TITLE: Quantification of HIV-1 DNA using Real-Time Recombinase Polymerase Amplification

AUTHORS:

Zachary Austin Crannell[1*], Brittany Rohrman[1*], Rebecca Richards-Kortum[1]

*These authors contributed equally

[1] Rice University

6500 Main Street

Houston, TX

ABSTRACT

Although recombinase polymerase amplification (RPA) has many advantages for the detection of pathogenic nucleic acids in point-of-care applications, RPA has not yet been applied for has not yet been implemented to quantify sample concentration using a standard curve. Here we describe a real-time RPA assay with an internal positive control and an algorithm that analyzes the real-time fluorescence data to quantify HIV-1 DNA. We show that DNA concentration and the onset of detectable amplification are correlated by an exponential standard curve. In a set of experiments in which the standard curve and algorithm were used to analyze and quantify additional DNA samples, the algorithm predicted an average concentration within one order of magnitude of the correct concentration for all HIV-1 DNA concentrations tested. These results suggest that qRPA may serve as a powerful tool for quantifying nucleic acids and may be adapted for use in single-sample point-of-care diagnostic systems.

INTRODUCTION

For sensitive and specific diagnosis of many infectious diseases, real-time polymerase chain reaction (PCR) is widely considered to be the gold standard method. In clinical settings, real-time PCR is used for both detection and quantification of pathogenic nucleic acids. For example, HIV-1 proviral DNA is detected for pediatric diagnosis, while HIV-1 viral RNA is quantified to monitor treatment efficacy and disease progression.[1,2] In areas where the infectious disease burden is highest, however, real-time PCR is often infeasible because of the requirement for expensive thermal cycling equipment, technical expertise, and other resources that may be unavailable in poor settings.

To address this problem, isothermal amplification platforms have been developed that rapidly amplify nucleic acids to detectable levels at a single temperature, alleviating the need for thermal cycling equipment.[3] Recombinase polymerase amplification (RPA) offers significant advantages over other isothermal amplification techniques for point-of-care applications: it requires a lower amplification temperature, is tolerant to impure samples, amplifies targets to detectable levels within minutes, and uses lyophilized enzymes to enable storage and transport at room temperature.[4,5] For these reasons, a number of recent reports have proposed RPA-based strategies for the detection of pathogens.[6–10] Although some papers demonstrate a relationship between nucleic acid concentration and onset of amplification,[10,11] to the best of our knowledge, RPA has not yet been implemented to quantify sample concentration using a standard curve. Moreover, the accuracy with which samples can be quantified with real-time RPA has not yet been characterized.

In this study, we explore the feasibility of developing a quantitative RPA assay using real-time fluorescence monitoring. With the ultimate goal of developing an assay for use with a point-of-care reader, we designed an assay to enable quantification of HIV-1 DNA in a single sample by referencing a standard curve determined previously (e.g. during calibration). This proof-of-concept quantitative RPA

(qRPA) assay detects HIV-1 DNA and an internal positive control (IPC) sequence, which are both amplified by the HIV-1 primers and detected using probe sequences conjugated to different fluorophores (Figure 1). Using a benchtop real-time PCR machine, RPA was performed to determine the correlation between the HIV-1 DNA target concentration and fluorescence intensity during amplification. Next, the IPC was generated and similarly assessed. Finally, both HIV-1 DNA and the IPC were amplified within the same reaction to generate data for the development, training, and validation of a custom algorithm for DNA quantification.

METHODS

Real-time RPA. In experiments detecting only one target sequence, RPA reactions were assembled according to the manufacturer's instructions using reagents from the TwistAmp exo kit (TwistDx, Ltd., Cambridge, United Kingdom). In experiments detecting both HIV-1 DNA and IPC DNA, each 50 µL reaction contained 29.5 µL rehydration buffer, 2.1 µL RPA forward primer [10 µM], 2.1 µL RPA reverse primer [10 µM], 0.6 µL HIV-1 HEX-labeled probe [10 µM], 0.6 µL IPC FAM-labeled probe [10 µM], 2.6 µL IPC DNA [10^4 copies/µL], 10 µL HIV-1 DNA at varying concentrations, 2.5 µL magnesium acetate, and one enzyme pellet. All primer and probe sequences can be found in **Table S1**. Previous work has shown that the RPA primers and HIV-1 probe target the HIV-1 *pol* gene with high sensitivity and specificity.[9] All DNA oligonucleotides were purchased from Integrated DNA Technologies (Coralville, USA) or BioSearch Technologies (Novato, USA). HIV-1 DNA samples contained a background of at least 10 ng of human genomic DNA and a total of 0, 10, 10^2, 10^3, 10^4, or 10^5 copies of the plasmid pHIV-IRES-eYFPΔEnvΔVifΔVpr, a generous gift from R. Sutton.[12] In all experiments, each concentration of HIV-1 DNA was tested in duplicate.

