The Pennsylvania State University
The Graduate School
The Huck Institutes of the Life Sciences

GENOMIC STRUCTURE OF AN E. COLI O157:H7 SUPER-SHEDDER ISOLATE
WITH INCREASED ADHERENCE TO BOVINE RSE CELLS

A Thesis in
Immunology and Infectious Diseases

by
Rebecca Ann Cote

© 2013 Rebecca Ann Cote

Submitted in Partial Fulfillment
of the Requirements
for the Degree of
Master of Science

May 2013
The thesis of Rebecca Ann Cote was reviewed and approved* by the following:

Vivek Kapur  
Professor of Veterinary and Biomedical Sciences  
Thesis Adviser

Subhashinie Kariyawasam  
Assistant Professor; Veterinary Microbiologist

Edward G. Dudley  
Associate Professor of Food Science

Chitrata DebRoy  
Director of the _E. coli_ Reference Center

Margherita T. Cantorna  
Professor of Molecular Immunology  
Chair of Intercollege Graduate Degree Program in Immunology and Infectious Diseases

*Signatures are on file in the Graduate School
ABSTRACT

Shiga toxin-producing *Escherichia coli* (STEC), in particular *E. coli* O157:H7, is recognized as a significant foodborne pathogen and a serious threat to public health worldwide. The major reservoir of *E. coli* O157:H7 are asymptomatic cattle which harbor the organism in the terminal recto-anal junction (RAJ) of their intestinal tract. Recently, isolates of *E. coli* O157:H7 have been obtained from cattle that are classified as “super-shredder”; excreting bacteria levels that are greater than $10^4$ CFU/g of feces. In this study, we examined the genome of a super-shredder isolate (SS17) and conducted comparative analysis with reference strains of *E. coli* O157:H7. Analysis of the genome revealed a genome size of 5.5 Mb, 5442 coding DNA sequences (CDS), and a G+C content of 50.5%. In addition, SS17 has a Shiga toxin gene profile of *stx1*-*stx2*+*stx2c*+ and contains two plasmids, pO157 and pSS17. Whole genome comparative analysis revealed a clustering of SS17 with the lineage I/II *E. coli* O157:H7 spinach outbreak strains, TW14359 and EC4115. In addition, analysis of phage regions in SS17 shows similar location and composition to the phage regions in TW14359 and EC4115. Of particular importance are virulence and adherence genes that may contribute to the super-shredder phenotype. A number of non-synonymous single nucleotide polymorphisms (nsSNPs) were identified in virulence and adherence genes including the adhesins *wzzB*, *fimA*, and *csgG*, along with a truncation of *cah*. Using a unique RAJ model, SS17 demonstrated an enhanced adherence pattern to bovine RSE cells as compared to other well-characterized *E. coli* O157:H7 strains. In addition, adherence of SS17 and reference O157 strains to RSE cells was not blocked in the presence of antisera to intimin, Tir, EspA, and EspB, suggesting a mechanism of adherence independent of these LEE-encoded proteins.
# TABLE OF CONTENTS

List of Figures ............................................................................................................. v
List of Tables ............................................................................................................. vi
List of Abbreviations ............................................................................................... vii
Acknowledgements ................................................................................................ ix

Chapter 1 Introduction .......................................................................................... 1

Chapter 2 Background and Significance ............................................................... 3

2.1 The Pathogenesis of *E. coli* O157:H7 ................................................................. 3
2.2 Definition of Super-Shedding ............................................................................ 5
2.2 Prevalence of Super-Shedding .......................................................................... 7
2.3 Contributing Factors to Super-Shedding ......................................................... 8
2.4 The RAJ and a Realistic *in vitro* Adherence Assay ......................................... 9

Chapter 3 Materials and Methods ........................................................................ 13

3.1 Bacterial Strains ................................................................................................ 13
3.2 Genomic DNA Extraction .............................................................................. 14
3.3 Genome Sequencing and Assembly ................................................................. 15
3.4 Annotation and Comparative Genomics ......................................................... 15
3.5 RSE Cell Adherence Assay ............................................................................ 16

Chapter 4 Results .................................................................................................. 18

4.1 SS17 Genome .................................................................................................... 18
4.2 Phage Regions in SS17 .................................................................................... 20
4.3 Virulence Genes .............................................................................................. 22
   4.3.1 Toxin Genes .............................................................................................. 22
   4.3.2 Adherence Genes ..................................................................................... 24
   4.3.3 Other Virulence-Associated Genes .......................................................... 25
4.4 Phylogenetic Analysis ...................................................................................... 27
4.5 Comparative Genomics ................................................................................... 27
   4.4.1 Whole Genome Comparison .................................................................. 28
   4.4.2 Single Nucleotide Polymorphisms ......................................................... 34
   4.4.3 Polymorphisms in Promoter and Termination Regions ....................... 38
   4.4.4 Other Types of Polymorphisms ............................................................... 38
4.6 RSE Adherence Assay ..................................................................................... 40

Chapter 5 Conclusion and Future Directions .......................................................... 44

References .............................................................................................................. 48
LIST OF FIGURES

Figure 2-1. Components that contribute to super-shedding .................................................. 9
Figure 2-2. Histological composition of the RAJ ................................................................. 10
Figure 2-3. RSE adherence patterns for two E. coli O157:H7 strains .................................. 12
Figure 4-1. Genome map of SS17 ....................................................................................... 19
Figure 4-2. Phage comparisons between SS17 and reference O157 strains ....................... 21
Figure 4-3. Cladogram of SS17 and E. coli O157:H7 outbreak strains .............................. 29
Figure 4-4. Mauve whole genome alignment of SS17 with reference O157 strains .......... 30
Figure 4-5. Mauve alignment showing sections of phage ..................................................... 32
Figure 4-6. SNP patterns of reference O157 genomes compared to SS17 ....................... 34
Figure 4-7. Adherence patterns of SS17 and reference O157 strains to RSE cells ............. 41
Figure 4-8. Adherence pattern of SS17 in the presence of antisera ................................. 43
LIST OF TABLES

Table 2-1. Features of *E. coli* pathotypes.................................................................3
Table 3-1. Characteristics of *E. coli* O157:H7 isolates used in this study......................14
Table 4-1. Genome statistics of SS17 and reference O157 strains ..................................18
Table 4-2. Virulence genes in SS17 with nsSNPs ..........................................................36
Table 4-3. Genes in SS17 with frameshifts and premature stop codons .............................39
LIST OF ABBREVIATIONS

A/E – Attaching and effacing lesions
AI – Autoinducer
ARS – Agricultural Research Service at the USDA
CDS – Coding DNA sequences
CFU – Colony forming unit
DAEC – Diffuse-adhering _Escherichia coli_
DAPI – 4’,6-diamidino-2-phenylindole
DNA – Deoxyribonucleic acid
_E. coli_ – _Escherichia coli_
EAEC – Enteroaggregative _Escherichia coli_
EHEC – Enterohemorrhagic _Escherichia coli_
EIEC – Enteroinvasive _Escherichia coli_
EPEC – Enteropathogenic _Escherichia coli_
ETEC – Enterotoxigenic _Escherichia coli_
FAE – Lymphoid follicle associated columnar epithelial
HEp-2 – Human epidermoid cancer cells derived from the larynx
HUS – Hemolytic-uremic syndrome
LEE – Locus of enterocyte effacement
LF – Lymphoid follicles
LSPA – Lineage-specific polymorphism assay
MAC-T – Bovine mammary epithelial cells
NADC – National Animal Disease Center at the USDA
nsSNP – Non-synonymous single nucleotide polymorphism
O157 – _Escherichia coli_ O157:H7 strains
ORF – Open reading frame
PC – Pseudocrypts
PFGE – Pulse field gel electrophoresis
PT – Phage type
QS – Quorum sensing
RAJ – Recto-anal junction
RAST – Rapid annotation using subsystem technology
rRNA – Ribosomal ribonucleic acid
RSE – Recto-anal junction squamous epithelial cells
SBI – Shiga toxin bacteriophage insertion site
SNP – Single nucleotide polymorphism
Sp – S-loops
SS17 – *Escherichia coli* O157:H7 super-shedder isolate 17
SSE – Stratified squamous epithelium
sSNP – Synonymous single nucleotide polymorphism
STEC – Shiga Toxin-producing *Escherichia coli*
T3SS – Type III secretion system
tRNA – Transfer ribonucleic acid
USDA – United State Department of Agriculture
I would first like to thank my adviser and committee members for their support and guidance with my studies and the preparation of my thesis. Secondly, I would like to thank and express my greatest appreciation to my lab members who have seen me through the good times and the bad. You have truly become like a second family, especially Tiffany Cunningham who kept me sane through this whole process. Finally, I would like to thank my family and friends who have encouraged and supported me throughout my graduate career and who I know will continue to support me throughout life.
Chapter 1

Introduction

Shiga toxin-producing *Escherichia coli* (STEC) are major foodborne pathogens that cause significant morbidity and mortality with symptoms ranging from bloody diarrhea and hemolysis to the development of life-threatening hemolytic-uremic syndrome (HUS) [1-3]. STEC infections cause over 265,000 infections in the United States each year with *E. coli* O157:H7 accounting for more than a third of these illnesses [4]. Unfortunately, effective control methods are lacking and the result is a number of large, deadly multi-state outbreaks caused primarily through the consumption of contaminated food products [4-8].

*E. coli* O157:H7 colonize the terminal end of the gastrointestinal tract in cattle which results in the asymptomatic carriage and excretion of varying levels of the bacteria. The majority of O157:H7 strains are shed at 10 to 100 CFU/g of feces [1]. Recently, isolates of *E. coli* O157:H7 have been found to be shed at greater than $10^4$ CFU/g of feces and are classified as “super-shedders” [9-15]. A major factor that has been found to contribute to super-shedding is the colonization of the recto-anal junction (RAJ), a unique junction located between two cell types which displays distinctive histology [9, 16, 17]. Little is known, however, about the microbial factors that contribute to super-shedding.

