11-13-2015

MOL-PCR For SNP Based Detection and Characterization of STEC non-O157 Strains

Heather M. Mendez

Heather M. Mendez

Candidate

Biomedical Engineering

Department

This thesis is approved, and it is acceptable in quality and form for publication:

Approved by the Thesis Committee:

Dr. Alina Deshpande, Co-Chairperson

Dr. Steven W. Graves, Co-Chairperson

Dr. Andrew Shreve

MOL-PCR FOR SNP BASED DETECTION AND CHARACTERIZATION OF STEC NON-O157 STRAINS

by

HEATHER M. MENDEZ

B.S., BIOLOGY UNIVERSITY OF NEW MEXICO, 2010

THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Master of Science
Biomedical Engineering

The University of New Mexico
Albuquerque, New Mexico

December, 2015

ACKNOWLEDGMENTS

The work presented in this thesis is in part due to the guidance and inspiration from many individuals. I would like to acknowledge the support from my Co-advisors Dr. Alina Deshpande and Dr. Steven Graves. I am extremely grateful for the tremendous amount of support and guidance that both of you have provided throughout the past two years. It is with your extensive knowledge and patience that have contributed to the success of this project. I am also very thankful to Dr. Andrew Shreve for serving on my defense committee.

I am very grateful to my colleague and mentor Travis Woods. I would like to thank Travis for designing this project in addition to his continuous encouragement and invaluable advice. Thank you for your time and guidance on this project. I would also like to thank the individuals in the Graves lab that provided feedback and support throughout the past two years. Additionally, thank you to Adan Myers Y Gutierrez and Aurora Fabry-Wood for providing feedback on my thesis.

Finally, I would like to thank my family and friends. I would like to thank my husband, Michael, for his love, support, and patience throughout my graduate career. In addition, I would like to thank my parents and my sisters for their love and words of encouragement.

MOL-PCR for SNP Based Detection and Characterization

of STEC non-O157 Strains

by

Heather M. Mendez

B.S., Biology, University of New Mexico, 2010
M.S., Biomedical Engineering, University of New Mexico, 2015

ABSTRACT

Shiga toxin-producing *Escherichia* coli (STEC) strains are important food-borne pathogens that have been linked to human illness. In the United States, STEC O157:H7 strains have been attributed to the majority of disease and outbreaks; however, the frequency of disease associated with STEC non-O157 strains have also increased and is often under recognized due to diagnostic limitations. In 2012, the Food Safety and Inspection Service (FSIS) implemented mandatory testing for non-O157 STEC serogroups O26, O45, O103, O111, O121, and O145. Single nucleotide polymorphisms (SNPs) have been identified in the O-antigen gene cluster for each of the non-O157 STEC serogroups that can differentiate between a STEC and non-STEC strain. In this study, the multiplex oligonucleotide ligation-PCR (MOL-PCR) detection platform was adapted for the detection and characterization of the above-described non-O157 STEC serogroups. The SNP-based MOL-PCR assay has demonstrated the ability to simultaneously detect the non-O157 serogroups and characterize the strain based on the SNP discriminates identified for each serogroup. This work provides a potential diagnostic platform that can reduce diagnostic limitations for the detection of non-O157 serogroups.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1: Introduction

Chapter 3: Materials and Methods

Chapter 5: Conclusions and Future Directions

.LIST OF TABLES

Chapter 3: Materials and Methods

Chapter 4: Results and Discussion

Chapter 5: Conclusion and Future Directions

Chapter 1: Introduction

1.1 Shiga Toxin Family

1.1.1 Structure of Shiga Toxins

The Shiga toxin family is a group of bacterial protein toxins that are comprised of *Shigella dysenteriae* (*S.dysenteriae*) serotype 1 and Shiga toxins produced by *Escherichia coli* (*E.coli*), otherwise known as Shiga toxin-producing *E.coli* (STEC). [1-2] The structure of the Shiga toxin (*stx*) protein consists of an AB_5 molecular configuration, with an A and B-subunit that are conserved across the Shiga toxin family.[3-4] The A-subunit is a 32-kDa protein that contains two cysteine residues (Cys-242 and Cys-261) that are covalently linked through a disulfide bond.[5-6] Through proteolytic cleavage by trypsin or furin, the A-subunit is divided into two components, A_1 (27.5 kDa) and A_2 (4.5 kDa) fragments, with the fragments remaining attached via a disulfide bridge.[3-8] The A_1 fragment contains the *N*-glycosidase catalytic site that is associated with inhibition of protein synthesis.[7] The A_2 fragment includes the carboxyl-terminal tail that functions to bind the A-subunit to the B-subunit, resulting in the formation of an α-helix that is enclosed by the B-subunit, as illustrated in Figure 1.1.[5, 8]The B-subunit is a pentameric protein that consists of five identical B-subunits (7.7-kDA) that are responsible for attaching to host cellular receptors (Figure 1.1).[8-9] The *stx* B-subunits bind to the receptor glycosphingolipid globotriaosylceramide (Gb3) located on the cellular surface.[9]

The molecular structure for each B-subunit is comprised of three Gb3 binding sites (carbohydrate binding sites), resulting in an *stx* B-subunit that can potentially bind to 15 Gb3 binding sites.[9] As a result, the potential *stx* B-subunit binding sites leads to high-affinity binding, with an approximate binding affinity of 10^9 M^{-1} (10^3 M^{-1} per Gb3 binding

site).[10] With the molecular configuration of the B-subunit, the conformation of the B-subunit allows for high-affinity binding that can be related to increased cellular toxicity.

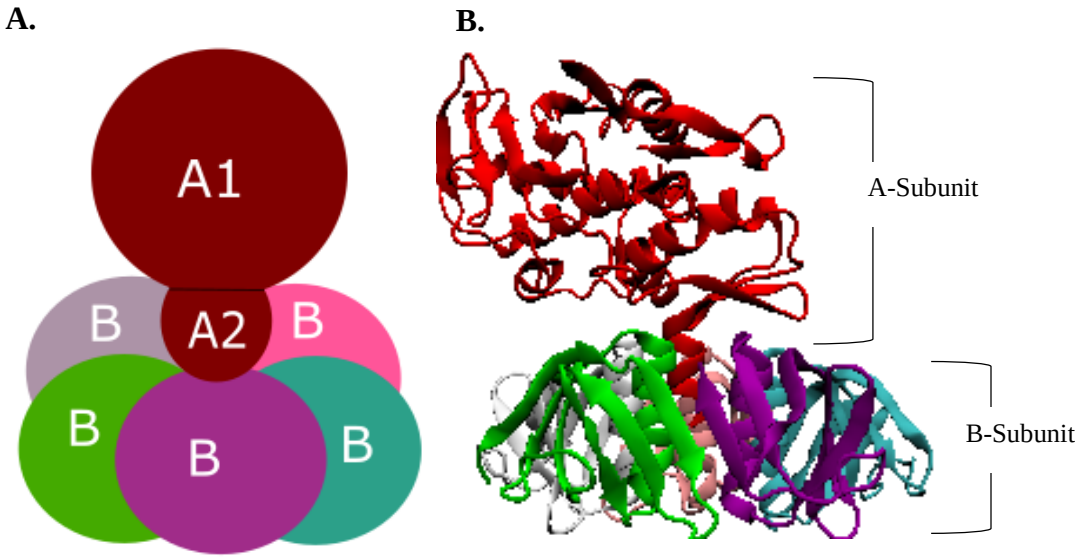


Figure 1.1: A) Graphical representation of the Shiga toxin structure. The A-subunit is shown in red and represents the A1 and A2 fragments. The B-subunit consists of a 5-pentameric ring. B) The structure of *stx* represented in a ribbon diagram.[8] Figure adapted from Johannes et al. 2010. Ribbon structure created in VMD [PDB 1DM0].[8]

1.1.2 AB_5 Toxin Structure for Stx1 and Stx2

Shiga toxins associated with *E.coli* contain a comparable A-B structure to that found in *S. dysenteriae*.[1] Within STEC, the existence of two *stx* proteins have been identified, Shiga toxin 1 (*stx*1) and Shiga toxin 2 (*stx*2).[11] Shiga toxins generated by *S. dysenteriae* and *stx*1 from *E.coli* have similar A-subunits and identical B-subunits.[12-13] The A-subunit of *stx*1 varies by three nucleotides, leading to an alteration in the amino acid sequence at the catalytic site of the A-subunit.[2, 12] *Stx*1 and *stx*2 contain a similar toxin structure and function; however, *stx*2 is comprised of only 56% of the same amino acid sequence as *stx*1.[14] Specifically, the A-subunit and B-subunit of *stx*2 are comprised of 55% and 57% of the amino acid sequence found in *stx*1, respectively.[14]

The molecular conformation for *stx*2 varies in both the A and B-subunit in comparison to the *stx* conformation.[5] The catalytic active site for *stx*2 is positioned on a

portion of the B-pentamer that allows for accessibility to the A_1 fragment.[5] The A_2 fragment extends through the pore of the B-pentamer, allowing for the formation of an α-helix at the surface of cellular receptors (Figure 1.2).[5] In contrast to the A-subunit of *stx*2, the accessibility of the A-subunit on *stx* and *stx*1 is blocked by the A_2 fragment; therefore, hindering catalytic activity prior to proteolytic cleavage of the A-subunit into A_1 and A_2 fragments.[3-8] As for the B-subunit, variation in the Gb3 binding sites have been identified between *stx* and *stx*2.

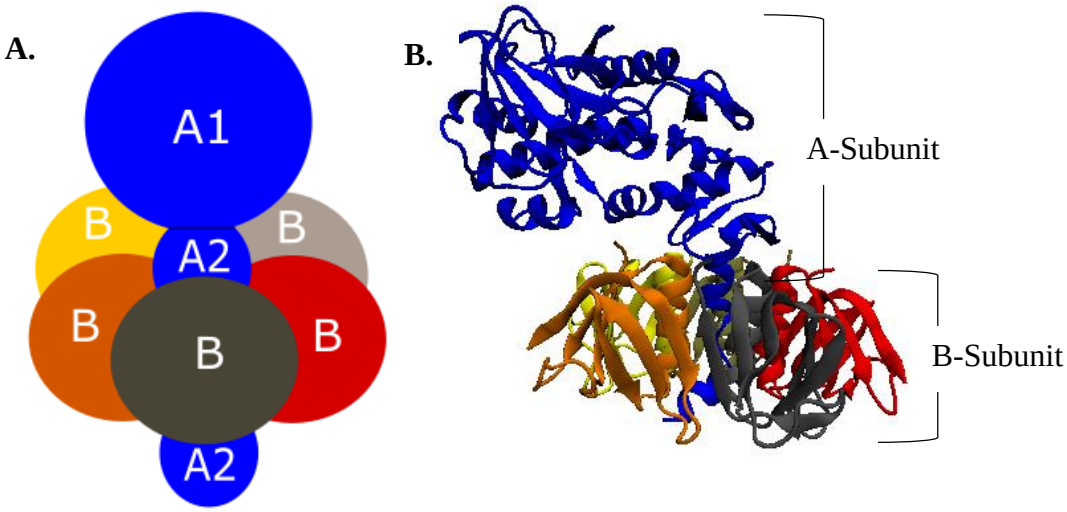


Figure 1. 2: A) Illustration of the structure of Shiga toxin 2 produced by *E.coli*. The A-subunit is represented in blue and shows the Al and A2 fragments, with the A2 fragment traversing the pore of the B-subunit. The B-subunit is comprised of a pentameric ring. B) A ribbon diagram of Stx2 with identical color scheme to graphical representation.[5] Figure adapted from Johannes et al. 2010. Ribbon structure created in VMD [PDB 1R4P].

*Stx*2 varies in conformation and amino acid sequence at the carbohydrate binding site 2, which may be attributed to variation in Gb3 binding affinities between s*tx*1 and *stx*2.[5] The distinction in Gb3 binding affinity between *stx*1 and *stx*2 may be associated with the role of pathogenicity in human illness.

1.2 Role of Stx in Human Illness

The role of *stx* in human illness begins with binding of the B-pentamer to the trisaccharide region of Gb3 located on the plasma membrane of the cell.[1] Expression of Gb3 is often observed in kidney epithelium and endothelium, microvascular endothelial cells located in intestinal lamina propria, and platelets.[1] After the B-pentameric moiety binds to Gb3, Shiga toxins are endocytosed into the cell.[1] A common method for the uptake of Shiga toxins is through clathrin-coated pits.[1] Clathrin has been identified as a protein that allows for the formation of coated vesicles and contributes to the uptake of s*tx* into the cell.[1] The toxin is then transported through the trans-Golgi network, the Golgi apparatus, and the endoplasmic reticulum.[1,3] The A-subunit is translocated into the cytoplasm, where the A_1 subunit cleaves an adenine residue from 28S rRNA of the 60S ribosomal unit. [1-7] As a result, protein synthesis is inhibited within the cell, leading to apoptosis.[1,3] Therefore, *stx* contributes to the disruption of protein synthesis and leads to human illness that could be potentially life threatening.

1.3 Human Illness Associated with STEC

Escherichia coli (*E.coli*) is a diverse bacterial species that is considered a part of the microbiota of the digestive tract of both humans and animals.[11] *E.coli* strains can be classified as commensal or pathogenic bacteria that have been associated with gastrointestinal disease.[3, 11] Pathogenic STEC strains have been linked to intestinal and extraintestinal disease, including hemorrhagic colitis (HC) and hemolytic-uremic syndrome (HUS).[3] The gastrointestinal disease HC includes symptoms accompanied by abdominal pain and watery diarrhea that may progress to bloody diarrhea.[3] Human STEC infections can progress to potentially life threatening HUS that is characterized by low

platelet count (thrombocytopenia), microangiopathic hemolytic anemia, and acute renal failure.[3]

STEC infections account for 265,000 infections in the United States annually, with 5 to 10 percent of infections developing into potentially life-threatening HUS.[16] Pathogenic STEC strains that are often associated with human illness and outbreaks consist of *E.coli* O157:H7 and non-O157 *E.coli* strains.[16] In the United States, *E.coli* O157:H7 strains contribute to 73,000 illnesses, while non-O157 strains contribute to 37,000 illnesses annually.[16-17] STEC infections lead to considerable economic cost to society. In 2005, the cost of STEC infections was estimated to be 405 million dollars.[18] The estimated cost of STEC infections was in relation to the cost of medical care, decrease in job productivity, and infections that lead to premature death.[18] The CDC estimates that 85 percent of STEC infections occur as a result of food contamination.[16] Significant cost of STEC infections have led to increased regulation in food processing and understanding the routes of transmission for STEC.

1.4 Reservoirs and Routes of Transmission for STEC

1.4.1 Potential Reservoirs for STEC

The epidemiology for STEC infections in humans can be evaluated with the farm-to-fork model. Ruminants represent a group of mammals that are considered a natural reservoir for STEC.[11] STEC is present in the gastrointestinal microbiota of healthy cattle and ruminants.[19] Dairy and beef cattle are considered a principal reservoir for STEC and have been implicated as a source of transmission for food-borne illness.[11] Major factors that contribute to prevalence of STEC in cattle include life stage of the cow and seasonality. Life stage is predictive of prevalence of STEC, with STEC often detected in calves and

heifers and decreasing in occurrence in adult cattle. [20] *E.coli* O157:H7 is often found in the excrement of cattle; specifically, weaned calves show higher prevalence (4.9 to 5.3%) of shedding in comparison to pre-weaned calves (1.5 to 2.9%).[20] Seasonality is another factor that contributes to the transmission of STEC. The patterns of STEC outbreaks are often correlated with patterns of *E.coli* O157:H7 and non-O157 shedding in herds.[21] In particular, excrement shedding of *E.coli* O157:H7 during the summer (12.9%) is higher than during winter months (0.3%).[21] As for non-O157, prevalence of non-O157 in cattle excrement was found to be at 19.3%, which was considerably higher than O157:H7.[21]

1.4.2 Routes of Transmission for STEC

The routes of transmission for STEC include infections associated with food-borne pathogens and exposure to the environment. Most often, STEC infections arise from food-borne pathogens, such as undercooked beef products or unpasteurized milk.[22] STEC infections have also been linked to unpasteurized apple juice, mayonnaise, and vegetables.[22] Determination of transmission of STEC can be attributed to food preparation. It has been estimated that 34% of *E.coli* O157:H7 related illnesses can be prevented by proper hand-washing to avoid cross-contamination.[23] In addition, water contamination with STEC is also a source of transmission, with potential water contamination arising from excrement that has been washed into the water supply, rivers, and dams.[23] Other sources of STEC transmission include direct contact with farm animals and person-to-person transmission.[23]

1.5 Characterization of STEC Strains

STEC that has been implicated as a source for human illness can be characterized by the serogroup. Identification of STEC serogroups are categorized by the O (Ohne)

antigen[11] The O-antigen is a component of the lipopolysaccharide (LPS) of Gram-negative *E.coli*.[24] Located on the outermost portion of LPS, the O-antigen consists of a repeating oligosaccharide unit that contains one to eight glycosyl residues (Figure 1.3).[25] The O-antigen varies by sugar composition, sugar sequences, chemical linkages, and substitution of sugar residues.[25] Structural variability within the O-antigen has led to identification of over 181 serogroups.[24] The variability of genes that compose the O-antigen are highly specific and allow for proper identification of *E.coli* serogroups.[26]

Characterization of STEC strains can also be performed by serotype.[11] The serotype is determined by the O-antigen and the H (Hauch) antigen.[11] The H-antigen identifies the presence of flagellar proteins.[11] Serotyping of STEC strains allows for the differentiation between STEC serotypes; including STEC O157:H7 and non-O157 STEC strains.[11] The O-antigen gene cluster used to differentiate STEC serogroups can also differentiate between STEC and non-STEC strains.[26] Specifically, single nucleotide polymorphisms (SNPs) have been identified within the O-antigen gene cluster for non-O157 *E.coli* serogroups O26, O45, O103, O111, O121,and O145.[26] SNPs located within the O-antigen gene cluster have been utilized for identification of serogroups and pathotypes.[26] SNP variation between serogroups have the ability to differentiate between STEC and non-STEC strains, which is predictive of variation amongst O-antigen gene clustering.[26] Furthermore, SNPs identified within specific genes located in specific serogroups have been identified for characterization of non-O157 strains.[26] Based on high specificity and selectivity for detecting the STEC allele in STEC strains, nucleic acid-based assays have been developed for the detection of STEC signatures.

