

ABSTRACT

MOORE, DAVID LEWIS. Accuracy and Sensitivity Comparison of ddPCR and qPCR for mtDNA. (Under the direction of Dr. Kelly Meiklejohn).

In forensic casework, analysis of nuclear DNA (nuDNA) is the gold standard to permit individualization. However, nuDNA is susceptible to degradation, and thus compromised samples containing low quantity and quality nuDNA are often encountered in forensic casework. In such cases, analysis of mitochondrial DNA (mtDNA) may be more successful given its high copy number even in highly degraded samples. While mtDNA analysis will not determine the individual to whom the questioned DNA belongs, it can be used to exclude individuals to whom their mtDNA does not share enough similarity with the questioned sample. Typical sample types that would require mtDNA analysis are bones, teeth, root-less hair, and samples with very degraded nuDNA. In order to adhere to the Federal Bureau of Investigation (FBI) quality assurance standards and to ensure optimal results, forensic laboratories are required to quantify the amount of DNA present in the sample before downstream analysis. DNA quantification is most routinely completed using quantitative PCR (qPCR). While several commercial nuDNA qPCR assays exist, no such assay exists for mtDNA. To address this gap, Dr. Kavlick at the FBI developed a human mtDNA qPCR triplex assay that amplifies both short (105 bp of ND5) and long (315 bp of 16s rRNA) mtDNA targets (Kavlick et al. 2019). By performing qPCR with primers and probes for these two targets, one can not only quantify the amount of mtDNA present initially in the sample but also determine the degree to which a sample may be degraded (based on the amplification of the short target compared to the long target). Quantification of an unknown is achieved by the triplex assay through comparison to a synthetic 105 bp oligonucleotide standard (dsT8sig; comprising the short target only) of known copy numbers.

While this qPCR assay has been demonstrated to be accurate and reproducible¹, sensitivity could be improved using new quantification approaches such as droplet digital PCR (ddPCR). ddPCR provides absolute quantification of target DNA templates in a sample via emulsion droplet generation and reading and offers several advantages over qPCR, including that a standard (and subsequent standard curves) are not required for copy number determination. In addition, ddPCR is less affected by the presence of PCR inhibitors including those commonly associated with forensic evidence (e.g., melanin and humic acid), as each DNA strand undergoes PCR in an individual droplet and not a bulk solution. In this study, the dsT8sig standard from Kavlick et al. (2011) as well as a new custom synthetic double-stranded gBlock[®] gene fragment standard (containing both the short and long targets) were used to quantify mtDNA. A side-by-side comparison between qPCR and ddPCR was completed using three types of samples (buccal, cell lines and environmental) containing DNA of varied quantity and quality. By comparing sensitivity and accuracy between qPCR and ddPCR, the overall goal is to help develop a more robust mtDNA quantification protocol for forensic purposes.

© Copyright 2022 by David Moore

All Rights Reserved

Accuracy and Sensitivity Comparison of ddPCR and qPCR for mtDNA

by
David Lewis Moore

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Biochemistry

Raleigh, North Carolina

2022

APPROVED BY:

Dr. Kelly Meiklejohn
Co-Chair of Advisory Committee

Dr. Melanie Simpson
Co-Chair of Advisory Committee

Dr. Josh Strable

DEDICATION

I dedicate this thesis primarily to my partner, Riley Schoonover. Without her support throughout this I would not have been able to focus my full attention and give my best effort to my work. Her kindness, compassion, and willingness to help know no bounds, and I am infinitely grateful to have her by my side. I also dedicate this to my parents, Dennis and Myra Moore, as a testament to the fact that my accomplishments are directly because of their openness, patience, and warmth during my upbringing. Also my little dog, Ellie, for being a source of joy (and sometimes annoyance, but the good kind) during this process.

BIOGRAPHY

I grew up in Lenoir, North Carolina, where I found a passion for science at an early age. I knew that I wanted to use science to help people and that led me to NC State University. I am currently concluding my Master of Science in Biochemistry degree at NC State University, where I also attended for my undergraduate career and graduated with a Bachelor of Science in Biochemistry degree in 2021. During my time at NC State University, I participated as a member, mentor, and ambassador in the Goodnight Scholars Program. I also conducted research starting in 2020 in the Meiklejohn lab and continued to do so when implementing my thesis project. When not in school or the lab, I love to read and exercise.

ACKNOWLEDGMENTS

I would foremost like to thank Dr. Kelly Meiklejohn, without whose direction I would surely be lost. Additional thanks to my other defense committee members, Dr. Josh Strable and Dr. Melanie Simpson. Also, I'd like to thank Melissa Scheible and Teresa Tiedge in the Meiklejohn Lab for their support during and outside of my thesis project. I would also like to thank Dr. Mark Kavlick at the FBI, for without his original work I would have had no basis for this project, and whose correspondence during was extremely helpful. There are many others who I was able to get assistance from in some way or another, but for brevity I will issue a general acknowledgement of thanks to all involved. If you helped me at all during this time, I greatly appreciate you.

TABLE OF CONTENTS

LIST OF TABLES	vi
LIST OF FIGURES	vii
Chapter 1: qPCR to ddPCR Transition and Comparison	1
Introduction.....	1
Materials and Methods.....	6
DNA Sources	6
Primers and Probes	6
Standards.....	7
qPCR.....	8
ddPCR.....	9
Calculation of Variation.....	10
Results and Discussion	10
qPCR Assay Replication.....	10
Optimization and gBlocks.....	14
Transitioning to ddPCR	16
qPCR and ddPCR Comparison	19
Conclusions.....	23
References.....	24
 APPENDIX.....	 28

LIST OF TABLES

Table 1	qPCR Primers and Probes	7
Table 2	ddPCR Master Mix Components	9
Table 3	Mean short Ct value comparison for dsT8sig and gBlock standards versus Kavlick's data	11
Table 4	ddPCR and qPCR (gBlock and Oligonucleotide) concentration	21
Table 5	Coefficient of Variation (CV) for sample types on ddPCR and qPCR	22
Table 6	Statistical Significance of Concentration Difference.....	23

LIST OF FIGURES

Figure 1	Amplification Plot of Short Standards Run.....	12
Figure 2	Standard Curve of Short Standards Run.....	13
Figure 3	Amplification Plot of Short gBlock Standards Run	15
Figure 4	Standard Curve of Short gBlock Standards Run	16
Figure 5	ddPCR Amplification Plot for Short gBlock.....	17
Figure 6	ddPCR Amplification Plot for Long gBlock.....	19
Figure 7	Comparison of ddPCR and qPCR quantification across sample types	20

CHAPTER 1: qPCR to ddPCR Transition and Comparison

Introduction

Most biological samples submitted as evidence to forensic laboratories contain nuclear DNA (nuDNA), that is, DNA found in the nucleus of the cell, of which there are only two copies per cell. In nuDNA analysis, total genomic DNA (both nuclear and mitochondrial) is extracted from biological samples and purified before nuDNA is analyzed. In forensic casework, 20 short tandem repeat (STR) regions of the nuclear genome that permit individualization are amplified via polymerase chain reaction (PCR). Once sufficient copies of these STRs have been generated, the underlying genotype profile is deduced via separation of alleles using capillary electrophoresis and subsequently compared either to a) evidence sample belonging to the suspect, or b) the Combined DNA Index System (CODIS) which contains STR profiles taken from a crime scene or collected from convicted offenders². When nuDNA of sufficient quantity and quality is available and there is a known reference profile to compare it to, identification of samples is not usually difficult. In fact, CODIS has been responsible for helping solve over half a million criminal cases since its inception more than two decades ago².