RPA reactions were assembled without magnesium acetate, avoiding direct exposure to light, and placed in a cold block cooled to 4 °C. Magnesium acetate, which is necessary for enzymatic activity, was

added immediately before transporting the RPA reactions to the real-time PCR machine. Amplification and real-time data collection were performed in a BioRad CFX96 Real Time qPCR machine (Hercules, USA). Reactions were pre-incubated for 1 min at 39 °C without fluorescence monitoring, followed by incubation for 20 min at 39 °C. The fluorescence of FAM, HEX, or both was monitored every 20 sec for 20 min following pre-incubation. Raw fluorescence data was collected and exported to a Microsoft Excel spreadsheet using the Bio Rad CFX Manager. Analysis was performed using a custom MATLAB algorithm.

Generation of internal positive control (IPC) DNA. *Cryptosporidium parvum*, an intestinal parasite, was chosen to serve as the template for IPC DNA generation because this organism is not found in human blood and shares minimal sequence similarity with HIV-1. To generate the IPC sequence, *C. parvum* DNA was first extracted and purified from oocysts (see Supplementary Information). PCR was then performed on the extracted DNA using primers with two regions, one complementary to *C. parvum,* and the other complementary to HIV-1 (**Table S1**). The PCR products, which served as the IPC, consisted of a *C. parvum* sequence flanked by the HIV-1 primer sequences (shown schematically in Figure 1). Generation of the 435 bp IPC was verified by gel electrophoresis, after which the DNA was extracted, purified, and diluted.

Quantification algorithm for analysis of real-time data. Raw fluorescence data were analyzed using a custom MATLAB script. To score a sample as positive or negative, the difference Δ_{sample} between the maximum and minimum fluorescence was determined for each sample. The average difference $\Delta_{background}$ and standard deviation $\sigma_{background}$ were calculated for all no-target control samples. A sample was considered positive if Δ_{sample} was more than $z^* \sigma_{background}$ above $\Delta_{background}$. Values for z varied from one to five to determine the effect of this parameter on the sensitivity and accuracy of the algorithm. Each sample was scored as negative or positive for both HIV-1 and the IPC. Samples that scored as negative for both HIV-1 and the IPC were considered invalid and removed from further analysis.

For each sample identified as positive for HIV-1, the raw fluorescence intensity from the HEX channel, which corresponded to amplification of the HIV-1 sequence, was smoothed with a five point running average. The slope was estimated at each time point by calculating the difference in fluorescence intensity between each time point and the previous time point. A slope threshold *S* was chosen to estimate the time at which detectable amplification began. The value of *S* was chosen such that the greatest slope of the no-target control samples was slightly less than *S*. The same value for *S* was used for both the training and the validation datasets. Preliminary analysis showed that minor changes in *S* did not significantly affect the sensitivity, specificity, or dynamic range of the algorithm (data not shown). The time *T* at which detectable amplification began was defined as the first time point at which the slope was greater than *S*. *T* was found for all positive samples from five experiments. $T_{average}$ was plotted for each concentration tested, and regression analysis was performed to construct a standard curve relating *T* to target concentration. To evaluate the accuracy of the algorithm, the concentration was predicted for samples from five additional experiments by calculating *T* and using the equation for the standard curve.

RESULTS

The generation of IPC DNA was verified by the presence of a 435 bp band on an agarose gel after electrophoresis (Figure S1). Preliminary RPA experiments detecting HIV-1 DNA and IPC DNA in separate reactions demonstrated that *T,* the time at which detectable amplification begins, increases with decreasing DNA concentration, suggesting that quantification of DNA with RPA is feasible.

