This is to certify that the
dissertation entitled

DEVELOPMENT AND VALIDATION OF A MULTIPLEX HAND-HELD GENE ANALYZER

presented by

ROBERT D. STEDTFELD

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Environmental Engineering

J. Alan Nason

Major Professor's Signature

Sep 16, 2009

Date

DEVELOPMENT AND VALIDATION OF A MULTIPLEX HAND-HELD GENE ANALYZER

By

Robert D. Stedtfeld

A DISSERTATION

Submitted to
Michigan State University
In partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Environmental Engineering

2009

ABSTRACT

DEVELOPMENT AND VALIDATION OF A MULTIPLEX HAND-HELD GENE ANALYZER

By

Robert D. Stedtfeld

The potential for microbial diagnostics to impact health care in developing regions of the world is well recognized. However, there is a large gap in the use of such technologies due to the unavailability of low cost point of care genetic screening tools. The device developed within this project fills this gap, with the potential for providing genetic screening at a cost that is well within reach of resource-limited regions (< $1,000). This was accomplished in a synergistic manner by means of: i) highly parallel validation of real time polymerase chain reaction (RT-PCR) amplification and hybridization-based molecular approaches, ii) development of a hand-held real-time gene analyzer, iii) validation of the device using isothermal loop mediated amplification (LAMP), and iv) embedding the component into a hand-held platform.

Validation of primers and probes was accomplished using an existing high throughput real time thermocycler for RT- PCR, and an in situ synthesized microarray for hybridization-based studies. The limits of the hybridization-based approach for detection of closely related fragments will be useful in developing assays for mutation detection. In addition to validating 220 primers sets targeting 30 microbial pathogens, a predictive equation for quantification of starting copies was experimentally established based on the sequence characteristics of primers and target organisms (as an efficient alternative to

generating standard curves). Validated primers can be used in the gene analyzer after it is fully developed.

Constraints in the development of truly portable and low cost gene analyzers are typically associated with the means of detection. Less costly components such as light emitting diodes (LEDs) and photodiodes are promising alternatives to charge coupled devices and lasers; however, simultaneous detection of multiple channels with spot detectors typically requires movement or assigning independent detectors to each channel on the device. To overcome this issue, an array of sixteen individually controlled LEDs (one LED for each reaction well) was used with optical fibers sending fluorescent signal to a single photodiode detector. Only one LED/well/optical fiber is turned "on" at a time and identity is determined by the associated sequence and timing of the photodiode reading. Optical cross talk between reaction wells was avoided using a thin-wall (shelled) microfluidic chip. The disposable chips were pressed in 100-μm thin Zeonor film by hot embossing with a male mold fabricated using stereolithography and a thermal plastic counter tool.

Components and control was embedded into a hand-held footprint through design of a custom printed circuit board, control with a microcontroller, and design of a cartridge to package the device. Comparable results were observed between the gene analyzer and a state-of-the-art real-time PCR instrument, demonstrating that this simple, small, low component cost system has the potential to play a key role in propagating the use of gene based screening devices in resource-limited settings.

This thesis is dedicated to my family for their support including: my Mother and Father, Brother and Sister, and to my Wife and Son for their patience and unconditional love.

ACKNOWLEDGEMENTS

I would like to thank my lab-mates for their contribution to this research and also for making the laboratory a fun and stimulating working environment. And thank Zheng Chen, a graduate student from the Department of Electrical Engineering, who helped me with basic electronics and control. I would also like to thank Scott Watson, Jimmy Johnson, Daphne Stoner, and Paul Meakin, whom I worked with at the Idaho National Laboratory for one year prior to attending graduate school at Michigan State University. They shared in my first experiences in research and showed me the fun side of science.

This work was supported by the National Institutes of Health (Grant R01 RR018625-01), Environmental Protection Agency (RD83162801-0 & RD83301001), and 21st Century Michigan Economic Development Corporation (GR-476 PO 085P3000517).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Images in this dissertation are presented in color.

CHAPTER ONE

INTRODUCTION

1.1 IMPACT OF LOW COST POINT-OF-CARE MICROBIAL DIAGNOSTIC DEVICES

Microbial agents are responsible for more than 75,000 deaths annually in the U.S. (11). Current screening practices consist of culture-based assays, smear tests, drug sensitivity testing, or real time PCR (RT-PCR). Culture based assays can take two to four days in most cases. For very slow growing organisms, e.g., multidrug or extensively drug resistance *Mycobacterium tuberculosis*, culture-based assay may take much longer (from weeks to months) because specimens must be shipped to a central testing facility. Genetic assays based on the presence/absence of quantification of a given gene or those capable of identifying specific mutations in a given gene have the potential to be much faster and more specific because: i) genes can be amplified by either polymerase chain reaction (PCR) with temperature cycling or by loop mediated isothermal amplification (LAMP). Based on the advancements taking place in the area of microfluidics, there are a number of efforts around the world to bring genetic assays into smaller packages and increase its ruggedness so that genetic marker-based assays can be deployed at the point-of-care. This is true for both amplification-based and hybridization-based assays. Combined with a suitable technique to break the microbial cells and expose the genomic DNA for amplification and/or hybridization, such devices have the potential to have a large impact on healthcare.

Reducing the time of genetic assays is expected to have a profound effect universally. Reducing the assay time from days to 15-20 minutes with a device that is not expensive (e.g., less than $1000), enable screening in clinical settings rather than

centralized labs. When the assay takes weeks to months because of its complex nature (e.g., detection of multiple mutations) and slow growth of the organisms (such as multiple or extensively drug-resistant tuberculosis MDR-XDR-TB), reducing the time and cost is equally critical. It is important to note that reducing the time has an indirect long-term (and perhaps intangible) benefit of minimizing the spread of antibiotic resistance. But often the immediate benefit alone is enough to justify the development of such approaches as evidenced by the outbreak of MDR-XDR TB in KwaZulu Natal during which 70 percent of those inflicted died within 30 days of initial sputum collection (12).

Essential characteristics of an ideal fieldable screening tool include: i) rapid time to completion, ii) rugged and user friendliness, iii) low cost, iv) hand-held or portable, v) high specificity, vi) superior detection sensitivity, vii) quantitative, and viii) simultaneous detection of multiple gene or mutation signatures. Individually, these characteristics are available in molecular diagnostics devices; however, no such device has been developed that implements all of these characteristics. Simultaneous detection of multiple signatures is important for two reasons, one for the added ability of incorporating redundant targets and controls, and the other for the need to diagnose mutations. Even for point-of-care devices, which are supposed to be inherently simple, the level of multiplexing of signatures will vary based on the application and/or organisms screened. For example, screening for MDR-XDR TB requires an assay that targets hundreds of genetic mutation associated with resistance against anti-TB drugs to determine the suitable means of treatment. On the contrary, an assay for methicillin resistant *Staphylococcus aureus* (MRSA) will require half a dozen assays with specificity at the

gene level. It is evident that to develop powerful point-of-care devices, amplification, hybridization, and even ligation based mutation detection approaches may need to be incorporated individually or in sequence to address various diagnostic needs. It is evident that the availability of portable and low cost devices at a point-of-care (POC), such as emergency first responders, healthcare centers, or physician's will aid in the realization of a more rapid means of suitable treatment, reduce fatality, and decrease emergence of strains with increasing resistance.

Besides the selected concerns noted above which are severe problems within the U.S. (e.g., MRSA) or considered a global health emergency (such as for MDR-XDR TB), there is also a broader range of diseases that may benefit from the development and availability of such devices. According to Peter Hotez, a leading expert in neglected tropical diseases (NTDs), globally the total number of lives and disability adjusted lifeyears lost from HIV/AIDS, NTDs, Malaria, and Tuberculosis exceeds 220 million (5). Based on the potential for massive social and economic impacts within and outside of third world countries, rural areas, and resource-poor regions, there is a growing interest for low cost POC devices (2, 4, 7, 9, 10, 16). Estimates for the global POC diagnostics market in 2005 was $32.2 billion, which was predicted to grow to $45.6 billion by 2010, with infectious disease testing contributing an estimated 37% of this market. With the US and Europe holding 90% and the rest of the world accounting for the remaining 10%, the availability of a low-cost point of care device could substantially increase markets in the developing world and low-resource settings.

1.2 COMPARATIVE EVALUATION OF SELECTED REAL TIME PCR DEVICES

Of the many genetic detection approaches, real time PCR (RT-PCR), first described in 1983, offers the highest level of sensitivity and quantification. Despite the advantages of this technique, there are still less than a dozen PCR diagnostic tests (17) that are approved by the Food and Drug Administration. This is due to the high cost of RT-PCR instruments, significant training to run the instrument, and need for storage of reagents. As listed in Table 1.1, a number of options are available to carry out this assay. This is not an exhaustive list and certain instruments that are still under development have been left out due to lack of information and also because they are known mostly through publications and have not yet arrived in the market.

Gene-based assays using RT-PCR generally take about 3-4 hours to complete and FDA has approved some of these assays for clinical use (e.g., for methicillin resistant *Staphylococcus aureus*). However, the cost of real-time PCR machines (tens of thousands of dollars; Table 1.1) generally requires that the specimens be sent to a centralized laboratory. This delays diagnosis causing the prescription of broad-spectrum antibiotics or delaying treatment, increasing potential for the emergence of a new more resistant strain and fatality. When portable options are available (albeit at very high prices), the validation of genes is done on a very limited set of organisms (e.g., anthrax for bio-Seeq and Bird Flu for Palm-PCR of Neuzil's group). For other pathogens of interest, users are expected to develop the assay on their own (perhaps in collaboration with the equipment suppliers). The high cost and the need for further development virtually make it impossible for these options to serve as truly low cost point-of-care options for most other applications.

Since all devices use amplification based strategies, parameters such as sensitivity are similar for all RT-PCR machines; however, only 2 of the portable devices listed in Table 1.1 are quantitative. The Hand-Held offers the same level of throughput as the other portable devices in terms of targets that can be assayed simultaneously. Potential cross contamination problems that can be associated with dispensing multiple samples in a small space make multiplexing with respect to samples less attractive. However, since both the equipment and the chip are inexpensive, and the time of assay is short, analyzing single samples is not a significant disadvantage. It is the gap in the availability of low cost genetic analysis point of care system (from here onward referred to as gene analyzer) that the current research aims to fill.

Table 1.1: Comparison of parameters for representative RT-PCR machines

	PCR Chip Hand-Held	Neuzil's microPCR	Smiths Detection (Bio-Seeq)	Idaho RAZOR	PCR-Chip Bench top	Idaho R.A.P.I.D.	BioTrove NTCycler
Validation Focus	MRSA	Bird Flu	Anthrax	Open	Open	Open	Open
Base Unit	$1,000	Unknown	$31,000	$35,000	$25,000	$55,000	$100,000
Consumable Price (U)	$10	Unknown	$38	$18	$100		$400
Time	5 min	7 min	20-30 min	30 min	30 min	30 min	3.5 hr
# of Samples	1	1	6	12	56	10	48
# of Targets	~5-10	~5-10	1-4	12	50	10	18-56
Sample volume	1000 nl	1000 nl	10 µl	5-20 µl	5-50 µl	5-20 µl	33 nl
Quantitation	Yes	Yes	No	No	Yes	Yes	Yes
Sensitivity	10 copies	10 copies	5 copies	5 copies	5 copies	5 copies	5 copies
Portability (Weight)	< 1 lb Goal*	5 lb	6.5 lb	9 lb	30 lb	50 lb	100 lb
							

1.3 SELECTION OF PARAMETERS IN DEVELOPING THE GENE ANALYZER

Numerous prototype microfluidic chips and analyzers have been described in literature that may be capable of carrying out PCR in hand-held or portable devices (1, 3, 6, 8, 13-15). Despite indications that development is near completion, few systems are commercially available that are hand-held, multiplex, and cost-efficient. This is due to several reasons including: i) the sensitivity required for RT-PCR amplification requires bulky, expensive, complex, and power consuming systems for excitation (mercury lamps or lasers with a number of optical filters) and detection of light (Photomultiplier tubes and high cost charged coupled devices), ii) lack of expertise in embedding devices within the settings of a single university laboratory, and iii) complexity and/or lack of portability associated with real time detection schemes for multiple assay amplification chips. In essence, innovation, adaptation, and cost reduction are still needed before hand-held microbial diagnostics technologies can impact health care in a point-of-care scenario with resource-limited settings (17).

As evident, a number of choices must be made and depending upon those choices, the cost, portability and power of the device will be determined (Figure 1.1). Often, the application may dictate that multiplexing with respect to gene (e.g., for MRSA) or mutation (e.g., for MDR/XDR TB) must be incorporated in order for the device to be useful for that application. Based on the molecular strategies adopted, essential characteristics, and the application-defined requirements (level of specificity and throughput), three separate products are being developed in a mutually synergistic manner, each suitable for a specific application. Table 1.2 (at the end of this chapter) lists the target value for many of the parameters that are necessary for a device of this nature. TB and MDR/XDR TB are taken as examples of two separate applications, one requiring

the detection of a limited set of genes (similar to gene analyzer) and the other requiring the detection of a large number of mutations.

1) **Photodiode-based gene analyzer**: Gene analyzer allows selection of LAMP or RT-PCR molecular amplification with an optical setup consisting of an array of sixteen individually controlled blue light emitting diodes (LEDs), and a single photodiode to read the sixteen reaction wells one by one using individual optical fibers. Figure 1.1 illustrates the approximate values chosen for the gene analyzer. ***Only the gene analyzer is the topic of this research.***
2) **CCD-based LAMP Array**: This product focuses on the measurement of approximately 1,000 LAMP reactions in real time (with options surface based hybridization to increase throughput) using a CCD camera with exposure control. Signals are extracted as SNR using custom software being developed as part of the project. Temperature control, fabrication technique of disposable chips, and microcontroller are based on the developments made for the gene analyzer.
3) **Single Photon Detection-based Rotating Disk Array**: This product uses a rotating disk instead of an optical fiber and LED harness system to move a circular chip underneath a fixed single photon detector set up. This allows rapid monitoring of the amplification product in multiple wells (the number of wells will depend upon the diameter of the chip but it could potentially be much larger in 100s to 1000s. Temperature control, fabrication technique of disposable chips, and microcontroller are similar to that used in the gene analyzer.

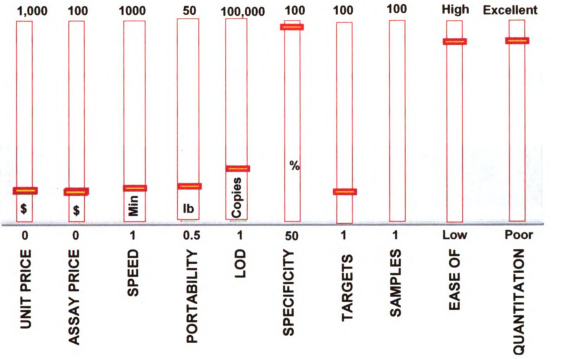


Figure 1.1. Depiction of the options for various parameters related to an analytical device illustrating various levels of equipment price, assay price, speed, portability, limit of detection, specificity, number of targets, number of samples, ease of use, and quantitation. Approximate values of these parameters are shown here for gene analyzer. Similar choices have been made for other instruments where speed, and the number of targets and number of samples are also maximized (e.g., single photon detector based rotating disk analyzer and CCD based LAMP microarray instrument).

1.4 Overview of the Project

Ideally an instrument must be developed first and the assay validation follows. However, to save time and due to the sheer volume of the assays to be validated, a parallel approach was adopted meaning the assays were validated on existing real time PCR instruments and hybridization systems with the assumption that when the hand-held gene analyzer is developed, these assays will be ported and if needed optimized on the portable system. As illustrated in Figure 1.2, the overall project was highly team oriented, which is generally not encountered in academic setting where students work on his/her topic with limited interaction to other students. However, this project required a

large number of interdependent tasks, ranging from assay design to microfluidic chip development and validation to selection of integrated circuits and printed circuit board design and its evaluation (Figure 1.2). Seven of eight team members worked at Michigan State University and one member worked at the University of Michigan.

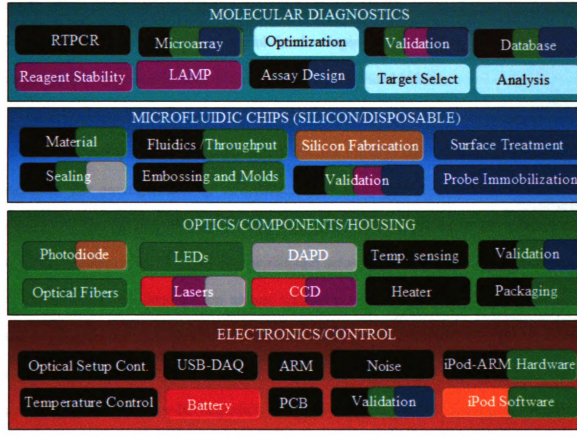


Figure 1.2. Major considerations for the devices are divided into four categories. These aspects have been subdivided into developmental tasks. Color codes refer to persons that have taken the lead on a task or share the task with other (light blue represents everyone, Robert Stedtfeld in black, Dieter Tourlousse in green, Tiffany Stedtfeld in dark blue, Farhan Ahmed in red, Grégoire Seyrig in purple, Deric Patton in gray, Aaron Thompson in orange, and Onnop Srivannavit in brown).

1.5 ORGANIZATION OF THIS THESIS

The thesis is divided into two main sections: ***Part I***: Validation of amplification and hybridization-based molecular approaches, and ***Part II***: Development and validation

of the hand-held real-time gene analyzer. Part I consist of two chapters that are already published (out of 15 publications). Chapter 2 focuses on the validation of primers using an existing high throughput real time PCR machine (Biotrove Inc). Primers similar to these will be used in the gene analyzer after it is fully developed. Chapter 3 evaluates the limits of hybridization-based approach for detection of closely related fragments. Lessons learned in this chapter will be useful in developing hybridization-based assays for mutation detection. A summary of these chapters is provided below.

Chapter 2 Development and experimental validation of a predictive threshold cycle (Ct) equation for quantification of virulence and marker genes using high-throughput nanoliter PCR on OpenArrays™: Development of quantitative PCR (QPCR) assays typically requires extensive screening within and across a given species to ensure specific detection and lucid identification among various pathogenic and nonpathogenic strains and to generate standard curves. To minimize screening requirements, multiple virulence and marker genes (VMGs) were targeted simultaneously to enhance reliability, and a predictive threshold cycle (Ct) equation was developed to calculate starting copies based on experimental Ct. The empirical equation was developed with SYBR Green detection in nanoliter QPCR chambers (OpenArray™) and tested with 220 previously unvalidated primer pairs targeting 200 VMGs from 30 pathogens. A high correlation ($R^2 = 0.816$) was observed between predicted and experimental Ct based on the organism's genome size, GC content, amplicon length, and stability of primer 3' end. The performance of the predictive Ct equation was tested using 36 validation samples consisting of pathogenic organisms spiked into genomic DNA extracted from three environmental waters. In

addition, primer success rate was dependent on GC content of the target organisms and primer sequences. Targeting multiple assays per organism and using the predictive Ct equation is expected to reduce the extent of validation necessary when developing QPCR arrays for a large number of pathogens or other targets.

Chapter 3 Influence of Dangling Ends and Surface-Proximal Tails of Targets on Probe-Target Duplex Formation in 16S rRNA Gene-Based Diagnostic Arrays: Dangling ends and surface-proximal tails of gene targets influence probe-target duplex formation and affect the signal intensity of probes on diagnostic microarrays. This phenomenon was evaluated using an oligonucleotide microarray containing 18-mer probes corresponding to the 16S rRNA genes of 10 waterborne pathogens and a number of synthetic and PCR-amplified gene targets. Signal intensities for Klenow/random primer-labeled 16S rRNA gene targets were dissimilar from those for 45-mer synthetic targets for nearly 73% of the probes tested. Klenow/random primer-labeled targets resulted in an interaction with a complex mixture of 16S rRNA genes (used as the background) 3.7 times higher than the interaction of 45-mer targets with the same mixture. A 7-base-long dangling end sequence with perfect homology to another single-stranded background DNA sequence was sufficient to produce a cross-hybridization signal that was as strong as the signal obtained by the probe-target duplex itself. Gibbs free energy between the target and a well-defined background was found to be a better indicator of hybridization signal intensity than the sequence or length of the dangling end alone. The dangling end (Gibbs free energy of –7.6 kcal/mol) was found to be significantly more prone to target-background interaction than the surface-proximal tail

(Gibbs free energy of –64.5 kcal/mol). This study underlines the need for careful target preparation and evaluation of signal intensities for diagnostic arrays using 16S rRNA and other gene targets due to the potential for target interaction with a complex background.

Part II consist of three chapters (Chapters 4-6). Chapters 4 and 5 focus on the development of gene analyzer and Chapter 6 is a summary of this thesis and perspectives toward further development and commercial use of the gene analyzer. A summary of Chapters 4 and 5 is provided below.

Chapter 4 Development of Embedded Control Systems for Gene Analyzer: The selection and control of components toward the development of a hand-held device for real time detection of loop mediated amplification is described in detail. Initial control of components was obtained using National Instruments LabVIEW with USB compatible data acquisition accessories and several integrated circuits (ICs). A rapid response thermocouple, PWM driver IC, and PID feedback-control was used for precise temperature control of a Kapton heater with a standard deviation from setpoint of 0.28 ^{o}C. The optical setup for capturing real time amplification signals in a multichamber well consisted of sixteen light emitting diodes (LEDs), optical fibers sending fluorescent light from each well to a single photodiode, which was used to acquire change in signal emission. Development of optics control included individual control of all 16 LEDs with a 4 to 16 line decoder IC, generating accurate and constant current for all LEDs with a LED driver IC, synchronized timing of LEDs with the photodiode (PD), and reducing and conditioning noise acquired from the PD. Once a control strategy was established and used to test components, a mechanism was selected and tested for embedded control.

This included the design of a custom PCB, replacing USB-DAQs and PC with an ARM microcontroller, designing a cartridge to package the device, and devising a circuit to communicate with an iPod Touch.

Chapter 5 Validation of Optical System for Gene Analyzer: The development of a hand-held real-time gene analyzer to detect amplification products from loop mediated isothermal amplification (LAMP) is described. The high yield of amplification product and isothermal characteristics of *Bst* polymerase allow the use of simple and low cost components to drive the reaction and detect the resulting product. The device consist of: i) a Kapton resistive heater to isothermally heat a chip with sixteen wells, ii) an array of sixteen blue light emitting diodes (LEDs), and iii) a single photodiode to read the sixteen wells one by one using individual optical fibers. The system is simple because it eliminates costly lenses, CCD imaging, or moving parts. The microfluidic chips consist of sixteen 3-mm long capillaries embossed in 100-μm thin Zeonor film using a modified heat embossing technique. A positive mold with shoulders fabricated using stereolithography and a counter tool consisting of acrylonitrite butadiene styrene (ABS) having low glass transition temperature was used to fabricate the disposable chips. The shoulders allowed a larger process window and reproducibility. The molded ABS tool served as both an alignment guide between the chips and optical components and optically isolated the reaction wells. Real time LAMP experiments performed with *Staphylococcus aureus* methicillin resistance gene (*mecA*) using calcein as a fluorescent dye yielded results comparable to a state-of-the-art real-time nucleic acid amplification

instrument. This simple, small, and low-cost system is expected to play a key role in propagating the use nucleic acid assays in resource-limited settings.

1.6 OBJECTIVES OF THIS RESEARCH

The goal of this research was to develop a low cost and rugged quantitative nucleic acid analyzer, offering a high level of specificity. The specific objectives were:

1. Validate a large set of real time PCR primers to be employed in any on-chip device.
2. Evaluate the extent of differentiation between two similar sequences by on-chip hybridization.
3. Develop a portable gene analyzer capable of quantitatively detecting gene signatures in DNA matrix. The boundary conditions for this device were:

   Component cost: less than $1000 in component cost

   Assay cost: less than $5 per sample

   Weight: less than 1 pound

   Size: a hand-held system similar in size to an iPod Touch

As part of the gene analyzer system, the following tasks were identified:

- Development of microfluidic chips.
- Evaluation and selection of temperature and optical components for heating and detecting amplification.
- Control of temperature and optical components with LabVIEW.
- Embedded control with ARM Microcontroller.
- Housing design

- Interfacing device with an iPod.
- Validation of all aspects.

I have taken the lead role and/or shared the following tasks interdependently:

- Target selection, database development, design, and validation of RT-PCR and microarray based assays.
- Development of microfluidic chip including design and fabrication of molds, optimization of hot embossing process for reproducible fabrication of chips, enclosing and sealing chips.
- Selection of heaters, temperature sensors, and development of feedback-control.
- Devising mechanism to control optical components.
- Selection and evaluation of ARM microcontroller for embedded control of components.
- Design of PCB to reduce size of circuit.
- Selection of ICs for regulating voltage to various components within packaged device.
- Validation and characterization of gene analyzer.

Table 1.2 Comparison of user needs specified in the Customer Requirement Document-Primary Care Level (provided by Foundation for New Innovative Diagnostics or FIND) to the proposed capabilities of various devices in this project.

	Desired	Minimum	Routine TB Assay (Gene Analyzer)	MDR/XDR TB Assay (SPD and CCD based Assay)
A. Workflow				
1. Sample				
1.1. Sample type	Sputum, skin, breath, urine (at least one of these sample types)	Sputum, skin, breath, urine, blood, serum (at least one of these sample types)	Sputum	Sputum
1.2. Sample processing	None	Simple 1 to 10 two-step procedure with less than 15 min hands-on time	None, sample processing integrated in the microfluidic chips/device	None, sample processing integrated in the microfluidic chips/device
2. Time-to-result	Less than 10 min, including sample processing and readout	Less than 2 hrs, including sample processing and readout	Less than 20 min, including sample processing and readout	Less than 30 min, including sample processing and readout

Table 1.2 continued

3. Instrumentation	None	None, or small, maintenance-free device	Portable and battery-powered device	Portable and battery-powered device
4. Additional equipment	None	Less than 2 devices with small footprint	None	None
5. Bio-safety	No need for biosafety cabinet, Disposal of clinical material without further treatment	No need for biosafety cabinet	No need for biosafety cabinet due to fully closed and self-contained assays in microfluidic chips	No need for biosafety cabinet due to fully closed and self-contained assays in microfluidic chips
B. Diagnostic performance				
1. Sensitivity				
1.1. Diagnostic sensitivity	Either: a) rule-in-test: >75% b) rule-out-test: >95%	Either: a) rule-in-test: >30% b) rule-out-test: >85%	Sensitivity: 90% or higher	Sensitivity: 90% or higher
2. Specificity				

Table 1.2 continued

2.1. Diagnostic specificity	Either: a) rule-in-test: >97% b) rule-out-test: >50%	Either: a) rule-in-test: >97% b) rule-out-test: >50%	Specificity: higher than 95%	Specificity: close to 100%
3. Effect of HIV co-infection	No change in performance	50% fall in sensitivity in HIV infected patients	Similar to that of other nucleic acid based assays	Similar to that of other nucleic acid based assays
C. Product design				
1. Stability/storage requirements				
1.1. Kit stability	36 months at 35°C and 70% humidity – including transport stress (48 hrs at 50°C)	9 months at 30°C and 3 months at 35°C including transport stress (48 hrs at 50°C)	At least 12 months at 30°C and 3 months at 35°C	At least 12 months at 30°C and 3 months at 35°C

Table 1.2 continued

1.2. Reconstituted reagents stability	Ready-to-use units, no reconstitution	1 week at 35°C	Ready-to-use and single-use microfluidic chips with stabilized reagents, no need to store reconstituted reagents	Ready-to-use and single-use microfluidic chips with stabilized reagents, no need to store reconstituted reagents
2. Calibrators	Fixed cut-off	Fixed cut-off	LAMP: time-to-positivity	LAMP: time-to-positivity Microarray: signal-to-noise ratio
3. Controls	Full-process positive control	No control	Internal cellular control (full-process positive control)	Internal cellular control (full-process positive control)

Table 1.2 continued

4. Determinations per 'reconstituted reagent unit'	Ready-to-use units, no reconstitution	Opened packaged should last for at 2 weeks	Ready-to-use and single-use microfluidic chips with stabilized reagents, single determination/chip	Ready-to-use and single-use microfluidic chips with stabilized reagents, single determination/ chip
5. Results capturing and documentation	Yes	No	Yes, CCD imaging and image/data analysis	Yes, CCD imaging and image/data analysis
6. Instrument design	No instrumentation or battery operated. Device platform can also be used for other infectious agents	115/220 V AC Operates at 35°C	Battery-operated. Can be used for other infectious agents by using microfluidic chips with different primer sets	Battery-operated. Can be used for other infectious agents by using microfluidic chips with different primer sets and probes

Table 1.2 continued

7. Training and education requirements	Less than 1 hour, high school doplaome or equivalent	Less than 2 days, nurse level skill set	Less than 1 hour due to full automation, high school doplaome or equivalent	Less than 2 days, nurse level skill set, most skills in data interpretation
D. Additional features				
			Routine TB: Quantitative Equipment cost: $1,000 $0.25 per assay for six genes	- Routine TB: Quantitative - MDR/XDR TB: Plus/Minus Equipment cost: $1,000 <$5 per assay

1.7 REFERENCES

1. **Bhattacharyya, A., and C. M. Klapperich.** 2007. Design and testing of a disposable microfluidic chemiluminescent immunoassay for disease biomarkers in human serum samples. Biomed. Microdevices. **9:**245-251.

2. **Herr, A. E., A. V. Hatch, D. J. Throckmorton, H. M. Tran, J. S. Brennan, W. V. Giannobile, and A. K. Singh.** 2007. Microfluidic immunoassays as rapid saliva-based clinical diagnostics. Proc. Natl. Acad. Sci. U.S.A. **104:**5268-73.

3. **Hofmann, O., X. Wang, A. Cornwell, S. Beecher, A. Raja, D. D. Bradley, A. J. Demello, and J. C. Demello.** 2006. Monolithically integrated dye-doped PDMS long-pass filters for disposable on-chip fluorescence detection. Lab Chip **6:**981-987.

4. **Holland, C. A., and F. L. Kiechle.** 2005. Point-of-care molecular diagnostic systems — past, present and future. Curr. Opin. Microbiol. **8.**

5. **Hotez, P.** 2009. Devastating global impact of neglected tropical diseases. Microbe **4:**366-372.

6. **Huen, W., A. Chow, C. McCloskey, Y. Tong, L. Hu, Q. You, C. P. Kelly, H. Kong, Y. W. Tang, and W. Tang.** 2008. Application of isothermal helicase-dependent amplification with a disposable detectio device in a simple sensitive stool test for toxigenic *Clostridium difficile*. J. Molec. Diag. **10:**452-458.

7. **Inganas, M., H. Derand, G. Eckersten, A.-K. Ekstrand, G. Honerud, G. Jesson, T. Thorsen, T. Soderman, and P. Andersson.** 2005. Integrated microfluidic compact disc device with potential use in both centralized and point-of-care laboratory settings Clin. Chem. **51:**1985–1989.

8. **Liao, C. S., G. B. Lee, H. S. Liu, T. M. Hsieh, and C. H. Luo.** 2005. Miniature RT–PCR system for diagnosis of RNA-based viruses Nucleic Acids Res. **33:**e156.

9. **Linder, V.** 2007. Microfluidics at the crossroad with point-of-care diagnostics. Analyst **132:**1186 - 1192.

10. **Malamud, D., H. Bau, S. Niedbala, and P. Corstjens.** 2005. Point detection of pathogens in oral samples. Adv. Dent. Res. **18:**12-16.

11. **Mokdad, A. H., J. S. Marks, D. F. Stroup, and J. L. Gerberding.** 2000. Actual causes of death in the United States. JAMA **291:**1238-1245.

12. **Panel, T. T. D. E. E.** 2008. Evaluation of diagnostic tests for infectous diseases: general principles.

13. **Pipper, J., Y. Zhang, P. Neuzil, and T. M. Hsieh.** 2008. Clockwork PCR including sample preparation. Angew. Chem. Int. Ed. Engl. **47:**3900–3904.

14. **Smith, M. C., G. Steimle, S. Ivanov, M. Holly, and D. P. Fries.** 2007. An integrated portable hand-held analyser for real-time isothermal nucleic acid amplification. Anal. Chim. Acta **598:**286-294.

15. **Wang, J., Z. Chen, P. L. Corstjens, M. G. Mauk, and H. H. Bau.** 2006. A disposable microfluidic cassette for DNA amplification and detection. Lab Chip **6:**46-53.

16. **Yager, P., T. Edwards, E. Fu, K. Helton, K. Nelson, M. R. Tam, and B. H. Weigl.** 2006. Microfluidic diagnostic technologies for global public health. Nature **442:**412–418.

17. **Yager, P., J. Gonzalo, G. J. Domingo, and J. Gerdes.** 2008. Point-of-Care diagnostics for global health. Annu. Rev. Biomed. Eng. **10:**107–44.

CHAPTER TWO

Stedtfeld, R. D., S. W. Baushke, D. M. Tourlousse, S. M. Mller, T. M. Stedtfeld, E. Gulari, J. M. Tiedje, and S. A. Hashsham. 2008. Development and experimental validation of a predictive threshold cycle equation for quantification of virulence and marker genes by high-throughput nanoliter-volume PCR on the OpenArray platform. Appl. Environ. Microbiol. **74:**3831-8.

DEVELOPMENT AND EXPERIMENTAL VALIDATION OF A PREDICTIVE THRESHOLD CYCLE (CT) EQUATION FOR QUANTIFICATION OF VIRULENCE AND MARKER GENES USING HIGH-THROUGHPUT NANOLITER PCR ON OPENARRAYS™

2.1 ABSTRACT

Development of quantitative PCR (QPCR) assays typically requires extensive screening within and across a given species to ensure specific detection and lucid identification among various pathogenic and nonpathogenic strains and to generate standard curves. To minimize screening requirements, multiple virulence and marker genes (VMGs) were targeted simultaneously to enhance reliability, and a predictive threshold cycle (Ct) equation was developed to calculate starting copies based on experimental Ct. The empirical equation was developed with SYBR Green detection in nanoliter QPCR chambers (OpenArray™) and tested with 220 previously unvalidated primer pairs targeting 200 VMGs from 30 pathogens. A high correlation (R^2 = 0.816) was observed between predicted and experimental Ct based on the organism's genome size, GC content, amplicon length, and stability of primer 3' end. The performance of the predictive Ct equation was tested using 36 validation samples consisting of pathogenic organisms spiked into genomic DNA extracted from three environmental waters. In addition, primer success rate was dependent on GC content of the target organisms and primer sequences. Targeting multiple assays per organism and using the predictive Ct equation is expected to reduce the extent of validation necessary when developing QPCR arrays for a large number of pathogens or other targets.