However, some biological samples collected either contain too little nuDNA or contain nuDNA that has been severely degraded, which limits successful STR analysis. When that happens, alternate methods of identification are utilized. Mitochondrial DNA (mtDNA) is one such alternate method in a forensic laboratory's toolbelt. The human mtDNA is 16,569 bp double-stranded circular molecule found in the mitochondria of cells, with usually 2,000-10,000 copies of mtDNA per cell³. Unlike nuDNA, mtDNA is only inherited maternally, meaning that all offspring have identical mtDNA as their mothers. This occurs because upon fertilization, the

mtDNA in sperm is tagged by the egg for destruction⁴. As such, mtDNA cannot be used to provide individual identification but instead help exclude or include an individual from a group of individuals. Since individuals descended from the same maternal line share mtDNA profiles, ancestry and familial matching services often use mtDNA to help others discover family members.² Forensic mtDNA typing typically includes sequencing a region of the mtDNA genome known as the control region or the D-loop⁵. This sequence contains two sections known as hypervariable (HV) regions 1 and 2. These regions are non-coding and are, as the name suggests, some of the most variable regions in the mtDNA genome. mtDNA analysis involves examining the control region sequence for single nucleotide polymorphisms (SNPs) to identify the haplogroup of the mtDNA and compare it to a reference. A reference sample is always needed during analysis, whether obtained from a direct relative or compared against a large mtDNA database like the European DNA Profiling Group's Mitochondrial DNA Population Database (EMPOP)⁶. Limitations of Sanger sequencing has meant that traditionally mtDNA typing requires amplification and sequencing of ~300 bp regions of the control region, but new sequencing technologies allow the entire control region to be sequenced at once. mtDNA has been used as a tool in several high-profile cases, including establishing a possible source of the remains of the famed astronomer Nicolaus Copernicus⁷.

With the advent of new sequencing technologies (also known as massively-parallel sequencing (MPS)), the standard of mtDNA typing may change to include the entire mtDNA genome for use in identification. However, even MPS methods require the entire mtDNA genome to be broken up into amplicons roughly 100-400 bp for sequencing on current short-read platforms commonly available in forensic laboratories⁸. Longer amplicons can be read, but the reads are not as robust

as short-read technologies. Amplicons are short generally to ensure highly efficient PCR amplification occurs to produce an accurate sequence with little variation. However, in cases of extreme degradation as can be common with highly processed forensic samples (e.g., burned bones, teeth from mass disasters, repatriated remains), amplification of even moderately small amplicons (>300 bp) may be impossible. In that case, shorter amplicons (~100 bp) would need to be used to amplify sections of the mtDNA that were not degraded. Of course, one cannot know the extent of degradation present in a biological sample without first analyzing it. In order to meet the FBI's quality assurance standards and to ensure optimal results, forensic laboratories are required to quantify the amount of DNA present in a biological sample before analysis. Currently the most common method of DNA quantification is quantitative PCR (qPCR), which both amplifies DNA from the target genome (nuclear or mitochondrial) and provides the copy number of the target relative to a DNA standard. To assist in casework where amplification of mtDNA would be necessary, Dr. Mark Kavlick of the FBI developed a human mtDNA triplex quantitation qPCR assay⁹.

This assay has several advantages over others, including its high species specificity, high accuracy, and reproducibility. It also includes an exogenous internal positive control (IPC) that would not be found in a biological sample if contamination is ever questioned. The IPC, under normal qPCR conditions, serves as proof that the reaction happened as it should amplify regardless of the quality of the DNA sample. In the presence of inhibitors, the lack of IPC amplification can alert one to inhibition¹⁰. The mtDNA targets chosen for the triplex assay were in the ND5 gene (105 bp) and the 16S rRNA gene (316 bp). The targets chosen were outside of the normally sequenced mtDNA regions like the control region, which later studies have found is

a poor target region for quantification purposes¹¹. By amplifying two targets of differing lengths (105 bp and 316 bp), one can compare the differences in cycle threshold (Ct) values between the standard (or ‘calibrator’) and the sample to determine the relative degree of mtDNA degradation. From Kavlick’s work, the formula for determining the degree of degradation is as follows:

$$\Delta\Delta Ct = \Delta Ct_{sample} - \Delta Ct_{calibrator} = (Ct_{316} - Ct_{105})_{sample} - (Ct_{316} - Ct_{105})_{calibrator}$$

Where values of $\Delta\Delta Ct$ of 1 would mean that there were roughly 2¹ or 2 times as many short targets as long, which Kavlick classifies as a low degree of degradation. $\Delta\Delta Ct$ s of 4 would indicate 16 times as many short targets as long—what Kavlick classifies as “higher state of sample degradation⁸.” Using the $\Delta\Delta Ct$ values that Kavlick provides the formula for, one can also determine the relative copy number of other parts of the mtDNA genome simply based on their relative size. For instance, the size of the entire control region is roughly 1,228 bp. According to Kavlick, one can calculate the copy number for any larger region by applying the following formula:

$$copy\ number_{large\ region} = copy\ number_{105} / (2^{\Delta\Delta Ct})^{(large\ region\ length / 316)}$$

This represents a remarkable way to calculate the copy number of desired targets for mtDNA sequencing based just on the Ct values of a short and long target. However, there are some drawbacks to qPCR in general that make this method less accurate than it could be. Firstly, qPCR is a relative quantification method, meaning that quantification is based on a standard curve using a dilution series of a standard of known concentration. This means that any errors in standard preparation or dilution result in an inaccurate standard curve and thus an inaccurate quantification of unknown samples. While uncommon, inaccuracies in standard curve dilutions can occur depending on the skill and experience of the technician diluting them. One alternative PCR method, which has gained popularity for use in quantification in recent years, is droplet

digital PCR (ddPCR). ddPCR has several distinct advantages over qPCR, the primary advantage being that ddPCR is an absolute quantification method. This means that ddPCR can determine the exact copy number of a target present in a sample without the use of a standard curve. ddPCR can utilize the same probes and primers that typical qPCR does— what differs is where each PCR reaction takes place. In ddPCR, DNA templates, polymerase, dNTPs, reaction mix and oil are combined to form an emulsion and partitioned into around 20,000 nanoliter droplets per sample, with each droplet containing either at least 1 or zero copies of DNA¹². PCR then proceeds within each droplet and a droplet reader later reads each droplet after amplification is completed and measures fluorescence to determine whether a droplet does or does not contain amplified target DNA. By using Poisson statistics after determining the number of positive and negative droplets in the sample, ddPCR can accurately determine the copy number of target DNA in the original sample. Whereas methods like qPCR rely on measuring fluorescence after each amplification cycle, ddPCR simply measures the fluorescence at the end of the entire process. This can generally result in slightly more efficient reactions and more accurate results than in qPCR¹². However, the biggest contributor of ddPCR accuracy relies on the fact that its quantification methods do not rely on extrapolating data from Ct values or a standard curve. In addition, ddPCR's use of massive partitioning helps prevent inhibitor effects on PCR, as inhibitors can only affect the PCR reaction of the droplet they inhabit instead of the entire bulk reaction as in qPCR. However, one drawback of ddPCR is that while it is more sensitive than qPCR, its dynamic range is lower, meaning that higher concentrations of DNA can completely saturate the droplets and fall outside the range of linear detection. Nevertheless, ddPCR is consistently used by different fields (cancer studies, pathogen identification, clinical diagnostics) for low copy number quantitative purposes and has generally been recognized as cheaper, more efficient, and

more accurate than qPCR for quantification¹³. As such, this study was focused on a) transitioning Kavlick's existing mtDNA qPCR triplex assay to ddPCR, and b) compare the accuracy and sensitivity of the two methods.

Materials and Methods

DNA Sources

For initial qPCR optimization of the mtDNA triplex assay, three commercially available DNA sources were used: K562 (Biochain, Newark, CA), Raji (Biochain), and Jurkat (Biochain). The primary sources of DNA used for ddPCR and qPCR comparison fall into one of three categories based on source: bloods, environmental indoor dusts, and buccal samples. 8 individual samples were chosen from each category and were run in duplicate in qPCR and singularly in ddPCR. Blood samples were purchased from Biochain. The indoor dust and buccal samples were collected under an approved NC State Approved IRB Protocol #14375.

Primers and Probes

The primers designed and published by Kavlick et al 2019 were synthesized and HPLC-purified by Integrated DNA Technologies (IDT, Coralville, IA)⁹. Each primer was reconstituted in IDTE (IDT) to 100 μ M. Probes contained a 3' minor groove binding non-fluorescent quencher (MGB-NFQ) and were synthesized by ThermoFisher (Waltham, MA); the short target was 5' labeled with the FAM dye and the long target probe was 5' labeled with the HEX dye. The sequences for the primers and probes are as shown in **Table 1** below.