Five experiments were then performed to characterize the performance of the qRPA assay in which both HIV-1 DNA and the IPC were amplified. Raw fluorescence data for a typical experiment are shown in Figure 2, where HIV-1 DNA was detected using a HEX-labeled probe (Figure 2A), and the IPC was detected using a FAM-labeled probe (Figure 2B). None of samples were scored as negative for both the

HIV sequence and control; therefore, all samples were classified as valid. For most experiments, *T* was approximately six minutes for the IPC, regardless of the concentration of HIV-1 DNA present in the sample. The results of the regression analysis, performed to determine the correlation between HIV-1 DNA concentration and *T*, demonstrated that an exponential fit yielded a higher correlation coefficient than first- and second-order polynomial fits. An exponential standard curve derived using data from the five experiments is plotted in Figure 3, in which samples were considered positive if Δ_{sample} was more than one $\sigma_{background}$ above $\Delta_{background}$ (z = 1). Standard curves using integer values of z from one to five had similar equations and correlation coefficients (data not shown).

The accuracy and precision of the algorithm in predicting the HIV-1 DNA concentration from the raw fluorescence data was evaluated with five additional experiments, using the exponential standard curves calculated for the training data. As for the training data, all samples were classified as valid. Table 1 shows the performance of the algorithm when z = 1, using the standard curve shown in Figure 3. The algorithm correctly identified nine out of ten of the no-target-control samples as negative. The no-target-control sample that was identified as positive was estimated to contain less than 10 copies of HIV-1 DNA and may have been a result of template contamination. The algorithm classified all HIV-1 DNA samples containing 10 copies or more as positive. For all HIV-1 DNA concentrations, the algorithm predicted an average concentration within one order of magnitude of the correct concentration; however, the average predicted concentrations were more accurate for lower concentrations of template DNA. The standard deviation of the predicted concentrations was less than 0.5 $\log_{10}$ (copies per reaction) for all concentrations, with the highest standard deviation at 100 copies due to two outlying samples. Variations in genomic background DNA concentration had little effect on the accuracy of predicted concentrations (see Supporting Information and Fig. S2).

DISCUSSION

These results provide proof-of-concept to support the use of RPA for accurate quantification of nucleic acid concentration with a standard curve determined previously. For example, when combined with a reverse transcriptase step, the qRPA assay described here may be adapted to determine HIV-1 RNA viral load. As most commercially available viral load assays have a limit of detection of at least 3 $\log_{10}$ (HIV-1 RNA copies/mL), and patient viral loads may exceed 6 $\log_{10}$ (copies/mL), a clinically useful HIV-1 RNA viral load test must be able to quantify plasma viremia over at least four orders of magnitude.[13] A viral load test should also have a precision of at least 0.5 $\log_{10}$ (copies/mL), which is considered to be a significant change in viral load.[14] Finally, a viral load test must be especially accurate at low viral loads because therapeutic failure due to drug resistance is characterized by viral loads between 500–1000 (copies/mL),[2] and successful suppression of viral replication is indicated by viral loads less than 200 (copies/mL).[14] This qRPA assay can quantify DNA concentrations over four orders of magnitude with a precision greater than 0.5 $\log_{10}$ (copies per reaction) and is most accurate at low DNA concentrations. These results suggest that a qRPA assay similar to the one described here may have the potential to accurately identify these clinical benchmarks.

This qRPA assay may also be adapted for quantification of other targets by modifying the assay and algorithm parameters. For example, the criteria for identifying positive samples can greatly affect the sensitivity and linear dynamic range of the algorithm. As shown in Figure 3 and Table 1, predicted DNA concentrations are more accurate for lower concentrations than for higher concentrations when $z = 1$. Alternatively, using higher values for z decreases the sensitivity for lower concentrations but increases the accuracy for higher concentrations. Figure 4 demonstrates how changing z alters the sensitivity at low concentrations, the accuracy across detectable concentrations, and the precision of predicted concentrations. In addition to altering the algorithm parameters, the assay could potentially be optimized for greater accuracy by collecting fluorescence data more frequently or decreasing the

reaction rate by either decreasing the concentration of magnesium acetate in the reaction or amplifying at a lower temperature.