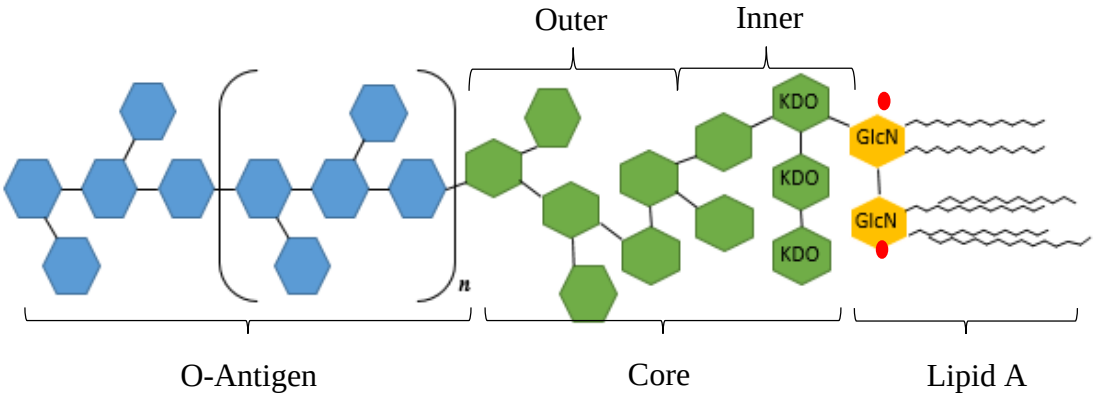


Figure 1.3: Schematic for the basic structure of lipopolysaccharide (LPS) that is located on the outer membrane of gram negative *E.coli*. The O-antigen is labeled in blue and contains repeating oligosaccharide subunits that contribute to variability in sugar composition and the identification of over 181 serogroups.[25] The Core is divided into the inner and outer core that consist of hexose sugars (glucose, galactose, etc.) and 3-deoxy-D-manno-octulosonic acid (Kdo), respectively.[25] Lipid A is responsible for endotoxic activity and is composed of β-D-GlcN-(1-6)-α-D-GlcN disaccharide each containing phosphate groups(red circles).[25] Figure adapted from Erridge et al. 2002.

1.6 Diagnostics for Detection of stx

1.6.1 Tissue Cytotoxicity Assays for stx Detection

Technologies for detection and characterization of STEC are important for both public notification of outbreaks and clinical management.[27] Rapid and sensitive diagnostic methods for STEC are essential for timely mitigation and treatment. Current diagnostic methods are designed to detect the presence of *stx* in fecal samples and STEC isolates.[27]The detection of *stx* by tissue culture is considered the "gold standard." [27] Specifically, tissue culture cytotoxicity assays that utilize African green monkey kidney cells or Vero cells have been employed, due to the high concentration of Gb3 and Gb4 receptors present within their plasma membranes.[27] Vero cells are plated in a monolayer and subject to fecal culture filtrates.[27] Strains of *E.coli* that contain *stx* produce a cytopathogenic effect on the cultured Vero cells, otherwise known as Vero cytotoxin-producing *E.coli* (VTEC) .[27] As a result, tissue culture with Vero cells is considered a highly sensitive diagnostic method for the detection of both *stx*1 and *stx*2 in fecal samples.[27] Although considered the "gold

standard," tissue culture cytotoxicity assays are not a rapid diagnostic method and require 48 to 72 hours for results.[27] In addition, tissue cytotoxicity assays are incapable of differentiating between different types of pathogens that produce Shiga toxins and do not have the ability to characterize the strain by serotype.

1.6.2 Enzyme-linked Immunosorbent Assays (ELISA) for stx Detection

Another method for STEC detection is the enzyme-linked immunosorbent assays (ELISA) that have the capability of detecting s*tx*1 and *stx*2 in clinical and food samples.[28] ELISA utilizes enzyme-labelled antibodies for STEC detection by using polyclonal or monoclonal antibodies that are specific to O and H antigens.[28] The type of ELISA often used for the detection of STEC is the sandwich ELISA.[28] The polyclonal or monoclonal antibodies are capture ligands that bind to the antigen of specific toxins.[28] The detection event for ELISA occurs when an *stx* detection antibody is added to the bound capture ligand, followed by incubation with an anti-immunoglobulin enzyme conjugate that signifies the presence of the antigen through color formation.[28] ELISA is considered less sensitive than Vero cytotoxicity assays due to the inability for ELISA to detect s*tx* at low concentrations in fecal extracts (<50 pg/mL).[28] In comparison, ELISA has the capability to detect s*tx* in fecal cultures where *stx* is present at higher concentrations (>100 pg/mL).[28] ELISA is considered both a rapid and sensitive detection system when evaluating fecal cultures, with results obtained within 1 day.[28]

1.7 Molecular Diagnostics for Detection of stx genes

Molecular diagnostics have been employed as a rapid and sensitive detection method for identifying pathogenic markers.[29] Assay development utilizing genomic markers to identify STEC genomic signatures, provides both a rapid and highly sensitive

detection system.[29] Identification of nucleotide sequences for *stx1* and *stx2* genes have allowed for the design of oligonucleotide primers for amplification of s*tx* genes through polymerase chain reaction (PCR).[28] Nucleic acid-based assays can be classified into two types of amplification; target amplification or signal amplification.[29] Target amplification techniques that are often associated with STEC detection are multiplex PCR and multiplex real-time PCR (RT-PCR). STEC detection methods for signal amplification include Multiplex ligase-dependent probe amplification (MLPA) and Multiplex oligonucleotide ligation-PCR (MOL-PCR). [29]

1.7.1 Multiplex PCR for Detection of stx genes

Multiplex PCR for the detection of s*tx*1 and *stx*2 genes in STEC isolates and fecal samples is the most common target amplification technique. The identification of nucleotide sequences for s*tx*1 and s*tx*2 have allowed for the design of synthetic oligonucleotide primers for target amplification through Multiplex-PCR.[30] Multiplex-PCR allows for simultaneous detection of STEC genomic markers (s*tx*1, *stx*2, eae, etc.) and characterization of STEC in a single tube reaction.[29-30] In the presence of a specific target, oligonucleotide sequences can amplify specific DNA sequences generating amplified copies of the target of interest.[31] Amplified products are often detected through end point detection by gel or capillary electrophoresis.[29]

The addition of multiple primer pairs for the detection of a specific genomic marker presents multiple disadvantages. With the presence of multiple oligonucleotide primers, primer sequences in multiplex may vary in GC content leading to variation in time of denaturation and reduction in amplification efficiency.[31] Differences in length of PCR primer and GC content lead to variation in melting temperature (Tm); therefore, designing

PCR primers with similar Tm is vital for successful amplification of PCR product.[31] In addition, designing multiple oligonucleotide primers requires careful selection of primer sequences to prevent primer dimer formation, due to portions of the primer sequence that are complementary to other primers.[31] As a result, the potential disadvantages for PCR primer design for multiplex-PCR contributes to a decreased level of multiplexing.[31]

Another disadvantage for multiplex-PCR is the analysis of multiplex PCR products. End point detection of PCR products is a cumbersome technique that is not considered a high throughput technique.[29] Gel electrophoresis separates PCR products based on the size of the DNA fragment.[31] Designing primers with similar amplicon size presents a challenge when differentiating between PCR products through gel electrophoresis.[31] To increase reliability of multiplex-PCR, multiplex-PCR is often coupled with nucleic acid microarrays and real-time PCR for the detection of amplified PCR products.[31]

1.7.2 Multiplex Real Time-PCR for Detection stx genes

Multiplex RT-PCR is a technology that allows for the detection and quantification of genomic signatures.[29] Real-time PCR assays utilize TaqMan or molecular beacon probes that are comprised of synthetic oligonucleotide sequences that are conjugated to a fluorophore (5') and a quencher (3').[33] The quencher functions to quench the fluorescent emission of the fluorophore prior to modification of the oligonucleotide probe, while the fluorophore serves as the reporter molecule for the quantification of pathogenic signatures during PCR amplification.[33] The modification of TaqMan probes exhibit increased fluorescence intensity, with the hydrolysis of the fluorophore by Taq polymerase during PCR.[33] The molecular beacon probe emits fluorescence as a result of hybridization to the complementary target of interest, causing a conformational change to the probe that allows

for increased fluorescence intensity.[33] RT-PCR has been applied for the detection and quantification for O157:H7 and non-O157 STEC serogroups, in addition to the detection of STEC virulence genes (*stx*1, *stx*2, and *eae*).[33-35]

Advantages to RT-PCR include the simultaneous detection and quantification of STEC pathogenic markers in a single tube reaction, leading to the elimination of post-PCR processing and a reduction in cross-contamination of PCR products.[32-33] RT-PCR is a high throughout technique that allows for multiplex detection of genomic signatures, with multiple signatures tested simultaneously.[32] RT-PCR is a highly sensitive technique that is considered 100-fold more sensitive than conventional end-point PCR.[35] Potential disadvantages for real-time PCR include limitations to multiplex level, cost of reagents due to limited multiplex capability, and PCR primer incomaptibilty.[29] Probe design considerations must also be taken into account when selecting fluorophores due to the overlap of emission spectra; therefore further complicating reconfiguration with the addition of genomic signatures and lending toward reduced signatures evaluated simulataneously.[32]

1.8 MOL-PCR for STEC Detection

Multiplex oligonucleotide ligation-PCR (MOL-PCR) is a nucleic acid-based detection platform that can be used for the detection of O-serogroups and STEC-associated virulence factors (*stx*1, *stx*2, *eae*). Signal amplification is incorporated into MOL-PCR assay chemistry that allows for the detection of unique signatures followed by amplification of the signal probe.[29] Specifically, MOL-PCR incorporates multiplex ligation into the assay chemistry, which allows for ligation of oligonucleotide probes adjacent to each other in the presence of a complementary target sequence.[29] MOL-PCR allows for

simultaneous detection of multiple nucleic acid signatures in a single tube reaction.[29] In addition, MOL-PCR assay chemistry has been combined with microsphere arrays and flow cytometry for interrogation of pathogenic markers. Therefore, MOL-PCR is a high throughput technique that allows for the simultaneous detection of multiple nucleic acid signatures; including single nucleotide polymorphisms (SNPs), insertions, deletions, and repeats.[36]

The integration of multiplex ligation into MOL-PCR assay chemistry offers many advantages over target-based amplification. The addition of the ligation step allows for not only the detection of the target of interest, but allows for the characterization of the strain based on the detection of a SNP.[36] 'MOL-PCR also allows for a high degree of multiplexing due to the incorporation of multiplex ligation.[29] The reconfiguration and expansion of the MOL-PCR assay is less cumbersome than other target-based amplification techniques.[29] The addition of target detection probes pairs to the MOL-PCR assay usually has limited interaction with existing probes; therefore, reducing the time and cost of reconfiguration in comparison to RT-PCR or multiplex-PCR.[29]

1.9 Previous Application of MOL-PCR for Detection of STEC-8

Molecular diagnostics that incorporate nucleic acid-based assays have been developed as diagnostic platforms for the detection of STEC. In particular, a high throughput MOL-PCR assay have been previously developed for the detection of STEC genomic signatures associated with human illness. The research in that study have been developed for the detection of 8-STEC serotypes (STEC-8) including; O26, O45, O103, O104, O111, O121, O145, O157:H7.[37] In addition, the study can identify STEC associated virulence genes *stx*1, s*tx*2, and *eae*. Published genomic sequences for the STEC-8 and

virulence genes were used to design target-specific probes for the detection of each unique signature.

The signatures for the detection of the STEC-8 and virulence genes was designed as an 11-plex MOL-PCR assay. This assay has the capability to identify each of the eight STEC serogroups and evaluates genomic signatures for the virulence genes (*stx*1, *stx*2, and *eae*).[37] The signatures associated with the virulence genes allow for the discrimination between a STEC and a non-STEC strain. The assay successfully evaluated isolated genomic DNA from both commercial sources and STEC-CAP collaborators.[37] The MOL-PCR STEC-8 assay provided the basis for the research presented in this thesis. With the success of the STEC-8 assay, a detection and characterization assay that encompasses SNPs identified by Norman et al. have been developed for the identification of non-O157 serogroups. The SNPs identified for each non-O157 serogroup are serogroup specific and are also unique to STEC strains.[26] Therefore, the SNPs allow for detection of a serogroup and characterization of the strain with a single genomic marker without having to detect the serogroup and characterize the strain with multiple genomic markers.

Chapter 2: Motivation and Overview of the MOL-PCR SNP Based Assay

The motivation for this research is to develop a SNP based MOL-PCR assay for the detection of non-O157 STEC serogroups that have recently been declared as adulterants in raw non-intact beef.[38] In 2012, the Food Safety and Inspection Service (FSIS) implemented mandatory testing for non-O157 STEC serogroups O26, O45, O103, O111, O121, and O145 (otherwise known as "big six" non-O157 serogroups).[38] The SNP based MOL-PCR assay allows for further characterization of non-O157 isolates after initial screening. The assay developed here is a high throughput, nucleic-acid based assay that utilizes MOL-PCR assay chemistry for the detection of non-O157 serogroups. This assay improves upon the current STEC-8 MOL-PCR assay because it not only has the capacity to identify the targeted serogroup, but can differentiate between a STEC and non-STEC strain with the direct detection of a single nucleotide polymorphism.

Characterization of non-O157 isolates is important for mandatory testing of the non-O157 serogroups because not all non-O157 strains are considered pathogenic.[17, 26] *E.coli* strains that have been identified in cattle contain the same O-antigen as a STEC strain, yet not all strains contain the virulence genes that encode Shiga toxins.[26] Therefore, the assay developed here can further characterize a non-O157 isolate by identifying the O-serogroup and differentiating between strains that contain STEC-associated alleles.

The development of the MOL-PCR SNP based assay for the detection of non-O157 serogroups is described in Chapter 3. The selection of DNA signatures for each of the non-O157 serogroups were based on the SNPs identified by Norman et al.[26] In their study, they used matrix-assisted laser desorption-ionization time-of-flight (MALDI-TOF) genotyping to determine the frequency of STEC associated alleles for a panel of non-O157 STEC

strains.[26] The MALDI-TOF assays were classified into three different classification groups based on the virulence profile (*stx* alone, *stx* with *eae*, *stx* with *eae* and *ehxA*).[26] The three classification groups provided sensitivity and specificity estimates for the STEC allele to accurately detect the STEC strains associated with the virulence profiles.[26] Sensitivity estimates refer to the proportion of strains that would result in a positive identification for producing the STEC-associated allele.[26] The specificity estimates measured the proportion of strains that would produce a negative result and is indicative of a strain that is lacking the STEC-associated allele.[26] Therefore, specificity estimates were used to correctly identify strains that do not contain the STEC-associated allele found in STEC strains.[26] Based on specificity and selectivity estimates, genes were selected for each of the six serogroups. Genes selected were used to computationally design PCR probes for both STEC and non-STEC signatures.

Two MOL-PCR SNP based assays have been designed for the detection of non-O157 serogroups and their associated SNPs. Originally designed as a 14-plex detection assay for both STEC and non-STEC, the competition at the SNP site present on MOLigo2 required the division of this assay into two detection assays. The STEC and non-STEC MOLigo pairs have been separated into two 8-plex probe panels, each containing an internal positive control. The sensitivity of the MOLigo pairs for both detection assays were evaluated against DNA panels provided by both collaborators and commercial sources. The 8-plex SNP based MOL-PCR assay correctly identifies the big six non-O157 serogroups of interest. The non-STEC SNP assay was evaluated against known STEC strains to test the selectivity of the SNP against non-target samples. The assay demonstrates the ability to be highly specific when in the presence of multiplexed MOLigo pairs that

contain a single nucleotide base change. Here I demonstrate the development of two high throughput SNP based MOL-PCR assays for the detection and characterization of non-O157 strains.

Chapter 3: Materials and Methods

3.1. E.coli non-O157 Panel

Two bacterial strain panels were used for the development of the SNP based MOL-PCR assay. The panel consisted of purified isolated genomic DNA for non-O157 serogroups O26, O45, O103, O111, O121, and O145. The strains for each serogroup contained the STEC-associated allele and were identified as strains that contained the STEC-associated virulence genes. The strains were obtained from both commercial sources and collaborators. A training panel consisting of complementary synthetic target DNA for the O26 serogroup was also designed to evaluate specificity and sensitivity of the SNP when both STEC and non-STEC probes were evaluated in multiplex.

3.2 SNP Selection

The selection of genomic signatures for this assay were designed from SNPs identified in a previous study.[26] MOL-PCR assay chemistry consists of two MOLigo probes that are designed to anneal at the SNP discriminate site. To ensure the success of the MOL-PCR assay, SNPs were chosen based on sensitivity and specificity estimates provided by Norman et al.[26] Categories for sensitivity and specificity estimates consisted of the SNPs identified for each serogroup, evaluated against strains that were positive for *stx* alone, *stx* with *eae*, and *stx* with both *eae* and *ehxA*.[26] The O26 gene *fnl1* contained a SNP that captured the O26 strains associated with *stx1*.[26] Emergent strains for O26 have recently been found to contain the *stx2* virulence gene.[26] Therefore, the selection of two genes for O26 was necessary to capture strains that were positive for either *stx1* or *stx2*. The gene *wbhW* 997 was initially selected for O45 based on the highest overall sensitivity and specificity estimates. After initial evaluation, O45 *wbhW* 997 was removed from the

probe panel due to cross reactivity between other MOLigo probes when evaluated in multiplex. O45 *wbhQ* 721 was selected to replace the gene O45 *wbhW* 997. O103 *wbtD* 937 and O145 *wzy* 37 were the only SNPs identified for each serogroup and were selected by default. The selection of O111 *wbdK* 687 and O121 *vioA* 313 were selected based on the highest overall sensitivity and specificity estimates for each category. Norman et al. characterized the SNPs based on the gene, the base pair position within the gene, and the change in allelic discrimination.[26] Table 3.1 lists the genes for the non-O157 serogroups that were selected for the design of oligonucleotide probes for the SNP based MOL-PCR assay. The referenced genome sequence for each gene appears in Appendix A.1-A.7.

Table 3.1: List of non-O157 single nucleotide polymorphisms used in MOL-PCR assay

Serogroup	Gene	Base Pair Position [a]	Polymorphism [b]
O26[c]	*fnl1*	88	G→A
	wzx	953	T→G
O45[d]	*wbhQ*	721	C→A
O103[e]	*wbtD*	937	C→T
O111[f]	*wbdK*	687	C→T
O121[g]	*vioA*	313	C→T
O145[h]	*wzy*	37	A→C

[a] Location of the SNP within the gene
[b] Non-STEC → STEC
[c] GenBank AY763106 Reference Genome
[d] GenBank AY771223 Reference Genome
[e] GenBank AY532664 Reference Genome
[f] GenBank AF078736 Reference Genome
[g] GenBank AY208937 Reference Genome
[h] GenBank AY647260 Reference Genome

3.3 MOLigo Probe Design

3.3.1 MOLigo Modular Components

The SNPs selected were used to computationally design highly specific single-stranded oligonucleotide probes for the detection of non-O157 serogroups, otherwise known as MOLigo pairs (MOL-PCR oligonucleotide pairs). MOLigo pairs, designated

MOLigo1 and MOLigo2, are comprised of modular components that are 40-90 nucleotides long. The modular components main functions are 1) target sequence detection, 2) amplification of ligation products, and 3) hybridization of amplified products onto a microsphere array. MOLigo1 contains two functional modular components that consist of a sequence that is complementary to the 5' dual biotinylated forward primer and a sequence that is complementary to the target sequence (Figure 3.1).