With this in mind, we developed the hypothesis that whole genome sequencing and comparative genomic analyses will help identify genetic features and polymorphisms that define super-shedder isolates of *E. coli* O157:H7. We sequenced the genome of a
super-shedder isolate and conducted comparative analysis from which polymorphisms unique to the super-shedder isolate were identified. A model of the RAJ was used to demonstrate the adherence pattern of the super-shedder isolate to bovine RSE cells. This enhanced adherence may play a role in the super-shedder phenotype and provide a plausible mechanism by which super-shedder isolates of *E. coli* O157:H7 are distinct from reference strains. The overall goal of the project is to discover the molecular mechanism or factors that contribute to super-shedding at the microbe level. This will enable the identification of the mechanism(s) involved in the adherence and colonization of the RAJ and allow the design of control methods that can target super-shedder isolates of *E. coli* O157:H7.
Chapter 2

Background and Significance

2.1 The Pathogenesis of *E. coli* O157:H7

*Escherichia coli* is a Gram-negative, rod-shaped bacterium found ubiquitously in the gastrointestinal tracts of humans and animals. The bacterium is also a facultative anaerobe, surviving with or without oxygen. There is great genetic diversity among this species as only about 20% of the genome is conserved among strains [18]. *E. coli* can be either a commensal or a pathogen. Commensal *E. coli* is found as part of the microflora and is thought to prevent the establishment of pathogenic bacteria in the gut [19]. Pathogenic *E. coli* have the ability to cause disease in humans and became pathogenic by the acquisition of virulence genes through the horizontal transfer of plasmids and by infection with bacteriophages [20]. There are six pathotypes of *E. coli* each grouped by pathogenesis, virulence factors, and clinical symptoms (Table 2-1) [3].

Table 2-1. Features of *E. coli* pathotypes

<table>
<thead>
<tr>
<th>Pathotype</th>
<th>Clinical Features</th>
<th>Virulence Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enteropathogenic (EPEC)</td>
<td>Watery diarrhea; vomiting</td>
<td>Bundle-forming pilus, attaching and effacing</td>
</tr>
<tr>
<td>Enterohemorrhagic (EHEC)</td>
<td>Watery to bloody diarrhea; hemolytic-uremic syndrome</td>
<td>Shiga toxins, attaching and effacing</td>
</tr>
<tr>
<td>Enterotoxigenic (ETEC)</td>
<td>Watery diarrhea</td>
<td>Pili, heat-labile and heat-stable enterotoxins</td>
</tr>
<tr>
<td>Enteroaggregative (EAEC)</td>
<td>Diarrhea with mucous</td>
<td>Pili, cytotoxins</td>
</tr>
<tr>
<td>Enteroinvasive (EIEC)</td>
<td>Dysentery; watery diarrhea</td>
<td>Cellular invasion, intracellular motility</td>
</tr>
<tr>
<td>Diffuse-adhering (DAEC)</td>
<td>Diarrhea</td>
<td>Induces the development of long cellular extensions</td>
</tr>
</tbody>
</table>

The most common pathotype of *E. coli* is the Shiga toxin-producing *E. coli* (STEC) which contains the subgroup, enterohemorrhagic *E. coli* (EHEC). Approximately 265,000 STEC infections occur each year in the United States [4]. Symptoms of a STEC infection include watery to bloody diarrhea, abdominal pain, and nausea and vomiting [1-3]. The most common complication of STEC infection is the development of hemolytic-uremic syndrome (HUS), characterized by the destruction of red blood cells (hemolytic anemia), acute kidney failure (uremia), and low platelet count (thrombocytopenia) [21]. HUS develops in less than 10% of people infected with STEC but carries a 5-10% mortality rate [22]. In addition, approximately 50% of surviving patients develop permanent renal damage.

STEC isolates carry a number of virulence factors on both the chromosome and a large ~90kb plasmid. Shiga toxins are the characteristic feature of STEC strains and cause localized and systematic damage to cells and can contribute to the development of HUS. Another virulence factor is the locus of enterocyte effacement (LEE) which is necessary for the development of the attaching and effacing (A/E) lesions characteristic of STEC infections [23, 24]. The plasmid encodes hemolysins which function to lyse erythrocytes to release hemoglobin and iron which the bacteria can utilize and contribute to the development of anemia in humans [1]. The STEC pathotype contains a number of serogroups which are divided into the O157 group and the non-O157 group which includes the outbreak serotypes O26, O45, O103, O104, O111, O121, and O145 [25].

*E. coli* O157:H7 itself is a common foodborne pathogen and causes ~36% of all STEC infections. One interesting feature of O157 strains is that they lack the ability to ferment sorbitol, allowing for the easy differentiation and detection of this pathogen. The
main source of transmission of *E. coli* O157:H7 and other STECs are ruminant animals, including cattle, sheep, and goats. Cattle are the principal animal reservoir in the United States and remain asymptomatic since the cells of the intestinal tract lack the receptor, globotriaosylceramide (Gb3), which binds to Shiga toxins [26]. Humans possess this receptor in the cells of the intestinal tract making them susceptible to the Shiga toxins and *E. coli* O157:H7.

### 2.2 Definition of Super-Shedding

Recently, a subset of *E. coli* O157:H7 isolates have been obtained from cattle and classified as “super-shedders”, designated by excretion levels at greater than $10^4$ CFU/g of feces [9-15]. Cattle carry the bacteria asymptomatically in their gastrointestinal tract, predominantly in the terminal recto-anal junction (RAJ). Colonization of this region has been determined to be necessary for the production of the high bacteria levels characteristic of super-shedding animals [9]. Since most *E. coli* O157:H7 strains are shed by cattle at 10 to 100 CFU/g of feces, there are several orders of magnitude greater potential of super-shedders to contaminate the surrounding environment and the food supply [1].

A recent study by our collaborator, Terrance Arthur from the USDA-ARS Meat Animal Research Center, determined strain-specific characteristics common to 102 super-shedders isolates from the United States including the isolate used in the current study (submitted for publication). Pulsed field gel electrophoresis (PFGE) segregated the isolates into 52 unique genotypes with the most populated containing 19 isolates and
matching one of the most common human illness genotypes. Phage typing was also conducted and revealed a predominance of 30% of the isolates for phage type (PT) 4 which differs from the PT 21/28 common among super-shedder isolates on Scotland farms [27]. To assess super-shedder relatedness, Arthur used lineage-specific polymorphism assay (LSPA) to determine the dominate lineage. Lineage I is associated with human disease outbreaks, lineage II with bovine isolates, and lineage I/II, which shares characteristics of both lineages, has been implicated in human *E. coli* O157:H7 outbreaks [28]. All three lineages were found in the 102 super-shedder isolates at roughly equal proportions, with the largest proportion of isolates (36%) belonging to the lineage I/II. Other strain typing methods include a nucleotide base substitution in the *tir* gene (A or T allele), Shiga toxin bacteriophage insertion sites (SBI), and the antiterminator *Q* gene [29-32]. The super-shedder isolates contained the T allele of *tir* more often than the A allele with all of the lineage I and I/II isolates harboring the T allele and lineage II isolates harboring the A allele most often at 86%. The T allele has been associated with human illness and isolates that harbor the allele may have a higher likelihood to cause disease in humans [29]. SBI genotyping was also used to sort the super-shedder isolates into clusters based on the insertion site of the Shiga bacteriophage. Analysis of clinical and bovine strains showed the majority of clinical isolates belong to clusters 1, 2, and 3 [30]. The super-shedder isolates were grouped into 12 clusters with the majority (55.9%) belonging to clusters 1-3. The antiterminator *Q* gene has been found to have a nonrandom distribution between human and bovine *E. coli* O157:H7 isolates and is a genetic marker that can be used to group the super-shedder isolates. More often, the *Q* gene of bacteriophage 933W (*Q*933) is detected in clinical isolates and
is in 59.8% of super-shedder isolates. The Q933 gene has been associated with increased Shiga toxin production when compared to isolates that harbor the Q21 variant or isolates harboring both gene variants [32]. The analysis of super-shedder relatedness revealed no exclusive clustering of isolates indicating a specific genotype is not responsible for the super-shedder phenotype. However, there is a tendency of the super-shedder isolates to cluster similar to human clinical isolates even though all of the super-shedder isolates were obtained from cattle.

2.2 Prevalence of Super-Shedding

In the process of mathematical modeling the transmission rates of E. coli O157:H7, a number of researchers have determined the best way to explain the high transmission rates yet low pathogen prevalence is if a small population of animals are high pathogen shedders or “super-shedders”. Indeed, researchers have shown that less than 10% of animals are super-shedders, yet greater than 96% of fecal E. coli O157:H7 isolates originate from super-shedding animals [33]. In addition, a survey of Scotland farms determined that super-shedders account for 80% of the transmission events within herds even with pathogen prevalence rates at less than 10% of animals [11].

These higher transmission rates are caused by the increased likelihood of susceptible cattle coming in contact with the high levels of bacteria, since as little as 300 CFU is needed to infect calves in a dose dependent manner [12]. With a low infectious dose of 100 CFU, humans are also at a 100-fold increased risk of being infected by a super-shedder isolates [12]. Thus, an animal that is a super-shedder and excreting large
numbers of *E. coli* O157:H7 will pose a greater risk of spreading the pathogenic bacteria, both to other cattle and into the food supply, than the combined output of many animals that shed bacteria at low levels [15]. Modeling of super-shedders suggests that if colonization could be prevented in the less than 10% of super-shedding animals, then the spread of *E. coli* O157:H7 could be prevented [11].