2.2 INTRODUCTION

Detection and identification of multiple pathogens requires simultaneously targeting a large number of virulence and marker genes (VMGs) (14, 31, 35, 43, 50, 54). This is because: i) use of multiple markers for the same pathogen enhances specificity and reliability of the assay, ii) often the allelic variability requires multiple primers for the same gene, iii) some of the markers may be less specific (i.e., they may also be found in nonpathogens, and iv) the VMGs for a given pathogen may be unevenly distributed among various strains. Use of a varied number of VMGs to differentiate clinical isolates has been described for many pathogenic organisms including: *Staphylococcus aureus* (3, 49), *Aeromonas* spp. (8), *Campylobacter jejuni* (39), *Pseudomonas aeruginosa* (13), *Vibrio* spp. (50), *Escherichia coli,* and *Shigella flexneri* (5, 53).

When many pathogens must be screened in parallel with the ability to characterize the uneven distribution of the associated VMGs and their allelic variability, optimization is necessary for high throughput tools with high sensitivity and specificity, e.g., quantitative PCR (QPCR) (6, 14, 19, 27, 33, 45, 51, 52). This optimization may be cumbersome because it requires generation of multiple standard curves with the caveat that all primer sets must perform under the same amplification conditions. Approaches that increase reliability of primer design or avoid the use of standard curves all together should be extremely useful in developing such parallel assays. Development of such approaches will undoubtedly depend upon the sequence

characteristics (GC content, Tm, amplicons length, etc.) of primers, amplicons, genome size, amplification conditions, and the matrix in which the target is present. However, the influence of these factors on the performance of primer sets, especially on the threshold cycle number (Ct) extensively used in QPCR to predict the abundance of pathogens has not been explored fully.

This study used the nanoliter volume BioTrove OpenArray™ platform (34) to examine its capacity for highly parallel diagnostics of human pathogens, and to systematically examine the influence of target and primer sequence characteristics on specificity, sensitivity and Ct value. Results were used to establish a predictive Ct equation to estimate starting copy without the use of standard curves. The study was performed with SYBR Green I and used approximately 220 primer sets targeting 200 VMGs for 30 human pathogens. Performance of the predictive Ct equation is examined and the success rate of previously un-validated pathogen targeted primer sets is also presented. These results have significance in developing high throughput and reliable screening tools for large numbers of pathogens without extensive validation.

2.3 Materials and Methods

2.3.1 DNA targets. Multiple organisms were used to develop and validate the predictive Ct equation. Genomic DNA (gDNA) from 21 bacterial pathogens (Table 2.1) was used including: *Clostridium perfringens, Enterococcus faecalis, Listeria monocytogenes, Legionella pneumophila, Pseudomonas aeruginosa, Salmonella enterica, Vibrio parahaemolyticus, Yersinia enterocolitica* type strains, which were

obtained from the American Type Culture Collection (ATCC; Manassas, VA) and grown as per the protocol provided. For *Helicobacter pylori, Campylobacter jejuni, Cryptosporidium parvum, Giardia intestinalis, Staphylococcus aureus, Vibrio cholerae, Mycobacterium genetalium, Haemophilus influenzae, Leptospira interrogans, Bacillus cereus, and Bordetella pertussis,* gDNA was obtained directly from ATCC. Genomic DNA for *Escherichia coli* was kindly provided by Dr. Thomas Whittam at Michigan State University. DNA from pure cultures was extracted using Promega's Wizard DNA Extraction Kit (Promega, Madison, WI).

2.3.2 Complex background gDNA. Genomic DNA from pure cultures was spiked into complex background gDNA extracted from environmental waters. River water was collected from the Red Cedar River and activated sludge and tertiary effluent from a wastewater treatment plant in East Lansing (MI). Samples were filtered through 0.45 μm nitrocellulose filters (Millipore, Billerica, MA) immediately after collection. Genomic DNA was extracted from the filters as instructed with the MegaPrep UltraClean™ Soil DNA Kit (Mo Bio Laboratories, Carlsbad, CA). Genomic DNA was purified further through ethanol precipitation. DNA quantification and purity was examined with the NanoDrop® ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE). Prior to QPCR experiments, background DNA was tested for inhibition by PCR amplification with a universal primer set targeting the 16S rRNA gene.

2.3.3 Design of PCR primers to establish the predictive Ct equation. This set of primers was designed to target 20 human pathogens, and was used to examine relationships between primer sequence characteristics and experimental Ct. Prior to primer design, consensus sequences were generated by aligning sequences for each of 96 VMGs using Kodon (Applied Maths, Austin, TX). The consensus sequence was used for primer design with Primer Express (Applied Biosystems, Foster City, CA). A majority of designed amplicons had a maximum length of 150 bases and primers had a theoretical melting temperature of 59 °C. Some genes required longer amplicons (<250 bases) for generating acceptable primer sets. From the list of primers provided by Primer Express, optimal primers were selected using a script that automatically highlighted unspecific primers. The script used the NCBI BLAST (2) to check specificity against the GenBank database. Specificity was based on the extent of 3' end perfect matches to non-targeted bacterial sequences. Sequences were selected manually based on results of BLAST output. When available, primers described in the literature for successful QPCR were also used. Overall, 110 primers pairs were designed and extracted from the literature. Including primers previously described in the literature, 3,687 VMG sequences were targeted with this primer set (determined by BLAST analysis with GeneBank, May 2006).

2.3.4 Design of PCR primers to validate correlations used in development of predictive Ct equation. To validate the predictive Ct equation and further examine the success rate of primers, 111 new primers were designed. The new group of primers were designed with the same criteria as primers designed to establish the

predictive CT equation. In addition, primers were filtered further to select sequences with lowest possible percentage of GC bases in the 3' end of the primer. One non-specific primer targeting *C. parvam* was removed from the set due to false positive observations. Thus, a total of 220 primers sets (109 used to establish the predictive Ct equantion, and 111 used to validate the predictive Ct equation) targeting 200 VMGs for 30 pathogenic bacteria were tested (Table 2.1, the new VMGs targeted with this primer set are in bold).

2.3.5 PCR on BioTrove OpenArray™ plates. Primer sets were tested simultaneously on the BioTrove OpenArray™ plates. Primers were synthesized by Sigma-Aldrich (St. Louis, MO) and pre-loaded (128 nM for primers designed to establish the predictive Ct equation and 400 nM for primers designed to validate the predictive Ct equation) into BioTrove OpenArray™ plates (Woburn, MA)(34, 47). Two to four subarrays (each with 64 wells for 56 separate assays and eight loading controls) were used for each PCR sample. PCR mixtures (5 µl for each sample array) consisted of 1 x LightCycler Fast Start DNA Master SYBR Green I Kit (Roche Applied Sciences, Indianapolis, IN), 1.6 x SYBR Green I, 0.5% Glycerol, 0.2% Pluronic F-68, 1 mg per ml BSA (New England Biolaboratories, Beverly, MA), 2.5 mM $MgCl_2$, 8% formamide, and a DNA mixture. After initial enzyme activation at 95 °C for 10 min, 36 cycles of the following program was used for amplification: denaturation at 95°C for 10 s, annealing at 53 °C for 10 s, elongation at 72 °C for 10 s.

2.3.6 Design of sample mixtures for development of predictive Ct equation. Samples were mixed to evaluate the dependence of sequence characteristics of primers, amplicons, genome size, amplification conditions, and the matrix in which the target is present on specificity and sensitivity. To evaluate specificity without a complex background, gDNA from 14 pathogenic organisms (pure cultures; ATCC number listed in Table 2.1) was tested individually (6 ng in total sample or 20 pg per reaction well). To develop standard curves and further evaluate specificity and sensitivity within a complex background, gDNA from 14 pathogens was mixed and spiked at various concentrations into gDNA from waste water tertiary effluent and river water samples (20 pg, 2 pg, 200 fg, 20 fg, 2 fg of each of 14 pathogens mixed together spiked into 66.6 pg of background gDNA per reaction well). The mixture of varying concentrations was also examined without a background to serve as a control. Complex background samples were also tested without spiking with gDNA from pure cultures. All of these samples were tested on the set of PCR primers designed to establish the predictive Ct equation. These samples were examined further to evaluate influence of sample inhibition and variation between OpenArray™ plates. All samples were tested in triplicate.

2.3.7 Design of validation sample mixtures for validation of predictive Ct equation. Validation samples were used to further evaluate the characteristics of template and primer sequences on primer success rate and to evaluate the predictive Ct equation. For the validation samples, gDNA from 21 organisms (ATCC number listed in Table 2.1, organism used solely with these mixtures are in bold) were spiked

at varying concentrations, for an absolute abundance of approximately 10, 100, and 1,000 genomic copies per reaction well, into gDNA extracted from either river water, tertiary effluent, or activated sludge (0.99 ng per reaction well). In total, 36 validation samples were prepared. All samples were tested in triplicate.

2.3.8 Data analysis. For all analysis, data was filtered to differentiate true and false positive and negative signals. Amplification was considered positive if the threshold cycle was less than 26 for all three replicates and experimental melting temperature was consistent. The influence of primer 3' end GC content on specificity was analyzed using the primer (either forward or reverse) with the highest GC content. Primers were grouped based on the number of GC bases within last five bases on the 3' end. For developing the predictive Ct equation, only the primer sets displaying true positive amplification were considered .

The GC content of primer and target organisms influence on success rate of novel primers was examined based on average sum of successes (taken between 36 sample mixtures for validating the predictive Ct equation) for all assays targeting an organism. Assays were considered successful if it displayed a true positive or true negative signal, and an organism was deemed present if two or more assays (targeting one organism) displayed amplification with a threshold cycle less than 26 for all three replicates. The average GC content of all primers used for a targeted organism was used for this figure. The 3D plot was generated with a loess smoother and 1.0 sampling proportion. This smoothing was performed to identify characteristics of the population.

To compare with the predictive Ct equation, standard curves were generated using an average slope and intercept from all three replicates in all three backgrounds (control, gDNA spiked in river and tertiary gDNA) from sample mixtures designed to develop a predictive Ct equation. PCR efficiency was examined to determine influence of sample background on quantitative values, and was calculated from the slope of the standard curves with the following equation:

$$PCR\ Efficiency = (-1 + 10^{-1/slope})$$

2.4 Results and Discussion

2.4.1 Sensitivity and specificity. The influence of primer and template sequence characteristics on specificity and sensitivity was examined using the primer sets and sample mixtures designed to develop the predictive Ct equation. The percentage of assays displaying false positive/negative signals for primer sets tested individually and within a complex background did not change for target concentrations above a range of 1 to 10 copies (Figure 2.1, top panel). Within the 1 to 10 copy range (i.e., with 20 fg of target gDNA present per reaction well), approximately half of targeted assays displayed positive amplification. Depending on the organisms, this corresponds to 1.5 to 11 genomic copies per reaction well. A majority of the assays displaying true positive amplification at this concentration were for the following organisms: *L. monocytogenes* (6.3 copies), *S. enterica* (3.7 copies), *V. parahaemolyticus* (3.6 copies), *H. pylori* (11.1 copies), *C. jejuni* (11.6 copies), and *V. cholerae* (4.6 copies). Assays targeting all other organisms within this concentration had a detection limit higher than 10 copies per reaction well, implying those organisms were not detectable at this concentration. In addition, a high

majority of the targeted assays not displaying amplification at higher concentrations (in Figure 2.1, top panel) were for primer sets targeting the same organisms. For example, with 2,000 fg (100 to 1,000 copies) of target gDNA per reaction spiked into 66.6 pg per reaction, the following organisms displayed false negative signals in at least two of the three replicates: *P. aeruginosa* (three out of five primer sets), *S. aureus* (1 out of 4 primer sets), *G. intestinalis* (4 out of 5 primer sets). This accounts for all assays (8 of 69) showing false negative amplification at 100 to 1,000 copies per reaction well. A closer look indicated that the GC content was either very low (37% for *S. aureus*) or very high (67% for *P. aeruginosa*). *G. intestinalis* had the largest genome of all the organisms targeted (12 Mbp).

Previous PCR studies on organisms with high GC content have found similar influences and suggest it is due to stabilizing secondary structure influencing polymerase extension or primer-template annealing, causing false negative observations (16, 24, 37). Solutions to minimize the effect include using a higher denaturation and annealing temperature, or using additives such as glycerol, betaine, ormamide, or dimethyl sulfoxide. The false negative signal observed with *S. aureus* was likely due to the low GC content of the organism requiring the length of the primers to be increased to maintain suitable stability. The reverse primer for the *S. aureus* assay displaying a false negative signal was longer (32 bp) than any other primer tested in this experiment. Internal complements within longer primers may reduce annealing to target sequences (40).

The false negative signal observed with *G. intestinalis* may be due to lower relative abundance caused by having a larger genome, and a potential reduction in availability of target. The influence of genome size on amplification potential has

been described previously (11, 15), and may be due to a decrease in relative abundance of template over nontarget DNA. Garner proposed that in addition to decreased relative abundance, there is an increased chance of non-specific annealing of primers to nontarget regions, diminishing annealing of primers to the target strand (15). Optimizing the PCR cycle conditions or concentration of reagents may alleviate false negative signals (41); however, changing these parameters may influence the specificity and sensitivity of targeted assays that behaved well.

Since primers were designed to have the same theoretical melting temperature (Tm) the terminal 3' end of primers was also examined. The GC content, Tm, and binding energy within the terminal 7, 5, 3 and 2 bases on the terminal 3' end of primers was considered for all false positive signals. Correlations were used to determine that both the Tm of the last 7 bases and the GC content of the last 5 bases had the highest influence on false positive signal (considering targets alone and spiked into a background). To demonstrate, primer sets were grouped based on the GC content within the last 5 bases of the terminal 3' end (Figure 2.1, bottom panel). For this figure, primer sets were grouped based on the GC content within the terminal 3' end. Two out of 7 primer sets (28%) with 5 GC within the last 5 bp at the 3' end of the primer displayed false positive amplification. The percentage of false positive signals decreased with the amount of GC in primers 3' end. Influence of the 3' end of a primer on the specificity of amplification has been described previously (32). As a result, many primer design software's now analyze the 3' end of potential primers, while Primer3 emphasizes stability of 5 base segments of the terminal 3' end (38).

It should be noted that results obtained with the OpenArray™ nanoliter volume reactions are comparable to conventional microliter QPCR. Low volume PCR has been optimized in nanoliter volume reactions by adjusting surface chemistry (34), ramping rates and decreased annealing temperatures (47), and adjusting PCR Master mix composition with extra SYBR Green I, BSA and formamide added to the standard PCR mixture. Cross platform comparisons between PCR performed in the microplate format (10 and 20 ul volume reactions in the 7900HT) and with the OpenArray have shown high similarity in PCR efficiency, and detection limit (Ortenberg, E. and Roberts D., unpublished data; 7). The comparison also demonstrated a high correlation (between the two platforms) of specific gene regulation patterns between experimental (diseased heart) and control (normal adult heart) tissues. The PCR efficiency observed with the experiments described in this study (described below) also demonstrates success of PCR primers with the OpenArray environment.

2.4.2 Development of predictive Ct equation. An empirical equation was developed using the sequence specific results observed with the primers designed to develop the predictive Ct equation. Multiple parameters were considered for the predictive Ct equation including size of genome, GC content of genome, and primer binding energy. Since all primers were designed with the same theoretical melting temperature (Tm) (for simultaneous amplification of all primer sets on the OpenArray plate), the terminal 3' end of primers was also considered. This included the binding energy, position of G and C bases, Tm, and GC content for the terminal 7, 5, 3 and 2 bases on the primers 3' end. Correlations were used to determine which parameters had a greater influence on the Ct

value. Inclusion of parameters other than those chosen either had no affect or decreased the correlation between the predictive and experimental Ct. In addition, using the general linear model requires that all variables are independent, thus only the top independent parameters were used. For example, parameters such as binding energy are not entirely independent of the Tm of the terminal 7 bases, as a primer with a high binding energy may have a higher Tm on terminal 3' end. Five parameters were identified that influenced the Ct of each primer set. These were: i) genome size of the targeted organism, ii) target organism concentration, iii) GC content of the targeted organism, iv) amplicon length, and v) and theoretical melting temperature of the last 7 bases on the primer 3' end. The correlation of these variables on the threshold cycle was in the order listed (i.e. genome size of the targeted organism had the highest correlation, and Tm of last 7 bases on primer 3' end had the lowest).

A multiple parameter linear regression curve was used to place a weighted influence on each of these parameters and the following equation was developed:

$$Y = 10^{\left[\frac{(Ct-25.43)+(7.755\times GC)+(0.01275\times al)+(0.0084638\times Tmlast7)+(0.254293\times Bbase)}{-3.6841}\right]}$$

where;

GC = GC content (e.g. an organism with 30%; GC = 0.3)

al = amplicon length (bases long)

Ct = experimental threshold cycle

Y = starting copies

$Bbase$ = Size of genome (Mbp)

Tmlast7 = melting temperature calculated with 7 terminal bases on primer 3' end; lowest temperature (^{o}C) of the two primers with base stacking calculation (42).

A high correlation (R^2= 0.816) between threshold cycle predicted with the predictive Ct equation and the experimental threshold cycle was observed (Figure 2.2). Since the number of a gene copies will vary based on the targeted organisms and gene, the predictive Ct equation solely predicts starting copies per reaction. The accuracy of the equation was tested and compared with standard curves using validation sample mixtures with the primer set designed to validate the empirical Ct equation.

It is likely that constants in the predictive Ct equation will change based on the chosen system, reaction conditions, and reagents from different vendors. Therefore, some validation is necessary for setting up novel highly parallel assays. BioTrove suggests using the same reaction conditions for all SYBR based diagnostics on Openarray™ plates. Therefore, users of BioTrove for diagnostics of microbial communities can readily apply the predictive Ct equation. Development of a predictive Ct equation for a new assay will consist of validating with organisms from targeted genera (with varying target characteristics), and establishing a multiple parameter linear regression curve to place a weighted influence on parameters. To normalize for the stability of the reagent (fresh one vs a less fresh one), lot number between reagents, sample loading, and inhibition within various samples, inter assay normalization with an internal control can be used with the predictive Ct equation. Internal controls can consist of DNA strands absent in the diagnostic target sample (1, 9, 30), and/or a universal and stable unregulated endogenous standard (28, 29, 36).

2.4.3. Validation of an organism's sequence characteristics on primer success rate. Validation sample mixtures were tested with primers designed to validate the predictive Ct equation to determine the success rate of novel primers. The influence of primers and targeted organism GC content on the success of primers was observed (Figure 2.3). Results show that the sequence of organisms with extreme GC content (low and high) had lower success rates. A study by Housley et al., (18) observed similar results concerning influence of GC content and success of primers, describing a success rate of 56.9% for primers designed to target an amplified region with GC content greater than 50%, and 74.2% success for primers designed to target an amplified region with GC content less than 50%. Organisms targeted with high GC content tend to frequently give weak signals in amplification (due to secondary structure and template-template annealing), and primers with high GC content will amplify non-targeted regions (due to high stability). In addition, organism with a low GC content in the genome will have a higher rate of false negative signals. The number of assays required for confidently calling presence or absence of an organisms is dependant on the success rate of designed primers. Therefore, targets with extreme GC content will require more assays for presence and absence calls.

These results suggest that validation requirements can be reduced if multiple primer sets are assayed simultaneously, overcoming issues in specificity and sensitivity observed with organisms and genes with extreme GC content. Typically, amplification reactions can be optimized to overcome issues with false positive or negative signals. However, large-scale PCR diagnostics such as the platform

described here require broad conditions to satisfy suitable reactions for the majority of targeted assays. Using redundant probes to increase confidence in presence and absence calls has been described for other high-throughput platforms such as microarrays (31, 54). It should be noted that these observations may not be observed using methods that provide greater differentiation between true positive and false positive amplification; such as TaqMan QPCR, dissociation analysis, or ligation based. However, cost and time to complete diagnostics significantly increased with these approaches, especially for high-throughput diagnostics.

2.4.4. Validation of predictive Ct equation. The predictive Ct equation was examined and compared with standard curves (obtained from dilution experiments performed with primers designed to develop the predictive Ct equation), by predicting the starting copy number in the 36 validation samples tested with the primers designed to validate the predictive Ct equation. A distribution of predicted values shows a high distribution of predicted starting copies around 20, 100, and 1,000, which is close to the actual starting copy numbers of 10, 100, and 1,000 spiked into the validation samples (Figure 2.4). Comparing predicted starting copies using the Wilcoxon Signed-Rank Test showed that the standard curve tended to predict higher starting copies than using the predictive Ct equation. Using the predictive Ct equation made differentiating between assays spiked at 10 and 100 starting copies unclear.

The large distribution observed with both quantitative measures may be due to differences in reaction constitute lot number (between experiments performed to

establish and validate the predictive Ct equation), manually mixing of samples, and the chosen calibration parameters. The influence of reagent formulation and calibration parameters (e.g. replication, serial x-fold dilution) on precision of standard curves have been documented previously (12, 25). Inaccuracies in quantifying gDNA using UV absorbance may have also occurred. The correlation of predictive starting copies calculated using the standard curve and predictive Ct equation was R = 0.87. This suggests that the error between predicted and actual starting copies may have been due to inaccuracies in mixing validation samples and/or variations caused by reagents and not the predictive methods themselves.

Inhibitions within various environmental samples may have also influenced quantification. The distribution of PCR efficiencies shifted when targets were spiked into gDNA from different environmental waters (Figure 2.5, left panel), as observed by testing primers and samples designed to validate the predictive Ct equation. Tertiary effluent background shifted the distribution of PCR efficiency above one and the river water shifted the distribution below one. An analysis of variance showed significant inequality involving the mean PCR efficiencies between the control and 2 backgrounds. Approximately 83% of targeted assays with 1,000 to 10,000 starting copies had a Ct standard deviation equal to or less than 0.35, 80% with 100 to 1,000 starting copies, 33% with 10 to 100 starting copies, and 15% with 1 to 10 copies. A Ct standard deviation of 0.35 corresponds to a coefficient of variation of 25% for estimated starting copies (assuming a standard curve slope of -3.3), which has been described in previous reports as a cutoff for acceptable precision in QPCR diagnostics (4, 20, 44). A correlation matrix was used to examine if primer characteristics

influenced assays with high standard deviations (data not shown). Analysis showed that size of genome had the highest weighted influence on standard deviation of the Ct, followed by number of starting copies and GC content of the target genome. Other studies have observed an influence of GC on primer success. A study by Vanichanon et al. (48), observed reduced repeatability with primers sets with high GC content. Thus, primers with GC content closer to 50% will be ideal for maintaining high reproducibility (10, 21, 23). This fluctuation in efficiency could be due to the designed assays, enzyme instability, and sample dependant inhibitions (22).

Random micro-scale defects of the interior polymer surface coating of OpenArray™ plates, as suggested by Morrison et al. (34), may have also contributed to differences between predicted and actual starting copies. Intra-assay reproducibility was examined with three replicates performed on three separate OpenArray™ plates ran simultaneously with primers and samples designed to establish the predicted Ct equation. Results illustrate that 93% of the assays displayed a coefficient of variation equal to or less than 25% between OpenArray™ plates for 1,000 to 10,000 starting copies, 93% for 100 to 1,000 starting copies, 70% for 10 to 100 starting copies, and 33% for 1 to 10 starting copies (Figure 2.5, right panel). Results suggest greater deviation was observed between sample backgrounds than between plates, and the accumulation of both aspects may have influenced the distribution between predicted and actual starting copies.

Future diagnostics will include both an exogenous internal positive control and a universal assay targeting the 16S rRNA gene to allow normalization for the potential influence of sample inhibitions, plate variation, lot constitutes, and assays on

threshold cycle (17, 26, 52). Cumulatively, the large distributions observed with predicting starting copies with standard curves along with the cumbersome requirement of validation make using the predictive Ct equation an attractive potential alternative. One other alternative strategy to quantify starting copies without generation of standard curves has been described using competitive PCR with a fluorescence quenching (46). This alternative has the potential to quantify DNA within presence of inhibitions as tested with high concentrations of gDNA from pure cultures spiked into Humic acid, however, the accuracy and precision of quantification and influence of inhibitors was not described with samples containing less than 1,000 starting copies.

2.5 Conclusion

The influence of sequence characteristics of primers, amplicons, target genome size, amplification conditions, and the matrix in which the target is present on the Ct was explored. A predictive Ct equation was experimentally established by examining PCR-based assays targeting multiple VMGs, and may be an efficient alternative to generating standard curves. These results are valuable when developing high throughput and reliable screening tools for a large number of pathogens without extensive validation.

2.6 Acknowledgments

We are grateful for the help from Elen Ortenberg with a proof of principal study with BioTrove.

This work was supported by the National Institutes of Health (Grant R01 RR018625-01), Environmental Protection Agency (RD83162801-0 & RD83301001), and 21st Century Michigan Economic Development Corporation (GR-476 PO 085P3000517).

Table 2.1 Pathogens and VMGs targeted with OpenArray™ plates and organisms used for validation. The VMGs in bold were only tested in validation sample mixtures prepared to validate the predictive Ct equation.

Genus	*Species*	Marker Genes Targeted	Designed Primer sets	Primer Sets from Literature	Organisms tested (ATCC)
Aeromonas	*hydrophila*	alt, exeF, tapA, arcV, **ascT**	5	0	-
Bacillus	*cereus, anthracis*	**atxA, capA, lef, cerA, clo, hlyIII, mprF, pagA, plcR**	9	0	**10987**
Bordetella	*pertussis*	**cyaC, dnt, ptx, tcfA**	4	0	**BAA-589D**
Burkholderia	*mallei, pseudomallei*	pilA, pilD, SL1631OMP, BPSS1407	4	0	-
Campylobacter	*jejuni, coli*	hipO, racR, mapA, gyrA, cdtC, cdtB, cdtA, **ciaB, cfrA, tonB**	8	2	700819
Clostridium	*perfringens, difficile, botulinum*	cpe, plc, pfo, cpbs1, etx, **bontA, ha33, ha70, cdtA, cdtB**	10	0	12916
Coryhebacterium	*diphtheria*	**dtxr**	1	0	-
Coxiella	*burnetii*	**dotA, icmP, icmQ, mip**	4	0	-
Enterococcus	*faecalis, faecium*	cylA, ace, cylLS, esp, gelE	6	0	**19433**
Escherichia	*coli* (including *Shigella*)	uidA, stx1, stx2, **eae**, papG, **lucC, ehxA, sfaA, ipaH, mxiH**	11	3	BAA-460D
Francisella	*tularensis*	**acpA, clpB, mglA**			-
Haemophilus	*influenzae*	**lic2A, tbp1, tbp2**	3	0	**51907D**

Table 2.1 continued

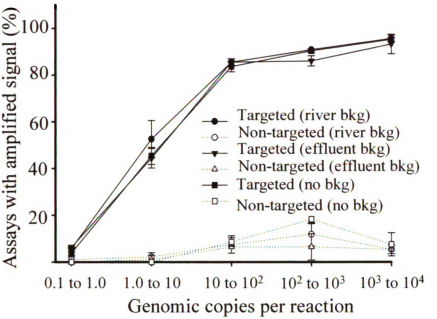
Helicobacter	*pylori*	cagA, cagE, ureA, virD4, **virB9, virB11, ureI, ureB, flaB**	10	2	700392
Klebsiella	*pneumoniae*	nuc, magA, kvgS, kci, kca	5	0	-
Legionella	*pneumophila*	lepB, icmQ, lepA, mip, **dotA, icmB, icmR, lspH, lssD**	9	1	33152
Leptospira	*interrogans*	hlyB, lipA, chpl, lip21, ompL1, **chpK, etpK**	6	1	**43642**
Listeria	*monocytogenes*	actA, hlyA, inlA, mpl, plcB, **lisA, plcA, iap, prfA**	8	1	15313
Mycobacterium	*MAC paratuberculosis, MTB tuberculosis, leprae*	erp, glnA1, mmpA, mmpB, IS900, **mmpC, plcA**	6	1	-
Mycoplasma	*genetalium*	**hww3, P1, P200, P30**	4	0	**33530D**
Proteus	*mirabilis*	**lpp, uca, zapA, zapD**	4	0	-
Pseudomonas	*aeruginosa*	exoS, iasA, pcrV, pilDXcpA, popD, **popB, las, etA, prfA, pilD**	11	-	10145
Salmonella		fimA, flicC, fljB, invA, **hiA, hiD, invE, prgH, sipB, sipC, spar, tyv**	15	1	13311, **19585**
Serratia	*marcescens*	**flhDC, hasD, lipA, safA**	4	0	-
Staphylococcus	*aureus*	alphahly, seA, seC, tsst1, **lukE, lukF, lytS, nuc, seA, tst**	10	1	700699

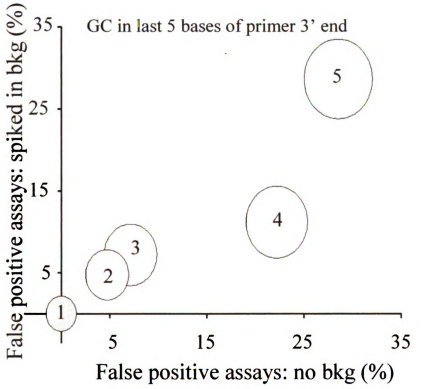
Table 2.1 continued

Strephocococcus	*pyogenes*	**ssa, mtsA, ska, speA, spec**	5		-
Vibrio	*cholerae, mimicus, parahaemolyticus, vulnificus*	ace, ompU, tdh, tdhS, tlh, toxR, ctxA, ctxB, mshA, tcpA, zot, trh1, **vpl, vvha, wp, tagA, trh1**	14	2	39315 43996
Yersinia	*enterocolitica, pestis, pseudotuberculosis*	crp, yopD, yscD, ystsA, bipA, **yadA, yspB, yscN, ail**	9	2	55075
Cryptosporidium	*parvum, hominis*	gp40, hsp70, COWP, cp23, cgd7	7	1	PRA-67
Giardia	*intestinalis, lamblia*	*B*-giardin, VSPH71, VSP4177, VSP4173, vsp, VSPH7	5	2	30888 A-1
Total: 30		200	(200 + 20) = 220		

2.7 FIGURES

Figure 2.1 Impact of gDNA from various environmental water samples and primer stability on specificity and sensitivity. Top panel: the percent of targeted and non-targeted primer assays displaying amplification at various dilutions of organisms spiked in gDNA from environmental samples. Errors bars represent standard deviation between replicates performed on 3 plates. Bottom panel: the sum of GC bases on the terminal 3' end (various size circles) versus the percentage of primer sets displaying false positive amplification when targets are spiked into background gDNA and not spiked into background gDNA.


Assays with amplified signal (%)
100
80
60
40
20
Targeted (river bkg)
Non-targeted (river bkg)
Targeted (effluent bkg)
Non-targeted (effluent bkg)
Targeted (no bkg)
Non-targeted (no bkg)
0.1 to 1.0
1.0 to 10
10 to 10²
10² to 10³
10³ to 10⁴
Genomic copies per reaction


False positive assays: spiked in bkg (%)
35
25
15
5
GC in last 5 bases of primer 3' end
5
4
3
2
1
5
15
25
35
False positive assays: no bkg (%)

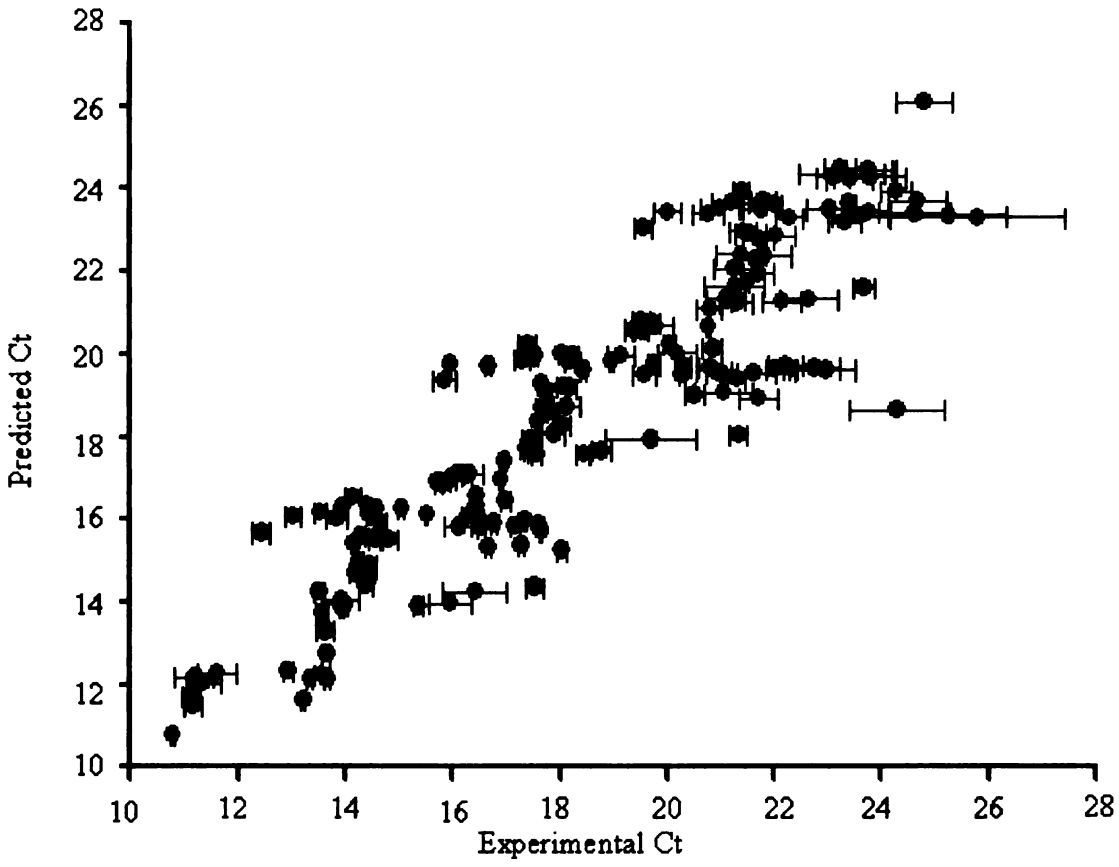


Figure 2.2 Experimental versus predicted Ct using predictive Ct equation, results obtained with primer designed to establish the predictive Ct equation. The equation is derived using amplicon length, starting genomic copies, number of base pairs in target organism's genome, GC content of target organism's genome, and theoretical melting temperature (Tm) of last 7 bases of primers 3' end. Errors bars represent standard deviation of experimental Ct between replicates on three plates.