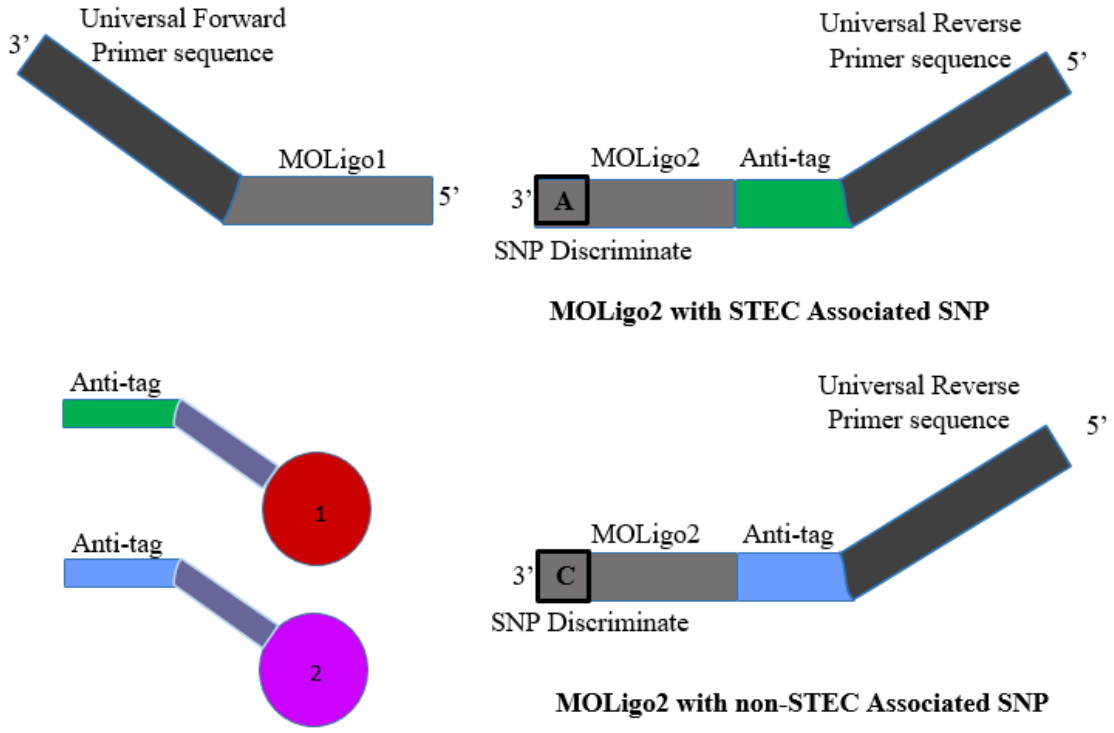


Figure 3.1: Schematic for the composition of the modular components for MOLigo1 and MOLigo2. MOLigo2 for the STEC and non-STEC probe differ by both a single base on 3' end and the anti-tag sequence. Depiction of two fluorescently labeled beads with 24-base capture sequence are used for discrimination between a STEC and non-STEC strain. Schematic adapted from Stucki et al.[39]

MOLigo2 contains three modular components that consist of a sequence that is complementary to the universal reverse primer, a 24-base pair anti-tag sequence for bead capture of the amplified products onto a microsphere array, and a complementary target

sequence that contains the allelic discriminate (SNP) on the 3' end of the probe (Figure 3.1).

3.3.2 MOLigoDesigner Tool

MOLigo probe pairs for the biallelic SNP genomic marker selected for MOL-PCR were computationally designed using the MOLigoDesigner software developed at Los Alamos National Laboratory (https://moligodesigner.lanl.gov/MoligoDesigner/). MOLigo pairs for each serogroup were designed to be highly specific to the target of interest. Reference sequences provided by Norman et al. were used as input sequences for the MOLigo design tool (Appendix A.1-A.7).[26] The input sequence consisted of the genomic sequence for each serogroup and the base pair position for the SNP for each non-O157 gene selected. Default tags listed in the software were selected in conjunction with beads that were compatible with the MAGPIX® instrument. The default conditions and parameters programmed in the MOLigoDesigner software were used for designing each MOLigo pair.

The MOLigoDesigner software provided two options for MOL-PCR assay chemistry; Classic MOL-PCR and Reverse MOL-PCR.[40] The version of MOL-PCR assay chemistry selected for this assay was Reverse MOL-PCR. Reverse MOL-PCR differs in the design of modular components in comparison to Classic MOL-PCR.[40] In Reverse MOL-PCR, the 24-base anti-tag sequence is located on MOLigo2 versus Classic MOL-PCR that contains the complementary tag sequence on MOLigo1.[40] Reverse MOL-PCR allows for greater flexibility when testing multiple alleles in a single tube reaction; furthermore, using a flow cytometer that has limitations in the number of fluorescent channels, such as the Luminex Magpix®.[40] Table 3.2 and Table 3.3 list the MOLigo pairs

and primer sequences that were ordered from Integrated DNA Technologies (http://www.idtdna.com/) and used for experimental testing for both non-O157 assays.

Table 3.2: List of MOLigo sequences for STEC non-O157 MOL-PCR assay.

Serogroup	xTAG	SNP	MOLigo	Sequence (5'-3')
O26 S-*fnl1*	A042	A	M1	AAT ATC TGT ATC TAG AAA ACG CTT AAG TAC **TCT CAC TTC TTA CTA CCG CG**
			M2	ACT CGT AGG GAA TAA ACC GTA TTT GTT ATG ATA AAT GTG TAG TGA TCC CTG CTA AAT ATT CGT ATT TCA G**T**
O26 S-*wzx*	A021	G	M1	CCC CCC CTA AAC TCA GGA C**TC TCA CTT CTT ACT ACC GCG**
			M2	ACT CGT AGG GAA TAA ACC GTA TTA AGT AAG AAT TGA GAG TTT GAG AAC CTT ATA TCC CAA TAT AGT ACC C**C**
O45 S-*wbhQ*	A057	A	M1	CAG AAC ACG TGA GTC TAG CTT TG**T CTC ACT TCT TAC TAC CGC G**
			M2	ACT CGT AGG GAA TAA ACC GTA GAG TAT TAG TAG TTA TTG TAA GTA GCA GAC AGA GTT CTG CGA **T**
O103 S-*wbtD*	A012	T	M1	CAG GTT AAT ATT CTT ATT GTT ATT GAT TTT TTT T**TC TCA CTT CTT ACT ACC GCG**
			M2	ACT CGT AGG GAA TAA ACC GTA GTA GAA AGT TGA AAT TGA TTA TGA GAT ATT ATA TAT AAT TGC ATT TTA GAT TTA ATA T**A**
O111 S-*wbdK*	A078	T	M1	CCT GTA ACC ATA TTC TCT TTT GGT AAA TT**T CTC ACT TCT TAC TAC CGC G**
			M2	ACT CGT AGG GAA TAA ACC GTG TAT AGT GTG ATT AGA TTT GTA AAA CTC TTC GAA AAT ATC ATC ACT CTT AGT **A**
O121 S-*vioA*	A063	T	M1	ATC AAT ATC GAC AAA AAC AGG TTT AAG C**TC TCA CTT CTT ACT ACC GCG**
			M2	ACT CGT AGG GAA TAA ACC GTT TTG TTG TTA AGT ATG TGA TTT AGT CGG ATC GAT ATT TAA GGT TTT GG**A**
O145 S-*wzy*	A076	C	M1	CAC TCC ATT TAT AAA CTT ATC TTT CTT TAT ATT CA**T CTC ACT TCT TAC TAC CGC G**
			M2	ACT CGT AGG GAA TAA ACC GTA AAG AAT TAG TAT GAT AGA TGA GAG TAA TAT AAC GAA GAA ATA ATT AAC CAA AAA AAA A**G**
Ligation Control	A056		M1	CTT TGG GGA TTA GCT CCG C**TC TCA CTT CTT ACT ACC GCG**
			M2	ACT CGT AGG GAA TAA ACC GTA ATT AGA AGT AAG TAG AGT TTA AGT GCA ATC CGA ACT GAG ATG GA

The forward primer, reverse primer, and 24-base anti-tag sequence are in bold, underlined, and italicized, respectively. The SNP discriminate is in bold and underlined. S denoted after the serogroup identifies the MOLigo that contains the STEC-associated allele.

Table 3.3: List of MOLigo sequences for non-STEC PCR assay.

Serogroup	xTAG	SNP	MOLigo	Sequence (5'-3')
O26 NS-*fnl1*	A018	G	M1	AAT ATC TGT ATC TAG AAA ACG CTT AAG TAC **TCT CAC TTC TTA CTA CCG CG**
			M2	ACT CGT AGG GAA TAA ACC GTG TAA TTG AAT TGA AAG ATA AGT GTC CCT GCT AAA TAT TCG TAT TTC AG**C**
O26 NS-*wzx*	A014	T	M1	CCC CCC CTA AAC TCA GGA C**TC TCA CTT CTT ACT ACC GCG**
			M2	ACT CGT AGG GAA TAA ACC GTA TTG TGA AAG AAA GAG AAG AAA TTG AAC CTT ATA TCC CAA TAT AGT ACC C**A**
O45 NS-*wbhQ*	A051	C	M1	CAG AAC ACG TGA GTC TAG CTT TG**T CTC ACT TCT TAC TAC CGC G**
			M2	ACT CGT AGG GAA TAA ACC GTG ATA AGA AAG TGA AAT GTA AAT TGA GCA GAC AGA GTT CTG CGA **G**
O103 NS-*wbtD*	A073	C	M1	CAG GTT AAT ATT CTT ATT GTT ATT GAT TTT TTT T**TC TCA CTT CTT ACT ACC GCG**
			M2	ACT CGT AGG GAA TAA ACC GTG TTG AGA ATT AGA ATT TGA TAA AGC AAG TTT AGA TAT TAT ATA TAA TTG CAT TTT AGA TTT AAT AT**G**
O111 NS-*wbdK*	A026	C	M1	CCT GTA ACC ATA TTC TCT TTT GGT AAA TT**T CTC ACT TCT TAC TAC CGC G**
			M2	ACT CGT AGG GAA TAA ACC GTT TTG ATT TAA GAG TGT TGA ATG TAC TCT TCG AAA ATA TCA TCA CTC TTA GT**G**
O121 NS-*vioA*	A053	C	M1	ATC AAT ATC GAC AAA AAC AGG TTT AAG C**TC TCA CTT CTT ACT ACC GCG**
			M2	ACT CGT AGG GAA TAA ACC GTG TTT GTG TTT GTA TAA GTT GTT AAT CGG ATC GAT ATT TAA GGT TTT GG**G**
O145 NS-*wzy*	A055	A	M1	CAC TCC ATT TAT AAA CTT ATC TTT CTT TAT ATT CA**T CTC ACT TCT TAC TAC CGC G**
			M2	ACT CGT AGG GAA TAA ACC GTG AAG ATA TTG AAA GAA TTT GAT GTA GTA ATA TAA CGA AGA AAT AAT TAA CCA AAA AAA AA**T**
Ligation Control	A056		M1	CTT TGG GGA TTA GCT CCG C**TC TCA CTT CTT ACT ACC GCG**
			M2	ACT CGT AGG GAA TAA ACC GTA ATT AGA AGT AAG TAG AGT TTA AGT GCA ATC CGA ACT GAG ATG GA

The forward primer, reverse primer, and 24-base anti-tag sequence are in bold, underlined, and italicized, respectively. The SNP discriminate is in bold and underlined. NS denoted after the serogroup identifies the MOLigo that contains the non-STEC associated allele.

3.4 MOL-PCR Assay

MOL-PCR is a nucleic-acid based diagnostic platform that allows for simultaneous detection of multiple genomic signatures.[29, 36] MOL-PCR assay chemistry incorporates multiplex ligation that allows of detection of multiple targets in a single reaction.[29-36] I have adapted and modified MOL-PCR assay chemistry by separating the ligation and PCR

reaction into two separate reactions. The MOL-PCR assay consists of the following steps (Figure 3.2):

1. MOL-PCR is comprised of single-stranded oligonucleotide probes that are designated as MOLigo pairs (MOL-PCR oligonucleotide pairs). MOLigo1 and MOLigo2 anneal adjacent to each other at the SNP discrimination site in the presence of a complementary target sequence. Thermostable DNA ligase covalently links MOLigo1 and MOLigo2 together, creating a single stranded DNA molecule that is 100-150 nucleotides in length. The MOLigo1+MOLigo2 complex serves as the template for singleplex PCR amplification. The ligation of MOLigo1 and MOLigo2 serves as the target detection event, with DNA ligase binding the MOLigo pairs only in the presence of a complementary target sequence.
2. The ligation product proceeds with singleplex PCR amplification with a set of universal primer pairs. MOLigos that did not anneal adjacent to one another will produce asymmetric amplified PCR products because the forward primer and reverse primer remain on separate MOLigo pairs. Ligated MOLigos contain modular components that are complementary to the universal forward and reverse primer sequences. The universal forward primer is composed of a 5’ dual biotinylated moiety that serves as the binding site for the reporter molecule Streptavidin-R-phycoerythrin (SAPE).
3. The amplified ligation products are hybridized to a Luminex xTAG® microsphere array using xTAG® sequences that are complementary to anti-tag sequence on MOLigo2. Luminex xTAG® microspheres are color coded microspheres and contain a 24-base capture sequence that is covalently attached to a polystyrene

microsphere. The bead PCR-products are incubated with SAPE forming the biotin-streptavidin complex that allows for production of fluorescence during signal detection on a Luminex platform. The hybridized product is analyzed by flow cytometer that measures fluorescence intensity emitted by the labeled PCR product.

3.5 MOL-PCR Protocol

STEC and non-STEC MOLigos were divided into two separate assays for the interrogation of the non-O157 SNPs. Potential competition at the SNP discriminatory site required the division of the designed 15-plex assay into two 8-plex assays. Each 8-plex assay consisted of 16 MOLigo probes and 8 MagPlex-xTAG® microspheres. The assay includes three controls; an internal positive control for ligation, a full length PCR amplification control, and a no template PCR control to evaluate cross reactivity in the absence of a DNA template. The MOL-PCR protocol is divided into four steps:

1. **Annealing and Ligation of MOLigo Pairs:** The ligation reaction was performed in a final reaction of 20 µL containing 1.25 units of Ampligase® Thermostable DNA Ligase (Epicentre), 1x Ampligase® reaction buffer (Epicentre), 2 nM of each MOLigo probe (Integrated DNA Technologies (IDT)), 1.25 units of Uracil-DNA Glycosylase (UDG) (ThermoFisher Scientific), 1.5 mg/mL Salmon Sperm DNA (ThermoFisher Scientific) and 2 µL of template DNA. The thermal cycling profile for the ligation reaction includes UDG treatment to prevent PCR carry-over contamination at 37°C for 30 minutes, DNA denaturation at 95°C for 3 minutes, 30 ligation cycles of 95°C for 25 seconds and 58°C for 2 minutes, UDG inactivation at 95°C for 30 minutes, and hold at 4°C until further processing.

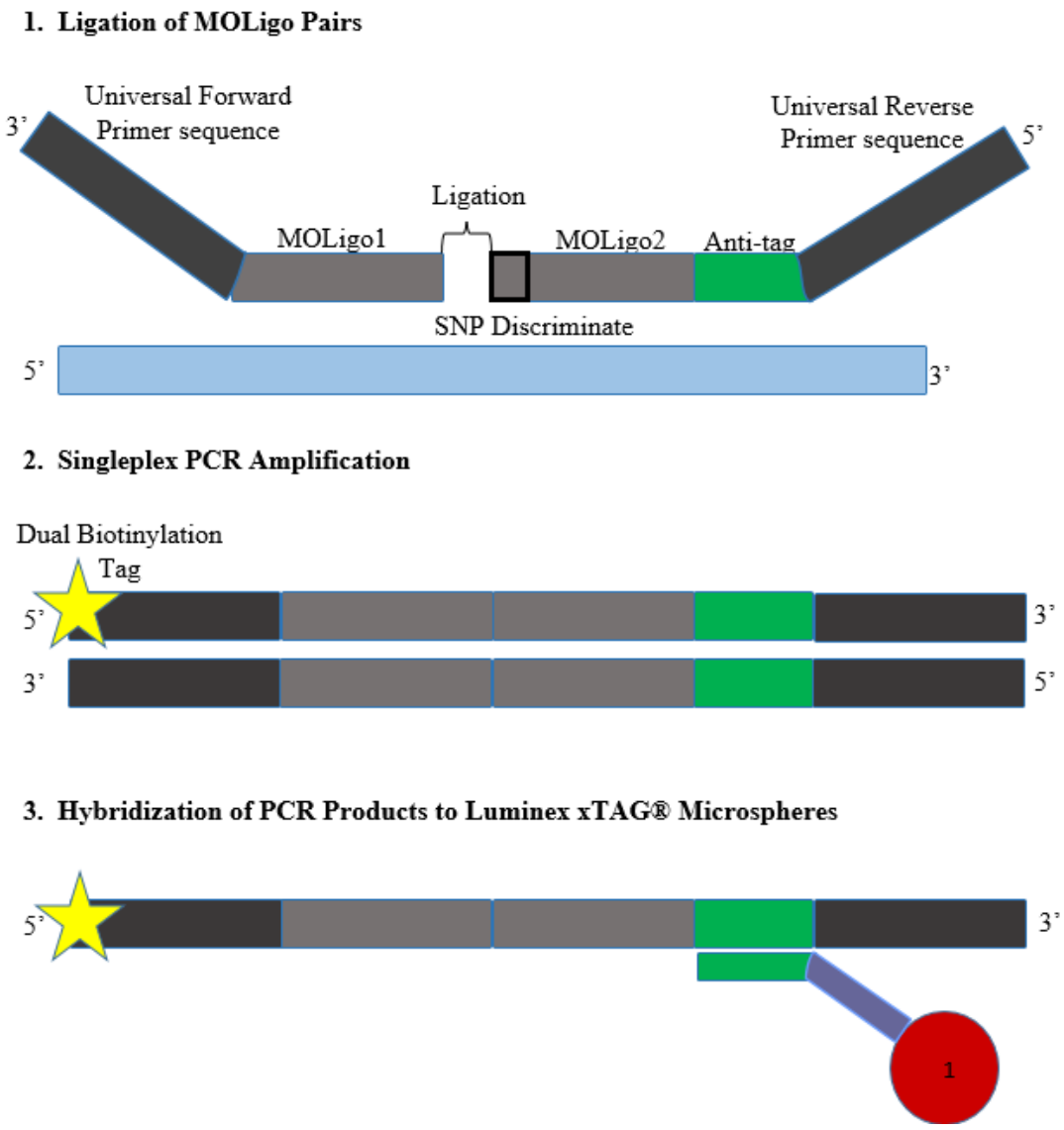


Figure 3.2: Schematic for MOL-PCR. 1) MOLigo pairs anneal adjacent to each other at the SNP discriminate site. DNA ligase recognizes the structure and covalently links MOLigo1 to MOLigo2, creating single stranded DNA molecule. 2) Ligated product proceeds through PCR amplification with biotinylated forward primer. 3) Amplified ligation product is hybridized to Luminex xTAG microsphere via 24-base anti-tag sequence. Bead-PCR complex is analyzed on Luminex platform. Schematic adapted from Stucki et al.[39]

2. **PCR Amplification of PCR Product:** The PCR reaction was performed in a final reaction of 20 µL containing GenAmp® 1X PCR Buffer II (ThermoFisher Scientific), 4 mM of $MgCl_2$ (ThermoFisher Scientific), 0.05 U of AmpliTaq Gold® DNA polymerase (ThermoFisher Scientific), dNTP/dUTP mix composed of 200

µM each of dATP, dCTP, dGTP and 400 µM dUTP (ThermoFisher Scientific), 0.5 µM of 5’ dual biotinylated forward primer (IDT) (Table 3.4),0.1 µM reverse primer (IDT)(Table 3.4), and 2 µL of the ligation product was added to the PCR master mix. The thermal cycling profile for PCR reaction is as follows; 95°C for 10 minutes, 45 cycles of 95°C for 15 seconds, 58°C for 20 seconds, 72°C seconds, and a final extension at 72°C for 7 minutes. The reaction were held at 4°C until proceeding with the bead-hybridization step.