2.3 Contributing Factors to Super-Shedding

What makes an isolate of *E. coli* O157:H7 a super-shedder is unknown; however, it is known that three principle components likely contribute to the super-shedder phenomenon. Those components include (1) the microbe (2) the host and (3) the environment (Figure 2-1). Factors that may influence super-shedding at the microbe level include strain-specific genomic characteristics such as the presence of virulence and adherence genes, strain lineage, and phage type. In addition, growth rates, biofilm formation, nutrient utilization, and the ability of the bacteria to survive and persistent in the environment may also contribute to super-shedding [9, 12, 13]. Cattle host factors consist of the genotype and both the innate and adaptive immune response which function to clear bacteria [10, 12]. The environmental factor can be further divided into the external environment of the host and the internal environment where the microbe inhabits. The external host environment includes factors such as the season and climate, both of which are known to effect shedding levels, the location the animal is housed and the proximate to other cattle, and management practices [13]. The internal environment contains factors such as the composition of the host microflora and the diet
Microbial, host, and environmental factors are all expected to contribute to the super-shedding phenomenon. Microbial factors include genetic makeup and the ability to adhere to cells while host factors include the genotype and immune response of the animal. The environmental factor can be divided into two groups; the external host environment which includes housing and management practices and the internal environment which is influenced by the surrounding microflora.

2.4 The RAJ and a Realistic in vitro Adherence Assay

The molecular mechanisms that underlie the super-shedder phenotype are unknown; although it is known that colonization of the terminal recto-anal junction (RAJ) in cattle is directly related to increased fecal shedding [1, 9, 12, 14, 15]. The RAJ is a unique region composed of lymphoid follicle-associated columnar epithelial (FAE) cells of the rectum and stratified squamous epithelial (SSE, also known as RAJ SSE or RSE) of the host, which both may affect the microbe and its niche [12]. Understanding the roles of each of these areas is important for the development of new methods to mitigate the high shedding and transmission of super-shedder isolates, and ultimately decrease food-borne illnesses associated with *E. coli* O157:H7.
cells of the anal canal (Figure 2-2) [9, 16]. Localization in the RAJ is unique to O157:H7 as other serotypes are found in similar levels throughout the large intestine without the congregation seen for O157:H7 at the RAJ mucosa [17].

![Image](https://via.placeholder.com/150)

Figure 2-2. Histological composition of the RAJ

The recto-anal junction (RAJ) is defined as the region between the stratified squamous epithelium (SSE) and columnar epithelial (FAE) with pseudocrypts (PC) adjacent to lymphoid follicles (LF). The SSE lines the anal canal while the FAE lines the rectum and large intestine. The hematoxylin-and-eosin-stained section is at 2.5x magnification.


Of particular interest are genes and proteins that contribute to the adherence and colonization of *E. coli* O157:H7 to the cell types present in the RAJ. Recent evidence has found that *E. coli* O157:H7 uses different mechanism of adherence to bind to and colonize the FAE cells and the RSE cells present at the RAJ. For FAE cells, experimental evidence has shown the LEE4 operon to be essential for the adherence and colonization of *E. coli* O157:H7 to these cells [34]. The LEE4 operon encodes for the proteins EspA, EspB, and EspD which forms the pore for the translocation of the intimin receptor, Tir and other secreted factors, included the LEE4-encoded EspF. LEE4
knockout mutants demonstrated decreased adherence to FAE cells in vitro and decreased fecal shedding in orally inoculated calves [34].

Since LEE encoded proteins mediate attachment to mucosal epithelial cells and are considered critical for the adherence of E. coli O157:H7 to FAE cells, a recent study determined the involvement of LEE proteins in the adherence of E. coli O157:H7 to RSE cells [35]. Using antisera against intimin, Tir, EspA, and EspB, and an intimin knockout mutant, Kudva, et al. determined that the lack of these proteins did not block the adherence of E. coli O157:H7 to RSE cells [35]. This suggests a mechanism of adherence that is independent of these LEE encoded proteins.

To elucidate possible mechanisms used by E. coli O157:H7 to adhere to RSE cells, the proteome of E. coli O157:H7 grown under the adherence assay conditions was defined. An evaluation of the proteome resulted in 684 expressed proteins of which 513 have been previously characterized and are involved in metabolism, cell division, regulation, transportation, and virulence. Of particular interest are proteins that are associated with adherence or virulence and include intimin, Tir, EspB, Iha, OmpA, LuxS, KatP, EspP, Stx1A and B, and Stx2B. In addition, 36 potential adhesins were identified in the proteome of which three potential adhesins, ChuA, TerE, and a putative outer membrane porin protein, are unique to E. coli O157:H7 [35].

Experimental evidence has shown that swabbing the RAJ directly with E. coli O157:H7 results in infection and carriage of the bacteria similar to cattle that have been naturally infected [36]. This indicates the importance of the RAJ in mimicking natural infections. Kudva, et al. has developed an in vitro assay using RSE cells obtained from the RAJ of cattle [16]. E. coli O157:H7 is incubated with the RSE cells and
immunofluorescent staining used to label the bacteria and cells. A unique scoring mechanism classifies the adherence strength as strongly positive, moderately positive, or non-adherent based on the number of bacteria attached to each RSE cell. In addition, the type of adherence pattern is described as aggregative, diffuse, or negative based on the arrangement of the bacteria attached to the RSE cells [16]. The *E. coli* O157:H7 strains, 86-24 and EDL933, displayed a strong, diffuse and a moderate, aggregative adherence pattern in the RSE *in vitro* assay, respectively (Figure 2-3). The strong, diffuse pattern is similar to the adherence pattern seen in experimentally infected calves. Since the RAJ is the primary colonization site of *E. coli* O157:H7, it is important to determine the ability of super-shedder isolates to adhere to and colonize this site.

Figure 2-3. RSE adherence patterns for two *E. coli* O157:H7 strains

*E. coli* O157:H7 strain 86-24 displays a strong, diffuse adherence pattern while the EDL933 strain displays a moderate, aggregative adherence pattern on RSE cells. Arrows point to bacteria (green fluorescence). The cytokeratin of the RSE cells have red fluorescence while their nuclei have blue fluorescence.

Chapter 3

Materials and Methods

3.1 Bacterial Strains

Super-shedder isolates of *E. coli* O157:H7 were obtained and initial characterized by Terrance Arthur at the USDA-ARS Meat Animal Research Center. Fecal samples were collected by swabbing the RAJ from approximately 1,500 cattle in commercial feedlots and 3,500 cattle in commercial processing plants at slaughter. Isolates were obtained in the summer months over a two year period in Midwestern States. Samples were placed into 4 mL of tryptic soy broth (TSB; Difco, Becton Dickenson, Sparks, MD) and shipped to the laboratory on ice. Upon arrival, samples were enumerated for *E. coli* O157:H7 by plating onto CHROMAgar O157 (Becton Dickenson, Sparks, MD), and colony forming units (CFU) per swab were calculated [37]. Animals were classified as super-shedders when counts were greater than $10^4$ CFU/swab [12]. Up to twenty colonies were picked for PCR to confirm that each *E. coli* isolate harbored genes for the O157 antigen, H7 flagella, γ-intimin, and at least one of the Shiga toxin genes [38]. Isolates were characterized by phage typing and *XbaI* PFGE as described [39]. This sampling produced 102 super-shedder isolates from which isolate 17 (SS17) was chosen to represent genetic diversity. SS17 is phage type 4 and had $3.1 \times 10^6$ CFU/swab.

Well-characterized strains of *E. coli* O157:H7 were used for comparative analysis (Table 3-1). Although it is unknown whether these reference genomes truly represent
non-super-shedder strains, they will be considered as such. In addition to EDL933, strain 86-24 was used as a control O157 in the RSE cell adherence assay.

Table 3-1. Characteristics of *E. coli* O157:H7 isolates used in this study

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Origin</th>
<th>Isolation Source</th>
<th>Outbreak Source</th>
<th>Shiga Gene Profile</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW14359</td>
<td>Michigan</td>
<td>Human</td>
<td>Spinach</td>
<td><em>stx</em>1-, <em>stx</em>2+, <em>stx</em>2c+</td>
<td>[40]</td>
</tr>
<tr>
<td>EC4115</td>
<td>Maine</td>
<td>Human</td>
<td>Spinach</td>
<td><em>stx</em>1-, <em>stx</em>2+, <em>stx</em>2c+</td>
<td>[41]</td>
</tr>
<tr>
<td>Sakai</td>
<td>Japan</td>
<td>Human</td>
<td>Radish Sprouts</td>
<td><em>stx</em>1+, <em>stx</em>2+, <em>stx</em>2c-</td>
<td>[42]</td>
</tr>
<tr>
<td>EDL933</td>
<td>Michigan</td>
<td>Ground Beef</td>
<td>Hamburger</td>
<td><em>stx</em>1+, <em>stx</em>2+, <em>stx</em>2c-</td>
<td>[43]</td>
</tr>
<tr>
<td>86-24</td>
<td>Washington</td>
<td>Human</td>
<td>Ground Beef</td>
<td><em>stx</em>1-, <em>stx</em>2+, <em>stx</em>2c-</td>
<td>[44]</td>
</tr>
<tr>
<td>SS17</td>
<td>Midwest</td>
<td>Bovine Feces</td>
<td>NA*</td>
<td><em>stx</em>1-, <em>stx</em>2+, <em>stx</em>2c+</td>
<td>This Study</td>
</tr>
</tbody>
</table>

*NA*-not applicable, SS17 is not an outbreak strain

3.2 Genomic DNA Extraction

Genomic DNA was isolated using QIAGEN DNeasy Blood and Tissue Kit (cat. no. 69504). Briefly, 1 mL of overnight culture was centrifuged for 5 min at 10,000xg in a microcentrifuge. The pellet was resuspend in provided buffer (ATL) and digested with 20 μl proteinase K (20mg/ml) for 30 min at 56°C with a 5 sec vortexing halfway through the incubation. Cells were then processed as dictated in the protocol and eluted from the column with 100 μl of ddH2O. DNA was quantified using Nanodrop (Thermo Scientific, Wilmington, DE) and aliquots stored at -20°C. Approximately 100 ng of purified DNA was submitted for sequencing.
3.3 Genome Sequencing and Assembly

Whole genome shotgun sequencing with genomic DNA was conducted by the Genomics Core Facility at Penn State University using the Ion Torrent PGM sequencer (Life Technologies, Grand Island, NY) [45]. Over 4.2 million reads with an average length of 165 ± 31 bp were generated in two runs using a 316 and a 318 chip to provide over 100-fold (137.5) coverage of the genome. Reference-guided assembly using *E. coli* O157:H7 strain EC4115 was conducted using DNASTAR SeqMan NGen 3.1 (Madison, WI). Lasergene SeqMan Pro 9 (DNASTAR) was used for further assembly and analysis.