Table 1. qPCR Primers and Probes. Sequences, concentrations, and volumes of primers and probes used for short and long target quantification. Most of this data was taken from Kavlick et al. 2019.

Target/ Reagent	Region/ Length	Primer/ Probe	Sequence (5'-3')	Stock Concentration (μ M)	Volume used in qPCR (μ L)	Final Concentration (μ M)
Short	13,288- 13,392 (ND5 gene)	Forward Primer	GGC ATC AAC CAA CAC CTA	1.25	0.8	50
		Reverse Primer	ATT GTT AAG GTT GTG GAT GAT GGA	22.5	0.8	900
		Probe (FAM)	CAT TCC TGC ACA TCT G	6.25	0.8	250
Long	2,332- 2,647 (16S rRNA gene)	Forward Primer	YGC ATA AGC CTG CGT CAG AT	7.5	0.8	300
		Reverse Primer	CCC TCG TGG AGC CAT TCA TA	22.5	0.8	900
		Probe (HEX)	AAC ACA GGC ATG CYC	6.25	0.8	250

Standards

Two different types of standards were used in these experiments: 1) A new custom double-stranded synthetic gBlock[®] gene fragment (IDT) designed during this study. The gBlock was 491 bp double-stranded DNA fragment that contains the sequence from both the short (ND5) and long (16S) targets in the triplex assay (**Table 1**), each bookended by 15 random bp and separated in the middle by ten 'T' nucleotides. 2) The custom 115 bp dsT8sig synthetic paired oligonucleotides developed in Kavlick et al 2011 (from IDT), which contained just the short target. Sequences for these can be found in **Appendix Table 1**.

Each standard had to be reconstituted to 100 μM in IDTE (IDT) prior to use. To enhance the annealing of the two complimentary dsT8sig strands, they were reconstituted in duplex buffer (IDT) and heated to 94°C for 2 minutes. The dsT8sig oligos were diluted down to 2 μM , combined and annealed to form a 1 μM stock, and then the copy number was determined via NanoDrop One quantification (ThermoFisher) using the extinction coefficient for the oligos at 260 nm. Using the molar extinction coefficient for each strand separately (each provided by the manufacturer) and then the extinction coefficient of the double-stranded product, the concentration of DNA in nanograms per microliter was determined. From there, the following calculation was applied:

$$\text{Number of Copies}/\mu\text{L} = \frac{\text{DNA Concentration (ng}/\mu\text{L}) \times 6.022 \times 10^{23}}{\text{Molecular Weight (g/mol)} \times 1 \times 10^9}$$

The dsT8sig oligos were then diluted to roughly 10^{10} copies/ μL . The gBlocks were reconstituted in 50 μL of IDTE, heated to 50°C for 20 minutes, and brought back to room temperature. They were then quantified on the NanoDrop One to determine their concentration (as described above), and then diluted to roughly 10^9 copies/ μL . In order to reduce contamination from the higher-quantity standard dilutions, serial dilutions were performed the day prior to use in downstream quantification, and all dilutions occurred in sterile O-ring screw top 2 mL tubes.

qPCR

qPCR was performed on a QuantStudio™ 5 Real-Time PCR System for Human Identification (ThermoFisher, Waltham, MA) instrument in 20 μL reactions using 10 μL TaqMan 2X Fast Advanced Master Mix (Applied Biosystems), 3.2 μL molecular-grade water, and 2 μL of DNA template. Primers and probe concentrations and characterization can be found in **Table 1**. The

following cycling conditions taken directly from Kavlick et al. 2019⁹ were used for this assay: 2 minutes at 50°C and 20 seconds at 90°C followed by 40 cycles of 3 seconds at 95°C and 30 seconds at 60°C. For qPCR a total of eight dsT8sig oligo or gBlock standards were included with each plate, with copy number ranging from 10⁸ to 10¹ copies/μL.

ddPCR

ddPCR was performed on a BioRad T100 thermal cycler (BioRad, Hercules, CA), using 12 μL ddPCR Supermix for Probes (no dUTP; BioRad), 7.2 μL molecular-grade water, and 2.4 μL of DNA template. Primer and probe concentrations can be found in **Table 2**. Droplets were generated using a QX200 Droplet Generator (BioRad) by adding 70 μL droplet generation oil and 20 μL master mix with DNA. The following cycling conditions were used on this assay: 95°C for 10 minutes followed by 40 cycles of 94°C for 30 seconds and 60 °C for 1 minute concluding with an enzyme deactivation step of 98°C for 10 minutes.¹⁵ Due to ddPCR's lower dynamic range, only five standards were used when using either gBlocks or dsT8sig oligos, ranging from 10⁵ to 10¹ copies/μL. The master mix was prepared as shown in **Table 2**. Droplets were read on a BioRad QX200 droplet reader (BioRad) using 20 μL reactions.

Table 2. ddPCR Master Mix Components. This contains a list of the concentrations of primers and probes used per ddPCR reaction. While the total volume is 24 μL, only 20 μL was used to generate droplets.

Reagent	Stock Concentration(μM)	Volume used in qPCR (μL)	Final Concentration(nM)
Short/Long Forward Primer	10	0.9	375
Short/Long Reverse Primer	10	0.9	375
Short/Long Probe	10	0.6	250

Calculation of Variation

When both the qPCR and ddPCR experiments were completed, a side-by-side comparison was necessary to compare the precision of each assay. To do this the coefficient of variation (standard deviation over the mean multiplied by 100) was calculated for each assay. Additionally to assess whether a statistical significance in mtDNA copy number was obtained between qPCR and ddPCR for each sample, a matched-pairs *t*-test was implemented.

Results and Discussion

qPCR Assay Replication

In order to transition the mtDNA qPCR triplex assay⁹ to ddPCR, it was first necessary to replicate the assay on the author's qPCR lab instrument. Considering other labs had reported issues in replicating the assay (*pers. comms.* with Dr. Kavlick), it was prudent to perform initial testing and optimization. While Kavlick et al. used a 7900HT FAST Sequence Detection System from ThermoFisher as well as fluorescent probes obtained from Applied Biosystems, qPCR was performed on a QuantStudio™ 5 from ThermoFisher using fluorescent probes obtained from Integrated DNA Technologies (IDT)⁹. After optimization and testing done, the Ct values that we obtained for the short dsT8sig Oligo standards were compared with what Kavlick et al. obtained in 2011. Notably, Ct comparisons for the long target were not possible given these were not reported by Kavlick in his 2019 paper.

Table 3. Mean short Ct value comparison for dsT8sig and gBlock standards versus Kavlick et al 2011. Kavlick's data was collected for 15 independent standard lots and runs, whereas the comparisons from this study (denoted by *) were generated from a single experiment with samples run in-duplicate.

Calculated Standard Copy Number(Copies/ μ L)	Kavlick et al. (2011). Mean dst8sig Short Ct	dsT8sig Mean Short Ct*	gBlock Mean Short Standards Ct*
10^8	13.2 ± 1.2	12.4 ± 0.026	12.4 ± 1.63
10^7	16.1 ± 1.1	16.3 ± 0.242	16.4 ± 0.16
10^6	20.0 ± 1.1	20.1 ± 0.045	20.1 ± 0.273
10^5	23.5 ± 0.9	23.9 ± 0.109	23.0 ± 0.023
10^4	26.9 ± 1.2	27.8 ± 0.124	27.0 ± 0.472
10^3	30.4 ± 1.5	31.5 ± 0.021	30.5 ± 0.194
10^2	33.6 ± 1.7	35.1 ± 0.101	34.5 ± 0.113
10^1	37.0 ± 2.3	39.1 ± 0.400	37.5 ± 0.165

The Ct values for the short target within the dsT8sig standard in this study fell within the standard deviation range of values that Kavlick et al. 2011 reported (**Table 3**). This result indicated that the attempted assay was producing results within acceptable ranges. While viewing the Ct values allowed for quick comparison to Kavlick et al. 2011's data, it was necessary to get a good representation of the reaction as a whole. For this reason, amplification plot for the short dsT8sig standard run was examined to look for overall dilution series separation as consistency among sample replicates (**Figure 1**).