Table 3.4: Forward and Reverse Primer Sequences

Forward Primer	/52-Bio/CGC GGT AGT AAG AAG TGA GA
Reverse Primer	ACT CGT AGG GAA TAA ACC GT

3. **Hybridization of beads:** The microsphere mixture was prepared using anti-tag Luminex xTAG® microspheres for each MOLigo pair that was being interrogated. Each bead mix contained 8 MagPlex x-TAG® microspheres (Luminex Corporation) for both the STEC and non-STEC assay. The bead mixture contained 2500 microspheres of each bead per reaction and was stored in 800 mM NaCl and 50 mM MES (Sigma Aldrich). For the hybridization reaction, 10 µL of bead mixture was added to the MOL-PCR reaction. The thermal cycling conditions include; 95°C for 3 minutes, a slow ramp down decreasing in temperature to 50°C in 10°C increments with a hold for 1 minute, 50°C for 30 minutes, and a ramp down decreasing in temperature to 25°C by 5°C increments. The reaction was held at 4°C for further processing.

4. **Labeling of PCR product and Analysis on Luminex Platform:** The bead-PCR product proceeded through a wash step before being labeled with SAPE. The wash step consists of separation of the supernatant from the bead-PCR product using a magnetic separator (EMD Millipore) for 5 minutes. SAPE is prepared at 10 ug/mL (BD Bioscience) in 1x TE Buffer (Affymetrix). The addition of 25 µL of SAPE is added to each reaction and incubated for 30 minutes. The labeled bead-PCR complex is washed to remove excess SAPE and re-suspended in 100uL of 1x TE Buffer (Affymetrix). The reaction is transferred into a 96-well v-bottom plate for the interrogation on the Luminex MagPix® (EMD Millipore).

3.6 Wilcoxon Rank Sum Test for SNP calling

The test samples and the control samples (no template control, ligation control, and PCR control) were interrogated using the Luminex MagPix®. Median fluorescence intensity (MFI) values were used in conjunction with the Wilcoxon rank sum test for statistical analysis. MFI values were based on the collection of 100 beads per region in each sample set. The Wilcoxon rank sum test is non-parametric statistical analysis tool that allows for the comparison of two samples that are either independent or unrelated. [41] The statistical analysis algorithm was developed by Woods et al. for the analysis of STEC samples using MOL-PCR assay chemistry.[37] The MFI values for the no template control and the test samples were combined and ranked in ascending order. The sum of the ranks were computed for both the test samples (W_A) and the no template control(W_B). Determination of the p-value for the Wilcoxon rank sum statistic was calculated by determining the distribution of the ranked sums (μ_A), the standard deviation (σ_A), and the

z-score (z). The sum of the ranks computed for the test samples (W_A) were used to calculate the z-score.

$$\mu_A = \frac{n_1(n_1 + n_2 + 1)}{2}$$

$$\sigma_A = \sqrt{\frac{n_1 n_2((n_1 + n_2 + 1)}{12}}$$

$$z = \frac{(W_A - 0.5) - \mu_A}{\sigma_A}$$

The approximate and exact Wilcoxon rank sum test were used for identification of the test samples. The approximate Wilcoxon rank sum test compared the p-value to the statistical significance level($\alpha = 0.01$). P-values greater than α= 0.01 resulted in a positive identification for the serogroup and P-values less than α= 0.01 resulted in a negative identification for the serogroup.

P-value $>$ 0.01=Positive

P-value $<$ 0.01=Negative

The exact Wilcoxon rank sum test compared the sum of ranks for the test samples (W_A) to the critical value. A one-tailed z-test and a statistical significance level of 0.05 was used for the exact Wilcoxon rank sum test. The critical values correspond to the statistical significance level and are listed in Appendix B. The critical value that were greater than W_A resulted in a negative call for the serogroup. If the critical value is less than W_A, the test sample resulted in a positive call for the serogroup.

Critical Value $<$ W_A =Positive

Critical Value $>$ W_A=Negative

Chapter 4: Results and Discussion

4.1 MOLigo Probe Selection and Assay Design

The MOL-PCR assay in this study evaluated 14 genomic markers for STEC and non-STEC; in addition to an internal ligation control and a PCR control. MOLigo pairs were designed for the biallelic SNP associated with each non-O157 serogroup. The MOL-PCR assay was originally designed as a 15-plex assay that consisted of MOLigo pairs for each biallelic SNP and an internal ligation control. Due to competition at the SNP discriminate site, the MOL-PCR 15-plex assay was divided into two 8-plex assays that evaluated the STEC and non-STEC associated alleles separately. MOLigo1 was identical for both biallelic SNPs, with MOLigo2 differing by both the SNP discriminate located on 3' end and the anti-tag sequence.

The selectivity for each biallelic SNP was associated with sensitivity and specificity estimates provided by Norman et al.[26] Estimates for each SNP were based on the capacity for the SNP to identify strains that contained *stx* alone, *stx* with *eae*, and *stx* with *eae* and *ehxA*.[26] Therefore, the SNPs for each serogroup were selected based on the highest sensitivity and specificity estimates for each category. SNPs for each non-O157 serogroup were evaluated in singleplex and multiplex format to determine overall performance of each SNP. The SNP selected for O45 *wbhW* 997 contained the highest overall sensitivity and specificity estimates for each category, yet underperformed in singleplex and exhibited cross reactivity in multiplex. The MOLigo pair for O45 *wbhW* 997 was not included in the test panel. The SNP for O45 *wbhQ* 721 contained the next overall highest sensitivity and specificity estimates for each category and was chosen to be evaluated in the probe panel. O45 *wbhQ* 721 increased in performance and cross-reactivity was eliminated upon the incorporation of this MOLigo pair into the probe panel.

4.2 Assay Development with Training Panel

The original MOL-PCR protocol included ligation and amplification in a single tube reaction.[36] The SNP based MOL-PCR protocol separated the ligation and amplification reactions to improve assay performance. The STEC and non-STEC MOL-PCR assays were originally evaluated as a 2-plex (O26 *fnl1* and *wzx*) against a training panel consisting of complementary synthetic target sequences (Appendix C). The training panel was used to optimize the ligation and PCR conditions for both MOL-PCR assays, as a result of poor signal-to-noise ratios when initially evaluating the assays against STEC DNA samples. The objective of this experiment was to evaluate the specificity of the MOLigo pairs to their intended target when evaluated in multiplex format. A ligation control was incorporated into this assay to ensure that the ligation reaction was performing. The assay was evaluated in quadruplicate for both samples and ligation controls.

Table 4.1 shows the ability for the O26 MOLigo pairs designed for both the STEC and non-STEC assays to correctly identify their synthetic complementary targets. The MFI values for this assay exhibit inconsistencies between replicates that may be attributed to issues with transferring the ligation product to the amplification reaction. The O26 STEC *wzx* MOLigo pair identifies the complementary target in the first and the fourth replicate, with second and third replicates failing to identify the target. The O26 STEC *fnl1* MOLigo identifies the complementary target in each replicate. The fourth replicate for O26 STEC *fnl1* exhibits cross reactivity between O26 STEC *fnl1* and *wzx* genes that may be the result of cross-reactivity between samples. The non-STEC MOLigos for both genes correctly identify each of their intended targets. Overall, the experiment shows the ability for each MOLigo pair to identify their complementary synthetic target when evaluated in multiplex. Further optimization of the assay is required to improve inconsistencies between replicates.

Sample	O26 S-wzx	O26 S-fnl1	O26 NS-wzx	O26 NS-fnl1	Positive Control
O26 wzx STEC Synthetic Template	30165	1071	1283	3085	1628
O26 wzx STEC Synthetic Template	2597	1077	1236	2303	1019
O26 wzx STEC Synthetic Template	2353	1404	1430	1981	1154
O26 wzx STEC Synthetic Template	16244	1001	706	1666	715
O26 fnl1 STEC Synthetic Template	3496	73132	1532	1383	928
O26 fnl1 STEC Synthetic Template	1777	43073	594	1435	885
O26 fnl1 STEC Synthetic Template	1160	14556	896	1317	1081
O26 fnl1 STEC Synthetic Template	32327	62278	580	921	929
O26 wzx non-STEC Synthetic Template	1400	1483	67641	1158	868
O26 wzx non-STEC Synthetic Template	1400	1318	23021	1317	1091
O26 wzx non-STEC Synthetic Template	1348	1306	4911	1435	865
O26 wzx non-STEC Synthetic Template	1281	1123	63303	1335	893
O26 fnl1 non-STEC Synthetic Template	1013	1440	1087	34368	1291
O26 fnl1 non-STEC Synthetic Template	3597	1251	1283	53504	951
O26 fnl1 non-STEC Synthetic Template	651	1389	1166	29066	705
O26 fnl1 non-STEC Synthetic Template	1996	703	719	35403	692
Positive Control	1270	736	962	1704	92504
Positive Control	2217	1057	668	1209	69227
Positive Control	2933	3277	764	938	66031
Positive Control	3326	5250	4351	3799	100922

Table 4.1: Synthetic target sequences were designed for assay development. Both STEC and non-STEC synthetic target sequences were designed for O26 genes *wzx* and *fnl1*. The STEC and non-STEC MOLigos were evaluated separately, consisting of the interrogation of the assay in multiplex (2-plex). The MOLigo pairs accurately identify their respective targets.

4.3 MOLigo Pairs Evaluated in Singleplex

The STEC non-O157 assay was evaluated as a singleplex assay to determine specificity of the MOLigo pairs to their intended targets. In this experiment, MOLigo pairs containing the STEC-associated allele were evaluated in singleplex by running multiple MOL-PCR assays in parallel. The probe panel consisted of purified genomic DNA for each non-O157 serogroup (Appendix D.1). The objective of this experiment was to evaluate the specificity of the STEC MOLigo pairs against their intended target in the absence of the other non-O157 MOLigo pairs.

The results in Table 4.2 show the ability for each STEC non-O157 MOLigo pair to correctly identify their intended target. The MFI values for O26 *wzx* and O45 *wbhW* identified the sample correctly, yet the MFI values are considerably lower than the other STEC non-O157 MOLigos. The virulence profile identified for BAA-2196 indicate the

strain is O26 positive and contains both *stx*1 and *stx*2 (Appendix D.1). The MOLigo pair for *wzx* correctly identifies the target sequence, unfortunately the signal-to-noise ratio is only slightly higher than the background. In subsequent experiments, O45 *wbhW* was eliminated from the MOLigo probe panel and replaced with O45 *wbhQ* due to poor MOLigo performance. This study provides insight as to how each MOLigo pair was performing for the STEC non-O157 serogroups. The range of MFI values between each samples is highly variable and could be a result in the variation of the bead counts associated with each MOLigo pair. The non-STEC assay was not evaluated in singleplex, due to absence of non-STEC strains.

Sample	O26 S-wzx	O26 S-fnl1	O45 S-wbhW	O103 S-wbtD	O111 S-wbdK	O121 S-vioA	O145 S-wzy	Ligation Control
O26 BAA-2196	2548	674	593	903	696	562	501	570
O26 BAA-2196	1009	64333	391	821	877	551	587	608
O45 BAA-2193	754	456	5664	706	637	519	671	292
O103 BAA-2215	868	770	800	42905	644	697	810	595
O111 BAA-2440	555	596	521	1778	39632	734	1966	738
O121 BAA-2219	690	536	434	1025	2308	87104	701	296
O145 BAA-2192	759	907	456	555	810	1875	29759	693
Ligation Control	829	320	406	822	905	612	772	98,394

Table 4.2: STEC MOLigo pairs evaluated in singleplex format. Experiment designed to evaluate MOLigo pair specificity for each non-O157 serogroup.

4.4 STEC non-O157 Assay Evaluated in Multiplex

The STEC non-O157 assay was evaluated as an 8-plex that consisted of the MOLigo pairs containing the seven STEC-associated alleles and an internal ligation control. The internal ligation control was incorporated into the assay to ensure the ligation reactions were working proficiently. A set of MOLigo pairs were designed for the ligation control to target a complementary synthetic oligonucleotide strand. The STEC MOLigo pairs designed for each assay were evaluated against an *E.coli* DNA panel that consisted of 6 known STEC strains. The validation of the 8-plex STEC assay examined each serotype identified in Appendix D.2 in triplicate.

Initial testing for the STEC non-O157 MOL-PCR assay is presented in Table 4.3. The PCR protocol initially included deoxythymidine triphosphate (dTTPs) incorporated into the dNTP mix for the PCR reaction. The MOLigo pair for O26 *wzx* identifies each strain as a positive for O26 *wzx*. This is a result of MOLigo probe contamination with the synthetic complementary target for O26 *wzx.* This assay also exhibits cross reactivity for strain CDC 96-3285 for each non-O157 MOLigo pair. The SNP for O45 *wbhW* was located near the end of the gene; therefore, the MOLigo pair designed was intergenic which may have contributed to the cross-reactivity seen in this experiment. Based on this experiment, O45 *wbhW* was redesigned and replaced with O45 *wbhQ*. Besides the synthetic target contamination, each of the STEC non-O157 MOLigo pairs correctly identify their intended targets. Due to the contamination, MOLigo pairs were redesigned and reordered for subsequent experiments and incorporation of UDG and deoxyuridine triphosphate (dUTPs) were added to reduce PCR contamination (Appendix E: Original MOLigo Design).

Strain ID	O26 S-wzx	O26 S-fnl1	O45 S-wbhW	O103 S-wbtD	O111 S-wbdK	O121 S-vioA	O145 S-wzy	Ligation Control
H30	49305	48101	723	963	870	661	1010	639
CDC 96-3285	33367	24224	48905	26060	20447	10052	28929	556
CDC 90-3128	65121	5982	5008	91851	4118	1685	5370	1,193
JB1-95	80781	909	1164	5267	55661	1113	1838	813
CDC 97-3068	101165	1503	1670	3738	3682	96254	2861	1,017
83-75	88282	956	1279	1866	1631	2672	60644	814
Ligation Control	2961	1627	1360	2606	2427	1493	3579	136296

Table 4.3: Initial testing for STEC non-O157 8-plex assay evaluated against DNA STEC panel.

MOLigo pairs were redesigned for O26 *wzx*, O103 *wbtD*, O111 *wbdK*, and O145 *wzy* to incorporate anti-tag sequences for Luminex xTAG® microspheres that were associated with the Luminex MagPix®. The PCR protocol was changed to replace dTTPs with dUTPs, to reduce carry-over PCR contamination (described in Appendix F). The STEC non-O157 assay was evaluated in multiplex against a STEC DNA panel from the STEC CAP collaborators. The STEC MOLigos were assembled into an 8-plex that

included the seven MOLigo pairs for the non-O157 serogroups and an internal control. A full length O26 oligonucleotide sequence was also included to evaluate PCR efficiency. The ligation and PCR reaction were performed in a 10 µL reaction.

Table 4.4 shows the results for the newly designed MOLigo pairs and the incorporation of both UDG and dUTPs into both the ligation and PCR reactions. The MFI value for the O103 MOLigo pair was considerably lower in comparison to the other STEC MOLigo pairs. This may be attributed to the addition of dUTPs in the PCR reaction to combat PCR carry-over contamination. The O103 MOLigo pair evaluated in singleplex and the initial evaluation of the 8-plex assay consisted of dTTPs in the PCR reaction and performed substantially better than the reactions that included dUTPs (Table 4.2 and Table 4.3). The assay exhibits cross-reactivity when evaluating the strains JB1-95, CDC 97-3068, and 83-75. This is the result of cross-reactivity between samples and is not indicative of cross-reactivity between MOLigo pairs. The negative strand for the O103 MOLigo pair is currently undergoing optimization to possibly replace the current positive strand for the O103 MOLigo pair.

Sample	O26 S-wzx	O26 S-fnl1	O45 S-wbhQ	O103 S-wbtD	O111 S-wbdK	O121 S-vioA	O145 S-wzy	PCR Control	Ligation Control
H30	378	2701	62	64	60	68	266	71	61
CDC 96-3285	65	62	4628	56	138	62	59	64	57
CDC 90-3128	86	78	84	218	76	77	81	85	73
JB1-95	64	45	1275	39	1916	65	378	42	39
CDC 97-3068	306	57	807	55	51	1956	53	55	49
83-75	77	78	79	75	77	1266	4534	82	69
PCR Control	123	113	122	108	119	103	118	7837	102
Ligation Control	92	97	110	101	91	90	99	95	7878

Table 4.4: Results for the STEC MOL-PCR assay evaluated against STEC DNA panel. The results of the 8-plex STEC MOL-PCR assay evaluated against known STEC strains. The results included here are indicative of single replicates that made correct calls for a single sample set.

Optimization of the reaction has continued with the evaluation of the assay in a 20 µL reaction. The assay was exhibiting inconsistencies between sample replicates and the no template control. Inconsistencies between replicates for the 10 µL reaction is most likely

associated with evaporation during ligation and PCR reactions. Table 4.5 shows the initial evaluation of the 20 µL reaction. Increasing the reaction volume increased the MFI values for O103 *wbtD* in this reaction. The MOLigos correctly identify each of their intended targets. This experiment shows cross-reactivity across each sample, with the exception of the ligation and PCR control. The spurious amplification between samples is believed to be contributed to sample transfer either from ligation to amplification reactions or possibly aerosol contamination when adding xTAG® microspheres to the hybridization reaction. The ligation control is positive in samples JB1-95, CDC 97-3068, and 83-75 that may also be a result of cross-reactivity between wells since it is not seen in the first three samples. Separating the ligation and amplification reactions creates an open system that opens many avenues to potential contamination.