A whole genome restriction optical map was generated using *BamH*I digestion by OpGen, Inc. (Gaithersburg, MD) to provide an isolate-specific reference on which to compare the assembly [46]. In brief, high molecular weight DNA was extracted, immobilized, and subjected to restriction digest preserving the fragment order. The fragment lengths are measured using fluorescent microscopy and assembled by overlapping fragments to generate a map of restriction cut sites.

The genome was closed and confirmation of ambiguous or low coverage areas was conducted using manual primer walking strategies and Sanger sequencing. The final assembly was compared to the optical map by generating an *in silico* restriction map of the sequence data.

3.4 Annotation and Comparative Genomics

Initial automated annotation was performed using RAST (Rapid Annotation using Subsystem Technology; [47]) with manual annotation using Artemis [48] and BLASTp
(NCBI) to confirm genes. Whole genome comparative analysis was conducted using Mauve [49] and the progressiveMauve algorithm [50] to identify SNPs, genomic islands, and relative relationships between SS17 and sequenced reference strains (Table 3-1). Lineage-specific polymorphism assay (LSPA) was conducted using primers described in Yang, et al. [28] and Sanger sequencing to confirm insertions and deletions. Further analysis of polymorphisms was conducted using a program developed by our collaborator, Michael Mwangi in the Departments of Veterinary and Biomedical Science, and Biochemistry and Molecular Biology at Penn State University. Circle images were generated using GenVision 10 (DNASTAR). The cladogram image was created using FigTree v1.3.1 (Institute of Evolutionary Biology, University of Edinburgh, http://tree.bio.ed.ac.uk/) from data generated by Mauve.

3.5 RSE Cell Adherence Assay

The RSE cell adherence assays were conducted by Indira Kudva at the USDA-ARS National Animal Disease Center (NADC). SS17 and control O157:H7 strains, EDL933 and 86-24, were cultured overnight in Dulbecco Modified Eagle Medium-Low Glucose (DMEM; Gibco/Invitrogen Corporation, Grand Island, NY) at 37°C without aeration, pelleted and resuspended in sterile saline as previously described [16]. RSE cells were suspended in 1 ml DMEM–No Glucose (DMEM-NG) ± 2.5% D + Mannose to a final concentration of $10^5$ cells/ml. For each strain, bacteria were mixed with RSE cells to achieve a final bacteria to cell ratio of 10:1. The mixture was incubated at 37°C with shaking for 4 hrs. The mixture was then pelleted, washed thoroughly, and reconstituted.
to 100 μl in ddH₂O. Drops of the suspension (2 μl) were placed on Polysine (Thermo Scientific Pierce, Rockford, IL) slides and dried overnight under direct light to quench non-specific fluorescence, before fixing in cold 95% ethanol for 10 min.

The slides were stained with 1% toluidine blue, or with DAPI and fluorescence-tagged antibodies specific to the O157 antigen and cytokeratins of the RSE cells as previously described [16]. Adherence patterns on RSE cells were recorded as aggregative, diffuse, or non-adherent [35]. The percent of RSE cells with or without bacteria adhering to them was determined and recorded as strongly positive when more than 50% of RSE cells had >10 bacteria attached, as moderately positive when 50% or less of the RSE cells contained 1–10 adherent bacteria or as non-adherent when less than 50% of the RSE cells had only 1–5 adherent bacteria.

In order to determine the involvement of LEE-encoded proteins, rabbit antisera generated against EspA, EspB, Tir, and intimin (stock from NADC), were pooled and tested at a 1:50 dilution [35]. Assays were preformed as described above with the antisera added to the resuspended bacterial pellets and incubated at 37°C for 30 min before mixing with the RSE cell suspension [35]. Human HEp-2 cells were also used in place of RSE cells for comparative purposes.
Chapter 4

Results

4.1 SS17 Genome

Recent research suggest super-shedder isolates play a major role in the prevalence and transmission of *E. coli* O157:H7 but little is known about what makes super-shedders different at the genomic level. With this in mind, we set out to completely sequence and annotate the genome of super-shedder isolate 17 (SS17). Sequencing revealed a common O157:H7 chromosome size of 5,523,848 bp with a G+C content of 50.5% (Table 4-1). Within the genome, SS17 encodes 5,442 CDS at an average length of 898bp. With 88.5% of the genome coding for proteins, SS17 has a gene density of 0.985 genes per kb. In addition, there are 107 tRNAs consisting of 8,019 bp and 22 rRNAs with 32,212 bp (Figure 4-1) which is consistent with other sequenced *E. coli* O157:H7 strains (Table 4-1).

<table>
<thead>
<tr>
<th>Table 4-1. Genome statistics of SS17 and reference O157 strains</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SS17</strong></td>
</tr>
<tr>
<td>-----------------</td>
</tr>
<tr>
<td>Length of Sequence (Mb)</td>
</tr>
<tr>
<td>G+C ratio (%)</td>
</tr>
<tr>
<td>Coding DNA Sequences (CDSs)</td>
</tr>
<tr>
<td>Protein Coding Region (%)</td>
</tr>
<tr>
<td>Average ORF length (bp)</td>
</tr>
<tr>
<td>No. rRNA</td>
</tr>
<tr>
<td>No. tRNA</td>
</tr>
<tr>
<td>No. Plasmids</td>
</tr>
</tbody>
</table>
In addition to the chromosomal DNA, we extracted plasmid specific sequences that revealed SS17 to contain two plasmids, pO157 (94,645 bp) and a smaller plasmid we labeled as pSS17 (37,446 bp) (Figure 4-1). The larger plasmid, pO157, contains 111 genes with an average length of 744 bp. Approximately 87% of the plasmid is protein coding to give a gene density of 1.172 genes per kb. The pO157 plasmid encodes for the hemolysin genes, *hlyABCD*, the cytotoxin, *toxB*, the type V secretion serine protease, *espP*, and the *etp* operon of type II secreted proteins (Figure 4-1). The pSS17 is similar to the second plasmid in EC4115, and contains 51 genes with an average length of 619
bp. This results in a density of 1.361 genes per kb to give a protein coding percentage of 84.3%. The genes in pSS17 encode for a number of conjugal transfer, type IV secretion, and hypothetical proteins.

4.2 Phage Regions in SS17

One of the most common methods of *E. coli* O157:H7 to acquire diversity is through the acquisition of bacteriophage genes and the potential virulence genes encoded in the phage. With this in mind, we set out to identify and describe the phage regions present in SS17. We also compared the insertion sites and makeup with reference *E. coli* O157:H7 genomes, and present those results in Section 4.4.1. Evaluation of the SS17 genome revealed 22 regions that house phage distributed over the entire genome (Figure 4-2). All of the phage regions contain an integrase and phage assembly proteins with most of the areas containing one or more transposases. Regions with additional genes of interest are described below.

The first phage region houses an integrase, phage assembly proteins, and an invertase. The second region contains the non-LEE-encoded secreted effectors, *nleB1*, *nleC*, *nleH1*, and *nleD*, a DNA repair gene and *lomK*, an attachment invasion locus protein. The third phage region contains 2 tRNAs and an attachment invasion protein. The fourth region contains the non-fimbrial adhesins, *cah* and *iha*, the *ure* and *ter* operons, and *espP*. The fifth region contains the non-LEE-encoded type III secreted effectors, *espX7* and *espN*. The sixth and seventh regions mainly house phage assembly proteins. The eighth region contains three tRNAs and *nleA*, *nleH2*, and *nleF*. The ninth
region contains 5 tRNAs, *lomU*, and a number of non-LEE-encoded secreted effectors, including *nleG7*. Region 11 contains 3 tRNAs, *espJ*, and *tccP*, the tir-cytoskeleton coupling protein. The 12\textsuperscript{th} region contains the *yeeT-W* genes which includes the YeeV-YeeU toxin-antitoxin system. The 13\textsuperscript{th} region contains the *stx2A* subunit, *stx2Bc* variant and 3 tRNAs while the 16\textsuperscript{th} region contains the *stx2A* and *stx2B* subunits, 3 tRNAs, and *lomW*, an outer membrane protein. Region 18 houses the *espM2* and *espW* genes with the 19\textsuperscript{th} region containing the *nleB2* and *nleE* genes. The 20\textsuperscript{th} region encompasses the LEE-encoded genes including *eae, tir*, and *ler*. The 21\textsuperscript{st} region contains *lomU* and a phage *eae* protein (Figure 4-2).

![Figure 4-2. Phage comparisons between SS17 and reference O157 strains](image)

Location of phage regions were determined in each strain and color coordinated based on insertion sites and similarities. From inner ring: number designation for SS17 phage regions, phage locations: SS17, EC4115, TW14359, EDL933 and Sakai, and S-loop numbers for Sakai phage. Black arrows indicate regions of phage that are present in SS17 and not in EDL933 and Sakai. Pink arrows indicated regions that are translocated in SS17 compared to EDL933 and Sakai.


4.3 Virulence Genes

In an analysis of the genes encoded by SS17, we identified 286 virulence related genes that have been previously described in *E. coli* O157:H7 strains. Of these, 267 genes are on the chromosome and 19 are in the pO157 plasmid, and can be divided into three distinct groups (toxin, adherence, and other virulence-associated) based on protein function. The toxin group contains 106 genes and the adherence group 102 genes. The virulence-associated group contains 78 genes and includes genes that function in motility, nutrient acquisition, and biofilm formation. Many of these genes serve multiple functions within the various virulence mechanisms employed by *E. coli* O157:H7 when invading a host.