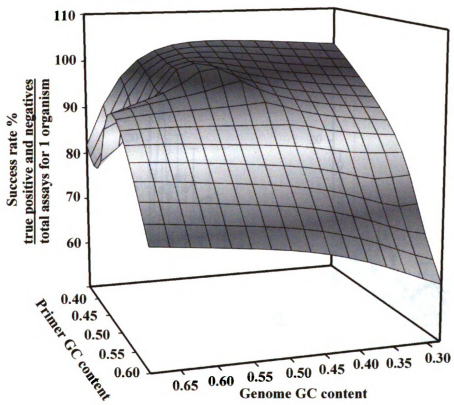


Figure 2.3 Success rate (%) of all assays targeting an organism. Success rate is defined as the sum of all true positive and negative assays divided by the sum of all assays targeting a single organism. The success rate was taken as an average of all 35 validation samples mixtures prepared to validate the predictive Ct equation.

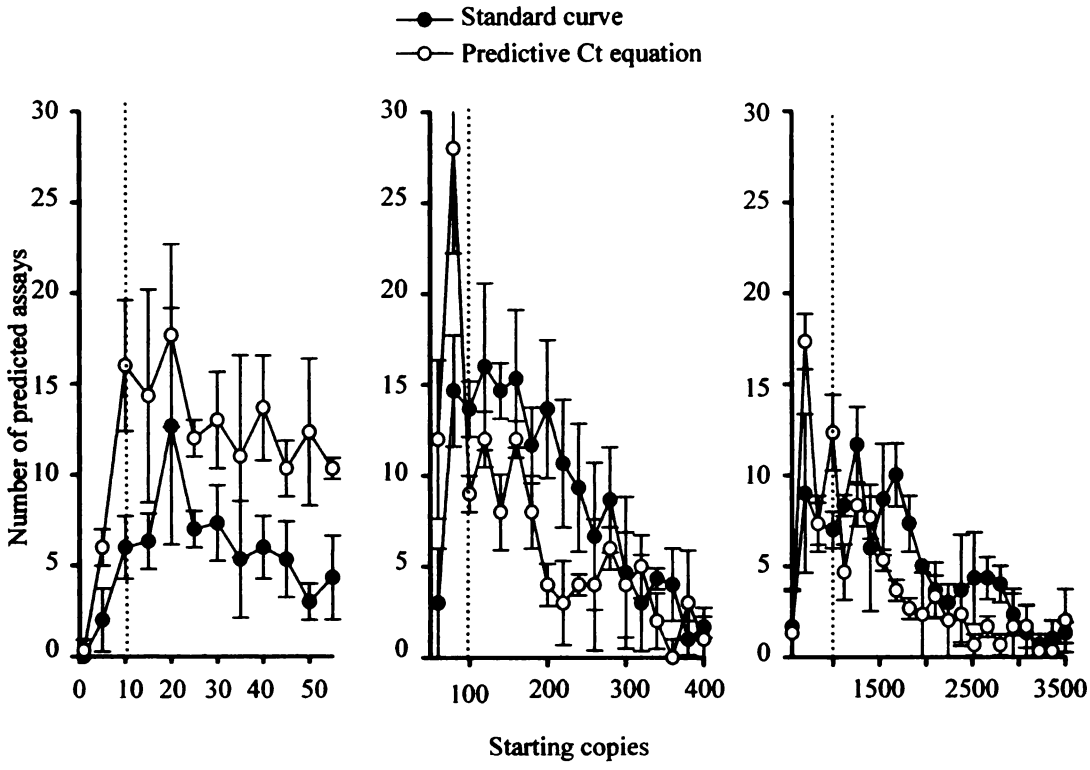


Figure 2.4 Distribution of predicted starting copies numbers using predictive Ct equation and standard curves for validation samples and primers designed to validated the predictive Ct equation. Note: x-axis of the 3 figures have different scales. Template of all organisms were spiked at either 10, 100, or 1,000 genomic copies per reaction well (indicated by dotted lines). Errors bars represent standard deviation between three replicates performed on the same OpenArray™ plates.

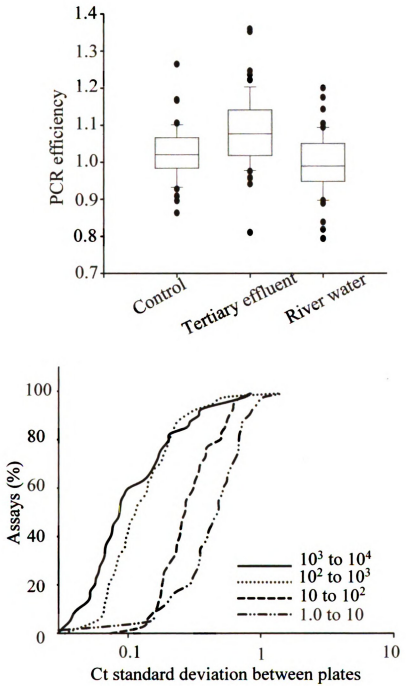


Figure 2.5 Factors potentially influencing quantitative inaccuracies as observed with primers and samples designed to establish the predictive Ct equation. Top panel: box plot showing distribution of PCR efficiency for target organisms spiked into gDNA from complex environmental waters and a control (no background). Bottom panel: the cumulative frequency distributions of the standard deviation of Ct determined between three separate OpenArray™ plates with various transcript concentrations.

2.8 REFERENCES

1. **Abdulmawjood, A., S. Roth, and M. Bulte.** 2002. Two methods for construction of internal amplification controls for the detection of *Escherichia coli* O157 by polymerase chain reaction. Mol. Cell. Probes **16:**335-339.

2. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25:**3389-402.

3. **Becker, K., A. W. Friedrich, G. Peters, and C. von Eiff.** 2004. Systematic survey on the prevalence of genes coding for staphylococcal enterotoxins SEIM, SEIO, and SEIN. Mol. Nutr. Food Res. **48:**488-495.

4. **Behets, J., P. Declerck, Y. Delaedt, L. Verelst, and F. Ollevier.** 2007. A duplex real-time PCR assay for the quantitative detection of *Naegleria fowleri* in water samples. Water Res. **41:**118-126.

5. **Blanco, J., M. Blanco, J. E. Blanco, A. Mora, E. A. Gonzalez, M. I. Bernardez, M. P. Alonso, A. Coira, A. Rodriguez, J. Rey, J. A. Alonso, and M. A. Usera.** 2003. Verotoxin-producing *Escherichia coli* in Spain: Prevalence, serotypes, and virulence genes of O157 : H7 and non-O157 VTEC in ruminants, raw beef products, and humans. Exp. Biol. Med. **228:**345-351.

6. **Bonants, P., M. Szemes, R. van Doom, and C. Schoen.** 2007. Quantitative multiplex detection of plant pathogens using PRI-lock probes and universal, ultra-high-throughput real-time PCR on OpenArrays (TM). BMC Genomics **97:**S12-S12.

7. **Brenan, C., and T. Morrison.** 2005. High throughput, nanoliter quantitative PCR. Drug Discovery Today: Technologies **2:**247-253.

8. **Chacon, M. R., M. J. Figueras, G. Castro-Escarpulli, L. Soler, and J. Guarro.** 2003. Distribution of virulence genes in clinical and environmental isolates of *Aeromonas* spp. Antonie Van Leeuwenhoek **84:**269-278.

9. **Cubero, J., J. van der Wolf, J. van Beckhoven, and M. M. Lopez.** 2002. An internal control for the diagnosis of crown gall by PCR. J Microbiol Methods. **51:**387-392.

10. **Dieffenbach, C. W., T. M. J. Lowe, and G. S. Dveksler.** 1993. General Concepts for PCR Primer Design. Pcr-Methods and Applications **3:**S30-S37.

11. **Farrelly, V., F. A. Rainey, and E. Stackebrandt.** 1995. Effect of genome size and rrn gene copy number on PCR amplification of 16s ribosomal-RNA genes from a mixture of gacterial species. Appl. Environ. Microbiol. **61:**2798-2801.

12. **Farriol, M., and X. Orta.** 2005. Influence of reagent formulation on mRNA quantification by RT-PCR using imported external standard curves. Acta Biochim. Pol. **52:**845-848.

13. **Feltman, H., G. Schulert, S. Khan, M. Jain, L. Peterson, and A. R. Hauser.** 2001. Prevalence of type III secretion genes in clinical and environmental isolates of *Pseudomonas aeruginosa*. Microbiology **147:**2659-2669.

14. **Fukushima, H., Y. Tsunomori, and R. Seki.** 2003. Duplex real-time SYBR green PCR assays for detection of 17 species of food- or waterborne pathogens in stools. J. Clin. Microbiol. **41:**5134-5146.

15. **Garner, T. W. J.** 2002. Genome size and microsatellites: the effect of nuclear size on amplification potential. Genome **45:**212-215.

16. **Grassi, M., E. Volpe, V. Colizzi, and F. Mariani.** 2006. An improved, real-time PCR assay for the detection of GC-rich and low abundance templates of *Mycobacterium tuberculosis*. J. Microbiol. Methods **64:**406-410.

17. **Hartman, L. J., S. R. Coyne, and D. A. Norwood.** 2005. Development of a novel internal positive control for Taqman based assays. Mol. Cell. Probes. **19:**51-59.

18. **Housley, D. J. E., Z. A. Zalewski, S. E. Beckett, and P. J. Venta.** 2006. Design factors that influence PCR amplification success of cross-species primers among 1147 mammalian primer pairs. BMC Genomics **7**.

19. **Hsu, C. F., T. Y. Tsai, and T. M. Pan.** 2005. Use of the duplex TaqMan PCR system for detection of Shiga-like toxin-producing *Escherichia coli* O157. J. Clin. Microbiol. **43:**2668-2673.

20. **Huang, H. Y., and T. M. Pan.** 2004. Detection of genetically modified maize MON810 and NK603 by multiplex and real-time polymerase chain reaction methods. J. Agric. Food Chem. **52:**3264-3268.

21. **Innis, M. A., and D. H. Gelfand.** 1990. Optimization of PCRs, p. 3-12. *In* M. A. Innis, D. H. Gelfand, J. J. Sninksy, and T. J. White (ed.), PCR protocols: A guide to methods and application. Academic Press, Inc., New York.

22. **Kains, P.** 2000. The PCR plateau phase-towards an understanding of its limitations. Biochim. Biophys. Acta **1494:**23-7.

23. **Kim, H. S., and O. Smithies.** 1988. Recombinant fragment assay for gene targetting based on the polymerase chain-reaction. Nucleic Acids Res. **16:**8887-8903.

24. **Kureishi, A., and L. E. Bryan.** 1992. Pre-boiling high GC content, mixed primers with 3' complementation allows the successful PCR amplification of *Pseudomonas aeruginosa* DNA. Nucleic Acids Res. **20:**1155-1155.

25. **Lai, K. K. Y., L. Cook, E. M. Krantz, L. Corey, and K. R. Jerome.** 2005. Calibration curves for real-time PCR. Clin. Chem. **51:**1132-1136.

26. **Layton, A., L. McKay, D. Williams, V. Garrett, R. Gentry, and G. Sayler.** 2006. Development of *Bacteroides* 16S rRNA gene TaqMan-based real-time PCR assays for estimation of total, human, and bovine fecal pollution in water. Appl. Environ. Microbiol. **72:**4214-4224.

27. **Lebuhn, M., M. Effenberger, A. Gronauer, P. A. Wilderer, and S. Wuertz.** 2003. Using quantitative real-time PCR to determine the hygienic status of cattle manure. Water Sci. Technol. **48:**97-103.

28. **Lee, C., S. Lee, S. G. Shin, and S. Hwang.** 2008. Real-time PCR determination of rRNA gene copy number: absolute and relative quantification assays with *Escherichia coli*. Appl. Microbiol. Biotechnol. **78:**371-376.

29. **Livak, K. J., and T. D. Schmittgen.** 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. Methods **25:**402-408.

30. **Malorny, B., J. Hoorfar, C. Bunge, and R. Helmuth.** 2003. Multicenter validation of the analytical accuracy of *Salmonella* PCR: towards an international standard. Appl. Environ. Microbiol. **69:**290-296.

31. **Miller, S. M., D. M. Tourlousse, R. D. Stedtfeld, S. W. Baushke, A. B. Herzog, L. M. Wick, J.-M. Rouillard, E. Gulari, J. M. Tiedje, and S. A. Hashsham.** 2007. An in situ synthesized virulence and marker genes (VMG) biochip for detection of waterborne pathogens. Appl. Environ. Microbiol. **In Review**.

32. **Miura, F., C. Uematsu, Y. Sakaki, and T. Ito.** 2005. A novel strategy to design highly specific PCR primers based on the stability and uniqueness of 3 '-end subsequences. Bioinformatics **21:**4363-4370.

33. **Molenkamp, R., A. van der Ham, J. Schinkel, and M. Beld.** 2007. Simultaneous detection of five different DNA targets by real-time Taqman PCR using the Roche LightCycler480: Application in viral molecular diagnostics. J. Virol. Methods **141:**205-211.

34. **Morrison, T., J. Hurley, J. Garcia, K. Yoder, A. Katz, D. Roberts, J. Cho, T. Kanigan, S. E. Ilyin, D. Horowitz, J. M. Dixon, and C. J. H. Brenan.** 2006. Nanoliter high throughput quantitative PCR. Nucleic Acids Res. **34.**

35. **Pannucci, J., H. Cai, P. E. Pardington, E. Williams, R. T. Okinaka, C. R. Kuske, and R. B. Cary.** 2004. Virulence signatures: microarray-based approaches to discovery and analysis. Biosens Bioelectron **20:**706-718.

36. **Pfaffl, M. W.** 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. **29:**e45.

37. **Raychaudhuri, A., and P. A. Tipton.** 2004. Protocol for amplification of GC-rich sequences from *Pseudomonas aeruginosa.* Biotechniques **37:**752, 754, 756.

38. **Rozen, S., and H. Skaletsky.** 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. **132:**365-86.

39. **Rozynek, E., K. Dzierzanowska-Fangrat, P. Jozwiak, J. Popowski, D. Korsak, and D. Dzierzanowska.** 2005. Prevalence of potential virulence markers in Polish *Campylobacter jejuni* and *Campylobacter coli* isolates obtained from hospitalized children and from chicken carcasses. J. Med. Microbiol. **54:**615-619.

40. **Rychlik, W., and R. E. Rhoads.** 1989. A computer-program for choosing optimal oligonucleotides for filter hybridization, sequencing and invitro amplification of DNA. Nucleic Acids Res. **17:**8543-8551.

41. **Sahdev, S., S. Saini, P. Tiwari, S. Saxena, and K. S. Saini.** 2007. Amplification of GC-rich genes by following a combination strategy of primer design, enhancers and modified PCR cycle conditions. Mol. Cell. Probes **21:**303-307.

42. **SantaLucia, J.** 1998. A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. Proc. Nat. Acad. Sci. U S A **95:**1460-1465.

43. **Sergeev, N., M. Distler, S. Courtney, S. F. Al-Khaldi, D. Volokhov, V. Chizhikov, and A. Rasooly.** 2004. Multipathogen oligonucleotide microarray for environmental and biodefense applications. Biosens Bioelectron **20:**684-698.

44. **Shindo, Y., H. Kuribara, T. Matsuoka, S. Futo, C. Sawada, J. Shono, H. Akiyama, Y. Goda, M. Toyoda, A. Hino, T. Asano, M. Hiramoto, A. Iwaya, S. I. Jeong, N. Kajiyama, H. Kato, H. Katsumoto, Y. M. Kim, H. S. Kwak, M. Ogawa, Y. Onozuka, K. Takubo, H. Yamakawa, F. Yamazaki, A. Yoshida, and T. Yoshimura.** 2002. Validation of real-time PCR analyses for line-specific quantitation of genetically modified maize and soybean using new reference molecules. J. AOAC Int. **85:**1119-1126.

45. **Skanseng, B., M. Kaldhusdal, and K. Rudi.** 2006. Comparison of chicken gut colonisation by the pathogens *Campylobacter jejuni* and *Clostridium perfringens* by real-time quantitative PCR. Mol. Cell. Probes **20:**269-279.

46. **Tani, H., T. Kanagawa, S. Kurata, T. Teramura, K. NakamuraO, S. Tsuneda, and N. Noda.** 2007. Quantitative method for specific nucleic acid sequences using competitive polymerase chain reaction with an alternately binding probe. Anal. Chem. **79:**974-979.

47. **van Doorn, R., M. Szemes, P. Bonants, G. A. Kowalchuk, J. F. Salles, E. Ortenberg, and C. Schoen.** 2007. Quantitative multiplex detection of plant pathogens using a novel ligation probe-based system coupled with universal, high-throughput real-time PCR on OpenArraysTM. BMC Genomics **8:**276.

48. **Vanichanon, A., N. K. Blake, J. M. Martin, and L. E. Talbert.** 2000. Properties of sequence-tagged-site primer sets influencing repeatability. Genome **43:**47-52.

49. **von Eiff, C., A. W. Friedrich, G. Peters, and K. Becker.** 2004. Prevalence of genes encoding for members of the staphylococcal leukotoxin family among clinical isolates of *Staphylococcus aureus*. Diagn. Microbiol. Infect. Dis. **49:**157-162.

50. **Vora, G. J., C. E. Meador, M. M. Bird, C. A. Bopp, J. D. Andreadis, and D. A. Stenger.** 2005. Microarray-based detection of genetic heterogeneity, antimicrobial resistance, and the viable but nonculturable state in human pathogenic *Vibrio* spp. Proc. Natl. Acad. Sci. U S A **102:**19109-19114.

51. **Wang, L. X., Y. Li, and A. Mustapha.** 2007. Rapid and simultaneous quantitation of *Escherichia coli* O157 : H7, *Salmonella*, and *Shigella* in ground beef by multiplex real-time PCR and immunomagnetic separation. J. Food Prot. **70:**1366-1372.

52. **Welti, M., K. Jaton, M. Altwegg, R. Sahli, A. Wenger, and J. Bille.** 2003. Development of a multiplex real-time quantitative PCR assay to detect *Chlamydia pneumoniae*, *Legionella pneumophila* and *Mycoplasma pneumoniae* in respiratory tract secretions. Diagn. Microbiol. Infect. Dis. **45:**85-95.

53. **Wick, L. M., W. H. Qi, D. W. Lacher, and T. S. Whittam.** 2005. Evolution of genomic content in the stepwise emergence of *Escherichia coli* O157 : H7. J. Bacteriol. **187:**1783-1791.

54. **Wilson, W. J., C. L. Strout, T. Z. DeSantis, J. L. Stilwell, A. V. Carrano, and G. L. Andersen.** 2002. Sequence-specific identification of 18 pathogenic microorganisms using microarray technology. Mol. Cell. Probes **16:**119-127.

CHAPTER THREE

Stedtfeld, R.D., Wick, L.M., Baushke, S.W., Tourlousse, D.M., Herzog, A.B., Xia, Y., Rouillard, J.M., Klappenbach, J.A., Cole, J.R., Gulari, E., Tiedje, J.M., Hashsham, S.A. 2007. Influence of Dangling End and Surface Proximal Tail of Targets on Probe-Target Duplex Formation in 16S rRNA Gene-Based Diagnostic Arrays. Appl Environ Microbiol. 73:380-9.

INFLUENCE OF DANGLING END AND SURFACE PROXIMAL TAIL OF TARGETS ON PROBE-TARGET DUPLEX FORMATION IN 16S rRNA GENE-BASED DIAGNOSTIC ARRAYS

3.1 ABSTRACT

Dangling ends and surface proximal tails of gene targets influence the probe-target duplex formation and affect the signal intensity of probes on diagnostic microarrays. This phenomenon was evaluated using an oligonucleotide microarray containing 18-mer probes corresponding to the 16S rRNA gene of 10 waterborne pathogens and a number of synthetic and polymerase chain reaction amplified gene targets. Signal intensities for Klenow/random primer labeled 16S rRNA gene targets vs. 45-mer synthetic targets were dissimilar for nearly 73% of the probes tested. Klenow/random primer labeled targets resulted in 3.7 times higher interaction with a complex mixture of 16S rRNA genes (used as background) compared to the interaction of 45-mer targets with the same mixture. An 7 bases long dangling end sequence with perfect homology to another single stranded background DNA sequence was sufficient to produce a cross hybridization signal that was as strong as the signal obtained by the probe-target duplex itself. Gibbs free energy between the target and a well defined background was found to be a better indicator of hybridization signal intensity than the sequence or length of the dangling end alone. The dangling end was found to be significantly more prone to target-background interaction (Gibbs free energy of -7.6 kcal/mol) than the surface proximal tail (Gibbs free energy of -64.5 kcal/mol). This study underlines the need for careful target preparation and evaluation of signal intensities for

diagnostic arrays using 16S rRNA and other gene targets due to the potential for target interaction with a complex background.

3.2 Introduction

There are many factors that influence the hybridization signal intensity between long strands of DNA (referred to as targets) and oligonucleotide probes in microarray-based diagnostics and gene expression studies (14). Some of these factors include characteristics of fluorescent molecule used to label the target, labeling efficiency, length of probe, target secondary structure, and hybridization conditions (3, 4, 6, 9, 15, 22, 28, 33-35, 37). However, factors such as interaction of **dangling end** and **surface proximal tail** of targets with background DNA has not yet been systematically characterized. In this study, the term dangling end refers to the sequence of target that extends beyond the distal end of the probe and surface proximal tail refers to the sequence of target that extends beyond the anchored end of the probe.

Target preparation and resulting length, sequence, double versus single stranded composition, position of fluorescent label, and secondary structure of the dangling end and surface proximal tail of targets influence the stability of duplex formation and resulting signal intensity (25, 31, 44). The targets may also interact with other non-target DNA sequences present in the background, previously referred to as "hitchhiking" (23). This is especially true for 16S rRNA gene-based diagnostic microarrays where dangling ends and surface proximal tails will generally represent conserved regions and may have greater potential to interact with the background DNA or other targets. Because background will be sample specific, comparative evaluation of two samples on the same microarray using target mixtures that interact with each other may pose problems. Low

reproducibility and false positive/negative signals with 16S rRNA gene based arrays is partly accredited to bias caused by characteristics of target molecules (8, 17), and target-background interaction may be one of the main reasons for this phenomenon. Poor statistical relationships between experimental and predicted signal intensities (from Gibbs free energy calculations) suggest that thermodynamic parameters between target and probe duplexes are not fully understood (26). Thus, an evaluation of the influence of dangling end and surface proximal tail on signal intensity and target-background interaction is needed in order to interpret the signal intensities on 16S rRNA gene based diagnostic and microbial community analysis arrays. The interaction among probe (P), target (T), and background (B) can be represented by:

$$Pn \Leftrightarrow Tn \Leftrightarrow Bn$$

Where, n represents the length of the corresponding DNA sequence. A given background DNA sequence may also interact with the probe sequence ($Pn \Leftrightarrow Bn$). But it was not considered significant in this study because signal due to background DNA when hybridized alone was absent for the targeted probes.

The objective of this study was to systematically examine the impact of dangling end and surface proximal tail of targets on hybridization signal intensity and evaluate target-background interaction by: i) comparing products of unknown length generated by random primer labeling with targets of known length, and ii) examining interactions between well defined targets and backgrounds. This study underlines the need for careful target preparation and evaluation of signal intensities for diagnostic arrays using 16S rRNA and other gene targets due to the potential for target interaction with a complex background.

3.3 Materials and Methods

Influence of background on dangling end and surface proximal tail was evaluated using an oligonucleotide microarray containing 18-mer probes corresponding to the 16S rRNA gene of 10 waterborne pathogens and a number of synthetic and polymerase chain reaction amplified gene targets. Using a set of 45-mer synthetic targets and Klenow/random primer labeled targets generated by PCR amplification (assumed to be approximately 200 to 1,500 bp long), the signal intensities were compared for a total of 95 probes targeting the 16S rRNA genes of 10 waterborne pathogens. Both types of targets were then spiked individually into a complex mixture of 16S RNA genes to study the influence of the type of background DNA that will be present as a result of target amplification using universal primers and random primer labeling. Using a Cy3 end-labeled 59-mer sequence as a model background DNA sequence designed to complement the dangling end of 5 different targets of variable length (and the associated Gibbs free energy), it was demonstrated that 8 to 12 bases of contiguous homology with the dangling end is sufficient to cause cross hybridization. A second 59-mer sequence served as a model background designed to complement the surface proximal tail of 11 different targets of various sequence similarity and Gibbs free energy. Sequences used for the model backgrounds were from the dangling ends of non-targeted 16S rRNA genes for which probes were also present on the array. Synthetic sequences used for the well defined model background complemented the dangling end of non-target probes.

3.3.1 Preparation of target DNA. Four different types of targets were prepared: i) mixture of Klenow/random primer labeled (with Cy5) targets prepared from 16S rRNA genes of 10 pathogens (Tk-m), ii) mixture of Cy3 end-labeled 45-mer synthetic targets corresponding to 95 probes for the 10 pathogens (T45-m), iii) one individual Cy3 end-labeled synthetic target of length 102 bases (Td) to examine dangling end length, 4 additional Cy3 end-labeled synthetic target of length 47, 60, 80, and 106 bases (Tn, where n is the length of the target) to study the effect of increasing the length of dangling end, and iv) 11 individual Cy3 end-labeled synthetic targets of lengths 89 to 102 bases (Ts) were used to study the effect of sequence similarity in the surface proximal tail of targets. The sequence of dangling end and surface proximal tail for all the synthetic targets (T45-m, Td, Tn, and Ts) matched the 16S rRNA gene sequence of the corresponding pathogens. The approach to prepare these targets is summarized below.

Klenow/random primer labeled target mixture of 16S rRNA genes of pathogens (Tk-m). Genomic DNA from 10 bacterial pathogens (Table 1) was used as the source of 16S rRNA genes. *E. coli, L. pneumophila, P. aeruginosa, S. enterica, S. typhimurium, Y. enterocolitica* type strains were obtained from the American Type Culture Collection (ATCC; Manassas, VA) and grown as per the protocol provided. For *H. pylori* and *C. perfringens,* only genomic DNA was obtained from the ATCC. *E. feacalis* and *C. jejuni* were kindly provided by Drs. Joan B. Rose and Vincent Young, respectively; both at Michigan State University. DNA from pure cultures and the environmental sample was extracted using Promega Wizard DNA Extraction Kit (Promega, Madison, WI).

The 16S rRNA gene was amplified from the respective genomic DNA using 27F (5'-AGAGTTTGATCMTGGCTCAG-3') and 1525R (5'-AAGGAGGTGWTCCARCC-

3') primer pairs (18) and Platinum Taq DNA Polymerase (Invitrogen; Carlsbad, CA). Thirty cycles of the following temperature program was used for amplification: denaturation at 94 °C for 30 s, annealing at 52 °C for 45 s, and elongation at 72 °C for 90 s. All PCR amplicons were cleaned using a Qiagen PCR clean up kit (Qiagen Inc., Valencia, CA).

The amplified 16S rRNA gene (250 ng) from each pathogen was individually labeled with Cy5 using the Bioprime DNA Labeling Kit (Invitrogen, San Diego, CA). Briefly, the protocol included a 90-min incubation of the amplicon with Klenow polymerase and 5:1 amino-allyl-dUTP:dTTP (Ambion, Austin, TX) followed by Cy5 labeling. All amino-allyl-dUTP-labeled products were cleaned using a Qiagen PCR clean up kit with modified phosphate wash buffer (5 mM K_2HPO_4, pH 8.0, 80% EtOH) and phosphate elution buffer (4 mM K_2HPO_4, pH 8.5). Cyanine dye was attached by incubating 3 to 5 μg of amino-allyl-dUTP-labeled DNA for 1 h in 50:50 mixture of 0.1 M sodium carbonate buffer (pH 9.3) and N-hydroxysuccinimide ester Cy dye (prepared in fresh dimethyl-sulfoxide). DNA product from dye coupling was cleaned using a Qiagen PCR clean up kit. A mixture of Cy5 labeled 16S rRNA gene was then prepared by mixing 10 pmol (157 to 357 ng DNA) of labeled product from each of the 10 pathogens. Klenow/random primer labeled targets were estimated to be between 200 to 1,500 bp (as assessed by gel electrophoresis). All Klenow-labeled targets are expected to have multiple Cy5 labels. Specific activity of Cy dye was calculated by dividing measured pmol of nucleotides with measured pmol of Cy3/Cy5 dye for each target.

Mixture of 45-mer synthetic targets (T45-m). A mixture of synthetic target was designed and synthesized to complement a set of 95 probes (18-mer) in the middle region

of the 45-mer targets (Figure 3.1a). Each 45-mer synthetic target had a 14 nucleotides surface proximal tail at the 5' terminus and 13 nucleotides dangling end at the 3' terminus. The sequences of both the overhanging ends matched the 16S rRNA gene sequence of the corresponding pathogen. The single stranded mixture of sequences was synthesized at the University of Michigan using an in situ oligonucleotide synthesis technology (11, 12), end labeled with Cy3 at the 5' terminus, and harvested from the solid substrate.

Synthetic target with dangling ends of variable lengths (Td). A 102-mer synthetic target was designed to hybridize to 9 different probes on the array (Td). After hybridization, Td was expected to have a dangling end of 8 to 16 bases at the 3' terminus and a surface proximal tail of 68 to 76 bases at the 5' terminus, depending upon the probe to which it was hybridized (Figure 3.2a). Td was synthesized and mono end-labeled with Cy3 at the 5' terminus by IDT (Coralville, IA). A 59-mer defined background DNA (Bd, described below) was designed to complement the dangling end sequence from position 86 to 98 with 4 to 12 base stretches of continuous homology. The sequence similarity between the surface proximal tail of Td and the Bd was less than 25%.

The dangling end interaction due to length of overhang was studied further using 4 synthetic targets of varying length T47, T60, T80, and T106 (Tn, where n refers to the length of the synthetic target, e.g., T60 is 60 bases long). Each Tn hybridized to 10 probes that had been designed from the same 16S rRNA sequence from the position 242 to 272 (starting at the 3' end). Each Tn was synthesized and mono end-labeled with Cy3 at the 5' terminus by IDT (Coralville, IA), and resulted in a 3' dangling end of 21, 34, 54, and 80 bases for the probe designed at start position 242, and 8, 21, 41, and 67 bases for

the probe designed at start position 254. At the 5' terminus, all 4 synthetic targets had a surface proximal tail of 8 to 21 bases depending on the probe it was hybridized to. Bd complemented the dangling end of Tn by 0, 13, 33, and 59 bases.

Synthetic targets (Ts) with 10 to 100% sequence similarity at the surface proximal tail. The effect of sequence similarity of the surface proximal tail of target to a defined background DNA sequence was studied using 11 different targets (Ts) of similar length (89 to 102 bases). Each Ts was synthesized and mono end-labeled with Cy3 at the 5' terminus by IDT (Coralville, IA). All the targets had 18 bases complementary to 52 probes on the microarray (Figure 3.3a); a dangling end of 1 to 16 bases and surface proximal tail of 67 to 76 bases depending on the probe and target. Ts were designed to have 10 to 100% sequence similarity to a single background sequence (Bs, described below).

3.3.2 Preparation of background DNA. Two types of background DNA were used in this study: i) a mixture of 16S rRNA genes obtained by PCR from an environmental sample referred to as complex background DNA (B16s), and ii) two defined background DNA of known sequences (both were 59-mers, referred to as Bd and Bs) with varying potential of interaction with dangling end of Td and surface proximal tail of Ts. For Td and Ts, synthetic sequences complemented the dangling end of non-target probes. These are described below.

Complex background DNA (B16s). A mixture of 16S rRNA genes amplified from genomic DNA of activated sludge from the wastewater treatment plant in East Lansing, MI) served as the source of complex background DNA. The genomic DNA was obtained

using the protocol described above except that a purification step was also performed using Wizard DNA clean-up kit (Promega, Madison, WI). A similar PCR protocol was used as in preparation of Tk-m, but only 25 cycles were used and annealing temperature was set to 50 °C for 45 s. The extent of sequence similarity of the complex background DNA to target sequences was unknown but expected to be substantial (because of conserved regions of 16S rRNA genes) and a function of the length of background and target DNA.

Synthetic background DNA of defined length. Two 59-mer synthetic defined background sequences (Bd and Bs) were designed to complement the dangling end of Td and surface proximal tails of Ts (11 different T_s sequences were used). The extent of sequence similarity for defined background DNA was carefully controlled to either result in continuous homology of variable length at the dangling end of Td or varying sequence similarity at the surface proximal tail of Ts. For Td, the sequence of Bd had 4 to 12 bases of continuous homology from position 86 to 98 considering the 5' end as 0 (Table 1 and Figure 3.2a). The Gibbs free energy between Td and Bd varied between -3.2 to -14.9 kcal/mol for 4 to 12 base stretches of continuous sequence homology at the dangling end. Gibbs free energy between Bd and P18 (complementing the $Td_)$ was between -4.3 and -4.6 kcal/mole.

When compared with Tn (T47, T60, T80, and T106), the sequence of Bd had 0, 13, 33, and 59 bases matching within the dangling end of Tn, and a Gibbs free energy of -2.3 to -74.9 kcal/mol. The Gibbs free energy between Bd and P18 (complementing the Tn) was between -3.2 and -5.3 kcal/mol.

The second defined background DNA sequence, Bs, was designed to complement the surface proximal tail of 11 different Ts (as shown in Table 1 and depicted in Figure 3.3a) with various degrees of sequence similarity. The surface proximal tail of Ts had an 8 to 16 base overhang (duplex dependent) with 0% sequence homology to Bs followed by a stretch of 59 bases with variable regions of contiguous perfect matches and sequence similarity with Bs. This resulted in a sequence similarity of 0 to 90% on the whole surface proximal tail and 10-100% sequence similarity considering only the 59 bases on the 5' end of Ts. The dangling end of Ts had 1 to 16 bases depending on the probe it was hybridized. For all interactions between Bs and Ts, the Gibbs free energy was always less negative than -5.7 kcal/mol. The Gibbs free energy between Bs and P18 (complementing the Ts) was between -3.8 and -8.1 kcal/mol.