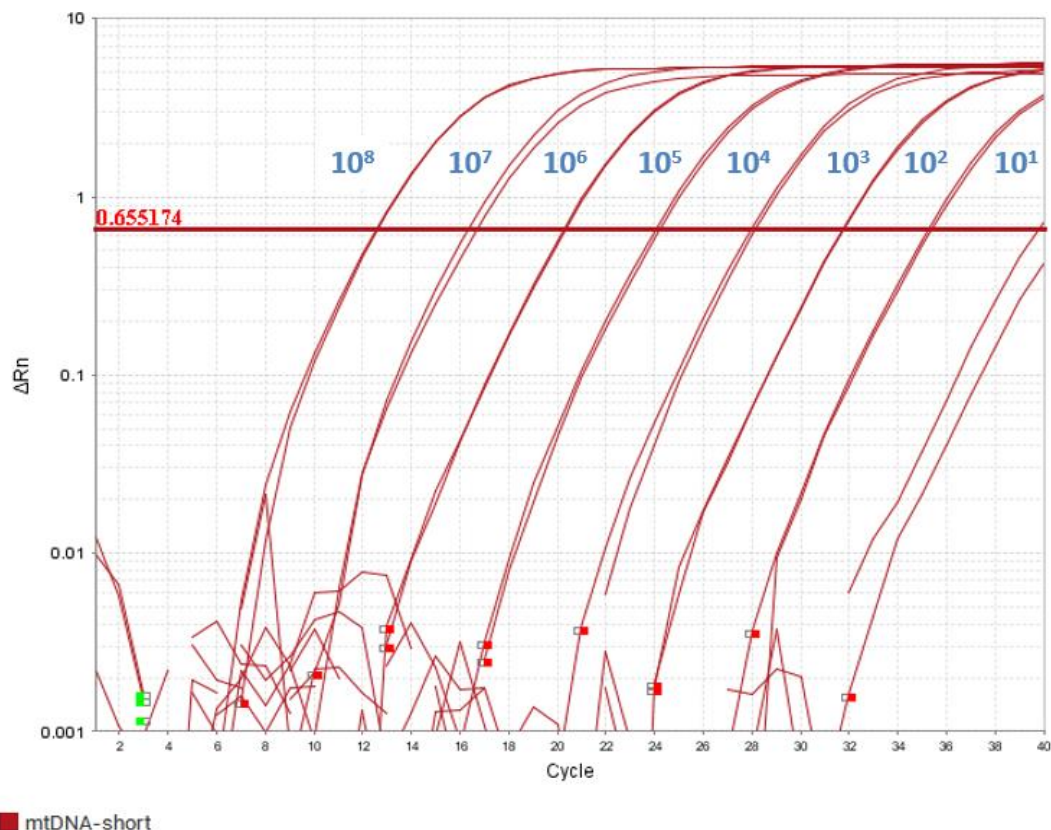


Figure 1. qPCR Amplification Plot of the Short Target with the dsT8sig Standard. A plot of the number of cycles versus the difference in fluorescence of the sample minus the baseline. Each dilution is labeled according to its determined copy number/ μ L. The horizontal line and value is the threshold line and corresponding threshold ΔRn value.

As one can see from **Figure 1**, the amplification plot of the short synthetic standards was well distributed with tight replicate amplification, with worse replicate performance at lower concentrations (e.g., 10^1 copies/ μ L standard; **Figure 1**). This is expected, as uniform dilutions tend to be more difficult as overall concentrations decrease. In order to understand the efficiency of the reaction and linearity of the dilution series quantification, a standard curve for the short standards run was also examined (**Figure 2**).

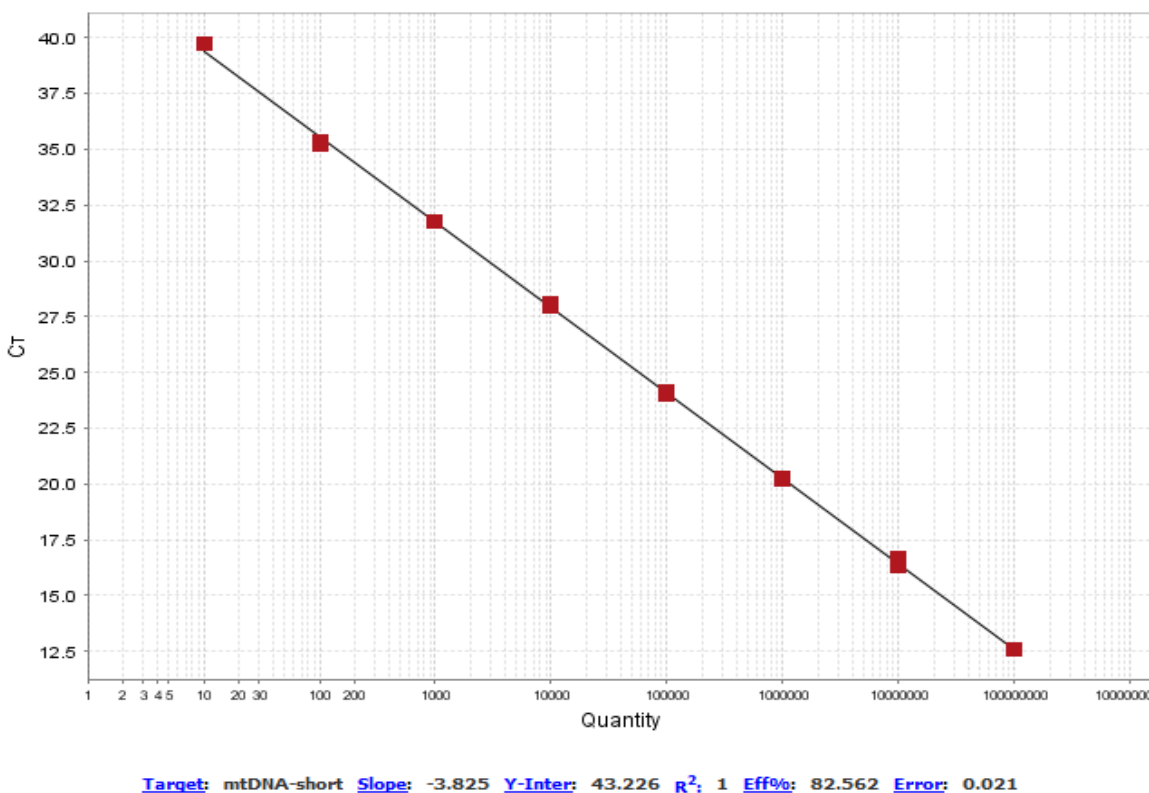


Figure 2. *qPCR Standard Curve of Short Target with the dsT8sig Standard.* The standard curve for the short target standard dilution series, with each red square representing a sample in the dilution series. Measures of slope, y-intercept, R^2 , efficiency, and error are included.

Typically, qPCR reactions with high-quality data results see a separation of 3.3 Ct values between 10-fold dilutions, which can be visually verified either in an amplification plot or by looking at the slope in a standard curve (-3.3). This is not the case in the author's replication of the mtDNA triplex (**Figure 2**). The standard curve has excellent linearity ($R^2 = 1$), but overall efficiency is less than desirable efficiency for qPCR, which is anywhere from 90-110 %¹⁴. An inefficient reaction can be caused by any number of things but considering that primers and reaction conditions were identical to that of Kavlick et al. 2019 one could most likely attribute the amplification issues to pipetting errors on behalf of the author or similar types of human error. Additionally, the consumables and materials used can affect dilutions based on their ability to separate the contents from the outside environment. Kavlick et al. 2011 found an average

efficiency value for their assay of around 96.5%. These comparisons can only be made for the short target as efficiency values for the long target were never published.