4.3.1 Toxin Genes

Toxins are molecules that either damage host cells or corrupt normal cellular activities and functions. SS17 encodes a number of these important toxins that have been previously described in reference *E. coli* O157:H7 strains. The locus of enterocyte effacement (LEE) is a well known feature of O157:H7 strains, and is critical for their ability to cause infection in humans [3]. This locus contains genes which are essential for the intimate cell-to-cell contact *E. coli* O157:H7 has with eukaryotic cells and the development of the attaching and effacing (A/E) phenotype [23, 24]. This is accomplished through a series of cellular effectors which disrupt host cell homeostasis, and proteins which mediate the expression, biogenesis, and delivery of these effector proteins, via the type III secretion system (T3SS). The LEE is arranged into five
polycistronic operons (LEE1 to LEE5) [51]. SS17 has 47 LEE encoded genes including the secreted effectors, escR and escS in LEE1, and espF in LEE4. In addition to the LEE encoded proteins, SS17 encodes 35 non-LEE encoded effectors that also utilize the type III secretion system. The few non-LEE proteins that have been studied have been shown to play a role in mediating host responses similar to the LEE operons [52]. SS17 contains the nleA-H genes that are translated into secreted effectors.

Another section of the toxin group are genes that are cytotoxic and include the Shiga toxin (stx) genes, and the plasmid encoded hemolysin (hly) genes. SS17 is positive for stx2 and stx2c but not stx1. Similar to EC4115 and TW14359, SS17 has two copies of stx2A, one located proximal to the stx2c variant and the other located near stx2B. An evaluation of the ability of SS17 to produce Shiga toxin revealed that the toxin was produced without induction (observational data). To elucidate the possible genomic factors that may influence the production of Shiga toxins, we examined the location of the anti-terminator Q genes which have been shown to influence Shiga toxin production [53, 54]. Studies have shown that the presence of only the Q933 variant results in higher production of Shiga toxin than the presence of the Q21 variant alone or the presence of both [32]. In SS17, Q933 is located upstream of stx2B and Q21 is located upstream of stx2c, indicating this may not be the cause of the high Shiga toxin production observed from the super-shedder isolate.

In addition to the cytotoxin stx genes found on the chromosome, the pO157 of SS17 encodes for the hemolysin genes, hlyABCD. Hemolysins are incredibly important for survival of the bacteria in the host [55]. They give the bacteria the ability to gain
access to an essential nutrient, iron, by freeing heme from hemoglobin. The free heme can be taken up and utilized by the bacteria and allow for survival in the host.

4.3.2 Adherence Genes

Of particular interest are genes related to adherence and colonization that may play a role in the unique phenotype exhibited by *E. coli* O157:H7 super-shedder isolates. SS17 encodes for 102 adherence related genes previously identified in O157:H7 strains. These include genes for fimbrial-like adhesins, non-fimbrial-like adhesins, ushers, chaperones, regulators, and effectors of adherence.

Fimbrial-like adhesins have been implicated in the attachment of bacterial cells to many different human cell types [56]. These filamentous adhesins are made up of a series of chaperones, ushers, and subunits to compose the fimbrial complex. Similar to the references strains of *E. coli* O157:H7, SS17 encodes for the *fim*, *sfm*, and *csg* operons. The *fim* operon encodes for the type I fimbriae and the *sfm* operon encodes for a putative chaperone-usher fimbria [57, 58]. The *csg* operon regulates the expression of curli and is not only important for adhesion to different surfaces including fibronectin and laminin found on eukaryotic cells, plant cells, and inert surfaces, but is also an important set of genes for biofilm formation [59].

Non-fimbrial-like adhesins are also important for adherence of O157:H7 strains. A number are encoded in the SS17 genome, and include the LEE5-encoded *eae* (intimin) and *tir* (translocated intimin receptor), and *tccP* (*tir* cytoskeleton coupling protein). These genes and their proteins are essential in the formation of attaching and effacing
(A/E) lesions on epithelial cells. Two other non-fimbrial-like adhesins, yeeJ and yfaL, have been shown to increase the efficiency of adherence to epithelial cells. Also encoded in SS17 is the iron regulated adhesin, iha which is not only important for adherence to epithelial cells but also serves as an enterochelin transporter [60]. This protein has the potential to aid in the adherence and colonization of the bacteria to host cells and to supply an additional mechanism by which the bacteria can obtain iron.

Numerous ushers, chaperones, and regulators associated with both the fimbrial-like and the non-fimbrial-like adhesins are found in the SS17 genome. Usher proteins that are encoded in SS17 include the fimbrial usher, htrE and the curli assembly and transport usher, csgG. Chaperone genes include the pilin chaperone, sfmC and the type I fimbriae, fimC. Two important global transcriptional regulators for many types of adhesins are crl and hha [61, 62]. The regulator crl, controls the curli operon which has a second regulator, csgD and an oscillating regulation, fur [62]. The positive regulator, hha, controls flagellar and curlin gene expression, but is also an important repressor of the hemolysin genes [61]. In addition to adhesins and their regulators, effectors of adherence also play an important role in adherence. The etp operon which is found in the pO157 of SS17, encodes for the type II secretion effectors.

4.3.3 Other Virulence-Associated Genes

The third group of virulence genes are other virulence-associated genes. SS17 encodes 78 genes that are involved in regulating virulence as well as survival in the host through evasion of host defenses and the formation of biofilms. The group contains a
wide variety of genes such as those for antimicrobial resistance, motility, signal response, and nutrient acquisition. Resistance genes include the ter and sap operons which are dedicated to protecting the bacteria against phage, colicins and tellurite, and human derived antimicrobial peptides, respectively. The mot and fli operons control the expression and function of the flagella which have implications in both motility and adherence.

An extremely important communication process is quorum sensing (QS) by which bacteria can communicate with one another and with the environment. Three QS systems, sdiA, luxS, qseE/F with auto-inducers 1, 2, and 3, respectively, are encoded in SS17. The most well known of these systems is the qseE/F, which responds to auto-inducer 3 (AI-3) and can also respond to human epinephrine and norepinephrine [63]. The system activates the LEE operons and initiates the A/E phenotype which in turn activates subsequent virulence factors [64]. Quorum sensing is also tightly linked to the development of biofilms [64]. SS17 also encodes for the major genes involved in biofilm and capsule biogenesis which include the wca and pga operons.

*E. coli* O157:H7 employ a series of nutrient acquisition genes to take advantage of the virulent effects it causes on host cells, such as the freeing of heme. The chu operon is expressed in response to low intracellular levels of iron and will compete for the free heme released by destroyed host cells, in concert with the array of other iron sequestering proteins.
4.4. Phylogenetic Analysis

Phylogenetic grouping allows the separation of *E. coli* strains into four groups (A, B1, B2, and D) based on the presence of *chuA*, *yjaA*, and an unnamed DNA fragment [65]. Studies have shown that more virulent strains belong to the B2 and D groups [65, 66]. SS17 belongs to the D group as the isolate contains the *chuA*, but not the *yjaA* gene or the unnamed DNA fragment. Further analysis of the genome sequence places SS17 in clade 8 [67]. Clade 8 strains have been shown to have 2-fold greater adherence to bovine epithelial cells (MAC-T) and increased expression of virulence genes, including the LEE-encoded genes, *espAB*, *tir*, and *eae, stx2*, and the plasmid encoded *hlyA, toxB*, and *tagA* [68]. Recent studies have shown that clade 8 strains of *E. coli* O157:H7 are also categorized as lineage I/II [69, 70]. Lineage classification is determined by the lineage-specific polymorphism assay (LSPA) which is based on six polymorphic loci to divide strains into three major lineages; lineage I, lineage II and lineage I/II [28]. LSPA classifies SS17 as lineage I/II (211111).

4.5 Comparative Genomics

One of the important aspects of this study is to determine genomic level factors that have the potential to contribute to the super-shedder phenotype, specifically genetic loci that can be used to distinguish and define super-shedder isolates. However, it is very difficult to distinguish a non-super-shedder from a super-shedder isolate since fecal shedding levels are temporal [12]. An isolate may be produced at high levels during the warmer months but if obtained in the colder months may not achieve the high fecal levels
needed to be classified as a super-shedder. In addition, a low shedding isolate may be obtained during the same season and even at the same time as a high shedder and still not be considered a non-super-shedder as the peak shedding for the low isolate may have been reached before sampling. With this in mind, we used the previously sequence *E. coli* O157:H7 genomes as references which were considered to be distinct from SS17 (Table 3-1). We preformed whole genome comparisons and identified polymorphism unique to SS17, concentration on those polymorphisms that are in virulence related genes.

### 4.4.1 Whole Genome Comparison

We examined the relatedness of the genome of SS17 and previously characterized O157:H7 genomes including those associated with foodborne outbreaks and those isolated from bovine sources. This whole genome analysis revealed a close clustering of SS17 with the lineage I/II genomes associated with spinach outbreaks and a segregation from the lineage I outbreak and the lineage II bovine isolates (Figure 4-3). This is consistent with the LSPA analysis which classifies SS17 into the lineage I/II and the close homology of SS17 with TW14359 and EC4115 seen in the whole genome comparisons described below.

To determine differences in SS17 as compared to the reference O157:H7 strains across the whole genome, we used Mauve to divide the genomes into blocks of homology [49, 50]. When SS17 is compared to the reference strains (TW14359, EC4115, EDL933, and Sakai), there are 5,121,820 bp conserved among all isolates (~93%). In addition,
there is approximately 42kb of non-homologous regions in SS17, ranging from 20bp to 6,296bp with an average length of 822bp. The majority of the differences seen in the alignment stem from phage coding areas.