This assay was designed for use with an inexpensive, point-of-care fluorescence reader, such as the commercially available Twista portable real-time fluorometer (TwistDx, Ltd., Cambridge, United Kingdom) or the ESEQuant Tube Scanner (Qiagen, Valencia, USA). In order to adapt this assay for use in the field, however, several factors must be considered. Because RPA lacks true 'cycles' to limit the rate of DNA amplification, and amplicons can be generated in a matter of minutes, the rate of amplification must be precisely controlled. A consistent amplification rate may be achieved by using the same reagent concentrations for each experiment and by precisely controlling the temperature during each reaction. Reagent consistency was achieved by using whole enzyme pellets as supplied by the manufacturer, and by using the same primer aliquots for both the experiments used to build the standard curve and the experiments using the standard curve to predict concentration of samples. The temperature was controlled within ±0.2^{o}C by using a bench-top thermal cycler according to the manufacturer's specifications (BioRad CFX96 Touch™ Real-Time PCR Detection System). Deviations from these specifications and their influence on consistent amplification should be explored in future studies. In addition, reaction components must be protected from heat and light to prevent enzyme activity loss and photobleaching of probes. To ensure accurate results, reactions must be started immediately before fluorescence monitoring, which may be accomplished by adding magnesium acetate immediately before data collection begins. If these conditions are achieved, qRPA may serve as a powerful tool for quantifying nucleic acids in single samples at the point of care.

REFERENCES

(1) Creek, T. L.; Sherman, G. G.; Nkengasong, J.; Lu, L.; Finkbeiner, T.; Fowler, M. G.; Rivadeneira, E.; Shaffer, N. *Am. J. Obstet. Gynecol.* **2007**, *197*, S64–71.

(2) Volberding, P. A.; Deeks, S. G. *Lancet* **2010**, *376*, 49–62.

(3) Asiello, P. J.; Baeumner, A. J. *Lab Chip* **2011**, *11*, 1420–1430.

(4) Piepenburg, O.; Williams, C. H.; Stemple, D. L.; Armes, N. A. *PLoS Biol.* **2006**, *4*, e204.

(5) Krõlov, K.; Frolova, J.; Tudoran, O.; Suhorutsenko, J.; Lehto, T.; Sibul, H.; Mäger, I.; Laanpere, M.; Tulp, I.; Langel, U. *J. Mol. Diagn.* **2014**, *16*, 127–135.

(6) Rohrman, B. A.; Richards-Kortum, R. R. A paper and plastic device for performing recombinase polymerase amplification of HIV DNA. *Lab on a Chip*, 2012, *12*, 3082.

(7) Crannell, Z. A.; Castellanos-Gonzalez, A.; Irani, A.; Rohrman, B.; White, A. C.; Richards-Kortum, R. *Anal. Chem.* **2014**, *86*, 2565–2571.

(8) Kersting, S.; Rausch, V.; Bier, F. F.; von Nickisch-Rosenegk, M. *Malar. J.* **2014**, *13*, 99.

(9) Boyle, D. S.; Lehman, D. A.; Lillis, L.; Peterson, D.; Singhal, M.; Armes, N.; Parker, M.; Piepenburg, O.; Overbaugh, J. *MBio* **2013**, *4*, 1–8.

(10) Abd El Wahed, A.; El-Deeb, A.; El-Tholoth, M.; Abd El Kader, H.; Ahmed, A.; Hassan, S.; Hoffmann, B.; Haas, B.; Shalaby, M. A.; Hufert, F. T.; Weidmann, M. *PLoS One* **2013**, *8*. e71648. doi:10.1371/journal.pone.0071642

(11) Escadafal, C.; Faye, O.; Sall, A. A.; Faye, O.; Weidmann, M.; Strohmeier, O.; von Stetten, F.; Drexler, J.; Eberhard, M.; Niedrig, M.; Patel, P. *PLoS Negl. Trop. Dis.* **2014**, *8*, e2730. doi:10.1371/journal.pntd.0002730

(12) Segall, H. I.; Yoo, E.; Sutton, R. E. *Mol. Ther. J. Am. Soc. Gene Ther.* **2003**, *8*, 118–129.

(13) Sollis, K. a; Smit, P. W.; Fiscus, S.; Ford, N.; Vitoria, M.; Essajee, S.; Barnett, D.; Cheng, B.; Crowe, S. M.; Denny, T.; Landay, A.; Stevens, W.; Habiyambere, V.; Perrins, J.; Peeling, R. W. *PLoS One* **2014**, *9*, e85869.

(14) *Dep. Heal. Hum. Serv.* **2012**, 1–166.

Acknowledgments

This research was funded by a grant from the Bill & Melinda Gates Foundation through the Grand Challenges in Global Health Initiative.

Conflict of Interest

The authors declare no competing financial interest.

