indel mutations within a population, but without the sequence, one cannot determine whether the indels present are more likely to conserve or inhibit the function of the gene of interest. The ddPCR method is perhaps suited best to analyze wild populations after a field release trial of a Cas9/gRNA-based drive organism because it can efficiently determine the frequency of introduction of the transgene into the native population and the generation of indels within the population in close to real-time. Due to the ddPCR quick turn-around time, it would be feasible to perform sampling and analyze the population in a field trial region weekly if materials were available locally. The start-up costs to purchase, import, and set up the ddPCR equipment would be high in remote laboratories but the benefits of being

able to assess rigorously a wild population as it is undergoing modification from a drive system would justify the costs.

Disclosures

The authors have no disclosures.

Acknowledgments

References

1. Gantz V. M. et al. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi. Proceedings of the National Academy of Sciences of the United States of America.* **112** (49), E6736-743 (2015).

2. Hammond, A. et al. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae. Nature Biotechnology.* **34**, 78-83 (2016).

3. Courtier-Orgogozo V., Morizot B., Boëte C. Agricultural pest control with CRISPR-based gene drive: time for public debate: Should we use gene drive for pest control? *EMBO Reports.* **18** (6), 878-880 (2017).

4. Carballar-Lejarazú, R. et al. Next-generation gene drive for population modification of the malaria vector mosquito, *Anopheles gambiae. Proceedings of the National Academy of Sciences of the United States of America.* **117** (37), 22805-22814 (2020).

5. Adolfi, A. et al. Efficient population modification gene-drive rescue system in the malaria mosquito *Anopheles stephensi. Nature Communications.* **11**, 5553 (2020).

6. Wang, Z. et al. Comparison of droplet digital PCR and direct Sanger sequencing for the detection of the *BRAF*^{V600E} mutation in papillary thyroid carcinoma. *Journal of Clinical Laboratory Analysis.* **33** (6), e22902 (2019).

7. Dong, L., Wang, S., Fu, B., Wang, J. Evaluation of droplet digital PCR and next generation sequencing for characterizing DNA reference material for *KRAS* mutation detection. *Scientific Reports.* **8**(1), 9650 (2018).

8. Bell, C. C., Magor, G. W., Gillinder, K. R., Perkins, A. C. A high-throughput screening strategy for detecting CRISPR_Cas9 induced mutations using next-generation sequencing. *BMC Genomics.* **15** (1), 1002 (2014).

9. Carballar-Lejarazú, R., Kelsey, A., Pham, T. B., Bennett, E. P., James, A. A. Digital droplet PCR and IDAA for the detection of CRISPR indel edits in the malaria species *Anopheles stephensi. Biotechniques.* **68** (4), 172-179 (2020).

10. Holt, R. A. et al. The genome sequence of the malaria mosquito *Anopheles gambiae. Science.* **298** (5591), 129-149 (2002).