Mauve alignments of SS17 with the four reference genomes revealed six blocks of homology of varying lengths ranging from ~50kb to ~2280kb. The smallest block contains the stx2 gene and is inverted in EDL933 and Sakai with SS17 similar to TW14359 and EC4115 (Figure 4-4, block 5). In addition, the inversion in EDL933, that is ~400kb and spans the replication terminus [43], is not seen in SS17 or in the other references (Figure 4-4, block 3). In the first block of ~1250kb, four areas in SS17 show
The six blocks of homology are outlined in different colors with lines connecting corresponding blocks in each reference strain (TW14359, EC4115, EDL933, and Sakai). The different colors within a block indicated homology between strains. Mauve-all strains, light green-SS17, TW14359, and EC4115, yellow-reference strains only, white-unique to isolate.

differences in phage regions among all reference strains. The first three areas are similar to TW14359 and EC4115 and dissimilar to Sakai and EDL933. The fourth area in block 1 is ~32kb, has similarity to TW14359, EC4115 and Sakai, and is part of the phage 3 region (Figure 4-2).

In the second homology block of ~460kb, five areas show non-homologous phage regions with the first area containing a transposase and a hypothetical protein common to SS17, TW14359, and EC4115. The second (phage region 5) and fourth area (phage 7) in
SS17 have similarity to TW14359, EC4115, and Sakai. The third region (phage 6) contains phage that is similar to TW14359 and EC4115 and dissimilar to EDL933 and Sakai. The fifth region is ~5kb at the end of the block, is similar in TW14539 and EC4115 and contains part of the phage region 8.

The third block of ~400kb, which is inverted in EDL933, contains three regions that display non-homologous regions of phage among all of the reference strains. The first region in SS17 is similar to TW14359 and EC4115 and is adjacent to the last region of block 2. The second region is similar to TW14359 and EC4115 and is also within phage region 8 which is ~45kb larger in SS17 when compared to Sp9 of Sakai. The last region (phage 9) is at the end of the block and is similar to TW14359 and EC4115 (Figure 4-4, block 3).

The fourth block is ~1090kb long and contains nine regions that are non-homology when SS17 is compared to the references. The first region contains a transposase, the second region a hypothetical protein and a transposase, and the fourth region two transposase, that are found in SS17, TW14359 and EC4115 and not in EDL933 and Sakai. The third region (phage 11) is similar in some regions to all of the references except EDL933, to Sakai and TW14359 only, or to just Sakai (Figure 4-5). The fifth region (phage 12) contains phage that is similar to all of the references except EDL933. Unlike Sakai and EDL933, the sixth area is the phage region 13 which is ~54kb and is located between dacD and sbcB (Figure 4-5). This area includes a copy of the nleG7 gene and the stx2c variant and is similar to the Sp10 in Sakai located between ydaO and uspF (Figure 4-2, pink arrow). The seventh (phage 14) region contains areas that are similar to TW14359 and EC4115, is located between genes yegQ and yegS, is
Whole genome alignment using Mauve shows two enlarged region of phage compared between SS17 and the four references. In the first panel, part of region 11 (Sp14 in Sakai) is highlighted showing the differences between SS17 and the references which can also be seen in the annotations (white bars underneath). The different colors represent the homologous regions when SS17 is compared to the references with the white indicating non-homology. In the second panel, region 13 is shown which has similarity to Sp10 is Sakai. This region is only at this location in SS17, TW14359 and EC4115. Outlined cursor shows corresponding locations.

~32kb and contains an integrase and phage assembly proteins. This phage region is one of the two not found in the genome of Sakai or EDL933 (Figure 4-2, black arrows). The eight region (phage 15) contains the most similarity to TW14359 with EC4115 containing ~47.5kb additional phage-related genes in the middle of the region. The ninth regions contains a hypothetical protein and a transposase that are found in all references except EDL933.

The smallest block (five) is inverted in EDL933 and Sakai, and contains 2 regions that are dissimilar across the references (Figure 4-4). The first region contains two
transposases that are not found in EDL933 and Sakai. The second region (phage 16) located at the end of the block is similar to TW14359 and EC4115, and contains the \( stx 2 \) A and B subunits. In SS17, this phage region is \( \sim 62 \text{kb} \), located between \( yfdC \) and a tRNA, and has similarity to the Sp5 of Sakai located between \( yccJ \) and \( yccG \) (Figure 4-2, pink arrow).

The sixth block at \( \sim 2280 \text{kb} \) is the largest and contains seven regions that are dissimilar when SS17 is compared to all of the references. The first region begins at the start of the block, covers phage region 16 and 17, and is similar to TW14359 and EC4115. The second and third regions have homology between SS17 and all of the references except EDL933. The fourth (phage 18) and fifth regions, which contains a transposase, are similar to TW14359 and EC4115. The sixth region includes two tRNAs that are not found in EDL933. The seventh region (phage 21) is the second phage region not found in Sakai or EDL933 but has similarity to TW14359 and EC4115 (Figure 4-2, black arrows). Phage region 21 is \( \sim 12 \text{kb} \) and located between \( yjbK \) and \( dusA \). Analysis shows this region, even though it is located proximal to Sakai’s Mu-like phage containing Sp18, does not have similarity to Sp18 which is lacking in SS17.

We also compared the plasmids found in SS17. The pO157 in SS17 is similar to the pO157 found in TW14359 and EC4115, and differs from the plasmid in Sakai by three regions. Two regions contain additional genes not found in Sakai (two transposase genes, and a transposase and two hypothetical proteins). The third region is absent in SS17 and contains a reverse transcriptase gene in Sakai. The pSS17 plasmid is similar to EC4115 and contained only minor differences.
4.4.2 Single Nucleotide Polymorphisms

Analyze revealed 310 SNPs in SS17 as compared with EC4115, 833 SNPs as compared with TW14359, 3860 as compared with Sakai, and 4847 as compared with strain EDL933 (Figure 4-6). These variations are expected as SS17 is more closely related to EC4115 and TW14359 than Sakai and EDL933 (Figure 4-3). There are 147 genes that have 262 non-synonymous SNPs (nsSNPs) unique to SS17 and not found in the reference strains. A few examples of genes with nsSNPs include the narQ, galP, yhjV, and nudE. The narQ gene along with narP belongs to the two component sensor-histidine kinase system and has one nsSNP of T452I. The galP gene, part of the galactose transport family, has two nsSNPs each with two nucleotide substitutions at L433H and W434V. The yhjV gene, an uncharacterised member of the STP family of transporters, has one sSNP and three nsSNPs at L103F, I109S, and N111I. The nudE gene belongs to the nudix hydrolase family of genes and has one nsSNP with two nucleotide substitutions at P146H.

Figure 4-6. SNP patterns of reference O157 genomes compared to SS17
Single nucleotide polymorphisms (SNPs) in SS17 compared to Sakai (blue), EDL933 (green), TW14359 (red), and EC4115 (purple).
There are 50 nsSNPs in SS17 with at least one reference in 41 genes related to adherence, toxin production, and other virulence functions. In each of these categories there are 20, 13, and 8 genes, respectively. The majority of the nsSNPs (27) are located in genes responsible for adherence, specifically non-fimbrial adhesins and include cah, yfaL, and toxB (Table 4-2). Of particular interest is the cah gene which encodes for the calcium-binding antigen-43 homologue. There is a nonsense mutation that causes the terminal end of the gene to be truncated 171 amino acids shorter than the gene in the reference E. coli O157:H7 strains. The yfaL gene is a type V AidA-like adhesin and contains 3 nsSNPs when compared to EDL933 and Sakai. In addition, there is one gene on the pO157 plasmid, toxB which contains two nsSNPs and three synonymous SNPs and works in concert with the efa1′ gene which also has a nsSNP (Table 4-2). When compared to all four references, there are 3 genes (cah, yfsS, and 2871) with nsSNPs.

There are also 13 nsSNPs in genes that function as toxins including the LEE-encoded type III secretion factors (escT & escR), the non-LEE mediated host cell effector (espR4 & nleH1), the lipid A biosynthesis gene (lpxB), and hemolysin genes encoded by the pO157 plasmid (hlyA & hlyC) (Table 4-2). The virulence-associated category has 8 nsSNPs and includes genes involved in nutrient acquisition (chuA & chuS), capsule biogenesis (wcaM), resistance (terD & terF), and signaling (phoB).
## Table 4-2. Virulence genes in SS17 with nsSNPs