Strain ID	O26 S-wzx	O26 S-fnlI	O45 S-wbhQ	O103 S-wbtD	O111 S-wbdK	O121 S-vioA	O145 S-wzy	PCR Control	Ligation Control
H30	173	2776	30	615	44	195	285	21	23
CDC 96-3285	169	43	3465	228	568	222	1241	39	254
CDC 90-3128	193	23	21	1014	168	313	395	17	31
JB1-95	251	27	25	107	2166	381	332	23	1475
CDC 97-3068	297	34	933	727	68	3402	605	31	893
83-75	154	35	29	305	45	967	1362	29	1979
Ligation Control	60	67	70	64	57	64	71	7642	57
PCR Control	61	63	60	62	69	60	60	65	8593

Table 4.5: STEC MOL-PCR assay evaluated in a 20 µL reaction against STEC DNA panel.

4.5 non-STEC Multiplex Assay

The non-STEC assay was evaluated as an 8-plex that consisted of the MOLigo pairs containing the non-STEC associated alleles and an internal ligation control. The non-STEC assay was examined against STEC samples provided by collaborators (Appendix D.2). The purpose of this experiment was to evaluate the specificity of the SNP identified for each non-O157 serogroup. Specificity estimates were defined by Norman et al. as the probability for the non-STEC associated allele to be unable to detect the STEC-associated allele.[26] To determine specificity of the non-STEC associated allele, the non-STEC MOLigo pairs were

evaluated against strains that contained the STEC-associated alleles. The internal ligation and PCR controls were included in this reaction. The ligation control consists of a set of MOLigo pairs that proceed through the ligation reaction, and are indicative that the ligation reaction is working. The full length O26 PCR control is a single stranded oligonucleotide that proceeds through amplification and serves as control to determine PCR efficiency. Therefore, the 8-plex non-STEC MOL-PCR was evaluated to determine both the specificity of the SNP and how efficiently the MOL-PCR assay chemistry could discriminate against the STEC-associated allele.

Table 4.6 shows the results for the non-O157 specificity experiment. The ligation control and the full length O26 PCR control exhibited high signal-to-noise ratios and were indicative of both the ligation and PCR reactions working accordingly. The non-STEC MOL-PCR assay did not identify the STEC strains, indicating high specificity for the SNP associated with non-STEC serogroups. The results in Table 4.6 show the ability for the non-STEC MOLigo pairs to accurately discriminate against the STEC-associated alleles. Therefore, the MOL-PCR assay chemistry has the ability to accurately discriminate between strains that differ by a single base change. To ensure the non-STEC MOLigo pairs are performing optimally, the non-STEC assay will need to be evaluated against non-STEC strains.

Sample	O26 NS-wzx	O26 NS-fnl1	O45 NS-wbhQ	O103 NS-wbtD	O111 NS-wbdK	O121 NS-vioA	O145 NS-wzy	PCR Control	Ligation Control
O26 KSU	954	495	1197	472	179	542	1054	23	1198
O45 KSU	221	256	1417	507	149	1252	1108	20	595
O103 KSU	173	233	934	322	143	657	1080	17	386
O111 KSU	130	159	1006	330	253	404	1138	18	725
O121 KSU	97	172	874	372	161	482	788	22	156
O145 KSU	136	309	1268	555	202	541	1546	21	345
PCR Control	48	47	50	54	51	56	50	3311	49
Ligation Control	30	38	240	49	43	74	75	39	5170

Table 4 6: Results for non-STEC SNP specificity assay evaluated against known STEC targets. The non-STEC MOL-PCR assay was evaluated against STEC strains. The results confirm specificity estimates for the non-STEC MOLigos to discriminate against the STEC-associated allele. The ligation and PCR control were included to confirm the reaction was performing optimally.

Chapter 5: Summary of Work and Future Directions

5.1 Summary of Work

Non-O157 STEC strains associated with human illness have increased in frequency and are considered a causative agent for hemorrhagic colitis and hemolytic uremic syndrome.[3] Determining the prevalence for non-O157 STEC strains associated with human illness is limited due to the absence of mandatory characterization of isolates and diagnostic tools that can detect and characterize non-O157 strains.[26] With the implementation of mandatory testing for O26, O45, O103, O111, O121, and O145, single nucleotide polymorphisms have been identified for each serogroup that can discriminate between a STEC and a non-STEC strain.[26] To overcome limitations of current diagnostic methods, a nucleic acid-based detection platform that incorporates the SNPs recently identified was successfully designed for the detection and characterization of non-O157 strains.

5.1.1 Development of MOL-PCR SNP Based Detection Assay

In this study, a high throughput MOL-PCR SNP based assay was successfully designed to detect and characterize non-O157 strains. The MOL-PCR SNP based assay offers improvement over conventional non-O157 PCR detection systems by incorporating SNPs into the MOLigo target detection probes. The SNPs are serogroup specific and can differentiate between a STEC and non-STEC strain. Therefore, eliminating the need for multiple target detection probes to detect the O-serogroup and virulence genes.

The computationally designed MOLigo pairs were able to accurately identify the O-serogroup and characterize the strain based on the STEC-associated alleles. STEC and non-STEC MOLigo pairs for O26 genes *fnl1* and *wzx*, were initially evaluated against synthetic complementary target sequences. Synthetic complementary target sequences

were designed to determine ligation and PCR conditions due to poor signal-to-noise ratios when evaluated against a STEC DNA panel. The O26 MOLigo pairs have demonstrated the capacity to accurately identify their complementary synthetic targets.

The STEC non-O157 assay was successfully developed and tested in both singleplex and multiplex format. In Chapter 4, the STEC non-O157 MOLigo pairs accurately identify each target when evaluated in singleplex. The singleplex assay determined which MOLigo pairs were under performing and provided insight as to which MOLigo pairs would need to be redesigned. Furthermore, the STEC non-O157 assay was evaluated as an 8-plex, consisting of MOLigo pairs for the 7 STEC-associated alleles identified for the non-O157 serogroups and an internal ligation control. The 8-plex assay identified each strain in the STEC DNA panel, yet the assay exhibited cross-reactivity between samples. Also, the MFI values for the O103 MOLigo pair are substantially lower than the other MOLigo pairs interrogated. The SNP identified for O103 is unfortunately the only SNP identified for this particular serogroup and could not be redesigned. To improve signal-to-noise ratios for O103, a titration of concentrations would need to be performed to determine if increasing the concentration of O103 MOLigo will increase the performance of this MOLigo pair. The STEC non-O157 assay is not performing consistently and further optimization and validation is needed.

The non-STEC assay was evaluated against known STEC strains. The assay was evaluated as an 8-plex that included the MOLigo pairs containing the non-STEC associated allele and an internal ligation control. The objective of this experiment was to test the specificity of the non-STEC MOLigo pairs against STEC strains. This experiment was performed to ensure that the non-STEC MOLigos were not producing false positives when

evaluated against a strain that differed by a single base. Therefore, the non-STEC MOLigos were evaluated against known STEC strains to determine both the specificity of the non-STEC associated allele and also to determine if the MOL-PCR assay chemistry was accurately discriminating against the STEC associated allele. The non-STEC MOLigo pairs accurately discriminate against the STEC-associated allele; in addition, this experiment provided further confidence in the ability for the MOL-PCR assay chemistry to accurately differentiate between strains that differ by a single base change. The non-STEC assay was not evaluated against non-STEC strains; thus, it cannot be confirmed from this study that the assay can correctly identify the non-O157 serogroups for a non-STEC strain.

5.2 Future Directions

The work presented in this thesis provides a potential diagnostic platform that can accurately detect and characterize non-O157 strains. The design and initial testing provide the framework for future validation of the SNP based MOL-PCR assays. The assays developed here can contribute to better understanding of true prevalence estimates for non-O157 strains in cattle and raw non-intact beef. Summarized in this section are future potential directions that will contribute to the overall improvement of the MOL-PCR assay.

5.2.1 Optimization of the STEC non-O157 Assay

Although the STEC non-O157 assay has the ability to correctly identify each of their intended targets, further optimization of reaction conditions is necessary. The negative strand for the O103 MOLigo pair is underperforming upon the incorporation of dUTPs. The O103 MOLigo pair contains a homopolymer of 9 thymines located on MOLigo1 and it appears that the assay may be experiencing PCR inefficiency with the present

concentration of the dNTP/dUTP mix. A solution to this problem would be to purchase dNTPs in a solution set rather than a mix, and possibly increase the concentration of dATPs to see if the signal-to-noise ratios will increase for O103. The positive strand for O103 has been evaluated for the O103 MOLigo pair; unfortunately, the strand has not shown improvement over the negative strand. The last option would be to incorporate dTTPs back into the PCR reaction, which the assay has previously worked during initial evaluation of the singleplex and the 8-plex assays. Stringent standard operation procedures would have to be followed to prevent PCR contamination (Appendix G). Table 5.1 shows the effect of the incorporation of dUTPs into the MOL-PCR SNP based assay. Furthermore, reproducibility between replicates and between operators was found to be poor. Determining optimal ligation and PCR conditions should reduce inconsistencies between sample replicates.

A) 7-plex with dTTPs in PCR reaction

Strain ID	O26 S-wzx	O26 S-fnl1	O45 S-wbhW	O103 S-wbtD	O111 S-wbdK	O121 S-vioA	O145 S-wzy	Ligation Control
H30	64959	62824	723	918	1294	1176	1112	1444
CDC 96-3285	39490	37303	83355	39762	34983	20748	45324	1337
CDC 90-3128	51677	2893	2492	79266	3048	1334	5347	751
JB1-95	79216	3533	5890	9498	51150	2269	4144	1161
CDC 97-3068	57008	3504	3024	2306	8170	33126	2923	1144
83-75	84647	2220	926	1808	1625	1988	54066	901
Ligation Control	1808	1365	713	1224	1245	1102	2041	98977

B) 8-plex with dUTPs in PCR reaction

Strain ID	O26 S-wzx	O26 S-fnl1	O45 S-wbhW	O103 S-wbtD	O111 S-wbdK	O121 S-vioA	O145 S-wzy	Ligation Control
H30	129512	267171	1958	2732	3128	1796	2890	154,408
CDC 96-3285	119684	16994	2881	4593	3172	2615	3247	165,281
CDC 90-3128	148638	5935	2382	5056	40134	10803	11202	354,537
JB1-95	100493	2164	1902	5701	207756	1933	3297	67664
CDC 97-3068	98984	2566	3285	5215	4983	323927	3859	69560
83-75	125641	6411	5826	7343	14615	13434	123856	258156
Ligation Control	28238	3602	2979	3451	4911	6949	9688	328160

Table 5.1: A) Assay evaluated as a 7-plex with the ligation control separated from the non-O157 MOLigos. The PCR reaction included dTTPs and also included the O45 *wbhW* gene before redesigning. O45 *wbhw* exhibits cross reactivity and was redesigned in subsequent experiments. B) The assay was evaluated as an 8-plex assay. The PCR reaction incorporates dUTPs into the reaction and shows a drastic reduction in signal for O103 *wbtD*. The O26 S-*wzx* MOLigo was positive in each sample for both assays due to synthetic O26 *wzx* DNA contamination.

5.2.2 Validation of STEC non-O157 Assay

Once the optimization for the STEC non-O157 has been achieved, validation of the STEC non-O157 assay will need to be assessed to determine robustness of the assay. Currently, the 8-plex STEC assay has only been evaluated against 6 STEC strains. This assay will need to be validated against a STEC DNA panel that consists of 220 samples from ATCC, Kansas State University (KSU), University of Nebraska-Lincoln (UNL), and the United States Department of Agriculture (USDA). An exclusivity panel consisting of 20 organisms that are distantly related to non-O157 strains will also need to be included to determine selectivity estimates for this assay. Moreover, the limit of detection (LOD) will need to be determined to better understand the minimum amount of nucleic acid copies that can be detected for this assay. With the validation of this assay and the determination of the LOD, the next step will be to evaluate the STEC MOL-PCR assay against spiked fecal samples provided by KSU. These particular samples will allow us to test the STEC MOL-PCR assay against complex samples that would mimic samples found in a clinical laboratory setting.

5.2.3 Optimization and Validation of the non-STEC Assay

The non-STEC MOL-PCR assay was evaluated against 6 known STEC strains. To determine if the assay is working optimally, the assay will need to be evaluated against non-STEC strains. The USDA is providing us with 13 non-STEC *E.coli* strains. The non-STEC MOLigo pairs for each serogroup will be evaluated against the *E.coli* strains provided to determine the specificity of the MOLigo pairs to their intended targets. This experiment will determine if the assay is optimized or if further optimization is needed. The strains provided will also allow us to determine LOD for the non-STEC assay.

5.2.4 MOL-PCR Incorporating Dual-Priming Oligonucleotides (DPO)

MOL-PCR assay chemistry offers numerous advantages over conventional and real-time PCR technologies. A potential future direction that could help overcome cross-reactivity and increase allelic discrimination for the STEC and non-STEC assays is the incorporation of a dual priming oligonucleotide system. This provides a unique modification to MOL-PCR assay chemistry that has been explored in previous studies for the interrogation of SNPs.[43] The introduction of the DPO system is incorporated into MOLigo2 containing the SNP discriminate and is designed to increase allelic specificity.[43] The DPO system separates MOLigo2 into two primer segments that are joined together by a polydeoxyinosine [poly (I) linker], typically consisting of five deoxyinosine (I) bases (Figure 5.1).[42] In comparison to the natural bases, deoxyinosine has a lower melting temperature and forms a "bubble-like" structure separating MOLigo2 into two distinct priming regions.[42-43]

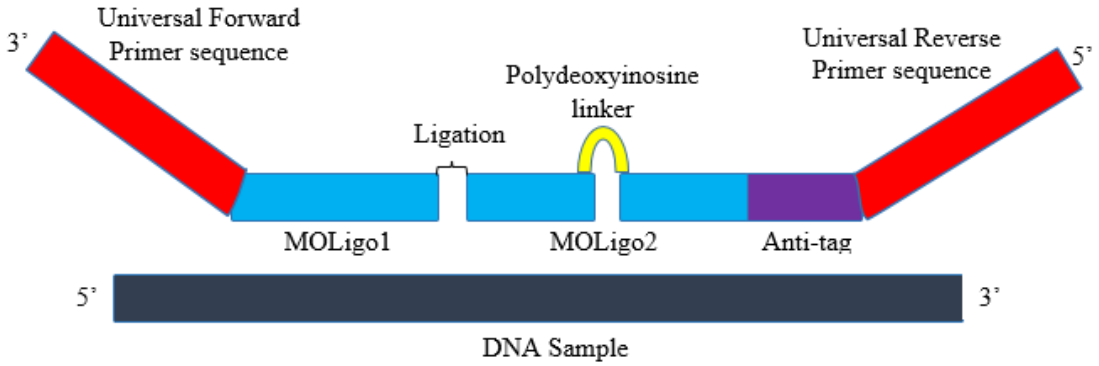


Figure 5.1: Modified modular components for MOL-PCR with integrated DPO system. Polydeoxyinosine linker incorporated into MOLigo2 design would reduce cross reactivity and increase allelic specificity.[42] MOLigo2 is separated into two components including; a 5'segment consisting of 11-23 nucleotides that allows for stable annealing and a 3'-segement consisting of 6-12 nucleotides that contains the SNP discriminate and determines target-specific extension.[42]

The design of MOLigo2 consists of a two separate regions; a 5'-segment (11-23 nucleotides in length) that binds to the target DNA site for stable annealing and a 3'-segment (6-12 nucleotides in length) that binds to the target DNA. Target-specific

extension will only occur if both segments of MOLigo2 bind to the target DNA; therefore, decreasing the extension of non-specific products and increase PCR specificity. This area of research is very interesting and would add to the increased specificity of the MOL-PCR SNP based assay. Also, the DPO system may allow for the STEC and non-STEC assay to be combined into a 15-plex assay that would contain the MOLigo pairs for both non-O157 allelic discriminates and the internal ligation control. Therefore, eliminating the need to evaluate two assays separately and ultimately reducing the cost of reagents.

5.2.5 Prevalence Studies of STEC non-O157 Associated with Cattle

Many outbreaks in the United States have been associated with beef products containing *E.coli* O157: H7.[3, 17] Epidemiological studies have mainly focused on the prevalence of *E.coli* O157:H7, with limited understanding on true prevalence estimates for non-O157 serogroups.[26] The development of the non-O157 MOL-PCR assay provides the framework for future prevalence studies of non-O157 STEC in cattle and raw non-intact beef. Better understanding of the prevalence of non-O157 STEC must be evaluated from pre-harvest to post-harvest.

Pre-harvest prevalence studies are important for decreasing non-O157 STEC strains in cattle feces and on cattle hides.[44] Prevalence studies could include how interventions such as probiotics, vaccines, and antimicrobials effect fecal shedding of non-O157 in feces.[45] It would be interesting to understand how these interventions influence seasonal shedding, with summer months often contributing to an increase in prevalence rates of non-O157 STEC in feces.[21] Cattle hides often test positive for O157:H7 and non-O157 STEC and are a potential source for contaminating carcasses during slaughter. [21, 44] The SNP based MOL-PCR assay could be used to better understand contaminants present

on hides, providing prevalence estimates for both STEC and non-STEC strains present on the cattle hides.

Post-harvest intervention are vital to ensure contaminated beef products are not distributed to the public. Highly variable prevalence estimates for non-O157 STEC are often reported consisting of rates between 2.1 to 70.1%.[44] STEC contamination usually results from removal of the hide or the gastrointestinal tract.[44] Many variables can influence prevalence rates including; STEC detection methods, pre-harvest intervention, and management practices.[44]

The assay presented in this research could aid in estimation of true prevalence of non-O157 by accurately differentiating between a STEC and non-STEC strain. The work presented here provides the framework for future studies that could aid in the understanding of the prevalence estimates for non-O157 STEC strains and provide preventative measures to ensure the safety of the consumer.

References

1. Bergan J, Lingelem ABD, Simm R, Skotland T, Sandvig K. Shiga toxins. Toxicon. 2012;60(6):1085-107.

2. Johannes L, Roemer W. Shiga toxins - from cell biology to biomedical applications. Nature Reviews Microbiology. 2010;8(2):105-16.

3. Nataro JP, Kaper JB. Diarrheagenic *Escherichia* coli. Clinical Microbiology Reviews. 1998;11(1):142-+.

4. Fagerquist CK, Sultan O. Top-Down Proteomic Identification of Furin-Cleaved alpha-Subunit of Shiga Toxin 2 from *Escherichia* coli O157:H7 Using MALDI-TOF-TOF-MS/MS. Journal of Biomedicine and Biotechnology. 2010.