Optimization and testing of the gBlocks standard

In the course of the author's attempts to replicate the original triplex assay, it was realized that while the dsT8sig oligonucleotide standard designed by Kavlick et al 2011 had performed well for its purposes as a short standard, it does not contain the long target sequence; the long target is quantified relative to the short. As the entire purpose of the long target's inclusion in the triplex assay was to allow comparison of Ct values between the long and short to determine a degradation index, it makes logical sense to include a standard with a known amount of both the long and short target. To address this shortcoming, Kavlick et al. 2019 included a "calibrator" DNA sample that contains known equal amounts of short and long target, therefore having a degradation index of 0. While the "calibrator" DNA used by Kavlick et al 2019 is commonly used by forensic laboratories as a positive control for mtDNA analysis (*pers. comms.*) (HL60; ATCC, Manassas, VA) it needs to be stored and used correctly to avoid degradation (*i.e.*, limited freeze thaws, fresh dilutions prepared monthly *etc.*).

To remove the need for including a calibrator sample, a new highly-stable standard that could include both the short and the long target was obtained. The custom gBlock standard designed in this study was tested using the same approach as the dsT8sig standard from Kavlick et al 2011. The Ct values for the short target within the gBlock standard in this study fell within the range of values that Kavlick et al. 2011 observed (**Table 3**).

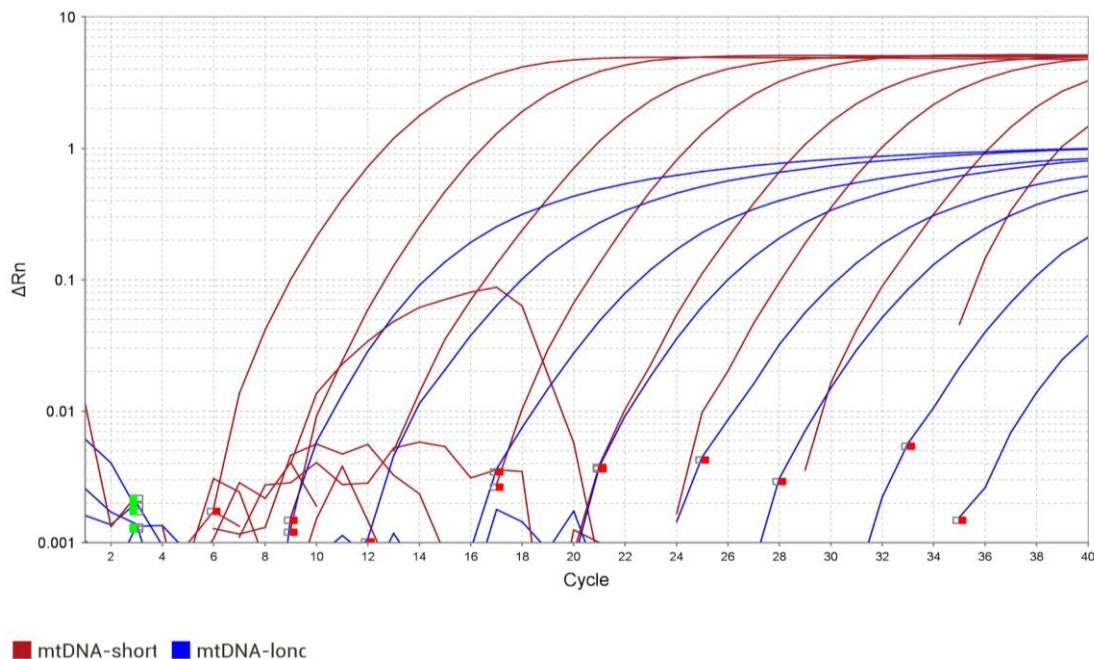


Figure 3. qPCR Amplification Plot of Short and Long targets with the custom gBlock Standard. In this reaction, the signal fluorescence minus the background fluorescence is plotted versus cycle number for both the short (105 bp; red) and long (316 bp; blue) targets.

When examining the amplification plot (**Figure 3**), it is apparent that the long target appeared on average 2.27 Cts later than its short target equivalent, amplification however does seem to be proceeding well and separation among ten-fold dilutions is clear. After corresponding with Dr. Kavlick, he communicated that the Ct values of the long target in the calibrator sample were generally 2 to 3 Ct values higher than the short target equivalent. This may very well have to do with the fluorescence from the dye used (given FAM should be used with targets that have less efficient amplification), or with the perhaps worse amplification efficiency for the long target. Nevertheless, both targets within the gBlock exhibited strong linearity comparable with that of the dsT8sig oligonucleotide standards as observed via the standard curve (**Figure 4**).

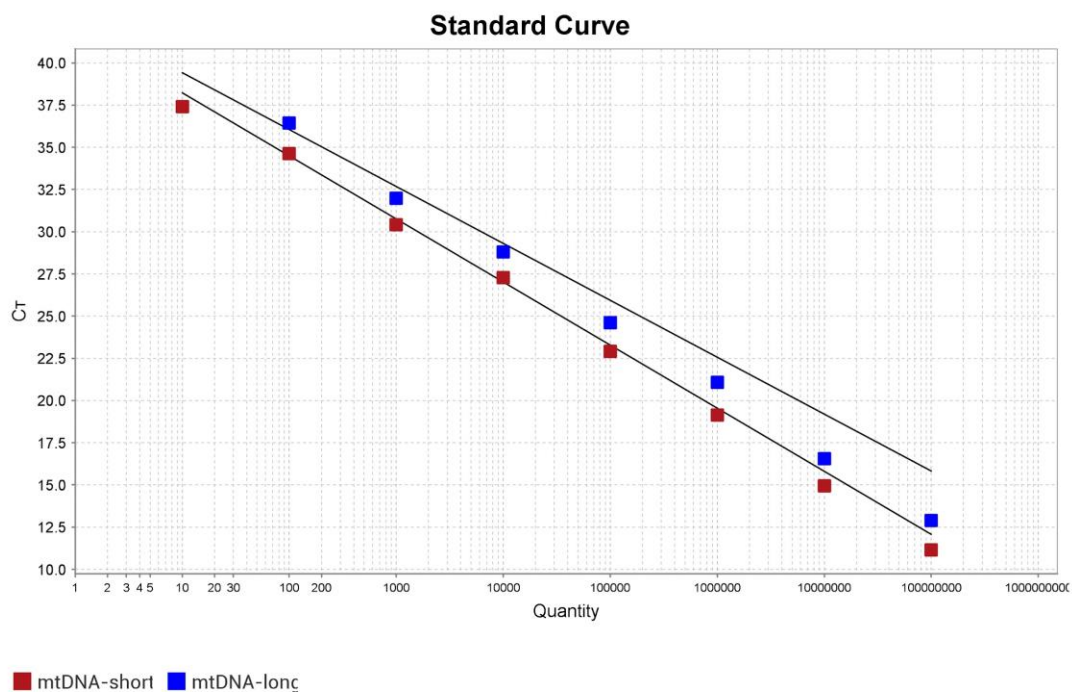


Figure 4. *qPCR Standard Curves of Short and Long targets with the custom gBlock Standard.* Quantity versus Ct for the short (105 bp; red) and long (316 bp; blue) targets. Notably, the long target standard equating to 10^1 copies/ μL was omitted as its amplification was undetermined.

As stated previously, the standard curve illustrates strong linearity, with R^2 values of 0.993 and 0.998 for the short and long target, respectively (**Figure 4**). Efficiency values of the gBlock standard were comparable to dsT8sig standards of approximately 85%, which while not ideal is consistent. This also lends credence to the idea that the standards themselves are not the issue with the amplification efficiency, but perhaps a component of the reaction.

Transitioning to ddPCR

When transitioning to ddPCR, the first steps involved running the assay with the recommended cycling conditions and reagent concentrations as published by the National Institute of Standards

and Technology in 2018¹⁵. Since it is documented that ddPCR has a dynamic range of 1 to 100,000 copies per reaction, the author initially tested the custom gBlock standard dilutions ranging from 10^5 copies/ μL to 10^1 copies/ μL ¹⁶. To assess the performance of the short target amplification, the 1D amplification plot provided in the QuantaSoft software was examined (**Figure 5**).