<table>
<thead>
<tr>
<th>Category &amp; Location</th>
<th>Gene</th>
<th>SNP</th>
<th>Reference</th>
<th>Product</th>
<th>Role in Virulence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Toxin</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chromosome</td>
<td><em>eivA</em></td>
<td>P33S</td>
<td>5</td>
<td>T3SS apparatus protein</td>
<td>LEE-encoded effector secretion</td>
</tr>
<tr>
<td></td>
<td><em>eivF</em></td>
<td>H163Q</td>
<td>3 &amp; 4</td>
<td>Putative regulatory protein for T3SS</td>
<td>Functional regulation</td>
</tr>
<tr>
<td></td>
<td><em>escE</em></td>
<td>H46N</td>
<td>4</td>
<td>LEE T3SS factor</td>
<td>LEE-encoded effector secretion</td>
</tr>
<tr>
<td></td>
<td><em>escR</em></td>
<td>G29V</td>
<td>4</td>
<td>LEE T3SS factor</td>
<td>LEE-encoded effector secretion</td>
</tr>
<tr>
<td></td>
<td><em>escT</em></td>
<td>F68L</td>
<td>1</td>
<td>LEE T3SS factor</td>
<td>LEE-encoded effector secretion</td>
</tr>
<tr>
<td></td>
<td><em>espR1</em></td>
<td>S57N</td>
<td>2</td>
<td>Predicted NLE T3SS effector</td>
<td>NLE mediated host cell effects</td>
</tr>
<tr>
<td></td>
<td><em>espR4</em></td>
<td>S93A</td>
<td>3 &amp; 4</td>
<td>Predicted NLE T3SS effector</td>
<td>NLE mediated host cell effects</td>
</tr>
<tr>
<td></td>
<td><em>espX6</em></td>
<td>F220L</td>
<td>3 &amp; 4</td>
<td>Predicted NLE T3SS effector</td>
<td>NLE mediated host cell effects</td>
</tr>
<tr>
<td></td>
<td>lpxB</td>
<td>A181R</td>
<td>5</td>
<td>Tetraacyldisaccharide-1-P synthase</td>
<td>Endotoxin Biosynthesis</td>
</tr>
<tr>
<td></td>
<td><em>nleH1</em></td>
<td>V245I</td>
<td>3 &amp; 4</td>
<td>NLE T3SS effector</td>
<td>NLE mediated host cell effects</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0524</td>
<td>3 &amp; 4</td>
<td>RTX-Family Toxin</td>
<td>Host-cell damage</td>
</tr>
<tr>
<td></td>
<td><em>hlyA</em></td>
<td>E157A</td>
<td>4</td>
<td>Hemolysin</td>
<td>Host-cell damage</td>
</tr>
<tr>
<td></td>
<td><em>hlyC</em></td>
<td>L208I</td>
<td>4</td>
<td>Hemolysin</td>
<td>Host-cell damage</td>
</tr>
<tr>
<td><strong>Adherence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chromosome</td>
<td><em>cah</em></td>
<td>G779W</td>
<td>5</td>
<td>Calcium-binding antigen-43 homologue</td>
<td>Non-fimbrial adhesin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>L781A</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>E782R</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T783Opal</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>csgG</td>
<td>P76T</td>
<td>3 &amp; 4</td>
<td>Curli production assembly component</td>
<td>Usher protein</td>
</tr>
<tr>
<td></td>
<td>eaeH</td>
<td>T1301K</td>
<td>3 &amp; 4</td>
<td>Putative adhesin with Ig-like class 3 domain</td>
<td>Non-fimbrial adhesin</td>
</tr>
<tr>
<td></td>
<td><em>efa1</em></td>
<td>T285P</td>
<td>1 &amp; 3 &amp; 4</td>
<td>Central fragment of <em>efa1</em></td>
<td>Non-fimbrial adhesin</td>
</tr>
<tr>
<td></td>
<td><em>fimA</em></td>
<td>V100A</td>
<td>4</td>
<td>Major type 1 subunit</td>
<td>Fimbrial-like adhesin</td>
</tr>
<tr>
<td></td>
<td><em>fimB</em></td>
<td>S80P</td>
<td>1</td>
<td>Tyrosine recombinase</td>
<td>Regulator</td>
</tr>
<tr>
<td></td>
<td><em>fimH</em></td>
<td>K156N</td>
<td>3 &amp; 4</td>
<td>Minor fimbrial subunit</td>
<td>Fimbrial-like adhesin</td>
</tr>
<tr>
<td></td>
<td><em>fmlC</em></td>
<td>D532G</td>
<td>4</td>
<td>F9 fimbriae usher</td>
<td>Usher Protein</td>
</tr>
<tr>
<td></td>
<td><em>htrE</em></td>
<td>H344Q</td>
<td>1</td>
<td>Putative fimbrial usher</td>
<td>Usher Protein</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T460A</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>ppdB</em></td>
<td>P158L</td>
<td>3 &amp; 4</td>
<td>Prepilin peptidase</td>
<td>Chaperone</td>
</tr>
<tr>
<td></td>
<td><em>wzzB</em></td>
<td>P136S</td>
<td>3 &amp; 4</td>
<td>O-antigen short chain regulator</td>
<td>Regulator</td>
</tr>
<tr>
<td>Category &amp; Location</td>
<td>Gene</td>
<td>SNP</td>
<td>Reference*</td>
<td>Product</td>
<td>Role in Virulence</td>
</tr>
<tr>
<td>----------------------</td>
<td>------</td>
<td>-----</td>
<td>------------</td>
<td>---------</td>
<td>------------------</td>
</tr>
<tr>
<td></td>
<td>ybgO</td>
<td>V46A</td>
<td>3 &amp; 4</td>
<td>Predicted fimbrial-like adhesin</td>
<td>Fimbrial-like adhesin</td>
</tr>
<tr>
<td></td>
<td>ycbF</td>
<td>R69S</td>
<td>3 &amp; 4</td>
<td>Predicted periplasmic pilin chaperone</td>
<td>Chaperone</td>
</tr>
<tr>
<td></td>
<td>yehC</td>
<td>A32V</td>
<td>3</td>
<td>Predicted periplasmic pilin chaperone</td>
<td>Chaperone</td>
</tr>
<tr>
<td></td>
<td>yfaL</td>
<td>S23P</td>
<td>3 &amp; 4</td>
<td>Type V AidA-like adhesin</td>
<td>Non-fimbrial adhesin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>D38N</td>
<td>3 &amp; 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>T104I</td>
<td>3 &amp; 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>yfcS</td>
<td>P163H</td>
<td>5</td>
<td>Predicted periplasmic pilus chaperone</td>
<td>Chaperone</td>
</tr>
<tr>
<td></td>
<td>yfcU</td>
<td>G746A</td>
<td>3 &amp; 4</td>
<td>Putative fimbrial usher</td>
<td>Usher Protein</td>
</tr>
<tr>
<td></td>
<td>0333</td>
<td>R218Q</td>
<td>3 &amp; 4</td>
<td>Putative adhesin</td>
<td>Non-fimbrial adhesin</td>
</tr>
<tr>
<td></td>
<td>2871</td>
<td>I282K</td>
<td>5</td>
<td>Putative fimbrial-like adhesin</td>
<td>Fimbrial-like adhesin</td>
</tr>
<tr>
<td></td>
<td>pO157</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>toxB</td>
<td>V237D</td>
<td>4</td>
<td>Cytotoxin B</td>
<td>Non-fimbrial adhesin</td>
</tr>
<tr>
<td></td>
<td></td>
<td>T1888I</td>
<td>4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Virulence-associated |
|----------------------|---------|---------|----------------------------------|------------------|
|                      | Chromosome | chuA | G259E, I577N | 3 & 4 | Outer membrane heme receptor | Nutrient acquisition |
|                      | chuS | K11T | 5 | Heme oxygenase | Nutrient acquisition |
|                      | phoB | Q179L | 3 & 4 | Phosphorus regulator | Signaling |
|                      | terD | M155R | 1 & 2 & 4 | Tellurite resistance | Resistance |
|                      | terF | Opal103S | 4 | Tellurite resistance | Resistance |
|                      | wcaJ | D216D | 4 | Colanic acid lipid carrier transferase | Biofilm/capsule biogenesis |
|                      | wcaL | R73S | 3 & 4 | Colanic acid glycosyl transferase | Biofilm/capsule biogenesis |
|                      | wcaM | L29F | 3 & 4 | Predicted colanic acid biosynthesis protein | Biofilm/capsule biogenesis |

* Reference with SNP: 1 - EC4115; 2 - TW14359; 3 - EDL933; 4 - Sakai; 5 - All four strains
4.4.3 Polymorphisms in Promoter and Termination Regions

In addition to analyzing genes for polymorphisms and protein changes, we looked at upstream (250bp) and downstream (100bp) sequences to determine any possible change in promoter or termination regions that could affect protein expression and contribute to the super-shedder phenotype. Our analysis revealed 108 genes with upstream sequences that differed from all four references. Genes include the adherence related eaeH, csgD, wzzB, and yraH, the LEE-encoded type III effector, espG, and the auto-inducer, luxS. In addition, a number of genes are associated with membrane function, including the membrane proteins, ytfF and imp, a putative membrane permeability protein, and an outer membrane receptor, fepA. Downstream, termination sequence analysis revealed 95 genes with differences compared to all four of the reference strains. Genes include the outer membrane proteins, ompF, yhfL and lomW, the iron-siderophore receptor precursor, fhuE, the regulatory protein, abgR, and yeeV which is the toxin component of the YeeV-YeeU toxin-antitoxin system. Of particular interest is the Q933 gene adjacent to the stx2 A and B subunits which has a single nucleotide deletion within the 100bp downstream region.

4.4.4 Other Types of Polymorphisms

There are 23 genes with one or more amino acids that have been deleted or inserted in SS17 and include mainly hypothetical and phage-related proteins. The mdtO gene, part of the multidrug efflux family of proteins which are involved in resistance to puromycin, acriflavine, and tetraphenylarsonium chloride, is one of the genes with a
deletion of 3 amino acids in the middle of the protein. In addition, the *appB* gene, a cytochrome d ubiquinol oxidase, has eight nsSNPs and one inserted amino acid.

Out of the 5,442 CDS that are in SS17, 18 genes (0.31%) have mutations that resulted in frameshifts or premature stop codons (Table 4-3). Sixteen genes have frameshifts and include a dehydrogenase, protease regulator, phage portal protein, a putative heme binding factor, and an inner membrane protein. Gene *hyfA*, a hydrogenase, contains a frameshift and is part of a ten-gene cluster that interacts with formate dehydrogenase [71]. In addition, two genes contained premature stop codons and include a galactosamine-6-phosphate isomerase (*agal*) and a putative ABC ATP-binding protein (*ytfR*) (Table 4-3).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>FdnG</td>
<td>Formate dehydrogenase N alpha subunit</td>
</tr>
<tr>
<td>hflK</td>
<td>FtsH protease regulator</td>
</tr>
<tr>
<td>hyfA</td>
<td>Hydrogenase 4 Fe-S subunit</td>
</tr>
<tr>
<td>mfd</td>
<td>Transcription-repair coupling factor</td>
</tr>
<tr>
<td>potA</td>
<td>Putrescine transport ATP-binding protein</td>
</tr>
<tr>
<td>1463</td>
<td>Phage-associated DNA primase</td>
</tr>
<tr>
<td>1796</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>2091</td>
<td>Phage minor tail protein</td>
</tr>
<tr>
<td>2678</td>
<td>Phage portal protein</td>
</tr>
<tr>
<td>3263</td>
<td>Putative heme binding factor</td>
</tr>
<tr>
<td>3265</td>
<td>Hypothetical protein</td>
</tr>
<tr>
<td>tap</td>
<td>Methyl-accepting chemotaxis protein IV</td>
</tr>
<tr>
<td>ybdR</td>
<td>Putative oxidoreductase</td>
</tr>
<tr>
<td>ybeS</td>
<td>Enzyme of polynucleotide modification</td>
</tr>
<tr>
<td>ybhF</td>
<td>ABC transporter of multidrug efflux pump</td>
</tr>
<tr>
<td>yhjT</td>
<td>Putative inner membrane protein</td>
</tr>
<tr>
<td>agal</td>
<td>Galactosamine-6-phosphate isomerase</td>
</tr>
<tr>
<td>ytfR</td>
<td>Putative ABC ATP-binding protein</td>
</tr>
</tbody>
</table>

Table 4-3. Genes in SS17 with frameshifts and premature stop codons
4.6 RSE Adherence Assay

Super-shedder isolates of *E. coli* O157:H7 display a unique phenotype that distinguishes them from other O157:H7 isolates. This phenotype, and the associated high shedding levels, is known to be dependent on colonization of the RAJ region in cattle [9]. With this in mind, we examined the ability of super-shedder isolate 17 (SS17) to bind and adhere to bovine RSE cells isolated from the RAJ region.