5. Fraser ME, Fujinaga M, Cherney MM, Melton-Celsa AR, Twiddy EM, O'Brien AD, et al. Structure of Shiga toxin type 2 (Stx2) from *Escherichia* coli O157 : H7. Journal of Biological Chemistry. 2004;279(26):27511-7.

6. Garred O, Dubinina E, Polesskaya A, Olsnes S, Kozlov J, Sandvig K. Role of the disulfide bond in Shiga toxin A-chain for toxin entry into cells. Journal of Biological Chemistry. 1997;272(17):11414-9.

7. Garred O, Vandeurs B, Sandvig K. Furin-Induced Cleavage and Activation of Shiga Toxin. Journal of Biological Chemistry. 1995;270(18):10817-21.

8. Fraser ME, Chernaia MM, Kozlov YV, James MNG. Crytsal-Structure of the Holotoxin From *Shigella*-Dysenteriae at 2.5-Angstrom Resolution. Nature Structural Biology. 1994;1(1):59-64.

9. Ling H, Boodhoo A, Hazes B, Cummings MD, Armstrong GD, Brunton JL, et al. 1BOS: Shiga-Like Toxin Complexed with its Receptor. Protein Data Bank. 1999;3.8.

10. Sthilaire PM, Boyd MK, Toone EJ. Interaction of the Shiga-Like Toxin Type-1 B-Subunit with Its Carbohydrate Receptor. Biochemistry. 1994;33(48):14452-63.

11. Gyles CL. Shiga toxin-producing *Escherichia* coli: An overview. Journal of Animal Science. 2007;85:E45-E62.

12. Strockbine NA, Jackson MP, Sung LM, Holmes RK, Obrien AD. Cloning and Sequencing of The Genes For Shiga Toxin From *Shigella*-Dysenteriae Type-1. Journal of Bacteriology. 1988;170(3):1116-22.

13. Degrandis S, Ginsberg J, Toone M, Climie S, Friesen J, Brunton J. Nucleotide-Sequence and Promoter Mapping of the *Escherichia*-coli Shiga-Like Toxin Operon of Bacteriophage-H-19B. Journal of Bacteriology. 1987;169(9):4313-9.

14. Jackson MP, Neill RJ, Obrien AD, Holmes RK, Newland JW. Nucleotide-Sequence Analysis and Comparison of the Structural Genes for Shiga-Like Toxin-I and Shiga-Like Toxin-II Encoded by Bacteriophages from *Escherichia*-Coli 933. Fems Microbiology Letters. 1987;44(1):109-14.

15. Johnson KE, Thorpe CM, Sears CL. The emerging clinical importance of non-O157 Shiga toxin - Producing *Escherichia* coli. Clinical Infectious Diseases. 2006;43(12):1587-95.

16. Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, et al. Food-related illness and death in the United States. Emerging Infectious Diseases. 1999;5(5):607-25.

17. Brooks JT, Sowers EG, Wells JG, Greene KD, Griffin PM, Hoekstra RM, et al. Non-O157 Shiga toxin-producing *Escherichia* coli infections in the United States, 1983-2002. Journal of Infectious Diseases. 2005;192(8):1422-9.

18. Frenzen PD, Drake A, Angulo FJ, Emerging Infections P. Economic cost of illness due to *Escherichia* coli O157 infections in the United States. Journal of Food Protection. 2005;68(12):2623-30.

19. Caprioli A, Morabito S, Brugere H, Oswald E. Enterohaemorrhagic *Escherichia* coli: emerging issues on virulence and modes of transmission. Veterinary Research. 2005;36(3):289-311.

20. Zhao T, Doyle MP, Shere J, Garber L. Prevalence of Enterohemorrhagic *Escherichia*-coli O157-H7 in a Survey of Dairy Herds. Applied and Environmental Microbiology. 1995;61(4):1290-3.

21. Barkocy-Gallagher GA, Arthur TM, Rivera-Betancourt M, Nou XW, Shackelford SD, Wheeler TL, et al. Seasonal prevalence of Shiga toxin-producing *Escherichia* coli, including O157 : H7 and non-O157 serotypes, and Salmonella in commercial beef processing plants. Journal of Food Protection. 2003;66(11):1978-86.

22. Verweyen HM, Karch H, Brandis M, Zimmerhackl LB. Enterohemorrhagic *Escherichia* coli infections: following transmission routes. Pediatric Nephrology. 2000;14(1):73-83.

23. Mead PS, Finelli L, LambertFair MA, Champ D, Townes J, Hutwagner L, et al. Risk factors for sporadic infection with *Escherichia* coli O157:H7. Archives of Internal Medicine. 1997;157(2):204-8.

24. D'Souza JA, Wang L, Reeves P. Sequence of the *Escherichia* coli O26O antigen gene cluster and identification of O26 specific genes. Gene. 2002;297(1-2):123-7.

25. Erridge C, Bennett-Guerrero E, Poxton IR. Structure and function of lipopolysaccharides. Microbes and Infection. 2002;4(8):837-51.

26. Norman KN, Strockbine NA, Bono JL. Association of Nucleotide Polymorphisms within the O-Antigen Gene Cluster of *Escherichia* coli O26, O45, O103, O111, O121, and O145 with Serogroups and Genetic Subtypes. Applied and Environmental Microbiology. 2012;78(18):6689-703.

27. Paton J, Paton A. Methods for detection of STEC in humans. E. coli.2003 (73) 9-26.

28. De Boer E, Heuvelink AE. Methods for the detection and isolation of Shiga toxin-producing *Escherichia* coli. Journal of Applied Microbiology. 2000;88:133S-43S.

29. Deshpande A, White PS. Multiplexed nucleic acid-based assays for molecular diagnostics of human disease. Expert Review of Molecular Diagnostics. 2012;12(6):645-59.

30. Paton AW, Paton JC. Direct detection and characterization of Shiga toxigenic *Escherichia* coli by multiplex PCR for stx(1), stx(2), eae, ehxA, and saa. Journal of Clinical Microbiology. 2002;40(1):271-4.

31. Mark JA, Green LD, Deshpande A, White PS, editors. System integration and development for biological warfare agent surveillance - art. no. 65401D. Conference on Optics and Photonics in Global Homeland Security III; 2007 Apr 10-12; Orlando, FL2007.

32. Klein D. Quantification using real-time PCR technology: applications and limitations. Trends in Molecular Medicine. 2002;8(6):257-60.

33. Sharma VK, Dean-Nystrom EA. Detection of enterohemorrhagic *Escherichia* coli O157 : H7 by using a multiplex real-time PCR assay for genes encoding intimin and Shiga toxins. Veterinary Microbiology. 2003;93(3):247-60.

34. Guy RA, Tremblay D, Beausoleil L, Harel J, Champagne MJ. Quantification of E. coli O157 and STEC in feces of farm animals using direct multiplex real time PCR

(qPCR) and a modified most probable number assay comprised of immunomagnetic bead separation and qPCR detection. Journal of Microbiological Methods. 2014;99:44-53.

35. Biassoni R, Raso A. Quantitative Real-Time PCR Methods and Protocols Preface. Quantitative Real-Time Pcr: Methods and Protocols. 2014;1160:V-V.

36. Deshpande A, Gans J, Graves SW, Green L, Taylor L, Kim HB, et al. A rapid multiplex assay for nucleic acid-based diagnostics. Journal of Microbiological Methods. 2010;80(2):155-63.

37. Woods TA, Mendez HM, Graves SW, Deshpande A. Development of 11-plex Assay for the Rapid Screening of Samples for Detection of Shiga toxin-producing E.coli. [Unpublished]

38. Federal Registar. Shiga toxin-producing *Escherichia* coli in certain raw beef products. 2011;76:58157-58165.

39. Stucki D, Malla B, Hostettler S, Huna T, Feldmann J, Yeboah-Manu D, et al. Two New Rapid SNP-Typing Methods for Classifying Mycobacterium tuberculosis Complex into the Main Phylogenetic Lineages. Plos One. 2012;7(7).

40. MOLigoDesigner [internet]; Los Alamos National Laboratory; 2008 [cited 2015 November 1st]. Available from: http://moligodesigner.lanl.gov/help.html

41. Salkind N. Encyclopedia of Research Design. Thousand Oaks. Sage Publications, Inc.; 2010.

42. Thierry S, Hamidjaja R, Girault G, Lofstrom C, et al. A multiplex bead-based suspension array for interrogation of phylogenetically informative single nucleotide polymorphisms for *Bacillus anthracis*. Journal of Microbiological Methods. 2013:95(3):357-365.

43. Chun JY, Kim KJ, Hwang IT, Kim YJ, Lee DH, Lee IK, et al. Dual priming oligonucleotide system for the multiplex detection of respiratory viruses and SNP genotyping of CYP2C19 gene. Nucleic Acids Research. 2007;35(6).

44. Hussein HS. Prevalence and pathogenicity of Shiga toxin-producing *Escherichia* coli in beef cattle and their products. Journal of Animal Science. 2007;85:E63-E72.

45. Sargeant JM, Amezcua MR, Rajic A, Waddell L. Pre-harvest interventions to reduce the shedding of E-Coli O157 in the faeces of weaned domestic ruminants: A systematic review. Zoonoses and Public Health. 2007;54(6-7):260-77.

46. LuminexCorp [internet]. Madison (WI): Equivalent Analytical Performance between the New MAGPIX® System and the Luminex® 100/200™ System. [White Paper] [cited 2015 November 1st]. Available from: https://www.luminexcorp.com/clinical/instruments/magpix/resources/

47. LuminexCorp [internet]; 2006[cited 2015 November 1st]. Available from: https://www.luminexcorp.com/clinical/instruments/magpix/product-details/

48. BD Biosciences [internet]; 2015 [cited 2015 November 1st]. Available from: http://m.bdbiosciences.com/us/reagents/research/antibodies-buffers/second-step-reagents/avidinstreptavidin/pe-streptavidin/p/554061

49. Longo MC, Berninger MS, Hartley JL. Use of Uracil DNA Glycosylase to Control Carry-over Contamination in Polymerase Chain-Reactions. Gene. 1990;93(1):125-8.

Appendix

Appendix A: Genomic Sequences for non-O157 Genes Selected

The reference genome for each gene selected for the non-O157 MOL-PCR assay is listed below (Appendix A.1-A.7). The genes selected for each serogroup were used as input sequences in the MOLigoDesigner software. As a result, MOLigo pairs were designed based on sequences both upstream and downstream from the SNP location. Referenced genomes listed in the Norman et al. study were obtained from GenBank. In parenthesis, the GenBank reference number is indicated and the location of the gene within the reference genome is also listed.

Appendix A.1: O26:H11 *fnl1* gene (GenBank AY763106:6715-7749)

ATGTTTAAGAATAAAACACTCGTTATCACTGGTGGCACTGGTTCTTTTGGTAATGCCGTACTTAAGCGTT
TTCTAGAT**A**CAGATATTACTGAAATACGAATATTTAGCAGGGATGAAAAAAAACAAGATGATATGCGGAA
AAAATATAATAACTCAAAATTAAAATTTTATATAGGTGATGTGCGAGACTATAATTCCGTTCTAAATGCA
ACGCGTGGTGCCGATTTTCTGTATCATGCAGCAGCCCTTAAACAAGTTCCTTCATGTGAATTTCACCCTA
TGGAGGCGGTTAAGACAAATGTTCTGGGTACGGAAAATGTTCTGGAGGCTGCTATTGCGAATGGGATTAA
ACGCGTGGTGTGCTTGAGTACCGATAAAGCCGTTTATCCTATCAATGCAATGGGCATATCTAAGGCAATG
ATGGAAAAAGTTATTGTTGCAAAATCACGTAATCTTGACAGTTCAAAAACAGTTATCTGTGGAACTCGTT
ATGGAAATGTAATGGCTTCACGTGGATCGGTCATCCCATTATTTGTTGATCTAATCAAAGCTGGTAAACC
ATTGACCATAACCGATCCCAATATGACTCGTTTCATGATGACGCTTGAGGATGCTGTCGATCTGGTCCTT
TATGCTTTCGAACATGGAAATAATGGTGACATTTTCGTTCAGAAAGCACCTGCGGCAACAATTCAAACAT
TAGCCATTGCACTTAAGGAATTGCTAAATGCCCATGAGCATCCAATCAATATTATTGGAACTCGACACGG
GGAAAAACTTTACGAAGCGTTATTGAGCCGAGAGGAAATGATTGCAGCGGAAGATATGGGTGATTATTAT
CGTGTTCCACCAGATCTCCGCGATTTGAACTATGGAAAATATGTGGAACATGGTGACCGTCGTATCTCGG
AAGTGGAAGATTATAATTCTCATAATACTGAGAGATTAGATGTTGAGGGAATGAAAGAATTACTGCTAAA
ACTTCCTTTTATCCGGGCACTTCGTTCTGGTGAAGATTATGAGTTGGATTCATAA

Appendix A.2: O26:H11 *wzx* gene (GenBank AY763106: 3543-4805)

ATGTTGAAAAAAAAACTTCAAAAAATAAAGGAATATCATTCAGTATTGGAGTTGGCAATAATTCAGGGTG
CGAATGCCATATTTCCTGTGTTGGTATTCCCATTTTTTCTTATTACCTTAGGGGAAAACATCTTTTCAAG
TATTGCTGTTGGTGAAGTACTAGCACTATATGTGCTTATATTTTCGCTATACAGTTTTGATATTATAAGT
GTGCAGAAGGTAATTTCAAGTGTGACAAAAGATGAAATATTTAAAGTTTACATTCTGACACTAATCTGTA
GGTTGTGTTTATTTGTTATTTCAGGAATATGTCTTTTATTTATAACGTATTTAATTAATAAAACATTAAG
TGTATACTTGGGATTGTTTTTATTGTACCCAGTAGGGATGATATTGCAATCTAATTATTTTTTTCAGGCT
ACGAATAACAATAGGCCATTGGCTGTTTTTGTACTAATTGCTCGTGGTATGTCATTATGTCTTATTTATT
TTTATAATGGACCAGCAGGCTATTTAACAAGTTATTATTATGTCATTTGTGTGTCTGGTTCGTATTTTTT
ATCTGGCGTGCTATCGCTTATATATATATATTATCAAAATAAGACTAATAAAGCTAAAATTCAATGGGCG
GAAATTTTAGAATATATATGCACAGGTTATCATCTGTTTATTGCTAATATATTTGTTATTCTATACAGAA
ATAGTAATATTATTATTCTTGGCACTCTTGCTTCGCCTGTTGCAACGTCTCTGTACGCGACGGCAGAGAA
AATTATTAAATGTATTCAGTCTATAGCAACCCCGTTAAATCAATACTATTTCACGAGGTTGATAAAGCAA
CATGAATTGAAATTAGAACCATACAAAGTTGGAGAATATAAAAGCCTGCTATATGCAAGCACAAATATTC
AGCTAAAGTTCATGGTTTTCATTGTCCTGAGTTTAGGGGGGGTGGGTACTATATTGGGATATAAGGTTCA

AAGTATCGCTGAAATTAGAAGCGCGTTCATCCCTTTATCAATAATGTCTTTTGCAATATTTATGGGGATA
TACAATTTTATGTTTGGTTCGGTTGGATTGTCCATAAGAGGGTATAAAAAAGAATTTTCTTATATAGTGG
CCATTACGGGTGTTTCAACTATTATTTTATCATTATGCCTGAGTTATTTCTTTGCTGAAATAGGCGCTGC
AATTGCTTATGTATTTGCTGAGTTTATCTTACTTATTCTCATACTTAGAATTTATAAAGTGAAACGATTA
TAA

Appendix A.3 O45 *wbhQ* gene (GenBank AY771223: 2196-3230)

ATGTTATATCTGGGCATTGTCGTTTTTGCATTCTTTATATCATGCGCGTTGACTTGGGGTTTAAGGTTAT
ATGCCATCAAGAATAATGTAATTGATCAACCAAACCAAAGAAGTTCTCACAGTGTTCCTACGCCTCGTGG
GGGAGGGGTGGCTATTGTATTAACATTGTTGGCTTCATTAATATGGCTGGTCCTTACAAATCATATAACT
CAGGCAACATTTTTAGGTTTCTTCGTGACTGGCTTATTGATTGCGGTTATTGGTTTCCTCGACGATCATG
GCCATATCGCTGCACGCTGGCGCTTGTTAATGCATTTCATCGCTGCAGCTATAGGGCTTTTTTATCTAGG
ATCTTTCCCAAGCATTAATATGTTTGGCTACGATGTTTCTTTGTCATGGTTTGGAATGATCCTCGGTAGT
ATCTATCTTGTATGGATGTTGAATTTATATAACTTTATGGATGGTATTAATGGTCTTGCAAGTGCGCAAG
CTATTACCTTTTCGCTTTGTAGTATCTTGATAATAACTATCAATAACTACTCTGATTCATCTGATGCAAT
GACTATGCTTGCATTAGCACTCGCTGGATCGGTCGCGGGATTCATCGTGTGGAACTTCCCTGTGGCAAGG
ATCTTCATGGGCGATGCTGGAAGCGGTTTTCTGGGTATCACCATTGGGCTGATGATCCTTTATTTTGCAA
AGCTAGACTCACGTGTTCTGATCGCAGAACTCTGTCTGCTGGGTGTTTTCATTGTTGATGCAACTACTAC
CTTACTGAGAAGATTACTGGCGGGTAAGAAAGTCTATGAAGCACATGCAAGTCATTGTTACCAAATTCTT
GCCCGTAAATATAAGAGCCATGTTCCTGTTACTATGGCGGCTATAGCTATTAACTTCATGTGGCTGCTCC
CTATTGCTTACTTAATCATCTCTGCAAAAATTGATGGAATTGTAGGCATTACCATCGCTTGGTTACCATT
GATGATTCTTGCATTCAGATGTGGTGCTGGAGTTAAAGATAAAGAGAGGGCATAA

Appendix A.4 O103 *wbtD* gene (GenBank AY532664: 5353-6339)