3.3.3 Probe design and microarray synthesis. Probes were designed using a Perl script developed in house that screened for all 18-mers of 16S rRNA genes that had at least two mismatches to every other sequence in the RDP-II (http://rdp.cme.msu.edu/). Screened probes for a given pathogen were ranked based on melting temperature (Tm) and GC content; following the parameters incorporated into OligoArray 2.0 probe design software (32). The final array included a set 146 probes (P18) targeting 21 different 16S rRNA gene sequences and mismatch probes; identical to the perfect match sequence except for 1 and 2 incorrect bases in the middle of the oligomer to measure the degree of cross hybridization. Microarrays containing the above set of probes were synthesized by Xeotron Corporation, Houston, TX (now part of Invitrogen, Carlsbad, CA) using a

proprietary in situ synthesis technology developed by the University of Michigan (11, 12).

3.3.4 Experimental approach. *A). Comparison of 45-mer synthetic target mixture vs. Klenow/random primer labeled target mixture.* Two-sample comparative hybridization, similar to that used in gene expression studies (5), was performed to evaluate the differences and similarities between synthetic targets and Klenow/random primer labeled targets. A similar experiment was conducted to study the impact of complex background 16S rRNA gene mixture on target signal intensity. A total of three comparative hybridizations were performed in triplicate (shown by double-arrows in Figure 3.1a). These were: i) T45-m vs. Tk-m - to compare the hybridization behavior of two types of target mixtures, ii) T45-m vs. B16s - to evaluate the impact of background DNA on 45-mer synthetic target mixture, and iii) Tk-m vs. B16s - to evaluate the impact of background DNA on Klenow/random primer labeled target mixture. The complex background mixture B16s, Bd, and Bs was also hybridized alone to ensure that target signals for the probes to be tested did not cross hybridize with various backgrounds. For hybridization (i), 20 pmol of T45-m (1,069 ng DNA) was mixed with 100 pmol (2,481 ng DNA) of Tk-m. For hybridization (ii), 20 pmol (1,069 ng of DNA) of Cy3-end-labeled T45-m was mixed with 200 pmol (6,550 ng DNA) of Cy5-labeled B16s. For hybridization (iii), 100 pmol of Cy5-labeled Tk-m was mixed with 200 pmol (6,550 ng DNA) of Cy3-labeled B16s. The last mixture was obtained by combining 10 pmol of labeled 16S rRNA gene from each of the 10 pathogens.

B. Evaluation of length and sequence of dangling end of synthetic targets. The effect of dangling end length was studied by varying the length of continuous sequence homology between Td and Bd from 4 to 12 bases as shown in Figure 3.2a. The dangling end had the sequence 5'-(86 bases)-CGACTTGCATGTGTTG-3', wherein bold type nucleotides complement the sequence of Bd. In order to generate the increasing dangling end overhang of Td from 8 to 16 bases, nine separate probes with increasing offsets relative to the target. The sequence of background DNA (Bd) included the complementary sequence for the 12 bases of dangling end to result in target-background interaction, which was expected to increase with an increase in dangling end length. The hybridization included 67 pmol of Cy3 labeled Td (3,557 ng) and 37 pmol of Cy5 labeled Bd (1,540 ng).

The dangling end interaction due to Gibbs free energy was examined further using four synthetic targets with 100% sequence similarity and varying length (T47, T60, T80, and T106) along with a 59-mer Bd that was designed to complement the dangling end of Tn by 0, 13, 33, and 59 bp (having Gibb's free energy of -8.8, -14.8, -43.8, -74.9 kcal/mol, respectively). The Bd for this examination was the same as before; however the Tn targeted 10 different probes than Td. The hybridization included 52 to 84 pmol of Cy3 labeled Tn (1522 to 4277 ng) and 37 pmol of Cy5 labeled Bd (1,540 ng).

C. Evaluation of sequence similarity in surface proximal tail of a synthetic target. The dependence of sequence similarity at the surface proximal tail of a synthetic target with background DNA was studied by synthesizing 11 different synthetic targets of similar length Ts (Figure 3.3a) and a new synthetic background DNA sequence (Bs). The Gibbs free energy of Ts sequences varied between -9.2 and -76.8 kcal/mol with sequence

similarity varying between 10 to 100% (Figure 3.3a). Each Ts was end-labeled with Cy3 at the 5' end and complemented 3 to 8 probes. The hybridization included a total of 160 pmol of Cy3 labeled Ts (7,895 ng) and 30 pmol of Cy5 labeled Bs (1,115 ng).

3.3.5 Hybridization and scanning. Target and background mixtures were prepared in 100 µl of hybridization solution containing 35% formamide, 0.4% Triton X-100, and 6X SSPE. SSPE buffers were made from a stock of 18X SSPE, which is 2.7 M NaCl, 180 mM Na_2PO_4, 18 mM Na_2EDTA (adjusted to 6.6 pH with HCl). The hybridization solution was heated at 95 °C for 3 min, cooled on ice for 1 min, and passed through 0.22 µm filter. All hybridizations were carried out in triplicate for 16 to 18 hr at 20 °C using a M-2 microfluidic station (Xeotron Corporation). A flow rate of 500 µl per min was used for recirculation of hybridization solution through the microfluidic array during hybridization (42). After overnight hybridization at 20 °C, the microarray was washed using wash buffer 2 (6X SSPE, 0.2% Triton X-100), wash buffer 4 (1X SSPE, 0.2% Triton X-100), and wash buffer 2 with no Triton X-100 in series for 2.2 minutes each (500 µl per min, 20 °C). A non-equilibrium thermal dissociation approach adapted to the Xeotron platform was used in all hybridization experiments (8, 20, 39, 42). The protocol was based on earlier studies utilizing a dissociation curve approach for diagnostic arrays (8, 20, 39, 42). The microarrays were washed with a high stringency wash buffer (10 mM Na_2HPO_4, 5 mM EDTA, pH 6.6; flow rate 500 µl per min) for 2.2 min at increasing temperatures from 20 to 70 °C at 2 °C intervals. Experimental conditions have been examined previously for optimal specificity (42). Signal intensities were quantified after each wash using a GenePix 4000B non-confocal laser scanner

(Axon Instruments, Inc, Foster City, CA) at a photomultiplier tube voltage setting of 650 V for Cy5 (635 nm laser) and 600 V for Cy3 (532 nm laser). For complex background DNA (B16s), this setting was 750 V (for Cy5) and 500 V (for Cy3). With this setting, photo-bleaching caused an average decrease in signal intensity of 0.85% for Cy5-labeled targets and 0.54% for Cy3-labeled targets between each of the 20 scans at 20 °C.

3.3.6 Data analysis. Data was analyzed using an XL script that imported raw GenePix dissociation curve data between 20 and 70 °C into Excel and generated a sigmoid curve for each microarray feature. Signal to noise ratio (S/N) at a given temperature was calculated from three replicate hybridization experiments. Noise was defined as the average signal of twenty empty spots (containing linker chemistry and no probes). The average signal of 20 randomly selected non-target probes (80.3 $\pm$ 10.6 arbitrary units with 635 nm, 148.3 $\pm$ 16.7 a.u. with 532 nm laser) was similar to average signal of 20 empty spots (83.6 $\pm$ 6.4 a.u. with 635 nm, 155.9 $\pm$ 10.6 a.u. with 532 nm laser) when microarray was hybridized with complex background. Compared to no background, the average noise increased slightly when a complex background DNA (B16s) was hybridized (1.19-fold for Cy3 and 1.54-fold for Cy5). For all experiments, an average S/N greater than 3.0 was considered positive to balance the maximum percentage of true positive calls of targeted probes and true negative calls of non-targeted probes. Figure 3.1c is presented as S/N for the area under the dissociation curve to include variation in both dissociation and signal intensity. For statistical comparison of probe signal intensity when hybridized with Klenow/random primer fragmented and 45-mer targets, a two-tailed inference about differences in population mean for independent

samples was performed with 95% confidence interval. The free energy of hybridization was computed using the two-state hybridization server (www.bioinfo.rpi.edu/) developed by Dimitrov and Zuker (7) with 43 °C instead of the actual 20 °C. This is because the hybridization buffer contained 35% formamide, which is expected to destabilize duplexes equivalent to increasing the temperature by 21 to 25 °C (2, 39).

3.4 Results and Discussion

3.4.1 Klenow/random primer labeled target mixture vs. end-labeled 45-mer synthetic target mixture. The hybridization behavior of Klenow/random primer labeled target mixture compared to the 45-mer synthesized target mixture, plotted as S/N of T45-m and Tk-m is shown in Figure 3.1b. Tested probes in Figure 3.1b were sorted based on the difference in S/N between T45-m and Tk-m. Considering a S/N greater than 3 as positive signal, 87 out of 95 tested probes yielded positive signal for T45-m while 91 probes gave positive signal for Tk-m. Approximately 5.4% of non-targeted probes gave false positive signal with T45-m, and 10.9% gave false positive signal with Tk-m. A higher amount of non-specific hybridization with Tk-m compared to single stranded DNA targets has been reported previously (10). Most of the probes having similar S/N with both the targets generally had a S/N of less than 10 (probes 21 to 46). The S/N ratio for 73 out of the 95 probes differed significantly between the two target mixtures (95% confident interval considering mean difference). There was no systematic pattern when signal intensities of all probes for a single 16S rRNA gene were compared for the two targets. Often the difference in S/N was as large as 33-fold (e.g., for probe numbers 80 to 95).

While a difference in signal intensity with various target mixtures is expected, higher signal with single stranded targets compared to Klenow/random primer labeled targets has not been reported previously. More specifically, probes 1 through 30 had much higher S/N with T45-m than with Tk-m, while the opposite was observed for probes 60 to 95. For 32 out of 95 probes, the S/N was significantly higher with T45-m than with Tk-m ($P \leq 0.05$). Dissimilar results were observed in a study by Franke-Whittle et al. (10), in which S/N was higher for a majority of duplexes targeted with Klenow fragmented targets compared to single stranded targets. Higher S/N with Klenow fragmented targets is expected because longer dangling ends in Tk-m provide opportunity for attachment of many more Cy5 dye molecules compared to only one Cy3 dye in T45-m. The average specific activity for Tk-m was 72 ± 20 nucleotides per Cy5 dye (calculated using arithmetic average of specific activity for each of 10 targets that were labeled separately). This implies that the average number of Cy5 dye was between 2 and 30 (assuming a target length between 200 and 1,500).

It is evident from Figure 3.1b that normalization of S/N for the higher number of dyes in Tk-m or dye bias would not result in a profile that overlaps completely with the S/N profile of T45-m. This implies that factors in addition to differences in label abundance and dye bias are responsible for the differences in S/N. These factors include: i) secondary structure of longer dangling ends of Tk-m making certain regions inaccessible for hybridization with probes (4, 19, 28), ii) Tk-m and T45-m displacement from probes due to complementing strands of double stranded Tk-m (25) known as a zipper effect (30), iii) bias in Klenow-labeling due to higher GC content (38), iv) probe sequence variability and linear probes influences on thermodynamic stability (31), v)

physical quenching of Cy dyes (6), and vi) conserved regions on overhang of hybridized Tk-m interacting with complementing strands from 1 or more of the 10 organisms also targeted in the mixture (23) creating chains of targets. Interactions occurring in solution between 16S rRNA gene targets from various organisms may influence the spot signal intensity by competing with the targets bound to the probe.

Further experiments were performed to examine and quantify the extent of target-background interaction on the two dissimilar targets by spiking both Tk-m and T45-m in background DNA. When the two target mixtures were spiked (separately) into a complex mixture of non-target 16S rRNA genes (B16s), the S/N of Tk-m changed significantly compared to the S/N of T45-m (Figure 3.1c). The X-axis in Figure 3.1c represents the factor by which the S/N ratio of the target mixture (T45-m or Tk-m) is influenced by the background, B16s, and Y-axis is frequency of probes with a given influence factor. For a given probe, influence factor was computed as S/N in absence of B16s divided by the S/N in the presence of B16s. For targets not impacted by the background DNA, this influence factor should be close to unity (implying no effect) and the spread of the bell-shaped curve should be minimal. For T45-m, the influence factor was close to 1 while for Tk-m, it was approximately 4. Similarly, the spread of T45-m influence curve was much smaller compared to that obtained for Tk-m. A simple comparison of the areas under each influence factor curve indicated that the interaction of Tk-m with the background was 3.7 times higher compared to T45-m (obtained by comparing the areas under the curve of Tk-m and T45-m which was 69.5 and 19 arbitrary units, respectively). A Wilcoxon signed-rank test was used to statistically conclude that distribution of influence factor was greater for Tk-m than T45-m with 95% confidence interval.

Therefore, background significantly influences signal intensity of Tk-m and not T45-m. Since the same probes were targeted with both target mixtures, the influence of background on signal intensity is dependent on the characteristics of the target molecules, and independent on the spatial region on the array. The B16s was also hybridized alone to ensure that targeted probes did not produce any signal with the background 16S rRNA gene mixture.

The influence of background on signal intensity of double stranded Tk-m may be the result of target interactions influenced by complementary and non-target strands in background DNA. In one study, greater hybridization efficiency was observed with single-stranded targets compared to double-stranded targets (13). This is due to competition between complementary strands and probes with double-stranded DNA. Related studies have reported increased competition between the complementing strand and probe as the length of the dangling end increased, while increasing the length of the surface proximal tail did not influence hybridization (25). Interactions between the conserved regions of target dangling ends and sequences from a second target have also been observed (23). These interactions were minimized by using single stranded RNA instead of double stranded DNA. If this observation can be extended to DNA sequences, the use of asymmetric PCR for target preparation may be advantageous in minimizing target-background interaction.

Target-background interaction on the dangling ends of the target may also influence signal intensity by affecting the secondary structure. As studied by Chandler et al. (4), signal intensity was substantially increased as target secondary structure was relieved using a second probe in solution that annealed to target dangling end.

Interaction between the target dangling end and a background DNA sequence may have a similar effect. Peplies et al. (24) observed similar results for a majority of tested probes. However, the signal of some probes decreased with a second probe in solution, suggesting that the opening of a selected binding site may lead to a reorganization of one or more secondary structures in other target regions.

There are additional motives for using shorter targets with diagnostic arrays. Shorter targets reduce intermolecular structures that occur more frequently with single stranded targets compared to stiffer double stranded products (36). Lane et al. (15) used amplicons of various lengths and concluded that shorter fragments reduced false negatives due to high intermolecular secondary structure in longer targets. Fluorophore position on a target DNA sequence can also influence signal intensity (44). Signal intensity decreases as the proximity of the fluorescent molecule from a probe target duplex increases. Thus, target preparation strategies should include amplifying shorter targets with specific asymmetric PCR, or fragment longer amplicons to produce a range of desired length.

Wilson et al. (43) suggested using fragments under 100 bp compared to full-length amplicons to obtain a stronger hybridization signal. A number of fragmentation strategies have been described for obtaining a size range between 35 and 200 bases with pre- and post-labeled targets and asymmetric PCR amplified products (1, 21, 27, 41). Subsequent studies comparing cost, efficiency, and influence on signal intensity for some of these methods are currently underway in our laboratory. Target-target and target-background interactions are expected to be less for targets other than 16S rRNA genes, especially if the targets have less conserved regions.

3.4.2 Synthetic targets of increasing dangling end length. The effect of dangling end length was evaluated using a target with an increasing stretch of continuous homology. A Cy3 end-labeled synthetic 102-mer target (Td) was designed to hybridize to 9 different probes on the array. After hybridization, Td was expected to have a dangling end of 8 to 16 bases (Figure 3.2a). A 59-mer synthetic background DNA sequence (Bd) end-labeled with Cy5 was designed to complement the dangling end sequence from position 86 to 98 with 4 to 12 base stretches of continuous homology. The signal to noise ratio of Bd and Td at 26 °C is shown for each of the 9 separate probes (Figure 3.2b). In order to generate the increasing dangling end overhang on the Td, probes were offset by one nucleotide relative to the target. Considering a S/N greater than 3 as positive signal, Bd hybridized to the dangling end of Td only when the dangling end stretch of continuous homology was 7 bases or longer. Signal from Bd was not observed when hybridized alone (i.e., without Td), and mismatch probes did not display signal for Bd or Td; directly implicating that Bd did not cross hybridize with the probe. As evident by the lack of signal for probes with target dangling end of 6 bases or less, interaction of surface proximal end of Td with Bd was insignificant. The S/N of Bd for 11 and 12 base overhangs was higher than the S/N given by Td. This may be due to dye bias between Cy5 and Cy3 and higher specific activity of Cy5-labeled Bd. The 59-mer Bd may have resulted in a greater percentage of dye labeled product compared to the 102-mer Td, as purity of oligonucleotide synthesis decreases with increasing length.

The dangling end interaction due to length of overhang was reexamined using four synthetic targets of varying lengths (T47, T60, T80, and T106 together represented

by Tn) along with the same 59-mer Bd that was designed to complement the dangling end of Tn by 0, 13, 33, and 59 bp. Dangling end interaction occurred between T60, T80, T106, and Bd but not between T47 and Bd (data not shown). The zero sequence homology between T47 and Bd was the obvious reason for this lack of interaction.

Gibbs free energy calculations incorporate influences such as length, sequence similarity, secondary structures and hairpin loops caused by near perfect match sequences with insertions or deletions (16). It was used to examine interaction between the Td, Tn and Bd. For Td, an overhang with 4 sequential perfect match bases had a Gibbs free energy of -3.2 kcal/mol, while an overhang with 12 sequential perfect match bases had a Gibbs free energy of -14.9 kcal/mol. The S/N of Bd became greater than 3 when the stretch of continuous bases on the dangling end was 7 or higher and Gibbs free energy was -7.6 kcal/mol or more negative. With the experimental conditions used in this study, target-background interaction on dangling end can occur with Gibbs free energy that is more negative than -7.6 kcal/mol. However, due to phenomenon such as difference in intermolecular structure and competitive influence of complementary strands (28), the interaction observed with this free energy may not be extensible to double-stranded DNA targets. For calculating the Gibbs free energy between Bd and Td, only the 3' end of Td was considered to circumvent the influence of probes on dangling end interaction. Neglecting the presence of the probe, the Gibbs free energy of the duplex formed between Td and Bd was -15.4 kcal/mol. Because the 5' end of Bd preferentially hybridizes to the 3' dangling end of Td, the interaction occurs solely on the dangling end rather than the surface proximal end of Td or the probe.

The S/N of Bd increased as Gibbs free energy became more negative and length of continuous homology on Td dangling end increased. Gibbs free energy between Bd and T47, T60, T80, and T106 was -2.3, -14.2, -43.8, -74.9 kcal/mol, respectively. Previous studies have also reported successful hybridization to 9-mer probes, with a Gibbs free energy of -6 to -10 kcal/mol producing weak signals and -8 to -12 kcal/mol producing strong signals (29). It should be noted that binding free energy has been shown to be a function of the surface material, surface charge density, length of linker molecule, and other experimental conditions (40).

3.4.3 Synthetic targets with surface proximal tail sequence complementary to background DNA. The effect of sequence similarity of the surface proximal tail of target to sequence of defined background DNA was studied using 11 different targets (Ts) of similar length (89 to 102 bases). All the targets had 18 bases complementary to 52 probes on the microarray (Figure 3.3a), a dangling end of 1 to 16 bases and surface proximal tail of 67 to 76 bases depending on the probe and target. Targets were designed to have 10 to 100% sequence similarity with Bs. Probes were not offset (as with dangling end experiments); however 11 targets were used resulting in a Gibbs free energy between -9.2 to -76.8 kcal/mol for Ts⇔ Bs duplex (Figure 3.3b). The Gibbs free energy for P18⇔ Ts varied between -17.5 and -24.1 kcal/mol. Continuous stretches of 38 and 45 perfect matches at the surface proximal tail of Ts (with free energy of -64.5 to -65.9 kcal/mole) displayed target-background interaction (Figure 3.3c), while continuous stretches of 19 or less perfectly matching bases (with free energy of -42.1 kcal/mol or

more positive) did not result in signals attributable to Bs. Targeted probes did not display signal when Bs was hybridized alone (i.e. without Ts).

These results suggest that target-background interaction at the surface proximal tail is different compared to a similar interaction at the dangling end. While a Gibbs free energy of -7.6 kcal/mol (stretch of 7 continuous perfect matches) was sufficient for dangling end interaction, a Gibbs free energy more positive than -64.5 kcal/mol (including a stretch of 19 continuous perfect matches) on surface proximal tail did not display interaction. Further examination of the Ts with a Gibbs free energy of -36.8 kcal/mol when hybridized with Bs (19 base stretch of continuous homology) showed no secondary structure on surface proximal tail. Thus external influences that are not incorporated into Gibbs free energy calculations may be causing greater destabilization on surface proximal end. A study examining the influence of target overhang length on hybridization efficiency showed that the dangling end of a target interacted with its complementary strand more than the surface proximal tail (25). It was suggested that this is because the dangling end is exposed to the liquid phase providing better access to complementary strands in solution. On the other hand, the surface proximal tail is closer to the surface, thus limiting availability to other strands in solution and reducing association with complementary strands. It is obvious that the difference in target-background interaction on surface proximal tail and dangling end requires further investigation.

3.5 Conclusions

The length and sequence of surface proximal tail and dangling end of targets and resulting target-background interaction influences signal intensity and decreases specificity of oligonucleotide probe-target hybridizations. The dangling end was found to be significantly more prone to target-background interaction compared to surface proximal tail. Gibbs free energy between the target and background was found to be a better indicator of hybridization signal intensity than the sequence or length of the dangling end alone. This study underlines the need for careful target preparation (fragmentation and asymmetric PCR) and evaluation of signal intensities for diagnostic arrays using 16S rRNA and other gene targets due to the potential for target interaction with a complex background.

3.6 Acknowledgments

We are grateful for the help from Jimmy Johnson and Weihong Qi with scripts for sequence analysis, and Alok Dhawan for critically reading this manuscript.

This work was supported by the National Institutes of Health (Grant R01 RR018625-01) and Michigan Economic Development Corporation (GR-476 PO 085P3000517).

Table 3.1 Identity of pathogens, number of probes and range of Gibbs free energy for the set of probes.

Pathogen Name	**ATCC OR STRAIN NO.**	**No. of Probes**	**Gibbs Free Energy range of P18 (kcal/mol)**
Burkholderia pseudomallei	1026b	2	-19.6
Campylobacter jejuni	700819	1	-21.1
Clostridium perfringens	13124	5	-21.3 to -26.7
Enterococcus faecalis	19433; B/66	15	-17.3 to -24.1
Escherichia coli	11775; M25588	15	-20.8 to -24.9
Helicobacter pylori	43504; 49396T	16	-17.9 to -20.6
Legionella pneumophila	33215	12	-18.1 to -21
	10145; TERI		
Pseudomonas aeruginosa	2014;	33	-18.6 to -23.
	DSM50071		
Salmonella enterica	13311; 76; 12-1	18	-20 to -23.4
Salmonella typhimurium	13314	5	-18.2 to -19.2
Salmonella enteritidis	SE22	3	-22.1 to -22.8
Vibrio parahaemolyticus	13; 17	11	-17.4 to -23.4
Yersinia enterocolitica	9610	10	-15.9 to -20.2

Table 3.2 Characteristics of various DNA sequences used to study the effect of dangling end and surface proximal tail on signal intensity.

	Free energy of duplex (kcal/mol)
A 45-mer mixture and probe sequence (in bold) for one of 95 targets TTACGACTT CACC**CCAATC** **GCTGACCCT** **ACT**GCGGCCG CTGCCTC 18-mer probe AGTAGGGTC AGCGATTGG	-15.9 to -26.7 between T45-m and P18
B 102-mer target (Td) and probe sequence 1-9 (in bold) GGACGTTAT CCCCCACTAC CAGGCAGATT CCTAGGCATT ACTCACCCGT CCGCCGCTGA ATCCAAGG**CA** **AGCTCCCTTC** **ATAAGCTCGA** **CTTGC**ATGTG TTG 59-mer background (Bd) AGAGTTTGA TCCTGGCTCA GAGTGAACGC TGGCGGCGTG CCTAATACAT GCAAGTCGAA	-3.2 to -14.9 between dangling end of Td and Bd -4.3 to -5.4 between Bd and P18 -18.6 to -21.4 between Td and P18

Table 3.2 Continued

C	
Ts and probe sequence (in bold) for one of 11 targets	
CCCGTTTCC GGACGTATCC CCCACTACCA GGCAGATTC CTAGGCATTA	-9.2 to -76.8 between surface proximal tail of Ts and Bs
CTCACGTCCG CCGCTGAANC CAG**GAGCAAC**	
GTCCGTCATC CGCTCGACT	-3.6 to -8.1 between Bs and P18
59-mer background (Bs)	
TCAGCGGCG GACGGGTGAG TAATGCCTAG	-17.4 to -24.9 between Ts and P18
GAATCTGCCT GGTAGTGGGG GATAACGTCC	

3.7 Figures

Figure 3.1 Klenow/random primer labeled target mixture vs. end-labeled 45-mer synthetic target mixture. a) Experimental strategy. b) Signal to noise ratio of synthetic targets T45-m (▲) and Klenow/random primer labeled targets Tk-m (■) at 26 °C. The average signal for three replicate experiments is plotted and the error bars are standard deviation. c) Ratio of signal from hybridizations with target alone over targets spiked in background DNA for T45-m (▲) and Tk-m (■). The X-axis represents the factor by which the S/N ratio of the target mixture (T45-m or Tk-m) is influenced by the background, B16s, and Y-axis is frequency of probes with a given influence factor. For a given probe, influence factor was computed as S/N ratio in absence of B16s divided by the S/N ratio in the presence of B16s. Signal in Figure 3.1c is S/N of area under the dissociation curve to include variation in both initial signal intensity and dissociation.

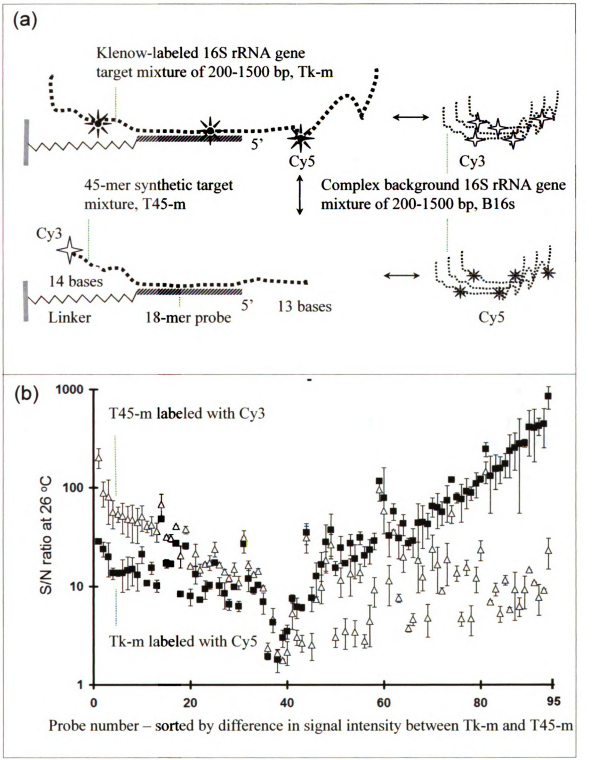

(a)
Klenow-labeled 16S rRNA gene
target mixture of 200-1500 bp, Tk-m
5'
Cy5
Cy3
45-mer synthetic target
mixture, T45-m
Complex background 16S rRNA gene
mixture of 200-1500 bp, B16s
Cy3
14 bases
Linker
18-mer probe
5'
13 bases
Cy5
(b)
T45-m labeled with Cy3
Tk-m labeled with Cy5
S/N ratio at 26 °C
1000
100
10
1
0
20
40
60
80
95
Probe number – sorted by difference in signal intensity between Tk-m and T45-m

Figure 3.1 continued

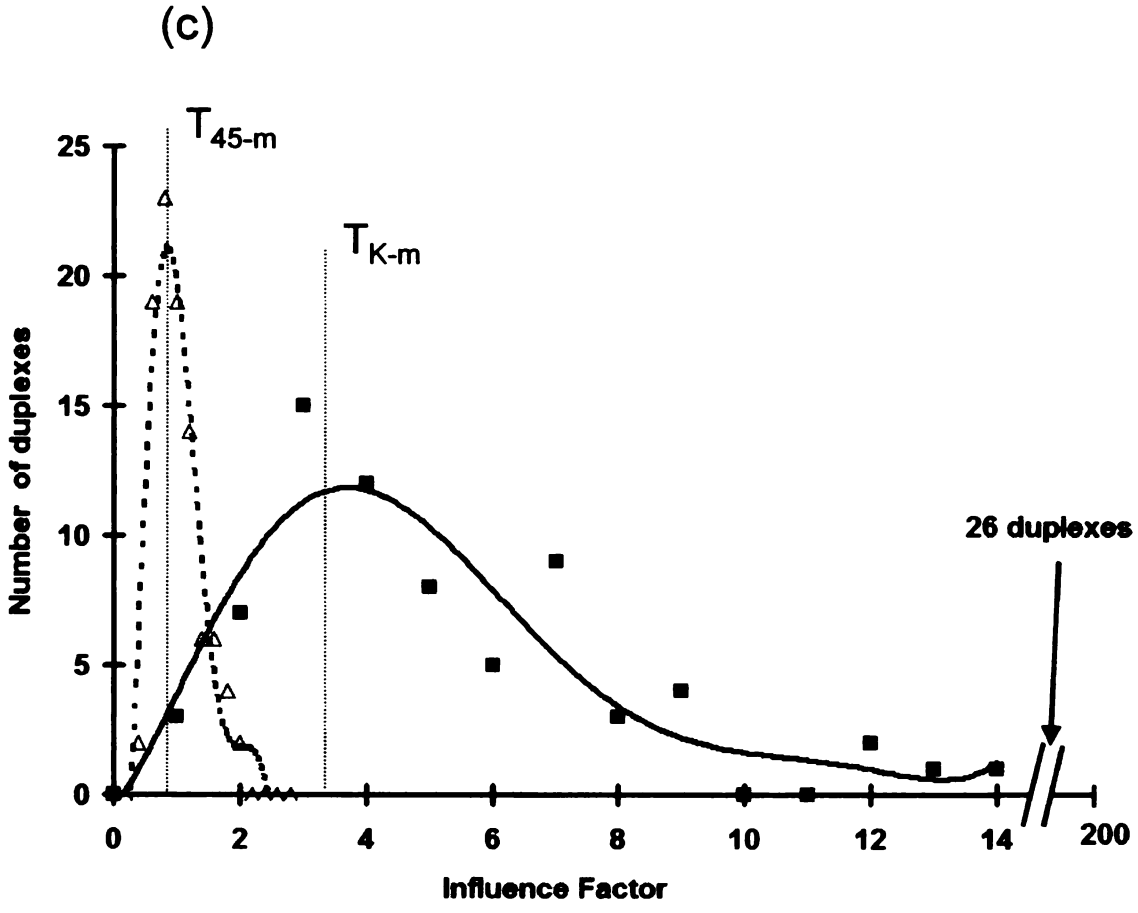
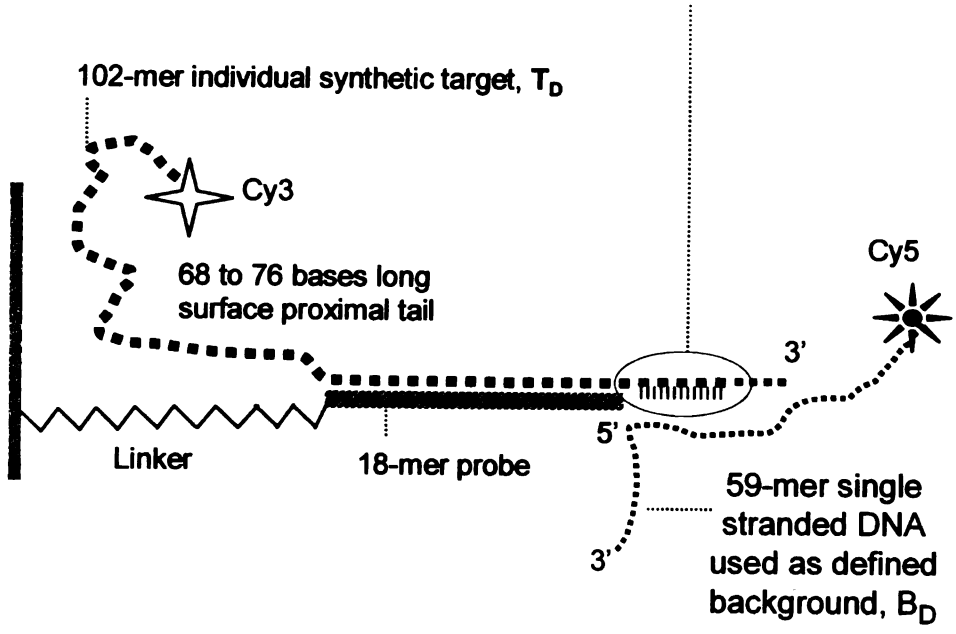


Figure 3.2 Synthetic targets of increasing dangling end length. a) Experimental strategy. b) The signal to noise ratio of Bd and Td at 26 °C (Y-axis) is shown for each of 9 separate probes offset by one nucleotide relative to the target, in order to generate the increasing dangling end overhang of Td (X-axis). Clear bars represent S/N of Td and dark bars are S/N of Bd. Signal of Bd is observed when target dangling end overhang has a stretch of continuous homology greater than 7 bases. The average signal for three replicate experiments is plotted and the error bars are standard deviation.