SS17 interacted with bovine RSE cells in a distinct adherence pattern when compared to the O157 control strains, EDL933 and 86-24 (Figure 4-7A and B). The results show that 100% of the RSE cells exposed to SS17 had a significantly (p < 0.01) larger number of cells with more than 10 bacteria per RSE cell as compared with only 16.5% for EDL933 and 80% with 86-24. These adherence levels classify SS17 as having a strong adherence to RSE cells. In addition, immunofluorescent microscopy determined SS17 forms aggregative clumps of bacteria on the cell surface (Figure 4-7B). Interestingly, this distinct strong, aggregative adherence pattern of SS17 on RSE cells was different from the moderate, diffuse adherence pattern observed for the same isolate on human HEp-2 cells (Figure 4-7A). In addition to SS17, eight other super-shedder isolates were examined for this unique adherence pattern. Indeed, all of the super-shedder isolates tested displayed the strong, aggregative adherence pattern seen with SS17 (Figure 4-7C). This suggests a unique mechanism used by super-shedder isolates that allows greater adherence to the bovine rectal epithelial cells.
Figure 4-7. Adherence patterns of SS17 and reference O157 strains to RSE cells

(A) Adherence levels of SS17, EDL933, and 86-24 to RSE and HEp-2 cells. SS17 displayed strong adherence with greater than 10 bacteria per RSE cell and only moderate adherence to HEp-2 cells. In contrast, EDL933 showed moderate adherence to both RSE and HEp-2 cells while 86-24 showed strong adherence to RSE cells and moderate adherence to HEp-2 cells. (B) Immunofluorescence stained slides revealed that both SS17 and EDL933 displayed an aggregative adherence pattern to RSE cells while 86-24 displayed a diffuse pattern. (C) Average of adherence levels for nine super-shedder isolates which all display strong, aggregative adherence to RSE cells. Arrows point to adherent bacteria (green). Slides are shown at 40x magnification with RSE cytokeratin (red) and nuclei (blue).
To elucidate possible mechanisms of the increased adherence to RSE cells, pooled antisera against the LEE-encoded proteins, intimin, Tir, EspA, and EspB, was used to inhibit this adherence pathway. The results showed that the pooled antisera did not block the adherence of SS17 to RSE cells but did block the adherence of SS17 to HEp-2 cells (Figure 4-8). While SS17 was rendered ‘non-adherent’ in the presence of the pooled antisera, a negligible (6%) decrease in adherence to RSE cells was observed. In addition, D-Mannose was used to determine the role of type I fimbriae in the adherence to RSE cells. Adherence to cells will be inhibited in the presence of D-Mannose if type I fimbriae are involved in adherence [16]. SS17 showed no differences in adherence levels in the presence or absence of D-Mannose, indicating type I fimbriae are not being utilized for adherence to RSE or HEp-2 cells. The results from both mechanistic tests are similar to what has been observed previously with other O157 strains suggesting unique mechanisms or factors are operating in the adherence of SS17 and other O157 strains to bovine RSE cells [35].
Figure 4-8. Adherence pattern of SS17 in the presence of antisera

Pooled antisera against LEE-encoded proteins (EspA, EspB, Tir, and intimin) was used to elucidate to mechanisms of adherence. (A) Adherence of SS17 on bovine RSE cells revealed a LEE-independent mechanism while decreased adherence of SS17 on human HEp-2 cells shows a LEE-dependent mechanism of attachment. (B) Immunofluorescence images show adherence of SS17 to RSE cells and HEp-2 cells in the absence of antisera (top) and the presence of antisera (bottom) to RSE cells with the toluidine blue stained slide depicting a lack of adherence to HEp-2 cells. All slides are at 40x magnification and have arrows pointing to bacteria (green) with RSE cytokeratins (red) and nuclei (blue).
Chapter 5

Conclusion and Future Directions

\textit{E. coli} O157:H7 remains a major public health concern even with the development of a number of control methods including vaccination and bacteriophage therapy [11, 13, 72, 73]. Unfortunately, these and other control methods are lacking and the result is a number of large, deadly multi-state outbreaks caused primarily through the consumption of contaminated food products.

With the major animal reservoir being asymptomatic cattle, the question becomes how can the pathogen be effectively controlled in such large populations of cattle housed in close proximity when no symptoms are present. Reducing the pathogenic bacteria load in cattle is the most significant way to reduce environmental contamination, human exposure, and ultimately foodborne illness and death [74]. While a number of approaches including vaccination, the use of probiotics, and the application of bacteriophage have been tested with the goal of reducing \textit{E. coli} O157:H7 in cattle, these have unfortunately met with only limited success [11, 13, 72, 73]. One possible reason is that these methods do not take into account the small population of cattle that are colonized with super-shedder isolates of \textit{E. coli} O157:H7 which are the largest contributor of \textit{E. coli} to the environment and account for the majority of transmission events [11].

An effective control method is one that is cost effective to be applied to a large population of cattle and yet specific enough to target super-shedder isolates. In order to
control super-shedder isolates in cattle, treatment methods will need to focus on preventing the bacteria from adhering and colonizing the recto-anal junction (RAJ), as this is a super-shedder related microbial factor that is known to contribute to high fecal shedding [36]. The overall goal of studying super-shedder isolates at the genomic and phenotypic level is to determine the factors that influence the super-shedding phenomena and aid in the development of the next generation of microbial control measures including the development of vaccines and preventative measures that can specifically target super-shedders.

Although various environmental and cattle host factors have been studied and are likely to contribute to super-shedding of *E. coli* O157:H7 by cattle, little is known about the microbial factors and molecular mechanisms that contribute to this phenotype [12]. With this in mind, we set out to analyze the genome of a super-shedder isolate and identify specific genomic factors and loci that could be further studied to determine their role in the super-shedder phenotype. Our analysis has revealed approximately 60 targets that deserve a closer look to determine the specific effect each has on the super-shedder phenotype. Targets include those that contain nsSNPs and other polymorphisms in virulence genes, specifically those involve in adherence. Given the distinct strong, aggregative adherence pattern observed with SS17 and other super-shedder isolates on RSE cells, it would be relevant to dissect mechanisms enabling this type of unique adherence. LEE-encoded proteins, critical to O157 interactions with FAE cells and HEp-2 cells [9, 34], were found not to contribute towards the adherence of SS17 to RSE cells. In addition, the role of type I fimbriae was ruled out by the lack of adherence inhibition seen in the presence of D-Mannose. This was not a surprise as other O157:H7 strains,
including EDL933 and 86-24, displayed adherence patterns that were not affected by LEE antisera and D-Mannose [35]. Understanding the adherence and colonization molecular mechanisms of super-shedders is critical for the development of control strategies that target adherence and can be used to eliminate *E. coli* O157:H7 from the RAJ.

We are a long way from knowing the cause of super-shedding and what factors the microbe contributes. Future studies will need to address the similarities and differences seen in the genome of the super-shedder isolate 17. Since SS17 has a genome arrangement and composition that is similar to reference *E. coli* O157:H7, large rearrangements and insertions or deletions are not expected to be the cause of the super-shedder phenotype. More likely the numerous nsSNPs is virulence and adherence genes will play a role in the phenotype. This will only be the case if the nsSNP or change in gene expression is conserved in other super-shedder genomes. Hence, the first step is to sequence and compare additional super-shedder genomes, to give a better picture of the genes and polymorphisms that may contribute to super-shedding.

Future studies also need to address the enhanced adherence pattern seen in the RSE cell *in vitro* adherence assay. The strong, aggregative adherence pattern was not only unique to SS17 but was also seen in eight other super-shedder isolates indicating this may be a conserved trait among super-shedders; however, the mechanism is still unknown. The first step will be to determine the proteome of SS17 and other super-shedders in the adherence assay with both RSE and HEp-2 cells since the bovine and human cell use different adherence mechanisms. From the proteome, information on possible adherence mechanisms used by the super-shedders can be gained based on the
proteins that have altered expression levels. The next step will be to block the mechanism(s) using either anti-sera or knockout methods.

Once the possible mechanism(s) of adherence are determined, the next endeavor will be to determine if genes that play a role in the mechanism have polymorphisms that effect the expression of the proteins and if the polymorphism is conserved in super-shedder genomes. Concurrently, another approach would be to knockout specific genes that have polymorphisms conserved in the super-shedder genomes to determine the effect on the adherence phenotype. This will elucidate any genes that may not have a known function in adherence and those that do not have altered expression levels.

The study presented here and the future proposed studies will enable the identification of the mechanism(s) involved in the adherence and colonization of the RAJ. These studies will also aid in the overall goal to determine the molecular mechanism or factors that could possibly contribute to super-shedding at the microbe level. Knowing the cause of super-shedding at the microbe level will enable the design of the next generation of control methods that can target super-shedder isolates of *E. coli* O157:H7 and reduce this pathogenic bacteria, prevent the transmission within cattle herd, and ultimately the spread to humans.
References