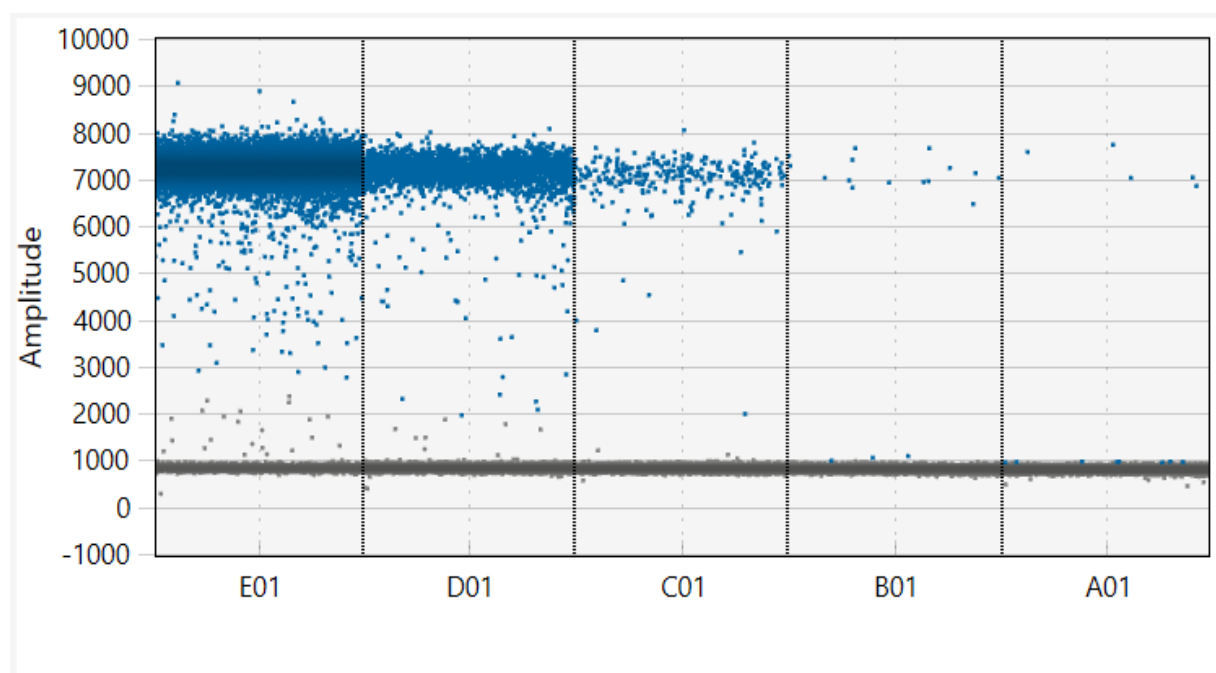


Figure 5. ddPCR Amplification Plot for Short gBlock. Samples E01, D01, C01, B01, and A01 correspond to samples containing 10^5 , 10^4 , 10^3 , 10^2 , and 10^1 copies/ μL respectively.

Since in ddPCR the roughly 20,000 droplets are separated binarily as either positive (containing amplified DNA) or negative (containing no amplified DNA), the blue dots in the plot reflect positive droplets and the gray dots negative droplets. From left to right in **Figure 5**, each sample is a 10-fold dilution (10^5 – to 10^1 copies), which were reported by ddPCR reads as 2,116, 182, 18.4, 1.1, and 0.817 copies/ μL respectively. This was concerning, since it was determined that each gBlock was reading at a concentration nearly 2 orders of magnitude lower than what their

concentrations had been previously determined to be. Upon speaking with BioRad technical support, it was determined that all ddPCR concentrations need to be multiplied by 10 due to the fact that the original DNA sample is diluted 1:10 when added to the entire 24 μ L reaction. While this explained a 1 order of magnitude discrepancy, it did not explain a 2 orders of magnitude error. As such, the gBlock standards were re-analyzed on a Qubit Fluorometer and compared to the original quantification on the NanoDrop One instrument. Results of Qubit quantification indicated that the original NanoDrop estimate had been overestimated by roughly an order of magnitude. As such, gBlock standards were re-labeled an order of magnitude lower and input appropriately on future qPCR and ddPCR assays.

Unfortunately, efficient amplification of the long target using ddPCR proved more difficult to reproduce. Only once out of 6 times tested was the author able to obtain an efficient enough reaction that the software was able to separate positive and negative droplets, and even then the separation was quite poor (**Figure 6**). In general, poor separation in ddPCR is caused by an inefficient PCR reaction. If the reaction is too inefficient, too few PCR products will be created in each droplet and the fluorescence will be too small to accurately distinguish between positive and negative droplets. As one can see below in **Figure 6**, separation is not nearly as clear or as large as was seen in the short target amplification, which saw a general separation between positive and negative droplets of nearly 6,000 relative fluorescence units (RFUs).

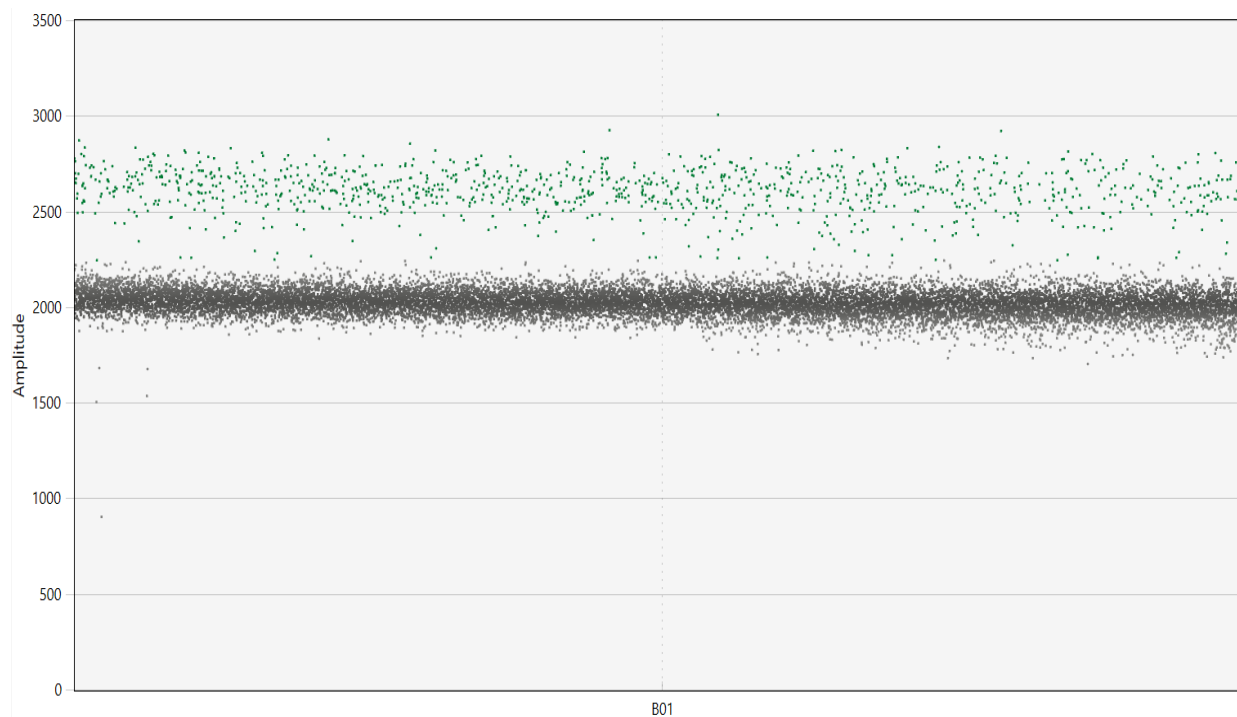


Figure 6. ddPCR Amplification Plot for Long gBlock. B01 represents a sample with a gBlock concentration estimated to be 10^4 copies/ μ L, which the software quantified correctly. While separation between positive and negative droplets exists, it is very small.

To attempt to improve the amplification efficiency for the long target, after conversations with BioRad technical support the author conducted an annealing temperature gradient and a primer concentration gradient. It was also noted that the suggested $2^\circ\text{C}/\text{sec}$ ramp rate was not being applied to all steps, and so that was remedied. This is a potential limitation of using shared equipment, since methods can be edited by any user. Unfortunately, even when testing temperatures and primer concentrations recommended by BioRad, reproducible separation of the long target was still not obtained (results not shown). Due to the poor separation observed in the long target, it was not included in subsequent comparisons of ddPCR and qPCR for mtDNA quantification.

qPCR and ddPCR Comparison

Following optimization of the short target on ddPCR, the gBlock standard, dsT8sig oligonucleotide standard, and three sample groups of unknown quantity were all quantified side-by-side using qPCR and ddPCR. Each sample measured on each instrument was from the same prepared dilution sample to prevent intra-sample variation. From this initial set of experiments, it was observed that ddPCR exhibits increased sensitivity compared to qPCR in direct quantification (**Figure 7**). The total concentration for each sample on each assay and standard are listed below in **Table 4**.