Supporting Information Available

The Supporting Information contains details about the DNA primers and probes, screening Internal Positive Control candidates, the specific protocol for extracting nucleic acids used to develop the Internal Positive Control, as well as details about the experiment mentioned at the end of the results section. This information is available free of charge via the Internet at http://pubs.acs.org/.

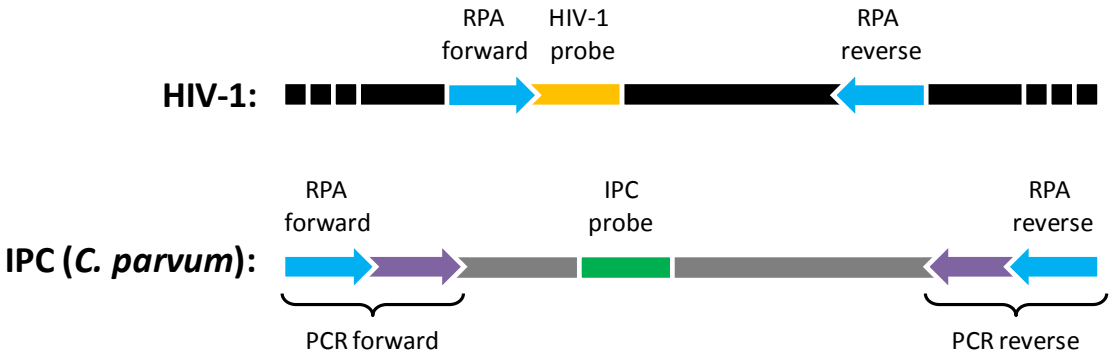

Figure 1: Quantitative RPA Assay Schematic. The qRPA assay detects two DNA sequences: a sequence within the HIV-1 genome ('HIV-1') and an internal positive control sequence ('IPC (*C. parvum*)'). Both sequences are amplified by the HIV-1 primers ('RPA forward' and 'RPA reverse'). The HIV-1 sequence is detected using a HEX-labeled probe ('HIV-1 probe'), while the IPC is detected using a FAM-labeled probe ('IPC probe'). The IPC consists of a sequence from *Cryptosporidium parvum* flanked by the HIV-1 primer sequences ('RPA forward' and 'RPA reverse') and is generated with PCR primers ('PCR forward' and 'PCR reverse') containing both a region complementary to *C. parvum* (purple) and a region complementary to HIV-1 (blue).

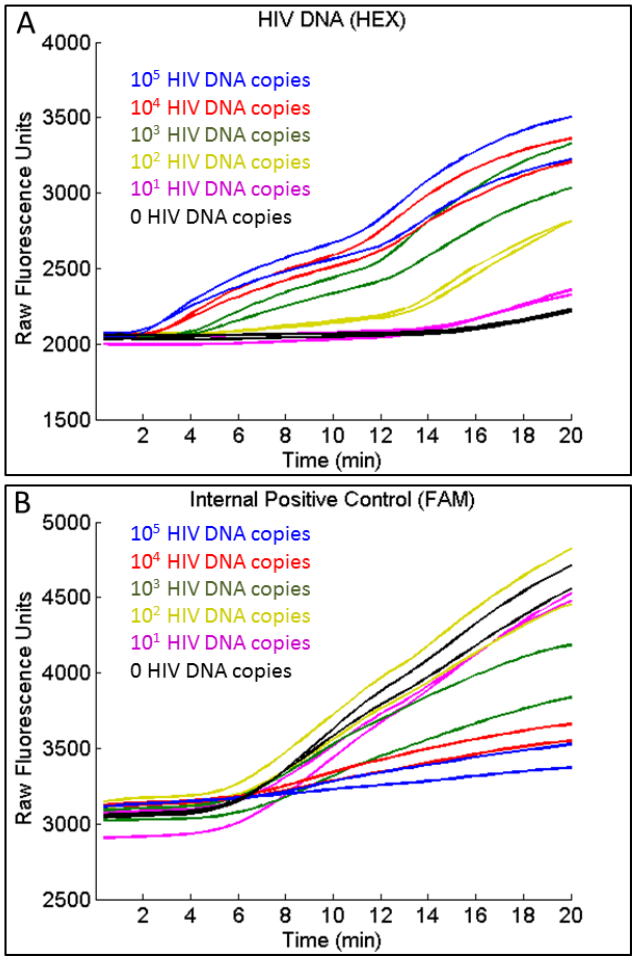

Figure 2: Typical raw fluorescence data for HIV-1 DNA and IPC amplification. (A) Raw fluorescence intensity of a HEX-labeled probe used for the detection of HIV-1 DNA at various concentrations. Note that the onset of detectable amplification occurred later for reactions containing fewer HIV-1 DNA copies. (B) Raw fluorescence of a FAM-labeled probe used for the detection of IPC DNA. The onset of detectable amplification for the IPC in all samples occurred at roughly the same time, but the rate of amplification was inversely proportional to the concentration of HIV-1 DNA.