4 to 12 bases long dangling end (5'-(86 bases)-**CGACTTGCATGT**GTTG-3') of the 102-mer synthetic target to interact with B_D

(a)



Gibbs Free Energy with B_D kcal/mol	T_D overhang
-3.2	ATGTGTTG-3'
-4.3	CATGTGTTG-3'
-6.9	GCATGTGTTG-3'
-7.6	TGCATGTGTTG-3'
-8.3	TTGCATGTGTTG-3'
-9.8	CTTGCATGTGTTG-3'
-11.1	ACTTGCATGTGTTG-3'
-12.0	GACTTGCATGTGTTG-3'
-14.9	CGACTTGCATGTGTTG-3'

Figure 3.2 continued

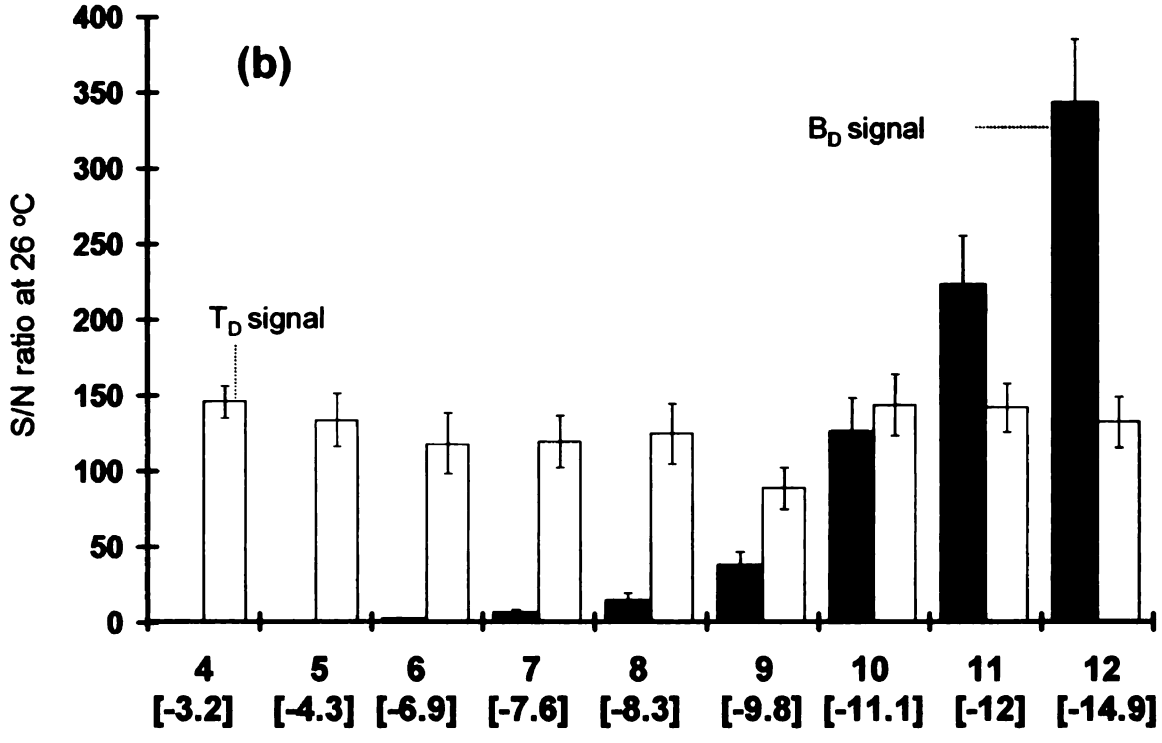
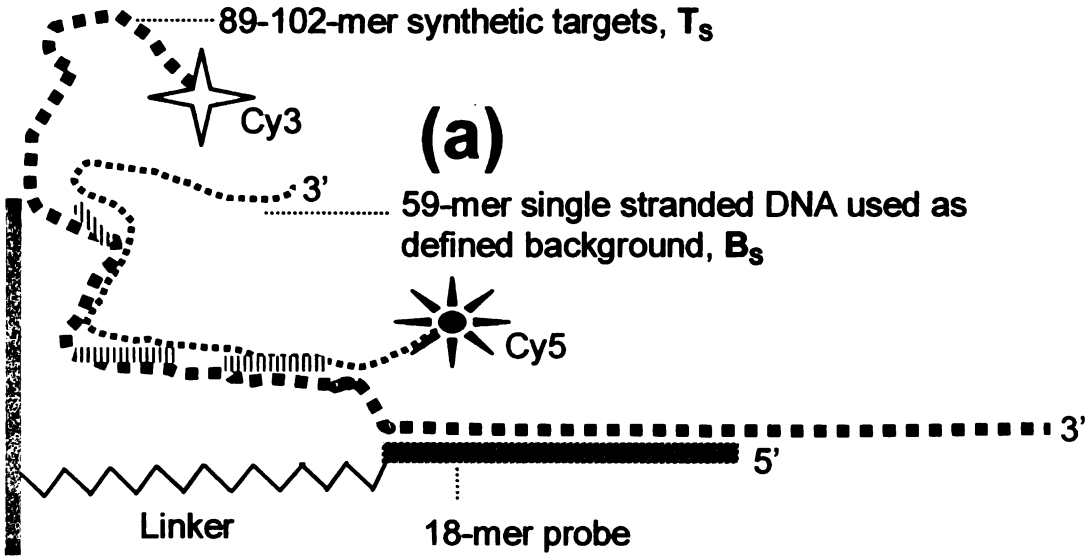


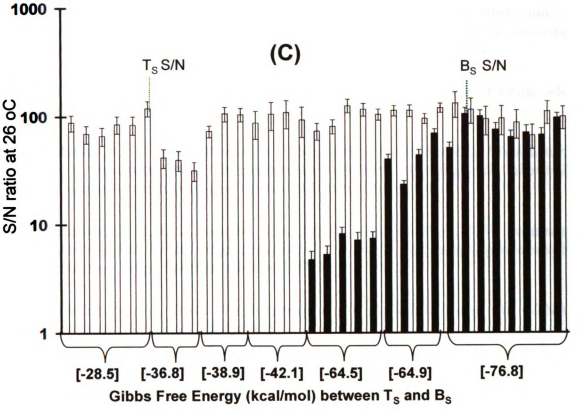
Figure 3.3 Synthetic targets with surface proximal tail sequence complementary to background DNA. a) Experimental strategy. b) Mismatches, insertions, and deletions influencing Gibbs free energy between the Bs and 11 different Ts, resulting in a Gibbs free energy for Ts⇔Bs duplex formation that ranged between -9.2 to -76.8 kcal/mol. c) S/N at 26 °C of Ts and Bs (Y-axis) versus Gibbs free energy (kcal/mole) between Ts and Bs (X-axis). Clear bars represent S/N of Ts and dark bars are S/N of Bs. Only 7 out of 11 results are displayed because the Bs did not interact with other Ts. The average signal for three replicate experiments is plotted and the error bars are standard deviation.



T_S Number	Gibbs free energy with B_S kcal/mol
1	- 76.8
2	- 64.9
3	- 64.5
4	- 42.1
5	- 38.9
6	- 36.8
7	- 28.5
8	- 13.5
9	- 12.5
10	- 11.1
11	- 9.2

Figure 3.3 continued

Figure 3.3 continued



3.8 REFERENCES

1. **Bekal, S., R. Brousseau, L. Masson, G. Prefontaine, J. Fairbrother, and J. Harel.** 2003. Rapid identification of *Escherichia coli* pathotypes by virulence gene detection with DNA microarrays J. Clin. Microbiol. **41:**2113–2125.

2. **Blake, R. D., and S. G. Delcourt.** 1996. Thermodynamic effects of formamide on DNA stability. Nucleic Acids Res. **24:**2095-2103.

3. **Bodrossy, L., N. Stralis-Pavese, J. C. Murrell, S. Radajewski, A. Weilharter, and A. Sessitsch.** 2003. Development and validation of a diagnostic microbial microarray for methanotrophs. Environ. Microbiol. **5:**566–582.

4. **Chandler, D. P., G. J. Newton, J. A. Small, and D. S. Daly.** 2003. Sequence versus structure for the direct detection of 16S rRNA on planar oligonucleotide microarrays. Appl. Environ. Microbiol. **69:**2950-2958.

5. **Churchill, G. A.** 2002. Fundamentals of experimental design for cDNA microarrays. Nat. Genet. **32:**490-495.

6. **Cox, W. G., M. P. Beaudet, J. Y. Agnew, and J. L. Ruth.** 2004. Possible sources of dye-related signal correlation bias in two-color DNA microarray assays. Anal. Biochem. **331:**243-254.

7. **Dimitrov, R., and M. Zuker.** 2004. Prediction of hybridization and melting for double-stranded nucleic acids. Biophys. J. **87:**215-216.

8. **El Fantroussi, S., H. Urakawa, A. E. Bernhard, J. J. Kelly, P. A. Noble, H. Smidt, G. M. Yershov, and D. A. Stahl.** 2003. Direct profiling of environmental microbial populations by thermal dissociation analysis of native rRNAs hybridized to oligonucleotide microarrays. Appl. Environ. Microbiol. **69:**2377-2382.

9. **Franke-Whittle, I. H., S. H. Klammer, and H. Insam.** 2005. Design and application of an oligonucleotide microarray for the investigation of compost microbial communities. J. Microbiol. Methods **62:**37-56.

10. **Franke-Whittle, I. H., S. H. Klammer, S. Mayrhofer, and H. Insam.** 2006. Compaison of different labeling methods for the production of labeled target DNA for microarray hybridization. J. Microbiol. Methods **65:**117-126.

11. **Gao, X. L., E. Gulari, and X. C. Zhou.** 2004. In situ synthesis of oligonucleotide microarrays. Biopolymers **73:**579-596.

12. **Gao, X. L., E. LeProust, H. Zhang, O. Srivannavit, E. Gulari, P. L. Yu, C. Nishiguchi, Q. Xiang, and X. C. Zhou.** 2001. A flexible light-directed DNA chip synthesis gated by deprotection using solution photogenerated acids. Nucleic Acids Res. **29:**4744-4750.

13. **Guo, Z., R. A. Guilfoyle, A. J. Thiel, R. Wang, and L. M. Smith.** 1994. Direct fluorescence analysis of genetic polymorphisms by hybridization with oligonucleotide arrays on glass supports. Nucleic Acids Res. **11:**5456–5465. .

14. **Hashsham, S. A., L. M. Wick, J. M. Rouillard, E. Gulari, and J. M. Tiedje.** 2004. Potential of DNA microarrays for developing parallel detection tools (PDTs) for microorganisms relevant to biodefense and related research needs. Biosens. Bioelectron. **20:**668-683.

15. **Kane, M. D., T. A. Jatkoe, C. R. Stumpf, J. Lu, J. D. Thomas, and S. J. Madore.** 2000. Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. Nucleic Acids Res. **28:**4552-4557.

16. **Koehler, R. T., and N. Peyret.** 2005. Thermodynamic properties of DNA sequences: characteristic values for the human genome. Bioinformatics **21:**3333-3339.

17. **Koizumi, Y., J. J. Kelly, T. Nakagawa, H. Urakawa, S. El-Fantroussi, S. Al-Muzaini, M. Fukui, Y. Urushigawa, and D. A. Stahl.** 2002. Parallel characterization of anaerobic toluene- and ethylbenzene-degrading microbial consortia by PCR-denaturing gradient gel electrophoresis, RNA-DNA membrane hybridization, and DNA microarray technology. Appl. Environ. Microbiol. **68:**3215-3225.

18. **Lane, D. J.** 1991. 16S/23S rRNA sequencing, p. 115–148. *In* E. Stackebrandt and M. Goodfellow (ed.), Nucleic Acid Techniques in Bacterial Systematics. John Wiley, New York.

19. **Lane, S., J. Evermann, F. Loge, and D. R. Call.** 2004. Amplicon secondary structure prevents target hybridization to oligonucleotide microarrays. Biosens. Bioelectron. **20:**728-735.

20. **Li, E. S. Y., J. K. K. Ng, J. H. Wu, and W. T. Liu.** 2004. Evaluating single-base-pair discriminating capability of planar oligonucleotide microchips using a non-equilibrium dissociation approach. Environ. Microbiol. **6:**1197-1202.

21. **Longo, M. C., M. S. Berninger, and J. L. Hartley.** 1990. Use of uracil DNA glycosylase to control carry-over contamination in polymerase chain reactions. Gene **93:**125-128.

22. **Loy, A., C. Schulz, S. Lucker, A. Schopfer-Wendels, K. Stoecker, C. Baranyi, A. Lehner, and M. Wagner.** 2005. 16S rRNA gene-based oligonucleotide microarray for environmental monitoring of the betaproteobacterial order "*Rhodocyclales*". Appl. Environ. Microbiol. **71:**1373–1386.

23. **Palmer, C., E. M. Bik, M. B. Eisen, P. B. Eckburg, T. R. Sana, P. K. Wolber, D. A. Relman, and P. O. Brown.** 2006. Rapid quantitative profiling of complex microbial populations. Nucleic Acids Res. **34:**e5.

24. **Peplies, J., F. O. Glockner, and R. Amann.** 2003. Optimization strategies for DNA microarray-based detection of bacteria with 16S rRNA-targeting oligonucleotide probes. Appl. Environ. Microbiol. **69:**1397-1407.

25. **Peytavi, R., L.-y. Tang, F. R. Raymond, K. Boissinot, L. Bissonnette, M. Boissinot, F. J. Picard, A. Huletsky, M. Ouellette, and M. G. Bergeron.** 2005. Correlation between microarray DNA hybridization efficiency and the position of short capture probe on the target nucleic acid. BioTechniques **39:**89-96.

26. **Pozhitkov, A., P. A. Noble, T. Domazet-Loso, A. W. Nolte, R. Sonnenberg, P. Staehler, M. Beier, and D. Tautz.** 2006. Tests of rRNA hybridization to microarrays suggest that hybridization characteristics of oligonucleotide probes for species discrimination cannot be predicted. Nucleic Acids Res. **34:**e66.

27. **Ramakrishnan, R., D. Dorris, A. Lublinsky, A. Nguyen, M. Domanus, A. Prokhorova, L. Gieser, E. Touma, R. Lockner, M. Tata, X. Zhu, M. Patterson, R. Shippy, S. T.J., and A. Mazumder.** 2002. An assessment of Motorola CodeLinkTM microarray performance for gene expression profiling applications. Nucleic Acids Res. **30:**e30.

28. **Ratushna, V. G., J. W. Weller, and C. J. Gibas.** 2005. Secondary structure in the target as a confounding factor in synthetic oligomer microarray design. BMC Genomics **6:**31.

29. **Reyes-Lopez, M. A., A. Mendez-Tenorio, R. Maldonado-Rodriguez, M. J. Doktycz, J. T. Fleming, and K. L. Beattie.** 2003. Fingerprinting of prokaryotic 16S rRNA genes using oligodeoxyribonucleotide microarrays and virtual hybridization. Nucleic Acids Res. **31:**779-789.

30. **Reynaldo, L. P., A. V. Vologodskii, B. P. Neri, and V. I. Lyamichev.** 2000. The kinetics of oligonucleotide replacements. J. Mol. Biol. **297:**511-520.

31. **Riccelli, P. V., K. E. Mandell, and A. S. Benight.** 2002. Melting studies of dangling-ended DNA hairpins: effects of end length, loop sequence and biotinylation of loop bases. Nucleic Acids Res. **30:**4088-4093.

32. **Rouillard, J.-M., M. Zuker, and E. Gulari.** 2003. OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. Nucleic Acids Res. **31**:3057-3062.

33. **Sengupta, S., K. Onodera, A. Lai, and U. Melcher.** 2003. Molecular detection and identification of influenza viruses by oligonucleotide microarray hybridization. J. Clin. Microbiol. **41**:4542-4550.

34. **Small, J., D. R. Call, F. J. Brockman, T. M. Straub, and D. P. Chandler.** 2001. Direct detection of 16S rRNA in soil extracts by using oligonucleotide microarrays. Appl. Environ. Microbiol. **67**:4708-4716.

35. **Southern, E., K. Mir, and M. Shchepinov.** 1999. Molecular interactions on microarrays. Nat. Genet. **21**:5-9.

36. **Southern, E. M., S. C. Case-Green, J. K. Elder, M. Johnson, K. U. Mir, L. Wang, and J. C. Williams.** 1994. Arrays of complementary oligonucleotides for analysing the hybridisation behaviour of nucleic acids. Nucleic Acids Res. **22**:1368-1373.

37. **Taroncher-Oldenburg, G., E. M. Griner, C. A. Francis, and B. B. Ward.** 2003. Oligonucleotide microarray for the study of functional gene diversity in the nitrogen cycle in the environment. Appl. Environ. Microbiol. **69**:1159-1171.

38. **Tiquia, S. M., L. Wu, S. C. Chong, S. Passovets, D. Xu, Y. Xu, and J. Zhou.** 2004. Evaluation of 50-mer oligonucleotide arrays for detecting microbial populations in environmental samples. BioTechniques **36**:664-70, 672, 674-5.

39. **Urakawa, H., P. A. Noble, S. El Fantroussi, J. J. Kelly, and D. A. Stahl.** 2002. Single-base-pair discrimination of terminal mismatches by using oligonucleotide microarrays and neural network analyses. Appl. Environ. Microbiol. **68**:235-244.

40. **Vainrub, A., and B. M. Pettitt.** 2003. Surface electrostatic effects in oligonucleotide microarrays: Control and optimization of binding thermodynamics. Biopolymers **68**:265-270.

41. **Wang, H., J. Li, H. Liu, Q. Liu, Q. Mei, Y. Wang, J. Zhu, N. He, and Z. Lu.** 2002. Label-free hybridization detection of a single nucleotide mismatch by immobilization of molecular beacons on an agarose film. Nucleic Acids Res. **30**:e61.

42. **Wick, L. M., J. M. Rouillard, T. S. Whittam, E. Gulari, J. M. Tiedje, and S. A. Hashsham.** 2006. On-chip non-equilibrium dissociation curve and dissociation

rate constant as methods to assess specificity of oligonucleotide probes. Nucleic Acids Res. **34:**e26.

43. **Wilson, W. J., C. L. Strout, T. Z. DeSantis, J. L. Stilwell, A. V. Carrano, and G. L. Andersen.** 2002. Sequence-specific identification of 18 pathogenic microorganisms using microarray technology. Mol. Cell Probes. **16:**119-127.

44. **Zhang, L., T. Hurek, and B. Reinhold-Hurek.** 2005. Position of the fluorescent label is a crucial factor determining signal intensity in microarray hybridizations. Nucleic Acids Res. **33:**e166.

CHAPTER FOUR

DEVELOPMENT OF EMBEDDED CONTROL SYSTEMS FOR GENE ANALYZER

4.1 INTRODUCTION

Development of embedded control systems and housing, generally assumed to be mature technologies, may pose a major challenge during development of a device in a university setting. Even though the technology is mature, the service is only available through expensive powerhouses requiring 6 to 24 months worth of consulting fee for a team of engineers amounting to several hundred thousand dollars. At universities, the skill may exist in the appropriate disciplines but not in a single laboratory. Thus the development of basic controls, internal structure, and housing becomes a problem in itself. Without these components, the device doesn't exist and work related to validation is not possible. Considering the above factors, a decision was made to develop the embedded system for the gene analyzer using basic skills in microcontroller programming, data communication, software programming, and CAD to design the housing. This chapter describes selection and control of components for the gene analyzer. Compared to the structure of previous chapters that were written as journal manuscripts, the structure of this chapter differs for two reasons: i) the content of this chapter is more suitable for a book chapter, and ii) there was a need to document the embedded system programming more thoroughly for future students.

Control and electronic related aspects in development of the gene analyzer consisted of the following tasks:

1. Temperature control using LabVIEW, USB-DAQs, and breadboard.
2. Detection system control using LabVIEW, USB-DAQs and breadboard.

3. Design of custom PCB to accommodate ARM microcontroller, ICs, and LED drivers.
4. Replacement of USB-DAQs and PC with custom ICs and ARM microcontroller.
5. Development of circuit to establish communication between ARM microcontroller and an iPod Touch, and
6. Design of housing.

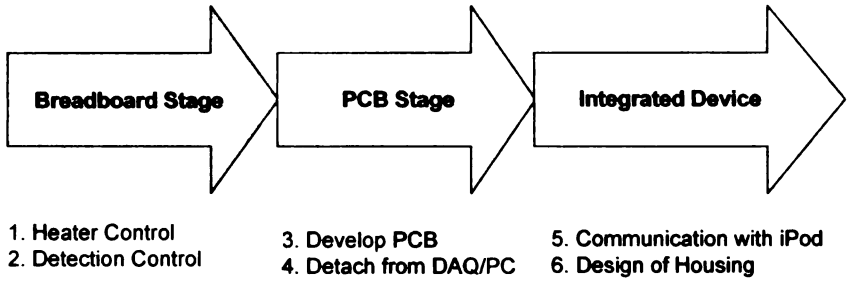


Figure 4.1 Schematic depicting the various stages during the development of gene analyzer.

4.2. TEMPERATURE CONTROL USING LABVIEW AND BREADBOARD

Initial control of components and data acquisition was obtained using National Instruments LabVIEW (Laboratory Virtual Instrumentation and Engineering Workbench) with USB compatible data acquisition accessories (USB-DAQ) as analog and digital inputs/outputs to and from a personal computer (PC). The USB-DAQ boards along with LabVIEW graphical interface provided a flexible means of control with minimal programming experience. Several integrated circuits (IC) were also used to aid in critical aspects of control (Figure 2). The small, single structured assembly of ICs made them ideal for deployment into a hand-held based platform.

to d

acqu

may

requ

deri

cont

ther

anal

belo

Figu

4.2.1

temp

accur

In general, a control strategy had to provide both accurate and precise temperature to drive the amplification reaction and synchronize control of a detector scheme to acquire amplification signal measurements in real-time. A means of feedback control may be required to drive the amplification reaction with precise temperature requirements. Use of pulse-width-modulation (PWM, 3, 8), or proportional–integral–derivative control (PID, 1, 5, 7) is well known for accurate and precise temperature control. The combined use of both PWM and PID together on a system with very small thermal mass has resulted in one of the most rapid cycling rates (4). For the gene analyzer, developed as part of this study, a coupled PID/PWM was used as described below.

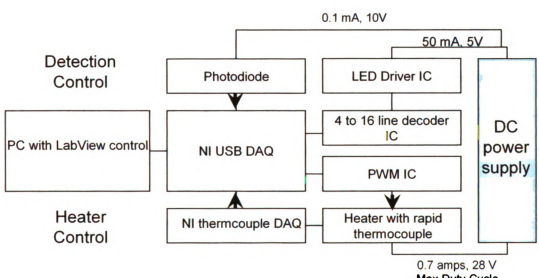


Figure 4.2 Schematic of temperature and detection system control for the gene analyzer.

4.2.1 Hardware for coupled PID/PWM implementation

Critical aspects of temperature control included using a rapid response temperature sensor, a heater that allowed rapid temperature cycling, and a strategy for accurate/precise feedback control. Signal from the temperature sensor had to be

amplified and conditioned for accurate control. Final hardware components for the device consisted of a rapid response thermocouple, a pulse width modulator IC (Texas Instruments DRV102T, $8.40), 2 USB DAQs from National Instruments (NI), the heater itself (2.54 x 7.62 cm Kapton heater, 40 ohm, Electro-Flex Heat, Inc., $40), and a LabVIEW VI developed in-house (Figure 3).

Rapid response temperature measurements were obtained using a type T thermocouple with a 0.004" sheath diameter (5SRTC-TT-T-40-36, Omega, $15.00) with a response time of approximately 0.1 sec and a specified standard limit of error less than 1 °C. An extension cable was used to increase the size of the thermocouple wires for easier handling and preventing damage that could influence the temperature measurement (TECT10-13, omega, $15.00). Initial evaluation of temperature sensors included surface mountable thermistors, however, these devices did not have the response time required for accurate temperature control causing drastic overshoot from the set-point before the senor responded to the correct temperature. Voltage output produced by the thermocouple was amplified and conditioned using the National Instruments USB-DAQ (9211, $530). The 9211 DAQ automatically converted the thermocouple output into temperature, which was read directly by the LabVIEW VI for feedback control. A LabVIEW PID subVI (graphical interface with a specific function inside of a graphical interface) provided an analog output (from multifunction NI USB-DAQ 6009, $280) based on the deviation from the set-point temperature (called using a signal generator VI described below). Large temperature deviations below the set point produced an output less than 0.5 V, small deviation below the set point produced a larger voltage (0.56 to 3.9 V), and deviation above the set point produced an output of 4 V. This output had to be

programmed into the DAQ assistant (described below) that was used for the analog output.

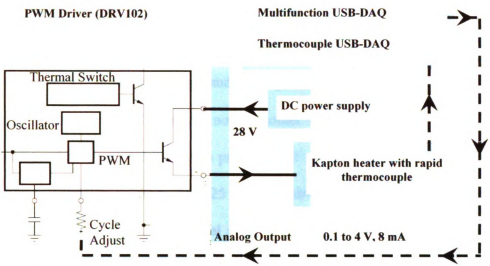


Figure 4.3 Detailed schematic of circuit used for feedback control of the heater on the bench-top device including description of how components were connected to USB DAQ and PWM IC. Dotted lines represent signal and solid lines represent power.

A pulse width modulation (PWM) driver was used for highly precise temperature control. The analog output from the DAQ was connected to the duty cycle pin of the PWM driver. When a larger duty cycle is required, a relay connecting the power supply (Vs pin) with the heater (Out pin) inside the PWM IC is closed for a larger fraction of the time, thus the average power delivered to the heater is increased. This allows the amount of power provided to the heater over time to be changed without using a potentiometer. One advantage of this strategy is that it reduces the total amount of power delivered to a load without losses normally incurred with sources limited by resistive means. All other pins on the PWM driver were left unconnected, except the ground. The selected PWM driver is ideal for smaller heaters and can supply a large range of voltages (8 to 60 V), but is not rated for a current above 2.7 A (the selected Kapton heater used a maximum

current of 0.45 A with 28 V). Therefore, larger heaters and devices that are driven with more power such as thermoelectric Peltiers will require a different means of PWM control. Some important notes for using the PWM IC include using a 1 to 10 k resistor connected in series between the analog output and duty cycle pin to limit swing to ground to be no less than 0.1V. In addition, a grounded 0.1 μF capacitor was placed in parallel between the power supply and IC to reduce noise to other sensors within the circuit. The PWM IC has a thermal shutdown switch to protect itself when close to overheating that will reset when temperature drops below 125 °C. The driver also has a pin that can be connected to a LED to signal when thermal shutdown has occurred. To date, thermal shutdown has never occurred when controlling a doped silicon heater or the Kapton heater.

4.2.2 Software for coupled PID/PWM implementation

The LabVIEW VI consisted of five main mechanisms, a PID SubVI, a signal generator to call the temperature profile, the analog inputs to read the thermocouples, analog output for feedback control of the PWM driver, and a program to write temperature data to a file (Figure 4). The Simulate Arbitrary Signal SubVI was used to generate a user specified temperature profile over time. The user can change the temperature profile by double clicking on the Arbitrary Signal SubVI, clicking on define signal, and manually entering time and temperature set-points in a tabular fashion. The signal generator was wired to the set-point of the PID SubVI. Analog inputs were obtained using a series of DAQmx SubVIs to obtain signal from the USB-DAQ 9211 and to report errors. Input from the DAQmx SubVI was wired to the process variable pin of

the PID SubVI. The output from the PID was wired to a DAQ assistant SubVI, which can be used for both analog and digital inputs/outputs. The DAQ assistant makes connection to the multifunction USB-DAQ simple, leading the user through the process of selecting if signal will be acquired or generated, what is to be measured (e.g. voltage, temperature, current, resistance), and suggests pin connection to the USB-DAQ. This is the first tool that should be used by an inexperienced LabVIEW user for accepting and generating signals from the PC via the multifunction USB-DAQ. The VI shown in Figure 3 also has the input for an IC thermocouple amplifier and conditioner, as a low cost alternative to the USB DAQ 9211 (described in section 4.4). The VI reads a voltage using the DAQ assistant, and a linear equation is used to convert the voltage to temperature. The conversion equation was provided within the specification of the IC conditioner. Finally, the temperature from all thermocouples is recorded along with a time stamp into a tab-delimited file that can be named by the user by double clicking on the Write to Measurement File SubVI.

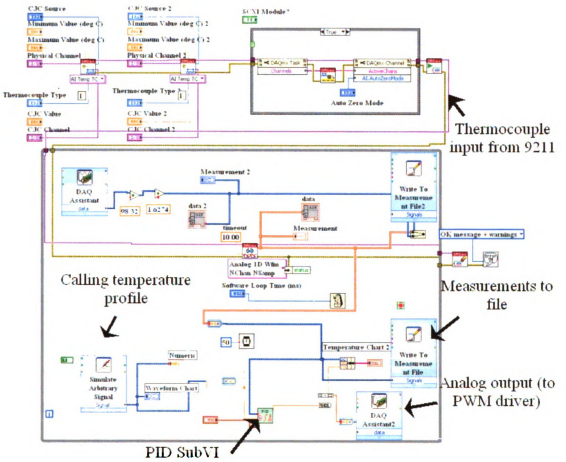


Figure 4.4 The temperature control VI developed to drive the amplification reaction, acquires analog input measurements from the thermocouple, and generates analog output with PID based feedback control

The front panel for the temperature control VI is user friendly in that only a few parameters need to be selected when starting the program (lower left corner, Figure 5). First, the user must set the optimal PID control parameters. The user must then select the physical channels for each of two thermocouples, and type of thermocouples (default selection shown if Figure 4). Two thermocouples can be used for dual measurements of either the difference in temperature across the surface of the amplification chip, or to examine the temperature change between the top and bottom of the amplification chip. However, only one of the thermocouples is used for feedback control. To determine

which thermocouple is connected to the PID controller, the user can start the VI, watch the graph on the lower right corner, and look for a temperature response by placing the thermocouple between their fingers. The thermocouple used for temperature control should be placed between the heater and the chip. The lower right graph also shows the set-point temperature and the voltage output from the PID used to increase duty cycle of the PWM IC (Figure 5). The top right graph also shows the set point, but over a longer period of time, and the top left graph shows the temperature of the second thermocouple (not controlling the PID). The axis of all graphs can be changed to reduce or increase temperature and time scales. This can become useful when manual tuning the PID to a single temperature. For example, when tuning the PID for maintaining temperature to drive loop mediated amplification; the scale of the y-axis on the lower right graph can be reduced to 60 to 66 ^{o}C. The VI can be stopped using the button on the front panel or the stop button on the toolbar.

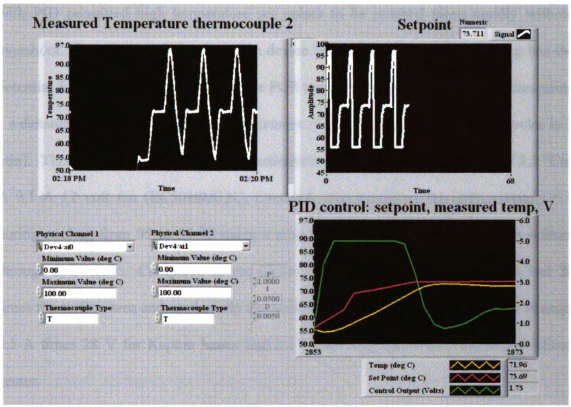


Figure 4.5 Front panel of temperature control VI. User changeable parameters including constants for PID control are in the lower left corner (default parameters are shown). The lower right graph shows the temperature measured by the thermocouple used for feedback control (yellow), set-point (red), and the output voltage from the PID (green). The upper right graph shows the set points over time, and the upper left graphs show the temperature measured with the second thermocouple

4.2.3 Performance of coupled PID/PWM implementation

Results of using PWM along with PID control showed that both control methods were important for precise control of the heater and to reduce overshoot and oscillation from the set points. Using a 1 x 1 x 0.05 cm piece of doped silicon as a test heater, the combined control of PWM and PID reduced overshoot and standard deviation of the set-point from 0.80 °C with PID alone to 0.28 °C with PWM and PID (measured during 72 °C set-point for 8 seconds, Figure 6 left panel). Compared to the sole use of PID (to open/close a relay between the heater and power supply), added precision using PWM

with PID permitted high temperature set points to be reached more rapidly without overshoot. For future refinement of the device for quantitative PCR, this setup has the potential for running a 30 cycle real-time PCR assays in approximately 12 minutes using 1 s denaturing, 1 s annealing, and 8 s extension (Figure 6 right panel shows 8 cycles in 3 min). This makes a heating rate of approximately 20 °C/s and a cooling rate of 3.5 °C/s. A 0.1 A 12 volt fan (EC5010M12C, Evercool) was turned on for approximately 6 s during cooling from 95 to 55 ^{0}C using a relay switch that was controlled with a timer through LabVIEW. Compared to the silicon doped heater, the Kapton heater had 20 times the surface area and used nearly the same amount of power for temperature change, 0.5 A times 28 V for Kapton heater and 1.13 amp times 10 V with the doped silicon heater.

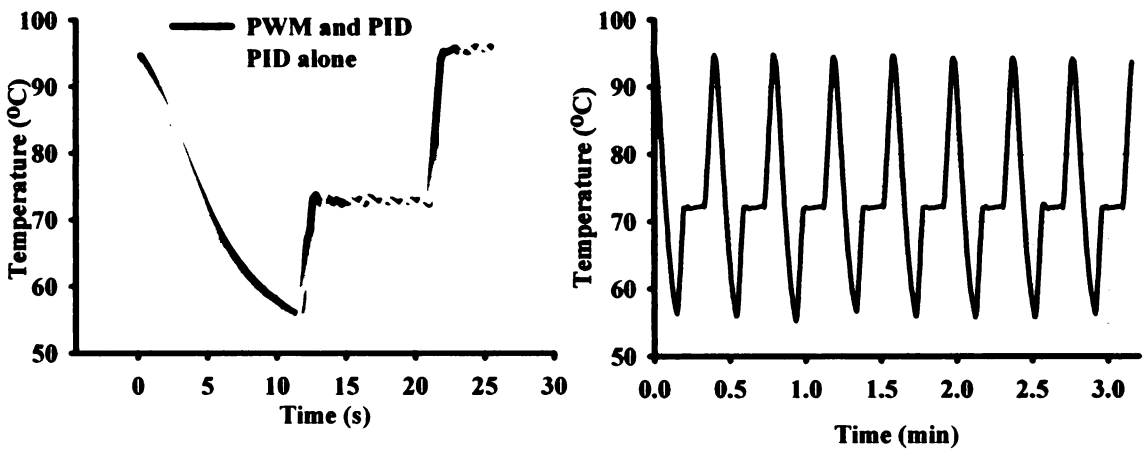


Figure 4.6 Left panel demonstrates a thermocycle temperature profile using PID/PWM feedback control and PID control without PWM (tested using doped silicon heater). The right panel shows 8 cycles performed in approximately 3 minutes using a 0.3 mm thick Kapton heater.