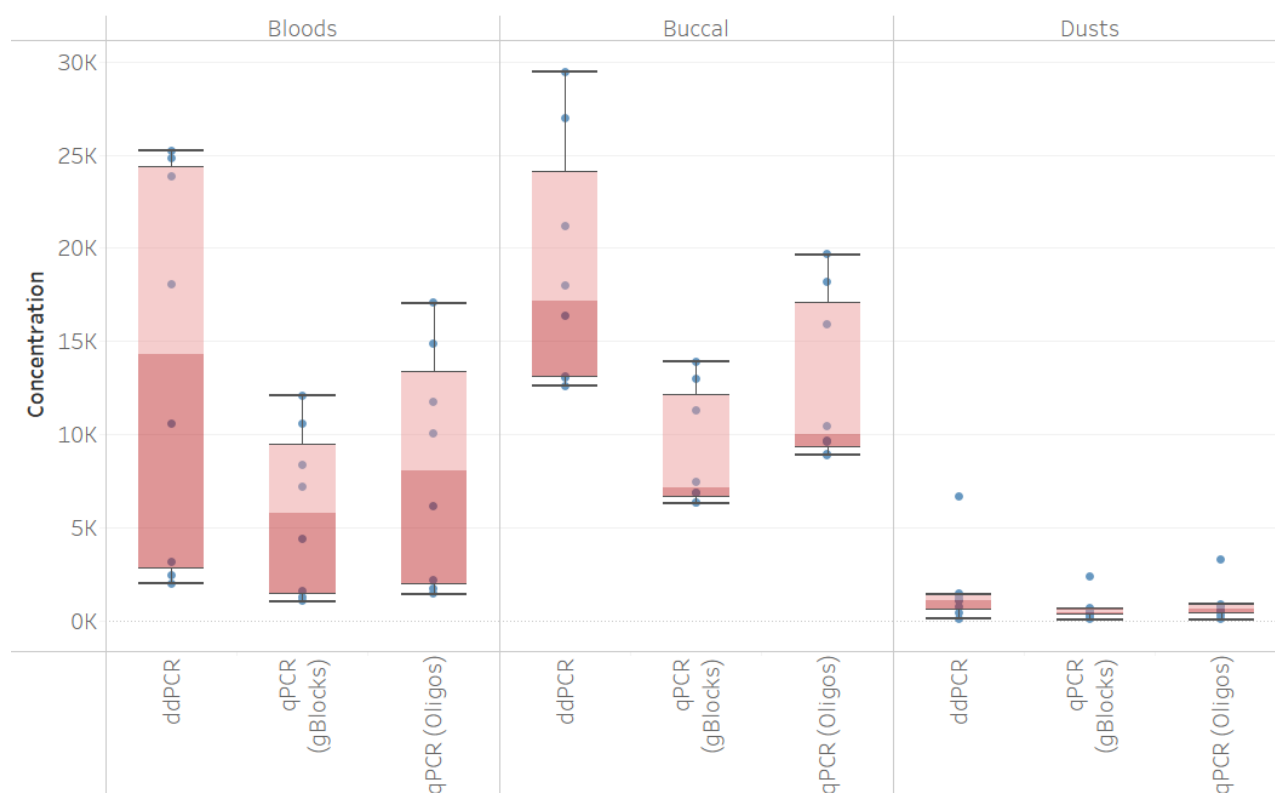


Figure 7. Comparison of ddPCR and qPCR quantification across sample types. Concentration (blue circles) for each sample broken down by sample type (blood, buccal, or dust) and instrument (ddPCR or qPCR based on gBlocks or oligonucleotides). The light red section contains the upper 50% of concentration and the darker red the lower 50%. The horizontal lines at the top and bottom of each plot are the upper and lower concentration ranges respectively.

The concentrations provided from using the dsT8sig oligonucleotide standards was manually reduced by a factor of 10, as it was determined later on ddPCR that the dsT8sig standards were originally quantified incorrectly by roughly a factor of 10.

Table 4. ddPCR and qPCR (gBlock and Oligonucleotide) concentration. Each sample and sample type as well as their concentrations on ddPCR and qPCR (based separately on gBlock and dsT8sig standards). The degradation index is calculated using the $\Delta\Delta C_t$ method outlined by Kavlick et al. 2019. * denotes that sample HS10-REF3 is an outlier and should be omitted, as a degradation index of 0 is not possible. Note that the qPCR concentration determined by the oligonucleotide standard is reduced by a factor of 10 from the original quantification. Re-quantification found the standards to be off by roughly an order of magnitude, so their actual concentration should be roughly 10-fold less.

Sample Type	Sample	ddPCR Short Concentration (copies/uL)	qPCR Concentration-gBlocks (copies/uL)	qPCR Concentration-Oligos (copies/uL)	Degradation Index (Short:Long)
Blood	D1	23860	12049	17041	1.1
	F1	2410	1220	1695	1.0
	B4	18060	7171	10004	1.3
	B6	10530	4371	6139	1.3
	F8	25200	10534	14872	1.4
	D10	3110	1547	2161	1.4
	H7	1970	1021	1421	1.2
	D8	24820	8325	11739	1.0
Dust	HS7-DS1	744	340	471	2.8
	HS7-DS2	1300	564	783	0.1
	HS7-DS4	77	35	48	1.2
	HS7-DS5	1130	502	696	5.3
	HS4-DS1	1420	620	861	7.5
	HS4-DS2	419	202	278	5.6
	HS4-DS4	6620	2342	3277	26.9
	HS4-DS5	1070	386	535	5.4
Buccal	HS1-REF1	13110	6298	8873	1.5
	HS8-REF1	21170	12965	18151	1.3
	HS9-REF2	16320	7434	10424	1.7
	HS9-REF4	12990	6815	9595	1.6
	HS10-REF1	29460	11234	15865	1.6
	HS10-REF3*	26940	13889	19637	0.0
	HS11-REF1	12590	6346	8933	2.0
	HS14-REF2	17940	6823	9608	1.5

In addition to the increased sensitivity observed on the box and whisker plot, it was also noted in calculated the coefficient of variation for each of the instruments (**Table 5**).

Table 5. Coefficient of Variation (CV) for sample types on ddPCR and qPCR. The values here are on the log scale because the CVs would be much larger on a linear scale and might lead to incorrect interpretation about the magnitude of variation. The qPCR concentrations were based on the mean of the gBlock and dst8sig standards.

Sample Type	Log ddPCR Concentration CV	Log qPCR Concentration CV
Bloods (n, 8)	12.265	12.246
Buccal (n, 8)	3.400	3.706
Dusts (n, 8)	18.545	20.009
Combined (n, 24)	18.946	20.887

The CV is the standard deviation according to each instrument and sample type over their means. In comparing two CVs, a CV higher than another simply means that concentrations for that group of samples were on average farther from the mean of those sample concentrations. As one can observe, ddPCR is on average less variable in its results than qPCR, which indicates that ddPCR is also more precise than qPCR. One may note that the CV (and the difference between CVs) vary among sample type. This may have to do with the fact that some samples types, like the dusts, contain lower quality DNA, which ddPCR has been shown to have an advantage quantifying over qPCR. In order to confirm that the difference in concentration for each sample was significantly different, a matched-pairs *t*-test was performed, the results of which are summarized below in **Table 6**.

Table 6. Statistical Significance of Concentration Difference. This matched-pairs *t*-test was conducted under the hypothesis that the mean difference between qPCR and ddPCR concentration was = 0. qPCR concentrations were calculated from the mean of the gBlocks and dst8sig standards. Calculations are across all 15 samples and based only on values for the short target.