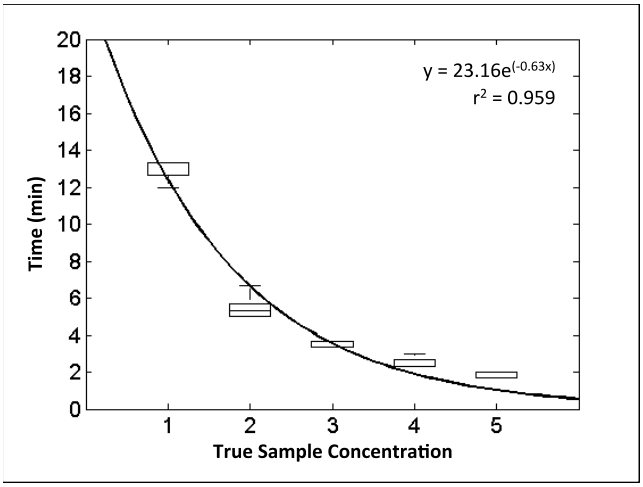

Figure 3: Generation of a standard curve. All positive samples from five experiments were plotted and used to generate exponential standard curves. All concentrations are given in units of $\log_{10}$ (copies per reaction). For the curve shown, samples were considered positive if Δ_{sample} was more than one $\sigma_{background}$ above $\Delta_{background}$ (z = 1). The quantification algorithm and standard curves were then used to estimate the number of HIV-1 DNA copies in each sample of five additional experiments.

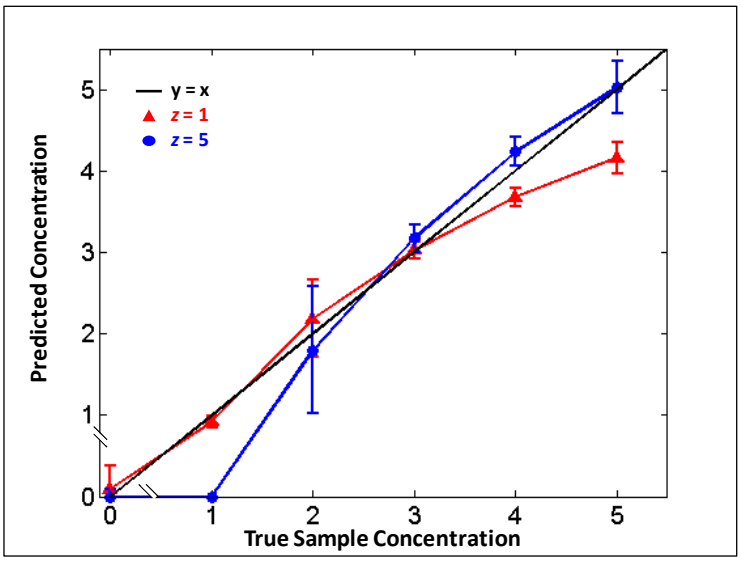

Figure 4: Effect of algorithm parameters on performance. The algorithm was used to generate standard curves and to predict the concentrations of samples from five experiments using $z = 1$ and $z = 5$. All concentrations are given in units of $\log_{10}$ (copies per reaction). When $z = 1$, predicted DNA concentrations are more accurate for lower concentrations than for higher concentrations. When $z = 5$, the sensitivity decreases for lower concentrations but the accuracy increases for higher concentrations. Thus, the algorithm can be tuned depending on clinical needs.

Sample Concentration	Average Predicted Concentration	Standard Deviation	Percent of Samples Identified as Positive
No target controls	0.1	0.3	10%
1	0.9	0.1	100%
2	2.2	0.5	100%
3	3.0	0.1	100%
4	3.7	0.1	100%
5	4.2	0.2	100%

Table 1: Performance of quantification algorithm. The algorithm was used to generate a standard curve and to predict the concentration of samples for five additional experiments (z = 1). All concentrations are given in units of $\log_{10}$ (copies per reaction).

For TOC Only