ATGATAATATCTTCAGATAACCTAAGTGTTATTATCCCAGTTTATAATGAAGAGAAGAATGTTATTAATA
TTCTTCAATCTCTTGAATGCCAGTCTTTAACGGGATTTAATGTTATTGTCATTGATGATGGTTCAAAAGA
TATGACTGCTTCTCTTGTAAAAGAATATAAACCATCATCATATAGTTTAAGCTTAATACAACAAAGTAAT
ATGGGGGCTGCTCGGGCAAGGGAGAATGCCATCAATTTTACTCATAGTGAGTATATTGCATTTATAGATA
GTGATGACTCTTTAAGTTCAGATGCTCTTGAAAAAGCATTAGCTCCGATGCTGGATAATAAAGATATCGA
CATCTCACTATTTGAACTGGTACATATTAAAAACTTAAATCATGATACTAATAACATATTTGTACCTTAT
TCTACAACCAAATTAATATGTGGAGAAGAGGCTTTTGCTAATTGTATTTCATATTGGGGATTGCATGGGT
TTGGTATCTACAAAAGAAAATTGATTCAAAAATCCTATGGTATTTATTATAAATATAATAAAAACAAAGA
AAATTATATAAATAATGATGAGGTTATTTCAAGAATCAGTTTTGGATTATCTAAAAATATCTATTTATCA
TCTGGAAGATATTTCTTTGTACAGAACATGGATTCAACAACAAGAAGAATAAATGAAAGTTATTATAAAG
TTATAAATAATGCTGTTTATTTAAGAGAGTATATAGATGAAGAAATAAAATATAATGATTTTGATTGTCT
AGGTGAAGCTAATAAACTCCTTGTAAGTACTATATGGGGAGTGTTTGTTCGTTATCAAAAATGGAAGAAA
AAATTCAGTGATGAAACAAACGGTAAATGGCGCGAAGCAATAAAAAAAGGAATGCAATATATAAAAAAAA
TCAATAACAATAAGAATATTAACCTGCATATTAAATCTAAAATGCAATTATATATAATATCTAAACTTGT
TAGCTAA

Appendix A.5 O111 *wbdK* gene (GenBank AF078736: 7441-8607)

ATGATTACATACCCACTTGCTAGTAATACTTGGGATGAATATGAGTATGCAGCAATACAGTCAGTAATTG
ACTCAAAAATGTTTACCATGGGTAAAAAGGTTGAGTTATATGAGAAAAATTTTGCTGATTTGTTTGGTAG
CAAATATGCCGTAATGGTTAGCTCTGGTTCTACAGCTAATCTGTTAATGATTGCTGCCCTTTTCTTCACT
AATAAACCAAAACTTAAAAGAGGTGATGAAATAATAGTACCTGCAGTGTCATGGTCTACGACATATTACC
CTCTGCAACAGTATGGCTTAAAGGTGAAGTTTGTCGATATCAATAAAGAAACTTTAAATATTGATATCGA
TAGTTTGAAAAATGCTATTTCAGATAAAACAAAAGCAATATTGACAGTAAATTTATTAGGTAATCCTAAT
GATTTTGCAAAAATAAATGAGATAATAAATAATAGGGATATTATCTTACTAGAAGATAACTGTGAGTCGA
TGGGCGCGGTCTTTCAAAATAAGCAGGCAGGCACATTCGGAGTTATGGGTACCTTTAGTTCTTTTTACTC
TCATCATATAGCTACAATGGAAGGGGGCTGCGTAGTTACTGATGATGAAGAGCTGTATCATGTATTGTTG
TGCCTTCGAGCTCATGGTTGGACAAGAAATTTACCAAAAGAGAATATGGTTACAGGCACTAAGAGTGATG
ATATTTTCGAAGAGTCGTTTAAGTTTGTTTTACCAGGATACAATGTTCGCCCACTTGAAATGAGTGGTGC
TATTGGGATAGAGCAACTTAAAAAGTTACCAGGTTTTATATCCACCAGACGTTCCAATGCACAATATTTT
GTAGATAAATTTAAAGATCATCCATTCCTTGATATACAAAAAGAAGTTGGTGAAAGTAGCTGGTTTGGTT

TTTCCTTCGTTATAAAGGAGGGAGCTGCTATTGAGAGGAAGAGTTTAGTAAATAATCTGATCTCAGCAGG
CATTGAATGCCGACCAATTGTTACTGGGAATTTTCTCAAAAATGAACGTGTTTTGAGTTATTTTGATTAC
TCTGTACATGATACGGTAGCAAATGCCGAATATATAGATAAGAATGGTTTTTTTGTCGGAAACCACCAGA
TACCTTTGTTTAATGAAATAGATTATCTACGAAAAGTATTAAAATAA

Appendix A.6 O121 *vioA* gene (GenBank AY208937: 4455-5558)

ATGGAAAAGCCAATCTTTGTAACGCAACCTAATTTACCACCGCTAGAGGAGTTTATACCATATCTGGAAA
TCATTTGGCAGAATAAGCAATTTACAAATAATGGTCCAATGCATCAAAAATTAGAAAAAAAATTATGCGA
GTTTCTTGGTGTTGAATACATTAGTCTATTTAATAATGGGACTATTGCGCTTATAACCGCAGTACAGGCT
TTGGGTGTTAAAGGCGAAGTAATTACCACACCATATTCCTTTGTAGCAACTGCACACTCCTTGGTCTTAA
ATGGGCTTAAACCTGTTTTTGTCGATATTGATTCCAAAACCTTAAATATCGATCCGAGAAGAATTGAAGA
GGCGATTACCCCTGAAACGCAGGCAATAATGCCGGTGCACTGCTATGGGAATCCTTGCGATACACAAGCT
ATTGCTGATATTGCGCAAAAATATAATTTAAAGGTCATTTATGATGCTGCGCATGCCTTTGGCGTTGAAG
ATGATGATGGAAGTGTTCTTCGCCATGGAGATCTAAGTGTCTTAAGTTTCCATGCAACTAAAGTGTTCAG
CACTTTTGAAGGCGGAGCTATTGTGTGTAATAGTAAAGAAATGAAAGAAAAAATTGATAGACTAAAAAAC
TTTGGTTATATCGATGAAACTAACATCAATATCATTGGCTCTAATGGAAAAATGAGCGAAGTTAATGCTG
CTTTCGGCTTGCTACAATTAGAACATATGGATACTTTTCTACGTGGTCGAATGAATGCTGACATGTTTTA
TCGGCAGAAACTTAAAGATATCACTGGTATAAGCATAGTAATTCCCAGCGGCCAGAAAATATCGAATTTT
TCATATTTCCCTATATTGGTTGAATCTGATTTTCCGTTATCTCGTGATGAACTATTTAATTATCTGAAGA
ACCAGAATATTTTTGCAAGACGTTATTTTTATCCTGTTATACCAGATTTTCAAGCGTATTTGAATGTAGG
TGAAGTCTGTGATGTCAAGAATGCCCGTGAAATAGCCTCGAAAGTGCTTTGCCTACCGATGCATGCAGAA
TTGAGCTCTGATATCTTAGAATATATTGTAAGTACGATTAGGGAGATTAAATGA

Appendix A.7 O145 *wzy* gene (GenBank AY647260: 5615-6802)

GTGAATATAAAGAAAGATAAGTTTATAAATGGAGTGATTTTTTTTTGGTTAATTATTTCTTCGTTATATT
ACTTAAATGCTATTTTTTCTGGTGTTGACACATTAAAATATAATGAAGATTTAACGCAAAAAATTATAAA
ATATATAGTTTGCTTAGTTATAAGTCTAAGTATCTTATTTATTTACAAGAAATTTAATTATTTTTTTGTA
TTGTTTTTTTTCTTGTTCCTGTCTGTTGCTTCAGCCCTTTTCAGTGGTGCGGTAACAATTTACGCAACAA
CAATGTTGATTATTGCAACTATGATCAGCTTTTGCCTGATTATTCCTCTATTTTCTTATAATATGGTGAA
AGTTAATAGAGTTCTTTTATGGACAGGAGTTATTGTAGGCACGATTTCTGTATTAGAATTAACGGTATTT
TATAATTATATGGTTTCATATTGGGCTGCCACTGATGGGATTAGGTCAATATCTTCTCTTCTGAATCCTA
CGAATAGTGGTGCTTATTCAGCGATTATTATTTTAATCGCCTTGGTGACAAATATAAAAAGTCTTTTTAA
AAGAGCTTTATTTCTTATAATGCCGATGATAACGTTAATTAGCAGTGGTTCGCGCACAGCATGGTTATCA
CTTGGTATGACACTTTTATTAACAGTAGTATTGAGAGACAGTGCCAGCATTCGCTTGCGAAAAAAAATAT
TTACTCTTGCAAGCATTGGCACTGTTTGCGGTGCATTGTACGCCATATTTTATATGGGCAGTATCTCTGG
TATTGAATCACAATATCGAGGTCTTAATACGTATACTGCATCAATTCGAGTTGAAAACTTTCTGACATAT
TTAAATTTAGTTGATCTGAATATGTTGCTACCTGATTTTTTAGATAAAAATATAAATCTCATTTCAGATA
ACTTTTATCTCGTAATGTTTAATTATGCCGGTCTAATCGGCTTTTTTATTGTTTTATTAATTTTATTGCT
GCTTATCTTCTGGAACATACAATTTAAAATATTTAATGAGTTAATGGCTGAAGATATAGCCATTTGGAGA
GTTGTTTTTATTTATTTCCTAATATCCGGGCTTTCAAATTCATTTATAAATTCTTTTCCTGTAAATCAAT
TGTTCTTTATCTCATGCGGATATTATATATATAAATATAAATTAGTTAAAAGCTCTATAGGAAGATAA

Appendix B: Critical Values for Wilcoxon Rank Sum

The Wilcoxon Rank Sum test was used as a statistical analysis tool for the MOL-PCR assay. The identification of each serogroup was determined by using either the approximate or the exact Wilcoxon Rank Sum Test. Listed below are the critical values for the exact Wilcoxon Rank Sum test based on a 0.05 one-tailed Z-test. The table lists the critical values for each n value. The critical value most often used in this assay was 27. The

N1 value corresponded to the number of sample replicates (3) and the N2 value corresponded to the number of no template control samples (8). Therefore, the chart was used to predict the critical value based on the number of sample replicates and the number of blanks. The critical value was compared against the sum of the ranks for the sample group to determine if the sample was positive or negative for a particular group (Chapter 3, Section 6).

	alpha	0.05 one-tailed														
N1	**3**		**4**		**5**		**6**		**7**		**8**		**9**		**10**	
N2																
3	6	15	7	17	7	20	8	22	9	24	9	27	10	**29**	11	31
4	7	17	12	24	13	27	14	20	15	33	16	36	17	**39**	18	42
5	7	20	13	27	19	36	24	40	22	43	24	46	25	**50**	26	54
6	8	22	14	30	20	40	28	50	30	54	32	58	33	**63**	35	67
7	9	24	15	33	22	43	30	54	39	66	41	71	43	**76**	46	80
8	9	27	16	36	24	46	32	58	41	71	52	84	54	**90**	57	95
9	10	29	17	39	25	50	33	63	43	76	54	90	66	**105**	69	11
10	11	31	18	42	26	54	35	67	46	80	57	95	69	**111**	83	127

Appendix C: Complementary Synthetic DNA Sequences for Assay Development

Appendix C contains the sequences for the synthetic complementary target sequences developed for the training panel. The training panel consists of synthetic oligonucleotides for both the STEC and non-STEC O26 genes (*fnl1* and *wzx).* The training panel aided in the optimization of both the ligation and the PCR conditions. The synthetic target sequences were developed from the reference genomic sequences listed in Appendix A.1 and A.2.

Serogroup	SNP	Sequence (5'-3')
O26 S-*wzx*	T	TCA TTG TCC TGA GTT TAG GGG GGG **G**GG GTA CTA TAT TGG GAT ATA AGG TTC AAA GT
O26 NS-*wzx*	G	TCA TTG TCC TGA GTT TAG GGG GGG **T**GG GTA CTA TAT TGG GAT ATA AGG TTC AAA GT
O26 S-*fnl1*	G	ATG CCG TAC TTA AGC GTT TTC TAG ATA CAG ATA TT**A** CTG AAA TAC GAA TAT TTA GCA GGG ATG AAA A
O26 NS-*fnl1*	A	ATG CCG TAC TTA AGC GTT TTC TAG ATA CAG ATA TT**G** CTG AAA TAC GAA TAT TTA GCA GGG ATG AAA A

Appendix D: Strain identification and virulence profiles for DNA panel

In Appendix D.1 and D.2 are the strains that were included in the STEC DNA panel. Each panel was tested against the computationally designed MOLigo pairs. Table 4.2-4.6 list the results of the MOL-PCR assays evaluated against each DNA panel. Described below are the strain identification numbers, serotypes, and virulence genes associated with each strain.

Appendix D.1: non-O157 Strains tested in Singleplex Assay

The strains listed in this table were obtained from ATCC. The strain is listed by ATCC identification number and serotype. The virulence profile for each strain is listed below. The ATCC samples were used for optimization of the STEC non-O157 assay. ATCC samples consist of purified genomic DNA and were instrumental to determine MOLigo pair performance.

Strain ID	**Serotype**	***stx1***	***stx2***	***eae***
BAA-2196	O26:H11	Positive	Positive	Positive
BAA-2193	O45:H2	Positive	Negative	Positive
BAA-2215	O103:H11	Positive	Negative	Positive
BAA-2440	O111	Positive	Positive	Positive
BAA-2219	O121:H19	Negative	Positive	Positive
BAA-2192	O145:NM	Positive	Positive	Positive

Appendix D.2: non-O157 Strains tested in 8-plex STEC and non-STEC Assays

The strains listed in this table were obtained from Kansas State University (KSU). Each strain is identified by serotype and strain ID number. The strains are identified based on their virulence profiles (*stx1*, *stx2*, and *eae*). The strains were obtained from clinical isolates associated with samples obtained from human stool. The strains were used to

optimize the STEC non-O157 assay and have also been used to evaluate the SNP specificity for the non-STEC assay.

Strain ID	**Serotype**	***stx1***	***stx2***	***eae***
H30	O26:H11	Positive	Negative	Positive
CDC 96-3285	O45:H2	Positive	Negative	Positive
CDC 90-3128	O103:H2	Positive	Negative	Positive
JB1-95	O111	Positive	Positive	Positive
CDC 97-3068	O121:H19	Negative	Positive	Positive
83-75	O145:NM	Negative	Positive	Positive

Appendix E: Original Design for MOL-PCR Assay

The table listed in this Appendix includes the original design of each MOLigo pair for both the STEC and non-STEC assays. The singleplex assay and the initial evaluation of the 8-plex assay were performed with the MOLigo pairs listed in this table. The MOLigo pairs had to be redesigned when the assay was evaluated on the Luminex MagPix®. The beads originally selected for O26 non-STEC *fnl1,* O26 STEC and non-STEC *wzx,* O103 STEC *wbtD,* O111 STEC *wbdK*, and O145 *wzy*, were not available on the Luminex MagPix®. Each MOLigo pair listed above was redesigned with beads that were available on the MagPix®.

Serogroup	xTAG	SNP	MOLigo	Sequence (5'-3')
O26 S-*fnl1*	A042	A	M1	/5Phos/AAT ATC TGT ATC TAG AAA ACG CTT AAG TAC TCT CAC TTC TTA CTA CCG CG
			M2	ACT CGT AGG GAA TAA ACC GTA TTT GTT ATG ATA AAT GTG TAG TGA TCC CTG CTA AAT ATT CGT ATT TCA GT
O26 NS-*fnl1*	B098	G	M1	/5Phos/AAT ATC TGT ATC TAG AAA ACG CTT AAG TAC TCT CAC TTC TTA CTA CCG CG
			M2	ACT CGT AGG GAA TAA ACC GTA AAT GTG TGT TTA GTA GTT GTA AAC CCT GCT AAA TAT TCG TAT TTC AGC
O26 S-*wzx*	B082	G	M1	/5Phos/CCC CCC CTA AAC TCA GGA CTC TCA CTT CTT ACT ACC GCG
			M2	ACT CGT AGG GAA TAA ACC GTA TTT GTA TAA AGT GAA GAA GAG AAG AAC CTT ATA TCC CAA TAT AGT ACC CC
O26 NS-*wzx*	B009	T	M1	/5Phos/CCC CCC CTA AAC TCA GGA CTC TCA CTT CTT ACT ACC GCG
			M2	ACT CGT AGG GAA TAA ACC GTG AAT TGT ATA AAG TAT TAG ATG TGG AAC CTT ATA TCC CAA TAT AGT ACC CA
O45 S-*wbhW*	A057	A	M1	/5Phos/TCT GAT ATG AGT CGA ACA TAA TGT TTT TAC TCT CAC TTC TTA CTA CCG CG
			M2	ACT CGT AGG GAA TAA ACC GTA GAG TAT TAG TAG TTA TTG TAA GTA GTT AAA TAG ATT AGA TAT CTT TAT TTA AAC AAC ATC TTA
O45 NS-*wbhW*	A051	C	M1	/5Phos/TCT GAT ATG AGT CGA ACA TAA TGT TTT TAC TCT CAC TTC TTA CTA CCG CG
			M2	ACT CGT AGG GAA TAA ACC GTG ATA AGA AAG TGA AAT GTA AAT TGG TTA AAT AGA TTA GAT ATC TTT ATT TAA ACA ACA TCT TC
O103 S-*wbtD*	B096	T	M1	/5Phos/CAG GTT AAT ATT CTT ATT GTT ATT GAT TTT TTT TAT ATA TTG TCT CAC TTC TTA CTA CCG CG
			M2	ACT CGT AGG GAA TAA ACC GTT ATG TGT ATG AAG ATT ATA GTT AGA CAA GTT TAG ATA TTA TAT ATA ATT GCA TTT TAG ATT TAA TAT A
O103 NS-*wbtD*	A073	C	M1	/5Phos/CAG GTT AAT ATT CTT ATT GTT ATT GAT TTT TTT TAT ATA TTG TCT CAC TTC TTA CTA CCG CG
			M2	ACT CGT AGG GAA TAA ACC GTG TTG AGA ATT AGA ATT TGA TAA AGC AAG TTT AGA TAT TAT ATA TAA TTG CAT TTT AGA TTT AAT ATG
	B086	T	M1	/5Phos/CCT GTA ACC ATA TTC TCT TTT GGT AAA TTT CTC ACT TCT TAC TAC CGC G