PID tuning was essential for reducing deviation from set-points and limiting overshoot during rapid temperature changes. For obtaining the rapid cycling demonstrated in figure 4.6, the PID parameters were manually tuned as follows. The K_P

parameter was the first to be tuned while the K_I and K_D were kept constant at zero. A higher K_P parameter increased the amount of change in the output for a given change in the error. Thus, a K_P that was too high caused the system to become unstable, producing large fluctuations in temperature from the set-point. In contrast, a small K_P resulted in a small output response to a large input error, making feedback less responsive and even prohibiting the heater from reaching the set-point temperature. For our setup, the K_P term heavily contributed to the majority of feedback control compared to the K_I and K_D constants. Next, the K_I constant was adjusted by finding the optimal constant between reaching steady-state more rapidly (higher K_I) and reducing overshoot from the set-point (lower K_I). The K_D is typically used to reduce overshoot, however, no observable change was measured, and therefore a low K_D constant was used. The PID constants used for the measurements in Figure 4.5 were a K_P of 1, a K_I of 0.05, and a K_D of 0.005.

4.3 Detection system control using LabVIEW and breadboard

An optical setup for capturing real time amplification signals in a multichamber assay will require synchronized control between the light source and the detection system. Multiple solutions exist for this purpose and have been in use in larger more expensive instruments. Key requirements dictated by the small footprint of gene analyzer were: i) the total detection system must not be larger than 20 cm^3 and the total cost must not exceed $200. This was achieved by the development of an optical system consisting of: i) 16 LEDs coupled to an IC that allowed multiplexing, ii) individual LED drivers to provide accurate and constant current to all LEDs, iii) synchronized control of LEDs with a photodiode (PD) used to acquire change in signal emission, and iv) an IC to reduce the

conditioning noise acquired from the PD. Details of the detection system are described below.

4.3.1 HARDWARE FOR DETECTION SYSTEM CONTROL

Critical aspects in controlling the detection setup included synchronized timing of components, controlling multiple LEDs with a limited number of digital outputs, obtaining a constant and programmable current to the LEDs, reducing noise from the photodiode, and using a strategy to prevent the photodiode from overloading. Control of the optical setup consisted of 16 light emitting diodes (LEDs) for exciting a dye in the amplification reaction, a photodiode (PD) for measuring emission, three separate ICs to aid in control, the same multifunction NI USB DAQ (6009) used for the temperature control, and a LabVIEW VI developed in-house. The following describes the hardware in more detail and connection between components, ICs, and the multifunction USB-DAQ (Figure 4.7).

The number of controllable LEDs can be limited to the number digital outputs on a processor (USB-DAQ has 12). Therefore, a 4-to-16 line decoder/demultiplexer (CD74HCT4514, Texas Instruments, $2.25) was used to individually control all 16 LEDs using only 5 digital outputs from the USB-DAQ. For communicating with the decoder, a high logic signal (3.3 V) was provided at all times to the LE pin. The on/off logic configuration of 4 other digital outputs from the USB-DAQ was used for gingival control of the LEDs. The small amount of current provided via the digital lines from the USB-DAQ was sufficient to power the decoder without connecting a power supply to the Vcc pin. The LE pin provides the possibility of cascading between multiple decoders using

only 1 additional line per decoder. To summarize, the first decoder requires five digital outputs and each new set of 16 LEDs only requires one additional digital output. In total, a maximum set of 128 addressable LEDs could be used with the 12 digital outputs on the USB-DAQ board.

Continuous current had to be delivered to the LEDs for reproducible and quantifiable measurements of the amplification reaction. This was obtained using a 50 mA stand-alone linear LED driver IC (STmicroelectronics, STLA01, $0.95). A high logic output of 3.3 V from the decoder (pins Y0 to Y15 of decoder connected to the EN pin of each driver) instructed drivers to turn on 1 of the 16 LEDs. In addition to providing constant current, the LED driver could be programmed to supply current based on the requirements of the user without the need for an external sensor. A power supply provided 5 V to the LED drivers, which were programmed (using a 4.6 kilo ohm resistor) to provide 26 mA with ± 5% accuracy to the 30 mA rated LEDs. A higher resistance resistor (connected to the PROG pin) could be used to reduce power to the LEDs and a lower resistance resistor could be used to increase the deliverable current up to 50 mA.

A critical aspect in using the photodiode for accurate, reproducible, and quantifiable detection of emitted light was reducing noise and preventing the PD from overloading. Overloading occurred when too much light was detected, causing the photodiode to generate inaccurate measurements, which could only be reset by briefly removing power. This problem was resolved by powering down the photodiode after every LED measurement cycle using a soft relay (Potter and Brumfield, T77S1D3-05, $1.80). A 5 V signal from the analog output of the USB-DAQ closed the relay switch, connecting the power supply to the photodiode for measurements during the LED

cycling. After each of 16 LEDs had been on for a user-defined amount of time and all PD measurements had been acquired, the relay would open, powering down the photodiode. Noise was also reduced using a grounded 0.1 μF capacitor in parallel with the V+ and V- pins (connected to a 10 volt power supply) as instructed in the photodiode specifications manual. The output pin of the photodiode was connected to the analog input on the USB-DAQ, which was also connected in parallel to a grounded 0.1 μF capacitor for additional noise reduction. Output reading from the photodiode and the circuit ground was connected to two analog inputs on the USB-DAQ for a obtaining a differential measurement (also implemented to reduce noise).

Through-hole configurations (DIP) were selected for all ICs to develop the prototype on a standard solderless breadboard (K and H products, SD-24, Figure 4.12A). The LED drivers were only available as a surface mountable configuration (DFN); therefore, they were mounted to DFN to DIP boards (E6-0091, ePBoard Design) and soldered using services provided by ePBoard Design. The total cost for the board and mounting was $12.15 per LED driver. It should be noted that all ICs were ordered from DigiKey. Ordering directly from the manufacturer typically required large volume, while single piece orders could usually be placed through DigiKey.

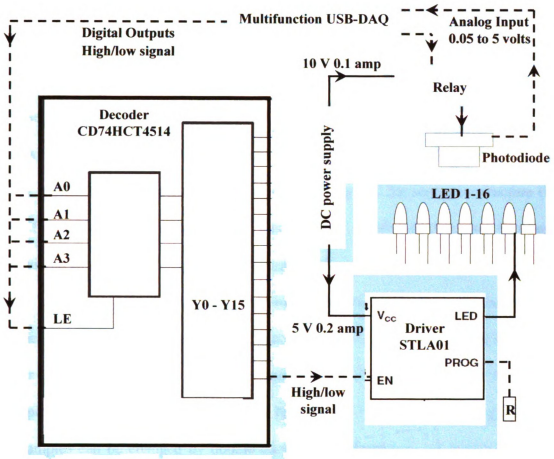


Figure 4.7 Detailed schematic of circuit used for the optical detection on the bench-top device including description of how components were connected to the multifunction USB DAQ and the ICs. Dotted lines represent signal and solid lines represent power.

4.3.2 SOFTWARE FOR DETECTION SYSTEM CONTROL

Precise timing of the device was obtained using a LabVIEW VI developed in-house. The VI for control of optical components consisted of six major mechanisms including a Boolean array for generating digital outputs to control LEDs, a DAQ assistant generating an analog output for controlling the relay between the PD and power supply, a DAQ assistant for analog input measurements from the photodiode, a VI to average all PD samples, flat sequence structure and loops for precise timing between LED and PD measurement acquisition, and a program to write PD data to a file (Figure 4.9). A

Boolean array (described in more detail in software section below) allowed user defined control of how many and which LEDs to utilize for a given experiment. The array of true/false calls was converted into an unsigned 32-bit integer and wired to a while loop to define how many LEDs to excite and PD readings to take per cycle. Inside and outside of the while loop was a DAQ assistant subVI that generated a signal to open and close the relay used between the photodiode and power supply. A flat sequence structure (looks like film reel) inside the while loop was used for synchronized control between the individual LEDs for a user specified amount of time, and reading the photodiode after the LED has been on for a user specified amount of time. The initial frame in the flat sequence loop generated digital outputs (to the decoder) from the unsigned integers. The latter of the two flat sequence frames acquired signals from the photodiode with a rate of sampling and number of samples defined by the user. The rate of sampling and number of samples can be changed by double clicking this DAQ assistant SubVI in the block diagram. An amplitude and level measurement subVI was used to generate a single data point by averaging all readings taken from the PD after each measurement. The average data point was time stamped and saved into a tab-delimited file named by the user. The data file can be directly imported into excel for analysis of real time amplification curves. An additional noise reduction strategy was implemented within the optical detection VI by changing the maximum and minimum anticipated voltage from 0 to 2.5 V (changed in DAQ assistant SubVI). The number of cycles can be selected by the user by entering a constant that is wired to a "for loop" that encompasses the entire block diagram.

The front panel of the optics control VI was designed to allow the user to change the following: 1. Number and position of LED readings using the Boolean

array, 2. Number of cycles, 3. Time for which the LEDs are on, 4. The amount of time to wait between measurement cycles, and 5. The time to allow the LED to warm prior to acquiring PD samples (Figure 4.8).

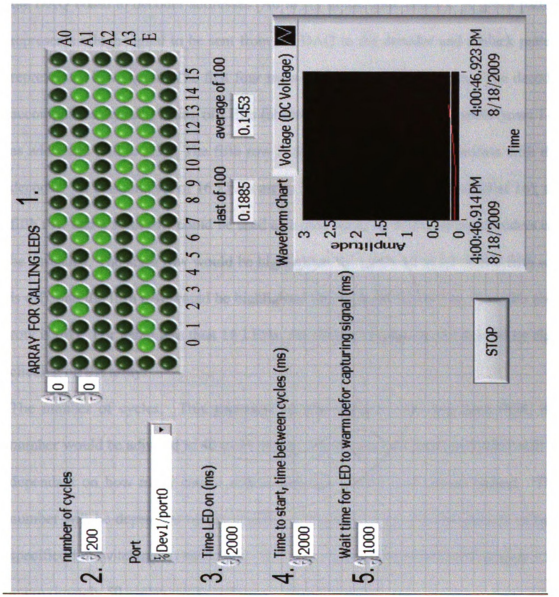


Figure 4.8 Front panel of optical component control VI. User changeable parameters include time LED is on, number of cycles, time to wait between cycles, time to wait for LED to warm before taking PD measurement, and the Boolean array for controlling individual LEDs. The chart shows the average PD measurement and the last measurement of multiple samples taken with the PD. Constants shown in the screen shot were used for experiments described in Chapter 5.

1. The LED array. This array is used to communicate with LED drivers to turn them on/off. Each column represents the sequential calling of the individual LEDs (0 refers to LED1, 15 refers to LED 16), and each row represents the digital output from the DAQ board to decoder addresses (A0 to A3 in the table below). A green pattern represents a high signal to be sent from the DAQ to the decoder and a black pattern represents a low signal. The first four rows are used to call the LEDs. The decoder accomplishes demultiplexing of LEDs (channel 0 to 15) using select inputs (rows 1-4) as addresses (A0 to A3). The fifth row is always on high to communicate with this decoder. If a second set of 16 LEDs are to be used (i.e. 32 LEDs instead of 16), the fifth row and a sixth row could be used as a chip select so that multiple decoders can be cascaded. The sixth row would be highlighted for LEDs 17 to 32 (while fifth row is off), and the fifth row would be highlighted to command LEDs 1 to 16 (while sixth row is off). For using less than 16 LEDs, the chosen column can be deleted by right clicking on the array.
2. The number of cycles. This parameter is user-defined. For real time PCR, this number would be adjusted to 40 to 45 cycles. The number of cycle for LAMP will be dependant on how many cycles will be obtained within a 30-minute period. This number will be dependant on the number of channels used, and the amount of time specified for switching on each LED. For the default numbers provided (Figure 4.8), approximately 50 cycles can be used for a 30 min reaction.

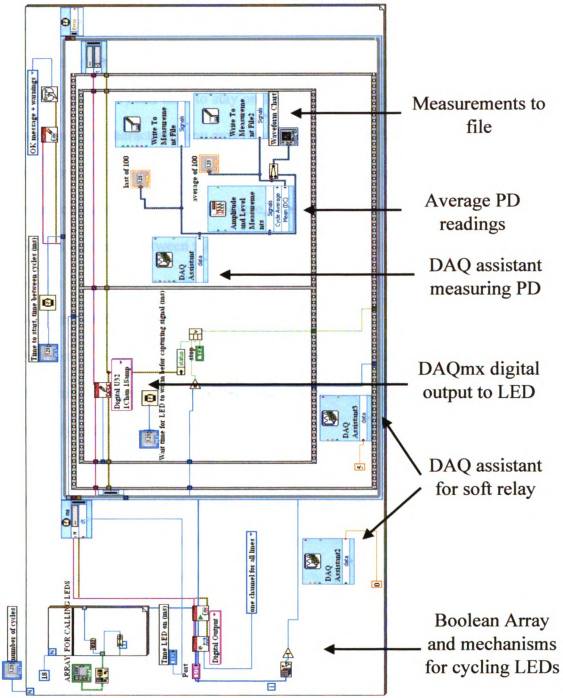


Figure 4.9 Block diagram of optical control VI developed to detect the amplification reaction, generate digital outputs for individual LED excitation using a Boolean array, and acquire PD measurements using DAQ assistant for analog input. A flat sequence structure loop was used to synchronize the LEDs with PD data acquisition.

3. The amount of time to turn on the LEDs (ms). This constant allows the user to determine how long each LED should be on. This includes time to allow LED to warm (control 5) and time to stay on for the PD reading. A default value of 2000 ms was used for characterization experiments described in chapter 5.
4. The time to wait before running next cycle. This time will be dependant on how many channels will be read per cycle, how long the LEDs will be on, and how much time is necessary between taking readings. For example, if 6 LEDs are measured and are on for 2 seconds each, and the user prefers 20 s between each measurement, a wait time of 8000 ms will be entered here.
5. Amount of time to wait before LED is measured. This control allows the user to determine how long to wait before measuring the PD to ensure that LEDs had reached a steady and high level of brightness.

4.3.3 PERFORMANCE OF THE DETECTION SYSTEM

The amount of time required for measurement of each well is critical in affecting the number of cycles (data points) that can be obtained to plot the real time amplification curve. The parameters used for experiments described in Chapter 5 are displayed in Figure 4.8. This included turning each LED on for 2 s, and acquiring a PD measurement after one second of allowing the LED to warm up. The average of 100 sample readings taken during a 0.2 sec time interval was recorded for the amplification curves. However, using these parameters reduced the amount of readings that could be taken during the exponential phase of amplification to 4 to 5 data points. Additionally, the amount of time

required to read all of the reaction wells may influence CV% of T_t between reactions on the same chip, subsequently reducing quantitative accuracy and reproducibility.

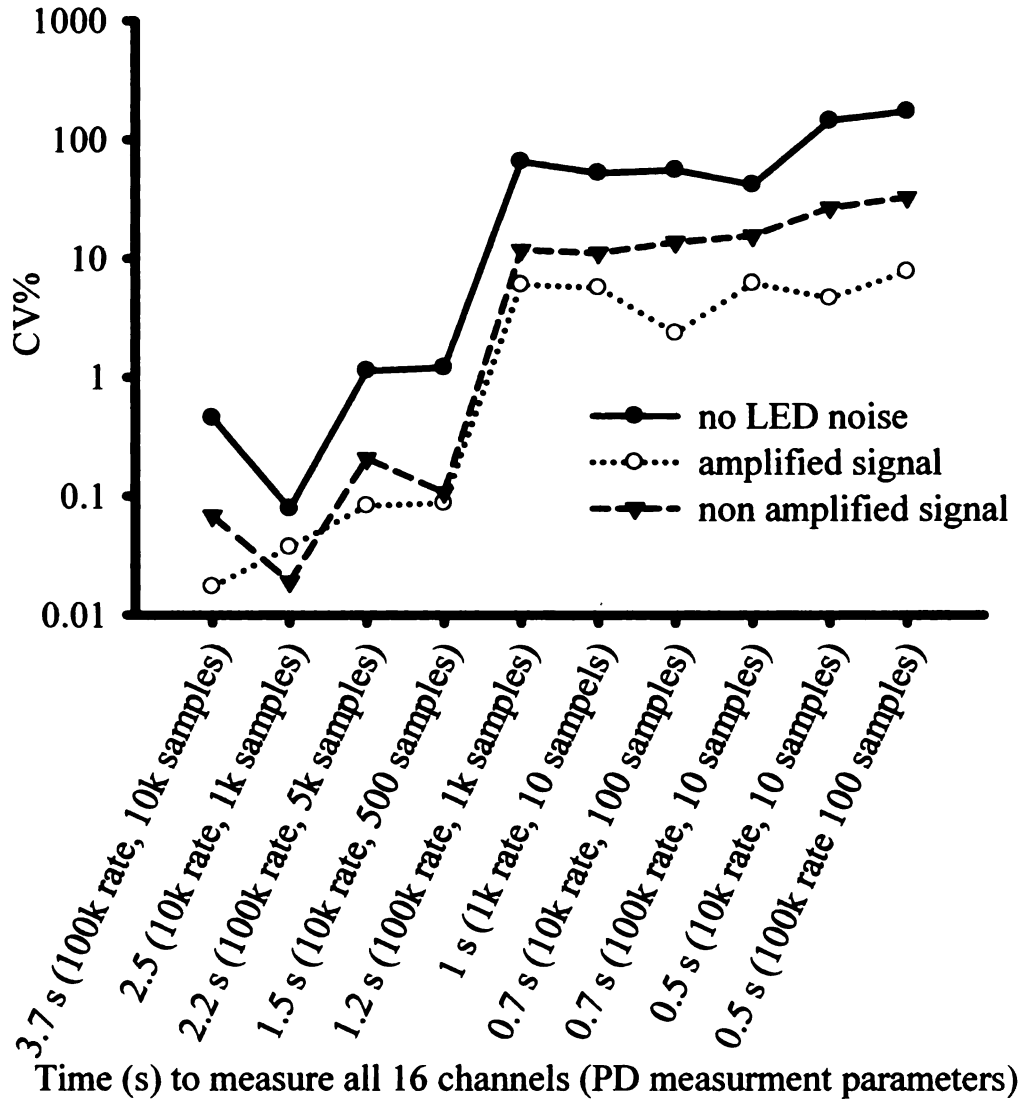


Figure 4.10 Coefficient of variation (%) for 4 replicate readings (y-axis on log scale) taken within the same reaction channel and a variation in the PD measurement parameters and therefore, the amount of time required for a one cycle measuring all 16 channels (x-axis). Readings were taken from 3 different channels including one without an LED, a channel with amplification, and a channel with no amplification

Further experiments were performed to examine the minimum amount of time that could be used for reproducible signal. Results demonstrated that a time lower than 1.5 s for acquiring reading from all 16 channels (0.09 s per channel) significantly

decreased the ability to obtain reproducible readings (Figure 4.10). Therefore, the 2 s cycle time per LED (used for data acquired for chapter 5) could be reduced to 0.09 s by obtaining 500 samples at a 10 kilo-hertz sampling rate. Irreproducible signal observed with reduced acquisition times may have been due to limitations of LEDs, PD, USB-DAQ, or timing of the LabVIEW VI. For example, the 10k sampling frequency of the PD always has a lower CV% than the 100k sampling frequency. Regardless of the cause, a total time of 1.5 s for measuring all 16-reaction wells should aid in the reduction of the CV% (described in Chapter 5).

4.4. Design of custom PCB to accommodate ARM microcontroller, ICs, and LED drivers

To reduce size of the circuit, a custom printed circuit board (PCB) was designed and fabricated to replace the breadboard. Using a PCB for the circuit also allows low cost mass production of the device (~$10 each for a 100 piece order). Issues related to PCB layout were critical for obtaining reproducible fabrication of the PCB with low noise, cost, and reducing accidental contact between two points in the circuit. A schematic of the circuit was drawn in ExpressSCH for initial determination of wire placement (Figure 4.11). An important point in using ExpressSCH, is that component footprints do not need to be accurate, however, a critical parameter for the development of the final PCB layout is the correct placement of wires between components. Therefore, each component (including resistors and capacitors) requires unique names with correctly identified pin numbers. Trace lines can be crossed as the software identifies connections via a node symbolized by a small square, which is only placed by

the user clicking on the line or pin. Simple components such as resistors, capacitor, ground, and ribbon connectors could be added from the component library. However, all ICs selected for the hand-held device had to be drawn manually. Once the schematic was completed, it was imported into the ExpressPCB software to aid in the correct connections between pins. This helped the user focus on critical PCB design such as spacing between traces instead of correct wiring. More specifically, clicking on a pin of a component caused all connections to be highlighted, thus traces could be drawn without continuously cross-referencing the schematic.

For prototyping purposes, the first set of PCB boards was designed using Express PCB's free software. Two separate PCBs were designed and fabricated using ExpressPCB. The first PCB functioned to hold all of the LEDs and consisted of a ribbon cable, and multiple through-holes for the LED pins (Figure 4.12, left panel). The second PCB was the circuit for the ICs and microcontroller and had a bottom and top layer of copper consisting of traces and pads for components, and a ground and power plane sandwiched in the middle (Figure 4.12, right panel). The 4-layer design was selected to reduce traces required for grounding and power, and for increased noise immunity for the PD and thermocouple sensors. In detail, the boards were designed with several important parameters including the following: vias (also called feed-through holes) were used to move traces from one layer to the other. The vias or through-holes can be connected to, or isolated from, the inner planes of the 4-layer board. A size of the hole used for all vias was 0.008". The vias become important when traces need to be crossed. An ideal layout for a board will consist of vertical traces on one side and horizontal traces on the other, with horizontal and vertical traces connected using the via. A width 0.010" was used for

most of the traces, as it is ideal for low current (less than 0.3 A) digital and analog signals. A trace of 0.025" (suggested for less than 1.0 A) was used to power the heater. While a minimum spacing of 0.010" was suggested between traces, a spacing of 0.050" was used to reduce the risk of a shorts developing in the board manufacturing process. One important design consideration is to avoid using sharp right angle turns when placing narrow traces (0.012" or less). A suitable alternative is to use two 45-degree bends with a short leg in between. A minimum of 0.021" space must remain between adjacent holes and the gap between every item (pad to pad, pad to trace, trace to trace) was 0.007" or greater. A 26-pin connection for a mountable ribbon cable was used to allow connection between the LED PCB and the component PCB.

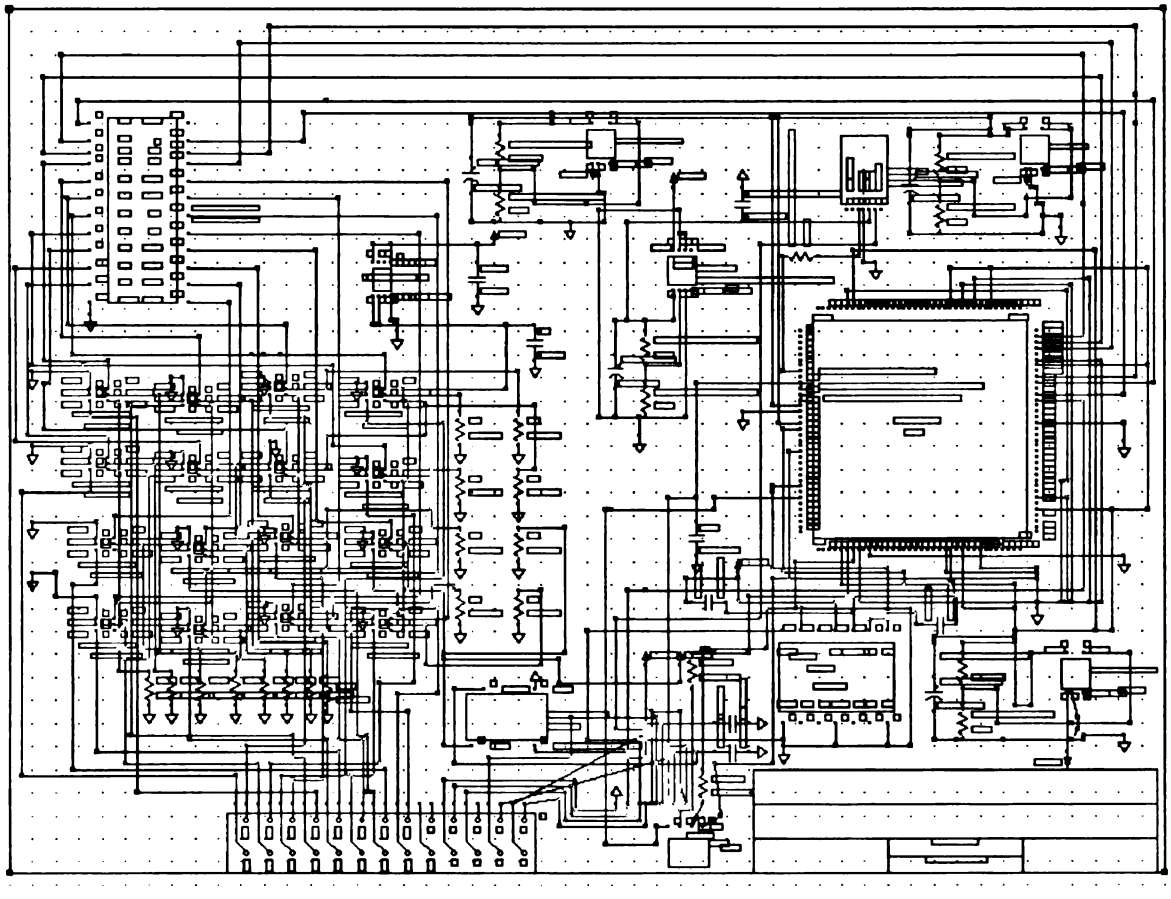


Figure 4.11 Schematic of circuit layout drawn in ExpressSCH

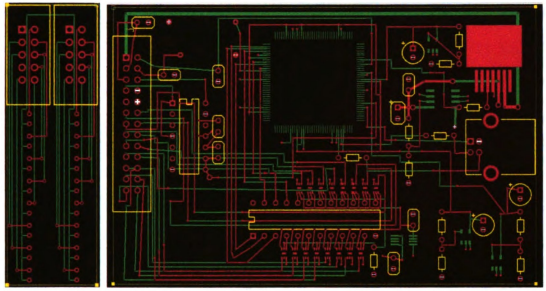


Figure 4.12 Screenshot of PCB layouts for the LED array and for ARM microcontroller and ICs

Many of the component ICs had standard sizing, therefore dimensions could be found in the component manager. However, some ICs were not in the component manager, requiring component specifications to be examined carefully for drafting in ExpressPCB. Pads for the majority of surface mountable ICs were placed on the bottom of the PCB while ICs requiring through-holes were placed on the top of the PCB (Figure 4.13B and 4.13C). Heat generated by the power regulating ICs and the PWM drivers was dissipated using copper pads with footprints suggested by specification manuals. Since most of the ICs can be purchased in several configurations (i.e. wide or thin, surface mountable or through-hole), it is important to ensure that the configuration selected for the PCB are not obsolete or require large quantity orders. For example, the decoder can be purchased in a thin configuration that is 7 mm wide or a configuration this is 14 mm

wide. However, the thin configuration can only be purchased in a minimum order of 285 pieces.

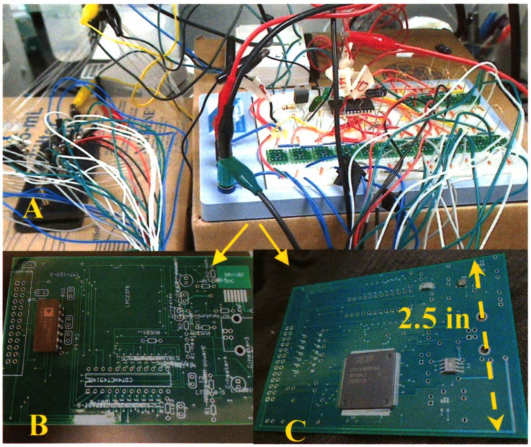


Figure 4.13 Picture of constructed circuit used to control and test components. A) bench-top control with circuit connected using a bread-board (USB-DAQ in the back). First version of custom PCB with B) front-side for through-hole ICs (only 1 is attached), and C) back-side of PCB for surface mountable ICs (ARM microcontroller and voltage regulators are shown).

This version of the PCB was designed with USB-B connections for communicating between the ARM device and a PC. Three identical component boards were printed and shipped for approximately $100. The MiniBoard manufacturing option was chosen which limits overall dimensions to 3.8 x 2.5", but reduces costs by $70.

These boards include the green solder mask and white silkscreen layers (with trace outlines and text to determine proper placement of ICs, Figure 4.13B and 4.13C).

4.5. Replacement of USB-DAQs and PC with custom ICs and ARM microcontroller

Embedded control of the LAMP hand-held device required replacing the LabVIEW Multifunction USB-DAQ, thermocouple conditioning USB-DAQ, and the PC with inexpensive alternatives suitable for a portable deployment. Microcontrollers are a cost and space-effective solutions for control compared to the large and costly USB-DAQ. Critical aspects in selection of a microcontroller include a simple and fast means of generating code. Familiarity with the LabVIEW programming environment dictated the selection of NI LabVIEW embedded module for ARM microcontrollers.

LabVIEW has recently released a module to aid in the programming of microcontrollers using their graphical interface (6). This module has been described as an ideal means of embedded control for engineers that are not accustomed to C/C++ programming (6). As of 2009, ARM (advanced reduced instruction set computer machine or advanced RISC machine) processors account for approximately 90% of all embedded 32-bit RISC processors, and are deployed in a wide range of applications such as iPods, game consoles, calculators, GPS navigators, industrial, and medical applications. Furthermore, approximately 98% of mobile phones sold each year use at least one ARM processor (2). In addition, the ARM microcontroller utilizes a CPU design with simplified instructions that can be executed quickly for obtaining a high level

of performance. Therefore, this processor was selected to control components and acquire data in the hand-held device.

Processor specifications for the 2 x 2 cm microcontroller (ARM7 NXP LPC2378) selected with the LabVIEW embedded module include the following: 58KB RAM, CPU clock up to 72 MHz, USB 2.0 Full Speed Device Controller, 10-bit analog-to-digital converter (ADC) with 8 channels, and a 10-bit digital-to-analog converter (DAC) with 1 channel. The module also contained an evaluation board (MBC2300), which allowed the functionality of the ARM microcontroller to be tested with the various components. Specification for the evaluation board include: serial interfaces, speaker, through-hole leads input/output pin on the ARM for prototyping, an LCD, USB connection, Ethernet connection, and 8 LEDs connected to digital outputs.

The LabVIEW ARM module automatically converted LabVIEW graphics based control VIs into C/C++ code and programmed the microcontroller for testing with the evaluation board. The workflow for this consisted of opening LabVIEW 8.6 and selecting ARM project target in the lower left portion of the "Getting Started" window. After saving the project filename, a "Project Explorer" window was opened that allowed the user to select digital, analog, PWM, input/outputs. After adding elemental inputs/outputs they could be dragged from the "Project Explorer" window into the block diagram and wired accordingly. Once the LabVIEW block diagram is completed, the run button is clicked and C++ code is generated and programmed to the ARM microcontroller via the Keil ULINK2 USB-JTAG Interface Adapter, a small hardware adapter that connects the USB port of the PC to the JTAG port on the evaluation board (Figure 4.14). Hand-held components were connected to the ARM controller using

through-holes in the prototyping area of the board. The board was powered through the USB, which also served to transmit measurements to the PC for ensuring proper control and measurements of the components with the ARM components.

The LabVIEW ARM module was highly efficient in terms of reducing time required for programming the microcontroller with a small learning curve. Basic LabVIEW VIs were written to test control of each component with the ARM module (Figure 4.15, 4.16). Sensor currents from the PD and thermocouple were converted to a digital value using the ARM microcontroller with an integrated analog-to-digital converter (ADC). Output from the temperature and PD sensors were displayed in real time on the LCD screen of the evaluation board and on LabVIEW charts (via the USB connection to the PC). Time that was saved in using the LabVIEW module, instead of manually writing code, became obvious when testing the components. For example, a simplistic LabVIEW block diagram for PID based control of the heater at a single temperature (Figure 4.15) generated 4,220 lines of code.

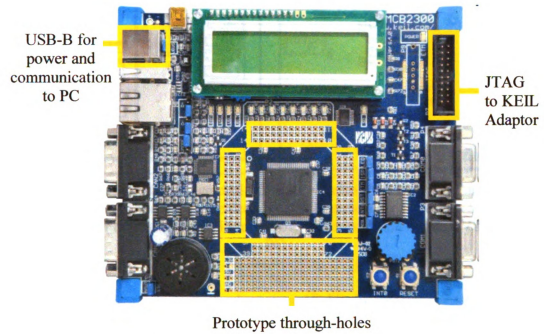


Figure 4.14 Picture of the ARM evaluation board used with the LabVIEW embedded control module. Sections of the board used to evaluate control of hand-held components are highlighted in yellow.

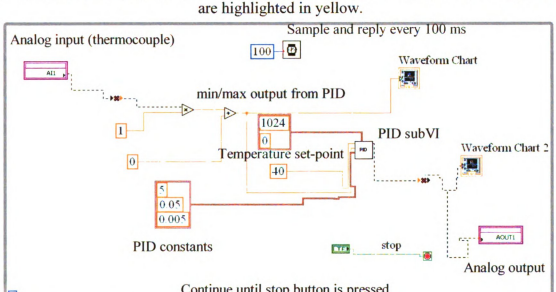


Figure 4.15 Block diagram of VI constructed to test PID-based feedback control of the heater with the LabVIEW embedded module. Approximately 4,220 lines of code were generated for this control scheme.

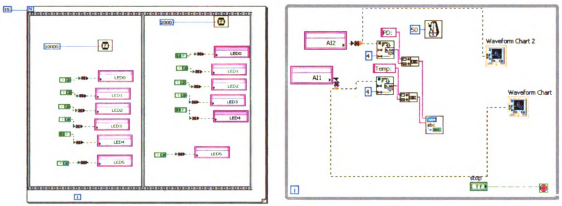


Figure 4.16 Left panel: Testing control of LEDs with Boolean control of multiple digital output channels connected to the decoder. Right panel: Analog inputs for PD and temperature sensor displayed on LED screen of evaluation board and LabVIEW charts on PC.

For feedback control of the heater, a 10 bit (minimum output set to 0 and maximum output set to 1024) resolution generated by the PID (subVI was converted to an analog signal that varied between 0 and 3.3 V for adjusting the duty cycle on the PWM IC. The evaluation board had one minor drawback in testing complete functionality of the PWM IC. This is due to the ARM drawing power for ADC and the DAC through the USB interface, allowing a maximum analog output voltage of 3.3 V. The PWM IC requires up to 4 V for a 0% duty cycle. This caused the heater to run at temperatures beyond the set-point. However, it was observed that once the temperature set-point was reached, the PID increased the analog voltage to 3.3 V and the amount of current drawn from the power supply was reduced. The maximum voltage to the ARM is approximately 4.6 V, therefore an operational amplifier can be used to provide the ARM DAC with enough voltage required for the PWM driver.