Mean ddPCR Concentration (copies/ μ L)	Mean qPCR Concentration (copies/ μ L)	Mean Difference (ddPCR-qPCR) (copies/ μ L)	P-value (Probability that ddPCR=qPCR)
11,385.8	6,169.64	5,216.2	0.0001

What a *p*-value of 0.0001 means is that one can very confidently say that the difference in concentration measured between samples on ddPCR and qPCR is statistically significant.

Looking at the test, one can also say that on average ddPCR is giving a higher concentration than qPCR (by roughly 1.8-fold).

Conclusions

Transitioning this triplex mtDNA qPCR assay to ddPCR has demonstrated that ddPCR is, on average, more sensitive and precise than qPCR, which makes it well-equipped for assays that might encounter low-copy number samples. Since mtDNA analysis is often used on degraded or aged samples, encountering degraded or low-copy mtDNA samples is expected and therefore ddPCR is a useful tool. The lack of a standard curve eliminated the careful preparation and quantification that is required of qPCR. An immediate future need for this assay is to optimize the amplification reaction of the long target to allow for ddPCR multiplexing and calculation of degradation indexes for samples. Without this, any indication as to the quality of the overall mitochondrial genome is lost and sequencing may not yield desired results without that knowledge. In addition, since the partitioning that occurs during ddPCR is known to diminish the effects of PCR inhibitors, one future direction would be to examine the effects of inhibitors on this mtDNA ddPCR assay and see if there is improved performance compared to qPCR.

REFERENCES

1. Sprouse ML, Phillips NR, Kavlick MF, Roby RK. Internal validation of human mitochondrial DNA quantification using real-time PCR. *Journal of Forensic Sciences*. 2014;59(4):1049–1056. doi:10.1111/1556-4029.12477
2. CODIS-NDIS Statistics. Law Enforcement. <https://le.fbi.gov/science-and-lab-resources/biometrics-and-fingerprints/codis/codis-ndis-statistics>
3. Amorim A, Fernandes T, Taveira N. Mitochondrial DNA in human identification: a review. *PeerJ*. 2019;7:e7314. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6697116/>. doi:10.7717/peerj.7314
4. Shitara H, Kaneda H, Sato A, Inoue K, Ogura A, Yonekawa H, Hayashi JI. Selective and continuous elimination of mitochondria microinjected into mouse eggs from spermatids, but not from liver cells, occurs throughout embryogenesis. *Genetics*. 2000;156(3):1277–1284. doi:10.1093/genetics/156.3.1277
5. Atig RK-B, Hsouna S, Beraud-Colomb E, Abdelhak S. [Mitochondrial DNA: properties and applications]. *Archives De l'Institut Pasteur De Tunis*. 2009;86(1–4):3–14.
6. Parson W, Gusmão L, Hares DR, Irwin JA, Mayr WR, Morling N, Pokorak E, Prinz M, Salas A, Schneider PM, et al. DNA Commission of the International Society for Forensic Genetics: Revised and extended guidelines for mitochondrial DNA typing. *Forensic Science International*:

Genetics. 2014 ;13:134–142.

<https://www.sciencedirect.com/science/article/pii/S1872497314001586>.

doi:10.1016/j.fsigen.2014.07.010

7. Bogdanowicz W, Allen M, Branicki W, Lembring M, Gajewska M, Kupiec T. Genetic identification of putative remains of the famous astronomer Nicolaus Copernicus. *Proceedings of the National Academy of Sciences*. 2009;106(30):12279–12282.

<https://pnas.org/doi/full/10.1073/pnas.0901848106>. doi:10.1073/pnas.0901848106

8. Ballard D, Winkler-Galicki J, Wesoly J. Massive parallel sequencing in forensics: advantages, issues, technicalities, and prospects. *International Journal of Legal Medicine*. 2020;134(4):1291–1303. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7295846/>. doi:10.1007/s00414-020-02294-0

9. Kavlick MF. Development of a triplex mtDNA qPCR assay to assess quantification, degradation, inhibition, and amplification target copy numbers. *Mitochondrion*. 2019;46:41–50. doi:10.1016/j.mito.2018.09.007

10. Kavlick MF, Lawrence HS, Merritt RT, Fisher C, Isenberg A, Robertson JM, Budowle B. Quantification of Human Mitochondrial DNA Using Synthesized DNA Standards*. *Journal of Forensic Sciences*. 2011;56(6):1457–1463. <https://onlinelibrary.wiley.com/doi/10.1111/j.1556-4029.2011.01871.x>. doi:10.1111/j.1556-4029.2011.01871.x

11. Li B, Kaushik S, Kalinowski P, Kim B, Gershon C, Ching J, Poburko D. Droplet digital PCR shows the D-Loop to be an error prone locus for mitochondrial DNA copy number determination. *Scientific Reports*. 2018;8(1):11392. <https://www.nature.com/articles/s41598-018-29621-1>. doi:10.1038/s41598-018-29621-1

12. Zhao Y, Xia Q, Yin Y, Wang Z. Comparison of Droplet Digital PCR and Quantitative PCR Assays for Quantitative Detection of *Xanthomonas citri* Subsp. *citri*. *PLoS ONE*. 2016;11(7):e0159004. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4948846/>. doi:10.1371/journal.pone.0159004

13. Li H, Bai R, Zhao Z, Tao L, Ma M, Ji Z, Jian M, Ding Z, Dai X, Bao F, et al. Application of droplet digital PCR to detect the pathogens of infectious diseases. *Bioscience Reports*. 2018;38(6):BSR20181170. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6240714/>. doi:10.1042/BSR20181170

14. ROGERS-BROADWAY K-R, KARTERIS E. Amplification efficiency and thermal stability of qPCR instrumentation: Current landscape and future perspectives. *Experimental and Therapeutic Medicine*. 2015;10(4):1261–1264. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4578049/>. doi:10.3892/etm.2015.2712

15. Romsos EL, Kline MC, Duewer DL, Toman B, Farkas N. Certification of standard reference material 2372a human DNA quantitation standard. Gaithersburg, MD: National Institute of Standards and Technology; 2018. p. NIST SP 260-189. Report No.: NIST SP 260-189.

<http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.260-189.pdf>.

doi:10.6028/NIST.SP.260-189

16. Planning Droplet Digital PCR Experiments. Bio-Rad Laboratories. <https://www.bio-rad.com/life-science/learning-center/introduction-to-digital-pcr/planning-ddpcr-experiments>

APPENDIX

Appendix Table 1. *gBlock and dsT8sig sequence.* Forward strand sequence for each standard set used. The dsT8sig standard sequence was taken from Kavlick et al. 2011.

Standard	Forward Strand Sequence (5'-3')
gBlock	ATA ATA GTT ACA ATC GGC ATC AAC CAA CCA CAC CTA GCA TTC CTG CAC ATC TGT ACC CAC GCC TTC TTC AAA GCC ATA CTA TTT ATG TGC TCC GGG TCC ATC ATC CAC AAC CTT AAC AAT GAA CAA GAT ATT CGA TTT TTT TTT TGA AAA CAT TCT CCT CCG CAT AAG CCT GCG TCA GAT TAA AAC ACT GAA CTG ACA ATT AAC AGC CCA ATA TCT ACA ATC AAC CAA CAA GTC ATT ATT ACC CTC ACT GTC AAC CCA ACA CAG GCA TGC TCA TAA GGA A..I\G GTT AAA AAA AGT AAA AGG AAC TCG GCA AAT CTT ACC CCG CCT GTT TAC CAA AAA CAT CAC CTC TAG CAT CAC CAG TAT TAG AGG CAC CGC CTG CCC AGT GAC ACA TGT TTA ACG GCC GCG GTA CCC TAA CCG TGC AAA GGT AGC ATA ATC ACT TGT TCC TTA AAT AGG GAC CTG TAT GAA TGG CTC CAC GAG GGT TCA GCT GTC TCT TA
dsT8sig	CAA TCG GCA TCA ACC AAC CAC ACC TAG CAT TCC TGC ACA TCT GTA CCC ACG CCT TCT TCA AAT AAC GAC TAT TTA TGT GCT CCG GGT CCA TCA TCC ACA ACC TTA ACA ATG AAC A