O111 S-*wbdK*			M2	ACT CGT AGG GAA TAA ACC GTG TGT GAT TAG TAA TGA AGT ATT TAA CTC TTC GAA AAT ATC ATC ACT CTT AGT A
O111 NS-*wbdK*	A026	C	M1	/5Phos/CCT GTA ACC ATA TTC TCT TTT GGT AAA TTT CTC ACT TCT TAC TAC CGC G
			M2	ACT CGT AGG GAA TAA ACC GTT TTG ATT TAA GAG TGT TGA ATG TAC TCT TCG AAA ATA TCA TCA CTC TTA GTG
O121 S-*vioA*	A063	T	M1	/5Phos/ATC AAT ATC GAC AAA AAC AGG TTT AAG CTC TCA CTT CTT ACT ACC GCG
			M2	ACT CGT AGG GAA TAA ACC GTT TTG TTG TTA AGT ATG TGA TTT AGT CGG ATC GAT ATT TAA GGT TTT GGA
O121 NS-*vioA*	A053	C	M1	/5Phos/ATC AAT ATC GAC AAA AAC AGG TTT AAG CTC TCA CTT CTT ACT ACC GCG
			M2	ACT CGT AGG GAA TAA ACC GTG TTT GTG TTT GTA TAA GTT GTT AAT CGG ATC GAT ATT TAA GGT TTT GGG
O145 S-*wzy*	B089	C	M1	/5Phos/CAC TCC ATT TAT AAA CTT ATC TTT CTT TAT ATT CAT CTC ACT TCT TAC TAC CGC G
			M2	ACT CGT AGG GAA TAA ACC GTT AGA TAA TGT GAA GTA ATA AGT GAA GTA ATA TAA CGA AGA AAT AAT TAA CCA AAA AAA AAG
O145 NS-*wzy*	A055	A	M1	/5Phos/CAC TCC ATT TAT AAA CTT ATC TTT CTT TAT ATT CAT CTC ACT TCT TAC TAC CGC G
			M2	ACT CGT AGG GAA TAA ACC GTG AAG ATA TTG AAA GAA TTT GAT GTA GTA ATA TAA CGA AGA AAT AAT TAA CCA AAA AAA AAT
Ligation Control	A056		M1	CTT TGG GGA TTA GCT CCG CTC TCA CTT CTT ACT ACC GCG
			M2	ACT CGT AGG GAA TAA ACC GTA ATT AGA AGT AAG TAG AGT TTA AGT GCA ATC CGA ACT GAG ATG GA
Forward Primer				/5Alex488N/CGC GGT AGT AAG AAG TGA GA
Reverse Primer				ACT CGT AGG GAA TAA ACC GT

Appendix E: Evaluation of Samples on Luminex MagPix®

The MOL-PCR assay was evaluated on the Luminex MagPix® instrument. The MagPix instrument is identified as a compact fluorescence-based detection system that uses light-emitting diodes (LEDs) and a CCD camera to detect 50 different analytes.[46] The MagPix® is not considered a flow cytometer, but is identified as a fluorescent imager.[46] The MagPix system® is used with xMAP magnetic microspheres that are color-coded and

allow for multiple analytes to be evaluated in a single sample.[46] The magnetic microspheres used in this assay were xTAG® microspheres that were coupled to a 24-base oligonucleotide sequence.[46]

The components of the MagPix® consists of a fluidics system and an optics module. The fluidics system has the following components; syringe pump, sample probe and a sample valve.[46] The optics module for this instrument consist of four main components: 1) sample chamber, 2) magnet, 3) Light emitting diodes (LEDs), and 4) CCD camera.[46] The samples are analyzed by a sample probe that draws the sample from the 96-well plate into the instrument.[47] The beads are transferred to the sample chamber. The magnetic plate is position at the back of the sample chamber where the microsphere are arranged in a monolayer along the magnetic plate.[47] Red LEDs illuminate the sample chamber and the CCD camera images the fluorescent microspheres based on the color-coded beads.[47] Green LEDs illuminate the sample chamber to capture whether the sample analyte is present.[47] The software then processes the images and determines concentrations of the analytes.[47] The magnet is moved away from the sample chamber allowing for the sample to be removed from the sample chamber.[47] The reporter channel excitation (LED excitation) is 511 nm. SAPE was used in this research and has an excitation of 496 nm and an emission maximum at 578 nm.[48]

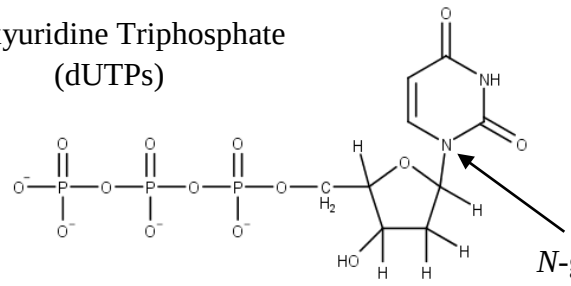
Appendix F: dUTPS and Uracil DNA- glycosylase Treatment

The polymerase chain reaction (PCR) is an essential component for molecular diagnostics. Amplification via target amplification or signal amplification generates a multitude of PCR products that allows for detection by end point detection, RT-PCR, or

by flow cytometry.[29, 49] Generating PCR products increases the susceptibility for potential problems associated with PCR contamination.[49] Contamination can occur as a result of cross-contamination between samples, contamination for laboratory environment, and carry-over contamination from previous experiments.[49] As a result, the three sources of contamination could lead to false positive results.[49] Often determining the source of contamination is time-consuming and can get very expensive.[49]

A method to control carry-over contamination involves the introduction of deoxyuridine triphosphate (dUTPs) and uracil DNA glycosylase (UDG) into the ligation and PCR reactions.[49] UDG is an enzyme that is capable of removing uracil bases from DNA by cleaving the *N*-glycosidic bond located between the base and the sugar-phosphate group.[49] This results in the production of an apyrimidinic site that prevents DNA polymerase from generating a new product.[49] The incorporation of dUTPs in the PCR reaction replace dTTPs, generating PCR products that contain uracil bases rather than thymine.[49] UDG treatment is employed prior to beginning the next reaction (during ligation), preventing amplification of previous PCR products. UDG is inactivated during the DNA denaturation step, preventing the enzyme from degrading the product that is being generated in the current experiment.[49] Identified below is the location of the bond cleaved by UDG.

Deoxyuridine Triphosphate
(dUTPs)



N-glycosidic bond cleaved
by UDG enzyme

Appendix G: Standard Operating Procedures for MOL-PCR Assay

The MOL-PCR assay chemistry is considered an open system. With the separation of the ligation and amplification reactions, the risk of contamination for this assay increases substantially. To limit the chances for contamination, a standard operating procedure is strictly followed. The laboratory contains three work stations where a unilateral workflow is employed. The first workstation (Pre-PCR Workstation) is used for mixing ligation and PCR reagents. In this workstation, reagents for ligation and PCR are mixed before template DNA is added to the reaction. The second pre-PCR workstation is where template DNA and ligation product is added to the reaction. The third PCR station is the post-PCR workstation, which is in a separate room on the opposite end of the lab. The addition of microspheres are added to the post-PCR product and all wash steps are completed in this workstation. The workstations are cleaned with 20% bleach after each use and all pipettes, tubes, tube racks, non-DNA solutions undergo UV irradiation for 30 minutes. The incorporation of the SOP has greatly reduced contamination issues that were previously experienced. Below is the standard operating procedure that is used for the MOL-PCR assays.

MOL-PCR Standard Operating Procedures

The entire STEC MOL-PCR assay will be done in three separate areas to prevent any contamination from one process into processes further down the line. There will be a Master Mix preparation area where all reagents that do not contain DNA other than the MOLigos will be made. There will be a separate area where the Template DNA and any positive control DNA will be added before ligation procedure. This area will also be where the transfer from ligation product to PCR reaction will take place. There will then be a third area where the post-PCR reactions and setup will take place that is in a completely separate laboratory.

Ligation Master Mix Preparation

1. Enter the pre-PCR lab space and immediately put on lab coat followed by nitrile gloves.
2. Open the MM PCR hood and place any tips, racks, Eppendorf tubes, water aliquots needed for the day inside and turn on the UV light inside for ~30 minutes, if not already done.
3. After the consumables have properly been irradiated by UV light then proceed on to the next steps.
4. Get the ligation freezer box from the freezer in the other room and bring into pre-PCR lab. Wipe the outside down with 10% bleach and place it inside of the hood.
5. Remove gloves and replace with new, clean nitrile gloves.
6. Remove the SSDNA, ligation buffer, and MOLigo mix and centrifuge them to help them thaw quicker. This should be done inside of the hood using only items that have been UV treated.
7. Once these items have thawed make sure to mix them well before dispensing volumes.
8. Dispense water into 3 tubes: 1 for MOLigo dilution, 1 for SSDNA dilution, and 1 for Master Mix.
9. Dilute the MOLigo Detection Mix for the STEC MOLigos (M1 and M2 are each at 25uM) to 100nM by adding 496uL of water and 2uL of each m1 and m2 STEC MOLigos. Make working concentration by diluting 100nM detection mix to 20nM (40uL MD+160uL H2O).
10. Dilute Salmon Sperm DNA in water to 12uL DNA to 68uL water. 1.5mg/ml SSDNA. Remove gloves and replace with new nitrile gloves
11. Use the chart below to add the correct volumes of components to the Master Mix tube.

Component	Volume per reaction (uL)	Total # reactions	Master Mix Volume (uL)
HPLC H2O	11.2	0	0
1.5mg/mL Salmons Sperm DNA	2	0	0
10X Pfu Buffer (1X final)	2	0	0
Non-STEC MOLigo(s) (20nM in.) (2nM fin.)(each)	2	0	0
Uracil-DNA Glycosylase 5u/uL (1U/20uL)	0.5	0	0
Ampligase (5U/uL) (1U/20uL reaction)	0.5	0	0
Master Mix Volume	18		0
Template DNA	2		0
Total Volume	20		0

12. Re-cap PCR Mix tube, briefly vortex, spin down and place back in workstation.
13. Take out a new 96 well plate from sealed bag.
14. Dispense equal volume of MM solution into 1 strip of 8-tube strip, for easy dispensing into PCR plate.
15. Use the multichannel pipette to dispense 18uL of MM from strip tubes into 96 well plate.
16. Inspect PCR plate to check that all appropriate wells have been filled and volumes appear consistent. Seal PCR plate.
17. Remove sealed plate from PCR hood and set on bench next to hood.
18. Seal the waste container bag and remove from hood, dispose of bag in large container.
19. Spray inside area of PCR hood with 10-20% bleach solution and wipe down thoroughly.
20. Follow bleach with water and dry towel.
21. Refill the tips, racks and other used items for next assay and UV irradiate for 30 minutes.

22. Return the freezer box to the -20 C freezer.
23. Remove gloves and replace with new nitrile gloves.

Ligation DNA Addition Station

24. Inspect the DNA addition PCR hood to make sure all the consumables needed are present.
25. Any items that are not present and have not been UV cleaned will need to be 20% bleach cleaned before being put into the PCR hood.
26. Move the MM containing PCR plate into the hood.
27. Remove seal from PCR plate, dispose and change gloves.
28. Using tips from a new box begin adding test reagents to plate.
29. To the first 8 wells of the plate add the negative control, then remove and replace gloves.
30. Add the DNA to be tested to each of their corresponding wells in triplicate, then remove and replace gloves.
31. Add the positive controls to the last 8 wells of the plate, then remove and replace gloves.

Plate Layout

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B												
C												
D												
E												
F												
G												
H												

32. Remove plate from the PCR hood and place in thermo cycler and run the ligation protocol.

    Ligation Conditions:

    1. 37°C 30 minutes (UDG Treatment)
    2. 95°C 3 minutes
    3. 95°C 25 seconds
    4. 58°C 2 minutes
    5. Repeat steps 3-4 x30 times
    6. 95°C 30 minutes (UDG Inactivation)
    6. Hold 4°C

33. Return to your workstation, dispose of all waste, change gloves, and bleach entire workstation.
34. Follow bleach with water and dry towel, then UV irradiate workstation.
35. Return to entry and remove PPE (gloves, then glasses and coat).

PCR Master Mix Preparation

1. Enter the pre-PCR lab space and immediately put on lab coat followed by nitrile gloves.
2. Open the MM PCR hood and place any tips, racks, Eppendorf tubes, water aliquots needed for the day inside and turn on the UV light inside for ~30 minutes, if not already done.
3. After the consumables have properly been irradiated by UV light then proceed on to the next steps.
4. Get the benchtop freezer box from the freezer in the other room and bring into pre-PCR lab. Wipe the outside down with 10% bleach and place it inside of the hood.
5. Remove gloves and replace with new, clean nitrile gloves.
6. Remove the 10X PCR Buffer, 25mM MgCl2, dUTP mix, Forward and Reverse Primer. Place 10X buffer and MgCl2 reagents into the centrifuge to reduce defrosting time. Centrifugation of reagents will be done inside the hood using the UV treated centrifuge.
7. After the reagents have thawed, make sure to mix them well before dispensing volumes

8. Use the chart below to add the correct volumes of components to the Master Mix tube

Component	Volume per	Total # reactions	Master Mix Volume
HPLC Water	10.05	0	0
10X PCR Buffer (1X final)	2	0	0
25mM MgCl2 (4mM final)	3.2	0	0
dUTP Mix (200uM(A,C,G), 400uM(U) final)	2	0	0
Forward Primer (0.5uM final)	0.5	0	0
Reverse Primer (0.1uM final)	0.1	0	0
Amplitaq Gold (0.05U/uL final)	0.15	0	0
Master mix Volume	18		0
Template DNA	2		0
Total Volume	20		0

9. Re-cap PCR Master Mix tube, briefly vortex, spin down and place back in workstation.
10. Take out a new 96 well plate from sealed bag.
11. Use the multichannel pipette to dispense 18uL of MM from strip tubes into 96 well plate
12. Inspect PCR plate to check that all appropriate wells have been filled and volumes appear consistent. Seal PCR plate.
13. Remove sealed plate from PCR hood and set on bench next to hood.
14. Take benchtop cooler back to main lab and put in -20°C freezer.
15. Seal the waste container bag and remove from hood, dispose of bag into large container.
16. Spray inside area of PCR hood with 10-20% bleach solution and wipe down thoroughly.
17. Follow bleach with water and dry towel.
18. Refill the tips, racks and other used items for the next assay and UV irradiate for 30 minutes
19. Remove nitrile glove and replace with new gloves

Amplification DNA Addition Station

20. Inspect the DNA addition PCR hood to make sure all the consumables needed are present.
21. Any items that are not present and have not been UV cleaned will need to be 20% bleach cleaned before being put into the PCR hood.
22. Move the MM containing PCR plate into the hood.
23. Remove the ligation product plate from the thermocycler and place it inside PCR hood. Remove and replace nitrile gloves.
24. Remove seal from PCR plate, dispose and change gloves.
25. Using the multichannel pipettor to dispense the ligation product from PCR plate. Remove 2uL of product from ligation plate and transfer it to MM PCR plate in the same location on plate.
26. Add the ligated template DNA to their corresponding wells, then remove and replace gloves.
27. Seal the plate using the foil adhesive sheets, making sure that all wells are firmly sealed.
28. Remove plate from the PCR hood and set on the bench next to the hood.
29. Seal the waste container bag and remove from hood, dispose of bag into large container.
30. Spray inside area of PCR hood with 10-20% bleach solution and wipe down thoroughly.
31. Follow bleach with water and dry towel
32. Refill the tips, racks and other used items for the next assay and UV irradiate for 30 minutes.
33. Remove gloves, followed by lab coat and leave in pre-PCR room.

34. Put on nitrile gloves before leaving pre-PCR lab.

Amplification Post-PCR Lab

1. Take sealed plate from pre-PCR lab into Post-PCR lab
2. Place PCR plate into thermo cycler and run Amplification

Amplification Conditions:

1.	95°C	10 minutes
2.	95°C	15 seconds
3.	58°C	20 seconds
4.	72°C	20 seconds
5.	Repeat Steps 2-4	45 cycles
6.	72°C	7 minutes
7.	Hold at 4°C	

Hybridization Post-PCR Lab

1. Enter the Post-PCR lab space and immediately put on lab coat, followed by nitrile gloves
2. Open the MM PCR hood and place any tips, racks, Eppendorf tubes, TE buffer, and MES buffer needed for the day inside and turn on the UV light inside for ~30 minutes, if not already done
3. After the consumables have properly been irradiated by UV light, then proceed on to the next steps
4. Retrieve Luminex beads from the refrigerator and place into MM PCR hood
5. Vortex Luminex beads for 30 seconds inside MM PCR hood using centrifuge that has been UV irradiated.
6. Aliquot 32 uL of each Luminex bead into one tube with 100uL MES buffer, centrifuge beads to a pellet, remove supernatant and suspend bead pellet in MES buffer.
7. Aliquot the Luminex bead multiplex into 8-strip tubes for dispensing into PCR product plate.
8. Add 10uL of 8-13plex Luminex beads to amplified PCR product
9. Seal PCR plate and transfer plate to thermo cycler for hybridization

Hybridization Conditions

1.	95°C	3 minutes
2.	85°C	1 minute (Decrease by 10°C until reaching 55°C)
3.	50°C	30 minutes
4.	40°C	1 minute (Decrease by 5°C until reaching 25°C)
6.	Hold at 25°C	1 hour

10. Return to Post-PCR MM PCR hood and proceed with clean up
11. Seal the waste container bag and remove from hood, dispose of bag in large container
12. Refill the tips, racks and other used items for the next assay and UV irradiate for 30 minutes
13. Remove gloves

Addition of SAPE Post-PCR Lab

1. Enter Post-PCR lab space and immediately put on lab coat, followed by nitrile gloves
2. Open the MM PCR hood and place any tips, racks, Eppendorf tubes, TE buffer needed for the day inside and turn on the UV light inside for ~30 minutes, if not already done.
3. Prepare SAPE by diluting 500ug/mL to 10ug/mL in TE buffer. (20uL SAPE in 980uL TE buffer)
4. Remove PCR plate from the thermal cycler and place into MM PCR workstation
5. Pour the diluted SAPE into origami reservoir.

6. Using multichannel to pull 50uL of SAPE from reservoir and use the tips to pierce the foil before dispensing into appropriate wells. Make sure the solution is at the bottom of the well.
7. Incubate at room temperature for 30 minutes, inside of PCR hood under a box to protect from light.

Wash Steps Post-PCR Lab

1. Place PCR plate with SAPE onto magnetic plate and leave there for 5 minutes inside of the PCR hood.
2. Get a flat bottom 96 well plate and place it inside of the PCR hood.
3. Inside of the PCR hood, dispense enough TE buffer into origami reservoir to add 100uL TE buffer into each well to be analyzed.
4. After 5 minutes on the magnetic plate, remove the SAPE supernatant by removing the whole assembly from the hood, walk over to the sink.
5. Place absorbent towels into the sink to catch the SAPE. Invert the entire assembly (plate and magnet) sharply over the towels once or twice to completely remove the SAPE. Try to keep all sample onto the towels. Dispose of towels into biological trash.
6. Once done put the plate back into the hood and put the magnetic plate away in drawer.
7. Use the multichannel pipettor to add 100uL of TE buffer to the first column on the plate to resuspend the pellet a few times. Use the same tips to remove the suspended pellets and place into the 96 well plate.
8. Once all samples have been transferred to the 96 well plate, set samples next to post-PCR hood. and proceed with clean up
9. Seal the waste container bag and remove from hood, dispose of bag in large container
10. Spray inside area of PCR hood with 10-20% bleach solution and wipe down thoroughly
11. Follow bleach with water and dry towel
12. Refill the tips, racks and other used items for the next assay and UV irradiate for 30 minutes
13. Remove gloves and replace with new nitrile gloves
14. Proceed with running samples on Magpix
15. After run, bleach (20%) soak all the counters inside and outside of this hood and wipe down the MagPix and counters near it with bleach solution.
16. After soaking for about 30 minutes, wipe down with water.