Portable deployment of the device also required replacing the $530 thermocouple conditioning USB-DAQ with an integrated circuit with similar function. The amplifier/conditioner is critical because raw signal generated by thermocouples produces

small voltage differences. For example, a type K thermocouple will generate a difference of 4 mV between 0 and 100 °C, therefore, this small change in voltage must be amplified and conditioned for accurate and precise measurement. While multiple ICs were tested for precise temperature acquisition with minimal noise, a monolithic chip from Analog Devices (AD595C, $20) with an instrumentation amplifier, a precalibrated accuracy of ±1°C, and thermocouple cold junction compensator offered the most accurate and precise control. Amplified and buffered signal from the AD595 IC provided a difference of 1,000 mV for temperatures between 0 and 100 °C. Both the NI USB-DAQ conditioner and Analog Device IC conditioner were tested at three separate temperatures to determine accuracy and precision. While both conditioners showed similar levels of precision, the integrated circuit had slightly better accuracy (Figure 4.17). The IC conditioner has an additional function in that it has a LED alarm to alert if the thermocouple has failed. Thus, the IC is a suitable alternative to the high cost and large USB-DAQ conditioner.

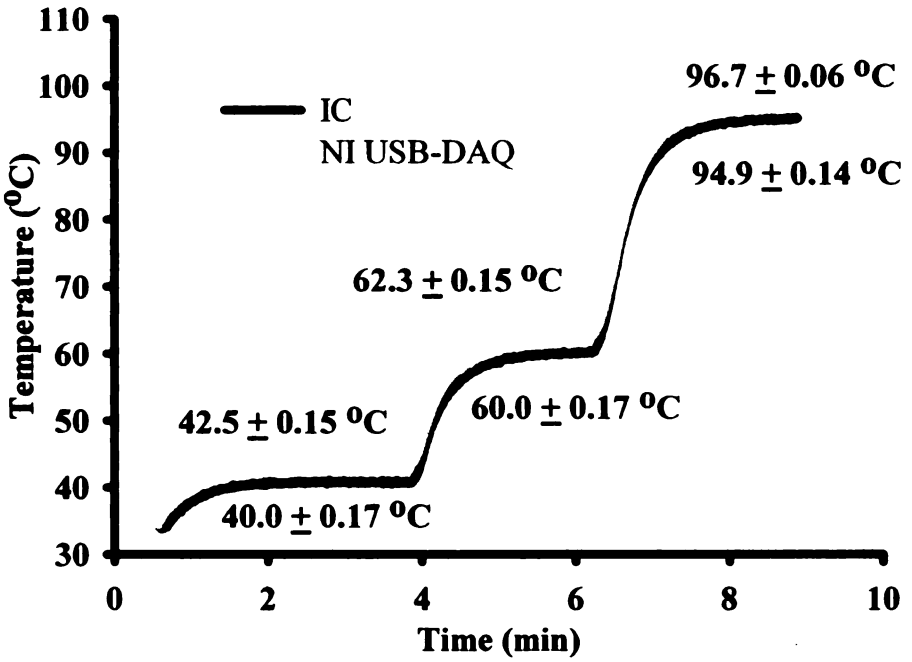


Figure 4.17 Temperature measured at three different set-points (40, 60, 95 °C) using a thermocouple connected to the National Instruments conditioning USB-DAQ (gray line) and Analog Devices IC (black line).

4.6. Development of circuit to establish communication between ARM microcontroller and an iPod Touch

The hardware and software for communicating between the ARM microcontroller and iPod Touch must also be developed and tested. A circuit for USB connection suggested by the specifications manual for the ARM microcontroller was drawn on the initial PCB, and will be tested. A more complex circuit may be necessary, especially if the final device will allow options for connecting to both the 30 pin docking device and USB-A. A potential circuit schematic used to connect USB-A to the ARM microcontroller may require one additional IC. Both configuration switch on a LED when connection between the ARM and host has been established. The LabVIEW VI must also be modified to allow communication between the ARM microcontroller and

the iPod Touch. Communication requirements include the microcontroller accepting information from user-supplied constants on the iPhone (such as time and temperature set-points), and the ARM transferring data to the iPod Touch. SubVIs for file I/O are supported through the LabVIEW embedded ARM module for accepting constants and writing data to a file off of the microcontroller. A means of programming the ARM microcontrollers directly on the customized PCB must also be established. One potential strategy is to include a circuit on the next version of the customized PCB for communication between the ARM and JTAG plug-in, which is used for programming the ARM microcontroller on the Keil evaluation board.

4.7. Design of Housing

Critical issues in development of component housing included arrangement of components around the amplification chip to reduce cross contamination of LED light between wells, and reproducible alignment between disposable amplification chip and LEDs/optical fibers. The cartridge for packaging the entire device also had to utilize a sliding top cover for removing/adding the chip that eliminates outside light and provides a latch for slight pressure, be astatically pleasing, allow plug-and-play capabilities with a iPod Touch to be used for initialization, data transfer, storage and analysis from the hand-held device.

Reproducible user alignment between the removable amplification chip and the LEDs/optical fibers was accomplished using a microstructered piece of ABS as an alignment guide for the reaction wells. Through-holes were drilled in the microstructured ABS piece for LEDs and optical fibers were permanently placed

within the ABS apparatus. Optical fibers (sending emitted light to the PD) contacted the sides of the shell structured amplification chip (Figure 4.18A shows 2-dimenstional side view, Figure 4.18D shows 3-dimensional isometric view). LEDs were placed beneath the chip to reduce the requirement of multiple wires being with the movable top cover when disposable amplification chips are loaded/unloaded. LEDs were also placed below the amplification chip and ABS holder to reduce scattered light from LEDs. A piece (44.6 x 41.1 x 15.9 mm) was designed and fabricated via rapid prototyping (black PC-ABS material, Stratasys, Inc.) to hold the ABS holder, amplification chip, secure optical fibers with a 5 mm bending radius, and mount LEDs (Figure 4.18B). A second piece (20.0 x 20.0 x 17.5 mm) was fabricated to house the PD and emission filter and securely hold all optical fiber ends around the active area of the photodiode (Figure 4.18C). The black PC-ABS was chosen for minimal light reflection, low transmission of light, ABS like properties that with high temperature stability similar to polycarbonate. The modular design of the holder allows components to be easily replaced if damaged, or changing LEDs and the filter to detect the amplification reactions at a different wavelength.

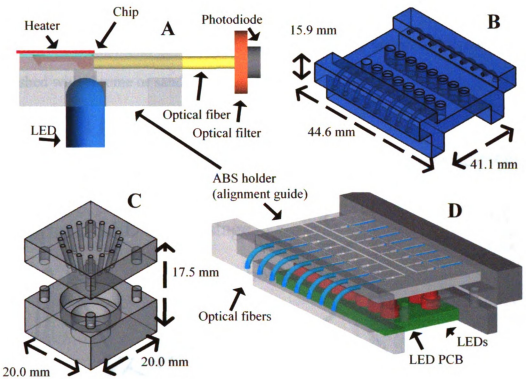


Figure 4.18 Component layout A) 2-dimensional view of components for the amplification chip, B) piece fabricated in PC-ABS to house the LEDs, optical fibers and ABS holder. C) piece fabricated to hold the PD, optical fibers. D) Three dimensional view of component layout for the amplification chip (PC-ABS piece is rendered transparent to view LED and fiber through holes in the structure.

A cartridge was designed to package all components into a small portable layout. The PCB was the largest component that had to be considered along with the foresight of a suitable battery. Attempts to decrease the layout with more efficient placement of traces and ICs would have been cumbersome. Therefore, the footprint required for the PCB was used as the base for an iPhone or iPod Touch (Figure 4.19). The top of the cartridge was designed to hold a majority of the components for detecting and heating the amplification reaction, and the base will serve to house the PCB and a battery. An of campus source was contracted to draft the cartridge due to the importance of designing it with a latch-able sliding cover (for inserting/removing the disposable amplification chip)

that does not allow ambient light into the holder (Vowin Model Design Co. Ltd, Schenzhen, China). Once designed, the cartridge will be fabricated out of ABS and finished with chrome or sandblasting.

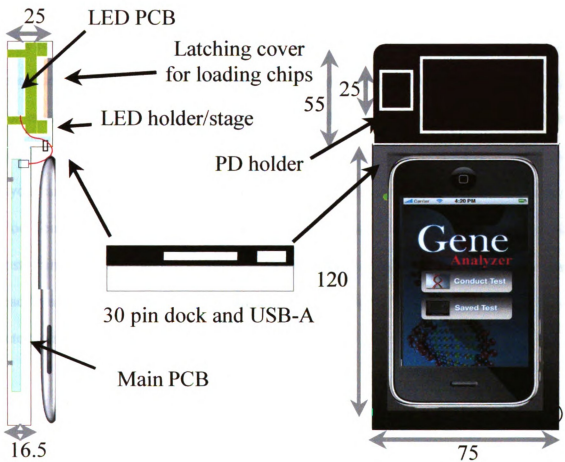


Figure 4.19 Schematic of the cartridge for packaging the LAMP based hand-held device (dimensions in mm). The back of the package, which encases the PCB, is used as a base for the iPod Touch and/or iPhone which is placed upside down and plugged into a male 30 pin docking port. The docking port connects the iPhone to the hand-held device for heating and real-time detection of the disposable amplification chip. The chip will be inserted/removed with a sliding self-latching door that will also serve to block all ambient light from entering the device during testing.

4.8 Concluding Remarks

A control setup for driving and detection loop mediated amplification was established and tested. Results showed that components and control setup allow precise and accurate temperature control for single set-point isothermal amplification and has the potential for rapid temperature cycling for quantitative PCR. Constant current LEDs, noise reduction strategies to the photodiode, and synchronized control strategy permitted reproducible signal measurements to be acquired in as little as 0.09 seconds per channel. For embedded control of components, a LabVIEW module was shown to efficiently program an ARM microcontroller. A custom PCB was used to replace a breadboard layout, and a cartridge has been designed for housing components and interfacing with an iPod Touch or iPhone. This control setup serves as a basic model for a multiassay, low-cost, miniaturized diagnostics system built with standard hardware components that are readily available and aids in the realization of a laboratory-independent tool for rapid detection of microbial pathogens.

4.9 References

1. **Easley, C. J., J. M. Karlinsey, J. M. Bienvenue, L. A. Legendre, M. G. Roper, S. H. Feldman, M. A. Hughes, E. L. Hewlett, T. J. Merkel, J. P. Ferrance, and J. P. Landers.** 2006. A fully integrated microfluidic genetic analysis system with sample-in–answer-out capability. Proc. Natl. Acad. Sci. U S A. **103:**19272-19277.

2. **Krazit, T.** 2006. ARMed for the living room, CNET News.

3. **Liao, C. S., G. B. Lee, H. S. Liu, T. M. Hsieh, and C. H. Luo.** 2005. Miniature RT–PCR system for diagnosis of RNA-based viruses Nucleic Acids Res. **33:**e156.

4. **Neuzil, P., J. Pipper, and T. M. Hsieh.** 2006. Disposable real-time microPCR device: lab-on-a-chip at a low cost. Mol. Biosyst. **2:**292-298.

5. **Pal, R., M. Yang, R. Lin, B. N. Johnson, N. Srivastava, S. Z. Razzacki, K. J. Chomistek, D. C. Heldsinger, R. M. Haque, V. M. Ugaz, P. K. Thwar, Z. Chen, K. Alfano, M. B. Yim, M. Krishnan, A. O. Fuller, R. G. Larson, D. T. Burke, and M. A. Burns.** 2005. An integrated microfluidic device for influenza and other genetic analyses. Lab Chip **5:**1024-1032.

6. **Roos, G.** 2008, posting date. National Instruments LabVIEW targets ARM microcontrollers. [Online.]

7. **Toriello, N. M., C. N. Liu, and R. A. Mathies.** 2006. Multichannel reverse transcription-polymerase chain reaction microdevice for rapid gene expression and biomarker analysis. Anal. Chem. **78:**7997-8003.

8. **Woolley, A. T., D. Hadley, P. Landre, A. J. deMello, R. A. Mathies, and M. A. Northrup.** 1996. Functional integration of PCR amplification and capillary electrophoresis in amicrofabricated DNA analysis device. Anal Chem. **68:**4081-4086.

CHAPTER FIVE

Validation of Optical System for Gene Analyzer

5.1 INTRODUCTION

Lab on a chip based formats are well suited for point of care screening of infectious diseases. Microbial agents are responsible for more than 75,000 deaths annually in the U.S. (24). A large part of this is due to the increase in antibiotic resistance among microbes over the past decade (18). One microbe of particular importance is methicillin-resistant *Staphylococcus aureus* (MRSA). Infections caused by this microbe have continuously increased in the U.S. from 108,600 in 1999 to 368,600 in 2005, and during the same period, MRSA-related deaths have increased from 11,000 to 19,000 (13). Current screening practices for MRSA consist of either culture-based assays or real time PCR (RT-PCR). Culture based assays take two to four days because specimens must be sent to a centralized laboratory and screened by culturing the specimen on MRSA media specific. RT-PCR assays are faster (e.g., BD GeneOhm™ Staph SR, made by BD Diagnostics and recently approved by FDA), taking about 2 hours. However, the cost of these machines (up to $30,000-$100,000) requires specimens to be sent to a centralized laboratory, delaying diagnosis, implementing the prescription of broad-spectrum antibiotics, and increasing potential for the emergence of a new more resistant MRSA strain. The availability of a portable and low cost device at a point-of-care, such as emergency first responders, healthcare centers, or physician's will aid in the realization of a more rapid means of identifying a suitable prescription of the correct antibiotic, avoiding the use of broad spectrum antibiotics.

Constraints in the development of truly portable and low cost nucleic acid analyzers are typically associated with the means of detection. The most common schemes for real time PCR amplification are based on fluorescence detection (6, 32), with components such as lasers and charge-coupled devices (CCDs). These apparatuses are considered to be too complex and large for miniaturization and integration into portable hand-held devices (14). Less costly components such as LEDs and photodiodes are promising alternatives, however, simultaneous detection of multiple channels with single point detectors typically requires movement or assigning independent detectors to each channel on the device (11). In contrast, an array of individually controlled LEDs can be used with optical fibers sending fluorescent signal to a single point photodiode detector. This mechanism for detection compliments loop mediated amplification (LAMP), as it produces more amplified product ~10 μg per 25 μl reaction (26) compared to conventional RT-PCR which produces approximately 200 ng per 25 μl reaction. Thus, a less sensitive, more compact, and cost efficient optical strategy can be utilized for real time fluorescent detection (20, 36). With this in mind, a device for fluorescence real time LAMP detection within a disposable microfluidic chip was developed (referred to as LAMP Analyzer). The setup uses an array of sixteen LEDs as an excitation source (one LED for each reaction well) and sixteen optical fibers to send fluorescent light from each well to a single photodiode sensor for detection. Only one LED/well/optical fiber is turned "on" at a time and identity is determined by the associated sequence and timing of the photodiode reading.

A critical component of the overall system is its low cost and user-sealable disposable microfluidic chip. Typical methods for hot embossing thin wall thermal plastic

structures consists of using a pair of matching dies which require very precise alignment of both embossing molds, and may lead to difficulties during ejection of the microchips and mold damage when alignment is not exact. To circumvent these issues, a low glass transition temperature thermal plastic (ABS) was used as a counter tool (12). This method provided numerous advantages over standard hot embossing for fabricating thin-walled structures, including the need for only a single embossing die, elimination of alignment issues, and ease of demolding as pressed chips could be deembossed from the counter tool instead of the mold itself. Thus, thin wall microfluidic chips were microstructured by hot embossing 100 μm Zeonor film (COP) with stereolithography molds. The process window for reproducible embossing of amplification chips was increased using a shoulder on the perimeter of the mold. The negative piece of microsturctured ABS was utilized further to: 1) aid in enclosing amplification chips under high pressure without damaging the shelled microstructures, 2) providing an alignment guide between optical components (optical fibers and LEDs) for replacing consumable amplification chips in the gene analyzer, and 3) reducing potential for crosstalk of light between adjacent reaction wells (17, 33).

Components for the gene analyzer were characterized by examining the reproducibility among reaction wells, and influence of light crosstalk between channels. Further validation included a comparison of quantitative capacity, sensitivity, and reproducibility between the gene analyzer and a standard laboratory real time PCR thermocycler. Real time LAMP characterization experiments were performed with calcein and an assay targeting the mecA gene of MRSA.

5.2 Materials and Methods

5.2.1 Disposable amplification chip fabrication and ABS holder: For embossing, male molds were drawn in Cobalt, a three dimensional computer aided design software (Ashlar-Vellum), and fabricated with 0.002 inch resolution out of Somos® NanoTool™ using stereolithography (FineLine Prototyping). NanoTool™ pieces underwent a thermal post-cure with UV to increase the glass transition temperature to 170 °C, and finished with nickel-plating (SLArmor) for increased strength and ease of de-embossing. For embossing, a 100 μm piece of ZeonorFilm (ZF14-100, Zeon Chemicals L.P) was sandwiched between the male mold and a piece of 2.2 mm thick low glass transition temperature plastic ABS (KS-8889, K-Mac Plastics), which was used as a soft counter mold (Figure 5.1 top panel). The embossing protocol consisted of preheating all materials without pressure to 150 °C for 3 min, followed by adding 2,000 lbs of pressure for 3 min using a Carver Hot Press (4386, Wabash). Subsequently, the press was cooled to 105 °C (well below the 136 °C glass transition temperature of COP) using manual cooling platens, while the microforming pressure was maintained. Pressure was released, and the formed polymer chips were deembossed from the formed ABS counter tool. The chips were cleaned by soaking in 1% Liqui-Nox (Alconox) for 5 min, thoroughly rinsed with distilled water for 1 min, rinsed in isoproponal, and air dried in a 65 °C oven.

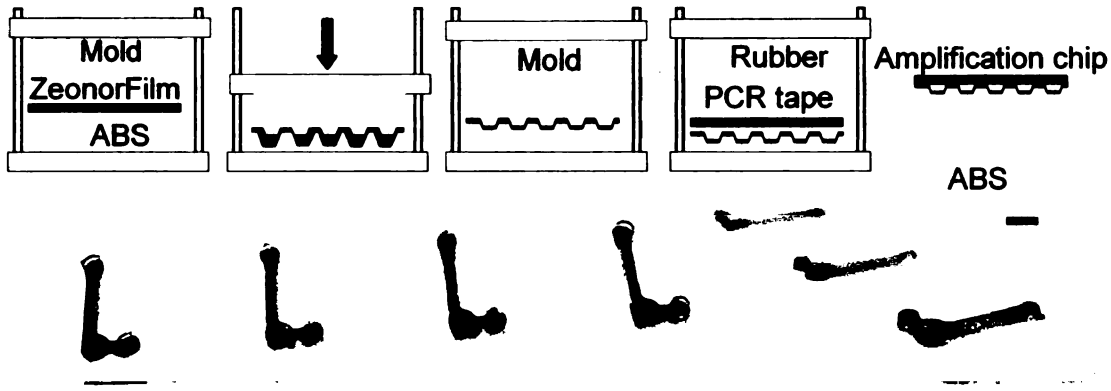


Figure 5.1 *Top panel*: Illustration of the low temperature plastic-assisted microforming and process for enclosing reaction wells. *Bottom panel*: Photograph of a microshelled 100-μm thick COP sheet containing 750-μm deep reaction wells enclosed with a top cover containing holes for adding samples and air ventilation.

To form enclosed features, the chip was bonded with 70-μm thick PCR tape (MicroAmp Optical Adhesive Film, Applied Biosystems). For introducing sample and air ventilation, circular holes were patterned into PCR tape using a Craft ROBO (CC33OL-20, 4). Patterned PCR tape was manually aligned onto the chips, and briefly pressed at 500 lbs (at room temperature) by sandwiching between a piece of silicon rubber (70 A Durometer Rating, McMaster-Carr) and the microstructured ABS. The silicon rubber was used to evenly distribute pressure over the full substrate area between the PCR tape and microstructured Zeonorfilm, and the ABS served to protect the shelled structure of the Zeonorfilm and maintain channel features. Samples were loaded into the channels, and a second piece of PCR tape was used to seal the sample ports.

5.2.2 Component setup: Components for heating the LAMP reaction and detecting amplification in real-time included a 2.54 x 7.62 cm Kapton heater (40 ohm, Electro-Flex Heat, Inc., Bloomfield, CT), 16 polymer optical fibers with a 1 mm core (IF C U1000, Industrial Fiber Optics, Tempe, AZ), a preamplified photodiode (ODA-6WB-500M, Opto

Diode Corp.), 16 blue 470 nm LEDs (RL3-B2030, Super Bright LEDs Inc), and a microstructured piece of ABS. The microsturctured piece of black ABS served as an alignment guide between the chip and the LEDs/optical fibers. For constructing the device, the microstructured piece of ABS was fastened to a piece of black nylon. One mm holes were drilled through the ABS directly underneath the reaction wells, and 3-mm holes were drilled into black nylon for mounting the LEDs. Reaction wells were individually illuminated by time staggered control of the LEDs (i.e., LEDs were turned "on" one at a time).

Polymer optical fibers, chosen due to their low cost and high numerical aperture, carried light from the reaction well to the PD. Optical fibers were butted against the sidewall of the microshelled reaction wells. An apparatus was designed and fabricated using stereolithography to receive all optical fibers with the same distance and angle from the active area of the photodiode. Different emission filters (colored glass longpass filters) were tested by placing in-between the optical fiber and the PD. The temperature of the chip was measured using a type T thermocouple with a 0.004 inch sheath diameter (5SRTC-TT-T-40-36, Omega) with a rapid response time of approximately 0.1 sec and a specified standard limit of error less than 1 ^{o}C. The control thermocouple was placed between the heater and the amplification chip. A more detailed description of component layout can be found in Chapter 4, Figure 4.18.

5.2.3 Bench-top control: Components were interfaced to a personal computer using a Multifunction USB Data Acquisition (USB-DAQ) card (6009, National Instruments, Inc.). A signal generating sub-virtual instrument (SubVI) was used to call a temperature

profile, which was controlled by LabVIEW software using proportional integrated derivative (PID). Voltage signals from the thermocouple were amplified and conditioned using the LabVIEW 9211 thermocouple input module. An analog output generated from the PID controller was used to adjust the duty cycle on a pulse width modulation (PWM) driver (DRV102T, Texas Instruments). The PWM driver acted as a gate between a power supply and the heater, providing fine control of power to the load without losses that occur when using resistive means of power management. A DC power supply set to 28 V provided a current of 0.5 A to initially heat the load, and an average current of 0.1 A was required to maintain isothermal temperature set-point for LAMP.

Individually controlled LEDs were turned on/off for obtaining signal with the PD of each reaction well. A 4-to-16 line decoder/demultiplexer (CD74HCT4514, Texas Instruments) was used to control the individual LEDs. The decoder is a high-speed silicon-gate that allows communication with each LED using only 5 digital outputs from the DAQ card. High logic output from the decoder signaled 1 of 16 LED drivers (STLA01) to provide constant current to the LEDs. The drivers were programmed to provide 26 mA with ± 5% accuracy. Prior to reading a signal with the PD, the LED was turned on for 1 s. The average of 100 sample readings taken during a 0.2 s time interval was recorded for the amplification curves. An analog switch was used to power down the photodiode between each cycle of sampling to circumvent issues related to overloading. The LabVIEW vi for both temperature control and optical detection are provided in Chapter 4.

5.2.4 Characterization of LAMP reaction using gene analyzer: Genomic DNA from methicillin-resistant *S. aureus* (ATCC 700699D-5) was obtained from the American Type Culture Collection (ATCC) and resuspended in diethylpyrocarbonate (DEPC)-treated and nuclease-free sterile water (Fischer Scientific, Pittsburgh, PA). The target was amplified using a set of four LAMP primers (F3-AAAAAACGAGTAGATGCTCAA, R3-TGGCCAATTCCACATTGT, FIP-TCCCAATCTAACTTCCACATACCATAAAACAAACTACGGTAACATTGA, BIP-CATAGCGTCATTATTCCAGGAATGCCGGTCTAAAATTTTACCACGT) along with a pair of loop primers (LF- TTTAACAAAATTAAATTGAACGTTGCGA, and LR-AGAAAGACCAAAGCATACATATTGAAAA). The primers were designed using Primer Explorer version 4 (http://primerexplorer.jp/e) from a consensus of 15 mecA alleles generated with Bioedit Sequence Alignment Editor (Ibis Biosciences, Carlsbad, CA). The specificity of the primers was ensured by checking against the GeneBank database using NCBI BLAST, and were ordered from Integrated DNA Technologies (IDT). LAMP reactions were performed in a total volume of 3 uL in the amplification chips and 12 µl in the benchtop Chromo4 Real-time PCR detector (Bio-Rad Laboratories), and consisted of 1.6 µM each of FIP and BIP primers, 0.2 µM each of F3 and B3 primers, 0.8 µM each of LF and LB primers, 0.8 M betaine (Sigma), 1.4 mM of each dNTPs (Invitrogen Corporation, Carlsbad, CA), 20 mM Tris/HCl (pH 8.8), 10 mM $(NH_4)_2SO_4$, 10 mM KCl, 8 mM $MgSO_4$, 8 mM Triton X-100, 0.64 units/uL of *Bst* DNA polymerase, large fragment (New England Biolabs Inc.) and 12.5 uM calcein and 0.25 mM $MnCl_2$ (half of concentration described (39)). Samples run with the Bio-Rad were loaded in 0.2 mL PCR tubes (VWR International) and incubated at 64°C for 36 min, and

fluorescence was measured every 45 s. Samples in the chip were incubated at 60 °C measured on the side of the chip furthest from the heater (a temperature of 70 °C was controlled using a thermocouple placed between the heater and amplification chip), and the PD was measured every 18 s for the 6 reaction experiment (dilution series experiments) and every 36 s for the 16 reaction experiments.

5.3 Results and Discussion

5.3.1 Microstructered amplification chip: Thin-wall chips were fabricated out of ZeonorFilm cyclic olefin polymer (COP) due to low water absorption, superior optical properties, inert surface, and no observable change in dimensions under high temperatures (2, 23, 43). Thin films were molded using hot embossing, because unlike injection molding, polymer flows a very short distance, resulting in parts with very little stress that are well suited for optical setups (16).

Hot embossing with molds fabricated using stereolithography provided a robust and suitable alternative compared to conventional stamping materials. Often, the embossing mold is made of a lithographic material such as silicon (8). However, this material is brittle resulting in a small process window for embossing microstructres with high-aspect ratios (44). Additional advantages using stereolithography fabricated molds include low cost and a rapid turnaround time (4 to 5 business days). Thermal expansion of nanotool caused slight bubbling in the layer of nickel after multiple uses, however, this did not influence hot embossing as the bubbles were pushed down during pressing. To our knowledge, this is the first manuscript describing the use of stereolithography fabricated molds for hot embossing microstructures (Figure 5.2).

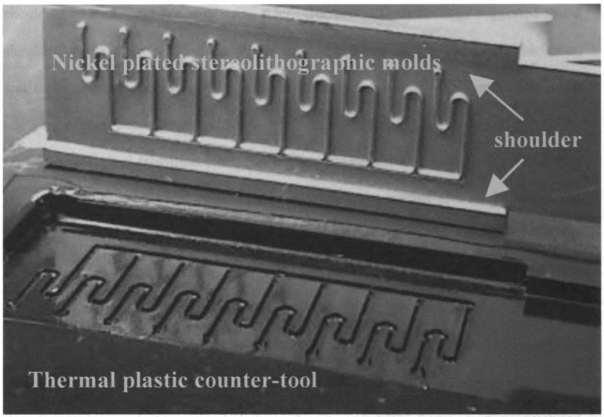


Figure 5.2 Picture of nickel-plated stereolithography mold pressed into a piece of low glass transition temperature thermoplastic (ABS) used as a counter-tool for microstructuring channels in ZeonorFilm COP. Shoulders were placed on the perimeter of the mold to increase the embossing process window.

A shoulder was placed on the perimeter of the mold to simplify the process of embossing (Figure 5.2). Specifically, the shoulder prevented lateral flow of the thermoplastic and allowed a more uniform buildup of pressure for pressing reproducible microstructures. The shoulder bordered two of four sides of the mold and was 100 um higher than the tallest features. The thickness of Zeonorfilm after pressing with molds was approximately 80 to 90 um where features had been pressed and was 100 um in areas without features. Pressing without a shoulder made the thickness of the film under the features irreproducible, often resulting in the mold pressing through both layers of ZeonorFilm and ABS. A study by Yao and coauthors (44) first described the use of

shoulders on molds for hot embossing microstructures. Their findings suggested that a shoulder on the mold reduced the dependence of pressure on the embossing thickness, and facilitated the attainment of a uniform pressure across the entire embossing area, allowing a larger processing window.

A low glass transition temperature thermoplastic (ABS) was used for thin film embossing of microstructures. The use of a soft support instead of a hard counter-tool has been described previously for simplifying embossing techniques by protecting the mold, allowing more uniform embossing pressure, and easier deembossing (12). Compared to rubber-assisted embossing techniques described previously (27), comparable features were only obtained when the surface area of the rubber counter mold was ~10 times larger than the mold itself. When counter molds were only slightly larger than the molds (in terms of spatial area), a remarkable difference in microstructured shells was observed (Figure 5.3). The less defined and irreproducible structures observed when using the smaller pieces of rubber may have been due to elongation of rubber reducing pressure and ability to obtain well defined shapes. However, elongation is reduced when the same amount of pressure is used on the larger pieces of rubber counter tool. In addition to obtaining well defined features, embossing with ABS had the added benefit of allowing microstructured COP to be deembossed from the negative ABS structure rather than the mold itself, preventing damage to the chip and eliminating the need to include small draft angles in the mold (5).

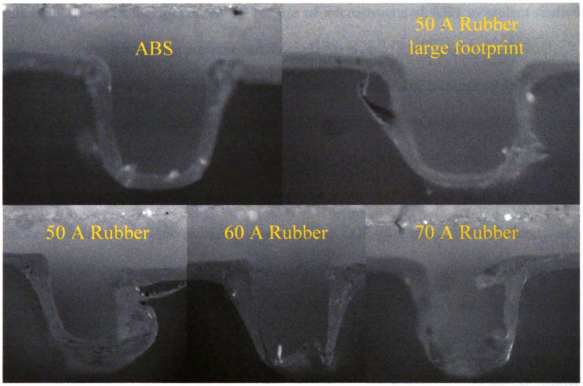


Figure 5.3 Cross sectional picture taken of 0.75 um wide microstructured channels pressed in 100 um ZeonorFilm with rubber and low glass transition temperature thermal plastic counter tools. All of chips (except 50 A rubber large footprint) were pressed using counter tools with spatial area slightly larger than the molds. For the 50 A rubber large footprint, a piece of rubber 10 x the spatial area of the mold was used as the counter tool. Chips were cut with scissors, taped to the side of a petridish and magnified 40 times using a fluorescent microscope to highlight shelled structure of COP.

The molded ABS tool was used further to enclose channels of the chip. The ABS served as a protective casing, allowing PCR tape to be pressed to the chip with high pressure without damaging or deforming shelled channel microfeatures. This circumvented the challenge of optimizing temperature and pressure to reduce deformation of microchannels for pressure driven bonding. Previous methods to overcome these issues have used costly programmable hot presses (9, 21, 28, 40), commercial bonding systems (1, 3), or laminators (15, 31). However, to our knowledge,

this is the first manuscript describing the use of microstructured ABS to protect shelled microstructures during high pressure bonding of adhesives.

A chip with 16 individual reaction channels (fluidically unconnected) was used for the characterization experiments so that the different sample mixtures (i.e. dyes, and concentration of gDNA) could be tested simultaneously. Microfluidic chips with a more simplistic mechanism for loading one sample into multiple reaction chambers are also under examination. This includes photografting COP to reduce hydrophobicity and increase capillary forces (22, 37). Strategies for eliminating cross contamination between reaction wells are also currently being examined including using valve systems with cover layers of soft PDMS (30, 35, 42), using a wax to eliminate primer dissolution and carryover during sample loading (19), and reducing cross contamination due to primer diffusion with increased lengths between reactions (7, 41). Alternative means of sealing reactions such as heat sealing and capping are also currently being explored. Previous bottlenecks for examinations of these strategies have been resolved with the establishment of a means to reproducibly fabricate chips.

5.3.2 Characterization of the gene analyzer: Prior to running comparative characterization experiments, the gene analyzer was tested to examine reproducibility between reaction wells on the same chip. For this, the same amount of *S. aureus* gDNA (1×10^4 genomic copies) was placed in all 16 reaction wells. While amplification was observed in all wells (Figure 5.5 left panel), the initial background signal varied from 0.05 to 0.82 V and the amplification signals was 1.19 to 1.87 times greater than the initial signal. To circumvent normalization issues caused by the variation in baseline signal, the

threshold time (Tt) was calculated as the inflection point (using the slope of 5 sequential data points) to determine when the change in fluorescence is at its greatest (Figure 5.4, (36). The variation in background is likely due to both a small difference in the distance between the capillary-end and side-wall of the amplification chip and the distance of the optical fiber from the active area of the photodiode. To reduce this variation, optical fiber glue (40-0005, Industrial Fiber Optics) will be used with subsequent prototypes of the device. In addition, multiple designs are being examined for placement of optical fibers with equal distance around the active area of the PD.

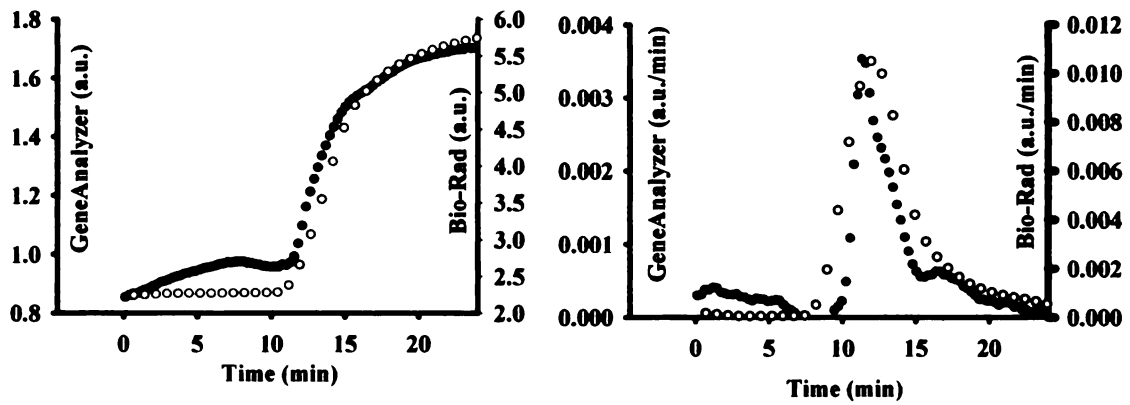


Figure 5.4 Difference in baseline signal between reaction wells and systems (left panel) prompted calculation of the inflection point (right panel) for determination of Tt, circumventing potential normalization bias.

Using the inflection point to determine Tt, the average Tt among 16 reaction wells was 12 min with a standard deviation of 0.94 min, and a coefficient of variation (CV) of 7.85%. This CV is high compared to studies described by Mori and coauthors in which a CV of 0.85% was observed using a turbidity meter for real time LAMP in 8 reaction wells (25). The CV obtained with the gene analyzer can be decreased with further optimization of reaction conditions in the low volume amplification chips. In addition, the 2 s LED pulse used to acquire signal between each reaction well may have

contributed to the high CV. Experiments (described in chapter 4) found that the scan rate could be reduced to 0.09 s per reaction well, which may serve to decrease CV in future experiments. No correlations were observed between the Tt and background signal, Tt and amplified signal over noise ratio, and Tt with respect to position of the reaction well on the chip.

Prior to running comparative studies, additional experiments were performed to ensure that the increase in signal was solely due to fluorescence and not turbidity generated due to precipitation of manganese pyrophosphate. Positive controls with and without calcein were tested with the hand-held components and results showed no observable increase in signal for the wells without calcein (Figure 5.5 right panel). Since the signal increase was caused solely by fluorescence, and there was no observable variation in Tt attributed to well position, further characterization experiments were conducted and compared to a standard laboratory real-time thermocycler.

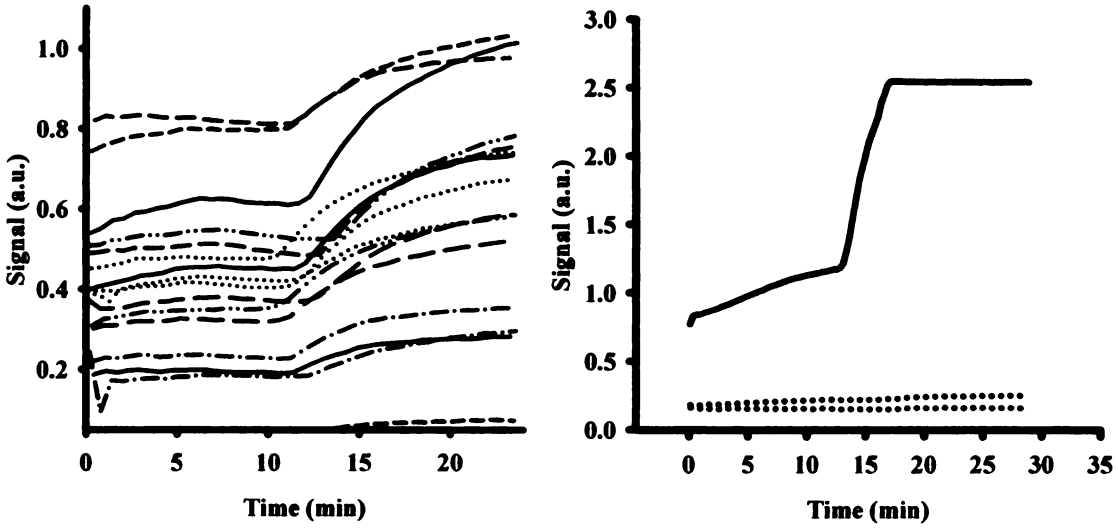


Figure 5.5 Left panel: LAMP in all sixteen-reaction wells spiked with 10,000 genomic copies of DNA. Right panel: positive control with calcein (solid line) and without calcein (dotted line) was examined to determine if turbidity influenced signal. The maximum input voltage (used to reduce noise in PD measurements) did not provide measurements above 2.5 V.

To examine and compare real time quantification and optical sensitivity of the hand-held device with a standard laboratory real-time thermocycler, a dilution series of 10 to 100,000 genomic copies of *S. aureus* per reaction well was prepared. The Tt was approximately 20 s lower with the gene analyzer for $1x\ 10^5$ and $1\ x\ 10^4$ genomic copies and 120 s higher with $1x10^3$ genomic copies compared to the laboratory real-time thermocycler (Figure 5.6 top left panel). A linear relationship with a correlation coefficient of $r^2 = 0.962$ was obtained with the gene analyzer (emission filter of 550 nm), and was $r^2 = 0.996$ with the Bio-Rad. Varying the emission filter with the gene analyzer had little influence on the correlation coefficient $r^2 = 0.999$ (570 nm) and $r^2 = 0.949$ (590 nm) (Figure 5.6 top right panel). As expected, the baseline signal was significantly reduced with the highest long pass emission filter (Figure 5.6 lower left panel), however, very little difference between amplified and baseline signal was observed between the 3 filters (Figure 5.6 lower right panel). Therefore, the 550 nm filter may be better suited for future experiments with the gene analyzer.

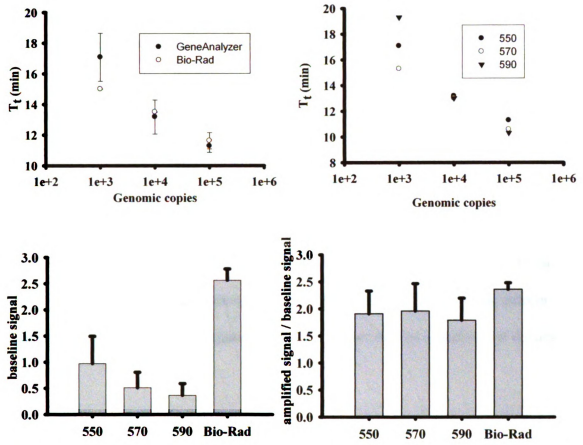


Figure 5.6 Top left panel: standard curves observed with the gene analyzer and standard laboratory real-time thermocycler (Bio-Rad) obtained using a dilution series with the same number of genomic copies per reaction well. Top right panel, standard curve obtained with 3 separate emission filters on the gene analyzer. Bottom left panel displays the baseline signal for all filters used on the gene analyzer and the Bio-Rad. Bottom right panel: the amplified signal divided by the baseline signal for all comparisons. Error bars represent standard deviation of three replicates.

The larger difference in Tt at lower concentration may have been due to interactions between the chip surface and biomolecules in the LAMP solution. This is primarily due to the increase of the surface-to-volume ratio and/or the PCR tape used to enclose the wells (45). Both a high concentration of BSA and Pluronic F68 were added to the amplification mixture to reduce the influence of the surface on the reaction (10, 38), however, the concentrations of this admixtures have yet to be optimized for use with LAMP or within COP chips. The low differences in Tt observed between the gene

analyzer and laboratory system suggest that the optical setup is well suited for quantitative detection of LAMP.

Amplification was not observed for the no template control and dilutions of 100 and 10 genomic copies per reaction in the laboratory thermocycler or gene analyzer. Therefore, the minimum amount of template DNA showing observable amplification for both systems was 1000 genomic copies. Prior studies have indicated that the LAMP reaction is capable of amplifying as little as six copies of template DNA (29), and studies examining the use of this primer set with intercalating dyes revealed a sensitivity of 100 copies (34). Further optimization of the reaction conditions such as calcein dye concentration and reaction temperature may serve to increase the sensitivity of the assay.

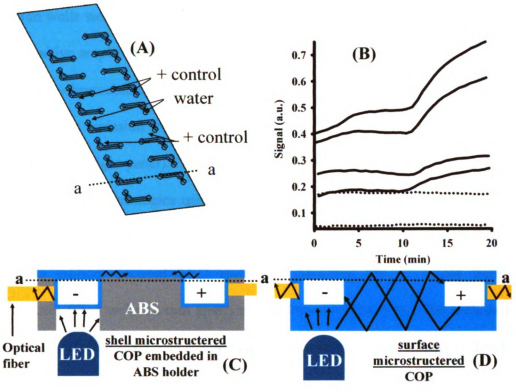


Figure 5.7 *Top left panel*: Computer aided draft of chip showing placement of samples for testing cross contamination of light, and cross sectional area used for schematics in lower panels. *Top right panels*; results of light scattering experiment for signal in positive control reaction wells (solid line) and signal in adjacent wells filled with water (dotted line). Lower left panel) cross sectional schematic of the chip illustrating cross contamination of light using the shell microstructered COP embedded in ABS holder versus the potential for light contamination in a surface microstructered amplification chip (Lower right panel).

Since COP is transparent and has a refractive index higher than that of air, it can behave as a waveguide for light, thus, experiments were also performed to investigate the fluorescence noise due to cross-talk between adjacent channels in the multi-channel amplification chip. The perceived mechanism for crosstalk in our setup was that fluorescent light emission from adjacent positive reaction wells (caused by scattering of light from LEDs) may influence signal when measurements are taken for an unamplified reaction well. To ensure that this was not influencing signal in the gene analyzer, two

reaction wells were filled with water and surrounding adjacent reaction wells were filled with positive amplification solutions (Figure 5.7 top left panel). Wells filled with water showed no observable increase in signal suggesting that crosstalk caused by the scattering of fluorescent emission from adjacent wells was not detectable with this optical setup (Figure 5.7 top right panel). Crosstalk of light between adjacent wells has been described for similar optical strategies using LEDs and optical fibers for detection in transparent chips (illustrated in Figure 5.7 lower right panel). More specifically, fluorescein filled channels without focused LED excitation displayed 5% of signal from adjacent fluorescein filled wells (2 mm apart) with a focused LED (17). The authors found the majority of the crosstalk was due to fluorescent emission (not scattered LED light from adjacent channels. Increasing distance between reaction channels and/or placement of a black piece of rubber between channels reduced observable cross contamination of light. In our setup, the ABS holder complemented the shelled chips, serving as a block between wells, thus reducing cross contamination of light (Figure 5.7 lower left panel).

*5.3.3 **Refining the gene analyzer for RT-PCR.*** A fluorescence means of detection was chosen (instead of turbidity) for future refinement of the gene analyzer for quantitative PCR. This would provide the option for detection of mutation. Issues related to thermal mass of the amplification chip have been circumvented through development of thin walled microfluidic chips with low mass. The heater control setup has also demonstrated accurate/precise thermocycling (Chapter 4, Figure 4.6). Therefore, two obstacles must be overcome for quantitative PCR with the gene analyzer including increasing sensitivity (with a potentially more complex and costly optical strategy) and overcoming issues with

thermal gradients between the heater, layers of the amplification chip, and the holder. Using a thin piece of black anodized aluminum, instead of microstructered ABS, may be a suitable means for reducing thermal gradients. The inability to obtain amplification profiles when testing green and blue intercalating dyes with the gene analyzer (instead of calcien) demonstrates the need for a more advanced series of excitation and emission filters, and perhaps a more sophistical means of transferring between consumable chips and optical fibers and the PD and optical fibers. Alternative means of increasing optical sensitivity may include increasing the emission intensity of the reaction by supplying a higher current to the LEDs. A study by Ren and coauthors observed a ten-fold increase in signal intensity when current to LEDs was increased from 20 to 200 mA (33).

A dilution series of fluorescein was prepared to compare optical sensitivity of the gene analyzer with the laboratory thermocycler. Results show a similar response in signal with an increasing amount of fluorescein between the two devices, with the Bio-Rad having a slightly increased response with higher concentration (Figure 5.9). A linear response in signal ($R^2 \geq 0.97$) with the concentration gradient of fluorescein (0.5 to 10 μM) further demonstrates the quantitative utility of the gene analyzer (Figure 5.8). A lower correlation ($R^2 = 0.95$) has been described for a hand-held device with similar optical components (36). With a response that is similar to the real-time laboratory thermocycler, the optical performance of the gene analyzer may only require slight modifications before it can be used for RT-PCR diagnostics.

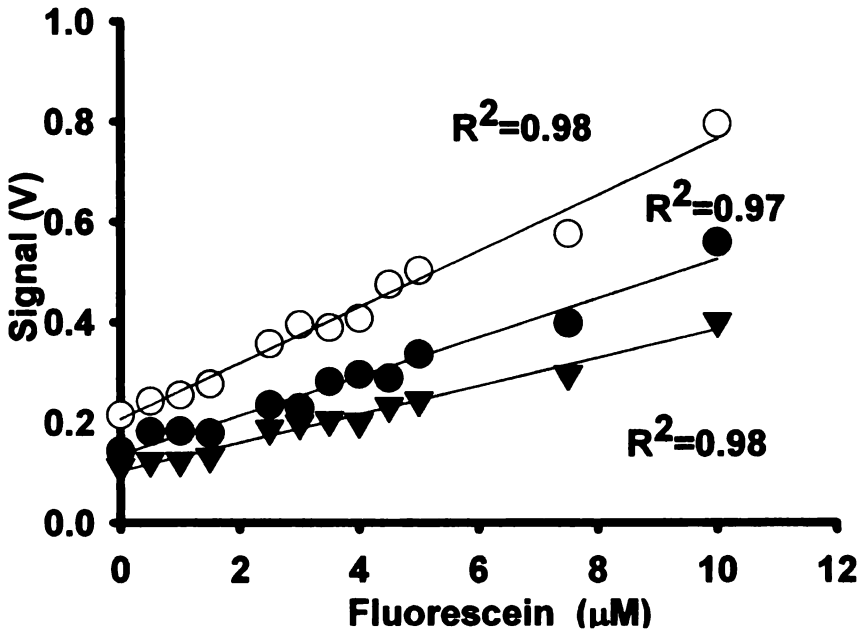


Figure 5.8 Dilution series of fluorescein with the gene analyzer, change in signal with varying concentrations of fluorescein is show for three separate reaction wells.

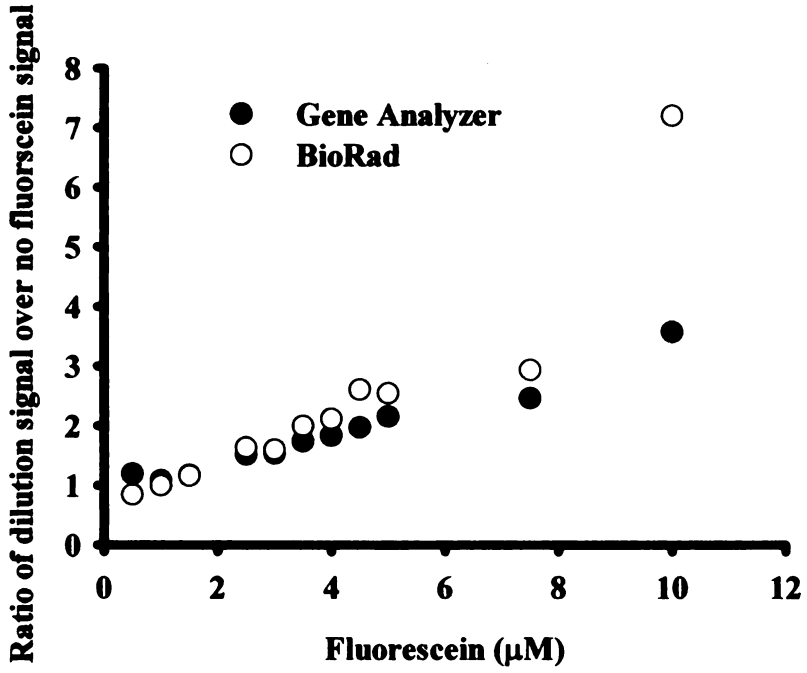


Figure 5.9 Ratio of dilution signal over signal of wells without fluorescein (water) tested with Bio-Rad (solid dots) and gene analyzer (open dots) shows similar response to increasing amount of fluorescein. This data represents the average ratio of 3 different reaction wells.

5.4 CONCLUSION

A device for real-time fluorescent-based detection of loop mediated amplification (LAMP gene analyzer) was developed with a simple optical setup that is capable of measuring 16 assays simultaneously. Strategies for hot embossing such as the use of stereolithographic molds with shoulders and low glass transition temperature counter tools were used to increase the process window for reproducible fabrication of amplification chips with well-defined microstructures. The simple and low cost optical setup consisting of individually pulsed LEDs, optical fibers for sending fluorescent signal, and a single photodiode provided comparable real time results with a laboratory thermocycler for calcein based LAMP detection. The small and low cost components assembled for the device make it a promising venture in the development of a nucleic acid analyzer that is portable, sensitive, quantitative, targets multiple assays simultaneously, and is cost efficient.

5.5 REFERENCES

1. **Abgrall, P., L. N. Low, and N. T. Nguyen.** 2007. Fabrication of planar nanofluidic channels in a thermoplastic by hot-embossing and thermal bonding. Lab Chip **7**:520-522.

2. **Ahn, C. H., J. W. Choi, G. Beaucage, J. H. Nevin, J. B. Lee, A. Puntambekar, and J. Y. Lee.** 2004. Disposable smart lab on a chip for point-of-care clinical diagnostics. Proceedings of the IEEE **92**:154-173.

3. **Arroyo, M. T., L. J. Fernandez, M. Agirregabiria, N. Ibanez, J. Aurrekoetxea, and F. J. Blanco.** 2007. Novel all-polymer microfluidic devices monolithically integrated within metallic electrodes for SDSCGE of proteins. J. Micromech. Microeng. **17**:1289-1298.

4. **Bartholomeusz, D. A., R. W. Boutté, and J. D. Andrade.** 2005. Xurography: rapid prototyping of microstructures using a cutting plotter. J. Micromech. Syst. **14**:1364-1374

5. **Becker, H., and U. Heim.** 2000. Hot embossing as a method for the fabrication of polymer high aspect ratio structures. Sensors and actuators. A: Physical **83**:130-135.

6. **Belgrader, P., S. Young, B. Yuan, M. Primeau, L. A. Christel, F. Pourahmadi, and M. A. Northrup.** 2001. A battery-powered notebook thermal cycler for rapid multiplex real-time PCR analysis. Anal Chem. **73**:286–289.

7. **Berthier, J., and F. Chatelain.** 2005. Presented at the Multiphysics User's Conference, Paris.

8. **Bhattacharyya, A., and C. M. Klapperich.** 2007. Design and testing of a disposable microfluidic chemiluminescent immunoassay for disease biomarkers in human serum samples. Biomed. Microdevices. **9**:245-251.

9. **Buch, J. S., C. Kimball, F. Rosenberger, W. E. Highsmith Jr., D. L. DeVoe, and C. S. Lee.** 2004. DNA mutation detection in a polymer microfluidic network using temperature gradient gel electrophoresis. Anal Chem. **76**:874-881.

10. **Cho, Y. K., J. Kim, Y. Lee, Y. A. Kim, K. Namkoong, H. Lim, K. W. Oh, S. Kim, J. Han, C. Park, Y. E. Pak, C. S. Ki, J. R. Choi, H. K. Myeong, and C. Ko.** 2006. Clinical evaluation of micro-scale chip-based PCR system for rapid detection of hepatitis B virus. Biosens. Bioelectron. **21**:2161-2169.

11. **Dishinger, J. F., and R. T. Kennedy.** 2008. Multiplexed detection and applications for separations on parallel microchips. Electrophoresis **29:**3296–3305.

12. **Dreuth, H., and C. Heiden.** 1999. Thermoplastic structuring of thin polymer films. Sensors and Actuators A: Physical **78:**198-204.

13. **Elixhauser, A., and C. Steiner.** 2007. Infections with methicillin-resistant *Staphylococcus aureus* (MRSA) in U.S. hospitals, 1993–2005. Healthcare Cost and Utilization Project Statistical Brief #35, Agency for Healthcare Research and Quality, Rockville, MD. .

14. **Fang, T. H., N. Ramalingam, D. Xian-Dui, T. S. Ngin, Z. Xianting, A. T. Lai Kuan, P. H. E.Y., G. Hai-Qing, and** 2009. Real-time PCR microfluidic devices with concurrent electrochemical detection. Biosens. Bioelectron. **24:**2131-2136.

15. **Fredrickson, C. K., Z. Xia, C. Das, R. Ferguson, F. T. Tavares, and Z. H. Fan.** 2006. Effects of fabrication process parameters on the properties of cyclic olefin copolymer microfluidic devices. J. Micromech. Syst. **15:**1060-1068.

16. **Heckele, M., and W. K. Schomburg.** 2004. Review on micro molding of thermoplastic polymers. J. Micromech. Microeng. **14:**R1-R14.

17. **Irawan, R., S. Chuan Tjin, P. Yager, and D. Zhang.** 2005. Cross-Talk problem on a fluorescence multi-channel microfluidic chip system. Biomed. Microdevices. **7:**205–211.

18. **Jones, R. N.** 2003. Global epidemiology of antimicrobial resistance among communityacquired and nosocomial pathogens: A five-year summary from the SENTRY Antimicrobial Surveillance Program (1997-2001). Semin. Respir. Crit. Care Med. **24:**121-134.

19. **Kim, J., D. Byun, M. G. Mauk, and H. H. Bau.** 2009. A disposable, self-contained PCR chip. Lab Chip **9:**606-612.

20. **Lee, S. Y., J. G. Huang, T. L. Chuang, J. C. Sheu, Y. K. Chuang, M. Holl, D. R. Meldrum, C. N. Lee, and C. W. Lin.** 2008. Compact optical diagnostic device for isothermal nucleic acids amplification. Sensors and Actuators B: Chemical **133:**493-501.

21. **Li, Y., J. S. Buch, F. Rosenberger, D. L. DeVoe, and C. S. Lee.** 2004. Integration of isoelectric focusing with parallel sodium dodecyl sulfate gel

electrophoresis for multidimensional protein separations in a plastic microfludic network. Anal. Chem. **76:**742-748.

22. **Lin, R., and M. A. Burns.** 2005. Surface-modified polyolefin microfluidic devices for liquid handling. J. Micromech. Microeng. **15** 2156–2162.

23. **Mela, P., A. van den Berg, Y. Fintschenko, E. B. Cummings, B. A. Simmons, and B. J. Kirby.** 2005. The zeta potential of cyclo-olefin polymer microchannels and its effects on insulative (electrodeless) dielectrophoresis particle trapping devices. Electrophoresis **26:**1792-1799.

24. **Mokdad, A. H., J. S. Marks, D. F. Stroup, and J. L. Gerberding.** 2000. Actual causes of death in the United States. JAMA **291:**1238-1245.

25. **Mori, Y., M. Kitao, N. Tomita, and T. Notomi.** 2004. Real-time turbidimetry of LAMP reaction for quantifying template DNA. J. Biochem. Biophys. Methods. **59:**145-57.

26. **Mori, Y., K. Nagamine, N. Tomita, and T. Notomi.** 2001. Detection of loop-mediated isothermal amplification reaction by turbidity derived from magnesium pyrophosphate formation. Biochem. Biophys. Res. Commun. **289:**150-4.

27. **Nagarajan, P., and D. Yao.** 2009. Rubber-assisted micro forming of polymer thin films. Microsyst. Technol. **15:**251-257.

28. **Nie, Z., and Y. S. Fung.** 2008. Microchip capillary electrophoresis for frontal analysis of free bilirubin and study of its interaction with human serum albumin. Electrophoresis **29:**1924-1931.

29. **Notomi, T., H. Okayama, H. Masubuchi, T. Yonekawa, K. Watanabe, N. Amino, and T. Hase.** 2000. Loop-mediated isothermal amplification of DNA. Nucleic Acids Res. **28:**e63.

30. **Oh, K. W., C. Park, K. Namkoong, J. Kim, K. S. Ock, S. Kim, Y. A. Kim, Y. K. Cho, and C. Ko.** 2005. World-to-chip microfluidic interface with built-in valves for multichamber chip-based PCR assays. Lab Chip **5:**845-850.

31. **Olivero, D., and Z. H. Fan.** 2008. Chips & tips: lamination of plastic microfluidic devices, Lab Chip.

32. **Ottesen, E. A., J. W. Hong, S. R. Quake, and J. R. Leadbetter.** 2006. Microfluidic digital PCR enables multigene analysis of individual environmental bacteria. Science **314:**1464–1467.

33. **Ren, K., Q. Liang, X. Mu, G. Luo, and Y. Wang.** 2009. Miniaturized high throughput detection system for capillary array electrophoresis on chip with integrated light emitting diode array as addressed ring-shaped light source. Lab Chip **9:**733-736.

34. **Seyrig, G., R. D. Stedtfeld, T. M. Stedtfeld, D. M. Tourlousse, J. M. Tiedje, S. A. Hashsham, and** 2009. Accelerated real-time loop mediated isothermal amplification with optimized fluorescent dyes and protocol for fast (less than 10 minutes) detection of nucleic acids. . For submission to Appl. Environ. Micrbiol.

35. **Shin, Y. S., K. Cho, S. H. Lim, S. Chung, S. J. Park, C. Chung, D. C. Han, and J. K. Chang.** 2003. PDMS-based micro PCR chip with Parylene coating. J. Micromech. Microeng. **13:**768-774.

36. **Smith, M. C., G. Steimle, S. Ivanov, M. Holly, and D. P. Fries.** 2007. An integrated portable hand-held analyser for real-time isothermal nucleic acid amplification. Anal. Chim. Acta **598:**286-294.

37. **Stachowiak, T. B., D. A. Mair, T. G. Holden, L. J. Lee, F. Svec, and M. J. Jean.** 2007. Hydrophilic surface modification of cyclic olefin copolymer microfluidic chips using sequential photografting. J. Sep. Sci. **30:**1088-1093.

38. **Stedtfeld, R. D., S. W. Baushke, D. M. Tourlousse, S. M. Mller, T. M. Stedtfeld, E. Gulari, J. M. Tiedje, and S. A. Hashsham.** 2008. Development and experimental validation of a predictive threshold cycle equation for quantification of virulence and marker genes by high-throughput nanoliter-volume PCR on the OpenArray platform. Appl. Environ. Microbiol. **74:**3831-8.

39. **Tomita, N., Y. Mori, H. Kanda, and T. Notomi.** 2008. Loop-mediated isothermal amplification (LAMP) of gene sequences and simple visual detection of products. Nat. Protoc. **3:**877-82.

40. **Tsao, C. W., L. Hromada, J. Liu, P. Kumar, and D. L. DeVoe.** 2007. Low temperature bonding of PMMA and COC microfluidic substrates using UV/ozone surface treatment. Lab Chip **7:**499-505.

41. **Wang, F., M. Yang, and M. Burns.** 2008. Microfabricated valveless devices for thermal bioreactions based on diffusion-limited evaporation. Lab Chip **8:**88-97.

42. **Warren, L., D. Bryder, I. L. Weissman, and S. R. Quake.** 2006. Transcription factor profiling in individual hematopoietic progenitors by digital RT-PCR. Proc. Natl. Acad. Sci. U S A. **103:**17807-17812.

43. **Yang, Y., C. Li, J. Kameoka, K. Lee, and H. G. Craighead.** 2005. A polymeric microchip with integrated tips and in situ polymerized monolith for electrospray mass spectrometry. Lab Chip **5:**869-876.

44. **Yao, D., and R. Kuduva-Raman-Thanumoorthy.** 2008. An enlarged process window for hot embossing. J. Micromech. Microeng. **18:**7.

45. **Zhang, C., and D. Xing.** 2007. Miniaturized PCR chips for nucleic acid amplification and analysis: latest advances and future trends. Nucleic Acids Res. **35:**4223-4237.

CHAPTER 6

ENGINEERING SIGNIFICANCE AND FUTURE PERSPECTIVES

6.1 ENGINEERING SIGNIFICANCE

The potential for microbial diagnostics technologies to impact health care in developing regions of the world is well recognized. However, there is a large gap in the use of such technologies due to the unavailability of a low cost genetic screening system that is point of care. The device developed within this project fills this gap, with the potential for providing genetic screening at a cost that is well within reach of resource-limited regions. This has been obtained through: 1) establishment of tools for efficient design and validation of high-throughput assays, 2) development and utilization of a low cost and user-sealable disposable microfluidic chip 3) integration of a simple and low cost optical setup that eliminates the complexity and size requirements associated with real time detection of multiple assays, and 4) acquisition of expertise for embedded control of components within a footprint that is truly portable. Overall, the gene analyzer has a total component cost of approximately $599 (Table 6.1). The use of a portable sample preperation device (1) to complement the gene analyzer will satisfy the the main objective of this project for providing a device with the potential to globally impact human health.

Table 6.1 Component costs for Gene Analyzer prototype with iPod Touch and Software

Component (number)	Low volume order cost
Optical fibers (16)	$10.00
Kapton heater	$35.00
Temperature sensor	$30.00
Photodiode	$64.00
High volume (100 piece) PCB order	$10.00
ARM microcontroller	$10.00
PWM driver IC	$8.00
Temperature conditioner IC	$19.00
4 to 16 line Decoder IC	$2.00
LED drivers (16)	$5.00
LEDs (16)	$11.00
Optical filter	$15.00
Ribbon cables, plug-in	$7.00
Voltage regulators (5)	$20.00
General resistor/capacitors	$10.00
Hardware for ARM-iPod connection	$15.00
Battery	$47.00
Housing (mass production)	$2.00
Total Device Cost	$320.00
iPod Touch	$229.00
GeneZ Software	$50.00
Grand Total	**$599.00**

Several novel strategies have been implemented toward the development of the device. In detail, a predictive equation for quantification of stating copies was experimentally established based on the sequence characteristics of primers and target organisms as an efficient alternative to generating standard curves. Development of the gene analyzer required the assembly of components that were not only small, and low cost, but provided the same level of sensitivity and quantitative capacity of a standard laboratory based RT-PCR device. The assembly of LEDs and photodiodes has been used previously for fluorescent detection, however, the implementation of this detection scheme with optical fibers for single point detection of multiple reactions without moving parts, along with the integration of a precisely controlled heater for genetic assay based amplification make the component setup and mechanism for utilization novel. To our

knowledge, the strategies used to reproducibly fabricate microfluidic chips including: use of stereolithography molds along with combined use of shoulders on positive molds and ABS counter tools to increase embossing process windows, and using microsturcutred ABS to protect channels of microfluidic chips during enclosure are also novel developments. One of the more critical aspects in development of Gene Analyzer is the electronics hidden within. Thus, the acquired expertise and collection of electronic component hardware such as ICs, LabVIEW embedded module for generating C++ code, and the library of developed LabVIEW software provides the laboratory with a distinct advantage in efficient development of future devices.

The advent of a prototype for the gene analyzer brings several more steps that must be implemented prior to commercialization including: further evaluation of the prototype using field trials, transfer of developments (gene analyzer and disposable chips) into manufacturing, improvement of analysis software, and development of after-sale services. Since the device will be low cost, the disposable and service based aspects will be critical aspects of commercialization. The process of transferring a prototype into a manufactured unit will slow availability as pilot builds, manufacturing process validation, regulatory submissions, and manufacturing scale-up can take 12 to 36 months. Once completed, the gene analyzer can be complemented with a sample preparation method to allow traditional laboratory analysis to be decentralized and taken into the field. Not only will such a device influence human health in clinical settings, but also in environmental, agricultural, and bio-defense applications.

6.2 Future Perspectives Toward Finalizing The Gene Analyzer Prototype

A few control related tasks remain before the field ready prototype of the LAMP based gene analyzer is complete. This includes testing ICs selected for power and voltage management and obtaining a suitable battery. Three separate ICs have been selected to convert 12 V DC power to the following: 3.3 V 20 mA power supplies for the ARM microcontroller (LT3014), 28 V and 700 mA max current for the heater (LT1129), and 5 V 50 mA max to the LED drivers (LT3010-5). For obtaining required voltages, all regulators can be programmed by replacing two resistors in the circuit. Pads for the power regulating devices were placed on the initial design of the PCB for testing.

A second remaining task is to program the ARM microcontroller with code for simultaneous control of all components, which may require optimization of the code compiler to reduce memory requirements. Due to the limited amount of memory that could be accessed using the evaluation kit of the LabVIEW embedded module (only 25% of 512 KB memory), the original LabVIEW VIs used for bench-top control had to be stripped to a very basic form. However, code for all tested VIs was generated using default debugging and compiling parameters, which can be changed to require less space, but also may run less efficiently.

Development of GeneZ (Objective C-based software) to be installed on the iPod Touch is the third and final component. A conceptual framework for this software is under development at Michigan State University with some help from an Apple software developer (Figure 6.1). The software is rather simple but extremely versatile. As illustrated in Figure 6.1, user specified functions include: adding any test (corresponding to a new chip), or changing the name of the gene (the gene list window is not shown in

detail but it is possible to change the identity of the gene), options for an isothermal or temperature cycling program, visualizing the results, calculating averages of identified wells and reporting the final result of the overall test. Because users can pick/download the list of genes for any given assay, the software is expected to be an ideal companion to the gene analyzer. When completed, this software will be distributed through the Apple Store with revisions provided through the same channel as necessary.

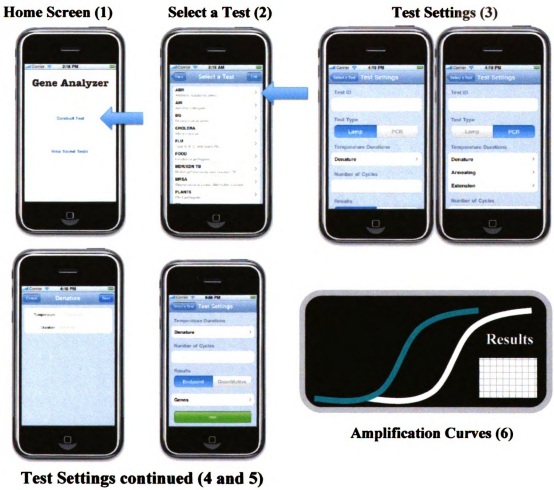


Figure 6.1 Conceptual representation of the screens for GeneZ software to be used to a) control the Gene Analyzer, b) visualize the amplification curves and c) report the final results.

6.3 REFERENCES

1. **Casper, E.T., Patterson, S.S., Bhanushali, P., Farmer, A., Smith, M., Fries, D.P., Paul, J.H.** A handheld NASBA analyzer for the field detection and quantification of *Karenia brevis.* 2007. Harmful Algae. **6**: 112.